JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 12 DECEMBER 1969

Internal Multiplicity Structure for the Chain $SU(n) \supset SU(n-1) \supset \cdots \supset SU(2)$

D. RADHAKRISHNAN
Department of Physics and Astrophysics, University of Delhi, Delhi-7, India

(Received 26 March 1968; Revised Manuscript Received 2 May 1969)

The partition function for $SU(n)$ is given in terms of that for $SU(n-1)$ through a recursion formula which is derived using the method of generating series. The usefulness of the expression is demonstrated in the cases of specific values of the rank.

I. INTRODUCTION

It is well known that the multiplicity of a weight in an irreducible representation (IR) of a complex semisimple Lie algebra is calculable using Kostant's formula.[1] This involves the explicit knowledge of a certain partition function, the connection of which with Diophantine equations has been established.[2] In Sec. II, the corresponding set of equations is written down for the algebra A_l [$\sim SU(l+1)$] and the solution by the method of generating series is indicated. In Sec. III, an expression is derived for the partition functions for A_l in terms of that for A_{l-1}. The usefulness of this recursion formula is demonstrated in the Appendix in the cases of specific values of the rank.

II. METHOD OF GENERATING SERIES

Let G be a complex semisimple Lie algebra and let the set of positive roots, with respect to a given Cartan subalgebra and fixed lexicographic ordering, be $\Delta \equiv \{\alpha_1, \cdots, \alpha_m\}$. The subset of simple roots is $\{\alpha_1, \cdots, \alpha_l\}$, with $l \leq m$. The multiplicity of a weight ν in a finite-dimensional IR $D(\lambda)$ of G with highest weight λ is given by Kostant's formula[1]

$$m_\lambda(\nu) = \sum_{S \in W} \pi_S P[S(\beta + \lambda) - (\beta + \nu)]. \tag{1}$$

Here, W is the Weyl group and $\pi_S = \pm 1$ according to whether S is an even or odd reflection, respectively. Also, β is half the sum of the positive roots. The partition function $P(\mu)$ is the number of ways[3] of writing the weight μ as a linear combination over Δ with nonnegative integral coefficients. Also, $P(0) = 1$, and $P(\mu) = 0$ unless $\mu = \sum_{i=1}^{l} k_i \alpha_i$, where the k_i are nonnegative integers. In short, the function is given by the number of ways of writing

$$\sum_{i=1}^{l} k_i \alpha_i = \sum_{i=1}^{m} a_i \alpha_i, \tag{2}$$

where the a_i are nonnegative integers.

It is known that

$$\alpha_j = \sum_{i=1}^{l} C_{ji} \alpha_i, \quad j = l+1, \cdots, m, \tag{3}$$

the C_{ji} being nonnegative integers. From (2) and (3), we see that $P(\mu) \equiv P(k_1, \cdots, k_l)$ is given by the number of solutions in nonnegative integers of the Diophantine equations

$$k_i = a_i + \sum_{j=l+1}^{m} a_j C_{ji}, \quad i = 1, \cdots, l. \tag{4}$$

In the case of the algebra A_l, the positive roots are of the form

$$e_i - e_k, \quad i < k = 1, \cdots, l+1,$$

and are $\frac{1}{2}l(l+1)$ in number. The simple roots are

$$e_i - e_{i+1}, \quad i = 1, \cdots, l,$$

where the e_i are unit vectors in a $(l+1)$-dimensional real linear space. The roots (and weights) are considered to lie in a hyperplane $\sum_{i=1}^{l+1} x_i = 0$ of this space.

Expression (2) in the case of A_l can be written as

$$\sum_{i=1}^{l} k_i \alpha_i = \sum_{r=1}^{l} \sum_{s=1}^{l-r+1} a_{rs} \sum_{t=1}^{r} \alpha_{s+t-1}, \tag{5}$$

where a double-index notation for the a's has been chosen for convenience. Then, Eq. (4) for A_l becomes

$$k_i = \sum_{r=1}^{i} \sum_{j=r}^{l-i+r} a_{j;i-r+1}, \quad i = 1, \cdots, l. \tag{6}$$

In the case of A_3 [$\sim SU(4)$], for example, we have

$$\begin{aligned} k_1 &= a_{11} + a_{21} + a_{31}, \\ k_2 &= a_{12} + a_{21} + a_{22} + a_{31}, \\ k_3 &= a_{13} + a_{22} + a_{31}. \end{aligned} \tag{6'}$$

[1] B. Kostant, Trans. Am. Math. Soc. 93, 53 (1959).
[2] D. Radhakrishnan and T. S. Santhanam, J. Math. Phys. 8, 2206 (1967).
[3] N. Jacobson, *Lie Algebras* (Interscience Publishers, Inc., New York, 1962), p. 260.

The number of solutions of (6) is given by the method of generating series[4] as follows: Let $x_1, \cdots, x_l$ be l real variables and define a generating function F_l for A_l as

$$F_l(x_1, \cdots, x_l) = \prod_{i=1}^{l} \left[\prod_{j=1}^{l-i+1} \left(1 - \prod_{k=0}^{i-1} x_{j+k} \right)^{-1} \right]. \quad (7)$$

Then the partition function $P(k_1, \cdots, k_l)$ is given by the coefficient of $\prod_{i=1}^{l} x_i^{k_i}$ in the power-series expansion in $x_1, \cdots, x_l$ of F_l. (The convergence of the series is assured by a suitable choice of values for the x-variables.)

The expression (7) in the case of A_2 is of the form

$$F_2(x_1, x_2) = \{(1 - x_1)(1 - x_2)(1 - x_1x_2)\}^{-1}.$$

III. RECURSION FORMULA

Obviously,

$$F_l(x_1, \cdots, x_l) = D_l F_{l-1}(x_1, \cdots, x_{l-1}), \quad (8)$$

where F_{l-1} is the generating function corresponding to the algebra A_{l-1} and

$$D_l = \prod_{j=1}^{l} \left(1 - \prod_{i=l-j+1}^{l} x_i \right)^{-1}. \quad (9)$$

We split D_l into partial fractions as follows:

$$D_l = (1 - x_l)^{-1} \prod_{k=1}^{l-1} \left(1 - \prod_{i=l-k}^{l-1} x_i \right)^{-1} + \sum_{j=2}^{l} T_j \left(1 - \prod_{i=l-j+1}^{l} x_i \right)^{-1}, \quad (10)$$

where

$$T_j = (-1)^{j-1} \prod_{i=1}^{j-1} \left[x_{l-i}^{i} \left(1 - \prod_{k=j-1}^{i} x_{l-k} \right)^{-1} \right] \times \prod_{s=1}^{l-j} \left(1 - \prod_{r=l-j-s+1}^{l-j} x_r \right)^{-1}.$$

(By convention, those x-variables with undefined indices are ignored. Thus the product involving s does not occur in T_l.)

Now, the coefficient of $x_l^{k_l}$ in D_l is

$$I_l = \prod_{k=1}^{l-1} \left(1 - \prod_{i=l-k}^{l-1} x_i \right)^{-1} + \sum_{j=2}^{l} (-1)^{j-1} \prod_{i=1}^{j-1} \left[x_{l-i}^{k_l+i} \left(1 - \prod_{k=j-1}^{i} x_{l-k} \right)^{-1} \right] \times \prod_{s=1}^{l-j} \left(1 - \prod_{r=l-j-s+1}^{l-j} x_r \right)^{-1}. \quad (11)$$

Noting that the first term on the right of (11) is D_{l-1}, we substitute for it from (10) and, after simplification, obtain the recurrence relation

$$I_l = \sum_{t=0}^{k_l} x_{l-1}^{t} I_{l-1}, \quad (12)$$

where we have used

$$\frac{1 - x^{k+1}}{1 - x} = \sum_{t=0}^{k} x^t, \quad x \neq 1.$$

It follows from (12), with $t = i_{l-1}$, that

$$I_l = \sum_{i_{l-1}=0}^{k_l} x_{l-1}^{i_{l-1}} \sum_{i_{l-2}=0}^{i_{l-1}} x_{l-2}^{i_{l-2}} \cdots \sum_{i_1=0}^{i_2} x_1^{i_1}. \quad (13)$$

The coefficient of $x_l^{k_l}$ in F_l is then given by $I_l F_{l-1}$. Let the coefficient of $\prod_{i=1}^{l-1} x_i^{k_i}$ in F_{l-1} be $P(k_1, \cdots, k_{l-1})$, the partition function corresponding to A_{l-1}. It follows that the coefficient of $\prod_{i=1}^{l} x_i^{k_i}$ in F_l is

$$P(k_1, \cdots, k_l) = \sum_{i_{l-1}=0}^{k_l} \sum_{i_{l-2}=0}^{i_{l-1}} \cdots \sum_{i_1=0}^{i_2} P(k_1 - i_1, k_2 - i_2, \cdots, k_{l-1} - i_{l-1}), \quad (14)$$

which is the required partition function.

Thus, knowing the function for a given l, that for $l + 1$ can be calculated recursively. Knowing P, one can calculate the multiplicity of a weight for A_l using Kostant's formula (1). We demonstrate the usefulness of (14) in the Appendix.

ACKNOWLEDGMENTS

The author is grateful to Professor R. C. Majumdar for his interest and support. Thanks are due to Dr. P. K. Menon for bringing this method to the author's notice, and to Dr. B. R. Srinivasan and Dr. T. S. Santhanam for useful discussions.

APPENDIX

The usefulness of the recursion formula (14) can be demonstrated vividly in the cases of A_2 and A_3. For the former, (14) takes the form

$$P(k_1, k_2) = \sum_{i=0}^{k_2} P(k_1 - i), \quad (A1)$$

where $P(k)$ on the right is the partition function corresponding to A_1 given by the coefficient of x^k in the expansion in power series of $F(x) = (1 - x)^{-1}$, and is always unity.

It is not difficult to see that

$$P(k_1, k_2) = 1 + \min(k_1, k_2), \quad (A2)$$

which is a well-known result.[5]

In the case of A_3, Eq. (14) becomes

$$P(k_1, k_2, k_3) = \sum_{j=0}^{k_3} \sum_{i=0}^{j} P(k_1 - i, k_2 - j), \quad (A3)$$

[4] P. A. MacMahon, *Combinatory Analysis* (Chelsea Publ. Co., New York, 1960), Vol. II.

[5] B. Gruber, J. Math. Phys. 7, 1797 (1966).

where $P(k, k')$ on the right is given by (A2). Since $k_2 - j \leq k_1 - i$ when $k_1 \geq k_2$, we find that, for $k_1 \geq k_2 \geq k_3$,

$$P(k_1, k_2, k_3) = \sum_{j=0}^{k_3} (j+1)(1+k_2-j)$$
$$= \tfrac{1}{6}(1+k_3)(2+k_3)(3+k_3)$$
$$+ \tfrac{1}{2}(k_2-k_3)(1+k_3)(2+k_3). \quad \text{(A4)}$$

Since P vanishes for negative arguments, the sum over j in (A3) is only up to k_2 when $k_3 > k_2$. Thus, for $k_1 \geq k_3 > k_2$ and $k_3 > k_1 \geq k_2$,

$$P(k_1, k_2, k_3) = \tfrac{1}{6}(1+k_2)(2+k_2)(3+k_2). \quad \text{(A5)}$$

When $k_2 > k_1 \geq k_3$, we have that

$$P(k_1, k_2, k_3)$$
$$= (1+k_1)(1+k_2-k_1) + \tfrac{1}{6}(k_2-k_1)$$
$$\times (k_2-k_1+1)(2k_1+k_2+2)$$
$$+ \tfrac{1}{2}[(k_1-k_2+k_3)\{k_1+k_2+2-(k_1-k_2)^2\}$$
$$+ k_2(k_1-k_2+k_3)(1-k_1+k_2+k_3)$$
$$- \tfrac{1}{3}k_3(1+k_3)(1+2k_3)], \quad \text{when} \quad k_3 > k_2 - k_1,$$
$$= (1+k_1)(1+k_3) + \tfrac{1}{6}k_3(1+k_3)(3k_1-k_3+1),$$
$$\text{when} \quad k_3 \leq k_2 - k_1. \quad \text{(A6)}$$

When $k_2 > k_3 > k_1$, we have that

$$P(k_1, k_2, k_3)$$
$$= \left(\sum_{j=0}^{k_1}\sum_{i=0}^{j} + \sum_{j=k_1+1}^{k_3}\sum_{i=0}^{k_1}\right) P(k_1-i, k_2-j). \quad \text{(A7)}$$

The first part on the right is given by Eq. (A6) with $k_3 \to k_1$. For $k_2 > k_3 > k_1$, the second part is given by

$$P'(k_1, k_2, k_3)$$
$$= \tfrac{1}{2}(k_3-k_1)\{2+k_2+k_1(k_1+2)-(k_1-k_2)^2$$
$$+ \tfrac{1}{2}(2k_2-2k_1-1)(1+k_1+k_3)\}$$
$$+ \tfrac{1}{12}\{k_1(k_1+1)(2k_1+1) - k_3(k_3+1)$$
$$\times (2k_3+1)\}, \quad \text{when} \quad k_3 > k_2 - k_1,$$
$$= \tfrac{1}{2}(1+k_1)(2+k_1)(k_3-k_1),$$
$$\text{when} \quad k_3 \leq k_2 - k_1. \quad \text{(A8)}$$

Finally, when $k_3 \geq k_2 > k_1$, the sum over j in (A3) is only up to k_2. Splitting (A3) as in (A7), the first part is known, while the second part is obtained from (A8) with $k_3 \to k_2$.

It has been correctly pointed out that the case $k_3 \geq k_2 > k_1$ can be derived from (A4) with $k_1 \leftrightarrow k_3$. Also, (A7) is given by (A6) under the same exchange. This is seen from the symmetry of the Diophantine equations (6′). However, the purpose of this Appendix is only to demonstrate the usefulness of Eq. (14). See, however, Ref. 6.

[6] B. Gruber, Nuovo Cimento **48A**, 23 (1967).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 12 DECEMBER 1969

Wavefunctions on Homogeneous Spaces

HENRI BACRY
Centre de Physique Théorique, Marseille, France
AND
ARNE KIHLBERG
Institute of Theoretical Physics, Göteborg, Sweden

(Received 24 February 1969)

The properties of a class of homogeneous spaces of the Poincaré group are discussed. An 8-dimensional space appears especially promising and the explicit unitary irreducible representations corresponding to physical particles are given using scalar wavefunctions on this space.

INTRODUCTION

The concept of wavefunction has played a paramount role in particle physics since the advent of quantum mechanics. A wavefunction is first of all a representative of the state of the physical system. Thus it is an element of the Hilbert space of state vectors. But it has not only this abstract property. It is an explicit function on coordinate space obeying certain differential equations. These equations often form the starting point for the introduction of interactions. The best known example in relativistic quantum mechanics is the Dirac wavefunction obeying the Dirac equation. It gives an explicit realization of the state vector in the Hilbert space defined by spin-$\frac{1}{2}$ mass-m (> 0) unitary irreducible representation of the Poincaré group. The great importance of the Dirac equation lies in the ease by which the electromagnetic interaction is introduced.

The impact of the Dirac theory on the research on wave equations and wavefunctions for free relativistic particles has been enormous. Thus almost all work in this field has started from the assumption that a wavefunction is a spinor- (tensor-) valued function on the Minkowski space.[1,2] All this work on relativistic wave equations has, despite its elegance and beauty, so far not contributed in a significant way to the solution of the theoretical problems of understanding elementary-particle interactions other than the electromagnetic ones. It is, therefore, natural to ask whether one could widen the concept of wavefunction. Two generalizations are close at hand. The first one consists in considering so-called infinite-component wavefunctions. Then we are concerned with functions on Minkowski space with values in an infinite-dimensional representation space of the homogeneous Lorentz group. During the last few years this possibility has been the subject of many investigations.[3] The second generalization consists in replacing the Minkowski space by a larger space on which the Poincaré group acts. If this action is to be transitive, one is lead to consider the homogeneous spaces of the Poincaré group. Such a generalization has been proposed by Finkelstein and others.[4–8] Of course, there is nothing preventing us from considering both generalizations at the same time, i.e., infinite-component functions on a homogeneous space. However, as we shall see in Sec. 1, the spin degree of freedom is naturally connected to certain coordinates in the homogeneous space and therefore we may completely avoid components or indices as long as we deal with one single particle. This possibility of having continuous variables describing the spin has been already touched upon by Bargmann and Wigner.[2] It may help to give the spin a dynamical role.[6,7,9,10] A more far-reaching generalization of the concept of wavefunction

[1] P. A. M. Dirac, Proc. Roy. Soc. (London) **A155**, 447 (1936); G. Petiau, thèse (Masson & Cie., Paris, 1936); R. G. Duffin, Phys. Rev. **54**, 1114 (1938); N. Kemmer, Proc. Roy. Soc. (London) **A173**, 91 (1939); M. Fierz, Helv. Phys. Acta **12**, 3 (1939); M. Fierz and W. Pauli, Proc. Roy. Soc. (London) **A173**, 211 (1939); W. Rarita and J. Schwinger, Phys. Rev. **60**, 61 (1941); L. de Broglie, *Théorie générale des particules à spin* (Gauthier-Villars, Paris, 1943); H. J. Bhabha, Rev. Mod. Phys. **17**, 300 (1945); **21**, 451 (1949); S. N. Gupta, Phys. Rev. **95**, 1334 (1954); H. Umezawa and A. Visconti, Nucl. Phys. **1**, 20 (1956); S. Weinberg, Phys. Rev. **133**, B1318 (1964); **134**, B882 (1964).

[2] V. Bargmann and E. P. Wigner, Proc. Natl. Acad. Sci. U.S. **34**, 211 (1948).

[3] E. Majorana, Nuovo Cimento **9**, 335 (1932); Harish-Chandra, Phys. Rev. **71**, 793 (1947); Proc. Roy. Soc. (London) **A192**, 195 (1948); I. M. Gel'fand and A. M. Yaglom, Zh. Eksp. Teor. Fiz. **18**, 703 (1948); G. Feldman and P. T. Matthews, Phys. Rev. **154**, 5 (1967); Y. Nambu, *ibid.* **160**, 1171 (1967); *Proceedings of the Eighth Nobel Symposium, 1968* (Almqvist & Wilksell boktryckeri, Stockholm, 1968), p. 105; C. Fronsdal, Phys. Rev. **156**, 1653 (1967); *Proceedings of the Eighth Nobel Symposium, 1968* (Almqvist & Wiksell boktryckeri, Stockholm, 1968), p. 119; T. Takabayasi, Progr. Theoret. Phys. Suppl., 339 (1965); *Proceedings of the Eighth Nobel Symposium, 1968*, (Almqvist & Wiksell boktryckeri, Stockholm, 1968), p. 157; I. T. Todorov, *Proceedings of the Eighth Nobel Symposium, 1968* (Almqvist & Wiksell boktryckeri, Stockholm, 1968), p. 133.

[4] D. Finkelstein, Phys. Rev. **100**, 924 (1955).

[5] A. Kihlberg, Arkiv Fysik **28**, 121 (1964); Nuovo Cimento **53A**, 592 (1968).

[6] H. Bacry and J. Nuyts, Phys. Rev. **157**, 1471 (1967).

[7] F. Lurçat, Phys. **1**, 95 (1964).

[8] J. Nilsson and A. Beskow, Arkiv Fysik **34**, 307 (1967).

[9] A. Kihlberg, Arkiv Fysik **39**, 77 (1969).

[10] A. Kihlberg, "On a New Field Theory," Report No. 68-9, Institute of Theoretical Physics, Göteborg (1968).

has been recently proposed by Finkelstein,[11] but we shall not consider it here.

The construction of wavefunctions for free relativistic particles is only the first step towards a particle theory. Interactions have to be defined. Since particles are created and destroyed, one has to perform second quantization and define fields using creation and annihilation operators in a Fock space. An example of such a field theory has been constructed and shown to lead to many unconventional features such as loss of the spin–statistics connection, nonlocality, and unusual analytic properties of the scattering amplitudes.[10] On the other hand, it seems promising for hadron physics since it has built-in decreasing form factors.

In Ref. 6 a homogeneous space is used to build wavefunctions for particles with nonzero mass and spin. These wavefunctions satisfy two differential equations corresponding to the two invariant operators of the Poincaré group. In other words, they satisfy one equation for the mass and one for the spin. A combination of the two equations is derived from a Lagrangian theory which leads to a mass–spin relation.

This paper contains two parts. The first one is general in character. We classify all homogeneous spaces of the Poincaré group P which "contain" the Minkowski space and which have continuous stabilizer groups. The natural action of P on these spaces is given. It is shown that the generators of P can always be split into an orbital part and a spin part which are mutually commuting. We examine which spaces admit an invariant measure and which admit half-integral-spin wavefunctions. The homogeneous space of lowest dimension having these properties is of dimension 8 and it is therefore natural to consider this space first. This is done in Sec. 2. After a description of the space with the aid of the coset variables defined in Sec. 1, an equivalent description is given in a spinor formalism. The generators and the fundamental invariants of the Poincaré group are calculated. After these general features, the different physical cases are discussed, namely massive spinning particles, massive spinless particles, and massless particles. In each case, a complete set of wavefunctions is explicitly given and a scalar product defined.

The so defined unitary irreducible representation of P can be extended to representations of the full Poincaré group including space and time reflections We do not discuss this extension here but refer to Refs. 5 and 6.

[11] D. Finkelstein, J. Math. Phys. 7, 1218 (1966).

1. CLASSIFICATION OF HOMOGENEOUS SPACES

A homogeneous space E of a group G has the following properties:

(a) It is a topological space on which the group G acts (continuously), i.e., let y be a point in E, then gy is defined and is again a point in E ($g \in G$).

(b) This action is transitive, i.e., if given any two points y_1 and y_2 in the space, it is always possible to find a group element $g \in G$ such that

$$y_2 = gy_1. \tag{1.1}$$

There is a one-to-one correspondence between the homogeneous spaces of G and the coset spaces of G. Denote by S_0 the maximal subgroup of G which leaves the point y_0 invariant,

$$gy_0 = y_0, \quad \text{for} \quad g \in S_0; \tag{1.2}$$

S_0 is called the stabilizer of y_0. Now by writing any group element of G in the form $g = g_c g_0$, where $g_0 \in S_0$ and $g_c \in G/S_0$, we see that, by virtue of the transitivity property (b), any point $y \in E$ can be given by

$$y = g_c g_0 y_0 = g_c y_0. \tag{1.3}$$

Thus the elements g_c of the coset space give a parametrization of E. The mapping $E \leftrightarrow G/S_0$ is, of course, continuous since the group multiplication is continuous and the action on E by definition is continuous. Now the stabilizers S and S_0 of two different points y and y_0 are conjugate, since from

$$S_0 y_0 = y_0,$$

$$y_0 = g^{-1}y, \tag{1.4}$$

it follows that

$$gS_0 g^{-1} y = y, \tag{1.5}$$

i.e.,

$$S = gS_0 g^{-1}. \tag{1.6}$$

Therefore the enumeration of the different E of the Poincaré group P amounts to an enumeration of the subgroups of P up to a conjugation.

We shall now make an important restriction on the class of homogeneous spaces we are going to consider in this paper. We require that E always contains the Minkowski space M which means that four parameters of E can be denoted by $x(x^\mu)$. P must also act on x in the usual way. This means that the stabilizer of a given point in E can never contain an element of the translation subgroup of P. The point $x = 0$, for instance, is invariant under all homogeneous Lorentz transformations, but not under any combination of

translations and Lorentz transformations. The stabilizer must therefore be a subgroup of the homogeneous Lorentz group L.

The restriction to homogeneous spaces containing the Minkowski space is done for physical reasons. We think that it would be very difficult to make an interpretation if the Minkowski space is not present. In this way we are also led to the starting point of Finkelstein.[4] Thus we can use his classification of homogeneous spaces. However, he considers only stabilizers which are generated from the Lie algebra and therefore, for instance, those spaces which have discrete stabilizers are missing. These spaces may be interesting since they can have finite invariant measures.[12] We hope to be able to consider these spaces in the near future.

It is also of interest to consider the homogeneous spaces of both P and $\bar{P}$, the covering group of P. If $E = P/S$, then we write $\bar{E} = \bar{P}/\bar{S}$, where $\bar{S}$ is the corresponding subgroup of $\bar{P}$. $\bar{E}$ may or may not be topologically isomorphic to E.

Let us now consider the action of P on the homogeneous space P/S. If we parametrize P in the form

$$g = g_x g_\Sigma, \tag{1.7}$$

where g_x is a translation x^μ and g_Σ is a homogeneous Lorentz transformation, then the points of E are parametrized by x and Σ modulo an element of S. The action of P is given by left multiplication

$$g_a g_\Lambda g_x g_\Sigma = g_{a+\Lambda x} g_{\Lambda\Sigma}, \quad \text{mod } S. \tag{1.8}$$

Denoting a point in L/S by z we, therefore, have

$$(x, z) \xrightarrow{(a,\Lambda)} (a + \Lambda x, \Lambda z) \tag{1.9}$$

and the action splits into an orbital part on x and an "internal" part on z. For the infinitesimal generators of P this means that they can be written as

$$P_\mu = i\frac{\partial}{\partial x^\mu},$$
$$M_{\mu\nu} = ix_\mu \frac{\partial}{\partial x^\nu} - ix_\nu \frac{\partial}{\partial x^\mu} + S_{\mu\nu}, \tag{1.10}$$

where the $S_{\mu\nu}$ are differential operators only in the variables z. The explicit expression for $S_{\mu\nu}$, of course, depends on the homogeneous space at hand.

We now introduce a parametrization of $L(\bar{L})$ which serves to induce suitable parametrizations of all the homogeneous spaces L/S or $\bar{L}/\bar{S}$. To this end we use the Iwasawa decomposition of L. It says that L can be written as a product of three subgroups[13]

$$L = KAN, \tag{1.11}$$

where K is the maximal compact subgroup $SO(3)$ [$SU(2)$ for $\bar{L}$], A is an Abelian one-parameter subgroup generated by an acceleration and N is a nilpotent (in fact, also Abelian in our case) 2-dimensional subgroup. Let the Lie algebra of L be spanned by $L_{\mu\nu}$ with the commutation relations

$$[L_{\mu\nu}, L_{\rho\sigma}] = i\{g_{\nu\rho}L_{\mu\sigma} + g_{\mu\sigma}L_{\nu\rho} - g_{\mu\rho}L_{\nu\sigma} - g_{\nu\sigma}L_{\mu\rho}\}. \tag{1.12}$$

Then we can let L_{12}, L_{23}, and L_{31} generate K, L_{03} generate A, and $L_{02} - L_{23}$, $L_{01} + L_{31}$ generate N. Introduce now parameters $(\varphi, \theta, \psi, s, t, u)$ in L through the formula

$$\Lambda = e^{-i\varphi L_{12}} e^{-i\theta L_{31}} e^{-i\psi L_{12}} e^{isL_{03}} \times e^{-it[L_{01}+L_{31}]} e^{-iu[L_{02}-L_{23}]}. \tag{1.13}$$

According to the foregoing, φ, θ, and ψ are parameters of K, s is the parameter of A, and t, u those of N. By putting

$$L_{ij} = \tfrac{1}{2}\epsilon_{ijk}\sigma_k,$$
$$L_{0k} = \tfrac{1}{2}i\sigma_k, \tag{1.14}$$

Λ becomes an $SL(2, c)$ matrix parametrized by $(\varphi, \theta, \cdots, u)$. Explicitly, we have

$$\Lambda = \begin{bmatrix} e^{-\frac{1}{2}s}\cos\tfrac{1}{2}\theta e^{-i\frac{1}{2}(\varphi+\psi)} - e^{\frac{1}{2}s}(t + iu)\sin\tfrac{1}{2}\theta e^{-i\frac{1}{2}(\varphi-\psi)}; & -e^{\frac{1}{2}s}\sin\tfrac{1}{2}\theta e^{i\frac{1}{2}(\varphi-\psi)} \\ e^{-\frac{1}{2}s}\sin\tfrac{1}{2}\theta e^{i\frac{1}{2}(\varphi-\psi)} + e^{\frac{1}{2}s}(t + iu)\cos\tfrac{1}{2}\theta e^{i\frac{1}{2}(\varphi+\psi)}; & e^{\frac{1}{2}s}\cos\tfrac{1}{2}\theta e^{i\frac{1}{2}(\varphi+\psi)} \end{bmatrix}. \tag{1.15}$$

One sees that $\varphi + \psi$ and $\varphi - \psi$ are defined only up to a multiple of 4π. It is, therefore, natural to require the ranges

$$0 \leq \varphi + \psi \leq 4\pi,$$
$$-2\pi \leq \varphi - \psi \leq 2\pi \tag{1.16}$$

for $\bar{L} = SL(2, c)$. By similar arguments it can be shown that, if (1.13) is considered to be a parametrization of $L = SO(1, 3)$, the φ and ψ can be chosen in the interval $[0, 2\pi]$. The ranges of the other parameters are in both cases

$$0 \leq \theta \leq \pi,$$
$$-\infty < s, t, u < \infty. \tag{1.17}$$

Notice that topologically the group space of $L(\bar{L})$ is the product of a Euclidean space in three dimensions

[12] L. Michel (private communication).

[13] K. Iwasawa, Ann. Math. **50**, 507 (1949).

and the group space of $SO(3)$ [$SU(2)$]. The parametrization (1.13) is, of course, also a parametrization of the homogeneous space belonging to the trivial stabilizer consisting only of the unit element. Let us calculate the action of L on this space. As an example, we choose an acceleration along the z axis. Then from

$$e^{i\epsilon L_{02}}\Lambda(\varphi, \theta, \cdots, u) = \Lambda(\varphi', \theta', \cdots, u'), \quad (1.18)$$

one gets

$$\begin{aligned}
\varphi' &= \varphi, \\
\psi' &= \psi, \\
\sin \tfrac{1}{2}\theta' &= \frac{e^{-\frac{1}{2}\epsilon} \sin \tfrac{1}{2}\theta}{(\cosh \epsilon + \cos \theta \sinh \epsilon)^{\frac{1}{2}}}, \\
e^{s'} &= e^{s}[\cosh \epsilon + \cos \theta \sinh \epsilon], \qquad (1.19) \\
t' &= t + e^{-s} \frac{\sin \theta \cos \psi \sinh \epsilon}{\cosh \epsilon + \cos \theta \sinh \epsilon}, \\
u' &= u - e^{-s} \frac{\sin \theta \sin \psi \sinh \epsilon}{\cosh \epsilon + \cos \theta \sinh \epsilon}.
\end{aligned}$$

The infinitesimal generators are

$$\begin{aligned}
S_{23} &= i\left(\frac{\cos \theta}{\sin \theta} \cos \varphi \frac{\partial}{\partial \varphi} + \sin \varphi \frac{\partial}{\partial \theta} - \frac{\cos \varphi}{\sin \theta} \frac{\partial}{\partial \psi}\right), \\
S_{31} &= i\left(\frac{\cos \theta}{\sin \theta} \sin \varphi \frac{\partial}{\partial \varphi} - \cos \varphi \frac{\partial}{\partial \theta} - \frac{\sin \varphi}{\sin \theta} \frac{\partial}{\partial \psi}\right), \\
S_{12} &= -i \frac{\partial}{\partial \varphi}, \\
S_{01} &= i\Big(-\frac{\sin \varphi}{\sin \theta} \frac{\partial}{\partial \varphi} + \cos \varphi \cos \theta \frac{\partial}{\partial \theta} \\
&\quad + \frac{\sin \varphi \cos \theta}{\sin \theta} \frac{\partial}{\partial \psi} + \sin \theta \cos \varphi \frac{\partial}{\partial s} \\
&\quad + e^{-s}(\sin \varphi \sin \psi - \cos \varphi \cos \theta \cos \psi) \frac{\partial}{\partial t} \\
&\quad + e^{-s}(\sin \varphi \cos \psi + \cos \varphi \cos \theta \sin \psi) \frac{\partial}{\partial u}\Big), \\
S_{02} &= i\Big(\frac{\cos \varphi}{\sin \theta} \frac{\partial}{\partial \varphi} + \sin \varphi \cos \theta \frac{\partial}{\partial \theta} \\
&\quad - \frac{\cos \varphi \cos \theta}{\sin \theta} \frac{\partial}{\partial \psi} + \sin \theta \sin \varphi \frac{\partial}{\partial s} \\
&\quad - e^{-s}(\cos \varphi \sin \psi + \sin \varphi \cos \theta \cos \psi) \frac{\partial}{\partial t} \\
&\quad - e^{-s}(\cos \varphi \cos \psi - \sin \varphi \cos \theta \sin \psi) \frac{\partial}{\partial u}\Big), \\
S_{03} &= i\Big(-\sin \theta \frac{\partial}{\partial \theta} + \cos \theta \frac{\partial}{\partial s} + e^{-s} \sin \theta \cos \psi \frac{\partial}{\partial t} \\
&\quad - e^{-s} \sin \theta \sin \psi \frac{\partial}{\partial u}\Big). \qquad (1.20)
\end{aligned}$$

They are defined through the equation

$$S_{\mu\nu} f[\Lambda(\varphi, \theta, \cdots)] = \lim_{\epsilon \to 0} (i/\epsilon)\{f(e^{i\epsilon L_{\mu\nu}}\Lambda) - f(\Lambda)\}, \quad (1.21)$$

where $L_{\mu\nu}$ is the corresponding $SL(2, c)$ generator and f is an arbitrary differentiable function of Λ.

Now the parametrization (1.13) also induces a parametrization of other homogeneous spaces. Let, for instance, the stabilizer be generated by $L_{02} - L_{23}$. Then we simply have to ignore the parameter u since we consider Λ only up to a right multiplication by $\exp[-iu(L_{02} - L_{23})]$. A parametrization of E is defined by $(\varphi, \theta, \psi, s, t)$. The action of the group is given by (1.19), omitting the last equation. The infinitesimal generators are given by (1.20), putting $\partial/\partial u = 0$. In the same way, one gets a parametrization of the homogeneous spaces with the stabilizers N, K_3N, AN, K_3AN, where $K_3 = \{\exp(-i\psi L_{12})\}$. Notice that K_3N and AN form groups with N as an invariant subgroup. Also K_3 and A commute. In this way, one finds the spaces [5], [4], [3_0], [3_π], and [2] in the Finkelstein classification. In Table I, all spaces are enumerated and we have also indicated the parametrization. The above procedure needs only a slight modification to cover also the other cases of Table I. Define new parameters $\tilde{t}$ and $\tilde{u}$ through

$$\Lambda = e^{-i\varphi L_{12}} e^{-i\theta L_{31}} e^{-i\tilde{t}(L_{01}+L_{31})} e^{-i\tilde{u}(L_{02}-L_{23})} e^{-i\psi L_{12}} e^{isL_{03}}. \quad (1.22)$$

Notice that ψ and s are unchanged, since N is an invariant subgroup of K_3AN. Comparing Eqs. (1.13) and (1.22), we find

$$\begin{aligned}
\tilde{t} &= e^{s}[t \cos \psi - u \sin \psi], \\
\tilde{u} &= e^{s}[u \cos \psi + t \sin \psi]. \qquad (1.23)
\end{aligned}$$

Equations (1.19) and (1.20) can be rewritten introducing $\tilde{t}$ and $\tilde{u}$ instead of t and u. The cases [5_0] and [4″] are now obtained by omitting ψ or ψ and s. The cases [5_f] and [3_f] need perhaps some further comments. Take first [5_f]. We want to single out a subgroup $\exp \alpha[\cos \tfrac{1}{2}fL_{12} + \sin \tfrac{1}{2}fL_{03}]$. Using the identity

$$\begin{aligned}
&\exp(-i[\psi L_{12} - sL_{03}]) \\
&= e^{-i\tilde{\psi}L_{12}} \exp(-i\alpha[\cos \tfrac{1}{2}fL_{12} + \sin \tfrac{1}{2}fL_{03}]), \quad (1.24)
\end{aligned}$$

where

$$\begin{aligned}
\alpha &= -s/\sin \tfrac{1}{2}f, \\
\tilde{\psi} &= \psi - s \cot \tfrac{1}{2}f, \qquad (1.25)
\end{aligned}$$

we find that $\tilde{\psi}$ is defined up to an angle 4π or 2π, as is ψ. Therefore, we may choose the same range for $\tilde{\psi}$ as for ψ and the space [5_f] is in fact parametrized by three Euler angles and two real parameters $\tilde{t}$ and $\tilde{u}$.

TABLE I. Homogeneous spaces of P and $\bar{P}$ belonging to continuous stabilizers of L and $\bar{L}$. The ranges of the parameters are $-\infty < x^\mu, s, t, u, \tilde{t}, \tilde{u} < \infty, 0 \leq \theta \leq \pi, 0 \leq \varphi + \psi \leq 4\pi, -2\pi \leq \varphi - \psi \leq 2\pi$ for $\bar{P}$ and $0 \leq \varphi, \psi \leq 2\pi$ for P.

Stabilizer generated by	Dimension	Parameters	Half-integral spin	Invariant measure	Notation of Finkelstein
0	10	$x^\mu, \varphi, \theta, \psi, s, t, u$	Yes	$d^4x e^{2s}\, ds\, dt\, du\, d\varphi d \cos\theta\, d\psi$	[6]
$L_{12} - L_{23}$	9	$x^\mu, \varphi, \theta, \psi, s, t$	Yes	$d^4x e^{2s}\, ds\, dt\, d\varphi d \cos\theta\, d\psi$	[5]
L_{12}	9	$x^\mu, \varphi, \theta, s, \tilde{t}, \tilde{u}$	No	$d^4x\, ds\, d\tilde{t}\, d\tilde{u}\, d\varphi d \cos\theta$	$[5_0]$
$\cos \frac{1}{2}fL_{12} + \sin \frac{1}{2}fL_{03}$ $0 < f \leq \pi$	9	$x^\mu, \varphi, \theta, \tilde{\psi}, \tilde{t}, \tilde{u}$	Yes	$d^4x\, d\tilde{t}\, d\tilde{u}\, d\varphi d \cos\theta\, d\tilde{\psi}$	$[5_f]$
$L_{02} - L_{23}$ $L_{01} + L_{31}$	8	$x^\mu, \varphi, \theta, \psi, s$	Yes	$d^4x e^{2s}\, ds\, d\varphi d \cos\theta\, d\psi$	[4]
$L_{02} - L_{23}, L_{03}$	8	$x^\mu, \varphi, \theta, \psi, \tilde{t}$	Yes	No	[4′]
L_{12}, L_{03}	8	$x^\mu, \varphi, \theta, \tilde{t}, \tilde{u}$	No	$d^4x\, d\tilde{t}\, d\tilde{u}\, d\varphi d \cos\theta$	[4″]
L_{12}, L_{23}, L_{31}	7	$x^\mu, q^\mu, q^\mu q_\mu = 1$	No	$d^4x\, (d^3q/q_0)$	[3]
L_{12}, L_{01}, L_{02}	7	$x^\mu, q^\mu, q^\mu q_\mu = -1$	No	$d^4x\, (dq^0\, dq^1\, dq^2/q^3)$	[3]
$L_{12}, L_{02} - L_{23}$ $L_{01} + L_{31}$	7	$x^\mu, \varphi, \theta, s$	No	$d^4x e^{2s}\, ds\, d\varphi d \cos\theta$	$[3_0]$
$\cos \frac{1}{2}fL_{12} + \sin \frac{1}{2}fL_{03}$ $L_{02} - L_{23}$ $L_{01} + L_{31}$ $0 < f \leq \pi$	7	$x^\mu, \varphi, \theta, \tilde{\psi}$	Yes	No	$[3_f]$
L_{12}, L_{03} $L_{02} - L_{23}$ $L_{01} + L_{31}$	6	x^μ, φ, θ	No	No	[2]
$L_{\mu\nu}$	4	x^μ	No	d^4x	[0]

Similar arguments lead to the parameters of $[3_f]$ in Table I. The cases [3] and [3′] are perhaps not so conveniently treated using the Iwasawa decomposition, but it is well known that the homogeneous spaces corresponding to the stabilizer $SU(2)$ $[SO(3)]$ is the timelike hyperboloid and that belonging to $SU(1, 1)$ $[SO(1, 2)]$ is a spacelike hyperboloid (cf. momentum space and Wigner's little groups).

When the stabilizer contains a subgroup of $SU(2)$ it also contains the two-element group Z_2 of matrices ± 1. This means that points in the $SL(2, c)$ group space differing by the element -1 must be identified so that one is lead to the $SO(1, 3)$ group space and further on to a homogeneous space of $L = SO(1, 3)$. Therefore, these spaces cannot carry half-integral-spin representations. In Table I, we have denoted those spaces which admit half-integral-spin representations.

Besides the topological properties, a homogeneous space is characterized by its invariant measure if it exists. Since we know explicitly how the group acts on the space through equations similar to (1.18), finding the explicit form of the measure is straightforward. They are given in Table I. The existence of such a measure may be important for the possibility of defining interactions in a field theory based on a homogeneous space.[10]

If we now choose to work with scalar wavefunctions on the homogeneous spaces, the number of parameters is related to the number of wave equations and the type of particle.[14] Call the dimension of the space d and the number of wave equations e. Then $d - e = 4$ for a massive particle with spin and $d - e = 3$ for a massive particle without spin and a massless particle. Looking at the table we then see that the case [4] requires a minimum number of wave equations if we insist on the existence of an invariant measure and half-integral-spin representations. This space is exactly the one which has been suggested by us[5,6] and which will be treated in more detail in the next section.

2. MODELS BASED ON THE 8-DIMENSIONAL HOMOGENEOUS SPACE [4]

As it has already been emphasized, the homogeneous space [4] is the smallest one which satisfies the two following conditions:

(a) it possesses an invariant measure,

(b) it can be used to describe half-integral spins by means of scalar wavefunctions.

Hereafter we will refer to this space as the space $\bar{H}$, each point of which is parametrized by the vector x and the set $\{s, \varphi, \theta, \psi\}$. The bar on $\bar{H}$ is to recall that we are concerned with a homogeneous space of $\bar{P}$, the covering group of the Poincaré group P. The corresponding homogeneous space H of P is obtained by identifying points of $\bar{H}$ pairwise. $\{s, \varphi, \theta, \psi\}$ and $\{s, \varphi', \theta, \psi'\}$ are considered as identical when $\varphi' - \varphi$ and $\psi' - \psi$ are multiples of 2π.

[14] H. Bacry, Commun. Math. Phys. 5, 95 (1967).

Besides the above parametrization of the space $\bar{H}$, another one appears to be very useful.[6] It consists in labeling each point of $\bar{H}$ by a two-dimensional spinor $\xi = \binom{\xi^1}{\xi^2} \neq 0$ in the following way:

$$\xi^1 = e^{\frac{1}{2}s} e^{i\frac{1}{2}\psi} \cos \tfrac{1}{2}\theta e^{i\frac{1}{2}\varphi},$$
$$\xi^2 = e^{\frac{1}{2}s} e^{i\frac{1}{2}\psi} \sin \tfrac{1}{2}\theta e^{-i\frac{1}{2}\varphi}. \tag{2.1}$$

Our four variables are now ξ^1, ξ^2, $\xi^{\dot{1}}$, and $\xi^{\dot{2}}$, where $\xi^{\dot{\alpha}}$ denotes the complex conjugate of ξ^α. The parametrization of the Minkowski space is made in an analogous way, using spinor indices

$$x = \begin{vmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 - x^3 \end{vmatrix} = \begin{vmatrix} x^{1\dot{1}} & x^{1\dot{2}} \\ x^{2\dot{1}} & x^{2\dot{2}} \end{vmatrix}. \tag{2.2}$$

The action of $\bar{P}$ is simple. Let $[a, \Lambda]$ be an element of $\bar{P}$. Then

$$[a, \Lambda](x, \xi) = (\Lambda x \Lambda^+ + a, \Lambda\xi) \tag{2.3}$$

or, in making use of spinor indices, we have

$$x'^{\alpha\dot{\beta}} = \Lambda^\alpha_\gamma x^{\gamma\dot{\delta}} \Lambda^{+\dot{\beta}}_{\dot{\delta}} + a^{\alpha\dot{\beta}},$$
$$\xi'^\alpha = \Lambda^\alpha_\gamma \xi^\gamma. \tag{2.4}$$

The invariant measure is the volume element $d^4x\, d^4\xi$, where $d^4\xi$ stands for $d\xi^1\, d\xi^2\, d\xi^{\dot{1}}\, d\xi^{\dot{2}}$.

Any function $f(x, \xi)$ is transformed as follows:

$$[a, \Lambda] f(x, \xi) = f([a, \Lambda]^{-1}(x, \xi)). \tag{2.5}$$

It is very easy to derive the following generators of $SL(2, c)$[15]:

$$S^{23} = \frac{1}{2}\left(\xi^1 \frac{\partial}{\partial \xi^2} + \xi^2 \frac{\partial}{\partial \xi^1} - \xi^{\dot{1}} \frac{\partial}{\partial \xi^{\dot{2}}} - \xi^{\dot{2}} \frac{\partial}{\partial \xi^{\dot{1}}}\right),$$
$$S^{31} = -\frac{i}{2}\left(\xi^1 \frac{\partial}{\partial \xi^2} - \xi^2 \frac{\partial}{\partial \xi^1} + \xi^{\dot{1}} \frac{\partial}{\partial \xi^{\dot{2}}} - \xi^{\dot{2}} \frac{\partial}{\partial \xi^{\dot{1}}}\right),$$
$$S^{12} = \frac{1}{2}\left(\xi^1 \frac{\partial}{\partial \xi^1} - \xi^2 \frac{\partial}{\partial \xi^2} - \xi^{\dot{1}} \frac{\partial}{\partial \xi^{\dot{1}}} + \xi^{\dot{2}} \frac{\partial}{\partial \xi^{\dot{2}}}\right),$$
$$S^{10} = S_{01} = \frac{i}{2}\left(\xi^1 \frac{\partial}{\partial \xi^2} + \xi^2 \frac{\partial}{\partial \xi^1} + \xi^{\dot{1}} \frac{\partial}{\partial \xi^{\dot{2}}} + \xi^{\dot{2}} \frac{\partial}{\partial \xi^{\dot{1}}}\right),$$
$$S^{20} = S_{02} = \frac{1}{2}\left(\xi^1 \frac{\partial}{\partial \xi^2} - \xi^2 \frac{\partial}{\partial \xi^1} - \xi^{\dot{1}} \frac{\partial}{\partial \xi^{\dot{2}}} + \xi^{\dot{2}} \frac{\partial}{\partial \xi^{\dot{1}}}\right),$$
$$S^{30} = S_{03} = \frac{i}{2}\left(\xi^1 \frac{\partial}{\partial \xi^1} - \xi^2 \frac{\partial}{\partial \xi^2} + \xi^{\dot{1}} \frac{\partial}{\partial \xi^{\dot{1}}} - \xi^{\dot{2}} \frac{\partial}{\partial \xi^{\dot{2}}}\right). \tag{2.6}$$

Two other operators, induced by linear transformations on the spinor space, can be defined. They are

$$\Delta = \frac{1}{2}\left(\xi^1 \frac{\partial}{\partial \xi^1} + \xi^2 \frac{\partial}{\partial \xi^2} - \xi^{\dot{1}} \frac{\partial}{\partial \xi^{\dot{1}}} - \xi^{\dot{2}} \frac{\partial}{\partial \xi^{\dot{2}}}\right) \tag{2.7}$$

and

$$D = \frac{1}{2}\left(\xi^1 \frac{\partial}{\partial \xi^1} + \xi^2 \frac{\partial}{\partial \xi^2} + \xi^{\dot{1}} \frac{\partial}{\partial \xi^{\dot{1}}} + \xi^{\dot{2}} \frac{\partial}{\partial \xi^{\dot{2}}}\right); \tag{2.8}$$

the latter is the generator of dilatations in the spinor space. Δ and D commute with $SL(2, c)$ and with each other. They form together with the $S^{\mu\nu}$ the Lie algebra of $GL(2, c)$, the group of all linear transformations on the spinor space. Using transformations (2.1), the generators (2.6)–(2.8) become

$$S^{23} = -i \frac{\cos\varphi}{\sin\theta} \frac{\partial}{\partial\psi} + i \sin\varphi \frac{\partial}{\partial\theta} + i \cos\varphi \frac{\cos\theta}{\sin\theta} \frac{\partial}{\partial\varphi},$$
$$S^{31} = -i \frac{\sin\varphi}{\sin\theta} \frac{\partial}{\partial\psi} - i \cos\varphi \frac{\partial}{\partial\theta} + i \sin\varphi \frac{\cos\theta}{\sin\theta} \frac{\partial}{\partial\varphi},$$
$$S^{12} = -i \frac{\partial}{\partial\varphi},$$
$$S^{10} = i \sin\varphi \frac{\cos\theta}{\sin\theta} \frac{\partial}{\partial\psi} + i \cos\varphi \cos\theta \frac{\partial}{\partial\theta} - i \frac{\sin\varphi}{\sin\theta} \frac{\partial}{\partial\varphi} + i \cos\varphi \sin\theta \frac{\partial}{\partial s},$$
$$S^{20} = -i \cos\varphi \frac{\cos\theta}{\sin\theta} \frac{\partial}{\partial\psi} + i \sin\varphi \cos\theta \frac{\partial}{\partial\theta} + i \frac{\cos\varphi}{\sin\theta} \frac{\partial}{\partial\varphi} + i \sin\varphi \sin\theta \frac{\partial}{\partial s},$$
$$S^{30} = -i \sin\theta \frac{\partial}{\partial\theta} + i \cos\theta \frac{\partial}{\partial s}, \tag{2.9}$$

$$\Delta = -i \frac{\partial}{\partial\psi}, \tag{2.10}$$

$$D = \frac{\partial}{\partial s}. \tag{2.11}$$

The relations (2.19) were already obtained in Sec. 1: Replace the terms $\partial/\partial t$ and $\partial/\partial u$ by zero in Eqs. (1.20). The generators of the Poincaré group are

$$P_\mu = i \frac{\partial}{\partial x^\mu}, \tag{2.12}$$

$$M_{\mu\nu} = x_\mu P_\nu - x_\nu P_\mu + S_{\mu\nu}. \tag{2.13}$$

The Pauli–Lubanski vector W^μ, defined as

$$W^\mu = \tfrac{1}{2}\epsilon^{\mu\nu\rho\lambda} M_{\rho\lambda} P_\nu = \tfrac{1}{2}\epsilon^{\mu\nu\rho\lambda} S_{\rho\lambda} P_\nu, \tag{2.14}$$

is easily computed in a spinor basis.[15] One gets

$$W^{\alpha\dot{\beta}} = (E^\alpha_\gamma - \tfrac{1}{2}\delta^\alpha_\gamma E^\delta_\delta) P^{\gamma\dot{\beta}} + (E^{\dot{\beta}}_{\dot{\gamma}} - \tfrac{1}{2}\delta^{\dot{\beta}}_{\dot{\gamma}} E^{\dot{\delta}}_{\dot{\delta}}) P^{\alpha\dot{\gamma}}, \tag{2.15}$$

where

$$E^\alpha_\gamma = \xi^\alpha \frac{\partial}{\partial \xi^\gamma}, \quad E^{\dot{\beta}}_{\dot{\gamma}} = \xi_{\dot{\gamma}} \frac{\partial}{\partial \xi_{\dot{\beta}}}. \tag{2.16}$$

[15] H. Bacry and A. Kihlberg, Commun. Math. Phys. 1, 150 (1965).

[To be more precise, one should write $E^{\alpha}{}_{\gamma} = \xi^{\alpha}(\partial/\partial\xi^{\gamma})$, which cannot be confused with $E_{\gamma}{}^{\alpha} = \xi_{\gamma}(\partial/\partial\xi_{\alpha})$.] One readily obtains

$$\begin{aligned} W^{1\dot{1}} &= S^{12}P^{1\dot{1}} + \xi^1\frac{\partial}{\partial\xi^2}P^{2\dot{1}} - \xi^{\dot{1}}\frac{\partial}{\partial\xi^{\dot{2}}}P^{1\dot{2}}, \\ W^{2\dot{2}} &= -S^{12}P^{2\dot{2}} + \xi^2\frac{\partial}{\partial\xi^1}P^{1\dot{2}} - \xi^{\dot{2}}\frac{\partial}{\partial\xi^{\dot{1}}}P^{2\dot{1}}, \\ W^{1\dot{2}} &= iS^{03}P^{1\dot{2}} + \xi^1\frac{\partial}{\partial\xi^2}P^{2\dot{2}} - \xi^{\dot{2}}\frac{\partial}{\partial\xi^{\dot{1}}}P^{1\dot{1}}, \\ W^{2\dot{1}} &= -iS^{03}P^{2\dot{1}} + \xi^2\frac{\partial}{\partial\xi^1}P^{1\dot{1}} - \xi^{\dot{1}}\frac{\partial}{\partial\xi^{\dot{2}}}P^{2\dot{2}}. \end{aligned} \tag{2.17}$$

A somewhat tedious calculation leads to the following expression for the invariant $W_\mu W^\mu = \frac{1}{2}W_{\alpha\beta}W^{\alpha\beta}$:

$$W_\mu W^\mu = -P_\mu P^\mu D(D+1) + (P_{\alpha\beta}\xi^\alpha\xi^\beta)\left(P^{\gamma\dot{\delta}}\frac{\partial}{\partial\xi^\gamma}\frac{\partial}{\partial\xi^{\dot{\delta}}}\right). \tag{2.18}$$

Note that this invariant is expressed in terms of four invariant operators, namely $P_\mu P^\mu$, D, $P_{\alpha\beta}\xi^\alpha\xi^\beta$, and $P^{\gamma\dot{\delta}}(\partial/\partial\xi^\gamma)(\partial/\partial\xi^{\dot{\delta}})$. Note also that $\xi^\alpha\xi^\beta$ are the components of a lightlike vector k in the future half-cone. The components of k are, according to Eqs. (2.1) and (2.2),

$$k = \tfrac{1}{2}e^s(1, \sin\theta\cos\varphi, -\sin\theta\sin\varphi, \cos\theta). \tag{2.19}$$

The angle ψ is not involved and, taking into account the range of φ, one can say that "a spinor up to a phase" is a double-valued function of a future lightlike vector. In other words, $\{x, k, \psi\}$ is a parametrization of the homogeneous space H which is covered twice by $\bar{H}$.

The fact that $W_\mu W^\mu$ is built with many invariants is not surprising: The number of degrees of freedom is larger than the one which is required for a single elementary particle. If we want our wavefunction to describe an elementary field, it is necessary to have an irreducible representation of the Poincaré group. This is obtained by having the wavefunction satisfy some differential equations. According to Ref. 14, it is necessary to require four equations for the general case (nonzero mass and nonzero spin) and five equations in the case of a massless or a spinless particle. These three cases are now discussed successively.

Case A: Mass and Spin Different from Zero. Even if mass and spin are fixed, that is to say even if the wavefunction f is an eigenfunction of $P_\mu P^\mu$ and $W_\mu W^\mu$,

$$P_\mu P^\mu f(x,\xi) = m^2 f(x,\xi), \tag{2.20}$$

$$W_\mu W^\mu f(x,\xi) = -m^2 j(j+1) f(x,\xi), \tag{2.21}$$

one is left with "internal" degrees of freedom. This means that the representation is reducible. One needs two extra conditions to get a specific particle in the Wigner sense. We require f to be an eigenfunction of D and Δ which both commute with the Poincaré group

$$Df = \frac{\partial f}{\partial s} = \alpha f, \tag{2.22}$$

$$\Delta f = -i\frac{\partial f}{\partial\psi} = nf. \tag{2.23}$$

Therefore, an elementary particle is labeled in this way by its mass m, its spin j, and the two quantum numbers α and n. The parameter α, which is related to a noncompact group, can take any complex value. The number n can only take integral or half-integral values because of the periodicity condition in the angle ψ. A complete set of solutions of Eqs. (2.20)–(2.23) can be written in the form $|m, j, n, \alpha; \mathbf{p}, \sigma\rangle$, where $\mathbf{p}$ and σ are eigenvalues of momentum and the third component of spin, respectively.

Two models have been built on such a scheme. In the first one,[9,10] a field describes a particle with a given mass and spin, but the field functions also depend on the new quantum numbers α and n. These may turn out to be connected to known elementary-particle quantum numbers or merely to be phenomenological parameters.

In Ref. 6, a slightly different aspect is discussed: One thinks of α and n not as arbitrary numbers but fixed by some natural condition. In fact, the wavefunction is required to be analytic in the ξ variables. In other words, (2.22) and (2.23) are replaced by

$$\frac{\partial f}{\partial\xi^{\dot{1}}} = \frac{\partial f}{\partial\xi^{\dot{2}}} = 0. \tag{2.24}$$

Using (2.18) and (2.21), one readily sees that D coincides with the spin operator. Equations (2.24) are equivalently written as

$$\frac{\partial f}{\partial s} = -i\frac{\partial f}{\partial\psi} = jf, \tag{2.25}$$

where j is the spin of the particle described by the wavefunction. Consequently, this theory appears as a particular case of the first one, choosing $n = \alpha = j$. The fundamental physical difference between the two models is that there is no place for internal quantum numbers in the second one.

As far as we are concerned with free particles, we must explicitly build the complete set of kets $|m, j, \alpha, n; \mathbf{p}, \sigma\rangle$ as functions of x and ξ and define an invariant scalar product. Any wavefunction $f(x, \xi)$ of a single

particle in the Wigner sense is a solution of the four Eqs. (2.20)–(2.23):

(a) $f(x, \xi)$ satisfies the Klein–Gordon equation. It is, therefore, convenient to introduce the Fourier transform $\tilde{f}(p, \xi)$. The Klein–Gordon equation requires $\tilde{f}$ to have its support on the mass shell. [Note that $\tilde{f}(p, \xi)$ is not a scalar function under the Poincaré group since under a translation it is multiplied by a phase factor. Moreover it is not a function on a homogeneous space since the quantity $p_{\alpha\dot{\beta}}\xi^{\alpha}\xi^{\dot{\beta}}$ is an invariant. The homogeneous space defined by the couples (p, ξ) with the restriction $p_{\alpha\dot{\beta}}\xi^{\alpha}\xi^{\dot{\beta}} = m^2\gamma$, where γ is a positive constant, is isomorphic to the group $SL(2, c)$.]

(b) Consider now the equation involving the $W_\mu W^\mu$ operator. It is convenient to "boost the ξ's to the rest frame." Let Λ_p be any transformation which maps a given four momentum p of mass m on the vector $(m, \mathbf{0})$ and let Λ_p act on ξ,

$$\xi'^{\alpha} = (\Lambda_p)^{\alpha}_{\beta}\xi^{\beta}. \tag{2.26}$$

Since Λ_p acts linearly on the spinor space and is a unimodular transformation, the operator D is unchanged. On the other hand, after this transformation, only $p_{1\dot{1}}$ and $p_{2\dot{2}}$ are nonvanishing. Consequently, the operator $W_\mu W^\mu$ of Eq. (2.18) becomes

$$W_\mu W^\mu = -m^2\left[D(D+1) - (|\xi'^1|^2 + |\xi'^2|^2) \times \left(\frac{\partial^2}{\partial\xi'^1\partial\xi'^{\dot{1}}} + \frac{\partial^2}{\partial\xi'^2\partial\xi'^{\dot{2}}}\right)\right] \tag{2.27}$$

or, equivalently, making use of the other parametrization,

$$W_\mu W^\mu = m^2\left[\frac{1}{\sin\theta'}\frac{\partial}{\partial\theta'}\left(\sin\theta'\frac{\partial}{\partial\theta'}\right) + \frac{1}{\sin^2\theta'} \times \left(\frac{\partial^2}{\partial\psi'^2} + \frac{\partial^2}{\partial\varphi'^2} - 2\cos\theta'\frac{\partial^2}{\partial\varphi'\partial\psi'}\right)\right]. \tag{2.28}$$

Equation (2.21) then shows that $\tilde{f}(p, s', \varphi', \theta', \psi')$ is a linear combination of Wigner functions $D^j_{\sigma n}(\psi', \theta', \varphi')$.[16] They also satisfy

$$-i\frac{\partial}{\partial\varphi'}D^j_{\sigma n} = \sigma D^j_{\sigma n}, \tag{2.29}$$

$$-i\frac{\partial}{\partial\psi'}D^j_{\sigma n} = nD^j_{\sigma n}. \tag{2.30}$$

(c) The wavefunction $\tilde{f}$ must be an eigenfunction of the operator $\Delta = -i(\partial/\partial\psi)$. This operator commutes with the Lorentz group and, consequently, it is not affected by the transformation (2.26). Therefore the equation $\Delta\tilde{f} = n\tilde{f}$ is given by (2.30).

(d) The wavefunction must be an eigenfunction of the dilatation operator D. Therefore, finally, one is led to the following expression for the kets (not yet normalized):

$$|m, j, \alpha, n; \mathbf{p}, \sigma\rangle = e^{\alpha s'}\delta(\mathbf{p} - \mathbf{q})D^j_{\sigma n}(\psi', \theta', \varphi'). \tag{2.31}$$

Let us look for a scalar product of two wavefunctions corresponding to the same mass. It must involve an integration over the p and ξ variables. The transformation (2.26) being unimodular, one has $d^4\xi = d^4\xi'$. In order to get a nondivergent scalar product, one can choose, as was done both in Refs. 5 and 6, the following definition:

$$(f, g) = \iint \frac{d^3p}{p_0}\, d^4\xi\,\delta(p_{\alpha\dot{\beta}}\xi^{\alpha}\xi^{\dot{\beta}} - m^2\gamma)\tilde{f}^*(p, \xi)\tilde{g}(p, \xi) \tag{2.32}$$

or, equivalently, using the coset variables $(s, \theta, \psi, \varphi)$,

$$(f, g) = \iiint \frac{d^3p}{p_0}\, e^{2s}\, ds\, d\Omega\,\delta(p\cdot k - m^2\gamma) \cdot \tilde{f}^*(p, s, \psi, \theta, \varphi)\tilde{g}(p, s, \psi, \theta, \varphi), \tag{2.33}$$

where k is the vector (2.19) and $d\Omega$ is the measure $\sin\theta\, d\theta\, d\varphi\, d\psi$. Since k is a future lightlike vector and p a future timelike one, the scalar product $(p\cdot k)$ is positive. Consequently, the constant γ must be positive. If we introduce the following vector κ collinear to k,

$$\kappa = (1, \sin\theta\cos\varphi, -\sin\theta\sin\varphi, \cos\theta) = 2e^{-s}k, \tag{2.34}$$

and if we perform the integration in s, we obtain

$$(f, g) = \gamma\iint \frac{d^3p}{p_0}\frac{d\Omega}{(p\cdot\kappa/m)^2}\tilde{f}^*\tilde{g}, \tag{2.35}$$

a formula which can be used to normalize the states (2.37).

Case B: Spinless Particles. This case is very close to the preceding one. The Eqs. (2.20)–(2.23) are kept (with $j = 0$). The Wigner functions $D^j_{\sigma n}$ for $j = 0$ are constants and $\sigma = n = 0$. This means that the $W_\mu W^\mu$ equation implies

$$\frac{\partial\tilde{f}}{\partial\varphi'} = \frac{\partial\tilde{f}}{\partial\theta'} = \frac{\partial\tilde{f}}{\partial\psi'} = 0.$$

The only extra quantum number left is α. For a more detailed discussion, see Refs. 5 and 9.

Case C: Massless Particles. We are interested only in the physical case where $W_\mu W^\mu = 0$. Looking at Eq.

[16] See, for instance, A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton University Press, Princeton, N.J., 1957), Chap. 4.

(2.8), one immediately sees that this equality is satisfied if one imposes the analyticity condition. This means that the wavefunctions do not depend on ξ^1 or ξ^2. However, this property is in itself not sufficient to guarantee that we have the physical case, since the requirement of having normalizable vectors may put further restrictions on the wavefunctions. Instead we shall start from the equation

$$W^\mu = \Sigma P^\mu, \tag{2.36}$$

where Σ is the helicity operator which has a given eigenvalue λ in an irreducible representation. In the special frame where $p^{22} = p^{12} = p^{21} = 0$ ($p^1 = p^2 = p^0 - p^3 = 0$), the helicity operator coincides with S^{12}, as can be seen from Eq. (2.17). It is always possible by performing a Lorentz transformation Λ (for instance, a rotation) to make the momentum satisfy the conditions $p^{22} = p^{12} = p^{21} = 0$. Such a transformation maps the spinor ξ into another spinor ξ'.

The little group associated with a massless particle is isomorphic to the 2-dimensional Euclidean group $E(2)$. The generators of $E(2)$ corresponding to our particular momentum are

$$\begin{aligned} \Sigma &= S^{12} = \frac{1}{2}\left(\xi'^1 \frac{\partial}{\partial \xi'^1} - \xi'^2 \frac{\partial}{\partial \xi'^2} - \text{c.c.}\right), \\ A^1 &= -i\left(\xi'^2 \frac{\partial}{\partial \xi'^1} + \text{c.c.}\right), \\ A^2 &= \left(\xi'^2 \frac{\partial}{\partial \xi'^1} - \text{c.c.}\right). \end{aligned} \tag{2.37}$$

In the physical case, A^1 and A^2 vanish. Therefore our wavefunctions do not depend on the ξ'^1 and $\xi'^{\dot 1}$ variables and the helicity operator Σ becomes

$$\Sigma = \frac{-1}{2}\left(\xi'^2 \frac{\partial}{\partial \xi'^2} - \text{c.c.}\right), \tag{2.38}$$

which coincides with the operator $-\Delta$. Since Σ is an invariant operator, the quantum number $-n$ coincides with the helicity λ.

Any wavefunction which is an eigenfunction of D must be of degree α in ξ'^2 and $\xi'^{\dot 2}$. As an eigenfunction of Σ, it must be of the form

$$\tilde f(p, \xi') = a(\mathbf{p})(\xi'^{\dot 1})^{\alpha-\lambda}(\xi'^{\dot 2})^{\alpha+\lambda}. \tag{2.39}$$

Therefore a complete set of states will be given by the kets $|\lambda, \alpha; \mathbf{p}\rangle$. Only one value for the helicity is permitted for an irreducible representation since we are not interested in parity transformations. Before looking for a scalar product for our wavefunctions, it is natural to write them in terms of the coset variables $(s, \psi, \theta, \varphi)$. One finds, up to a normalization factor,

$$|\lambda, \alpha; \mathbf{p}\rangle = e^{\alpha s'}(1 - \cos\theta')^\alpha e^{i\lambda(\varphi' - \psi')}\delta(\mathbf{p} - \mathbf{q}). \tag{2.40}$$

As in Case A, the analyticity condition requires the equality $\alpha = -\lambda$.

In building the scalar product of two wavefunctions, we cannot introduce the δ function already used in the massive case, namely

$$\delta(p_{\alpha\beta}\xi^\alpha\xi^\beta - m^2\gamma) = \delta(p \cdot k - m^2\gamma),$$

since $p \cdot k$ is now a product of two lightlike vectors which vanishes when p becomes parallel to k. Fortunately, there exists another invariant δ function, namely $\delta^4(p - k)$, and we define the following scalar product:

$$(f, g) = \int\int d^4p\, d^4\xi \delta^4(p - k)\tilde f^*(p, \xi)\tilde g(p, \xi). \tag{2.41}$$

The use of this δ function suggests to us a new kind of model for the wavefunction of a massless particle, which does not satisfy our basic requirement but which seems to be interesting enough to be mentioned with some details in this paper. Let us consider functions on a spinor space. Such a function can always be expanded in terms of Wigner functions as follows:

$$f(\xi) = \sum_{j,\sigma,\lambda} f^j_{\sigma\lambda}(s) D^j_{\sigma\lambda}(\psi, \theta, \varphi). \tag{2.42}$$

Now let us make the Poincaré group act on these functions as follows. We describe the action for the infinitesimal generators. Put first

$$P^{\alpha\dot\beta} = \xi^\alpha \xi^{\dot\beta}, \tag{2.43}$$

i.e., the momenta are multiplicative operators. The $M^{\mu\nu}$'s are chosen to be the $S^{\mu\nu}$'s given by Eqs. (2.6). They are no longer split into an orbital and a spin part. What kind of representations do we get? Obviously $P_\mu P^\mu$ is zero. Using (2.17), we compute easily the components of W^μ. One gets

$$W^{\alpha\dot\beta} = \xi^\alpha \xi^{\dot\beta}\Delta, \tag{2.44}$$

where $\Delta = -i(\partial/\partial\psi)$. Consequently, only physical particles are obtained. The only nontrivial invariant is the helicity operator Δ. Any wavefunction of helicity λ will be given by

$$f_\lambda(\xi) = \sum_{j,\sigma} f^j_\sigma(s) D^j_{\sigma\lambda}(\psi, \theta, \varphi) \tag{2.45}$$

or equivalently

$$f_\lambda(\xi) = f(p)e^{i\lambda\psi}. \tag{2.46}$$

The "usual" wavefunction on Minkowski space is obtained by Fourier transformation,

$$f_\lambda(x) = \frac{1}{(2\pi)^3}\int d^4\xi e^{i\xi^\alpha\xi^\beta x_{\alpha\beta}} e^{-i\lambda\psi} f_\lambda(\xi). \tag{2.47}$$

It is natural to look for an analogous treatment of massive particles. Since it is not possible to build a timelike four-momentum from only one spinor, one must take a larger homogeneous space. One possibility consists in taking the Lorentz group itself. A complete set of functions on this space has been given by Ström,[17] in terms of matrix elements $D^{1\varphi\gamma}_{11'mm'}(\varphi, \theta, \alpha, \psi, \tilde{\theta}, \tilde{\varphi})$. (Ström's angles β and γ are here replaced by $\tilde{\theta}$ and $\tilde{\varphi}$, respectively.) The Lorentz group is acting on this set of functions according to the generators given in the appendix of Ref. 17. If we choose the translation generators P^μ to be the following multiplicative operator :

$$P^\mu = m(\cosh\alpha, \sinh\alpha \sin\theta \sin\varphi, \sinh\alpha \sin\theta \cos\varphi, \sinh\alpha \cos\theta), \quad (2.48)$$

where the mass m is given, one has a complete prescription for the action of the Poincaré group. By computing the helicity operators and the spin operators, one gets

$$\Sigma = \frac{\frac{1}{2}\epsilon_{ijk}P^i M^{jk}}{|\mathbf{P}|} = i\frac{\partial}{\partial\psi}, \quad (2.49)$$

$$W_\mu W^\mu = m^2\left[\frac{1}{\sin\tilde{\theta}}\frac{\partial}{\partial\tilde{\theta}}\left(\sin\tilde{\theta}\frac{\partial}{\partial\tilde{\theta}}\right) + \frac{1}{\sin^2\tilde{\theta}}\left(\frac{\partial^2}{\partial\psi^2} + \frac{\partial^2}{\partial\tilde{\varphi}^2} - 2\cos\tilde{\theta}\frac{\partial^2}{\partial\psi\partial\tilde{\varphi}}\right)\right]. \quad (2.50)$$

This last operator is identical to the Wigner operator (2.28). A complete set of invariant operators is given by $P_\mu P^\mu$, $W_\mu W^\mu$, $i\,\partial/\partial\psi$, the eigenvalues of which are m^2, $-m^2 j(j+1)$, and n, respectively. Therefore the wavefunctions of an elementary particle characterized by m, j, n are of the form

$$f^{m,j}_n(p, \psi, \tilde{\theta}, \tilde{\varphi}) = \sum_\lambda f^{m,j}_\lambda(p) D^j_{\lambda n}(\tilde{\varphi}, \tilde{\theta}, \psi), \quad (2.51)$$

where λ is the eigenvalue of the helicity operator.

It is interesting to note that these wavefunctions are closely related to those one obtains by using the Mackey method of induced representations.[18] This remark is not too surprising. All ways of building wavefunctions for elementary particles are equivalent since we are concerned with the kinematics of free particles only. This is in fact a consequence of the Wigner definition of an elementary particle as an irreducible representation of the Poincaré group. Nevertheless, the explicit way of constructing wavefunctions may have important consequences when interactions are introduced. Different kinematical frameworks may suggest different dynamical postulates. In order to test our model, it is therefore necessary to look for dynamical schemes. An attempt in this direction was made in Ref. 11. We hope to continue these investigations.

ACKNOWLEDGMENTS

One of the authors (H. B.) wants to acknowledge Nordita for financial support and Professor N. Svartholm for the hospitality extended to him at the Institute of Theoretical Physics in Göteborg.

[17] S. Ström, Arkiv Fysik **29**, 467 (1965).

[18] G. Mackey, *The Theory of Group Representations*, Lecture Notes, Univ. of Chicago, 1955; P. Moussa and R. Stora, *Lectures in Theoretical Physics*, W. E. Brittin and A. O. Barut, Eds. (Univ. of Colorado Press, Boulder, Col., 1965), Vol. 7A.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 12 DECEMBER 1969

Treatment of Degeneracies in the Schrödinger Perturbation Theory by Partitioning Technique*

JONG H. CHOI
Department of Chemistry, University of Alberta, Edmonton, Alberta

(Received 10 February 1969)

The degenerate case in the Schrödinger perturbation theory has been treated by use of the partitioning technique developed by Löwdin. In order to simplify the concept and treatment, the *repeated partitioning technique* is utilized. This repeated partitioning allows us to use a one-dimensional reference space and to determine the correct zero-order wavefunction ϕ_A.

1. INTRODUCTION

One of the convenient ways of solving the Schrödinger equation is to use a partitioning process.[1,2] The principal concept embedded in the partitioning technique is that one may first concentrate on a particular subspace, called a reference space. From this reference space one obtains eigenvectors by way of an eigenoperator[2] and the eigenvalues of a given Hamiltonian. For this purpose, we can use an iteration method or an expansion method.

The method of expansion is convenient when the expansion converges rapidly. The choice of the reference space $S(O)$ for this expansion is arbitrary so long as the eigenvector has some component in this reference space. Let us consider a system whose Hamiltonian operator $\mathcal{H} = \mathcal{H}^0 + V$ is only slightly different from the Hamiltonian operator $\mathcal{H}^0$ of some problem which has already been solved. In this case, the method of expansion is attractive if we take our reference subspace $S(O)$ to be that composed of all eigenvectors of $\mathcal{H}^0$ whose eigenvalue is E^0. Then the eigenoperator reduces to a wave operator and the energy is expressed in terms of the reaction operator.[2]

Löwdin has given the relations between various types of perturbation and his partitioning technique in a series of papers on perturbation theory.[1-3] His treatment of the Schrödinger perturbation theory has been primarily for nondegenerate cases.

In this paper, the degenerate case of $\mathcal{H}^0$ is considered mainly for the Schrödinger perturbation theory.[4] There are two principal types of perturbation, Schrödinger and Brillouin.[5] Even if their expanded forms are different from each other beyond the first order, their accuracies in terms of the order of perturbation V are the same.

The Schrödinger perturbation scheme is manipulated in such a way that the reduced resolvent[2] R_0 should always exist. Therefore, the *reference space* $S(O)$ should include all[6] eigenvectors Ψ_i^0 of $\mathcal{H}^0$ with an eigenvalue E^0 so that the reduced resolvent R_0 exists in the *complementary space* $S(P)$. Since $S(O)$ is multidimensional in our consideration, it is necessary, first of all, to obtain the zero-order eigenvector ϕ_A,

$$\phi_A = \sum_{i=1}^{g} C_i \Psi_i^0, \tag{1.1}$$

where the C_i are constants and

$$\mathcal{H}^0 \Psi_i^0 = E^0 \Psi_i^0, \quad i = 1, 2, \cdots, g, \tag{1.2}$$

$$\langle \phi_A \mid \phi_A \rangle = 1. \tag{1.3}$$

The constants C_i are to be determined by considering perturbed terms in the Hamiltonian. In order to determine C_i uniquely except for a phase factor, it is necessary that the degeneracy is completely resolved for the branch we are considering. The problem of obtaining ϕ_A is conveniently attacked by the introduction of the *repeated partitioning*. We perform the *repeated partitioning* until we are able to determine C_i uniquely.

The process of this *repeated partitioning* is described in Sec. 2 and its applications to degenerate cases of the Schrödinger perturbation theory is given in Sec. 4 with accompanying examples in Secs. 4 and 5.

An advantage of using the partitioning technique in perturbation theory is the flexibility of our manipulation and the use of simple notations. This aspect is most outstanding in the treatment of degeneracy, as

* Aided by Research grants to the University of Florida from the National Science Foundation and in part assisted by grants from the National Research Council of Canada.

[1] P.-O. Löwdin, J. Mol. Spectr. **10**, 12 (1963).
[2] P.-O. Löwdin, J. Math. Phys. **3**, 969 (1962).
[3] P.-O. Löwdin, J. Chem. Phys. **19**, 1396 (1951).
[4] E. Schrödinger, Ann. Physik (4) **80**, 437 (1928).
[5] L. Brillouin, J. Phys. Radium **33**, 373 (1932).

[6] In the definition of O, we do not often have to introduce all the Ψ_i^0 associated with the eigenvalue E^0 of $\mathcal{H}^0$. By considering the symmetry of $\mathcal{H}^0$ and $\mathcal{H}$, we could treat the problem in a subspace, the dimension of which is h, where $h \leq g$. As an example, we can consider the Stark effect in the excited state of the hydrogen atom.

can be seen from the expressions in Sec. 4 where the wavefunction is given correct up to third order.

2. REPEATED PARTITIONING PROCESS

For procedural convenience, we first introduce the *partitioning technique* developed by Löwdin.[1,2]

Let us divide the Hilbert space in which we are interested into two subspaces $S(O)$ and $S(P)$, associated with the Hermitian projection operators O and P, respectively. O and P have the following properties:

$$\begin{gathered}\mathbf{1} = O + P, \quad OP = PO = 0, \\ O^2 = O = O^\dagger, \quad P^2 = P = P^\dagger, \\ \mathrm{Tr}\,(O) = h,\end{gathered} \tag{2.1}$$

where $\mathbf{1}$ is the identity operator for the Hilbert space we are considering.

If we are interested in particular symmetry with respect to $\mathcal{K}$, then we need only concentrate on the subspace associated with this symmetry so that $S(O)$ and $S(P)$ together form this particular subspace.

We consider a Schrödinger equation

$$\mathcal{K}\Psi = E\Psi. \tag{2.2}$$

If $O\Psi \equiv \phi = 0$, one can treat this problem in $S(P)$ only. However, if $O\Psi = \phi \neq 0$, then one can replace (2.2) by

$$EO\Psi = O(\mathcal{K} + \mathcal{K}T_{(E)}\mathcal{K})O\Psi, \tag{2.3}$$

where

$$T_{(E)} = P[\alpha \cdot O + P(E - \mathcal{K})P]^{-1}P, \tag{2.4a}$$

with $\alpha \neq 0$.

It is convenient to use a symbolic notation,

$$T = T_{(E)} = P/(E - \mathcal{K}). \tag{2.4b}$$

Denoting

$$\mathcal{K}_1 \equiv \mathcal{K} + \mathcal{K}T\mathcal{K}, \tag{2.5}$$

we can rewrite (2.3) as

$$E\phi = O\mathcal{K}_1 O\phi, \tag{2.6}$$

where

$$\phi \equiv O\Psi. \tag{2.7}$$

Since (2.2), which is an eigenvalue problem in an infinite-dimensional Hilbert space, is contracted to (2.6), which is an eigenvalue problem in a finite-dimensional space $S(O)$ with $\mathrm{Tr}\,(O) = h$, we shall call $S(O)$ an h-dimensional reference space. Equation (2.6) is equivalent to an h-dimensional secular equation.

A. Repeated Partitioning

An eigenvector Ψ of $\mathcal{K}$, with an eigenvalue E, satisfies (2.6) if $\phi \neq 0$. One of the problems left to us in the partitioning process, when $S(O)$ is multi-dimensional, is how to find ϕ. In principle, we could obtain ϕ by solving a secular equation equivalent to (2.6). However, before solving this secular equation we have to know the value of E; hence, some other method which circumvents this difficulty is necessary.

If there is any good reason for one to expect that the vector ϕ is heavily localized in a particular subspace, say $S(O_1)$ in $S(O)$, then it is convenient to apply the partitioning process again to partition $S(O)$ into two mutually orthogonal subspaces $S(O_1)$ and $S(P_1)$ and associate them with Hermitian projection operators O_1 and P_1, respectively, so that

$$\begin{gathered}O = O_1 + P_1, \quad O_1P_1 = 0, \\ O_1^2 = O_1^\dagger = O_1, \quad P_1^2 = P_1^\dagger = P_1.\end{gathered} \tag{2.8}$$

For $\mathrm{Tr}\,(O_1) \geq 2$, an occasion might arise that a further partitioning of $S(O_1)$ into mutually orthogonal $S(O_2)$ and $S(P_2)$ would facilitate the treatment of our problems. Projection operators associated with $S(O_2)$ and $S(P_2)$ are denoted by O_2 and P_2, respectively. They have the properties

$$\begin{gathered}O_1 = O_2 + P_2, \quad O_2P_2 = 0, \\ O_2^2 = O_2 = O_2^\dagger, \quad P_2^2 = P_2 = P_2^\dagger.\end{gathered} \tag{2.9}$$

For the sake of convenience, let us call the partitioning, which divides a Hilbert space into $S(O)$ and $S(P)$, the *unipartitioning*; the partitioning, which subdivides $S(O)$ into $S(O_1)$ and $S(P_1)$, we call the *bipartitioning*; and the partitioning, which again subdivides $S(O_1)$ into $S(O_2)$ and $S(P_2)$ we call *the tripartitioning*. We can repeat this process further. From now on, a bipartitioning, or a partitioning of higher degree than bipartitioning, will be called the *repeated partitioning*.

B. Bipartitioning

We now define a resolvent T_1 in $S(P_1)$ as

$$T_1 \equiv P_1[\beta \cdot (O_1 + P) + P_1(E - \mathcal{K}_1)P_1]^{-1}P_1, \tag{2.10}$$

where $\beta \neq 0$. Then by the bipartitioning we obtain, from (2.6),

$$EO_1\phi = O_1(\mathcal{K}_1 + \mathcal{K}_1T_1\mathcal{K}_1)O_1\phi, \tag{2.11}$$

where it is understood that $\mathcal{K}_1$ and T_1 involve the eigenvalue E. As in the case of T, we can denote T_1 symbolically as

$$T_1 \equiv \frac{P_1}{E - \mathcal{K}_1} = \frac{P_1}{E - (\mathcal{K} + \mathcal{K}T\mathcal{K})}. \tag{2.12}$$

Similar symbolic notations T_l, $l \geq 2$, will be used in the following for further partitionings, i.e.,

$$T_l = P_l/(E - \mathcal{K}_l), \tag{2.13}$$

where

$$\mathcal{H}_l = \mathcal{H}_{l-1} + \mathcal{H}_{l-1}T_{l-1}\mathcal{H}_{l-1}. \tag{2.14}$$

It is shown,[2] in the case of unipartitioning, that there exists an eigenoperator Ω satisfying the relations

$$\mathcal{H}\Omega = E\Omega, \quad \Omega^2 = \Omega, \tag{2.15}$$

provided that

$$EO = O\mathcal{H}_1 O = O\mathcal{H}\Omega. \tag{2.16}$$

In order to establish relations similar to (2.15) and (2.16) for the bipartitioning, it is convenient to define

$$\Omega_0 \equiv (1 + T\mathcal{H})O, \tag{2.17}$$

$$\Omega_1 \equiv (1 + T_1\mathcal{H}_1)O_1, \tag{2.18}$$

$$\Omega \equiv \Omega_0\Omega_1, \tag{2.19}$$

so that

$$\begin{aligned} &\Omega_0^2 = \Omega_0, \quad \Omega_1^2 = \Omega_1, \quad \Omega^2 = \Omega, \\ &O\Omega_1 = \Omega_1 O = \Omega_1, \\ &\Omega_1\Omega_0 = \Omega_1, \\ &O_1\Omega_1 = O_1. \end{aligned} \tag{2.20}$$

Then it can be shown that

$$\mathcal{H}\Omega = E\Omega, \tag{2.21}$$

provided that

$$E\Omega_1^\dagger\Omega_1 = \Omega_1^\dagger\mathcal{H}_1\Omega_1 \tag{2.22a}$$

or

$$\begin{aligned} EO_1\phi &= O_1\mathcal{H}_1\Omega_1\phi \\ &= O_1\mathcal{H}_2O_1\phi, \end{aligned} \tag{2.22b}$$

where

$$\mathcal{H}_2 = \mathcal{H}_1 + \mathcal{H}_1T_1\mathcal{H}_1, \tag{2.23}$$

according to (2.14).

Using the relations in (2.20) and (2.21), this is seen as the following:

$$\begin{aligned} E\Omega_0\Omega_1 &= \mathcal{H}\Omega_0\Omega_1 \\ &= \mathcal{H}(1 + T\mathcal{H})O(1 + T_1\mathcal{H}_1)O_1 \\ &= \mathcal{H}(1 + T\mathcal{H})(1 + T_1\mathcal{H})O_1 \\ &= \mathcal{H}_1(1 + T_1\mathcal{H}_1)O_1 \\ &= \mathcal{H}_1\Omega_1. \end{aligned} \tag{2.24}$$

Multiplying $\Omega_1^\dagger$ or O_1 from the left on both sides of (2.24) we obtain (2.22).

It is important to note here that the condition expressed in (2.22) not only determines the eigenvalue E, but it also determines a correct vector $O_1\Psi$. If the reference space on which the eigenoperator acts is multidimensional, an arbitrary vector in that reference space, in general, will not satisfy the relation (2.21) unless Tr $(O_1) = 1$.

C. Tripartitioning

Following the definitions in (2.13) and (2.14), for the case of the tripartitioning, one obtains

$$\begin{aligned} EO_2\Psi &= O_2(\mathcal{H}_2 + \mathcal{H}_2T_2\mathcal{H}_2)O_2\Psi \\ &= O_2\mathcal{H}_3O_2\Psi. \end{aligned} \tag{2.25}$$

The eigenoperator which satisfies the relation (2.21) can be defined as

$$\Omega \equiv \Omega_0\Omega_1\Omega_2, \tag{2.26}$$

where Ω_1 is given by (2.18) and

$$\Omega_2 = (1 + T_2\mathcal{H}_2)O_2. \tag{2.27}$$

Relations similar to (2.20) and (2.22) are satisfied for Ω_2.

The partitioning process could be repeated further until one finally obtains a one-dimensional reference vector. If the partitioning is performed repeatedly $n + 1$ times, then one can write, in general,

$$\Omega = \Omega_0\Omega_1 \cdots \Omega_n, \tag{2.28}$$

$$EO_n\Psi = O_n\mathcal{H}_{n+1}O_n\Psi \tag{2.29a}$$

or

$$E\Omega_n^\dagger\Omega_n = \Omega_n^\dagger\mathcal{H}_n\Omega_n. \tag{2.29b}$$

3. DEGENERATE PERTURBATION THEORY

We consider that the Hamiltonian is decomposable into two parts $\mathcal{H}^0$ and V, so that $\mathcal{H} = \mathcal{H}^0 + V$, where V is small and $\mathcal{H}^0$ has known eigenvectors and known eigenvalues. Let us consider a particular eigenvalue E^0 which has degeneracy of order g,

$$\mathcal{H}^0\Psi_i^0 = E^0\psi_i^0, \quad i = 1, \cdots, g, \tag{3.1}$$

where

$$\langle\Psi_i^0 \mid \Psi_j^0\rangle = \delta_{ij}. \tag{3.2}$$

For this case, it is convenient to define O in Sec. 2 by[6]

$$O = \sum_{i=1}^{h} |\Psi_i^0\rangle\langle\Psi_i^0|. \tag{3.3}$$

Then (2.6) reads

$$E\phi = O(E^0 + V + VTV)O\phi \tag{3.4}$$

or

$$\begin{aligned} E\phi &= O(E^0 + VW_0)O\phi \\ &\equiv O(E^0 + t_0)O\phi, \end{aligned} \tag{3.5}$$

where

$$W_0 = 1 + TV, \tag{3.6}$$

$$t_0 = VW_0. \tag{3.7}$$

The operator t_0 is called the *reaction operator* and W_0 is called the wave operator.[2]

We restrict ourselves to the case where V is small enough to allow the expansion[2]

$$E = E^0 + \sum_{i=1}^{\infty} E^{(i)} \tag{3.8}$$

and

$$T = R_0 \sum_{n=0}^{\infty} (V' R_0)^n, \tag{3.9}$$

where $E^{(i)}$ denotes the ith-order energy with respect to the power of V,

$$V' = V - \sum_{i=1}^{\infty} E^{(i)}, \tag{3.10}$$

$$R_0 = P/(E^0 - \mathcal{H}^0). \tag{3.11}$$

For simplicity, let us introduce the notation $\mathcal{H}^{(i)}$ i.e.,

$$\begin{aligned}
\mathcal{H}^{(0)} &= E^0, \\
\mathcal{H}^{(1)} &= E^0 + V, \\
\mathcal{H}^{(2)} &= \mathcal{H}^{(1)} + V R_0 V, \\
\mathcal{H}^{(3)} &= \mathcal{H}^{(2)} + [-E^{(1)} V R_0^2 V + V(R_0 V)^2], \\
\mathcal{H}^{(4)} &= \mathcal{H}^{(3)} + [-E^{(2)} V R_0^2 V + (E^{(1)})^2 V R_0^3 V \\
&\quad - E^{(1)}(V R_0 V R_0^2 V + V R_0^2 V R_0 V) + V(R_0 V)^3],
\end{aligned} \tag{3.12}$$

where $\mathcal{H}^{(i)}$ is an approximate operator of $\mathcal{H}_1$ correct to the ith order.

Under the assumption of the validity (3.8), one can see that

$$O\mathcal{H}_1 O = \lim_{i \to \infty} O\mathcal{H}^{(i)} O. \tag{3.13}$$

Sometimes the degeneracy remains[7] in all orders of V. If a degeneracy of order $f, f \leq h$, remains in all orders of V, one can reduce[1,2] the dimension of $S(O)$ so that $\mathrm{Tr}\,(O) = h - f + 1$. Hence, in the following process we can assume that the degeneracy is completely resolved at a certain order of perturbation.

In the application of the partitioning technique to the degenerate case of perturbation theory, we have first to decide in which branch of the splitting we are interested. Let us assume that we are interested in the branch A. We repeat the partitioning process until the degeneracy is completely resolved for the branch A. Once the degeneracy is completely resolved for the branch we are interested in, we are able to determine the zero-order wavefunction ϕ_A uniquely except for a possible phase factor. Thus the expression for the wavefunction and the energy depends on how many times we perform the partitioning process, that is to say, the *degree of partitioning*, and also on the degree of resolution at each stage of partitioning. The degree of partitioning depends on the resolution and persistency of the degeneracy of the branch A only. Even if the degree of partitioning is the same, the resultant expressions for the wavefunction and the energy vary according to the character of persistency. This property can be seen by comparing the results of cases 2 and 3 of our examples.

In order to facilitate the explanation, let us consider three different cases. We first introduce h mutually orthogonal vectors $\phi_1, \cdots, \phi_h$ in $S(O)$ which diagonalize an operator OVO so that

$$\langle \phi_i \mid \phi_j \rangle = \delta_{ij}, \quad V_{ij} = V_{ii}\delta_{ij}, \tag{3.14}$$

where

$$V_{ij} \equiv \langle \phi_i | \, V \, | \phi_j \rangle. \tag{3.15}$$

For convenience, it is assumed that ϕ_i is the zero-order eigenvector for the branch i. Before considering the examples, we identify ϕ_A in (1.1) with ϕ_1 and define projection operators Q_i as

$$\phi_A = \phi_1 \tag{3.16}$$

$$Q_i \equiv |\phi_i\rangle \langle \phi_i|. \tag{3.17}$$

Case 1: $V_{11} \neq V_{jj}$ for $j = 2, \cdots, h$: We take[8] $S(O_1)$ to be $S(Q_1)$ and apply the bipartitioning technique with respect to the reference vector $S(Q_1)$.

Since the degeneracy is completely resolved at the first order, the bipartitioning is sufficient for our purpose. This case is considered in Sec. 4.

Case 2: $V_{11} = V_{jj}$ for all $j = 2, \cdots, h$. For this case, the degeneracy persists at the first order for all branches, and the operator $O\mathcal{H}^{(2)}O$ has to be considered in order to determine ϕ_1. We assume the degeneracy for the branch A is completely resolved at the second order, then the bipartitioning technique gives us the desired expressions. For this case, it can be seen that

$$\begin{aligned}
&O_1 V P_1 = 0 \\
&\text{and} \\
&O_1 V R_0 V P_1 = 0.
\end{aligned} \tag{3.18}$$

The results of this case are given in Sec. 5.

Case 3: It is assumed that the degeneracy for the branch A is partially resolved at the first order by considering OVO and completely resolved at the second order by $O_1 V R_0 V O_1$. If the degree of remaining degeneracy after the application of unipartitioning process is m, one can denote

$$\begin{aligned}
O_1 &= \sum_{i=1}^{m} |\phi_i\rangle \langle \phi_i|, \quad P_1 = \sum_{i=m+1}^{h} |\phi_i\rangle \langle \phi_i|, \\
O_2 &= |\phi_1\rangle \langle \phi_1|, \qquad P_2 = \sum_{i=2}^{m} |\phi_i\rangle \langle \phi_i|,
\end{aligned} \tag{3.19}$$

so that

$$O = O_1 + P_1, \quad O_1 = O_2 + P_2. \tag{3.20}$$

[7] As an example, we can consider the Stark effect in the rigid rotator, where $h = f = 2$.

[8] It is understood in this paper that, if Q is a Hermitian projection operator with $\mathrm{Tr}\,(Q) = n$, the $S(Q)$ is an n-dimensional space associated with Q.

We perform the bipartitioning process with respect to the reference space $S(O_1)$, and then apply the tripartitioning technique with respect to the reference vector $S(O_2)$. The results of this case are given in Sec. 5.

In the partitioning technique, we have dealt with Hermitian projection operators which divide a Hilbert space into mutually orthogonal subspaces. Hence, it is convenient to have the intermediate normalization

$$\langle \phi_A | \phi_A \rangle = 1, \quad \langle \phi_A | \Psi \rangle = 1. \tag{3.21}$$

4. APPLICATION OF REPEATED PARTITIONING TO DEGENERATE PERTURBATION (CASE 1)

In this section, we consider the case where the degeneracy for the branch A is completely resolved at the first order of the perturbation V.

Since it is assumed that $\phi_A = \phi_1$, according to (3.17) and (2.8) one can write

$$O_1 = Q_1, \tag{4.1a}$$

$$P_1 = O - Q_1. \tag{4.1b}$$

Using the relations (2.22), (3.3), and (3.6), one can reduce (2.17), (2.18), (2.19), and (2.24) to

$$\Omega_0 = (1 + TV)O, \tag{4.2}$$

$$\Omega_1 = (1 + T_1VTV)O_1, \tag{4.3}$$

$$\Omega = (1 + TV)(1 + T_1VTV)O_1, \tag{4.4}$$

$$EO_1\phi_1 = [(E^0 + V) + (VTV) + (VTV)T_1(VTV)]O_1\phi_1, \tag{4.5}$$

where

$$\phi_1 = O_1\Psi/[\langle \Psi | O_1 | \Psi \rangle]^{\frac{1}{2}}. \tag{4.6}$$

After the definition of Ω in (2.19), it is natural to define a *wave operator* W for this case by

$$W \equiv W_0W_1, \tag{4.7}$$

where W_0 is given by (3.6) and

$$W_1 \equiv 1 + T_1(VTV). \tag{4.8}$$

Using the property $O_1VP_1 = 0$ and relations (3.6), (4.7), and (4.8), one can write (4.5) as

$$EO_1\phi_1 = O_1[E^0 + V(1 + TV) \times (1 + T_1VTV)]O_1\phi_1, \tag{4.9}$$

$$E\phi_1 = O_1[E^0 + VW]\phi_1 = O_1[E^0 + t]\phi_1, \tag{4.10}$$

where

$$t = VW. \tag{4.11}$$

We can compare (3.5), (4.5), and (4.9) as follows:

(a) The multidimensional reference space $S(O)$ is reduced to a one-dimensional reference vector $S(O_1)$. Thus we are able to define W which will give us an eigenvector of $\mathcal{H}$. We note here that W_0 by itself is not able to give us an eigenvector of $\mathcal{H}$.

(b) For (3.5), V is a perturbation with a reference space $S(O)$ consisting of eigenfunctions of $\mathcal{H}^0$ with the eigenvalue E^0, whereas, for (4.5), VTV is a perturbation with a reference space, or a reference vector, $S(O_1)$ consisting of an eigenfunction of $O\mathcal{H}O$ with the eigenvalue $(E^0 + V_{11})$.

(c) In (4.5), we have an additional term

$$(VTV)T_1(VTV)$$

which does not appear in the nondegenerate case.

(d) The operator W_1 determines a correct reference vector in $S(O)$ from ϕ_1, and W_0 determines an eigenvector Ψ from ϕ.

Eigenvector and Eigenvalue

Here, we introduce

$$\delta_m = \sum_{i=m}^{\infty} E^{(i)}, \tag{4.12}$$

where $E^{(i)}$ is the ith-order energy as can be seen from (3.8) and $E^{(1)} = V_{11}$ for the present consideration. Using the notations defined in (3.8) and (3.11), and using the operator identity

$$(A - B)^{-1} = A^{-1} + A^{-1}B(A - B)^{-1},$$

one can write

$$T = R_0 + R_0[(V - E^{(1)}) - \delta_2]R_0 + R_0[(V - E^{(1)}) - \delta_2]R_0[(V - E^{(1)}) - \delta_2]T \tag{4.13}$$

and

$$\delta_2 = E^{(2)} + \delta_3, \qquad \delta_i = E^{(i)} + \delta_{i+1}. \tag{4.14}$$

The relations in (4.13) and (4.14) give us

$$TV = R_0V + R_0(V - E^{(1)})R_0V + R_0(V - E^{(1)})R_0(V - E^{(1)})R_0V - E^{(2)}R_0^2V + O(V^4), \tag{4.15}$$

$$VTV = VR_0V + VR_0(V - E^{(1)})R_0V + VR_0(V - E^{(1)})R_0(V - E^{(1)})R_0V - E^{(2)}VR_0^2V + O(V^5), \tag{4.16}$$

$$\begin{aligned} T_1 = \frac{P_1}{E - \mathcal{H}_1} &= \frac{P_1}{E - E^0 - V - VTV} \\ &= \frac{P_1}{(E^{(1)} - V) - (VTV - \delta_2)} \\ &= R_1 + R_1(VTV - \delta_2)R_1 + R_1(VTV - \delta_2)R_1(VTV - \delta_2)T_1 \\ &= R_1 + R_1(VR_0V - E^{(2)})R_1 + R_1[VR_0(V - E^{(1)})R_0V - E^{(3)}]R_1 \\ &\quad + R_1(VR_0V - E^{(2)})R_1(VR_0V - E^{(2)})R_1 + O(V^2), \end{aligned} \tag{4.17}$$

where

$$R_1 \equiv \frac{P_1}{E^{(1)} - V} = \sum_{i=2}^{h} \frac{|\phi_i\rangle\langle\phi_i|}{V_{11} - V_{ii}}. \quad (4.18)$$

Introducing the relations in (4.15)–(4.18) into (4.7), one obtains

$$W = W^{(0)} + W^{(1)} + W^{(2)} + W^{(3)} + \mathcal{O}(V^4), \quad (4.19)$$

where the superscript i denotes the order of V,

$$\begin{aligned}
W^{(0)} &= 1,\\
W^{(1)} &= R_0V + R_1VR_0V,\\
W^{(2)} &= R_0(V - E^{(1)})R_0V + R_0VR_1VR_0V\\
&\quad + R_1VR_0(V - E^{(1)})R_0V\\
&\quad + R_1(VR_0V - E^{(2)})R_1VR_0V,\\
W^{(3)} &= R_0(V - E^{(1)})R_0(V - E^{(1)})R_0V\\
&\quad - E^{(2)}R_0^2V + R_0VR_1VR_0(V - E^{(1)})R_0V\\
&\quad + R_0(V - E^{(1)})VR_1VR_0V\\
&\quad + R_0VR_1(VR_0V - E^{(2)})R_1VR_0V\\
&\quad + R_1VR_0(V - E^{(1)})R_0(V - E^{(1)})R_0V\\
&\quad - E^{(2)}R_1VR_0^2V\\
&\quad + R_1(VR_0V - E^{(2)})R_1VR_0(V - E^{(1)})R_0V\\
&\quad + R_1[VR_0(V - E^{(1)})R_0V - E^{(3)}]R_1VR_0V\\
&\quad + R_1(VR_0V - E^{(2)})R_1(VR_0V - E^{(2)})R_1VR_0V.
\end{aligned} \quad (4.20)$$

Equation (4.10) can be written as

$$E = \langle\phi_1|\, E^0 + VW\,|\phi_1\rangle. \quad (4.21)$$

Introducing the results of (4.20) into (4.21), one obtains

$$E = E^0 + E^{(1)} + E^{(2)} + E^{(3)} + E^{(4)} + \cdots, \quad (4.22)$$

where

$$\begin{aligned}
E^{(1)} &= V_{11} = \langle V\rangle,\\
E^{(2)} &= \langle VR_0V\rangle,\\
E^{(3)} &= \langle VR_0(V - E^{(1)})R_0V\rangle + \langle VR_0VR_1VR_0V\rangle,\\
E^{(4)} &= \langle VR_0(V - E^{(1)})R_0(V - E^{(1)})R_0V\rangle\\
&\quad - E^{(2)}\langle VR_0^2V\rangle\\
&\quad + \langle VR_0VR_1VR_0(V - E^{(1)})R_0V\rangle\\
&\quad + \langle VR_0(V - E^{(1)})R_0VR_1VR_0V\rangle\\
&\quad + \langle VR_0VR_1(VR_0V - E^{(2)})R_1VR_0V\rangle,
\end{aligned} \quad (4.23)$$

with

$$\langle A\rangle = \langle\phi_1|\, A\,|\phi_1\rangle. \quad (4.24)$$

For the nondegenerate case $h = 1$, $R_1 = 0$, every term which includes R_1 in the expression for the wave operator and energy drops out and the resultant expressions are identical with those of the nondegenerate case given by Löwdin.[2]

5. EIGENVECTORS AND EIGENVALUES FOR CASES 2 AND 3

Case 2: In this case the relations (3.14) and (3.15) are automatically satisfied for any choice of orthonormal set $\{\phi_i\}$ in $S(O)$. The set $\{\phi_i\}$ are chosen to satisfy the relations

$$\langle\phi_i|\, VR_0V\,|\phi_j\rangle = \langle\phi_i|\, VR_0V\,|\phi_i\rangle\,\delta_{ij}. \quad (5.1)$$

Here we can use the same relations from (4.1) through (4.15) as developed in Sec. 4. However, the expression for T_1 is different from Case 1; now,

$$\begin{aligned}
T_1 &= S_1 + S_1[VR_0(V - E^{(1)})R_0V - E^{(3)}]S_1\\
&\quad + S_1[VR_0(V - E^{(1)})R_0(V - E^{(1)})R_0V\\
&\quad - E^{(2)}VR_0^2V - E^{(4)}]S_1\\
&\quad + S_1[VR_0(V - E^{(1)})R_0V - E^{(3)}]\\
&\quad \times S_1[VR_0(V - E^{(1)})R_0V - E^{(3)}]S_1 + \mathcal{O}(V),
\end{aligned} \quad (5.2)$$

where

$$S_1 \equiv \frac{P_1}{E^{(2)} - VR_0V} = \sum_{i=2}^{h} \frac{|\phi_i\rangle\langle\phi_i|}{E^{(2)} - \langle\phi_i|\, VR_0V\,|\phi_i\rangle}. \quad (5.3)$$

Using (3.18), the expanded forms for the wave operator W and the eigenvalue E are

$$\begin{aligned}
W &= W_0W_1\\
&= W^{(0)} + W^{(1)} + W^{(2)} + \mathcal{O}(V^3),
\end{aligned} \quad (5.4)$$

where

$$\begin{aligned}
W^{(0)} &= 1,\\
W^{(1)} &= R_0V + S_1VR_0(V - E^{(1)})R_0V,\\
W^{(2)} &= R_0(V - E^{(1)})R_0V + R_0VS_1VR_0(V - E^{(1)})R_0V\\
&\quad + S_1[VR_0(V - E^{(1)})R_0V - E^{(3)}]\\
&\quad \times S_1VR_0(V - E^{(1)})R_0V\\
&\quad + S_1[VR_0(V - E^{(1)})R_0(V - E^{(1)})R_0V\\
&\quad - E^{(2)}VR_0^2V],
\end{aligned} \quad (5.5)$$

$$\begin{aligned}
E &= \langle\phi_1|\, E^0 + VW\,|\phi_1\rangle\\
&= E^0 + E^{(1)} + E^{(2)} + E^{(3)} + \mathcal{O}(V^4),
\end{aligned} \quad (5.6)$$

where, using the property (3.18),

$$\begin{aligned}
E^{(1)} &= \langle V\rangle,\\
E^{(2)} &= \langle VR_0V\rangle\\
E^{(3)} &= \langle VR_0(V - E^{(1)})R_0V\rangle,
\end{aligned} \quad (5.7)$$

with

$$\langle A\rangle = \langle\phi_1|\, A\,|\phi_1\rangle.$$

Case 3: For this case, since we use the tripartitioning process, Ω is given by (2.26); accordingly, the wave operator has the form

$$\begin{aligned} W &= W_0 W_1 W_2 \\ &= [1 + TV][1 + T_1(VTV)] \\ &\quad \times [1 + T_2\{(VTV) + (VTV)T_1(VTV)\}], \end{aligned} \tag{5.8}$$

where W, W_1, R_0, and R_1 are given by (3.6), (4.8), (3.11), and (4.18), respectively, and

$$W_2 = 1 + T_2[(VTV)T_1(VTV)], \tag{5.9}$$

$$\begin{aligned} T_2 = R_2 &+ R_2[VR_0(V - E^{(1)})R_0V \\ &+ VR_0VR_1VR_0V - E^{(3)}]R_2 + \mathcal{O}(V^0), \end{aligned} \tag{5.10}$$

$$R_2 = P_2/(E^{(2)} - VR_0V). \tag{5.11}$$

Using the relations in (3.18), we expand W according to the power of V,

$$W = W^{(0)} + W^{(1)} + W^{(2)} + \mathcal{O}(V^3), \tag{5.12}$$

where

$$\begin{aligned} W^{(0)} &= 1, \\ W^{(1)} &= R_0V + R_1(VR_0V) + R_2(VR_0V)R_1(VR_0V) \\ &\quad + R_2VR_0(V - E^{(1)})R_0V, \\ W^{(2)} &= R_0(V - E^{(1)})R_0V + R_0VR_1(VR_0V) \\ &\quad + [R_0V + R_1VR_0V]R_2[(VR_0V)R_1(VR_0V) \\ &\quad + VR_0(V - E^{(1)})R_0V] \\ &\quad + R_1VR_0(V - E^{(1)})R_0V \\ &\quad + R_1(VR_0V - E^{(2)})R_1VR_0V \\ &\quad + R_2(VR_0V)R_1VR_0(V - E^{(1)})R_0V \\ &\quad + R_2VR_0(V - E^{(1)})R_0VR_1(VR_0V) \\ &\quad + R_2(VR_0V)R_1(VR_0V - E^{(2)})R_1(VR_0V) \\ &\quad + R_2VR_0(V - E^{(1)})R_0(V - E^{(1)})R_0V \\ &\quad - E^{(2)}R_2VR_0^2V \\ &\quad + R_2[VR_0(V - E^{(1)})R_0V - E^{(3)}] \\ &\quad \times R_2[(VR_0V)R_1(VR_0V) \\ &\quad + VR_0(V - E^{(1)})R_0V]. \end{aligned} \tag{5.13}$$

Hence,

$$\begin{aligned} E &= \langle\phi_1|\, E^0 + VW\,|\phi_1\rangle \\ &= E^0 + E^{(1)} + E^{(2)} + E^{(3)} + \mathcal{O}(V^4), \end{aligned} \tag{5.14}$$

where

$$\begin{aligned} E^{(1)} &= \langle V\rangle, \\ E^{(2)} &= \langle VR_0V\rangle, \\ E^{(3)} &= \langle VR_0(V - E^{(1)})R_0V\rangle + \langle(VR_0V)R_1(VR_0V)\rangle, \end{aligned} \tag{5.15}$$

with

$$\langle A\rangle = \langle\phi_1|\, A\,|\phi_1\rangle.$$

6. DISCUSSION

If the reference space $S(O)$ is multidimensional, we have to first decide on a proper reference vector ϕ in $S(O)$ since any arbitrary vector in $S(O)$ will not satisfy the condition given by (2.15). This problem is technically solved by introducing the repeated partitioning and accompanying *subeigenoperators* $\Omega_1, \Omega_2, \cdots$, etc. In the Schrödinger perturbation theory, we take $S(O)$ to be the space composed of all eigenvectors of $\mathcal{H}^0$ whose eigenvalues are E^0 in order to eliminate the singularity in R_0. Usually, we can reduce the dimension of $S(O)$ by considering the symmetry properties of $\mathcal{H}^0$ and $\mathcal{H}$. The repeated partitioning is performed in such a way that the resolvents R_1, R_2, S_1, etc., exist. The most important problem is to find a zero-order eigenvector ϕ_A. This becomes complicated if the degeneracy is not resolved at the first order. We have obtained the expressions for the eigenvector and the energy by simply expanding the wave operator, whereas in the conventional Schrödinger perturbation treatment one has to deal with the equations of various orders.[9,10] The treatment and notations we have used are rather compact which is a characteristic of the partitioning technique. One can treat degenerate perturbation theory in a multidimensional reference space $S(O)$ without resorting to the repeated partitioning technique.[11] However, the use of the latter process is more simple and convenient.

ACKNOWLEDGMENTS

The author would like to express his gratitude to Professor Per-Olov Löwdin for many stimulating discussions. The author is most grateful to Professor Darwin W. Smith and Professor Fraser W. Birss for helpful comments.

[9] J. O. Hirschfelder, W. Byers Brown, and S. T. Epstein, Advan. Quant. Chem. **1**, 256 (1964).
[10] A. Dalgarno, Quant. Theory **1**, 197 (1961).
[11] Jong H. Choi, Preprint No. 89, Quantum Theory Project, University of Florida, 1966.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 12 DECEMBER 1969

Quasiparticle Formalism and Atomic Shell Theory*

M. J. CUNNINGHAM AND B. G. WYBOURNE
Physics Department, University of Canterbury, Christchurch, New Zealand

(Received 29 April 1969)

The quasiparticle formalism developed by Armstrong and Judd for atomic shells is extended to expose the complete group structure of the quasiparticle eigenfunctions of the equivalent electron l shell. A simple method for relating quasiparticle states to determinantal states and for calculating quasiparticle matrix elements is developed. The need for fractional parentage coefficients in calculating these matrix elements is entirely eliminated.

I. INTRODUCTION

The theory of continuous groups has been used extensively both to classify eigenfunctions and interactions.[1–4] The orthonormal set of antisymmetrized eigenfunctions associated with an equivalent electron l^N span the $\{1^N\}$ representation of the unitary group U_{4l+2}. The members of the set of eigenfunctions may be classified by considering their transformation properties under the operations of the various subgroups of U_{4l+2}. The chain of groups

$$\begin{array}{ccc} & \nearrow SU_2 \times SU_{2l+1} & \\ U_{4l+2} & & \searrow SU_2 \times (R_{2l+1} \to R_3) \\ & \searrow Sp_{4l+2} & \nearrow \end{array} \qquad (1)$$

may be used to distinguish completely the eigenfunctions in the d-shells ($l = 2$). If the exceptional Lie group G_2 is embedded in R_7, an almost complete classification of the eigenfunctions of the f shell is possible.[1] For higher values of l, the above classification scheme rapidly proves to give an inadequate number of classification symbols to distinguish eigenfunctions associated with the same SL quantum numbers. Thus, in the t^{13} configuration, no distinction is possible among the 30 598 ^{14}E states.

The chain of groups given in Eq. (1) sheds little light on relationships between the properties of configurations involving different numbers of electrons. This has led Judd[4] to consider the various subgroups of the group $U_{2^{4l+2}}$, which comprises the group of all unitary transformations among the 2^{4l+2} multielectron states of the l shell. He has shown that the chain of groups given in Eq. (1) may be replaced by the chain

$$U_{2^{4l+2}} \to R_{8l+5} \to R_{8l+4} \to SU_2^Q \times (Sp_{4l+2} \to SU_2 \times (R_{2l+1} \to R_3)). \qquad (2)$$

This chain of groups provides no additional classificatory symbols, but it does, through the introduction of the quasispin group SU_2^Q, display the N-dependence of the multi-electron states in a transparent manner.

The eigenfunctions of the l shell all transform according to the $\{1\}$ representation of $U_{2^{4l+2}}$, which under restriction to the subgroup R_{8l+5} goes down irreducibly into the basic spin representation $\Delta \equiv [(\frac{1}{2})^{4l+2}]$ of R_{8l+5}. To obtain further classificatory symbols it is necessary to explore the subgroups of R_{8l+5} that are alternatives to those of Eq. (2).

Judd[5] has shown that a much richer classification scheme is obtained decomposing the representation $\{1^N\}$ of U_{4l+2} under the chain of groups

$$U_{4l+2} \to U_{2l+1}^{\uparrow} \times U_{2l+1}^{\downarrow} \to R_3^{\uparrow} \times R_3^{\downarrow} \to R_3, \qquad (3)$$

where the orbitally antisymmetrized eigenfunctions associated with electrons having their spins "up" ($m_s = +\frac{1}{2}$) transform under $U_{2l+1}^{\uparrow}$ and those with their spin "down" ($m_s = -\frac{1}{2}$) transform under $U^{\downarrow}{}_{+1}$. This chain of groups gives a complete classification of the eigenfunctions of the configurations l^N up to $l = 3$, and for $l \geq 4$ proves much more successful in distinguishing states with the same SL quantum numbers than that of Eq. (1), though it is of course no more successful in distinguishing the 30 598 ^{14}E states of t^{13}.

These group structures have all been studied[4,5] by representing the Lie algebras in terms of the annihilation and creation operators of the method of second quantization.[6] Armstrong and Judd[7] have recently showed that it is possible to develop an alternative classification scheme of the l shell by considering the

* Research sponsored in part by the Air Force Office of Scientific Research, Office of Aerospace Research, U.S. Air Force, under AFOSR Grant No. 1275-67.

[1] G. Racah, Phys. Rev. **76**, 1352 (1949).

[2] G. Racah, *Ergebnisse der Exacten Naturwissenschaften, Vol. 37: Group Theory and Spectroscopy* (Springer-Verlag, Berlin, 1965).

[3] B. R. Judd, *Operator Techniques in Atomic Spectroscopy* (McGraw-Hill Book Co., Inc., New York, 1963).

[4] B. R. Judd, "Group Theory in Atomic Spectroscopy," in *Group Theory and its Applications*, E. M. Loebel, Ed. (Academic Press Inc., New York, 1968).

[5] B. R. Judd, Phys. Rev. **162**, 28 (1967).

[6] B. R. Judd, *Second Quantization and Atomic Spectroscopy* (Johns Hopkins Press, Baltimore, Md., 1967).

[7] L. Armstrong, Jr., and B. R. Judd, "Quasiparticles in Atomic Shell Theory," Proc. Roy. Soc. (London) (to be published).

properties of the basic tensor operators

$$\begin{aligned}\lambda_q^{(l)} &= (2)^{-\frac{1}{2}}[a^+_{\frac{1}{2}q} + (-1)^{l-q}a_{\frac{1}{2}-q}],\\ \mu_q^{(l)} &= (2)^{-\frac{1}{2}}[a^+_{\frac{1}{2}q} - (-1)^{l-q}a_{\frac{1}{2}-q}],\\ \nu_q^{(l)} &= (2)^{-\frac{1}{2}}[a^+_{-\frac{1}{2}q} + (-1)^{l-q}a_{-\frac{1}{2}-q}],\\ \xi_q^{(l)} &= (2)^{-\frac{1}{2}}[a^+_{-\frac{1}{2}q} - (-1)^{l-q}a_{-\frac{1}{2}-q}],\end{aligned} \tag{4}$$

where the subscripts to the annihilation and creation operators a and a^+ specify m_s and m_l for an electron. Their results lead to the conclusion that the states of the l shell can be classified according to the scheme

$$R^{\uparrow}_{\lambda}(2l+1) \times R^{\uparrow}_{\mu}(2l+1) \times R^{\downarrow}_{\nu}(2l+1) \times R^{\downarrow}_{\xi}(2l+1) \tag{5}$$

and its subgroups. There is no basic difficulty in extending their method to mixed configurations.

In the present paper we propose to demonstrate that the group structure suggested by Armstrong and Judd may be embedded in the group R_{8l+5} in a natural manner which sheds considerable light on representation of atomic shells in the quasiparticle methods familiar in nuclear and superconductivity physics.[8] After careful consideration of the defining of the vacuum states appropriate to the quasiparticle operators, a method for expanding the quasiparticle states in terms of the familiar determinantal states is outlined. Finally, a method for calculating quasiparticle matrix elements is developed. The need for fractional parentage coefficients in calculating these matrix elements is entirely eliminated.

II. THE BASIC GROUP STRUCTURE

Armstrong and Judd[7] derived the group structure of Eq. (5) by showing that the coupled products

$$\tfrac{1}{2}(\lambda^{(l)}\lambda^{(l)})^{(k)}_q, \quad -\tfrac{1}{2}(\mu^{(l)}\mu^{(l)})^{(k)}_q, \quad \tfrac{1}{2}(\nu^{(l)}\nu^{(l)})^{(k)}_q,$$

and

$$-\tfrac{1}{2}(\xi^{(l)}\xi^{(l)})^{(k)}_q,$$

with k *odd*, form the generators of the rotation groups $R_{\lambda}(2l+1)$, $R_{\mu}(2l+1)$, $R_{\nu}(2l+1)$, and $R_{\xi}(2l+1)$, respectively. In this section we show that the group scheme of Armstrong and Judd is part of a larger structure of the form

$$\begin{aligned}R_{8l+5} \rightarrow R^{\uparrow}_{4l+2} \times R^{\downarrow}_{4l+2} \rightarrow R^{\uparrow}_{\lambda(2l+1)} \times R^{\uparrow}_{\mu(2l+1)}\\ \times R^{\downarrow}_{\nu(2l+1)} \times R^{\downarrow}_{\xi(2l+1)},\end{aligned} \tag{6}$$

where, as before, the arrow $\uparrow$ is associated with eigenfunctions with $m_s = +\frac{1}{2}$ and the arrow $\downarrow$ with those associated with $m_s = -\frac{1}{2}$. The individual spin spaces may be further reduced by examining the subgroups of the R_{2l+1} groups, e.g., in the spin-up space we have

$$R^{\uparrow}_{\lambda(2l+1)} \times R^{\uparrow}_{\mu(2l+1)} \rightarrow R^{\uparrow}_{\lambda(3)} \times R^{\uparrow}_{\mu(3)} \rightarrow R^{\uparrow}_{\lambda\mu(3)}. \tag{7}$$

Equations (4) can be re-expressed as

$$\begin{aligned}a^+_{\frac{1}{2}q} &= (2)^{-\frac{1}{2}}(\lambda_q^{(l)} + \mu_q^{(l)}),\\ a_{\frac{1}{2}-q} &= (2)^{-\frac{1}{2}}(-1)^{l-q}(\lambda_q^{(l)} - \mu_q^{(l)}),\\ a^+_{-\frac{1}{2}q} &= (2)^{-\frac{1}{2}}(\nu_q^{(l)} + \xi_q^{(l)}),\\ a_{-\frac{1}{2}-q} &= (2)^{-\frac{1}{2}}(-1)^{l-q}(\nu_q^{(l)} - \xi_q^{(l)}).\end{aligned} \tag{8}$$

Judd[4] has shown that for the l shell the $8l+4$ annihilation and creation operators $a_{m_s m_l}$ and $a^+_{m_s m_l}$, together with their $(4l+2)(8l+3)$ distinct nonzero anticommutators, close on the commutation algebra of the group R_{8l+5} and thus the construction of the group R_{8l+5} from the operators of Eq. (8) is self-evident.

Furthermore, the coupled products $\frac{1}{2}(\lambda^{(l)}\lambda^{(l)})^{(k)}_q$ and $-\frac{1}{2}(\mu^{(l)}\mu^{(l)})^{(k)}_q$ with k *odd*, together with the linear combinations $\frac{1}{2}(\lambda\mu)^{(k)}_q - (-1)^k\frac{1}{2}(\mu\lambda)^{(k)}_q$ for all k *even* or *odd* have the same commutation relations as those of the tensor operators $v^{(k)}_q(l_\lambda, l_\lambda)$ and $v^{(k)}_q(l_\mu, l_\mu)$ with k *odd* and $v^{(k)}_q(l_\lambda, l_\mu) + (-1)^k v^{(k)}_q(l_\mu, l_\lambda)$ with k *even* or *odd* and thus form the generators of the group[9] $R_{2(l_\lambda + l_\mu + 1)}$, i.e., R_{4l+2}. Since the annihilation and creation operators used in constructing this group are all associated with $m_s = +\frac{1}{2}$, we designate this group as $R^{\uparrow}_{4l+2}$. The generators of the corresponding group $R^{\downarrow}_{4l+2}$ can likewise be readily constructed from the annihilation and creation operators with $m_s = -\frac{1}{2}$. To develop the group structure further, we must consider the group representations that arise in the classification of the multi-electron eigenfunctions.

III. THE BASIC SPIN REPRESENTATIONS

Having established the existence of the group chain $R_{8l+5} \rightarrow R^{\uparrow}_{4l+2} \times R^{\downarrow}_{4l+2}$, we show that the basic spin representation[10,11] Δ of R_{8l+5}, under restriction, decomposes into the conjugate spin representations $\Delta_1 = [\frac{1}{2} \cdots \frac{1}{2}]$, $\Delta_2 = [\frac{1}{2} \cdots -\frac{1}{2}]$ of the R_{4l+2} groups in the following manner:

$$\Delta \rightarrow (\Delta_1 + \Delta_2)^{\uparrow} \times (\Delta_1 + \Delta_2)^{\downarrow}, \tag{9}$$

where Δ_1 is to be associated with an *even* number of quasiparticles and Δ_2 with an *odd* number.

To establish the above result, we first consider the

[8] R. D. Mattuck, *A Guide to Feynman Diagrams in the Many-Body Problem* (McGraw-Hill Publ. Co., New York, 1967).

[9] J. P. Elliott, Proc. Roy. Soc. (London) **A245**, 128 (1958).

[10] D. E. Littlewood, *The Theory of Group Characters* (Oxford University Press, London, 1950), 2nd ed.

[11] P. H. Butler and B. G. Wybourne, "Reduction of the Kronecker Products for Rotation Groups," J. Phys. (to be published).

group $R^{\uparrow}_{4l+2}$ and introduce the quasiparticle annihilation and creation operators

$$\begin{aligned}
\beta_q^{(l)+} &= (2)^{-\frac{1}{2}}[a^+_{\frac{1}{2}q} + (-1)^{l-q}a_{\frac{1}{2}-q}] = \lambda_q^{(l)},\\
\beta_q^{(l)} &= (2)^{-\frac{1}{2}}[a_{\frac{1}{2}q} + (-1)^{l-q}a^+_{\frac{1}{2}-q}] = (-1)^{l-q}\lambda^{(l)}_{-q},\\
\gamma_q^{(l)+} &= (2)^{-\frac{1}{2}}[a_{\frac{1}{2}q} - (-1)^{l-q}a^+_{\frac{1}{2}-q}] = -(-1)^{l-q}\mu^{(l)}_{-q},\\
\gamma_q^{(l)} &= (2)^{-\frac{1}{2}}[a^+_{\frac{1}{2}q} - (-1)^{l-q}a_{\frac{1}{2}-q}] = \mu_q^{(l)}, \qquad (10)
\end{aligned}$$

all of which have $q > 0$. We then use these operators to define a normalized quasiparticle vacuum state $|\tilde{0}\rangle$, which is related to the normal-particle vacuum state $|0\rangle$ by

$$|\tilde{0}\rangle = 2^{(l/2)}\prod_{q>0}\beta_q^{(l)}\prod_{q>0}\gamma_q^{(l)}|0\rangle. \qquad (11)$$

The above definition of the vacuum state has the necessary property that $\beta_q^{(l)}|\tilde{0}\rangle = 0$ for all the quasiparticle annihilation operators.

By using a natural extension of Judd's[3] operators

$$w_{\nu\mu} = \sum_{k,p}(-1)^{l-\nu}[1-(-1)^k]\begin{pmatrix} l & k & l\\ -\nu & q & \mu\end{pmatrix}v_q^{(k)}(ll),$$

we find that Weyl's commuting operators[12] for the group $R^{\uparrow}_{4l+2}$ are, in terms of the quasiparticle operators,

$$\begin{aligned}
H_q^{\lambda} = \tfrac{1}{2}[\beta_q^+, \beta_q], \quad H_q^{\mu} = -\tfrac{1}{2}[\gamma_q^+, \gamma_q],\\
H_0^{\lambda\mu} = \tfrac{1}{2}(-1)^l[\beta_0, \gamma_0], \qquad (12)
\end{aligned}$$

where for convenience we now drop the quantum number l of the electrons. H_q^{λ} gives eigenvalues $+\frac{1}{2}$ or $-\frac{1}{2}$ according to whether or not β_q^+ is or is not contained in the quasiparticle state $\beta_\mu^+\cdots\beta_\tau^+\gamma_\eta^+\cdots\gamma_\rho^+|\tilde{0}\rangle$, while H_q^{μ} gives eigenvalues $-\frac{1}{2}$ or $+\frac{1}{2}$ if γ_q^+ is or is not in the state. Our definition of the quasiparticle vacuum makes it legitimate to include β_0 or γ_0 in the quasiparticle state, implying that if the state contains both β_0 and γ_0 or neither of these, $H_0^{\lambda\mu}$ gives eigenvalues of $\frac{1}{2}(-1)^l$, while if the state contains only one of these, the eigenvalues become $-\frac{1}{2}(-1)^l$.

The operators of $R^{\uparrow}_{4l+2}$, being coupled products of the λ's and μ's, connect only states differing by two quasiparticles (or two ordinary particles) or none. It follows that if a state contains an *even* number of quasiparticles (including either β_0 or γ_0), it transforms according to the representation $\Delta_1 = [\frac{1}{2}\cdots\frac{1}{2}]$ of $R^{\uparrow}_{4l+2}$, and if an *odd* number, according to the conjugate representation $\Delta_2 = [\frac{1}{2}\cdots-\frac{1}{2}]$. We note that our quasiparticle vacuum is proportional to

$$\prod_q[1+(-1)^{l-q}a^+_{-q}a^+_q]|0\rangle,$$

i.e., contains an even number of real particles, and consequently if the quasiparticle number is even (odd), then the real particle number is even (odd).

Under the restriction

$$R_{4l+2} \to R_\lambda(2l+1)\times R_\mu(2l+1),$$

both Δ_1 and Δ_2 decompose irreducibly into the product representations $\Delta^{\uparrow}_\lambda \times \Delta^{\uparrow}_\mu$, where Δ_λ and Δ_μ are the basic spin representations of the R_{2l+1} rotation group. Similarly, under restriction of $R^{\downarrow}_{4l+2}$, Δ_1 and Δ_2 decompose irreducibly into the product representation $\Delta^{\downarrow}_\nu \times \Delta^{\downarrow}_\xi$ of $R_\nu(2l+1)\times R_\xi(2l+1)$. The coupled product $\frac{1}{2}(\lambda^{(l)}\lambda^{(l)})^{(1)}_q$ supplies the generators of the group $R^{\uparrow}_\lambda(3)$, which is a subgroup of $R^{\uparrow}_\lambda(2l+1)$, and similarly $-\frac{1}{2}(\mu^{(l)}\mu^{(l)})^{(1)}_q$ gives the generators of $R^{\uparrow}_\mu(3)$. Thus we may write the complete group structure for describing the transformation of the l-shell eigenfunctions in terms of the chain of groups:

$$\begin{aligned}
U_{2^{4l+2}} &\to R_{8l+5} \to R^{\uparrow}_{4l+2}\times R^{\downarrow}_{4l+2}\\
&\to (R_\lambda(2l+1)\times R_\mu(2l+1))^{\uparrow}\\
&\quad\times(R_\nu(2l+1)\times R_\xi(2l+1))^{\downarrow}\\
&\to (R_\lambda(3)\times R_\mu(3))^{\uparrow}\times(R_\nu(3)\times R_\xi(3))^{\downarrow}\\
&\to R^{\uparrow}_{\lambda\mu}(3)\times R^{\downarrow}_{\nu\xi}(3) \to R_{(3)} \to R_{(2)}. \qquad (13)
\end{aligned}$$

If, under the restriction $R_{2l+1} \to R_3$, the basic spin representation Δ decomposes into the representations of R_3 without repetitions, then the chain of groups given in Eq. (13) will give a complete set of classificatory symbols to uniquely label all the eigenfunctions of the l shell. Methods of determining these branching rules have been discussed by Butler and Wybourne.[11] As noted by Armstrong and Judd, the classification is indeed unique for $l \leq 8$. It is interesting to note that even in the t shell ($l = 14$), no representation of R_3 occurs more than 15 times in the decomposition of the basic spin representation under $R_{29} \to R_3$.

It is evident from the nature of the chain of groups given in Eq. (13) that eigenfunctions constructed with these transformation properties will be characterized by well-defined L and M_L quantum numbers, but will generally not correspond to a definite number of particles N or have accessible S and M_S quantum numbers. In trivial cases, some identification may still be possible, as may be seen when the quasiparticle states appropriate to the d shell are expanded as linear combinations of determinantal states.

IV. QUASIPARTICLE AND DETERMINANTAL STATES

The quasiparticle scheme is unlikely to be of calculational value if the quasiparticle states have

[12] H. Weyl, Math. Z. 23, 271 (1925).

to be first expanded as the linear combinations of the usual determinantal states. Rather we would like to calculate directly in the quasiparticle scheme. Nevertheless, it is useful in some applications to be able to make the transformation from quasiparticle states to determinantal states.

Consider the segment

$$R_\lambda^\uparrow(2l+1) \times R_\mu^\uparrow(2l+1) \to R_\lambda^\uparrow(3) \times R_\mu^\uparrow(3) \to R_{\lambda\mu}^\uparrow(3) \tag{14}$$

of the group chain described by Eq. (13). Suppose under the restriction $R_{2l+1} \to R_3$ we have

$$\Delta_\epsilon \to \sum g_{j_\epsilon}[j_\epsilon],$$

where g_{j_ϵ} is the number of times the $[j_\epsilon]$ representation of $R_\epsilon(3)$ arises in the decomposition; then the same branching rule holds for both $R_\lambda^\uparrow(2l+1)$ and $R_\mu^\uparrow(2l+1)$. Thus the problem of forming eigenfunctions where every electron in the l shell has $m_s = +\frac{1}{2}$ reduces to constructing eigenfunctions $|j_\lambda j_\mu; L_{\lambda\mu}M_{L_{\lambda\mu}}\rangle^\uparrow$ for pseudo two-particle configurations $(j_\lambda j_\mu)^\uparrow$, where j_λ and j_μ will either be both half-integral or integral angular momentum.

For example, for the g shell we find, under $R_9 \to R_3$, $\Delta \to d + h$, and hence the LM_L states associated with all $m_s = +\frac{1}{2}$ will be the same as those arising in the three pseudo two-particle configurations $d_\lambda d_\mu$, $d_\lambda h_\mu$, $h_\lambda d_\mu$, and $h_\lambda h_\mu$.

We now wish to develop a systematic method for expanding the typical pseudo two-particle state $|j_\lambda j_\mu LM_L\rangle^\uparrow$ as a linear combination of determinantal states. To do this we first construct a set of angular-momentum ladder operators in the spaces of $R_\lambda(3)$ and $R_\mu(3)$ in terms of the quasiparticle annihilation and creation operators. These operators L_q^λ and L_q^μ must clearly be simply proportional to the coupled products $(\lambda^{(l)}\lambda^{(l)})_q^{(1)}$ and $(\mu^{(l)}\mu^{(l)})_q^{(1)}$ and, in detail,

$$\begin{aligned} L_0^\lambda &= \sum_{q>0} qH_q^\lambda = \tfrac{1}{2}\sum_{q>0} q[\beta_q^+, \beta_q], \\ L_1^\lambda &= \sum_{q>0}(-1)^{l-q}[\tfrac{1}{2}(l+q)(l-q+1)]^{\frac{1}{2}}\lambda_q^{(l)}\lambda_{1-q}^{(l)} \\ &= \sum_{q>0}[\tfrac{1}{2}(l+q)(l-q+1)]^{\frac{1}{2}}\beta_q^+\beta_{q-1}, \end{aligned} \tag{15a}$$

and

$$\begin{aligned} L_0^\mu &= -\sum_{q<0} qH_q^\mu = \tfrac{1}{2}\sum_{q>0} q[\gamma_q^+, \gamma_q], \\ L_{-1}^\mu &= -\sum_{q<0}(-1)^{l-q}[\tfrac{1}{2}(l-q)(l+q+1)]^{\frac{1}{2}}\mu_q^{(l)}\mu_{-1-q}^{(l)} \\ &= \sum_{q>0}[\tfrac{1}{2}(l+q)(l-q+1)]^{\frac{1}{2}}\gamma_q^+\gamma_{q-1}. \end{aligned} \tag{15b}$$

The form of L_0 shows that the quasiparticle vacuum state is the state having the highest j_λ or j_μ and the lowest m_{j_λ} or highest m_{j_μ} in the $R_\lambda(3)$ and $R_\mu(3)$ spaces, respectively. In obtaining these results we have required that the quasiparticle vacuum state $|\tilde{0}\rangle_\lambda$ in the λ space and $|\tilde{0}\rangle_\mu$ in the μ space may be coupled to yield the vacuum state $|\tilde{0}\rangle_\lambda|\tilde{0}\rangle_\mu \equiv |\tilde{0}\rangle$ defined earlier for the $R_{4l+2}^\uparrow$ space. As a result, we have, to within an inconsequential phase,

$$|\tilde{0}\rangle_\lambda = 2^{(l/4)}\prod_{q>0}\beta_q|0\rangle$$

and

$$|\tilde{0}\rangle_\mu = 2^{(l/4)}\prod_{q>0}\gamma_q|0\rangle.$$

Having constructed the angular-momentum ladder operators, it becomes a trivial matter to construct the angular-momentum states $|j_\lambda m_\lambda\rangle$ and $|j_\mu m_\mu\rangle$ for highest j_λ and j_μ appropriate to the $R_\lambda(3)$ and $R_\mu(3)$ spaces in terms of an *even* number of quasiparticles (remembering to count a β_0 or a γ_0 if it occurs).

The angular-momentum states for values of $j_\lambda' < j_\lambda$ or $j_\mu' < j_\mu$ may be constructed by requiring that the states $|j_\lambda' - m_{j_\lambda'}\rangle$ and $|j_\mu' m_{j_\mu'}\rangle$ be orthogonal to the states $|j_\lambda - m_{j_\lambda'}\rangle$ and $|j_\mu m_{j_\mu'}\rangle$, respectively, and to any other states having the same $-m_{j_\lambda'}$ or $m_{j_\mu'}$, i.e., we construct an orthonormal set.

The states involving an *odd* number of quasiparticles in either space may be built up by applying the ladder operators to $(2)^{\frac{1}{2}}\beta_0|\tilde{0}\rangle_\lambda$ and $(2)^{\frac{1}{2}}\gamma_0|\tilde{0}\rangle_\mu$, instead of to $|\tilde{0}\rangle_\lambda$ and $|\tilde{0}\rangle_\mu$. In each case, the normalization is chosen so that $\langle j-j|j-j\rangle_\mu$ and $\langle jj|jj\rangle_\mu$ are equal to $2^{-(l/2)}$.

As an example, consider the case of the d shell, where for $R_5 \to R_3$ we have $[\frac{1}{2}\frac{1}{2}] \to [\frac{3}{2}]$. The proportionality constant required to recover the ladder operators is $(10)^{\frac{1}{2}}$ and we have

$$L_0^\lambda = H_1^\lambda + 2H_2^\lambda \quad \text{and} \quad L_1^\lambda = (2)^{\frac{1}{2}}\beta_2^+\beta_1 + (3)^{\frac{1}{2}}\beta_1^+\beta_0 \tag{16a}$$

and

$$L_0^\mu = -H_1^\mu - 2H_2^\mu \quad \text{and} \quad L_{-1}^\mu = (2)^{\frac{1}{2}}\gamma_2^+\gamma_0 + (3)^{\frac{1}{2}}\gamma_1^+\gamma_0. \tag{16b}$$

As a consequence, for n_λ *even* (where n_λ is the number of quasiparticles in the λ space) we have

$$\begin{aligned} &|\tfrac{3}{2}-\tfrac{3}{2}\rangle_\lambda = |\tilde{0}\rangle_\lambda, \quad |\tfrac{3}{2}-\tfrac{1}{2}\rangle_\lambda = -(2)^{\frac{1}{2}}\beta_1^+\beta_0|\tilde{0}\rangle_\lambda, \\ &|\tfrac{3}{2}\,\tfrac{1}{2}\rangle_\lambda = (2)^{\frac{1}{2}}\beta_2^+\beta_0|\tilde{0}\rangle_\lambda, \text{ and } |\tfrac{3}{2}\,\tfrac{3}{2}\rangle_\lambda = \beta_1^+\beta_2^+|\tilde{0}\rangle_\lambda, \end{aligned} \tag{17a}$$

while for n_μ even in the μ space, we have

$$|\tfrac{3}{2}\,\tfrac{3}{2}\rangle_\mu = |\tilde{0}\rangle_\mu, \quad |\tfrac{3}{2}\,\tfrac{1}{2}\rangle_\mu = (2)^{\frac{1}{2}}\gamma_1^+\gamma_0\,|\tilde{0}\rangle_\mu,$$
$$|\tfrac{3}{2}\,-\tfrac{1}{2}\rangle_\mu = (2)^{\frac{1}{2}}\gamma_2^+\gamma_0\,|\tilde{0}\rangle_\mu, \quad |\tfrac{3}{2}\,-\tfrac{3}{2}\rangle_\mu = \gamma_1^+\gamma_2^+\,|\tilde{0}\rangle_\mu. \tag{17b}$$

Having constructed the states given in Eq. (17), we may perform a vector coupling to produce states characterized by the orbital quantum numbers L and M_L. For an even or odd total number of particles N,

$$|j_\lambda j_\mu L_{\lambda\mu} M_{L_{\lambda\mu}}\rangle = \sum_{m_\lambda, m_\mu} \langle m_\lambda m_\mu \mid L_{\lambda\mu} M_{L_{\lambda\mu}}\rangle\, |j_\lambda m_\lambda, j_\mu m_\mu\rangle, \tag{18}$$

where $\langle m_\lambda m_\mu \mid L_{\lambda\mu} M_{L_{\lambda\mu}}\rangle$ is the usual Clebsch–Gordan coefficient and

$$|j_\lambda m_\lambda, j_\mu m_\mu\rangle = [(L_+^\lambda)^{j_\lambda - m_\lambda}(L_-^\mu)^{j_\mu + m_\mu}/AB]\,|j_\lambda m_{j_\lambda}, j_\mu - m_{j_\mu}\rangle, \tag{19}$$

with

$$A = [(2j_\mu)!\,(j_\mu - m_\mu)!/(j_\mu + m_\mu)!]^{\frac{1}{2}}$$

and

$$B = [(2j_\lambda)!\,(j_\lambda + m_\lambda)!/(j_\lambda - m_\lambda)!]^{\frac{1}{2}}.$$

If N is *even* and n_λ and n_μ are both even or odd, then, apart from an inconsequential phase,

$$|j_\lambda - m_{j_\lambda}, j_\mu m_{j_\mu}\rangle = |\bar{0}\rangle,$$

while if N is *odd* and n_λ and n_μ are of opposite parity,

$$|j_\lambda - m_{j_\lambda}, j_\mu m_{j_\mu}\rangle = (2)^{\frac{1}{2}}\beta_0\,|\bar{0}\rangle.$$

Use of the above results for the particular case of the d shell gives

$$|\Delta_1(\tfrac{3}{2})_\lambda(\tfrac{3}{2})_\mu F3\rangle = \beta_1^+\beta_2^+\,|\bar{0}\rangle = a_1^+a_2^+\,|0\rangle$$

and

$$|\Delta_2(\tfrac{3}{2})_\lambda(\tfrac{3}{2})_\mu F2\rangle = 2(\beta_2^+\beta_0 - \gamma_1^+\gamma_0\beta_1^+\beta_2^+)\beta_0\,|\bar{0}\rangle = a_1^+a_2^+a_{-1}^+\,|0\rangle,$$

while

$$|\Delta_2(\tfrac{3}{2})_\lambda(\tfrac{3}{2})_\mu D2\rangle^\uparrow = 2(\beta_2^+\beta_0 + \gamma_1^+\gamma_0\beta_1^+\beta_2^+)\beta_0\,|\bar{0}\rangle_R = -a_2^+\,|0\rangle,$$

where the creation operators a_q^+ are all associated with $m_s = +\tfrac{1}{2}$.

The construction of the states $|j_\nu j_\xi L M_L\rangle^\downarrow$ in the spin-down space proceeds in exactly the same manner except that the creation operators a_q^+ are now all associated with $m_s = -\tfrac{1}{2}$. The states constructed for the spin-up and spin-down spaces may be coupled by the usual vector-coupling method to give the final l-shell eigenfunctions

$$|(j_\lambda j_\mu)L_{\lambda\mu}^\uparrow(j_\nu j_\xi)L_{\nu\xi}^\downarrow; LM_L\rangle = \sum_{M_{L_{\lambda\mu}}, M_{L_{\nu\xi}}} \langle M_{L_{\lambda\mu}} M_{L_{\nu\xi}} \mid LM_L\rangle \times |(j_\lambda j_\mu)L_{\lambda\mu}M_{L_{\lambda\mu}}\rangle^\uparrow\,|(j_\nu j_\xi)L_{\nu\xi}M_{L_{\nu\xi}}\rangle^\downarrow. \tag{20}$$

These resultant eigenfunctions will not in general correspond to a definite number of particles, but will, however, involve either *even* or *odd* numbers of particles.

It should be apparent from the preceding discussion that no significant complications arise when treating pseudo two-particle configurations where $j_\lambda \neq j_\mu$ or where j_λ and j_μ occur more than once in the decomposition of the basic spin representation of R_{2l+1} to the group R_3.

V. CALCULATION OF MATRIX ELEMENTS

Any interaction may be expanded in terms of sums of products of tensor operators.[3] In the case of equivalent electron configurations l^N, the interactions may be expressed in terms of sums of products of the double tensor operators $\mathbf{W}^{(\kappa k)}$, where

$$\mathbf{W}^{(\kappa k)} = \sum_{i=1}^{N} \mathbf{w}_i^{(\kappa k)}$$

and

$$\langle sl \|\, w^{(\kappa k)} \,\| s'l'\rangle = [(2\kappa + 1)(2k + 1)]^{\frac{1}{2}}\delta(s, s')\delta(ll').$$

Judd[6] has shown that the double tensors $\mathbf{W}^{(\kappa k)}$ may be related to coupled products of the usual annihilation and creation operators, viz.,

$$W_{\pi q}^{(\kappa k)} = -(a^+a)_{\pi q}^{(\kappa k)}. \tag{21}$$

The annihilation and creation operators in Eq. (21) may be re-expressed in terms of the quasiparticle annihilation and creation operators using Eq. (8) and then decoupled to expose their spin dependence; for k *odd* this gives

$$\begin{aligned} W_{\pi q}^{(\kappa k)} = -\tfrac{1}{2}[&\langle s\tfrac{1}{2}\, s\tfrac{1}{2} \mid \kappa\pi\rangle\{(\lambda\nu)_q^{(k)} - (\mu\xi)_q^{(k)}\} \\ &+ \langle s - \tfrac{1}{2}\, s\tfrac{1}{2} \mid \kappa\pi\rangle\{(\nu\nu)_q^{(k)} - (\xi\xi)_q^{(k)}\} \\ &- \langle s\tfrac{1}{2}\, s - \tfrac{1}{2} \mid \kappa\pi\rangle\{(\lambda\lambda)_q^{(k)} - (\mu\mu)_q^{(k)}\} \\ &- \langle s - \tfrac{1}{2}\, s - \tfrac{1}{2} \mid \kappa\pi\rangle\{(\nu\lambda)_q^{(k)} - (\xi\mu)_q^{(k)}\}] \end{aligned} \tag{22a}$$

and, for k *even* this gives

$$\begin{aligned} W_{\pi q}^{(\kappa k)} = -\tfrac{1}{2}[&\langle s\tfrac{1}{2}\, s\tfrac{1}{2} \mid \kappa\pi\rangle \\ &\times \{(\mu\nu)_q^{(k)} - (\lambda\xi)_q^{(k)} + \delta(k, 0)(2l + 1)^{\frac{1}{2}}\} \\ &+ \langle s - \tfrac{1}{2}\, s\tfrac{1}{2} \mid \kappa\pi\rangle \\ &\times \{(\xi\nu)_q^{(k)} - (\nu\xi)_q^{(k)} + \delta(k, 0)(2l + 1)^{\frac{1}{2}}\} \\ &- \langle s\tfrac{1}{2}\, s - \tfrac{1}{2} \mid \kappa\pi\rangle \\ &\times \{(\mu\lambda)_q^{(k)} - (\lambda\mu)_q^{(k)} + \delta(k, 0)(2l + 1)^{\frac{1}{2}}\} \\ &- \langle s - \tfrac{1}{2}\, s - \tfrac{1}{2} \mid \kappa\pi\rangle \\ &\times \{(\xi\lambda)_q^{(k)} - (\nu\mu)_q^{(k)} + \delta(k, 0)(2l + 1)^{\frac{1}{2}}\}]. \end{aligned} \tag{22b}$$

All one-particle interactions can be expressed in terms of the quasiparticle operators of Eqs. (22a) and (22b) and the relevant matrix elements evaluated in the λ, μ, ν, ξ spaces, i.e., for the pseudo four-particle configurations $j_\lambda j_\mu j_\nu j_\xi$ with states as defined in Eq. (20).

The two-particle interactions will involve terms of the form[13]

$$\sum_{i\neq j}[\mathbf{w}_i^{(\kappa_1 k_1)}\mathbf{w}_j^{(\kappa_2 k_2)}]^{(\kappa k)K}_{Q} = [\mathbf{W}^{(\kappa_1 k_1)}\mathbf{W}^{(\kappa_2 k_2)}]^{(\kappa k)K}_{Q} - (-1)^{2l+2s+\kappa+k} \times [(2\kappa_1+1)(2\kappa_2+1)(2k_1+1)(2k_2+1)]^{\frac{1}{2}} \times \begin{Bmatrix} k_1 & k_2 & k \\ l & l & l \end{Bmatrix}\begin{Bmatrix} \kappa_1 & \kappa_2 & \kappa \\ s & s & s \end{Bmatrix}\mathbf{W}^{(\kappa k)K}_{Q}. \quad (23)$$

For scalar interactions where $\kappa = k = K = Q = 0$, Eq. (23) reduces to

$$\sum_{i\neq j}(\mathbf{w}_i^{(\kappa k)}\cdot\mathbf{w}_j^{(\kappa k)}) = \mathbf{W}^{(\kappa k)}\cdot\mathbf{W}^{(\kappa k)} - \left[\frac{(2\kappa+1)^2(2k+1)^2}{(4l+2)}\right]^{\frac{1}{2}}\mathbf{W}^{(00)}, \quad (24)$$

where $\kappa_1 = \kappa_2$ and $k_1 = k_2$ and we drop the subscripts.

Having expressed the one- and two-particle interactions in terms of quasiparticle operators, the matrix elements may be directly evaluated using the standard methods of tensor operators and angular-momentum recoupling,[3] thus entirely eliminating the need for the fractional parentage coefficients that arise in the traditional evaluation of matrix elements.

The use of tensor-operator methods requires the evaluation of the reduced matrix elements of the quasiparticle operators between quasiparticle states. The reduced matrix elements of the quasiparticle operators will be zero between states containing the same number of quasiparticles. The reduced matrix elements may be generally evaluated by expanding a particular component of the operator, say $\lambda_0^{(l)}$, together with the states of the bra and ket in terms of ordinary annihilation and creation operators, remembering that the states are orthogonal, though not usually orthonormal. For the particular case of maximum $j = j_M$ in the space under consideration, a simple formula can be found, viz.,

$$\begin{aligned}\langle\Delta_1 j_M\|\, A^{(l)}\,\|\Delta_2 j_M\rangle &= \pm(-1)^l\langle\Delta_2 j_M\|\, A^{(l)}\,\|\Delta_1 j_M\rangle \\ &= \pm(-1)^{l+2j_M}\left[2^{\frac{1}{2}}\begin{pmatrix} j_M & l & j_M \\ j_M & 0 & j_M\end{pmatrix}\right]^{-1} \\ &= \pm[(2j_M-l)!\,(2j_M+l+1)!/2]^{\frac{1}{2}}/(2j_M)!, \quad (25)\end{aligned}$$

where the plus sign is taken if A is λ or ν and the minus sign if A is μ or ξ. For example, we may readily deduce from Eq. (25) that

$$\langle\Delta_1\tfrac{3}{2}\|\,\lambda^{(2)}\,\|\Delta_2\tfrac{3}{2}\rangle = (10)^{\frac{1}{2}}. \quad (26)$$

The matrix elements of the Coulomb repulsion may be readily calculated in the quasiparticle scheme by first noting that, for equivalent electron configurations l^N, we have[3]

$$\sum_{i>j}\frac{e^2}{r_{ij}} = e^2\sum_k\frac{r_<^k}{r_>^{k+1}}\frac{(l\|\, C^{(k)}\,\|l)^2}{2k+1}\sum_{i\neq j}(\mathbf{w}_i^{(0k)}\cdot\mathbf{w}_j^{(0k)})$$

and, from Eq. (24),

$$\sum_{i\neq j}(\mathbf{w}_i^{(0k)}\cdot\mathbf{w}_j^{(0k)}) = \mathbf{W}^{(0k)}\cdot\mathbf{W}^{(0k)} - \frac{(2k+1)}{(4l+2)^{\frac{1}{2}}}\mathbf{W}^{00}, \quad (27)$$

and then expressing the matrix elements of $\mathbf{W}^{(0k)}$ in terms of the quasiparticle operators via Eq. (22b).

The scalar operator $W_{00}^{(00)}$ is proportional to the number operator[6] $\sum a_\tau^+ a_\tau$ and for the d shell in the "spin-up" space will have eigenvalues $N/(4l+2)^{\frac{1}{2}}$, where N is the number of particles, since in this case every state corresponds to a definite number of particles. For more general cases of particles, the eigenvalues of $W_{00}^{(00)}$ will reflect the mixing of particle numbers in the quasiparticle state. This sometimes leads to a simple method for expanding the pseudoparticle states as linear combinations of the usual $|l^N SLM_S M_L\rangle$ states to within a phase.

As an example of the above, consider the states of the f shell in the spin-up space. The states may be constructed in terms of pseudoparticles having 3 and 0 units of angular momentum, since, under $R_7 \to R_3$, we have $\Delta \to [0] + [3]$. Evaluation of the matrix element of $W_{00}^{(00)}$ for the S-states for an even number of f-electrons proceeds in the pseudoparticle scheme as follows:

$$\begin{aligned}&\langle\Delta_1(3)_\lambda(3)_\mu;\, S0|\, W_{00}^{(00)}\,|\Delta_1(3)_\lambda(3)_\mu;\, S0\rangle \\ &= (2)^{-\frac{1}{2}}\langle\Delta_1(3)_\lambda(3)_\mu;\, S0|\,(\mu\lambda)_0^{(0)}\,|\Delta_1(3)_\lambda(3)_\mu;\, S0\rangle + \frac{(14)^{\frac{1}{2}}}{4} \\ &= (2)^{-\frac{1}{2}}\langle\Delta_1(3)_\lambda(3)_\mu;\, S\|\,(\mu\lambda)^{(0)}\,\|\Delta_1(3)_\lambda(3)_\mu;\, S\rangle + \frac{(14)^{\frac{1}{2}}}{4} \\ &= (2)^{-\frac{1}{2}}\begin{Bmatrix}3 & 3 & 3\\ 3 & 3 & 3\\ 0 & 0 & 0\end{Bmatrix}\langle\Delta_1(3)_\mu\|\,\mu^{(3)}\,\|\Delta_2(3)_\mu\rangle \\ &\quad\times\langle\Delta_2(3)_\lambda\|\,\lambda^{(3)}\,\|\Delta_1(3)_\lambda\rangle + \frac{(14)^{\frac{1}{2}}}{4} \\ &= \frac{(14)^{\frac{1}{2}}}{28};\end{aligned}$$

[13] B. G. Wybourne, J. Chem. Phys. **48**, 2596 (1968).

and since we must necessarily have

$$|\Delta_1(3)_\lambda(3)_\mu; S0\rangle = a\,|f^4\,{}^5SM_S = 2M_L = 0\rangle + b\,|f^0\,{}^1SM_S = 0M_L = 0\rangle,$$

we find $a^2 = \frac{1}{8}$ and $b^2 = \frac{7}{8}$. Similarly, for an *odd* number of f-electrons,

$$|\Delta_2(3)_\lambda(3)_\mu; S0\rangle = a\,|f^3\,{}^4SM_S = \tfrac{3}{2}M_L = 0\rangle + b\,|f^7\,{}^8SM_S = \tfrac{7}{2}M_L = 0\rangle,$$

where again $a^2 = \frac{1}{8}$ and $b^2 = \frac{7}{8}$. The above two states will be orthogonal to the states $|\Delta_1(0)_\lambda(0)_\mu; S0\rangle$ and $|\Delta_2(0)_\lambda(0)_\mu; S0\rangle$.

To summarize, the matrix elements of any interaction may be calculated in the quasiparticle formalism by the following steps: (1) Express the interaction in terms of sums of products of the tensor operators $\mathbf{W}^{(\kappa k)}$; (2) express the sums of products of the tensor operators $\mathbf{W}^{(\kappa k)}$ as sums and products of the coupled products of the quasiparticle operators $(\boldsymbol{\lambda\mu})^{(k)}$, etc.; (3) calculate the matrix elements of the quasiparticle operators within the pseudo four-particle configuration $j_\lambda j_\mu j_\nu j_\xi$, evaluating the reduced matrix elements as required.

VI. CONCLUSIONS

The establishment of the complete group chain in Eq. (13) sheds further light on the role of the quasiparticle formalism in atomic shell theory. The principal disadvantage of the quasiparticle scheme would seem to be the abandonment of the spin quantum numbers SM_S and the formation of eigenfunctions involving an indefinite number of particles.

The shortcomings of the quasiparticle formalism are partially compensated by the establishment of a remarkably rich classification scheme. Furthermore, the calculation of matrix elements in the quasiparticle scheme requires little more than a knowledge of the theory of angular-momentum recoupling coefficients, and, if combined with the powerful diagrammatic methods of Jucys *et al.*,[14] becomes a trivial problem readily amenable to machine calculation without recourse to the usual coefficients of fractional parentage.

Finally, we should note that while the methods outlined here have been devoted solely to the case of shells of equivalent electron orbitals, there is no difficulty in extending these methods to mixed configurations.

ACKNOWLEDGMENTS

We are indebted to Dr. L. Armstrong, Jr., and Professor B. R. Judd for an advance copy of their manuscript. We are particularly grateful to Dr. J. P. Elliott for numerous stimulating discussions which led to this work.

[14] A. Jucys, J. Levinsonas, and V. Vanagas, *Mathematical Apparatus of the Theory of Angular Momentum* (Vilna, 1960). A translation by R. Sen and R. N. Sen has been published in the Israel Program for Scientific Translations (S. Monson, Jerusalem, 1962).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 12 DECEMBER 1969

Character Analysis of $U(N)$ and $SU(N)$

STEPHEN BLAHA*
The Rockefeller University, New York, New York

(Received 2 May 1969)

A symmetric group analysis of the characters of $U(N)$ and $SU(N)$ representations yields formulas for (i) the multiplicities of weights in irreducible and tensor product representations, (ii) the coefficients occurring in the Clebsch–Gordan series decomposition of Kronecker products with an arbitrary number of factors, (iii) the content of irreducible and tensor product representations of $U(\sum_i N_i)$ with respect to representations of its direct product subgroup, $U(N_1) \otimes U(N_2) \otimes \cdots \equiv \otimes_i U(N_i)$, and (iv) the content of irreducible representations of $U(NM)$ with respect to irreducible representations of $U(N) \otimes U(M)$. In particular, we exhibit formulas for (i), (ii), and (iii) containing only irreducible characters and Frobenius compound characters of the symmetric group. Under the application of an operator of the subgroup, $\otimes_i U(N_i)$ with $\sum_i N_i < N$, a vector in a representation of $U(N)$ transforms as a linear combination of vectors in irreducible representations of the subgroup. We give formulas for determining the vectors occurring in such a linear combination. They are derived in a similar fashion to the formulas for (i), (ii), and (iii). In terms of weight diagrams, the formulas give the number of times a weight diagram of the subgroup's algebra occurs in the hyperplane generated by the application of the algebra to the weight of the $U(N)$ vector in question.

I. INTRODUCTION

The generators of the $U(N)$ algebra A_k^i, $i, k = 1, \cdots, N$, satisfy the commutation relations

$$[A_k^i, A_m^j] = \delta_m^i A_k^j - \delta_k^j A_m^i. \tag{1}$$

One form of the generators of $SU(N)$ also satisfies (1), but in addition must meet the requirements of tracelessness and

$$\sum_{i=1}^{N} A_i^i = 0. \tag{2}$$

Another form of the generators and commutation relations of $SU(N)$ is the Cartan form

$$\begin{aligned} [H_i, H_j] &= 0, \\ [\mathbf{H}, E_\alpha] &= \mathbf{r}_\alpha E_\alpha, \\ [E_\alpha, E_{-\alpha}] &= \mathbf{r}_\alpha \cdot \mathbf{H}, \\ [E_\alpha, E_\beta] &= N_{\alpha\beta} E_{\alpha+\beta}, \end{aligned}$$

with $\mathbf{H} = (H_1, H_2, \cdots, H_{N-1})$ an $(N-1)$-tuple of the generators of the Cartan subalgebra, E_α the generator corresponding to the root $\mathbf{r}_\alpha$, and $N_{\alpha\beta}$ a set of numbers. The generators A_j^i and H_i of $SU(N)$ are simply related by

$$H_i = \sum_j S_{ij} A_j^j, \tag{3}$$

where

$$\begin{aligned} S_{ij} &= 0, && \text{for } j > i+1, \\ S_{ij} &= -[i/(i+1)]^{\frac{1}{2}}, && \text{for } j = i+1, \quad S^T S = 1, \\ S_{ij} &= [i(i+1)]^{-\frac{1}{2}}, && \text{for } j < i+1. \end{aligned}$$

* Research sponsored in part by the Air Force Office of Scientific Research, Office of Aerospace Research, U.S. Air Force, under AFOSR contract grant number 69-1629.

In $U(N)$, the matrices representing the generators A_i^i, for $i = 1, 2, \cdots, N$, are diagonal, and we define the weight of a vector x in a $U(N)$ representation to be $\mathbf{W} = (W_1, W_2, \cdots, W_N)$ with $A_i^i x = W_i \mathbf{x}$ for $i = 1, 2, \cdots, N$. In $SU(N)$, the weight of a vector x with respect to the generators A_i^i is defined in the same manner as for $U(N)$. However, the weight of $\mathbf{x}$ with respect to the generators H_i, for $i = 1, 2, \cdots, N-1$, is $\mathbf{W}' = (W_1', W_2', \cdots, W_{N-1}')$ with $H_i \mathbf{x} = W_i' \mathbf{x}$. These definitions and Eq. (3) imply

$$W_i' = \sum_j S_{ij} W_j. \tag{3'}$$

As a result of this simple relationship, we need only consider the weights with respect to the generators A_i^i in the remainder of the paper.

In Sec. II we discuss the structure of the weights occurring in irreducible representations of $U(N)$ and $SU(N)$, and emphasize their inherent symmetric group (Weyl group) structure for later use in our analysis of characters.

The characters of $U(N)$ representations are analyzed in Sec. III in order to derive formulas for (i) the multiplicity of a weight in a reducible or irreducible representation [Eq. (21) below], (ii) the coefficients occurring in the Clebsch–Gordan series decomposition of Kronecker products with an arbitrary number of factors [Eqs. (27) and (33) below], (iii) the content of any irreducible or tensor product representation of $U(\sum_i N_i)$ with respect to representations of its direct product subgroup $U(N_1) \otimes U(N_2) \otimes \cdots \equiv \otimes_i U(N_i)$ [Eqs. (24) and (35) below], and (iv) the content of an irreducible representation of $U(NM)$ with respect to representations of $U(N) \otimes U(M)$ [Eq. (36) below].

Of course, the results we obtain are not indispensable for present-day physics since the group representations of current physical interest can be handled quite well by traditional techniques. However, besides being useful for more complex cases, our results do demonstrate the interesting fact that one may obtain all class properties (i.e., properties unaffected by similarity transformations) of the unitary groups from purely symmetric group considerations. To be more precise, all formulas developed in Secs. III and V can be derived strictly within the theory of symmetric functions[1] and expressed solely in terms of characters of the symmetric group. Furthermore, one may use the formulas of Sec. III to find the content of Kronecker products of symmetric group representations, and the content of symmetric group representations with respect to subgroups.

The orthogonal and symplectic groups may be analyzed in a similar fashion to our analysis of the unitary groups with the role of the symmetric group assumed by the "octahedral" groups.

Several papers have appeared[2] on formulas for weight multiplicities and Clebsch–Gordan series coefficients. Kostant derived a formula for the multiplicity of a weight in an irreducible representation of a semisimple Lie algebra. Straumann has given formulas for the decomposition of irreducible representations of semisimple Lie algebras with respect to semisimple subalgebras. Steinberg developed a formula for Clebsch–Gordan series coefficients. Our corresponding results [Eqs. (21) and (25)] for multiplicities and Clebsch–Gordan series coefficients are expressed in a simpler form and appear to be more useful for computational purposes. We obtained simplified equations because we dealt only with the unitary groups. In addition, we were able to express our results solely in terms of symmetric group characters. There has been other work done on specific Lie groups. Gruber derived formulas for Clebsch–Gordan series coefficients for the groups $SU(N)$, $SO(2N+1)$, $SO(2N)$, and G_2. Biedenharn, Gruber, and Weber have extended Kostant's formula to the noncompact group $SU(2, 1)$.

Hagen and MacFarlane have developed a recursive procedure for calculating the subgroup content of representations which is based on an analysis of symmetric functions. Our results, besides applying to more general cases, are in the form of closed formulas and are not of a recursive nature.

In Sec. IV we discuss an orthogonal tableau basis for $U(N)$ irreducible representations (the Gel'fand basis of Baird and Biedenharn[3]) for use in Sec. V. We define a $U(N)$ state to be a vector in the orthogonal basis of some $U(N)$ irreducible representation. Each $U(N)$ state is an eigenvector of the Casimir operators of the chain of subgroups $U(N-1), U(N-2), \cdots, U(1)$, and the set of these eigenvalues together with the state's weight constitute a complete set of labels for the state. In addition, each $U(N)$ state x transforms as a state in an irreducible representation with partition μ^p of the subgroup $U(p)$ with respect to the operators of $U(p)$ for $p = N-1, N-2, \cdots, 1$. The set of these $(N-1)$ partitions μ^p are an alternate set of labels to the subgroup Casimir operator eigenvalues. We show how to find the $(N-1)$ partitions and subgroup Casimir operator eigenvalues from the numbered tableau labeling a state.

In Sec. V we find the transformation properties of a $U(N)$ state with respect to a subgroup of the form $\otimes_i U(N_i)$ with $\sum_i N_i \leq N$. In general, the state x transforms like a linear combination of states in irreducible representations of the subgroup if one applies an operator of the subgroup to x. We develop a procedure for finding the states occurring in that linear combination. The procedure results from an analysis of the character of $U(N)$ in a similar fashion to the analysis of Sec. III. In terms of weight diagrams, our procedure finds the weight diagrams of irreducible representations of the subgroup lying in any hyperplane of the weight diagram of the $U(N)$ irreducible representation.

II. THE STRUCTURE OF THE WEIGHTS IN REPRESENTATIONS OF $U(N)$ AND $SU(N)$

In this section, we develop a simple method for finding the weights occurring in an irreducible representation of $U(N)$ or $SU(N)$ for later use in Secs. III and V.

To the defining representation of the group $U(N)$ corresponds an $N \times N$ matrix representation of its algebra in which the matrix corresponding to the generator A^i_k is given by

$$(a^i_k)_{\mu\nu} = \delta_{i\nu}\delta_{k\mu}, \tag{4}$$

with μ and ν labeling rows and columns respectively. As a result, the N weights $\mathbf{V}^j$, for $j = 1, 2, \cdots, N$,

[1] S. Blaha, "The Calculation of the Irreducible Characters of the Symmetric Groups in Terms of the Compound Characters," J. Combinatorial Theory (to be published).

[2] N. Jacobson, *Lie Algebras* (Interscience Publishers, Inc., New York, 1962), Chap. 8; B. Kostant, Trans. Am. Math. Soc. **93**, 53 (1959); N. Straumann, Helv. Phys. Acta **38**, 56 (1965); N. Straumann, Helv. Phys. Acta **38**, 481 (1965); B. Gruber, Ann. Inst. Henri Poincaré **8**, 43 (1968); L. C. Biedenharn, B. Gruber, and H. J. Weber, Proc. Roy. Irish Acad. **67A**, 1 (1968); C. R. Hagen and A. J. MacFarlane, J. Math. Phys. **6**, 1355 (1965); C. R. Hagen and A. J. MacFarlane, J. Math. Phys. **6**, 1366 (1965).

[3] G. E. Baird and L. C. Biedenharn, J. Math. Phys. **4**, 1449 (1963).

called the fundamental weights of $U(N)$, are given by

$$(\mathbf{V}^j)_i = \delta_{ij}, \tag{5}$$

with $i = 1, 2, \cdots, N$ labeling the components of the vector. For $SU(N)$ one need only subtract the trace $\delta_{ik}\delta_{\mu\nu}/N$ on the right side of (4) to obtain the corresponding representation of its generators A^i_k. Consequently, the N fundamental weight vectors of $SU(N)$, $\mathbf{W}^j$, have their ith coordinate given by

$$(\mathbf{W}^j)_i = \delta_{ij} - 1/N. \tag{5'}$$

Before dealing with the weights of other irreducible representations of $U(N)$ and $SU(N)$, we make the following definitions and conventions regarding partitions and Young tableaux. A partition of the integer m is a set of integers whose sum is m and is written in two forms (following the notation of MacMahon[4]),

$$\mu = \mu_1\mu_2\cdots \equiv 1^{k_1}2^{k_2}\cdots,$$

with $\mu_1\mu_2\cdots$ the set of integers placed in nonincreasing order going to the right (i.e., $\mu_1 \geq \mu_2 \geq \cdots$) and with k_1 of the integers being 1's, k_2 of the integers being 2's, $\cdots$. Each of the integers μ_i is called a part of the partition. We define the following addition and subtraction of partitions for $\mu = \mu_1\mu_2\cdots = 1^{k_1}2^{k_2}\cdots$ and $\nu = \nu_1\nu_2\cdots = 1^{m_1}2^{m_2}\cdots$ by

$$\mu + \nu = \alpha_1\alpha_2\cdots, \tag{6a}$$

$$\mu - \nu = \beta_1\beta_2\cdots, \tag{6b}$$

$$[\![(\mu)(\nu)]\!] = 1^{k_1+m_1}2^{k_2+m_2}\cdots, \tag{6c}$$

with $\alpha_i = \mu_i + \nu_i$ and $\beta_i = \mu_i - \nu_i$. We write $\mu > \nu$ if the first nonzero difference $\mu_i - \nu_i$ is greater than zero.

Young frames are labeled by the partition of their row lengths. A standard tableau is any frame whose boxes have been filled with integers in such a way that the integers are nondecreasing going to the right in any row, and strictly increasing going down any column. We define τ^N_μ to be the set of all distinct standard tableaux which can be formed from the frame μ and all choices of integers (with repetitions of integers allowed) from the set $\{1, 2, \cdots, N\}$.

To each irreducible representation of $U(N)$ there corresponds a unique Young frame; and to each irreducible representation of $SU(N)$ there also corresponds a unique Young frame modulo columns of length N. In each case we use the frame's partition to label the irreducible representation. It has been shown[3] that one can label each of the basis vectors of representation μ of $U(N)$ or $SU(N)$ with a tableau from τ^N_μ such that the weight of the basis vector is given by

$$\mathbf{V} = \sum_{i=1}^{N} \rho_i \mathbf{V}^i$$

in $U(N)$, and

$$\mathbf{W} = \sum_{i=1}^{N} \rho_i \mathbf{W}^i$$

in $SU(N)$, with ρ_i being the number of boxes in the labeling tableau containing the integer i. If

$$r = \sum_{i=1}^{N} \rho_i = \sum_{i=1}^{N} \mu_i, \quad \mu = \mu_1\mu_2\cdots,$$

then the jth coordinates of $\mathbf{V}$ and $\mathbf{W}$ are related by

$$(\mathbf{W})_j = (\mathbf{V})_j - r/N. \tag{7}$$

Due to the simple relation between the weights of the irreducible representations μ of $U(N)$ and $SU(N)$, the remainder of Sec. II applies equally well to $U(N)$ and $SU(N)$, though we confine our discussion to $U(N)$.

In Table I we present the representation 21 of $U(3)$ as an example of the labeling of states with τ^3_{21} and their corresponding weights. It is apparent from the example and the definition of τ^N_μ that more than one state in an irreducible representation may have the

TABLE I. Tableau states and weights of the representation 21 of $U(3)$.

State	Weight
$1\ \begin{array}{|c|c|}\hline 1 & 1 \\ \hline 3 \\ \cline{1-1}\end{array}\ >$	(2, 0, 1)
$1\ \begin{array}{|c|c|}\hline 1 & 1 \\ \hline 2 \\ \cline{1-1}\end{array}\ >$	(2, 1, 0)
$1\ \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 2 \\ \cline{1-1}\end{array}\ >$	(1, 2, 0)
$1\ \begin{array}{|c|c|}\hline 2 & 2 \\ \hline 3 \\ \cline{1-1}\end{array}\ >$	(0, 2, 1)
$1\ \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 3 \\ \cline{1-1}\end{array}\ >$	(0, 1, 2)
$1\ \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 3 \\ \cline{1-1}\end{array}\ >$	(1, 0, 2)
$1\ \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array}\ >$	(1, 1, 1)
$1\ \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array}\ >$	(1, 1, 1)

[4] P. A. MacMahon, *Combinatory Analysis* (Chelsea Publ. Co., New York, 1960), p. 1.

TABLE II. Weyl class structure of the representations 42 of $U(5)$ and $SU(5)$.

Weyl class ν	M_ν^{42}	Dominant weights $U(N)$	Dominant weights $SU(N)$	Dominant weights $SU(N)$ (Cartan form)
42	1	(4, 2, 0, 0, 0)	$\frac{1}{5}(4, 4, -6, -6, -6)$	$\left(\sqrt{2}, \sqrt{6}, \sqrt{3}, \frac{3}{\sqrt{5}}\right)$
41^2	1	(4, 1, 1, 0, 0)	$\frac{1}{5}(14, -1, -1, -6, -6)$	$\left(\frac{3}{\sqrt{2}}, \sqrt{\frac{3}{2}}, \sqrt{3}, \frac{3}{\sqrt{5}}\right)$
3^2	1	(3, 3, 0, 0, 0)	$\frac{1}{5}(9, 9, -6, -6, -6)$	$\left(0, \sqrt{6}, \sqrt{3}, \frac{3}{\sqrt{5}}\right)$
321	2	(3, 2, 1, 0, 0)	$\frac{1}{5}(9, 4, -1, -6, -6)$	$\left(\frac{1}{\sqrt{2}}, \sqrt{\frac{3}{2}}, \sqrt{3}, \frac{3}{\sqrt{5}}\right)$
31^3	3	(3, 1, 1, 1, 0)	$\frac{1}{5}(9, -1, -1, -1, -6)$	$\left(\sqrt{2}, \sqrt{\frac{2}{3}}, \frac{1}{\sqrt{3}}, \frac{3}{\sqrt{5}}\right)$
2^3	3	(2, 2, 2, 0, 0)	$\frac{1}{5}(4, 4, 4, -6, -6)$	$\left(0, \sqrt{\frac{2}{3}}, \sqrt{3}, \frac{3}{\sqrt{5}}\right)$
2^21^2	4	(2, 2, 1, 1, 0)	$\frac{1}{5}(4, 4, -1, -1, -6)$	$\left(0, \sqrt{\frac{2}{3}}, \frac{1}{\sqrt{3}}, \frac{3}{\sqrt{5}}\right)$
21^4	6	(2, 1, 1, 1, 1)	$\frac{1}{5}(4, -1, -1, -1, -1)$	$\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{12}}, \frac{1}{\sqrt{20}}\right)$

same weight. The multiplicity of a weight in an irreducible representation is the number of states which have it as a weight. Thus we have a method[3] of computing the multiplicity of a weight $\mathbf{V} = (V_1, V_2, \cdots, V_N)$ in irreducible representation μ by counting the number of standard tableaux which can be formed from frame μ, and V_1 ones, V_2 twos, $\cdots$, and V_N N's. In Sec. III we present explicit formulas for computing multiplicities which do not rely on counting tableaux.

The weights of $U(N)$ representations may be divided into equivalence classes using the Weyl group which is the group of all permutations of the coordinates of a weight. For $U(N)$ the Weyl group is S_N, the symmetric group on N objects. All weights of an equivalence class are equal to within a permutation of the Weyl group. As a result, we can define the partition of a Weyl class to be the partition whose parts are the coordinates of any weight in the class. The dominant weight of the Weyl class $\nu = \nu_1\nu_2\cdots$ is $\sum_i \nu_i \mathbf{V}^i$. The highest weight of the representation $\mu = \mu_1\mu_2\cdots$ is $\sum_i \mu_i \mathbf{V}^i$.

We define M_ν^μ to be the multiplicity of each weight in the Weyl class ν of the representation μ. (All weights in any Weyl class of a representation have the same multiplicity.) As examples, we give tables of M_ν^μ in Appendix A for each irreducible representation of $U(N)$ corresponding to a partition of any integer less than eight.

At this point, we have a simple method of obtaining all weights of an irreducible representation μ. The Weyl class ν (ν of course has less than $N+1$ parts) occurs in the representation μ if M_ν^μ is nonzero, and, as a result, all weights may be obtained by applying permutations to the dominant weights of the classes occurring in the representation. One can obtain the weights of the two forms of $SU(N)$ through the use of Eqs. (7) and (3′). As an example, we present in Table II the partitions, multiplicities, and dominant weights of the representations 42 of $U(5)$ and $SU(5)$ which have dimension 420.

III. CHARACTER ANALYSIS OF $U(N)$ REPRESENTATIONS

In this section, we analyze the characters of $U(N)$ representations in order to obtain formulas for multiplicities, the content of representations of $U(N)$ with respect to subgroups of the general form $U(N_1) \otimes U(N_2) \otimes \cdots \otimes U(N_p) \equiv \otimes_i U(N_i)$ $(\sum_i N_i = N)$, the content of irreducible representations of $U(NM)$ with respect to representations of $U(N) \otimes U(M)$, and the coefficients occurring in Clebsch–Gordan series reductions.

The character[5] of the irreducible representation μ of $U(N)$ is a function of N variables $\boldsymbol{\phi} = (\phi_1, \phi_2, \cdots, \phi_N)$:

$$\chi_\mu = \sum \exp(i\mathbf{V}\cdot\boldsymbol{\phi}), \tag{8a}$$

where the sum is over all weights $\mathbf{V}$ in μ. For the corresponding $SU(N)$ representation μ, the sum is over all $SU(N)$ weights of μ with the variables $\boldsymbol{\phi}$

[5] H. Weyl, *The Classical Groups* (Princeton University Press, Princeton, N.J., 1939), p. 134.

restricted by the condition $\sum_{i=1}^{N} \phi_i = 0$. Owing to this restriction and Eq. (7), the characters of representation μ of $U(N)$ and $SU(N)$ are identical if μ has no columns of length N; they are equal to within a phase $\exp(ik \sum_{P=1}^{N} \phi_P)$ with k the number of columns of μ of length N, otherwise. Therefore, all succeeding results apply to $SU(N)$ as well as $U(N)$ if one drops columns of length N from frames (and partitions) *after* performing all indicated operations.

In Sec. II we showed that the weights of an irreducible representation may be divided into Weyl classes in which each weight of the class may be obtained by applying some S_N permutation to the dominant weight. In particular, one need only apply a subset of the permutations of S_N to obtain all weights of the class: For the Weyl class $\nu = 1^{\alpha}2^{\beta}3^{\gamma}\cdots$ with $\alpha + \beta + \gamma + \cdots = k$, one applies

$$N!/(N-k)!\,\alpha!\,\beta!\,\gamma!\cdots$$

permutations, which is the number of distinct weights in the class. We call S_ν this minimal set of permutations. If the weight $\mathbf{V} = \sum_i \beta_i \mathbf{V}^i$ and we let $x_k = \exp(i\mathbf{V}^k \cdot \boldsymbol{\phi}) = \exp(i\phi_k)$, then

$$\exp(i\mathbf{V}\cdot\boldsymbol{\phi}) = x_1^{\beta_1} x_2^{\beta_2} \cdots x_N^{\beta_N}.$$

We may therefore rewrite (8a) in the form

$$\chi_\mu = \sum_\nu M_\nu^\mu \sum_{Q\in S_\nu} x_1^{Q\nu_1} x_2^{Q\nu_2} \cdots x_N^{Q\nu_N} = \sum_\nu M_\nu^\mu \sum_{Q\in S_\nu} x^{Q\nu}, \tag{8b}$$

where we sum over partitions ν and the permutations Q of S_ν, and we let $x_1^{Q\nu_1} x_2^{Q\nu_2} \cdots x_N^{Q\nu_N} = x^{Q\nu}$.

Another well-known[6] form for $U(N)$ characters is

$$\chi_\mu = \frac{\sum_{P\in S_N} \delta_P \exp[iP(\mathbf{V}_\mu + R_N)\cdot\boldsymbol{\phi}]}{\sum_{Q\in S_N} \delta_Q \exp(iQR_N\cdot\boldsymbol{\phi})},$$

where $R_N = (N-1, N-2, \cdots, 0)$, $\mathbf{V}_\mu = \sum_i \mu_i \mathbf{V}^i$, and δ_P is the signature of the permutation P. In the notation of (8b) this becomes

$$\chi_\mu = \sum_{P\in S_N} \delta_P x^{P(\mu+R_N)} \Big/ \sum_{Q\in S_N} \delta_Q x^{QR_N}. \tag{9}$$

Equations (8) and (9) can be shown to be equivalent using the Frobenius equations in the theory of symmetric functions. In particular, if

$$s_r = \sum_{i=1}^{N} x_i^r \quad \text{for} \quad r = 1, 2, \cdots, N,$$

if $\kappa = 1^{k_1}2^{k_2}\cdots$ is a partition, and if

$$D_N(x) = \prod_{i<j} (x_i - x_j) = \sum_{Q\in S_N} \delta_Q x^{QR_N},$$

then[7] [in the notation of (8b)]

$$S_\kappa = s_1^{k_1} s_2^{k_2} \cdots = \sum_\mu \phi_\kappa^\mu \sum_{Q\in S_\mu} x^{Q\mu} \tag{10}$$

and

$$S_\kappa D_N(x) = \sum_\lambda \chi_\kappa^\lambda \sum_{P\in S_N} \delta_P x^{P(\lambda+R_N)}, \tag{11}$$

where ϕ_κ^μ is a Frobenius compound character and χ_κ^λ is an irreducible character of the symmetric group. It has been shown[8] that

$$\phi_\kappa^\mu = \sum_\nu M_\mu^\nu \chi_\kappa^\nu, \tag{12}$$

where the integers M_μ^ν will turn out to be multiplicities previously defined. From Eqs. (10)–(12), we have

$$\sum_{\mu,\nu} M_\mu^\nu \chi_\kappa^\nu \sum_{Q\in S_\mu} x^{Q\mu} = \sum_\lambda \chi_\kappa^\lambda \sum_{P\in S_N} \delta_P x^{P(\lambda+R_N)} / D_N(x).$$

If we multiply by $(g_\kappa/g)\chi_\kappa^\beta$ (g_κ is the number of elements in class κ, and $g = N!$ is the number of elements in S_N) and sum over κ, we obtain

$$\sum_\mu M_\mu^\beta \sum_{Q\in S_\mu} x^{Q\mu} = \sum_{P\in S_N} \delta_P x^{P(\beta+R_N)} / D_N(x) \tag{13}$$

(using the orthogonality of irreducible characters of the symmetric group), which demonstrates the equivalence of (8) and (9) and establishes a formula for multiplicities,

$$M_\nu^\rho = (1/g) \sum_\alpha g_\alpha \chi_\alpha^\rho \phi_\alpha^\nu, \tag{14}$$

using Eq. (12) and the orthogonality of irreducible characters. Below, we will derive another method for calculating M_ν^ρ via recurrence relations.

The integration volume element of $U(N)$ is[9]

$$\int_0^{2\pi} \cdots \int_0^{2\pi} \frac{d\phi_1\, d\phi_2 \cdots d\phi_N}{N!\,(2\pi)^N} \overline{D_N}(x) D_N(x) \equiv \int dU \tag{15}$$

and the orthogonality condition for the characters is

$$\int dU \bar{\chi}_\nu \chi_\mu = \delta_{\mu\nu},$$

with $\delta_{\mu\nu}$ a Kronecker delta in partitions, $\overline{D_N}(x)$ the complex conjugate of $D_N(x)$, and $\bar{\chi}_\nu$ the complex conjugate of χ_ν.

Next, we show that, in the reduction of the irreducible representation μ of $U(M+N)$ with respect to the

[6] Reference 5, p. 200.

[7] D. E. Littlewood, *The Theory of Group Characters* (Oxford University Press, Oxford, England, 1950), pp. 63, 67.

[8] Reference 7, p. 71.

[9] Reference 7, p. 219.

direct product subgroup $U(N) \otimes U(M)$, the number of times $\alpha \otimes \beta$ appears [α is an irreducible representation of $U(N)$ and β of $U(M)$] is given by

$$\mathcal{N}^{\mu}_{\alpha\otimes\beta} = \sum_{\substack{P\in S_N \\ Q\in S_M}} \delta_P \delta_Q M^{\mu}_{[(\alpha+R_N-PR_N)(\beta+R_M-QR_M)]}, \quad (16)$$

where $[(\alpha + R_N - PR_N)(\beta + R_M - QR_M)]$ is evaluated in accordance with the definitions of Sec. II, and terms in the sum are ignored when either expression $\alpha + R_N - PR_N$ or $\beta + R_M - QR_M$ has negative parts. Note that the sum of the parts of α and β must equal the sum of the parts of μ. For example, the representation 321 of $U(6)$ contains the representation 21 $\otimes$ 21 of $U(3) \otimes U(3)$ twice,

$$\begin{aligned}\mathcal{N}^{321}_{21\otimes 21} &= \sum_{\substack{P\in S_3 \\ Q\in S_3}} \delta_P \delta_Q M^{321}_{[(420-P(210))(420-Q(210))]} \\ &= M^{321}_{2^2 1^2} - 2M^{321}_{321} + M^{321}_{3^2} = 2,\end{aligned}$$

since $M^{321}_{2^2 1^2} = 4$, $M^{321}_{321} = 1$, and $M^{321}_{3^2} = 0$.

The proof of (16) starts with the standard character expression,

$$\begin{aligned}\mathcal{N}^{\mu}_{\alpha\otimes\beta} &= \int dU\, dV \bar{\chi}_{\alpha}(U)\bar{\chi}_{\beta}(V)\chi_{\mu} \\ &= \int \frac{d\boldsymbol{\phi}\, d\boldsymbol{\psi} D_N(x) D_M(y)}{N!\, M!\, (2\pi)^{N+M}} \sum_{S\in S_N} \delta_S x^{-S(\alpha+R_N)} \\ &\quad \times \sum_{T\in S_M} \delta_T y^{-T(\beta+R_M)} \sum_{\nu} M^{\mu}_{\nu} \sum_{Q\in S_\nu} z^{Q\nu}, \end{aligned} \quad (17)$$

where $\boldsymbol{\phi} = (\phi_1, \phi_2, \cdots, \phi_N), \boldsymbol{\psi} = (\psi_1, \psi_2, \cdots, \psi_M)$, $x_j = \exp(i\mathbf{V}^j \cdot \boldsymbol{\phi}) = \exp(i\phi_j)$ for $j = 1, 2, \cdots, N$, $y_j = \exp(i\mathbf{V}^j \cdot \boldsymbol{\psi}) = \exp(i\psi_j)$ for $j = 1, 2, \cdots, M$, and $z^{\gamma} = x_1^{\gamma_1} x_2^{\gamma_2} \cdots x_N^{\gamma_N} y_1^{\gamma_{N+1}} \cdots y_M^{\gamma_{N+M}}$ for $\gamma = (\gamma_1, \gamma_2, \cdots, \gamma_{N+M})$. The product

$$\begin{aligned}&\sum_{P\in S_N} \delta_P x^{PR_N} \sum_{V\in S_M} \delta_V y^{VR_M} \sum_{S\in S_N} \delta_S x^{-S(\alpha+R_N)} \\ &\qquad \times \sum_{T\in S_M} \delta_T y^{-T(\beta+R_M)} \sum_{\nu} M^{\mu}_{\nu} \sum_{Q\in S_\nu} z^{Q\nu} \end{aligned} \quad (18)$$

can be written as a sum over monomials:

$$x_1^a \cdots x_N^b y_1^c \cdots y_M^d.$$

No monomial with one or more of the exponents $a, \cdots, b, c, \cdots, d$ nonzero, will contribute in the integral of (17). Thus,

$$\begin{aligned}&N!\, M!\, \mathcal{N}^{\mu}_{\alpha\otimes\beta} \\ &\quad = \sum_{P\in S_N} \sum_{V\in S_M} \sum_{S\in S_N} \sum_{T\in S_M} \sum_{\nu} \sum_{Q\in S_\nu} \delta_P \delta_V \delta_S \delta_T M^{\mu}_{\nu}, \end{aligned} \quad (19)$$

where the summations are restricted by the condition $O = PR_N - S(\alpha + R_N) + Q\nu_{(N)}$ and $O = VR_M - T(\beta + R_M) + Q\nu_{(M)}$ with $Q\nu_{(N)} = (Q\nu_1, Q\nu_2, \cdots, Q\nu_N)$ and $Q\nu_{(M)} = (Q\nu_{N+1}, Q\nu_{N+2}, \cdots, Q\nu_{N+M})$. The restriction on the summations can be rewritten as $Q\nu = [(S(\alpha + R_N) - PR_N)(T(\beta + R_M) - VR_M)]$. Thus, Q and ν are determined if α, β, S, P, T, and V are given. Therefore, the restriction on the summations may be implemented by dropping the sums over ν and Q after making the convention to ignore the order of entries in the subscript ν of M^{μ}_{ν} (i.e., $M^{\mu}_{\nu} = M^{\mu}_{Q\nu}$ for all Q) and substituting the expression for $Q\nu$ in M^{μ}_{ν} in Eq. (19):

$$\begin{aligned}N!\, M!\, \mathcal{N}^{\mu}_{\alpha\otimes\beta} &= \sum_{P\in S_N} \sum_{S\in S_N} \sum_{V\in S_M} \sum_{T\in S_M} \delta_P \delta_S \delta_V \delta_T \\ &\quad \times M^{\mu}_{[(S(\alpha+R_N)-PR_N)(T(\beta+R_M)-VR_M)]} \\ &= \sum_{P} \sum_{S} \sum_{V} \sum_{T} \delta_{S^{-1}P} \delta_{T^{-1}P} \\ &\quad \times M^{\mu}_{[(\alpha+R_N-S^{-1}PR_N)(\beta+R_M-T^{-1}VR_M)]}, \end{aligned} \quad (20)$$

using $\delta_S \delta_P = \delta_{S^{-1}P}$, $\delta_V \delta_T = \delta_{T^{-1}V}$ and the convention previously stated to write $M^{\mu}_{[(S\gamma)(T\sigma)]} = M^{\mu}_{[(\gamma)(\sigma)]}$ [with $S(\alpha + R_N) - PR_N = S(\alpha + R_N - S^{-1}PR_N)$ and

$$T(\beta + R_M) - VR_M = T(\beta + R_M - T^{-1}VR_M)].$$

If we let $H = S^{-1}P$ and $G = T^{-1}V$, then

$$\sum_{P\in S_N} \sum_{S\in S_N} = N! \sum_{H\in S_N} \quad \text{and} \quad \sum_{V\in S_M} \sum_{T\in S_M} = M! \sum_{G\in S_M}$$

in Eq. (20), and we obtain Eq. (16).

In the case $M = 0$, Eq. (16) can be used to compute multiplicities. Since we may drop β, $\beta + R_M - QR_M$, and the sum over Q, Eq. (16) results in

$$\delta^{\mu}_{\alpha} = \sum_{P\in S_N} \delta_P M^{\mu}_{\alpha+R_N-PR_N}, \quad (21)$$

with $\mathcal{N}^{\mu}_{\alpha} = \delta^{\mu}_{\alpha}$, a Kronecker delta in partitions. The reduction of any irreducible representation μ of $U(N)$ with respect to the trivial subgroup $U(N)$ yields μ and only μ exactly once, resulting in the Kronecker delta in (21). An example of Eq. (21) is

$$\delta^{\mu}_{2^3} = M^{\mu}_{2^3} - 2M^{\mu}_{321} - M^{\mu}_{42} + M^{\mu}_{41^2} + M^{\mu}_{3^2}.$$

From this example it is clear that one may choose N to be the number of parts of α without loss of generality. Equation (21) gives a recurrence relation for M^{μ}_{ν} in terms of M^{μ}_{α} where $\alpha > \nu$. Since Eq. (21) implies $M^{\mu}_{\nu} = 0$ for $\nu > \mu$ and one for $\mu = \nu$, we can find the multiplicities in any representation μ recursively. The matrix of M^{μ}_{ν} for all pairs of partitions μ and ν of the integer m is presented in Appendix A for $m = 2, 3, \cdots, 7$, and, as a result, one has the multiplicities and Weyl classes for each representation of $U(N)$, corresponding to a partition of any integer less than eight by inspection of these tables.

Equation (21) could have been written in the form[1]

$$\delta_\alpha^\mu = \sum_\nu B_\alpha^\nu M_\nu^\mu, \tag{21'}$$

where B_α^ν is the matrix inverse to the M_ν^μ matrix (with partitions labeling rows and columns). In general, any equation of the form

$$Z_\nu = \sum_{P\in S_N} \delta_P Y_{\nu+R_N-PR_N} \tag{22a}$$

has the equivalent form

$$Z_\nu = \sum_\beta B_\nu^\beta Y_\beta. \tag{22b}$$

In Appendix B, we present tables of B_ν^μ to aid in the evaluation of formulas of the preceding type. Explicit formulas for certain classes of the coefficients B_ν^μ have been given elsewhere.[1] One can thus evaluate (16) in the following form:

$$\mathcal{N}^\mu_{\alpha\otimes\beta} = \sum_{\delta,\gamma} B_\alpha^\delta B_\beta^\gamma M^\mu_{[(\delta)(\gamma)]}. \tag{16'}$$

For typographic convenience, we use the equivalence of (22a) and (22b) to write formulas in terms of the coefficients B_ν^μ whenever possible in the remainder of this paper. Some general properties of the coefficients M_ν^μ are the following:

(i) $M_\nu^\mu = 0$ if $\nu > \mu$;

(ii) $M_{1^m}^\mu = \chi_{1^m}^\mu$, the dimension of the irreducible representation μ of S_N;

(iii) $M_\alpha^m = 1$ for all α;

(iv) $M_\nu^\mu = 1$ if $\mu = \nu$;

(v) $\sum_{k\le N} \frac{N!}{(N-k)!} \sum_{\substack{\alpha,\beta,\cdots \\ \alpha+\beta+\cdots=k}} \frac{M^\mu_{1^\alpha 2^\beta\cdots}}{\alpha!\,\beta!\cdots}$ = the dimension of representation μ of $SU(N)$.

Property (iii) implies that all weights of $U(2)$ have multiplicity one, as is well known.

If we apply Eq. (16) to the case $U(N-1)\otimes U(1)$, we have

$$\mathcal{N}^\mu_{\alpha\otimes j} = \sum_\nu B_\alpha^\nu M^\mu_{[(\nu)(j)]}, \tag{23}$$

where j is an integer, $0 \le j \le \mu_1$.

Equation (16) can be generalized to give the content of the reduction of the representation μ of $U(\sum_{i=1}^k N_i)$ with respect to the direct product subgroup $\otimes_{i=1}^k U(N_i)$. The number of times μ contains $\rho^1\otimes\rho^2\otimes\cdots\otimes\rho^k = \otimes_i \rho^i$ [where ρ^i is an irreducible representation of $U(N_i)$] is

$$\mathcal{N}^\mu_{\otimes_i \rho^i} = \sum_{\nu_1,\nu_2,\cdots,\nu_k} B_{\rho^1}^{\nu_1} B_{\rho^2}^{\nu_2}\cdots B_{\rho^k}^{\nu_k} M^\mu_{[(\nu_1)(\nu_2)\cdots(\nu_k)]}. \tag{24}$$

A special case of this is

$$\mathcal{N}^\mu_{\alpha_1\otimes\alpha_2\otimes\cdots\otimes\alpha_k} = M_\alpha^\mu, \tag{24'}$$

where α_i is an integer for all i and $\alpha = \alpha_1\alpha_2\alpha_3\cdots\alpha_k$ is a partition.

We now consider the Clebsch–Gordan series reduction of tensor products of irreducible representations in $U(N)$. We will establish that the number of times the irreducible representation μ occurs in the tensor product of α and β, denoted by $\alpha\otimes\beta$, with μ, α, and β irreducible representations of $U(N)$, is

$$\mathcal{N}_\mu^{\alpha\otimes\beta} = \sum_\gamma B_\mu^\gamma M_\gamma^{\alpha\otimes\beta}, \tag{25}$$

where

$$M_\gamma^{\alpha\otimes\beta} = \sum_\nu \sum_{Q\in S_\nu} M^\alpha_{\gamma-Q\nu} M_\nu^\beta \tag{26}$$

is the multiplicity of the Weyl class γ in the tensor product representation $\alpha\otimes\beta$ of $U(N)$. Equation (26) reflects the fact that the weight γ in $\alpha\otimes\beta$ is the sum over all possible pairs of weights $\gamma - Q\nu$ in α and $Q\nu$ in β. The contribution to the multiplicity $M_\gamma^{\alpha\otimes\beta}$ for a particular $\gamma - Q\nu$ and $Q\nu$ is $M^\alpha_{\gamma-Q\nu} M^\beta_{Q\nu} = M^\alpha_{\gamma-Q\nu} M_\nu^\beta$.

The proof of Eq. (25) starts with the standard expression

$$\begin{aligned}\mathcal{N}_\mu^{\alpha\otimes\beta} &= \int \frac{d\Phi}{(2\pi)^N N!}\, \overline{D_N}(x) D_N(x) \bar\chi_\alpha \bar\chi_\beta \chi_\mu \\ &= \int \frac{d\Phi}{(2\pi)^N N!} \sum_{P\in S_N} \delta_P x^{-PR_N} \sum_\nu M_\nu^\alpha \\ &\quad\times \sum_{Q\in S_\nu} x^{-Q\nu} \sum_\gamma M_\gamma^\beta \sum_{T\in S_\gamma} x^{-T\gamma} \sum_{S\in S_N} \delta_S x^{S(\mu+R_N)}.\end{aligned}$$

As in the previous proof, we have

$$N!\,\mathcal{N}_\mu^{\alpha\otimes\beta} = \sum_{P\in S_N}\sum_\nu \sum_{Q\in S_\nu}\sum_\gamma \sum_{T\in S_\gamma}\sum_{S\in S_N} \delta_P\delta_S M_\nu^\alpha M_\gamma^\beta,$$

where the sums are restricted to satisfy

$$0 = S(\mu + R_N) - PR_N - Q\nu - T\gamma.$$

If we implement the restriction by letting

$$M_\gamma^\beta = M^\beta_{T^{-1}[S(\mu+R_N)-PR_N-Q\nu]} = M^\beta_{S(\mu+R_N)-PR_N-Q\nu}$$

and by dropping the sums over γ and TeS_γ, then

$$\begin{aligned}N!\,\mathcal{N}_\mu^{\alpha\otimes\beta} &= \sum_{P\in S_N}\sum_{S\in S_N}\sum_\nu\sum_{Q\in S_\nu} \delta_P\delta_S M_\nu^\alpha M^\beta_{S(\mu+R_N)-PR_N-Q\nu} \\ &= \sum_{P\in S_N}\sum_{S\in S_N} \delta_{S^{-1}P} M^{\alpha\otimes\beta}_{\mu+R_N-S^{-1}PR_N}.\end{aligned}$$

Using Eq. (26), $\delta_P\delta_S = \delta_{S^{-1}P}$ and $M^{\alpha\otimes\beta}_{S\pi} = M^{\alpha\otimes\beta}_\pi$ for any permutation S and partition π. If we let $U = S^{-1}P$, then $\sum_{P\in S_N}\sum_{S\in S_N} = N!\sum_{U\in S_N}$ in the previous equation, which proves Eq. (25).

Equations (25) and (26) have simple generalizations to tensor products with an arbitrary number of factors ($\otimes \alpha^i = \alpha^1 \otimes \alpha^2 \otimes \cdots \otimes \alpha^{k+1}$):

$$\mathcal{N}_\mu^{\otimes \alpha^i} = \sum_\beta B_\mu^\beta M_\beta^{\otimes \alpha^i}, \tag{27}$$

$$\begin{aligned} M_\beta^{\otimes \alpha^i} &= \sum_{\nu_1} \sum_{Q_1 \in S_{\nu_1}} \sum_{\nu_2} \sum_{Q_2 \in S_{\nu_2}} \cdots \sum_{\nu_k} \sum_{Q_k \in S_{\nu_k}} M_{\nu_1}^{\alpha^1} \\ &\quad \times M_{\nu_2}^{\alpha^2} \cdots M_{\nu_k}^{\alpha^k} M_{\beta - Q_1\nu_1 - Q_2\nu_2 - \cdots - Q_k\nu_k}^{\alpha^{k+1}} \\ &= \sum_{\nu_1} \sum_{Q_1 \in S_{\nu_1}} \cdots \sum_{\nu_k} \sum_{Q_k \in S_{\nu_k}} M_{\beta - Q_1\nu_1}^{\alpha^1} \\ &\quad \times M_{\nu_1 - Q_2\nu_2}^{\alpha^2} \cdots M_{\nu_{k-1} - Q_k\nu_k}^{\alpha^k} M_{\nu_k}^{\alpha^{k+1}}. \end{aligned} \tag{28}$$

Another formula for $\mathcal{N}_\mu^{\alpha \otimes \beta}$ which can be derived in a similar fashion to Eq. (25) and has been given by Weyl[10] is

$$\mathcal{N}_\mu^{\alpha \otimes \beta} = \sum_{P \in S_N} \delta_P M_{\mu + R_N - P(\beta + R_N)}^\alpha. \tag{29}$$

Equation (26) can be expressed solely in terms of characters of the symmetric group

$$\begin{aligned} M_\alpha^{\mu \otimes \nu} &= \sum_\gamma \sum_{Q \in S_\gamma} M_{\alpha - Q\gamma}^\mu M_\gamma^\nu \\ &= \sum_\gamma \sum_{Q \in S_\gamma} \sum_\kappa \sum_\lambda \frac{g_\kappa g_\lambda'}{g g'} \chi_\lambda^\mu \phi_\lambda^{\alpha - Q\gamma} \chi_\kappa^\nu \phi_\kappa^\gamma \\ &= \sum_\kappa \sum_\lambda \frac{g_\kappa g_\lambda'}{g g'} \chi_\kappa^\nu \chi_\lambda^\mu \sum_\gamma \sum_{Q \in S_\gamma} \phi_\kappa^\gamma \phi_\lambda^{\alpha - Q\gamma} \\ &= \sum_\kappa \sum_\lambda \frac{g_\kappa g_\lambda'}{g g'} \chi_\kappa^\nu \chi_\lambda^\mu \phi_{[(\kappa)(\lambda)]}^\alpha, \end{aligned} \tag{30}$$

using Eq. (14) with g_κ and g from the symmetric group of which χ_κ^ν is a character, and g_λ' and g' from the symmetric group of which χ_λ^μ is a character. We also use the identity

$$\phi_{[(\kappa)(\lambda)]}^\alpha = \sum_\gamma \sum_{Q \in S_\gamma} \phi_\kappa^\gamma \phi_\lambda^{\alpha - Q\gamma},$$

which has been proven elsewhere.[1] In general,

$$\begin{aligned} M_{\mu^{\otimes_{i=1}^p} \alpha^i} &= \sum_{\lambda_1, \lambda_2, \cdots, \lambda_p} \frac{g_{\lambda_1} g_{\lambda_2} \cdots g_{\lambda_p}}{g_1 g_2 \cdots g_p} \\ &\quad \times \chi_{\lambda_1}^{\alpha^1} \chi_{\lambda_2}^{\alpha^2} \cdots \chi_{\lambda_p}^{\alpha^p} \phi_{[(\lambda_1)(\lambda_2)\cdots(\lambda_p)]}^\mu. \end{aligned} \tag{31}$$

Using the identity[1]

$$\chi_\nu^\mu = \sum_\beta B_\mu^\beta \phi_\nu^\beta, \tag{32}$$

Eqs. (31) and (27) yield

$$\begin{aligned} \mathcal{N}_{\mu^{\otimes_{i=1}^p} \alpha^i} &= \sum_{\lambda_1, \lambda_2, \lambda_3, \cdots, \lambda_p} \frac{g_{\lambda_1} g_{\lambda_2} \cdots g_{\lambda_p}}{g_1 g_2 \cdots g_p} \\ &\quad \times \chi_{\lambda_1}^{\alpha^1} \chi_{\lambda_2}^{\alpha^2} \cdots \chi_{\lambda_p}^{\alpha^p} \chi_{[(\lambda_1)(\lambda_2)\cdots(\lambda_p)]}^\mu. \end{aligned} \tag{33}$$

[10] Reference 5, p. 231.

A special case of (31) is ($\otimes_{i=1}^p 1 = 1 \otimes 1 \otimes \cdots \otimes 1$),

$$M_{\alpha^{\otimes_{i=1}^p} 1} = \phi_{1^p}^\alpha,$$

which implies

$$\mathcal{N}_{\alpha^{\otimes_{i=1}^p} 1} = \chi_{1^p}^\alpha$$

and shows that the $\chi_{1^p}^\alpha$ Young projection operators corresponding to the frame α are sufficient to project out all $\mathcal{N}_\alpha^{\otimes_i 1}$ (differently constructed) representations α in a totally unsymmetrized tensor with p subscripts.

Now we show that

$$M_{\nu^{\otimes_{i=1}^p} \alpha^i} = \sum_\mu \mathcal{N}_{\mu^{\otimes_{i=1}^p} \alpha^i} M_\nu^\mu, \tag{34}$$

which expresses the multiplicity of a Weyl class in a tensor product representation as a sum over the multiplicities of the Weyl class in irreducible representations occurring in its Clebsch–Gordan series decomposition. Multiplication of Eq. (27) by M_ν^μ and summation over μ yields Eq. (34) immediately by Eq. (21′).

We now calculate the content of the tensor product representation $\alpha \otimes \beta$ of $U(N+M)$ [α and β are irreducible representations of $U(N+M)$] when reduced with respect to the subgroup $U(N) \otimes U(M)$. In particular, the number of times the representation $\mu \otimes \nu$ of $U(N) \otimes U(M)$ [μ is a representation of $U(N)$ and ν of $U(M)$] occurs in the reduction of the tensor product $\alpha \otimes \beta$ is

$$\mathcal{N}_{\mu \otimes \nu}^{\alpha \otimes \beta} = \sum_\lambda \sum_\kappa B_\mu^\lambda B_\nu^\kappa M_{[(\lambda)(\kappa)]}^{\alpha \otimes \beta}. \tag{35}$$

The proof of (35) starts with the identity

$$\mathcal{N}_{\mu \otimes \nu}^{\alpha \otimes \beta} = \sum_\gamma \mathcal{N}_\gamma^{\alpha \otimes \beta} \mathcal{N}_{\mu \otimes \nu}^\gamma, \tag{35′}$$

where the sum is over all representations γ of $U(N+M)$ occurring in the Clebsch–Gordan series decomposition of $\alpha \otimes \beta$. If we use Eq. (16′), we obtain

$$\begin{aligned} \mathcal{N}_{\mu \otimes \nu}^{\alpha \otimes \beta} &= \sum_\gamma \mathcal{N}_\gamma^{\alpha \otimes \beta} \sum_\lambda \sum_\kappa B_\mu^\lambda B_\nu^\kappa M_{[(\lambda)(\kappa)]}^\gamma \\ &= \sum_\lambda \sum_\kappa B_\mu^\lambda B_\nu^\kappa \sum_\gamma \mathcal{N}_\gamma^{\alpha \otimes \beta} M_{[(\lambda)(\kappa)]}^\gamma, \end{aligned}$$

which results in Eq. (35) due to Eq. (34). Equation (35) can be easily generalized in the manner of Eqs. (24) and (27).

The number of times the representation $\alpha \otimes \beta$ of $U(N) \otimes U(M)$ occurs in the reduction of the irreducible representation μ of $U(NM)$ is given by

$$\mathcal{M}_{\alpha \otimes \beta}^\mu = \sum_\lambda \sum_{Q \in S_\lambda} B_\alpha^{(Q\lambda)_N} B_\beta^{(Q\lambda)_M} M_\lambda^\mu, \tag{36}$$

where the partitions

$$(Q\lambda)_N = \left(\sum_{j=0}^{M-1} \nu_{jN+1}, \sum_{j=0}^{M-1} \nu_{jN+2}, \cdots, \sum_{j=0}^{M-1} \nu_{(j+1)N} \right)$$

and

$$(Q\lambda)_M = \left(\sum_{k=1}^{N} \nu_k, \sum_{k=1}^{N} \nu_{N+k}, \cdots, \sum_{k=1}^{N} \nu_{(M-1)N+k} \right)$$

with

$$(\nu_1, \nu_2, \cdots, \nu_{NM}) = Q(\lambda_1, \lambda_2, \cdots, \lambda_{NM}).$$

Equation (36) follows from the expression

$$\mathcal{M}^{\mu}_{\alpha\otimes\beta} = \int \frac{d\phi\, d\psi}{N!\, M!\, (2\pi)^{N+M}} \sum_{\lambda} M^{\mu}_{\lambda} \sum_{Q\in S_\lambda} x^{-Q\lambda} \overline{D_N}(y)\, \overline{D_M}(z) \times \sum_{P\in S_N} \delta_P y^{P(\alpha+R_N)} \sum_{Q\in S_M} \delta_Q z^{Q(\beta+R_M)},$$

with $x_{jN+k} = y_k z_{j+1}$ for $j = 0, 1, \cdots, (M-1)$ and $k = 1, 2, \cdots, N$; $y_k = \exp(i\phi_k)$ for $k = 1, 2, \cdots, N$ and $z_j = \exp(i\psi_j)$ for $j = 1, 2, \cdots, M$. Equation (36) can be directly generalized to the reduction of irreducible and tensor product representations of $U(\prod_i N_i)$ with respect to $\otimes_i U(N_i)$.

IV. SUBGROUP LABELING OF $U(N)$ STATES

Baird and Biedenharn[3] have shown that a complete set of labels for a state x in a $U(N)$ irreducible representation is given by its weight and eigenvalues with respect to the Casimir operators of the chain of subalgebras $U(N-1)$, $U(N-2), \cdots, U(1)$ [with the subalgebra $U(p)$ defined to have the generators A^i_j, for all i and j less than $p+1$]. Since x is an eigenstate of the Casimir operators of the subalgebra $U(p)$, it transforms as a state in an irreducible representation μ^p of $U(p)$ with respect to the operators in $U(p)$ for $p = N-1, N-2, \cdots, 1$. The $N-1$ partitions μ^p constitute an alternate set of labels to the Casimir operators' eigenvalues. We show how to obtain these partitions from the tableau labeling x for later use in Sec. V. In addition, one can calculate the Casimir operator eigenvalues of x from the partitions μ^p with the formula of Perelomov and Popov[11]:

$$C_q = \sum_{i=1}^{p} L_i^q - \sum_{\alpha+\beta=q-1} \sum_{i<j} L_i^\alpha L_j^\beta + \cdots + (-1)^q \sum_{i_1<i_2<\cdots<i_q} 1,$$

where $q \leq p$, and $L_i = \mu^p_i + p - i$ with $\mu^p = \mu^p_1 \mu^p_2 \cdots$. C_q is the eigenvalue of the $U(p)$ Casimir operator of qth degree.

The development of the tableaux basis in Sec. II only specified the tableaux labeling of $U(N)$ states to the extent that it assigned a set of tableaux to each set of states having a given weight, with the number of objects in each set being the same. We now complete the labeling of states in irreducible representations by requiring the tableau labeling of a state x to be such that the removal of all boxes containing integers greater than p from the tableau results in a new tableau whose frame μ^p gives the transformation properties of x with respect to $U(p)$ for $p = 1, 2, \cdots, N-1$. As a result, one may read off the partitions μ^p of the subgroup irreducible representations from the tableau labeling x.

The proof that one may consistently label the orthogonal states of an irreducible representation of $U(N)$ with tableaux as described above follows from a simple construction. In the irreducible representation λ of $U(N)$, let $x_1, x_2, \cdots, x_b$ be the states with weight V. Each state x_i is labeled by $N-1$ partitions μ^p_i, with $p = 1, 2, \cdots, N-1$. Construct a tableau for each x_i by (1) placing μ^1_i ones in the first row of frame λ, (2) adding twos in such a way that the boxes containing ones and twos correspond to frame μ^2_i, and (3) adding threes in such a way that the numbered boxes correspond to frame $\mu^3_i, \cdots$. That there exists a way of placing the integer p in boxes such that one can go from partition μ^{p-1}_i to μ^p_i in the construction follows from (i) the theorem[12] giving the irreducible representations of $U(p-1)$ lying in the irreducible representation μ^p_i of $U(p)$ and (ii) the block diagonality of the $U(p-1)$ operators with respect to the $U(p)$ operators in $U(N)$.

Each tableau constructed above is a standard tableau in τ^N_λ which is distinct from the tableaux of the other states with weight V. Thus our method of obtaining the subgroup partitions labeling a state from the state's tableau is valid. In fact, it is a direct generalization of the procedure for finding the Yamanouchi symbol[13] of a tableau in a symmetric group representation.

V. TABLEAU STATES AND ARBITRARY UNITARY SUBGROUPS OF $U(N)$

In Sec. IV, we developed a procedure for finding how a tableau state of $U(N)$ transformed with respect to the subgroup $U(p)$ consisting of the generators A^i_k, for $i, k = 1, 2, \cdots, p$. It was noted that each tableau state transformed as a state in one irreducible representation of $U(p)$, for $p = 1, 2, \cdots, N-1$. However, if we examine the transformation properties of a $U(N)$ state with respect to an arbitrary subgroup

[11] A. Perelomov and V. Popov, JETP Letters **1**, 160 (1965).

[12] H. Boerner, *Representations of Groups* (North-Holland Publ. Co., Amsterdam, 1963), pp. 161, 164.

[13] M. Hamermesh, *Group Theory* (Addison-Wesley Publ. Co., Reading, Mass., 1962), p. 221.

of the form $U(N_1) \otimes U(N_2) \otimes \cdots \otimes U(N_r)$, with $\sum_{i=1}^{r} N_i \leq N$, we find that the $U(N)$ state transforms as a linear combination of states in different irreducible representations of the subgroup. Furthermore, an irreducible representation may occur more than once in the linear combination, just as an irreducible representation may occur more than once in the Clebsch–Gordan series of a tensor product. In this section we develop methods for finding the irreducible representations and the number of times they occur in the linear combination transforming like a $U(N)$ tableau state with respect to a subgroup.

The subalgebra $U(p)$ in the integers $i_1, i_2, \cdots, i_p$ is defined to contain only the generators A_k^j of $U(N)$ such that j and k are integers from the set $i_1, i_2, \cdots, i_p$. Given a weight V in an irreducible representation of $U(N)$, we define the $U(p)$ hyperplane of V to be the set of all weights which can be generated from V by the application of operators in the $U(p)$ algebra. In terms of weight diagrams, a $U(p)$ hyperplane is on a geometrical hyperplane in the weight diagram of the $U(N)$ representation, and, in general, it contains the superimposed weight diagrams of several $U(p)$ irreducible representations.

Given a weight $V = (V_1, V_2, \cdots, V_N)$ of the irreducible representation μ of $U(N)$, we wish to find how the states associated with its transform with respect to the subalgebra $\otimes_{j=1}^{k} U(N_j)$, with $\sum_j N_j = N$ and $U(N_j)$ a subalgebra in the integers $p_1^j, p_2^j, \cdots, p_{N_j}^j$. Let ρ^j be the partition whose parts are the integers $V_{p_1^j}, V_{p_2^j}, \cdots, V_{p_{N_j}^j}$; for $j = 1, 2, \cdots, k$. The partition ρ^j labels the Weyl class of the weight with respect to $U(N_j)$.

If we apply Eq. (21′) to Eq. (24) k times, we obtain

$$M^{\mu}_{[(\rho^1)(\rho^2)\cdots(\rho^k)]} = \sum_{\mu^1,\mu^2,\cdots,\mu^k} M^{\mu^1}_{\rho^1} M^{\mu^2}_{\rho^2} \cdots M^{\mu^k}_{\rho^k} \mathcal{N}^{\mu}_{\otimes_{i=1}^{k} \mu^i} \tag{37}$$

with the sums over k partitions, $\mu^1, \mu^2, \cdots, \mu^k$. On the right side of Eq. (37) one can read off the linear combination of states, giving the transformation properties of V's states with respect to the subalgebra. In particular, the irreducible representation $\mu^1 \otimes \mu^2 \otimes \cdots \otimes \mu^k$ occurs $\mathcal{N}^{\mu}_{\otimes_i \mu^i}$ times [Eq. (24)] in the linear combination. Furthermore, $M^{\mu^1}_{\rho^1} M^{\mu^2}_{\rho^2} \cdots M^{\mu^k}_{\rho^k}$ states of each representation $\mu^1 \otimes \mu^2 \otimes \cdots \otimes \mu^k$ appear in the linear combination. For example, we find the transformation properties of the states of $V = (2, 1, 0, 1, 1, 1)$ in the $U(6)$ representation, 42, with respect to the subalgebra $U(3) \otimes U(3)$ in the integers 1, 3, 4 and 2, 5, 6, respectively. All $\mathcal{N}^{42}_{\mu^1 \otimes \mu^2}$ are zero except $\mathcal{N}^{42}_{3 \otimes 3} = \mathcal{N}^{42}_{3 \otimes 21} = \mathcal{N}^{42}_{21 \otimes 3} = \mathcal{N}^{43}_{21 \otimes 21} = 1$, and therefore, each tableau state of V transforms as the linear combination:

$$a\left|\begin{array}{|c|c|c|}\hline 1&1&4\\\hline\end{array} \otimes \begin{array}{|c|c|c|}\hline 2&5&6\\\hline\end{array}\right\rangle + b\left|\begin{array}{|c|c|c|}\hline 1&1&4\\\hline\end{array} \otimes \begin{array}{|c|c|}\hline 2&5\\\hline 6\\\cline{1-1}\end{array}\right\rangle$$

$$+ c\left|\begin{array}{|c|c|c|}\hline 1&1&4\\\hline\end{array} \otimes \begin{array}{|c|c|}\hline 2&6\\\hline 5\\\cline{1-1}\end{array}\right\rangle + d\left|\begin{array}{|c|c|}\hline 1&1\\\hline 4\\\cline{1-1}\end{array} \otimes \begin{array}{|c|c|c|}\hline 2&5&6\\\hline\end{array}\right\rangle$$

$$+ e\left|\begin{array}{|c|c|}\hline 1&1\\\hline 4\\\cline{1-1}\end{array} \otimes \begin{array}{|c|c|}\hline 2&5\\\hline 6\\\cline{1-1}\end{array}\right\rangle + f\left|\begin{array}{|c|c|}\hline 1&1\\\hline 4\\\cline{1-1}\end{array} \otimes \begin{array}{|c|c|}\hline 2&6\\\hline 5\\\cline{1-1}\end{array}\right\rangle,$$

with a, b, c, d, e, and f being numerical coefficients. In terms of weight diagrams, we found the weight diagrams of $3 \otimes 3$, $3 \otimes 21$, $21 \otimes 3$, and $21 \otimes 21$ lying in the $U(3) \otimes U(3)$ hyperplane of V in the weight diagram of 42 in $U(6)$.

If we had chosen the $U(3) \otimes U(3)$ algebra to be in the integers 1, 2, 3 and 4, 5, 6, respectively, then, because of our labeling of tableau states in Sec. IV, we have the states

$$\left|\begin{array}{|c|c|c|c|}\hline 1&1&2&4\\\hline 5&6\\\cline{1-2}\end{array}\right\rangle, \quad \left|\begin{array}{|c|c|c|c|}\hline 1&1&2&5\\\hline 4&6\\\cline{1-2}\end{array}\right\rangle, \quad \left|\begin{array}{|c|c|c|c|}\hline 1&1&2&6\\\hline 4&5\\\cline{1-2}\end{array}\right\rangle$$

transforming as

$$a\left|\begin{array}{|c|c|c|}\hline 1&1&2\\\hline\end{array} \otimes \begin{array}{|c|c|c|}\hline 4&5&6\\\hline\end{array}\right\rangle + b\left|\begin{array}{|c|c|c|}\hline 1&1&2\\\hline\end{array} \otimes \begin{array}{|c|c|}\hline 4&5\\\hline 6\\\cline{1-1}\end{array}\right\rangle$$

$$+ c\left|\begin{array}{|c|c|c|}\hline 1&1&2\\\hline\end{array} \otimes \begin{array}{|c|c|}\hline 4&6\\\hline 5\\\cline{1-1}\end{array}\right\rangle,$$

while

$$\left|\begin{array}{|c|c|c|c|}\hline 1&1&4&5\\\hline 2&6\\\cline{1-2}\end{array}\right\rangle, \quad \left|\begin{array}{|c|c|c|c|}\hline 1&1&4&6\\\hline 2&5\\\cline{1-2}\end{array}\right\rangle, \quad \left|\begin{array}{|c|c|c|c|}\hline 1&1&5&6\\\hline 2&4\\\cline{1-2}\end{array}\right\rangle$$

transform as

$$a\left|\begin{array}{|c|c|}\hline 1&1\\\hline 2\\\cline{1-1}\end{array} \otimes \begin{array}{|c|c|c|}\hline 4&5&6\\\hline\end{array}\right\rangle + b\left|\begin{array}{|c|c|}\hline 1&1\\\hline 2\\\cline{1-1}\end{array} \otimes \begin{array}{|c|c|}\hline 4&5\\\hline 6\\\cline{1-1}\end{array}\right\rangle$$

$$+ c\left|\begin{array}{|c|c|}\hline 1&1\\\hline 2\\\cline{1-1}\end{array} \otimes \begin{array}{|c|c|}\hline 4&6\\\hline 5\\\cline{1-1}\end{array}\right\rangle,$$

with the coefficients a, b, c depending on the $U(6)$ tableau state under consideration.

Given a weight $\mathbf{V} = (V_1, V_2, \cdots, V_N)$ of the irreducible representation μ of $U(N)$, we now find the transformation properties of the states of weight

V with respect to the subalgebra $\otimes_j U(N_j)$, with $\sum_{j=1}^{r} N_j < N$ and $U(N_j)$ a subalgebra in the integers $p_1^j, p_2^j, \cdots, p_{N_j}^j$. Let ρ^j be the partition with parts $V_{p_1^j}, V_{p_2^j}, \cdots, V_{p_{N_j}^j}$; and if $a+1$ is the minimum integer in the set of integers p_i^j for $i = 1, 2, \cdots, N_j$ and $j = 1, 2, \cdots, r$, then let α be the partition with parts $V_1, V_2, \cdots, V_a$, and let β be the partition whose parts are all coordinates of V not previously taken as a part of a partition. Let $y^\alpha = x_1^{V_1} x_2^{V_2} \cdots x_a^{V_a}$ and

$$z^\beta = x_{a+1}^{V_{a+1}} x_{a+2}^{V_{a+2}} \cdots x_N^{V_N} \Big/ \prod_{j=1}^{r} \prod_{i=1}^{N_j} x_{p_i^j}^{V_{p_i^j}}.$$

To the subgroup $U(N_j)$ we assign the character $\chi_{\mu_j}(U_{N_j})$ in the variables $x_{p_1^j}, x_{p_2^j}, \cdots, x_{p_{N_j}^j}$ (instead of $x_1, x_2, \cdots, x_{N_j}$, respectively).

The number of states of $\mu_1 \otimes \mu_2 \otimes \cdots \otimes \mu_r = \otimes_j \mu_j$, appearing in the linear combination expressing the transformation properties of states of V with respect to the subgroup $\otimes_j U(N_j)$, is given by

$$M_{\rho^1}^{\mu_1} M_{\rho^2}^{\mu_2} \cdots M_{\rho^r}^{\mu_r} \tilde{N}_{\otimes_j \mu_j}^{\mu, V}, \tag{38}$$

where

$$\tilde{N}_{\otimes_j \mu_j}^{\mu, V} = \int dU_{N_1}\, dU_{N_2} \cdots dU_{N_r} \times \bar{\chi}_\mu \chi_{\mu_1}(U_{N_1}) \chi_{\mu_2}(U_{N_2}) \cdots \chi_{\mu_r}(U_{N_r}) y^\alpha z^\beta = \sum_{\nu_1, \nu_2, \cdots, \nu_r} B_{\mu_1}^{\nu_1} B_{\mu_2}^{\nu_2} \cdots B_{\mu_r}^{\nu_r} M^{\mu}_{[[\nu_1)(\nu_2)\cdots(\nu_r)(\alpha)(\beta)]}, \tag{39}$$

with the summation over r partitions $\nu_1, \nu_2, \cdots, \nu_r$. Equation (39) is derived in the same way as similar relations in Sec. III. Equation (39) also gives the number of times the weight diagram of $\otimes_j \mu_j$ appears in the $\otimes_j U(N_j)$ hyperplane of V. For the tableau state of V transforming like the representation λ of $U(a)$, we use

$$\tilde{N}_{\otimes_j \mu_j, \lambda}^{\mu, V} = \sum_{\delta, \nu_1 \nu_2, \cdots, \nu_r} B_\lambda^\delta B_{\mu_1}^{\nu_1} B_{\mu_2}^{\nu_2} \cdots B_{\mu_r}^{\nu_r} M^{\mu}_{[[(\delta)(\nu_1)(\nu_2)\cdots(\nu_r)(\beta)]}$$

instead of (39).

As an example of the above equations, we find the transformation property of a state of weight $V = (0, 0, 0, 1, 2, 1, 3, 0)$ in the representation 421 of $U(8)$ with respect to the subalgebra $U(3)$ in the integers 5, 7, and 8. Expression (38) becomes

$$M_{32}^{\mu_1} \tilde{N}_{\mu_1}^{421, V},$$

with (39) giving

$$\tilde{N}_{\mu_1}^{421, V} = \sum_{\nu} B_{\mu_1}^{\nu} M^{421}_{[[(\nu)(1)(1)]]}.$$

Since $M_{32}^{\mu_1} = 0$ for $32 > \mu_1$, we need only consider the following possibilities:

$$\tilde{N}_5^{421, V} = 0 \quad \text{and} \quad \tilde{N}_{41}^{421, V} = \tilde{N}_{32}^{421, V} = 2.$$

As a result, we have each of the four states with weight V transforming as

$$a \left| \begin{array}{|c|c|c|c|} \hline 5 & 5 & 7 & 7 \\ \hline 7 \\ \cline{1-1} \end{array} \; 1 \right\rangle + b \left| \begin{array}{|c|c|c|c|} \hline 5 & 5 & 7 & 7 \\ \hline 7 \\ \cline{1-1} \end{array} \; 2 \right\rangle + c \left| \begin{array}{|c|c|c|} \hline 5 & 5 & 7 \\ \hline 7 & 7 \\ \cline{1-2} \end{array} \; 1 \right\rangle + d \left| \begin{array}{|c|c|c|} \hline 5 & 5 & 7 \\ \hline 7 & 7 \\ \cline{1-2} \end{array} \; 2 \right\rangle,$$

with respect to the $U(3)$ subgroup where the additional integers in the kets serve to distinguish multiple-occurring irreducible representations.

For the analysis of the states of weight V in a tensor product representation of $U(N)$, $\otimes_i \alpha_i$, one need only substitute $M^{\otimes_i \alpha_i}_{[[(\nu_1)(\nu_2)\cdots(\nu_r)(\alpha)(\beta)]}$ in equation (39) for $M^{\mu}_{[[(\nu_1)(\nu_2)\cdots(\nu_r)(\alpha)(\beta)]}$.

ACKNOWLEDGMENTS

The author wishes to thank Dr. A. Pais for his advice and encouragement. I would also like to thank Professor L. C. Biedenharn for drawing my attention to recent work on weight multiplicity and Clebsch–Gordan series coefficient formulas.

APPENDIX A: THE MULTIPLICITIES M_ν^μ

ν \ μ	2	1^2
2	1	0
1^2	1	1

ν \ μ	3	21	1^3
3	1	0	0
21	1	1	0
1^3	1	2	1

ν \ μ	5	41	32	31^2	2^21	21^3	1^5
5	1	0	0	0	0	0	0
41	1	1	0	0	0	0	0
32	1	1	1	0	0	0	0
31^2	1	2	1	1	0	0	0
2^21	1	2	2	1	1	0	0
21^3	1	3	3	3	2	1	0
1^5	1	4	5	6	5	4	1

ν \ μ	6	51	42	41^2	3^2	321	31^3	2^3	2^21^2	21^4	1^6
6	1	0	0	0	0	0	0	0	0	0	0
51	1	1	0	0	0	0	0	0	0	0	0
42	1	1	1	0	0	0	0	0	0	0	0
41^2	1	2	1	1	0	0	0	0	0	0	0
3^2	1	1	1	0	1	0	0	0	0	0	0
321	1	2	2	1	1	1	0	0	0	0	0
31^3	1	3	3	3	1	2	1	0	0	0	0
2^3	1	2	3	1	1	2	0	1	0	0	0
2^21^2	1	3	4	3	2	4	1	1	1	0	0
21^4	1	4	6	6	3	8	4	2	3	1	0
1^6	1	5	9	10	5	16	10	5	9	5	1

ν \ μ	4	31	2^2	21^2	1^4
4	1	0	0	0	0
31	1	1	0	0	0
2^2	1	1	1	0	0
21^2	1	2	1	1	0
1^4	1	3	2	3	1

ν \ μ	7	61	52	51^2	43	421	41^3	3^21	32^2	321^2	31^4	2^31	2^21^3	21^5	1^7
7	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
61	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
52	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
51^2	1	2	1	1	0	0	0	0	0	0	0	0	0	0	0
43	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0
421	1	2	2	1	1	1	0	0	0	0	0	0	0	0	0
41^3	1	3	3	3	1	2	1	0	0	0	0	0	0	0	0
3^21	1	2	2	1	2	1	0	1	0	0	0	0	0	0	0
32^2	1	2	3	1	2	2	0	1	1	0	0	0	0	0	0
321^2	1	3	4	3	3	4	1	2	1	1	0	0	0	0	0
31^4	1	4	6	6	4	8	4	3	2	3	1	0	0	0	0
2^31	1	3	5	3	4	6	1	3	3	2	0	1	0	0	0
2^21^3	1	4	7	6	6	11	4	6	5	6	1	2	1	0	0
21^5	1	5	10	10	9	20	10	11	10	15	5	5	4	1	0
1^7	1	6	14	15	14	35	20	21	21	35	15	14	14	6	1

APPENDIX B: THE COEFFICIENTS B^{μ}_{ν}

ν \ μ	2	1^2
2	1	0
1^2	−1	1

ν \ μ	3	21	1^3
3	1	0	0
21	−1	1	0
1^3	1	−2	1

ν \ μ	4	31	2^2	21^2	1^4
4	1	0	0	0	0
31	−1	1	0	0	0
2^2	0	−1	1	0	0
21^2	1	−1	−1	1	0
1^4	−1	2	1	−3	1

ν \ μ	5	41	32	31^2	2^21	21^3	1^5
5	1	0	0	0	0	0	0
41	−1	1	0	0	0	0	0
32	0	−1	1	0	0	0	0
31^2	1	−1	−1	1	0	0	0
2^21	0	1	−1	−1	1	0	0
21^3	−1	1	2	−1	−2	1	0
1^5	1	−2	−2	3	3	−4	1

ν \ μ	6	51	42	41^2	3^2	321	31^3	2^3	2^21^2	21^4	1^6
6	1	0	0	0	0	0	0	0	0	0	0
51	−1	1	0	0	0	0	0	0	0	0	0
42	0	−1	1	0	0	0	0	0	0	0	0
41^2	1	−1	−1	1	0	0	0	0	0	0	0
3^2	0	0	−1	0	1	0	0	0	0	0	0
321	0	1	0	−1	−1	1	0	0	0	0	0
31^3	−1	1	1	−1	1	−2	1	0	0	0	0
2^3	0	0	−1	1	1	−2	0	1	0	0	0
2^21^2	0	−1	1	1	0	0	−1	−1	1	0	0
21^4	1	−1	−2	1	−1	4	−1	1	−3	1	0
1^6	−1	2	2	−3	1	−6	4	−1	6	−5	1

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 12 DECEMBER 1969

Functional Integration and the Generalized Matthews-Salam Equations

ROBERT L. ZIMMERMAN
Institute of Theoretical Science, University of Oregon, Eugene, Oregon

(Received 29 July 1968)

Various properties of Feynman functional integrals that appear in quantum field theory are studied. An indefinite functional integral is constructed. For the indefinite functional integral we prove a relation which is analogous in ordinary Riemann integrals to integration by parts. A special case of this relation gives an integration-by-parts formula for the Feynman functional integrals. In addition, various relations for integrating over variationals and variational derivatives are obtained. Application of these relations gives, among other things, a set of generalized Matthews–Salam equations.

I. INTRODUCTION

Various formulations of quantum field theories have been developed in order to go beyond the framework of perturbation theory. Among these attempts is the one based on the use of functional integrals. The advantage of this approach is that it is possible to obtain closed expressions for the complete Green's functions including all the radiation corrections. The lack of formal properties and approximate methods for solving functional integrals has hampered the development of this formalism. It is the purpose of this paper to investigate various properties for functional integrals over variationals and variational derivatives of functionals.

The concept of a functional integral or an integral over an infinite-dimensional space was first considered by Daniell.[1] Wiener later developed the Wiener integral[2] which is used in studying Browning motion. Feynman, in his space–time approach to nonrelativistic quantum mechanics,[3] developed another functional integral, sometimes referred to as the Feynman functional integral. The Feynman functional integral is not an integral in the true mathematical sense; however, its close relationship to the Wiener integral has been shown by Cameron.[4] We will not concern ourselves in this paper with the outstanding question on the mathematical meaning of the Feynman integral. Our main concern will be directed towards its properties and various means of evaluating it. Because of the nebulous definition of this integral, the results of this paper can at most be considered formal and mere reflection of the properties of Wiener integrals onto the Feynman integrals.

The extension of Feynman's formulation of nonrelativistic quantum mechanics to quantum field theory was first accomplished by Davison,[5] Matthews and Salam,[6] and Polkinghorne.[7] In extending the functional formalism to quantum field theory, there appear two distinct classes of integrals: those over fields obeying Bose–Einstein statistics and those over fields obeying Fermi–Dirac statistics. In this paper we will consider both kinds of integrals; however, we will limit ourselves to the more common fields, i.e., scalar, photon, and spin-$\frac{1}{2}$ fields.

The scalar functional integral over the field $\phi(x)$, whose integrand is $F(\phi)$, is denoted by

$$\langle \phi'', \sigma'' |\, T(F(\phi))\, | \phi', \sigma' \rangle = \frac{1}{N} \int_{[\phi'',\phi']} F(\phi) e^{iI_{\sigma'',\sigma'}(\phi)} \delta\phi, \tag{1}$$

where

$$N = \int_{[\phi'',\phi']} e^{iI_{\sigma'',\sigma'}(\phi)} \delta\phi,$$

$$I_{\sigma''\sigma'}(\phi) = -\tfrac{1}{2} \int_{\sigma'}^{\sigma''} \phi(x)(\Box + m^2)\phi(x)\, d^4x.$$

ϕ', ϕ'' are the eigenvalues of a complete commuting set of operators, which specify the state of the system on the two surfaces σ' and σ''. The support of the functional integral is over all continuous fields defined between the two spacelike surfaces σ' and σ'', which are restricted to the values ϕ' and ϕ'', respectively, on these surfaces. The $1/N$ factor is a normalization factor. We have chosen our matrix $g_{\mu\nu}$ such that $A_\mu B^\mu = A_0 A^0 - \mathbf{A} \cdot \mathbf{B}$. Here $\Box$ is the d'Alembertian operator

$$\Box = \frac{\partial}{\partial x_\mu} \frac{\partial}{\partial x^\mu}.$$

Let us now define what we mean by the indefinite Feynman integral over scalar fields. We denote the

[1] P. J. Daniell, Ann. Math. **19**, 279 (1918); **20**, 281 (1918); **21**, 30 (1919).
[2] N. Wiener, Ann. Math. **22**, 66 (1920).
[3] R. P. Feynman, Rev. Mod. Phys. **20**, 367 (1948).
[4] R. H. Cameron, J. Math. & Phys. **39**, 126 (1960).
[5] B. Davison, Proc. Roy. Soc. (London) **A225**, 252 (1954).
[6] P. T. Matthews and A. Salam, Nuovo Cimento **2**, 120 (1955).
[7] J. C. Polkinghorne, Proc. Roy. Soc. (London) **A230**, 272 (1955).

indefinite Feynman integral by

$$\langle \phi'', \sigma''|\, T(F(\phi))\,|\phi', \sigma'\rangle_{\mu(x)} = \frac{1}{N}\int_{[\phi'',\phi']}^{\phi(x)<\mu(x)} F(\phi)e^{iI_{\sigma''\sigma'}(\phi)}\delta\phi. \quad (2)$$

The indefinite Feynman integral is meant to indicate that the integral is taken over all continuous fields $\phi(x)$ such that $\phi(x) < \mu(x)$ and $\phi = \phi' < \mu$ on σ' and $\phi = \phi'' < u$ on σ''.

Likewise, the functional integral over the electromagnetic field $A(x)$ with the integrand $F(A)$ is denoted by

$$\langle A'', \sigma''|\, T(F(A))\,|A', \sigma'\rangle = \frac{1}{N}\int_{[A'',A']} F(A)e^{iI_{\sigma'',\sigma'}(A)}\delta A_0\delta A_1\delta A_2\delta A_3, \quad (3)$$

where

$$I_{\sigma'',\sigma'}(A) = -\tfrac{1}{4}\int_{\sigma'}^{\sigma''} F_\mu F^\mu\, d^4x,$$

$$F_\mu = \frac{\partial A_\mu}{\partial x^\nu} - \frac{\partial A_\nu}{\partial x^\mu},$$

and

$$N = \int_{[A'',A']} e^{iI_{\sigma'',\sigma'}(A)}\delta A_0\delta A_1\delta A_2\delta A_3.$$

The indefinite integral over $A(x)$ is

$$\langle A'', \sigma''|\, T(F(A))\,|A', \sigma'\rangle_{\mu(x)} = \frac{1}{N}\int_{[A'',A']}^{A(x)<\mu(x)} F(A)e^{iI_{\sigma'',\sigma'}(A)}\delta A_0\delta A_1\delta A_2\delta A_3. \quad (4)$$

The functional integral over Fermi fields $\psi(x)$ and $\bar\psi(x)$ is the same as for Boson integrals, except that the Fermi fields anticommute. The Feynman functional integral for Fermi fields is denoted by

$$\langle \psi'', \bar\psi'', \sigma''|\, T(F(\psi\bar\psi))\,|\psi', \bar\psi', \sigma'\rangle = \frac{1}{N}\int_{[\psi'',\psi'];[\bar\psi'',\bar\psi']} F(\psi\bar\psi)e^{iI_{\sigma'',\sigma'}(\psi,\bar\psi)}\delta\psi\delta\bar\psi. \quad (5)$$

Likewise, the indefinite integral is

$$\langle \psi'', \bar\psi'', \sigma''|\, T(F(\psi, \bar\psi))\,|\psi', \bar\psi', \sigma'\rangle_{\mu(x);\mu(x)} = \frac{1}{N}\int_{[\psi'',\psi'];[\bar\psi'',\bar\psi']}^{\psi(x)<\mu(x);\bar\psi(x)<\bar\mu(x)} F(\psi, \bar\psi)e^{iI_{\sigma''\sigma'}(\psi\bar\psi)}\delta\psi\delta\bar\psi, \quad (6)$$

where

$$N = \int_{[\psi'',\psi'];[\bar\psi'',\bar\psi']} e^{iI_{\sigma''\sigma'}(\psi\bar\psi)}\delta\psi\delta\bar\psi$$

and

$$I(\psi\bar\psi) = \int_{\sigma'}^{\sigma''} \bar\psi(x)(i\not{\nabla} - m)\psi(x)\, d^4x.$$

The notation used in the Dirac equation is

$$i\not{\nabla} = \not{p} = i\gamma_0\frac{\partial}{\partial t} + i\boldsymbol{\gamma}\cdot\frac{\partial}{\partial \mathbf{x}}$$

and the γ matrices satisfy the anticommutation relations

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}.$$

We would now like to prove some formal properties of these integrals. In particular, we would like to consider variationals of the indefinite Feynman integral and integrals over variational derivatives of the integrand. As a specific example of our results, we will get a relation which corresponds to integration by parts for ordinary Riemann integrals. From this and similar relations will follow the variational formulas of Schwinger and the set of differential equations for the tau functions as developed by Matthews and Salam.[8] The derivation of the Matthews–Salam equations from the functional integral formalism was observed by Polkinghorne[7]; however, he had to assume that it was valid to integrate by parts. The analogous properties that occur for the Wiener integral have been derived by Cameron[9] and Owchar.[10]

The outline of the paper is as follows. In Sec. II we will consider the Boson integrals and construct various properties for these functional integrals over variations and variational derivatives. Similar properties will be given in Sec. III for the Fermi field. In Sec. IV we will give explicit examples in order to illuminate how these results may be used to solve complicated functional integrals. The Matthews–Salam equations will be shown in Sec. V to be a special case of the relations derived in Secs. II and III. In the final section we will also construct a generalized set of Matthews–Salam equations.

II. BOSON INTEGRALS

In this section we will limit our investigation to Boson fields. In particular, the first part of this section will be limited to scalar fields. At the end of the section we will extend the significant properties of scalar fields to zero-mass vector fields.

Let us denote the first variational of the functional $F(\phi)$ by $\delta_f F(\phi)$. It is defined by the relation

$$\delta_f F(\phi) = \frac{d}{dh}\, F(\phi(x) + hf(x))\big|_{h=0}. \quad (7)$$

[8] P. T. Matthews and A. Salam, Proc. Roy. Soc. (London) **A221**, 128 (1954).
[9] R. H. Cameron, Proc. Am. Math. Soc. **2**, 914 (1951).
[10] M. Owchar, Proc. Am. Math. Soc. **3**, 459 (1952).

The nth variational is defined in a similar manner as

$$\delta^n_{f_n,\cdots,f_1}F(\phi) = \frac{d}{dh}\,\delta^{n-1}_{f_{n-1},\cdots,f_1}F(\phi(x) + hf_n(x))\big|_{h=0}. \tag{8}$$

The first functional (Volterra) derivative of $F(\phi)$ at the point y is denoted by $\delta F(\phi(x))/\delta\phi(y)$ and is implicitly defined by the relation

$$\delta_f F(\phi(x)) = \int_{\sigma'}^{\sigma''} \frac{\delta F}{\delta\phi(y)}(\phi(x))f(y)\,d^4y. \tag{9}$$

For the sake of convenience and simplicity, it will be assumed throughout the remainder of the paper that $f(y)$ vanishes on the surfaces σ'' and σ'.

In a similar fashion the nth functional derivative is

$$\frac{\delta^n}{\delta\phi(x_n)\cdots\delta\phi(x_1)}F(\phi(x)) = \frac{\delta}{\delta\phi(x_n)}\left\{\frac{\delta^{n-1}}{\delta\phi(x_{n-1}),\cdots,\delta\phi(x_1)}F(\phi(x))\right\}. \tag{10}$$

Let us now consider the variation with respect to the field $\mu(x)$ of the indefinite Feynman integral defined in Eq. (2):

$$\delta_f\langle\phi'',\sigma''|\,T(F(\phi))\,|\phi',\sigma'\rangle_{\mu(x)} = \frac{d}{dh}\langle\phi'',\sigma''|\,T(F(\phi))\,|\phi',\sigma'\rangle_{\mu(x)+hf(x)}\big|_{h=0}. \tag{11}$$

We will proceed to show that this is equivalent to

$$\delta_f\langle\phi'',\sigma''|\,T(F(\phi))\,|\phi',\sigma'\rangle_{\mu(x)} = \langle\phi'',\sigma''|\,T(\delta_f F(\phi))\,|\phi',\sigma'\rangle_{\mu(x)} + i\langle\phi'',\sigma''|\,T(F(\phi)\delta_f I_{\sigma''\sigma'}(\phi))\,|\phi',\sigma'\rangle_{\mu(x)}, \tag{12}$$

where the variations on the right-hand side of Eq. (12) are taken with respect to the fields $\phi(x)$, while the one on the left is taken with respect to $\mu(x)$. Equation (12) follows from the fact that

$$\langle\phi'',\sigma''|\,T(F(\phi))\,|\phi',\sigma'\rangle_{\mu+hf} = \frac{1}{N}\int_{[\phi'',\phi']}^{\phi_1<\mu+hf} F(\phi_1)e^{iI(\phi_1)}\delta\phi_1. \tag{13}$$

Making a linear transformation of the variables in Eq. (13),

$$\phi(x) = \phi_1(x) - hf(x), \tag{14}$$

we obtain

$$\langle\phi'',\sigma''|\,T(F(\phi))\,|\phi',\sigma'\rangle_{\mu(x)+hf(x)} = \int_{[\phi'',\phi']}^{\phi(x)<\mu(x)} F(\phi + hf)e^{iI(\phi+hf)}\delta\phi. \tag{15}$$

Using Eqs. (11) and (15), we get

$$\begin{aligned}\delta_f\langle\phi'',\sigma''|\,T(F(\phi))\,|\phi',\sigma'\rangle_{\mu(x)} &= (d/dh)\langle\phi'',\sigma''|\,T(F(\phi))\,|\phi',\sigma'\rangle_{\mu(x)+hf(x)}\big|_{h=0}\\ &= \frac{1}{N}\int_{[\phi'',\phi']}^{\phi(x)<\mu(x)}\frac{d}{dh}(F(\phi + hf)e^{iI(\phi+hf)})\delta\phi\big|_{h=0}.\end{aligned} \tag{16}$$

In the ordinary manner we differentiate the integrand with respect to h; Eq. (16) becomes

$$\begin{aligned}\delta_f\langle\phi'',\sigma''|\,T(F(\phi))\,|\phi',\sigma'\rangle_{\mu(x)} &= \frac{1}{N}\int_{[\phi'',\phi']}^{\phi<\mu(x)}\delta_f F(\phi)e^{iI(\phi)}\delta\phi\\ &\quad+\frac{1}{N}\int_{[\phi'',\phi']}^{\phi(x)<\mu(x)}F(\phi)\delta_f I(\phi)e^{iI(\phi)}\delta\phi.\end{aligned} \tag{17}$$

This is precisely Eq. (12), which we wanted to justify.

A special case of Eq. (12) is when we let $\mu(x) = +\infty$. In this case we integrate over the whole space and the indefinite integral becomes the ordinary Feynman functional integral. Notice that

$$\delta_f\langle\phi'',\sigma''|\,T(F(\phi))\,|\phi',\sigma'\rangle_{\infty} = \delta_f\langle\phi'',\sigma''|\,T(F(\phi))\,|\phi',\sigma'\rangle = 0 \tag{18}$$

since the integral is no longer a function of the variable $\mu(x)$, and Eq. (12) reduces to

$$0 = \langle\phi'',\sigma''|\,T(\delta_f F(\phi))\,|\phi',\sigma'\rangle + i\langle\phi'',\sigma''|\,T(F(\phi)\delta_f F_{\sigma''\sigma'}(\phi))\,|\phi',\sigma'\rangle. \tag{19}$$

Using the fact that

$$\delta_f I_{\sigma''\sigma'}(\phi) = -\int_{\sigma'}^{\sigma''} f(x)(\Box + m^2)\phi(x)\,d^4x, \tag{20}$$

Eq. (19) becomes

$$\langle\phi'',\sigma''|\,T(\delta_f F(\phi))\,|\phi',\sigma'\rangle = i\int d^4x f(x)(\Box + m^2)\langle\phi'',\sigma''|\,T(F(\phi)\phi(x))\,|\phi',\sigma'\rangle. \tag{21}$$

Let us now generalize the results to the nth variational. The nth variational of $F(\phi)$ satisfies

$$\langle\phi'',\sigma''|\,T(\delta^n_{f_{n_1}}f_1F(\phi))\,|\phi',\sigma'\rangle = -i\langle\phi'',\sigma''| \times TF(\phi)\prod_{j=2}^{n}(-i[\delta_{f_j}I_{\sigma'',\sigma}(\phi)] - \delta_{f_j})\delta_{f_j}I(\phi)\,|\phi',\sigma'\rangle. \tag{22}$$

The proof of Eq. (22) follows directly from induction, as will be shown below.

For $n = 1$ Eq. (22) reduces to Eq. (21), which was already shown to be valid. Let us assume Eq. (21) holds for $n = N$ and we will show it is true for $n = N + 1$.

Applying Eq. (22) for $n = N$ to the functional $\delta_{f_1}F(\phi)$, we have

$$\begin{aligned}&\langle\phi'', \sigma''|\, T(\delta^{N+1}_{f_{N+1}\cdots f_1}F(\phi))\,|\phi', \sigma'\rangle\\ &\quad= \langle\phi'', \sigma''|\, T\delta^{N}_{f_N\cdots f_2}(\delta_{f_1}F(\phi))\,|\phi', \sigma'\rangle\\ &\quad= -i\,\langle\phi'', \sigma''|\, T\delta_{f_1}f(\phi)\\ &\qquad\times\prod_{i=3}^{N+1}(-i\delta_{f_1}I_{\sigma''\sigma'}(\phi) - \delta_{f_1})\delta_{f_2}I(\phi)\,|\phi', \sigma'\rangle. \quad (23)\end{aligned}$$

Using Eq. (19) in the right-hand side of Eq. (23), we get

$$\begin{aligned}&\langle\phi'', \sigma''|\, T(\delta^{N+1}_{f_{N+1},\cdots,f_1}F(\phi))\,|\phi', \sigma'\rangle\\ &= -i\,\langle\phi'', \sigma''|\, TF(\phi)\prod_{i=3}^{N+1}(-i[\delta_{f_i}F(\phi)] - \delta_{f_i})\\ &\quad\times\delta_{f_2}I_{\sigma'',\sigma'}(\phi)i\delta_{f_1}I_{\sigma''\sigma'}(\phi)\,|\phi', \sigma'\rangle\\ &\quad- i\,\langle\phi', \sigma'|\, TF(\phi)\delta_{f_2}\prod_{i=3}^{N+1}(-i[\delta_{f_i}F(\phi)] - \delta_{f_i})\,|\phi', \sigma'\rangle\\ &= -i\,\langle\phi'', \sigma''|\, TF(\phi)\prod_{i=2}^{N+1}(-i[\delta_{f_i}F(\phi)] - \delta_{f_i})\\ &\quad\times\delta_{f_1}I_{\sigma'',\sigma'}(\phi)\,|\phi', \sigma'\rangle, \quad (24)\end{aligned}$$

and this is just Eq. (22). This completes the proof by induction of Eq. (22).

Let us now consider integration over functional derivatives. Substituting Eq. (9) into Eq. (21), we get

$$\begin{aligned}&\int d^4x f(x)\,\langle\phi'', \sigma''|\, T\left(\frac{\delta F(\phi)}{\delta\phi(x)}\right)|\phi', \sigma'\rangle\\ &= i\int d^4x f(x)(\Box + m^2)_x\,\langle\phi'', \sigma''|\, T(F(\phi)\phi(x))\,|\phi', \sigma'\rangle.\end{aligned} \quad (25)$$

Consequently it follows that

$$\begin{aligned}&\langle\phi'', \sigma''|\, T\left(\frac{\delta F(\phi)}{\delta\phi(x)}\right)|\phi', \sigma'\rangle\\ &\quad+ i(\Box + m^2)_x\,\langle\phi'', \sigma''|\, T(F(\phi)\phi(x))\,|\phi', \sigma'\rangle. \quad (26)\end{aligned}$$

The integration over the nth functional derivative of $F(\phi)$ satisfies the relation

$$\begin{aligned}&\langle\phi'', \sigma''|\, T\left(\frac{\delta^n F(\phi)}{\delta\phi(x_n), \cdots, \delta\phi(x_1)}\right)|\phi', \sigma'\rangle\\ &\quad= (i)^n\prod_{j=1}^{n}(\Box + m^2)_{x_j}\,\langle\phi'', \sigma''|\, T(F(\phi)\phi(x_j))\,|\phi', \sigma'\rangle.\end{aligned} \quad (27)$$

The proof of Eq. (27) follows in a trivial manner from induction and will not be shown.

A special case follows from Eq. (21) if we let

$$F(\phi) = G(\phi)R(\phi). \quad (28)$$

Substituting the first variational of Eq. (28),

$$\phi_f F(\phi) = \phi_f G(\phi)R(\phi) + G(\phi)\delta_f R(\phi), \quad (29)$$

into Eq. (21), we obtain

$$\begin{aligned}&\langle\phi'', \sigma''|\, T(G(\phi)\delta_f R(\phi))\,|\phi', \sigma'\rangle\\ &\qquad- \langle\phi'', \sigma''|\, T(\delta_f G(\phi)R(\phi))\,|\phi', \sigma'\rangle\\ &\qquad+ i\int d^4x f(x)(\Box + m^2)_x\\ &\qquad\times\langle\phi'', \sigma''|\, T(G(\phi)R(\phi)\phi(x))\,|\phi', \sigma'\rangle.\end{aligned} \quad (30)$$

In a similar manner we get an analogous relation from Eq. (26) for functional derivatives:

$$\begin{aligned}&\langle\phi'', \sigma''|\, T\left(G(\phi)\frac{\delta R}{\delta\phi(x)}\right)|\phi', \sigma'\rangle\\ &\quad= -\langle\phi'', \sigma''|\, T\left(\frac{\delta G(\phi)}{\delta\phi(x)}R(\phi)\right)|\phi', \sigma'\rangle\\ &\qquad+ i(\Box + m^2)_x\,\langle\phi'', \sigma''|\, T(G(\phi)R(\phi)\phi(x))\,|\phi', \sigma'\rangle.\end{aligned} \quad (31)$$

Equations (30) and (31) are similar to integration by parts in ordinary integration theory.

Let us generalize Eq. (31) to the analogous case of integrating by parts N times. Using Eq. (31) over and over again, we get

$$\begin{aligned}&\langle\phi'', \sigma''|\, T\left(G(\phi)\frac{\delta^N R(\phi)}{\delta\phi(x_N), \cdots, \delta\phi(x_1)}\right)|\phi', \sigma'\rangle\\ &\quad= i(\Box + m^2)_{x_N}\,\langle\phi'', \sigma''|\, T\left(G(\phi)\frac{\delta^{N-1}R(\phi)}{\delta\phi(x_{N-1}), \cdots, \delta\phi(x_1)}\phi(x_N)\right)|\phi', \sigma'\rangle\\ &\qquad- i(\Box + m^2)_{x_{N-1}}\,\langle\phi'', \sigma''|\, T\left(\frac{\delta G(\phi)}{\delta\phi(x_N)}\frac{\delta^{n-2}R(\phi)}{\delta\phi(x_{N-2}), \cdots, \delta\phi(x_1)}\phi(x_{N-1})\right)|\phi', \sigma'\rangle\\ &\qquad+ (-1)^{N+1}i(\Box + m^2)_{x_1}\,\langle\phi'', \sigma''|\, T\left(\frac{\delta^{N-1}G(\phi)}{\delta\phi(x_2), \cdots, \delta\phi(x_N)}R(\phi)\phi(x_2)\right)|\phi', \sigma'\rangle\\ &\qquad+ (-1)^N\,\langle\phi'', \sigma''|\, T\left(\frac{\delta^N G(\phi)}{\delta\phi(x_1), \cdots, \delta\phi(x_N)}R(\phi)\right)|\phi', \sigma'\rangle.\end{aligned} \quad (32)$$

The Matthews–Salam equations are special examples of integration by parts in Eq. (31). Their generalization will follow from Eq. (32).

The pertinent relations derived for the scalar integral are Eqs. (21), (22), (26), (27), (30), and (32). These equations generalize in a trivial manner to arbitrary Boson fields. In particular, for the electromagnetic field $A_\mu(x)$ we have

$$\langle A'', \sigma''|\, T(\delta_{f_\mu} F(A_\mu))\,|A', \sigma'\rangle = i\int d^4x f_\mu(x)\Box_x \langle A'', \sigma''|\, T(f(A)A^\mu(x))\,|A', \sigma'\rangle, \tag{33}$$

$$\langle A'', \sigma''|\, T(\delta^n_{f_{\mu_n}\cdots f_{\mu_1}} F(A))\,|A', \sigma'\rangle = -\,i\,\langle A'', \sigma''|\, T\left(F(A)\prod_{j=2}^{n}(-i\delta_{f_{\mu_j}} I(A) - \delta_{f_{\mu_j}})\delta_{f_{\mu_1}} I(A)\right)|A', \sigma'\rangle, \tag{34}$$

$$\langle A'', \sigma''|\, T\left(\frac{\delta F(A)}{\delta A_\mu(x)}\right)\,|A', \sigma'\rangle = i\Box_x \langle A'', \sigma''|\, T(F(A)A^\mu(x))\,|A', \sigma'\rangle, \tag{35}$$

$$\langle A'', \sigma''|\, T\left(\frac{\delta F(A)}{\delta A_{\mu_1}(x)\cdots\delta A_{\mu_n}(x_n)}\right)\,|A', \sigma'\rangle = (i)^n\prod_{j=1}^{n}\Box_{x_j}\langle A'', \sigma''|\, T(F(A)A^\mu(x_j))\,|A', \sigma'\rangle, \tag{36}$$

$$\begin{aligned}&\langle A'', \sigma''|\, T(G(A)\delta_{f_\mu} R(A))\,|A', \sigma'\rangle\\ &\quad = -\langle A'', \sigma''|\, T(\delta_{f_\mu} G(A)R(A))\,|A', \sigma'\rangle + i\int d^4x f_\mu(x)\Box_x\langle A'', \sigma''|\, T(G(A)R(A)A^\mu(x))\,|A', \sigma'\rangle,\end{aligned} \tag{37}$$

$$\begin{aligned}&\langle A'', \sigma''|\, T\left(G(A)\frac{\delta^n R(A)}{\delta A_{\mu_n}(x_n)\cdots\delta A_{\mu_1}(x_1)}\right)\,|A', \sigma'\rangle\\ &\qquad = i\Box_{x_n}\langle A'', \sigma''|\, T\left(G(A)\frac{\delta^{n-1}R(A)}{\delta A_{n-1}(x_{n-1})\cdots\delta A(x_1)}\right)A_{\mu_n}(x_n)\,|A', \sigma'\rangle\\ &\qquad\quad\times\cdots\times i(-1)^{n+1}\Box_{x_1}\langle A'', \sigma''|\, T\left(\frac{\delta^{n-1}G(A)}{\delta A_{\mu_2}(x_2)\cdots\delta A_{\mu_n}(x_n)}R(A)A_{\mu_1}(x_1)\right)\,|A', \sigma'\rangle\\ &\qquad\quad\times(-1)^n\langle A'', \sigma''|\, T\left(\frac{\delta^n G(A)}{\delta A_{\mu_1}(x_1)\cdots\delta A_{\mu_n}(x_n)}R(A)\right)\,|A', \sigma'\rangle.\end{aligned} \tag{38}$$

This completes our discussion of Boson integrals. Illustration of these relations will be found in Secs. IV and V.

III. FERMI FUNCTIONAL INTEGRALS

The significant formulas of Sec. II will be extended to Fermi fields. This problem is slightly more complex because of the noncommutivity of the functional integrals. Care must be exercised with the order of the fields; otherwise the extension is straightforward.

The first variation of $F(\psi, \bar\psi)$ with respect to ψ is defined as

$$\delta_\eta F(\psi, \bar\psi) = \frac{d}{dh}F(\psi + h\eta, \bar\psi)\big|_{h=0}. \tag{39}$$

The auxiliary spinor η anticommute among themselves and with all ψ and $\bar\psi$.

The nth variation of $F(\psi, \bar\psi)$ with respect to ψ is just

$$\delta^n_{\eta_n\cdots\eta_1}F(\psi, \bar\psi) = \frac{d}{dh}\delta^{n-1}_{\eta_{n-1}\cdots\eta}F(\psi + h\eta_n, \bar\psi)\big|_{h=0}. \tag{40}$$

The variational of $F(\psi, \bar\psi)$ with respect to $\bar\psi$ follows in an analogous manner:

$$\delta_{\bar\eta}F(\psi, \bar\psi) = \frac{d}{dh}F(\psi, \bar\psi + h\bar\eta)\big|_{h=0} \tag{41}$$

and

$$\delta^N_{\eta_{n_1}\cdots\eta_1}F(\psi, \bar\psi) = \frac{d}{dh}\delta^{N-1}_{\eta_{n-1}\cdots\bar\eta_1}F(\psi, \bar\psi + h\bar\eta_N)\big|_{h=0}, \tag{42}$$

where $\bar\eta$ anticommutes with all ψ, $\bar\psi$, and $\bar\eta$.

The first functional derivative of $F(\psi, \bar\psi)$ with respect to ψ is defined by the implicit relation

$$\delta_\eta F(\psi(y), \bar\psi) = \int\frac{\delta F}{\delta\psi(x)}(\psi(y), \bar\psi)\eta(x)\,d^4x. \tag{43}$$

The nth functional derivative is

$$\frac{\delta^n F(\psi, \bar\psi)}{\delta\psi(x_1)\cdots\delta\psi(x_n)} = \frac{\delta}{\delta\psi(x_1)}\left\{\frac{\delta^{n-1}F(\psi, \bar\psi)}{\delta\psi(x_2)\cdots\delta\psi(x_n)}\right\}. \tag{44}$$

The first functional derivative of $F(\psi, \bar\psi)$ with respect to $\bar\psi$ is

$$\delta_{\bar\eta}F(\psi, \bar\psi) = \int\bar\eta(x)\frac{\delta F(\psi, \bar\psi)}{\delta\bar\psi(x)}\,d^4x, \tag{45}$$

and its nth functional derivative is generalized in the usual manner.

The Feynman functional integral over the first variational of $F(\psi, \bar{\psi})$ with respect to ψ is

$$\langle \psi'', \bar{\psi}'', \sigma'' | \, T(\delta_\eta F(\psi, \bar{\psi})) \, | \psi', \bar{\psi}', \sigma' \rangle = -i \langle \psi'', \bar{\psi}'', \sigma'' | \, T(F(\psi, \bar{\psi}) \delta_\eta I_{\sigma'', \sigma'}(\psi)) \, | \psi', \bar{\psi}', \sigma' \rangle. \tag{46}$$

The generalization to the nth variation is

$$\langle \psi'', \bar{\psi}'', \sigma'' | \, T(\delta^N_{\eta_N \cdots \eta_1} F(\psi, \bar{\psi})) \, | \psi', \bar{\psi}', \sigma' \rangle = -i \langle \psi'', \bar{\psi}'', \sigma'' | \, TF(\psi, \bar{\psi}) \times \prod_{j=2}^{N} (-i \delta_{\eta_j} I_{\sigma'', \sigma'}(\psi) - \delta_{\eta_j}) \delta_{\eta_1} I_{\sigma'', \sigma'}(\psi) \, | \psi', \bar{\psi}, \sigma' \rangle. \tag{47}$$

The verification of Eqs. (46) and (47) is equivalent to the proof of Eqs. (19) and (22) and will not be repeated.

Likewise, for variations with respect to $\bar{\psi}$ we get

$$\langle \psi'', \bar{\psi}'', \sigma'' | \, T(\delta_{\bar{\eta}} F(\psi, \bar{\psi})) \, | \psi', \bar{\psi}', \sigma' \rangle = -i \langle \psi'', \bar{\psi}'', \sigma'' | \, T(F(\psi, \bar{\psi}) \delta_{\bar{\eta}} I_{\sigma'', \sigma'}(\bar{\psi})) \, | \psi', \bar{\psi}', \sigma' \rangle \tag{48}$$

and

$$\langle \psi'', \bar{\psi}'', \sigma'' | \, T(\delta^N_{\bar{\eta}_N \cdots \bar{\eta}_1} F(\psi, \bar{\psi})) \, | \psi', \bar{\psi}', \sigma' \rangle = -i \langle \psi'', \bar{\psi}'', \sigma'' | \, TF(\psi, \bar{\psi}) \times \prod_{j=2}^{N} (-i \delta_{\bar{\eta}_j} I_{\sigma'', \sigma'}(\bar{\psi}) - \delta_{\bar{\eta}_j}) \delta_{\bar{\eta}_1} I_{\sigma'' \sigma'}(\bar{\psi}) \, | \psi', \bar{\psi}', \sigma' \rangle. \tag{49}$$

Let us now consider integration over functional derivatives. Using the fact that

$$\delta_\eta I_{\sigma'', \sigma'}(\psi) = \int_{\sigma'}^{\sigma''} d^4x \bar{\psi}(x)(-i\overleftarrow{\not\partial} - m)\eta(x) \tag{50a}$$

and

$$\delta_{\bar{\eta}} I_{\sigma'', \sigma'}(\bar{\psi}) = + \int_{\sigma'}^{\sigma''} d^4x \bar{\eta}(i\overrightarrow{\not\partial} - m)\psi, \tag{50b}$$

relations (46) and (48) become

$$\langle \psi'', \bar{\psi}'', \sigma'' | \, T(\delta_\eta F(\psi, \bar{\psi})) \, | \psi', \bar{\psi}', \sigma' \rangle = -i \int_{\sigma'}^{\sigma''} d^4x \, \langle \psi'', \bar{\psi}'', \sigma'' | \times T(F(\psi, \bar{\psi})\bar{\psi}(x)) \, | \psi', \bar{\psi}', \sigma' \rangle (-i\overleftarrow{\not\partial} - m)\eta \tag{51a}$$

and

$$\langle \psi'', \bar{\psi}'', \sigma'' | \, T(\delta_{\bar{\eta}} F(\psi, \bar{\psi})) \, | \psi', \bar{\psi}', \sigma' \rangle = -i \int d^4x \, \langle \psi'', \bar{\psi}'', \sigma'' | \times T(F(\psi, \bar{\psi})\bar{\eta}(-i\overrightarrow{\not\partial} - m)\psi(x)) \, | \psi', \bar{\psi}', \sigma' \rangle. \tag{51b}$$

Let us define the sign indicator S_F such that S_F is $+1$ or -1 when $F(\psi, \bar{\psi})$ transforms like the product of an even or odd number of spinors, respectively. We can now write Eq. (51b) as

$$\langle \psi'', \bar{\psi}'', \sigma'' | \, T(\delta_\eta F(\psi, \bar{\psi})) \, | \psi', \bar{\psi}', \sigma' \rangle = -iS_F \int d^4x \eta(x)(i\overrightarrow{\not\partial} - m)_x \times \langle \psi'', \bar{\psi}'', \sigma'' | \, T(F(\psi, \bar{\psi})\psi(x)) \, | \psi', \bar{\psi}', \sigma' \rangle. \tag{51c}$$

Using the definitions of functional derivatives in Eqs. (43) and (45), it follows directly from Eqs. (51a) and (51c) that

$$\langle \psi'', \bar{\psi}'', \sigma'' | \, T\left(\frac{\delta F(\psi, \bar{\psi})}{\delta \psi(x)}\right) | \psi', \bar{\psi}', \sigma' \rangle = -i \langle \psi'', \bar{\psi}'', \sigma'' | \, T(F(\psi, \bar{\psi})\bar{\psi}(x)) \, | \psi', \bar{\psi}', \sigma' \rangle \times (-i\overleftarrow{\not\partial} - m) \tag{52a}$$

and

$$\langle \psi'', \bar{\psi}'', \sigma'' | \, T\left(\frac{\delta F(\psi, \bar{\psi})}{\delta \psi(x)}\right) | \psi', \bar{\psi}', \sigma' \rangle = -i(i\overrightarrow{\not\partial} - m) \langle \psi'', \bar{\psi}'', \sigma'' | \, T(\psi(x)F(\psi, \bar{\psi})) \, | \psi', \bar{\psi}', \sigma' \rangle. \tag{52b}$$

Equations (52a) and (52b) can be generalized for the nth functional derivative:

$$\langle \psi'', \bar{\psi}'', \sigma'' | \, T\left(\frac{\delta^N F(\psi, \bar{\psi})}{\delta \psi(x_N) \cdots \delta \psi(x_1)}\right) | \psi', \bar{\psi}', \sigma' \rangle = (-1)^N \langle \psi'', \bar{\psi}'', \sigma'' | \, T(F(\psi, \bar{\psi})\bar{\psi}(x_1) \cdots \bar{\psi}(x_n)) \times | \psi', \bar{\psi}', \sigma' \rangle (-i\overleftarrow{\not\partial} - m)_{x_1} \cdots (i\overrightarrow{\not\partial} - m)_{x_N} \tag{53a}$$

and

$$\langle \psi'', \bar{\psi}'', \sigma'' | \, T\left(\frac{\delta^N F(\psi, \bar{\psi})}{\delta \bar{\psi}(x_N) \cdots \delta \bar{\psi}(x_1)}\right) | \psi', \bar{\psi}', \sigma' \rangle = (-i)^N (i\overrightarrow{\not\partial} - m)_{x_N} \cdots (i\overrightarrow{\not\partial} - m)_{x_1} \times \langle \psi'', \bar{\psi}'', \sigma'' | \, T(\psi(x_N) \cdots \psi(x_1)F(\psi, \bar{\psi})) \, | \psi', \bar{\psi}', \sigma' \rangle. \tag{53b}$$

The formulas for integration by parts follow directly from (52a) and (52b) by letting

$$F = GR. \tag{54}$$

Substituting the first variational derivative

$$\frac{\delta}{\delta \psi} G(\psi, \bar{\psi})R(\psi, \bar{\psi}) = S_R \frac{\delta G}{\delta \psi}(\psi, \bar{\psi})R(\psi, \bar{\psi}) + G(\psi, \bar{\psi}) \frac{\delta R}{\delta \psi}(\psi, \bar{\psi}) \tag{55}$$

into Eq. (52a), we get

$$\langle \psi'', \bar{\psi}'', \sigma''| \, T\left(G(\psi, \bar{\psi}) \frac{\delta R(\psi, \bar{\psi})}{\delta \psi}\right) |\psi', \bar{\psi}', \sigma'\rangle$$
$$= -S_R \langle \psi'', \bar{\psi}'', \sigma''| \, T\left(\frac{\delta G}{\delta \psi}(\psi, \bar{\psi}) R(\psi, \bar{\psi})\right) |\psi', \bar{\psi}', \sigma'\rangle$$
$$- i \langle \psi'', \bar{\psi}'', \sigma''| \, T(G(\psi, \bar{\psi}) R(\psi, \bar{\psi}) \psi(x)) |\psi', \bar{\psi}', \sigma'\rangle$$
$$\times (-i\overleftarrow{\not{\nabla}} - m)_x. \tag{56a}$$

Likewise, for $\bar{\psi}(x)$ we get

$$\langle \psi'', \bar{\psi}'', \sigma''| \, T\left(\frac{\delta G(\psi, \bar{\psi})}{\delta \bar{\psi}(x)} R(\psi, \bar{\psi})\right) |\psi', \bar{\psi}', \sigma'\rangle$$
$$= -S_G \langle \psi'', \bar{\psi}'', \sigma''| \, T\left(G(\psi, \bar{\psi}) \frac{\delta R(\psi, \bar{\psi})}{\delta \bar{\psi}(x)}\right) |\psi', \bar{\psi}', \sigma'\rangle$$
$$- i(i\vec{\not{\nabla}} - m) \langle \psi'', \bar{\psi}'', \sigma''|$$
$$\times T(\psi(x) G(\psi, \bar{\psi}) R(\psi, \bar{\psi})) |\psi', \bar{\psi}', \sigma'\rangle. \tag{56b}$$

Let us now consider the general case of integrating by parts N times. Using the fact that, for $N > 1$,

$$G(\psi, \bar{\psi}) \frac{\delta^N R(\psi, \bar{\psi})}{\delta \bar{\psi}(x_1) \delta \bar{\psi}(x_N)}$$
$$= + S_G \frac{\delta}{\delta \bar{\psi}(x_1)}\left(G(\psi, \bar{\psi}) \frac{\delta^{N-1} R(\psi, \bar{\psi})}{\delta \bar{\psi}(x_2) \cdots \delta \bar{\psi}(x_N)}\right)$$
$$+ (S_G)^2 \frac{\delta}{\delta \bar{\psi}(x_2)}\left(\frac{\delta G(\psi, \bar{\psi})}{\delta \bar{\psi}(x_1)} \frac{\delta^{N-2} R(\psi, \bar{\psi})}{\delta \bar{\psi}(x_3) \cdots \delta \bar{\psi}(x_N)}\right)$$
$$- (S_G)^3 \frac{\delta}{\delta \bar{\psi}(x_3)}\left(\frac{\delta^2 G(\psi, \bar{\psi})}{\delta \bar{\psi}(x_2) \delta \bar{\psi}(x_1)} \frac{\delta^{N-3} R(\psi, \bar{\psi})}{\delta \bar{\psi}(x_4) \cdots \delta \bar{\psi}(x_N)}\right)$$
$$+ \cdots + (-1)^N (S_G)^N \frac{\delta}{\delta \bar{\psi}(x_N)}$$
$$\times \left(\frac{\delta^{N-1} G(\psi, \bar{\psi})}{\delta \bar{\psi}(x_{N-1}) \cdots \delta \bar{\psi}(x_1)} R(\psi, \bar{\psi})\right)$$
$$- (-1)^N (S_G)^N \frac{\delta^N G(\psi, \bar{\psi})}{\delta \bar{\psi}(x_N) \cdots \delta \bar{\psi}(x_1)} R(\psi, \bar{\psi}) \tag{57a}$$

and

$$\frac{\delta^N G(\psi, \bar{\psi})}{\delta \bar{\psi}(x_1) \cdots \delta \bar{\psi}(x_N)} R(\psi, \bar{\psi})$$
$$= S_R \frac{\delta}{\delta \psi(x_1)}\left\{\frac{\delta^{N-1} G(\psi, \bar{\psi})}{\delta \psi(x_2) \cdots \delta \psi(x_N)} R(\psi, \bar{\psi})\right\}$$
$$+ (S_R)^2 \frac{\delta}{\delta \psi(x_2)}\left\{\frac{\delta^{N-2} G(\psi, \bar{\psi})}{\delta \psi(x_3) \cdots \delta \psi(x_N)} \frac{\delta R(\psi, \bar{\psi})}{\delta \psi(x_1)}\right\}$$
$$+ (-1)^N (S_R)^N \frac{\delta}{\delta \psi(x_N)}$$
$$\times \left\{G(\bar{\psi}, \psi) \frac{\delta^{N-1} R(\psi, \bar{\psi})}{\delta \psi(x_{N-1}) \cdots \delta \psi(x_1)}\right\}$$
$$- (-1)^N (S_R)^N G(\psi, \bar{\psi}) \frac{\delta^N R(\psi, \bar{\psi})}{\delta \psi(x_N) \cdots \delta \psi(x_1)}, \tag{57b}$$

Eqs. (56a) and (56b) generalize to

$$\langle \psi'', \bar{\psi}'', \sigma''| \, T\left(\frac{\delta^N G(\psi, \bar{\psi})}{\delta \psi(x_1) \cdots \delta \psi(x_N)} R(\psi, \bar{\psi})\right) |\psi', \bar{\psi}', \sigma'\rangle$$
$$= (-1)^{N+1} S_R^N \langle \psi'', \bar{\psi}'', \sigma''| \, T\left(G(\psi, \bar{\psi}) \frac{\delta^N R(\psi, \bar{\psi})}{\delta \psi(x_N) \cdots \delta \psi(x_1)}\right) |\psi', \bar{\psi}', \sigma'\rangle$$
$$- i S_R \langle \psi'', \bar{\psi}'', \sigma''| \, T\left(\frac{\delta^{N-1} G(\psi, \bar{\psi})}{\delta \psi(x_2) \cdots \delta \psi(x_N)} R(\psi, \bar{\psi}) \bar{\psi}(x_1)\right) |\psi', \bar{\psi}', \sigma'\rangle (-i\not{\nabla} - m)_1$$
$$- i S_R^2 \langle \psi'', \bar{\psi}'', \sigma''| \, T\left(\frac{\delta^{N-2} G(\psi, \bar{\psi})}{\delta \psi(x_3) \cdots \delta \psi(x_N)} \frac{\delta R(\psi, \bar{\psi})}{\delta \psi(x_1)} \bar{\psi}(x_2)\right) |\psi', \bar{\psi}', \sigma'\rangle (-i\not{\nabla} - m)_2$$
$$- i(-1)^N S_R^N \langle \psi'', \bar{\psi}'', \sigma''| \, T\left(G(\psi, \bar{\psi}) \frac{\delta^{N-1} R(\psi, \bar{\psi})}{\delta \psi(x_{N-1}) \cdots \delta \psi(x_1)} \bar{\psi}(x_N)\right) |\psi', \bar{\psi}', \sigma'\rangle (-i\not{\nabla} - m)_N \tag{58a}$$

and

$$\langle \psi'', \bar{\psi}'', \sigma''| \, T\left(G(\psi, \bar{\psi}) \frac{\delta^N R(\psi, \bar{\psi})}{\delta \bar{\psi}(x_1) \cdots \delta \bar{\psi}(x_N)}\right) |\psi', \bar{\psi}', \sigma'\rangle$$
$$= -(-1)^N S_G^N \langle \psi'', \bar{\psi}'', \sigma''| \, T\left(\frac{\delta^N G(\psi, \bar{\psi})}{\delta \bar{\psi}(x_N) \cdots \delta \bar{\psi}(x_1)} R(\psi, \bar{\psi})\right) |\psi', \bar{\psi}', \sigma'\rangle$$
$$- i S_G (i\vec{\not{\nabla}} - m) \langle \psi'', \bar{\psi}'', \sigma''| \, T\left(\psi(x_1) G(\psi, \bar{\psi}) \frac{\delta^{N-1} R(\psi, \bar{\psi})}{\delta \bar{\psi}(x_2) \cdots \delta \bar{\psi}(x_N)}\right) |\psi', \bar{\psi}', \sigma'\rangle$$
$$- i S_G^2 (i\vec{\not{\nabla}} - m)_2 \langle \psi'', \bar{\psi}'', \sigma''| \, T\left(\psi(x_N) \frac{\delta G(\psi, \bar{\psi})}{\delta \bar{\psi}(x_1)} \frac{\delta^{N-2} R(\psi, \bar{\psi})}{\delta \bar{\psi}(x_3) \cdots \delta \bar{\psi}(x_N)}\right) |\psi', \bar{\psi}', \sigma'\rangle$$
$$- i(-1)^N S_G^N (i\vec{\not{\nabla}} - m)_N \langle \psi'', \bar{\psi}'', \sigma''| \, T\left(\psi(x_N) \frac{\delta^{N-1} G(\psi, \bar{\psi})}{\delta \bar{\psi}(x_{N-1}) \cdots \delta \bar{\psi}(x_1)} R(\psi, \bar{\psi})\right) |\psi', \bar{\psi}', \sigma'\rangle. \tag{58b}$$

IV. APPLICATIONS

In this section we would like to illustrate the use of the previous relations by solving some specific examples.

A. Example of Variational Formula

Let us consider the evaluation of the functional integral

$$\int d^4y g(y) \times \langle \phi'', \sigma''| T\left(\log\left[\int f(x)e^{\phi(x)/\beta(x)}\, d^4x\right]\phi(y)\right)|\phi', \sigma'\rangle, \tag{59}$$

where

$$(\Box + m^2)\beta(x) = g(x). \tag{60}$$

This functional integral can be solved by using Eq. (21). Consider the first variation of the functional

$$F(\phi) = \log\{f(x)e^{\phi(x)/\beta(x)}\, d^4x\}. \tag{61}$$

Its value is

$$\delta_\beta F(\phi) = 1. \tag{62}$$

Thus we obtain the solution of (59) by substituting Eq. (62) into Eq. (21):

$$\begin{aligned} 1 &= i\int d^4y \langle \phi'', \sigma''| \\ &\quad \times T\left(\log\left\{\int f(x)e^{\phi(x)/\beta(x)}\, d^4x\right\}\right. \\ &\quad \left.\times \phi(g)(\Box + m^2)_y\beta(y)\right)|\phi', \sigma'\rangle \\ &= i\int d^4y g(y) \langle \phi'', \sigma''| \\ &\quad \times T\left(\log\left\{\int f(x)e^{\phi(x)/\beta(x)}\, d^4x\right\}\phi(y)\right)|\phi', \sigma'\rangle. \end{aligned} \tag{63}$$

B. Example of Functional Derivative Formula

Let us illustrate the use of Eq. (26) by solving for the ratio

$$R(\phi) = \frac{\langle \phi'', \sigma''| T\left(\phi(x)\exp\left(i\int g(y)\phi(y)\, d^4y\right)\right)|\phi', \sigma'\rangle}{\langle \phi'', \sigma''| T\left(\exp\left(i\int g(y)\phi(y)\, d^4y\right)\right)|\phi', \sigma'\rangle}. \tag{64}$$

Let

$$F(\phi) = \exp\left(i\int g(y)\phi(y)\, d^4y\right). \tag{65}$$

Taking its functional derivative, we obtain

$$\frac{\delta f(\phi)}{\delta\phi(x)} = ig(x)F(\phi). \tag{66}$$

Substituting Eq. (66) in Eq. (26), we get

$$\begin{aligned} i(\Box + m^2)\langle \phi'', \sigma''| T(F(\phi)\phi(x))|\phi', \sigma'\rangle \\ = ig(x)\langle \phi'', \sigma''| T(F(\phi))|\phi', \sigma'\rangle. \end{aligned} \tag{67}$$

Therefore the ratio in Eq. (64) is given by the solution of the simple differential equation

$$(\Box + m^2)_x R(\phi(x)) = g(x) \tag{68}$$

with the appropriate boundary conditions.

C. Example of Integration by Parts

We will now solve for the ratio

$$B[\phi(x)] = \frac{\langle \phi'', \sigma''| T\left(\phi(x)a\sin\left(\eta\int\phi(y)\, d^4y\right) - \eta\cos\left(\eta\int\phi(y)\, d^4y\right)\exp\left(a\int\phi(y)\, d^4y\right)\right)|\phi', \sigma'\rangle}{\langle \phi'', \sigma''| T\left(\sin\left(\eta\int\phi(y)\, d^4y\right)\exp\left(a\int\phi(y)\, d^4y\right)\right)|\phi', \sigma'\rangle} \tag{69}$$

by means of Eq. (32), where $N = 2$.

Let

$$G(\phi) = \exp\left(a\int\phi(y)\, d^4y\right), \quad R(\phi) = \sin\left(\eta\int\phi(y)\, d^4y\right). \tag{70}$$

Taking the appropriate functional derivatives, we have

$$\frac{\delta G}{\delta\phi(x_1)} = aG, \tag{71a}$$

$$\frac{\delta^2 G}{\delta\phi(x_1)\delta\phi(x_2)} = a^2 G, \tag{71b}$$

$$\frac{\delta R}{\delta\phi(x_1)} = \eta\cos\left(\eta\int\phi(y)\, d^4y\right), \tag{71c}$$

$$\frac{\delta^2 R}{\delta\phi(x_1)\delta\phi(x_2)} = -\eta^2\sin\left(\eta\int\phi(y)\, d^4y\right). \tag{71d}$$

Substituting the above relations into Eq. (32), we obtain

$$-\eta^2 \langle \phi'', \sigma''| T\left(\exp\left(a\int \phi(y)\,d^4y\right)\sin\left(\eta\int \phi(y)\,d^4y\right)\right)|\phi', \sigma'\rangle$$
$$= i(\Box + m^2)_{x_2}\langle \phi'', \sigma''| T\left(\phi(x_2)\eta \exp\left(a\int \phi(y)\,d^4y\right)\cos\left(\eta\int \phi(y)\,d^4y\right)\right)|\phi, \sigma'\rangle$$
$$- i(\Box + m^2)_{x_1}\langle \phi'', \sigma''| T\left(\phi(x_1)a \exp\left(a\int \phi(y)\,d^4y\right)\sin\left(\eta\int \phi(y)\,d^4y\right)\right)|\phi', \sigma'\rangle$$
$$+ a^2\langle \phi'', \sigma''| T\left(\exp\left(a\int \phi(y)\,d^4y\right)\sin\eta\int \phi(y)\,d^4y\right)|\phi', \sigma'\rangle. \quad (72)$$

Letting $x_1 = x_2$, it follows immediately that

$$i(\Box + m^2)B(\phi(x)) = (a^2 + \eta^2). \quad (73)$$

Consequently, the explicit relation for Eq. (69) follows directly from the solution of Eq. (73).

V. GENERALIZED MATTHEWS–SALAM EQUATIONS

The Matthews–Salam equations are special cases of the previous theorems. In order to show this, let us relate the functional integrals to the vacuum expectation value of a time-ordered product.

Let $L_i(\cdot)$ be the interacting Lagrangian of the system. Then the vacuum expectation value of the time-ordered product of $F(\cdot)$ is defined as

$$\langle F(\cdot)\rangle = \frac{\frac{1}{N}\int_{[0,0]} F(\cdot)e^{i(I_{\sigma''\sigma'}+I_{\sigma''\sigma'^{,i}})}\delta(\cdot)}{\frac{1}{N}\int_{[0,0]} e^{i(I_{\sigma''\sigma'}+I_{\sigma''\sigma'^{,i}})}\delta(\cdot)}, \quad (74)$$

where

$$I^i_{\sigma''\sigma'} = \int L_i(\cdot)\,d^4x \quad (75)$$

and $t' \to -\infty$, $t'' \to +\infty$. The integration is performed over the appropriate fields.

The Matthews–Salam equations[8] follow from integrating the numerator in Eq. (74) by parts once. Using Eqs. (31), (38), (56a), (56b), and letting

$$G(\cdot) = F(\cdot),$$
$$R(\cdot) = \exp\left(i\int L_i(\cdot)\,d^4x\right), \quad (76)$$

we get the ordinary Matthews–Salam equations:

(a) Scalar fields:

$$i(\Box + m^2)_x\langle\phi(x)F(\phi)\rangle = i\langle F(\phi)j_\phi\rangle + \left\langle\frac{\delta F(\phi)}{\delta\phi(x)}\right\rangle, \quad (77)$$

where

$$j_\phi = \int \frac{\delta L_i(\phi)}{\delta\phi(x)}\,d^4x; \quad (78)$$

(b) Massless vector fields:

$$i\Box_x\langle A_\mu(x)F(A)\rangle = i\langle F(A)j_{A_\mu}\rangle + \left\langle\frac{\delta F(A)}{\delta A_\mu(x)}\right\rangle, \quad (79)$$

where

$$j_A = \int \frac{\delta L_i(A)}{\delta A^\mu(x)}\,d^4x; \quad (80)$$

(c) Spin-$\frac{1}{2}$ fields:

$$-i(i\overrightarrow{\not\nabla} - m)\langle\psi(x)F(\psi,\bar\psi)\rangle = S_F\langle F(\psi,\bar\psi)j_{\bar\psi}\rangle + \left\langle\frac{\delta F}{\delta\bar\psi(x)}\right\rangle \quad (81)$$

and

$$-i\langle F(\psi,\bar\psi)\bar\psi(x)\rangle(i\overleftarrow{\not\nabla} - m) = S\langle F(\psi,\bar\psi)j_\psi\rangle + \left\langle\frac{\delta F}{\delta\psi(x)}\right\rangle, \quad (82)$$

where

$$j_{\bar\psi} = \int \frac{\delta L_i(\psi,\bar\psi)}{\delta\bar\psi(x)}\,d^4x,$$
$$j_\psi = \int \frac{\delta L_i(\psi,\bar\psi)}{\delta\psi(x)}\,d^4x. \quad (83)$$

This equation can be generalized by integrating by parts N times. It follows from equations (32), (38), (58a), and (58b) that the generalized Matthews–Salam equations are:

(a) Scalar field:

$$i(\Box + m^2)_{x_N}\left\langle\phi(x_N)F(\phi)\frac{\delta^{N-1}e^{iI^i(\phi)}}{\delta\phi(x_{N-1})\cdots\delta\phi(x_n)}e^{-iI^i(\phi)}\right\rangle$$
$$- i(\Box + m^2)_{x_{N-1}}\left\langle\phi(x_{N-1})\frac{\delta F(\phi)}{\delta\phi(x_N)}\frac{\delta^{N-2}e^{iI^i(\phi)}}{\delta\phi(x_{N-2})\cdots\delta\phi(x_1)}e^{-iI^i(\phi)}\right\rangle$$
$$+ i(-1)^{N+1}(\Box + m^2)_{x_1}\left\langle\phi(x_1)\frac{\delta^{N-1}F(\phi)}{\delta\phi(x_2)\cdots\delta\phi(x_N)}\right\rangle$$
$$= \left\langle F(\phi)\frac{\delta^N e^{iI^i(\phi)}}{\delta\phi(x_N)\cdots\delta\phi(x_1)}e^{-iI^i(\phi)}\right\rangle - (-1)^N\left\langle\frac{\delta^N F(\phi)}{\delta\phi(x_1)\cdots\delta\phi(x_N)}\right\rangle; \quad (84)$$

(b) Massless vector field:

$$i\Box_{x_N}\left\langle A_{\mu_N}(x_N)F(A)\frac{\delta^{N-1}e^{iI^i(A)}}{\delta A^{\mu_{N-1}}(x_{N-1})\cdots\delta A^{\mu_1}(x_1)}e^{-iI^i(A)}\right\rangle$$
$$+ i(-1)^{N+1}\Box_{x_1}\left\langle A_\mu(x_1)\frac{\delta^{N-1}F(A)}{\delta A^{\mu_2}(x_2)\cdots\delta A^{\mu_N}(x_N)}\right\rangle$$
$$= \left\langle F\frac{\delta^N e^{iI^i(A)}}{\delta A^{\mu_N}(x_N)\cdots\delta^{\mu_1}A(x_1)}e^{-iI(A)}\right\rangle - (-1)^N\left\langle\frac{\delta^N F(A)}{\delta A^{\mu_1}(x_1)\cdots\delta A^{\mu_N}(x_N)}\right\rangle; \quad (85)$$

(c) Spin-$\frac{1}{2}$ field ($N > 1$):

$$-iS_F(i\vec{\not\partial} - m)\left\langle\psi(x_1)F(\psi,\bar\psi)\frac{\delta^{N-1}e^{iI^i(\psi,\bar\psi)}}{\delta\bar\psi(x_2)\cdots\delta\bar\psi(x_N)}e^{-iI^i(\psi,\bar\psi)}\right\rangle$$
$$- iS_F^2(i\vec{\not\partial} - m)\left\langle\psi(x_2)\frac{\delta F(\psi,\bar\psi)}{\delta\bar\psi(x_1)}\frac{\delta^{N-2}e^{iI^i(\psi,\bar\psi)}}{\delta\bar\psi(x_3)\cdots\delta\bar\psi(x_N)}e^{-iI^i(\psi,\bar\psi)}\right\rangle$$
$$\vdots$$
$$- i(-1)^N S_F^N(i\vec{\not\partial} - m)_N\left\langle\psi(x_N)\frac{\delta^{N-1}F(\psi,\bar\psi)}{\delta\bar\psi(x_{N-1})\cdots\delta\bar\psi(x_1)}\right\rangle$$
$$= \left\langle F(\psi,\bar\psi)\frac{\delta^N e^{iI^i(\psi,\bar\psi)}}{\delta\bar\psi(x_1)\cdots\delta\bar\psi(x_N)}e^{-iI^i(\psi,\bar\psi)}\right\rangle + (-1)^N\left\langle\frac{\delta^N F(\psi,\bar\psi)}{\delta\bar\psi(x_N)\cdots\delta\bar\psi(x_1)}\right\rangle \quad (86a)$$

and

$$-iS_F\left\langle e^{-iI^i(\psi,\bar\psi)}\frac{\delta^{N-1}e^{+iI^i(\psi,\bar\psi)}}{\delta\psi(x_2)\cdots\delta\psi(x_N)}F(\psi,\bar\psi)\bar\psi(x_1)\right\rangle(-i\overleftarrow{\not\partial} - m)$$
$$-iS_F^2\left\langle e^{-iI^i(\psi,\bar\psi)}\frac{\delta^{N-2}e^{+iI(\psi,\bar\psi)}}{\delta\psi(x_3)\cdots\delta\psi(x_N)}\frac{\delta F(\psi,\bar\psi)}{\delta\psi(x_1)}\bar\psi(x_2)\right\rangle(-i\overleftarrow{\not\partial} - m)_2$$
$$\vdots$$
$$- i(-1)^N S_F^N\left\langle\frac{\delta^{N-1}F(\psi,\bar\psi)}{\delta\psi(x_{N-1})\cdots\delta\psi(x_1)}\bar\psi(x_N)\right\rangle(-i\overleftarrow{\not\partial} - m)_2$$
$$= \left\langle e^{-iI^i(\psi,\bar\psi)}\frac{\delta^N e^{iI^i(\psi,\bar\psi)}}{\delta\psi(x_1)\cdots\delta\psi(x_N)}F(\psi,\bar\psi)\right\rangle + (-1)^N S_F^N\left\langle\frac{\delta^N F(\psi,\bar\psi)}{\delta\psi(x_N)\cdots\delta\psi(x_1)}\right\rangle. \quad (86b)$$

Higher-order differential equations for the vacuum expectation values may also be obtained from Eqs. (27), (36), (53a), and (53b).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 12 DECEMBER 1969

Application of Perturbation Theory to Many-Body Systems with Localized Particles*

C. C. ROUSSEAU
Department of Physics, Baylor University, Waco, Texas
AND
D. P. SAYLOR
Department of Physics, Texas A & M University, College Station, Texas

(Received 29 April 1969)

We investigate the possibility of using perturbation theory to compute the binding energy for infinite systems in which the particles are localized. For the case of the linear chain of coupled harmonic oscillators, we prove that the perturbation series for the ground-state energy per particle is convergent. Exact expressions for the generalized Padé approximants are derived. The generalized approximants provide a manifestly convergent sequence of approximations to the energy.

1. INTRODUCTION

Consider the problem of computing the cohesive energy of one of the rare gas solids. A method which works reasonably well except for helium is the Hartree method.[1] The Hartree method is based on the trial function

$$\Psi(r_1, r_2, \cdots, r_N) = \prod_{i=1}^{N} \phi(r_i - R_i). \quad (1.1)$$

The single-particle functions which give the lowest total energy for the system are solutions of

$$-(\hbar^2/2M)\nabla^2\phi(r) + V_H(r)\phi(r) = \epsilon\phi(r), \quad (1.2)$$

where $\phi(r)$ and $V_H(r)$ must be determined self-consistently. Thus, the Hartree solution is an eigenstate of

$$H_0 = \sum_{i=1}^{N} [p_i^2/2M + V_H(r_i)], \quad (1.3)$$

so we can write

$$H = H_0 + \lambda H_1, \quad (1.4)$$

where

$$H_1 = \sum_{j>k=1}^{N} V(r_j - r_k) - \sum_{i=1}^{N} V_H(r_i). \quad (1.5)$$

Then we contemplate an application of perturbation theory in order to account for the lowering of the energy due to H_1. It is obvious that the practical problems associated with such a calculation are formidable. Therefore, before seriously considering such an attack, it is worthwhile to know the mathematical properties of such a perturbation expansion when applied to the simplest possible model problem. For the purpose of this study, we consider the linear chain of coupled harmonic oscillators. With rigid wall boundary conditions applied to the N-particle system, the Hamiltonian is

$$H = \sum_{i=1}^{N} \frac{p_i^2}{2M} + \frac{k}{2}\sum_{i=1}^{N-1}(x_{i+1} - x_i)^2 + \frac{k}{2}(x_1^2 + x_N^2). \quad (1.6)$$

The normal-mode frequencies for this system are

$$\omega_j = 2\left(\frac{k}{M}\right)^{\frac{1}{2}} \sin\left(\frac{j\pi}{2(N+1)}\right), \quad j = 1, 2, \cdots, N, \quad (1.7)$$

and the ground-state energy is

$$E_0(N) = \hbar\left(\frac{k}{M}\right)^{\frac{1}{2}} \sum_{j=1}^{N} \sin\left(\frac{j\pi}{2(N+1)}\right) = \frac{\hbar}{2}\left(\frac{k}{M}\right)^{\frac{1}{2}}\left[\cot\left(\frac{\pi}{4(N+1)}\right) - 1\right]. \quad (1.8)$$

The energy per particle of the infinite chain is

$$\lim_{N\to\infty} \frac{E_0(N)}{N} = \frac{2}{\pi}\hbar\left(\frac{k}{M}\right)^{\frac{1}{2}}. \quad (1.9)$$

Now we ask whether this result can be calculated by a straightforward application of Rayleigh–Schrödinger perturbation theory.

2. EXPLICIT CALCULATION OF THE PERTURBATION SERIES THROUGH FOURTH ORDER

In order to apply the perturbation method to this problem, we write (1.6) in the form

$$H = H_0 + \lambda H_1, \quad (2.1)$$

where

$$H_0 = \sum_{i=1}^{N} \frac{p_i^2}{2M} + k\sum_{i=1}^{N} x_i^2 \quad (2.2)$$

and

$$H_1 = -k\sum_{i=1}^{N-1} x_i x_{i+1}. \quad (2.3)$$

As usual, we expand in powers of λ and then set $\lambda = 1$. We shall simply compute the expansion for the ground-state energy using the standard technique of Rayleigh–Schrödinger perturbation theory. We note

* Supported in part by the United States Air Force Office of Scientific Research under Contract AF 918-67.

[1] L. H. Nosanow and G. L. Shaw, Phys. Rev. **128**, 546 (1962).

at this point that the application of perturbation theory to this model problem has been discussed previously, but from a completely different point of view.[2]

Our system consists of N distinguishable particles. An excited state of the unperturbed system is specified by noting the harmonic oscillator level occupied by each of the N particles. Such an excited state can be denoted by

$$|n_1, n_2, \cdots, n_N\rangle.$$

It is clear that there is no first-order contribution to the energy. In second order

$$E_0^{(2)} = \langle 0| H_1 \frac{P}{\epsilon_0 - H_0} H_1 |0\rangle = \sum_\alpha \frac{\langle 0| H_1 |\alpha\rangle \langle \alpha| H_1 |0\rangle}{\epsilon_0 - \epsilon_\alpha}, \quad (2.4)$$

where

$$|0\rangle = |0, 0, \cdots, 0, 0, \cdots, 0, 0\rangle \quad (2.5)$$

and

$$|\alpha\rangle = |0, 0, \cdots, 1, 1, \cdots, 0, 0\rangle. \quad (2.6)$$

Now

$$\langle \alpha| H_1 |0\rangle = \frac{\hbar}{4}\left(\frac{2k}{M}\right)^{\frac{1}{2}} \quad (2.7)$$

and

$$\epsilon_\alpha - \epsilon_0 = 2\hbar(2k/M)^{\frac{1}{2}} \quad (2.8)$$

for each of the $N - 1$ possible intermediate states. Thus

$$E_0^{(2)} = -\hbar\left(\frac{2k}{M}\right)^{\frac{1}{2}} \frac{N-1}{32}. \quad (2.9)$$

Once again, it is clear that the third-order contribution vanishes. In fourth order we shall need to compute the "regular" term,

$$\langle 0| H_1 \frac{P}{\epsilon_0 - H_0} H_1 \frac{P}{\epsilon_0 - H_0} H_1 \frac{P}{\epsilon_0 - H_0} H_1 |0\rangle, \quad (2.10)$$

and one "irregular" term,

$$-\langle 0| H_1 \frac{P}{\epsilon_0 - H_0} H_1 |0\rangle \langle 0| H_1 \frac{P}{(\epsilon_0 - H_0)^2} H_1 |0\rangle. \quad (2.11)$$

All of the other "irregular" terms vanish since $\langle 0| H_1 |0\rangle = 0$.

We shall need a systematic way of keeping track of all the contributions to the regular terms. Let us enumerate the various contributions by simply listing in columns the individual particle levels which occur in intermediate states and grouping the contributions according to how many particles are involved. We shall refer to such a listing as a "diagram." For an n-particle diagram there will be a certain weight, which is the number of times such a diagram can occur for an N-particle system. To illustrate, in second order there is only one diagram:

00
11
00

This is a two-particle diagram and its weight is $N - 1$.

For the fourth-order calculation we list all of the diagrams below.

Two-Particle Diagrams

00	00	00
11	11	11
22	20	02
11	11	11
00	00	00

Three-Particle Diagrams

000	000	000	000	000	000	000	000	000	000	000	000
110	110	011	011	110	110	011	011	110	110	011	011
121	121	121	121	101	101	101	101	020	020	020	020
110	011	110	011	110	011	110	011	110	011	110	011
000	000	000	000	000	000	000	000	000	000	000	000

Four-Particle (Unlinked) Diagrams

00 · · · 00	00 · · · 00
11 · · · 00	11 · · · 00
11 · · · 11	11 · · · 11
11 · · · 00	00 · · · 11
00 · · · 00	00 · · · 00

[2] R. D. Mattuck, Ann. Phys. (N.Y.) 27, 216 (1964).

It is a straightforward matter of counting to see that the appropriate weights are as shown below:

n	Weight for an n-Particle Diagram,
2	$N-1$,
3	$N-2$,
4	$(N-2)(N-3)=N^2-5N+6$.

Note that we must have $N \geq 2$ in order to make sense here. For higher orders in perturbation theory, the minimum N will increase and, in general, for $2n$th order we must have $N \geq n$ in order to simply specify a weight without going into more detailed considerations. This point will be made more transparent later on in our discussion of the problem.

Before proceeding to write down the regular contribution to $E_0^{(4)}$, let us look at the irregular term. Since the matrix elements, energy denominators, and weights are just exactly the same as for $E_0^{(2)}$, we can see immediately that

$$-\langle 0|\, H_1 \frac{P}{\epsilon_0 - H_0} H_1 |0\rangle \langle 0|\, H_1 \frac{P}{(\epsilon_0 - H_0)^2} H_1 |0\rangle = \hbar\left(\frac{2k}{M}\right)^{\frac{1}{2}} \frac{(N-1)^2}{2048}. \quad (2.12)$$

Since we have already listed all of the diagrams and their weights, it is simply a matter of keeping track of factors which occur in the matrix elements and energy denominators to compute the total contribution of the regular term. When this is added to the irregular term, the result is

$$E_0^{(4)} = -\hbar\left(\frac{2k}{M}\right)^{\frac{1}{2}} \frac{1}{2048}[6(N-1) + 12(N-2) + (N^2 - 5N + 6) - (N-1)^2]. \quad (2.13)$$

The two parts which are proportional to N^2 cancel and the result is

$$E_0^{(4)} = -\hbar\left(\frac{2k}{M}\right)^{\frac{1}{2}} \frac{15N-25}{2048}. \quad (2.14)$$

Our purpose for computing the fourth-order contribution in detail is to demonstrate that the calculation is perfectly routine. The essential point is that the number of nonvanishing matrix elements is limited. Where this feature can be assumed, the entire process of enumerating the various diagrams and computing their contribution is amenable to electronic computation. By writing a computer code for this purpose, it is conceivable that one may be able to obtain many terms in the perturbation expansion.

3. THE ENTIRE SERIES

Now we ask whether such an expansion is of any utility. In particular, for the case of infinitely many particles, we wish to know whether the series is convergent for $\lambda = 1$. To this end, we first find the exact ground-state energy for the case where the Hamiltonian is given by (2.1)–(2.3). By means of the usual analysis of the coupled chain,

$$E_0(N,\lambda) = \frac{\hbar}{2}\left(\frac{2k}{M}\right)^{\frac{1}{2}} \sum_{j=1}^{N}\left[1 - \cos\left(\frac{j\pi}{N+1}\right)\lambda\right]^{\frac{1}{2}}. \quad (3.1)$$

Now we formally expand each term in the sum using the binomial theorem and then interchange the orders of summation, thus obtaining

$$E_0(N,\lambda) = \frac{\hbar}{2}\left(\frac{2k}{M}\right)^{\frac{1}{2}} \sum_{n=0}^{\infty}\binom{\frac{1}{2}}{2n}\lambda^{2n}\sum_{j=1}^{N}\cos^{2n}\left(\frac{j\pi}{N+1}\right). \quad (3.2)$$

By noting that

$$\cos^{2n} x = \frac{1}{2^{2n}}\left[\binom{2n}{n} + 2\sum_{m=1}^{n}\binom{2n}{n+m}\cos(2mx)\right] \quad (3.3)$$

and

$$\sum_{j=1}^{N}\cos\left(\frac{2mj\pi}{N+1}\right) = \begin{cases} -1, & m \neq k(N+1), \\ N, & m = k(N+1), \end{cases} \quad (3.4)$$

for $k = 0, 1, 2, \cdots$, we find

$$\sum_{j=1}^{N}\cos^{2n}\left(\frac{j\pi}{N+1}\right) = \frac{N+1}{2^{2n}}\left[\binom{2n}{n} + 2\sum_{k=1}^{L}\binom{2n}{n+k(N+1)}\right] - 1, \quad (3.5)$$

where

$$L = [n/N+1] \quad (3.6)$$

and it is understood that, if $L = 0$, the sum over k is simply omitted. Thus, the complete perturbation expansion is

$$E_0(N,\lambda) = \frac{\hbar}{2}\left(\frac{2k}{M}\right)^{\frac{1}{2}} \sum_{n=0}^{\infty}\binom{\frac{1}{2}}{2n}\left\{\frac{N+1}{2^{2n}}\left[\binom{2n}{n} + 2\sum_{k=1}^{L}\binom{2n}{n+k(N+1)}\right] - 1\right\}\lambda^{2n}. \quad (3.7)$$

In particular,

$$E_0^{(2)} = -\hbar\left(\frac{2k}{M}\right)^{\frac{1}{2}} \frac{N-1}{32} \quad (3.8)$$

and

$$E_0^{(4)} = -\hbar\left(\frac{2k}{N}\right)^{\frac{1}{2}} \frac{15N-25}{2048}, \quad N \geq 2, \quad (3.9)$$

in agreement with (2.9) and (2.14), respectively. We note at this point that the condition $L = 0$ coincides

exactly with the ability to specify simply a "weight" for each diagram. This condition is $n \leq N$, and, if it is satisfied, we can write $E_0^{(2n)} = A_{2n}N + B_{2n}$, where A_{2n} and B_{2n} are independent of N.

4. ANALYTICITY AND CONVERGENCE

By inspection of (3.1), we see that $E_0(N, \lambda)$ is analytic except for branch points on the real axis at $\sec(j\pi/N + 1)$, $j = 1, 2, \cdots, N$. Thus, for finite N the radius of convergence of the perturbation series is equal to $\sec(\pi/N + 1)$. The expansion for the energy per particle of the infinite system is clearly

$$\lim_{N\to\infty} E_0(N, \lambda)/N = \frac{\hbar}{2}\left(\frac{2k}{M}\right)^{\frac{1}{2}} \sum_{n=0}^{\infty} \frac{1}{2^{2n}} \binom{\frac{1}{2}}{2n} \binom{2n}{n} \lambda^{2n} \quad (4.1)$$

and it is convergent within the unit circle. If we set

$$C_n = \frac{1}{2^{2n}} \binom{\frac{1}{2}}{2n} \binom{2n}{n}, \quad (4.2)$$

it is easy to show that

$$\lim_{n\to\infty} n\left[1 - \frac{C_{n+1}}{C_n}\right] = 2. \quad (4.3)$$

By Raabe's test,[3] then,

$$\sum_{n=0}^{\infty} \frac{1}{2^{2n}} \binom{\frac{1}{2}}{2n} \binom{2n}{n} \quad (4.4)$$

is convergent. Consequently, by Abel's theorem,[4] (4.1) is uniformly convergent for $0 \leq \lambda \leq 1$. Although (4.4) is convergent, its rate of convergence is somewhat slow, and it behooves us to investigate methods for obtaining a more rapidly convergent sequence of approximations than we can realize from the partial sums of the series. To this end we investigate the generalized Padé approximant.

5. THE GENERALIZED APPROXIMANT

The problem of extending the usefulness of perturbation expansions is one which is common to a number of areas of theoretical physics. Thus, it is not surprising that the Padé approximant has received much attention and has been applied in such areas as statistical physics,[5] the many-body problem,[6] scattering theory,[7] and elementary-particle physics.[8] The Padé approximant can be applied when one has only a formal power series expansion and no additional information. Often, however, one has additional information which is relevant to the problem of finding an approximate analytic continuation of the power series. In order to make use of this additional information, several generalizations of the Padé approximant method have been proposed. These include the two-point Padé approximant[9] and the method to be discussed in this section.

Given a formal power series expansion $\sum a_n z^n$ and a known function,

$$g(z) = \sum_{n=0}^{\infty} b_n z^n, \quad (5.1)$$

we define the $[N, N + j]$ generalized Padé approximant to be the expression

$$F_{N,j}(z) = \sum_{n=1}^{N} \alpha_n g(\beta_n z) + \sum_{k=0}^{j} \gamma_k z^k, \quad (5.2)$$

with the α, β, and γ coefficients chosen so that

$$\sum_{n=0}^{\infty} a_n z^n - F_{N,j}(z) = O(z^{2N+j+1}). \quad (5.3)$$

It is understood that, if $j = -1$, the sum over k in (5.2) is omitted. The mathematical properties of generalized approximant have been discussed in some detail elsewhere.[10,11] At this point we simply state the method by which it is computed. In order to compute the $[N, N + j]$ generalized approximant, one must simply perform the following calculations:

(1) Compute the $[N, N + j]$ Padé approximant to the series

$$\sum_{n=0}^{\infty} (a_n/b_n) z^n, \quad (5.4)$$

i.e., find the rational function

$$\frac{P_{N+j}(z)}{Q_N(z)} = \frac{p_0 + p_1 z + \cdots + p_{N+j} z^{N+j}}{1 + q_1 z + \cdots + q_N z^N}, \quad (5.5)$$

which satisfies

$$\sum_{n=0}^{\infty} (a_n/b_n) z^n - \frac{P_{N+j}(z)}{Q_N(z)} = O(z^{2N+j+1}). \quad (5.6)$$

(2) Write the $[N, N + j]$ Padé approximant in partial-fraction form:

$$\sum_{n=1}^{N} \frac{A_n}{1 - B_n z} + \sum_{k=0}^{j} C_k z^k. \quad (5.7)$$

Then $\alpha_n = A_n$, $\beta_n = B_n$, $n = 1, 2, \cdots, N$, and $\gamma_k = b_k C_k$, $k = 0, 1, \cdots, j$.

[3] P. Dienes, *The Taylor Series* (Dover Publications, Inc., New York, 1957), p. 78.
[4] E. C. Titchmarsh, *The Theory of Functions* (Oxford University Press, London, 1932), p. 9.
[5] G. A. Baker, Jr., Phys. Rev. **124**, 768 (1961).
[6] G. A. Baker, Jr., J. L. Gammel, and B. J. Hill, Phys. Rev. **132**, 1373 (1963).
[7] S. Tani, Phys. Rev. **139**, B1011 (1965).
[8] D. Bessis and M. Pusterla, Phys. Letters **25B**, 279 (1967).
[9] G. A. Baker, Jr., "The Theory and Application of the Padé Approximant Method," in *Advances in Theoretical Physics*, K. A. Brueckner, Ed. (Academic Press, New York, 1965), Vol. I, p. 54.
[10] J. L. Gammel, C. C. Rousseau, and D. P. Saylor, J. Math. Anal. Appl. **20**, 416 (1967).
[11] G. A. Baker, Jr., Phys. Rev. **161**, 434 (1967).

In the original work on this method, the function $g(z)$ was called the "model" function.[12] The reason for this designation is that $g(z)$ may be derived from the solution of some simpler model problem. Alternatively, $g(z)$ may be chosen on the basis of its analytic structure or its asymptotic behavior. We expect that the sequences of $[N, N+j]$ generalized approximants which will be of most value are the $j=-1$ and $j=0$ sequences. Now let us compute these two sequences of generalized approximants for the series given by (4.1), choosing as our "model" function

$$g(\lambda) = \tfrac{1}{2}[(1+\lambda)^{\frac{1}{2}} + (1-\lambda)^{\frac{1}{2}}] = \sum_{n=0}^{\infty} \binom{\frac{1}{2}}{2n} \lambda^{2n}. \tag{5.8}$$

In order to compute the generalized approximants, we must find the partial-fraction form of the Padé approximants of

$$\sum_{n=0}^{\infty} \frac{1}{2^{2n}} \binom{2n}{n} z^n = (1-z)^{-\frac{1}{2}}. \tag{5.9}$$

This problem can be solved exactly with the result that

$$[N, N-1](z) = \frac{1}{N} \sum_{n=1}^{N} \left[1 - \left(\cos^2 \frac{(2n-1)\pi}{2N}\right) z\right]^{-1} \tag{5.10}$$

and

$$[N, N](z) = \frac{2}{2N+1} \sum_{n=1}^{N} \left[1 - \left(\cos^2 \frac{(2n-1)\pi}{2(2N+1)}\right) z\right]^{-1} + \frac{1}{2N+1}. \tag{5.11}$$

(These results are proved in the Appendix to this paper.) Making use of these results, as well as some simple trigonometric relations, we find

$$F_{N,-1}(\lambda) = \frac{\hbar}{2}\left(\frac{2k}{M}\right)^{\frac{1}{2}} \frac{1}{2N} \sum_{n=1}^{2N} \left[1 - \left(\cos \frac{(2n-1)\pi}{4N}\right) \lambda\right]^{\frac{1}{2}} \tag{5.12}$$

and

$$F_{N,0}(\lambda) = \frac{\hbar}{2}\left(\frac{2k}{M}\right)^{\frac{1}{2}} \frac{1}{2N+1} \times \sum_{n=1}^{2N+1} \left[1 - \left(\cos \frac{(2n-1)\pi}{2(2N+1)}\right) \lambda\right]^{\frac{1}{2}}. \tag{5.13}$$

Our major concern, of course, is the sum of the series given by (4.1) for $\lambda = 1$. Performing the sums in (5.12) and (5.13), we obtain the particularly simple results

$$F_{N,-1}(1) = \hbar\left(\frac{k}{M}\right)^{\frac{1}{2}} \frac{1}{4N} \csc\left(\frac{\pi}{8N}\right) \tag{5.14}$$

and

$$F_{N,0}(1) = \hbar\left(\frac{k}{M}\right)^{\frac{1}{2}} \frac{1}{2(2N+1)} \csc\left(\frac{\pi}{4(2N+1)}\right). \tag{5.15}$$

Both sequences of approximants are manifestly convergent to the limit $(2/\pi)\hbar(k/M)^{\frac{1}{2}}$, in agreement with (1.9). From the Laurent expansion for $\csc(z)$, we find the approximate errors of $F_{N,-1}(1)$ and $F_{N,0}(1)$ to be

$$\hbar\left(\frac{k}{M}\right)^{\frac{1}{2}} \frac{\pi}{48(2N)^2} \quad \text{and} \quad \hbar\left(\frac{k}{M}\right)^{\frac{1}{2}} \frac{\pi}{48(2N+1)^2},$$

respectively.

6. DISCUSSION

The application of perturbation theory to an infinite system in which the particles are more or less localized yields some interesting results. In the first place, the series for the energy particle is convergent for $\lambda = 1$. Secondly, the generalized Padé approximant provides a usefully convergent sequence of approximations, thus adding to the utility of the perturbation approach. It remains to be seen whether these results can be of any benefit to those seeking accurate calculations of the binding energies of the rare gas solids. In any case, we suggest that this work provides a different perspective to the use of perturbation theory in connection with infinite systems.

APPENDIX

In order to prove (5.10) and (5.11), we make use of simple trigonometric formulas and we appeal to the uniqueness property of the Padé approximant.[13] Formally expanding the right-hand side of (5.10), we have

$$\frac{1}{N} \sum_{j=1}^{N} \left[1 - \left(\cos^2 \frac{(2j-1)\pi}{4N}\right) z\right]^{-1} = \sum_{n=0}^{\infty} \left(\frac{1}{N} \sum_{j=1}^{N} \cos^{2n} \frac{(2j-1)\pi}{4N}\right) z^n. \tag{A1}$$

Now making use of (3.3) and noting that

$$\sum_{j=1}^{N} \cos\left(\frac{m(2j-1)\pi}{2N}\right) = \begin{cases} 0, & m \neq k(2N), \\ -N, & m = k(2N), \end{cases} \tag{A2}$$

for $k = 1, 2, \cdots$, we see immediately that

$$\frac{1}{N} \sum_{j=1}^{N} \cos^{2n} \frac{(2j-1)\pi}{4N} = \frac{1}{2^{2n}} \binom{2n}{n}, \quad n < 2N. \tag{A3}$$

[12] D. P. Saylor, J. L. Gammel, and C. Rousseau, Bull. Am. Phys. Soc., Ser. II 12, 83 (1967).

[13] H. S. Wall, *Analytic Theory of Continued Fractions* (Chelsea Publ. Co., New York, 1967), p. 377.

Similarly, let us formally expand the right-hand side of (5.11):

$$\frac{2}{2N+1}\sum_{j=1}^{N}\left[1-\left(\cos^2\frac{(2j-1)\pi}{2(2N+1)}\right)z\right]^{-1}+\frac{1}{2N+1}$$
$$=1+\sum_{n=1}^{\infty}\left(\frac{2}{2N+1}\sum_{j=1}^{N}\cos^{2n}\frac{(2j-1)\pi}{2(2N+1)}\right)z^n. \quad \text{(A4)}$$

Once again, we make use of (3.3) and note that

$$\sum_{j=1}^{N}\cos\left(\frac{m(2j-1)\pi}{2N+1}\right)$$
$$=\begin{cases}\frac{1}{2}, & m\neq k(2N+1) \text{ and odd},\\ -\frac{1}{2}, & m\neq k(2N+1) \text{ and even},\\ -N, & m=k(2N+1),\end{cases} \quad \text{(A5)}$$

for $k=1,2,\cdots$. Thus

$$\sum_{j=1}^{N}\cos^{2n}\frac{(2j-1)\pi}{2(2N+1)}$$
$$=\frac{1}{2^{2n}}\left[\binom{2n}{n}N+\sum_{\substack{m=1\\ \text{odd}}}^{n}\binom{2n}{n+m}-\sum_{\substack{m=2\\ \text{even}}}^{n}\binom{2n}{n+m}\right]$$
$$=\frac{1}{2^{2n}}\binom{2n}{n}\left(\frac{2N+1}{2}\right), \quad 1\leq n<2N+1. \quad \text{(A6)}$$

Finally, noting (A3) and (A6) and using the fact that the Padé approximant is unique, we have established (5.10) and (5.11).

In connection with the results given in (5.10) and (5.11), it is well to mention that exact expressions for the $[N, N]$ Padé approximants have been obtained for a certain class of hypergeometric functions,[14] and that $(1-z)^{-\frac{1}{2}}=F(1,\frac{1}{2};1;z)$ is a member of that class. The results obtained in Ref. 14 are very beautiful. However, the partial-fraction form of the Padé approximant, which for the special function $(1-z)^{-\frac{1}{2}}$ is so simple, is not discussed in that work.

The method of proof given in this Appendix has been selected because it is straightforward and requires no preliminary mathematical development. However, if one recognizes the connection between the generalized Padé approximant and the method of Gaussian quadrature, then (5.12) and (5.13) follow in a particularly simple way and this development is perhaps to be preferred over the one given in the text. We begin by noting the integral representation of the energy per particle of the infinite chain,

$$\lim_{N\to\infty}\frac{E_0(N,\lambda)}{N}=\frac{\hbar}{2}\left(\frac{2k}{M}\right)^{\frac{1}{2}}\frac{1}{\pi}\int_{-1}^{+1}(1-u\lambda)^{\frac{1}{2}}\frac{du}{(1-u^2)^{\frac{1}{2}}}. \quad \text{(A7)}$$

Now let us approximate the integral in (A7) by the $2N$-point Gauss–Chebyshev quadrature formula. Since the Gaussian quadrature formula is exact if the integrand is any polynomial of degree $4N-1$ or less, its application to (A7) must agree with the exact power series through $O(\lambda^{2(2N-1)})$ and the result is thus equivalent to $F_{N,-1}(\lambda)$. In a similar fashion, we can conclude that the $(2N+1)$-point Gauss–Chebyshev formula applied to (A7) is equivalent to $F_{N,0}(\lambda)$. It is well known that, for M-point Gauss–Chebyshev quadrature, the weights are all equal to π/M and the abscissas are the zeros of $T_M(x)$, namely,

$$\cos\left((2n-1)\pi/2M\right).$$

Thus we obtain (5.12) and (5.13).

[14] Y. L. Luke J. Math. & Phys. 37, 110 (1958).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 12 DECEMBER 1969

Physical Regions of Six-Particle Processes

RICHARD P. MCNEIL* AND RICHARD A. MORROW
Wilder Laboratory, Dartmouth College, Hanover, New Hampshire

(Received 28 August 1968, Revised Manuscript Received 8 May 1969)

The physical regions of six-particle processes are constructed in all planes of pairs of Lorentz-invariant variables. As a matter of course, the permissible ranges of the eight independent variables are established. Thus, one application is the determination of the integration limits in phase-space integrals that occur in calculations involving two-to-four and one-to-five particle processes.

I. INTRODUCTION

A previous analysis[1] of the physical regions of general five-particle relativistic processes is extended here to general six-particle reactions. Specifically, construction of the physical regions for any six-particle reaction in all possible topologically distinct planes of pairs of Lorentz-invariant kinematic variables (namely the Mandelstam variables) is shown. This knowledge could be useful in the analysis in such planes of the many-sheeted singularity structure of perturbation-theory amplitudes, since only under special circumstances can singularities lie in the physical regions.[2] Also, such knowledge is useful in establishing the domain of validity of single-variable dispersion relations and partial-wave dispersion relations in perturbation theory.[2] Furthermore, recent work in formal S-matrix theory concerning the nature of singularities following from unitarity makes a special point of beginning first with physical-region singularities[2] and so it is desirable to know the extent of these regions.

Another reason for making this study, from which the limits on the ranges of all kinematic variables can be obtained in a systematic manner, is that the result enables one to develop efficient numerical-integration techniques for use in calculating, for example, two-to-four particle transition probabilities and one-to-five particle decay probabilities. As applied to the latter example, the results below encompass those of Nyborg.[3]

Finally, the techniques developed below can be regarded as an explicit example of and introduction to the more abstract n-particle formulation.[4]

It should be pointed out that only conditions on the invariants, in order that a general six-particle reaction be physical, are of interest here. The actual ranges of the invariants for any particular process, i.e., any particular channel, easily follow, however, and, when used in conjunction with the expressions for the phase-space volume element (in invariant variables, given by Byers and Yang[5]), a complete and concise description of the phase-space integral results.

In the next section, the Gram determinantal conditions which serve to define the physical regions are summarized and the manner of implementing them is outlined. Various determinantal identities and symbolic operations are then discussed in Sec. III and are applied in Sec. IV to the Gram determinantal conditions to obtain closed sets of determinantal conditions from which the allowed ranges of all invariants may be deduced.

All quantities throughout are real.

II. DEFINITION OF PHYSICAL REGIONS

The general six-particle process to be discussed is shown in Fig. 1 where the ith particle of mass m_i has its 4-momentum p_i ($p_i^2 = m_i^2$) directed inwards for convenience and the convention used is that if the particle is incoming (outgoing), its 4-momentum is positive (negative) timelike. A convenient, symmetrically chosen set[6] of Lorentz-invariant variables with which to decribe the kinematics is the following:

$$\begin{aligned} r &= (p_6 + p_1)^2, \quad u = (p_3 + p_4)^2, \\ & \qquad\qquad a = (p_1 + p_2 + p_3)^2, \\ s &= (p_1 + p_2)^2, \quad v = (p_4 + p_5)^2, \\ & \qquad\qquad b = (p_2 + p_3 + p_4)^2, \quad (1) \\ t &= (p_2 + p_3)^2, \quad w = (p_5 + p_6)^2, \\ & \qquad\qquad c = (p_3 + p_4 + p_5)^2. \end{aligned}$$

* National Science Foundation Undergraduate Research Participant supported by NSF grant GY-4373 to Dartmouth College, Hanover, N.H.

[1] R. A. Morrow, J. Math. Phys. 7, 844 (1966), hereafter referred to as Paper I.

[2] R. J. Eden, P. V. Landshoff, P. I. Olive, and J. C. Polkinghorne, *The Analytic S-Matrix* (Cambridge University Press, London, 1966). See particularly p. 84ff and p. 204ff.

[3] P. Nyborg, Phys. Rev. **140**, B921 (1965).

[4] R. A. Morrow, Ann. Phys. (N.Y.) (to be published).

[5] N. Byers and C. N. Yang, Rev. Mod. Phys. **36**, 595 (1964).

[6] V. E. Asribekov, Nucl. Phys. **34**, 461 (1962); Phys. Letters **2**, 284 (1962).

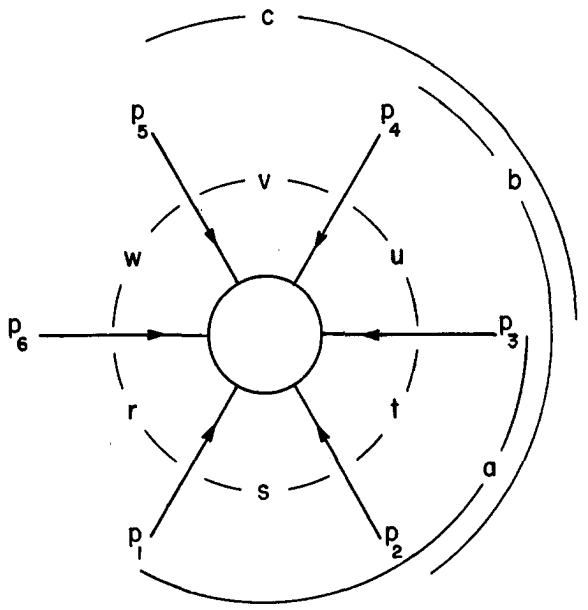


FIG. 1. This depicts a general six-particle process and indicates schematically the nine Mandelstam variables used in the analysis.

Since there are, however, only eight independent variables for a six-particle reaction, these nine variables are not all independent. It is convenient not to specify a dependent variable at this stage, however, since a choice will enter naturally later on.

Because any eight of these variables have interdependent ranges, they cannot all be treated on an equal basis. Rather, the physical regions in planes of pairs of them will be set up with the remaining variables treated as parameters. Conditions on these latter parameters are then sought so that physical regions do exist in the plane of interest. A brief survey of Fig. 1 shows that any plane is topologically equivalent to one of the following six: a–c, a–t, c–t, v–s, v–r, and s–r. Therefore only these six will be treated.

The general Gram determinantal conditions from which the physical regions may be found were established in Paper I. For the particular case of six-particle processes, they are, with the Gram determinant of the 4-vectors $p_i, p_j, \cdots, p_k$ defined as

$$\Delta(p_i, p_j, \cdots, p_k) \equiv \det(p_a \cdot p_b),$$

with $a, b = i, j, \cdots, k$:

$$\text{(i)} \qquad \Delta_2(p_i, p_j) \leq 0, \tag{2a}$$

$$\text{(ii)} \qquad \Delta_3(p_i, p_j, p_k) \geq 0, \tag{2b}$$

$$\text{(iii)} \qquad \Delta_4(p_i, p_j, p_k, p_l) \leq 0, \tag{2c}$$

$$\text{(iv)} \qquad \Delta_5(p_i, p_j, p_k, p_l, p_m) = 0, \tag{2d}$$

subject to

$$\sum_{i=1}^{6} p_i = 0.$$

As in Paper I, conservation of 4-momentum will be imposed on $\Delta_5 = 0$, which can then be put in terms of the nine Mandelstam invariants. Specifically, Δ_5 may be brought into the form $L/2^5$, where L is the symmetric determinant

$$L \equiv \begin{vmatrix} 2w & w+b-m_1^2 & w+u-s & w-a+m_4^2 & -w-m_5^2+m_6^2 \\ \cdots & 2b & u+b-m_2^2 & b-t+m_4^2 & r-b-m_5^2 \\ \cdots & \cdots & 2u & u-m_3^2+m_4^2 & c-u-m_5^2 \\ \cdots & \cdots & \cdots & 2m_4^2 & v-m_4^2-m_5^2 \\ \cdots & \cdots & \cdots & \cdots & 2m_5^2 \end{vmatrix}. \tag{3}$$

Note that the six variables of interest r, s, t, v, a, and c each occur in only one element of L (and, of course, in the transposed element). As will be seen, this allows an easy solution of the equation $L = 0$ for any of them in terms of the other eight. Thus numerous choices of the dependent variable can be made.

The over-all aim is now to solve $L = 0$ subject to conditions (2) on Δ_2, Δ_3, and Δ_4. These Gram determinantal conditions are necessary and sufficient for the invariants to lie in a physical region. However, given two invariants there is no guarantee that a physical region exists in their plane if (2) is satisfied. For this, further conditions on the remaining six independent variables, logically derivable from those on the Gram determinants in (2), are necessary. It remains to find these logical structures.

III. IDENTITIES AND SYMBOLIC OPERATIONS

In order to find the implications of (2), it will prove convenient to first develop some logical operations (almost identities) which will be applied to (2) in the next section. To begin, the notation used in this and

the following section is that of Paper I: L is the determinant of a symmetric matrix whose elements a_{ij} are defined by (3); $L_{ij\cdots k}$ is the principal minor of L obtained by deleting the ith, jth, $\cdots$, kth rows and columns of L; $V(ij\cdots k)_{mn}$ is the cofactor (signed minor) of the element a_{mn} of $L_{ij\cdots k}$; $V(ij\cdots k)_{mn_0}$ is $V(ij\cdots k)_{mn}$ with $a_{nm}=0$.

Numerous formulas, useful in dealing with symmetric determinants, are summarized in the appendix of Paper I. The three that will be used in the present work are

(a) the identity

$$L_{i\cdots j}=[L_{i\cdots jk}L_{i\cdots jm}-V(i\cdots j)^2_{km}]/L_{i\cdots jkm}; \quad (4)$$

(b) the solution of $L_{i\cdots j}=0$ for a_{km}:

$$a_{km}=[V(i\cdots j)_{km_0}\pm(L_{i\cdots jk}L_{i\cdots jm})^{\frac{1}{2}}]/L_{i\cdots jkm}$$
$$\equiv a_{km}(i\cdots j,\pm); \quad (5)$$

(c) a concise expression for $L_{i\cdots j}$:

$$L_{i\cdots j}=-L_{i\cdots jkm}[a_{km}-a_{km}(i\cdots j,+)]$$
$$\times[a_{km}-a_{km}(i\cdots j,-)]. \quad (6)$$

From these relations and from the structure of L there follow a number of useful properties:

(A) $L=0$ and $L_i\leq 0$ imply $L_j\leq 0$. This results from a straightforward application of Eq. (4).

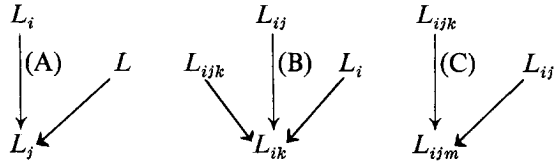
(B) $L_i\leq 0$, $L_{ij}\geq 0$, and $L_{ijk}\leq 0$ imply $L_{ik}\geq 0$. This also results from applying Eq. (4).

(C) $L_{ij}\geq 0$ and $L_{ijk}\leq 0$ imply $L_{ijm}\leq 0$.[7] The argument here is a bit more involved. If $L_{ijkm}>0$ then an application of (4) shows that $L_{ijm}\leq 0$. On the other hand if $L_{ijkm}\leq 0$, then the functional form of L_{ijm} requires it to be negative. To see this, notice that each $L_{ijm}=-\lambda(x,y,z)$, where

$$\lambda(x,y,z)=x^2+y^2+z^2-2xy-2yz-2zx, \quad (7)$$

and each of x, y, and z is either a variable or the square of a mass. Then, L_{ijkm} is either $2x$, $2y$, or $2z$, depending on the index m. From Eq. (7) it follows that $\lambda(x,y,z)\geq 0$ if x, y, and z are neither all positive nor all negative. Now, by inspection of (3), at least one of x, y, and z is the square of a mass [except in the case $L_{245}=-\lambda(s,u,w)$, which will not enter the present analysis] while, because of $L_{ijkm}\leq 0$, another is nonpositive. Hence $\lambda(x,y,z)\geq 0$ and so $L_{ijm}\leq 0$ as asserted.

These properties may be symbolized as logical operations or implications:



(8)

where $L=0$, $L_i\leq 0$, $L_{ij}\geq 0$, and $L_{ijk}\leq 0$. Determinants at the tails of arrows are *assumed* to have these signs, whence the identities *require* the determinants at the heads of arrows to have these signs, as shown above. The manner in which this requirement on any particular determinant is actually met is by varying one of the invariants that the determinant depends on. To see how this operates, suppose $L_{i\cdots j}\leq 0$ is wanted. Consider the solutions of $L_{i\cdots j}=0$ for a_{km} as given in Eq. (5). It will always turn out in the present work, by proper choice of the indices k and m, that $L_{i\cdots jk}L_{i\cdots jm}\geq 0$. Thus $L_{i\cdots j}=0$ has real solutions $a_{km}(i\cdots j,\pm)$ and there is, therefore, a range of real values of a_{km} for which $L_{i\cdots j}\leq 0$, found by using (5). Concrete examples of these arguments are given in the next section.

IV. LOGICAL STRUCTURES OF DETERMINANTAL CONDITIONS

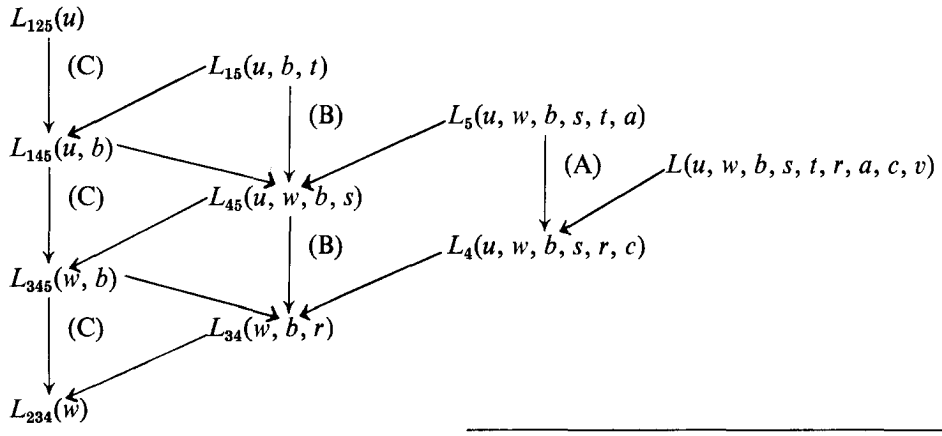
The arguments of Sec. II may now be continued using the symbolic operations set up in the previous section. The aim is to build upon a set of Gram determinantal conditions (2), using the operations in (8), a structure of other determinantal conditions from which the ranges of all invariants in (1) may be found in a systematic manner. It turns out that, although many such structures are possible, two particular ones are sufficient for a study of the physical regions in all of the six planes proposed in Sec. II. Only these two structures will be shown.

The a–c, a–t, and c–t Planes

Take as the set of Gram determinants $\Delta_2(p_3,p_4)$, $\Delta_3(p_2,p_3,p_4)$, and $\Delta_4(p_1,p_2,p_3,p_4)$. By adding appropriate rows and columns and by using the conservation-of-momentum condition in (2), these three Gram determinants become $L_{125}/2^2$, $L_{15}/2^3$, and $L_5/2^4$, respectively. Thus conditions (2) become $L_{125}\leq 0$, $L_{15}\geq 0$, $L_5\leq 0$, and $L=0$. Beginning with these and using the operations (8), the following structure

[7] An exception to this—which will not be encountered in the present work, however—is pointed out below.

can be obtained:



(9)

with $L = 0$, $L_i \leq 0$, $L_{ij} \geq 0$, and $L_{ijk} \leq 0$. By construction, the conditions in (9) are necessary and sufficient for a set of values of the kinematic variables (1) to describe a physical process. How they determine the physical regions in the various planes will now be described.

Physical Regions in the a–c Plane

The conditions (9) may be satisfied as follows:

(i) pick u so that $L_{125} \leq 0$;

(ii) pick w and b so that L_{145}, L_{345}, $L_{234} \leq 0$;

(iii) pick t, s, and r so that $L_{15} \geq 0$, $L_{45} \geq 0$, and $L_{34} \geq 0$, respectively;

(iv) pick a and c so that $L_5 \leq 0$ and $L_4 \leq 0$, respectively;

(v) pick v so that $L = 0$.

In the a–c plane, the physical regions exist and are given by those regions where L_5, $L_4 \leq 0$. Use of (6) shows that there is only one such region and that its shape is rectangular.

That the above choices can always be made was argued at the end of Sec. III. For example, consider satisfying $L_{45} \geq 0$ by choice of s. From (3), s occurs only in a_{13} and (5) shows that $L_{45} = 0$ can be solved for a real value of a_{13}, since $L_{145}L_{345} \geq 0$ by prior choice of u, w, and b. Thus real values of s exist that make $L_{45} \geq 0$, the range being found by use of (6).

Essentially the same argument can be applied to the choice of each invariant. Note in particular, from (5), that *two* values of the dependent variable v satisfy $L = 0$. For any particular channel (selected by prior choice of the eight independent variables), both values of v are physical. In other words, specification of the eight independent invariants does not lead to a unique configuration of momentum in Minkowski space. In fact, two additional pieces of information[8] are needed before the configuration is unique; one is the value of v, while the other is not really important, since it just distinguishes between configurations related by a spatial reflection. These remarks also hold in the following cases.

Physical Regions in the a–t Plane

Conditions (9) can be satisfied precisely as above. To then find the form of the physical regions in the a–t plane, consider using (5) to solve $L_5 = 0$ for $a_{14}(a)$ and $a_{24}(t)$:

$$a_{14} = [V(5)_{14_0} \pm (L_{15}L_{45})^{\frac{1}{2}}]/L_{145},$$
$$a_{24} = [V(5)_{24_0} \pm (L_{25}L_{45})^{\frac{1}{2}}]/L_{245}.$$

This shows that $L_{15} = 0$ and $L_{25} = 0$ are tangent lines to $L_5 = 0$, parallel to the coordinate axes. Furthermore, since $L_{45} \geq 0$, the curve $L_5 = 0$ lies only in those regions where L_{15}, $L_{25} \geq 0$. Use of (6) on L_{15} shows that if $L_{1245} \equiv 2u \geq 0$, then $L_{15} \geq 0$ for t between the tangent lines $L_{15} = 0$; otherwise t must be taken outside these lines. The two possibilities are shown in Fig. 2. Whether the hyperbola lies in the first and third quadrants as shown or in the second and fourth quadrants depends on the values of the other invariants.

Physical Regions in the c–t Plane

The order of choosing invariants to satisfy (9) is for the most part the same as before. The exception is that a must be chosen before t or c. This is easily handled, for from Fig. 2 it can be noted that if a is chosen so that $L_{25}(u, w, s, a) \geq 0$ then any t such that $L_5 \leq 0$ also gives $L_{15} \geq 0$. Thus conditions (9) can be satisfied by the scheme:

(i) pick u so that $L_{125} \leq 0$;

(ii) pick w and b so that L_{145}, L_{345}, $L_{234} \leq 0$;

(iii) pick s and r so that L_{45}, $L_{34} \geq 0$, respectively;

(iv) pick a so that $L_{25} \geq 0$;

(v) pick t and c so that L_5, $L_4 \leq 0$, respectively;

(vi) pick v so that $L = 0$.

[8] F. Rohrlich, Nucl. Phys. **67**, 659 (1965); Nuovo Cimento **38**, 673 (1965).

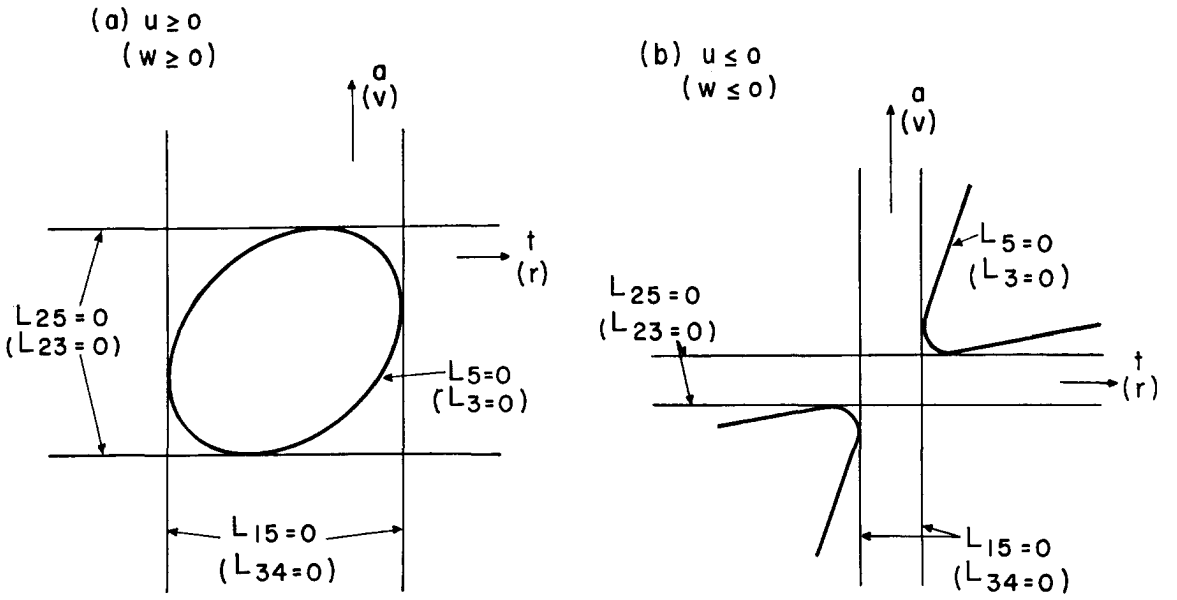
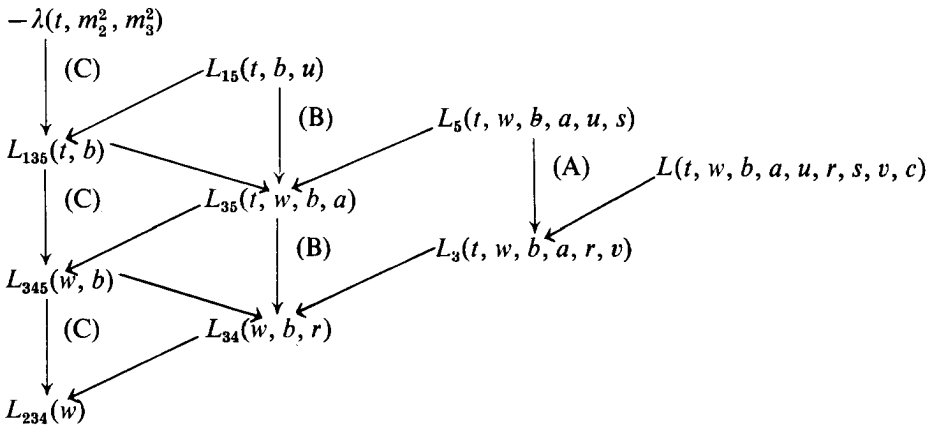


FIG. 2. The form of the physical regions in the a–t (v–r) plane for (a) $u(w) \geq 0$ and (b) $u(w) < 0$. In (a) there is only one physical region, the interior of the ellipse. In (b) there are two physical regions, inside each branch of the hyperbola.

As in the case of the a–c plane there is only one physical region in the c–t plane given by the overlap of $L_5 \leq 0$ and $L_4 \leq 0$. Its shape is also rectangular.

The v–s, v–r, and s–r Planes

Take as the set of Gram determinants $\Delta_2(p_2, p_3)$, $\Delta_3(p_2, p_3, p_4)$, and $\Delta_4(p_1, p_2, p_3, p_4)$, the only difference with the previous choice being Δ_2, which here may be written in the form $-\frac{1}{4}\lambda(t, m_2^2, m_3^2)$. Thus conditions (2) become $-\lambda(t, m_2^2, m_3^2) \leq 0$, $L_{15} \geq 0$, $L_5 \leq 0$, and $L = 0$. The operations (8) can then be used to build the following structure:



(10)

with $L = 0$, $L_i \leq 0$, $L_{ij} \geq 0$, $L_{ijk} \leq 0$, and

$$\lambda(t, m_2^2, m_3^2) \geq 0.$$

Although $-\lambda(t, m_2^2, m_3^2)$ is not one of the L_{ijk}, operation (C) of (8) can still be used to require $L_{135} \leq 0$, because L_{15} can be put in the (symmetric) form

$$L_{15} = \begin{vmatrix} 2t & t - m_2^2 + m_3^2 & b - t - m_4^2 \\ \cdots & 2m_3^2 & u - m_3^2 - m_4^2 \\ \cdots & \cdots & 2m_4^2 \end{vmatrix},$$

and application of (4) gives

$$L_{15} = [-\lambda(t, m_2^2, m_3^2)L_{135} - (t - m_2^2 + m_3^2)^2(b - t - m_4^2)^2]/2t$$

from which it follows, by repeating the arguments used to establish operation (C), that

$$L_{15} \geq 0 \quad \text{and} \quad -\lambda(t, m_2^2, m_3^2) \leq 0$$

imply $L_{135} \leq 0$.

Physical Regions in the v–s Plane

This case is quite similar to that of the a–c plane. The conditions (10) may be satisfied as follows:

(i) pick t so that $\lambda(t, m_2^2, m_3^2) \geq 0$;
(ii) pick w and b so that $L_{135}, L_{345}, L_{234} \leq 0$;
(iii) pick u, a, and r so that $L_{15}, L_{35}, L_{34} \geq 0$, respectively;
(iv) pick s and v so that $L_5, L_3 \leq 0$, respectively;
(v) pick c so that $L = 0$.

The single, rectangular-shaped, physical region in the s–v plane is given by the overlap of $L_5 \leq 0$ and $L_3 \leq 0$.

Physical Regions in the v–r Plane

The conditions (10) may be satisfied precisely as in the previous case. The situation here is quite analogous to that in the a–t plane case and so to find the form of the physical regions first use (5) to solve $L_3 = 0$ for $a_{45}(v)$ and $a_{25}(r)$:

$$a_{45} = [V(3)_{45_0} \pm (L_{34}L_{35})^{\frac{1}{2}}]/L_{345},$$
$$a_{25} = [V(3)_{25_0} \pm (L_{23}L_{35})^{\frac{1}{2}}]/L_{235}.$$

This shows that the lines $L_{34} = 0$ and $L_{23} = 0$ are tangent to $L_3 = 0$ and that the latter curve exists only in regions in the v–r plane where $L_{34}, L_{23} \geq 0$ since $L_{35} \geq 0$. Use of (6) on L_{34} then shows that if $L_{2345} \equiv 2w \geq 0$, then $L_{34} \geq 0$ for r between the tangent lines $L_{34} = 0$; otherwise r must be taken outside these lines. Thus Fig. 2 applies to this case with the replacements $a \to v$, $t \to r$, $u \to w$, $L_{25} \to L_{23}$, $L_{15} \to L_{34}$, and $L_5 \to L_3$.

Physical Regions in the s–r Plane

This case is analogous to that in the c–t plane. The order of choosing the variables can be the same as in the previous two cases with the exception that v must be chosen before r or s. This is easily done, for, from Fig. 2, if v is chosen so that $L_{23}(w, a, v) \geq 0$, then any r that makes $L_3 \leq 0$ will give $L_{34} \geq 0$. Conditions (10) are thus satisfied by the following:

(i) pick t so that $\lambda(t, m_2^2, m_3^2) \geq 0$;
(ii) pick w and b so that $L_{135}, L_{345}, L_{234} \leq 0$;
(iii) pick u and a so that $L_{15}, L_{35} \geq 0$, respectively;
(iv) pick v so that $L_{23} \geq 0$;
(v) pick r and s so that $L_5, L_3 \leq 0$, respectively;
(vi) pick c so that $L = 0$.

Since $L_5 = 0$ and $L_3 = 0$ are each two straight lines in the s–r plane, there is only one physical region and it is rectangular, being given by the overlap $L_5, L_3 \leq 0$.

V. CONCLUSIONS

It has been shown how to systematically construct the physical regions of six-particle processes in all topologically distinct planes of pairs of Lorentz-invariant variables. Instructions have been given as to how to select the ranges of the first six variables which are viewed as parameters.

There was no need in this analysis to specify which particles were incoming and which were outgoing. This information is imposed when the ranges of the variables are chosen. That is, many of the variables have more than one nonoverlapping range, each range corresponding to a different process. Thus all 25 possible reactions[9] (plus any decays) are included in the formulation.

As an important application, the ranges of the variables as found here can be taken as the integration ranges in the phase-space integral set up by Byers and Yang.[5] The resulting expression, which is straightforward to obtain, is particularly well suited to peripheral reactions.[10]

[9] For an n-particle reaction there are $2^{n-1} - n - 1$ (if decays are forbidden) disjoint connected physical regions, each corresponding to a particular channel; see D. A. Jacobson, Nuovo Cimento **45A**, 905 (1966).

[10] A Monte Carlo numerical-integration program based on this scheme has been used successfully in calculations involving the multi-Regge model. Its great usefulness rests on the efficiency with which events at small values of the momentum-transfer variables can be generated, a prime reason for this being the fact that such variables occur as integration variables.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 12 DECEMBER 1969

Solution of a Three-Body Problem in One Dimension

F. CALOGERO*
Physics Department, Imperial College, London SW7, England

(Received 27 January 1969)

The problem of three equal particles interacting pairwise by inversecube forces ("centrifugal potential") in addition to linear forces ("harmonical potential") is solved in one dimension.

1. INTRODUCTION

It has been known for some time that the one-dimensional three-body problem with linear ("harmonical") and inverse-cube ("centrifugal") pair forces is separable,[1,2] but apparently there has been no attempt at its actual solution. In this paper this problem is solved in the case of equal particles: The complete energy spectrum is determined, and all the corresponding eigenfunctions are explicitly written out.

The particles may satisfy Boltzmann, Bose, or Fermi statistics; in fact, the nature of the problem is such that the type of statistics does not modify the energy spectrum and affects the wavefunctions only in a trivial way. The problem which obtains from that described above eliminating the inverse-cube force between two pairs (so that it acts only between one pair) is also solved.

In Sec. 2 we discuss the two-body problem with the same "oscillator plus centrifugal" forces. This treatment is useful both as a preliminary for the solution and as a model for the interpretation of the three-body problem, which is discussed in Sec. 3. The last section contains comments on possible extensions of the results of this paper.

Units are chosen so that $2m\hbar^{-2} = 1$, where m is the mass of the particles.

2. THE TWO-BODY PROBLEM

The Schrödinger equation for the two-body problem under consideration is

$$\left[-\frac{\partial^2}{\partial x_1^2} - \frac{\partial^2}{\partial x_2^2} + \tfrac{1}{8}\omega^2(x_1 - x_2)^2 + g(x_1 - x_2)^{-2}\right]\psi = E\psi. \quad (2.1)$$

Here x_1 and x_2 indicate, of course, the coordinates of the two particles, and we assume that $g > -\frac{1}{2}$ to avoid "fall to the center."[3] Going over to the center-of-mass (CM) and relative coordinates,[4]

$$R^{(2)} = \tfrac{1}{2}(x_1 + x_2), \quad (2.2a)$$

$$x = 2^{-\frac{1}{2}}(x_1 - x_2), \quad (2.2b)$$

and eliminating the center-of-mass motion, we get

$$\left[-\frac{d^2}{dx^2} + \tfrac{1}{4}\omega^2 x^2 + \tfrac{1}{2}gx^{-2} - E\right]\psi = 0, \quad (2.3)$$

where now E is the energy in the CM system.

The physically acceptable solutions of this equation (in the interval $0 \leq x < \infty$) are

$$\psi_n(x) = x^{a+\frac{1}{2}} \exp(-\tfrac{1}{4}\omega x^2) L_n^a(\tfrac{1}{2}\omega x^2), \quad n = 0, 1, 2, \cdots, \quad (2.4)$$

with

$$a = \tfrac{1}{2}(1 + 2g)^{\frac{1}{2}}. \quad (2.5)$$

Here L_n^a is a generalized Laguerre polynomial, defined as in Ref. 5. By changing the sign of a, namely, by taking the negative determination of the square root in Eq. (2.5), one would still obtain a solution of the Schrödinger equation, but it would not be an acceptable one owing to its behavior at $x = 0$.[6]

The corresponding energy levels are

$$E_n = \omega(2n + a + 1), \quad n = 0, 1, 2, \cdots. \quad (2.6)$$

Because, for $g \neq 0$, both $\psi_n(x)$ and $\psi_n(x)\psi_n'(x)$ vanish at $x = 0$, the physically acceptable solutions in the whole interval $-\infty < x < \infty$ are obtained by supplementing Eq. (2.4) with the simple prescription[6]

$$\psi(-x) = \pm\psi(x), \quad x \geq 0. \quad (2.7)$$

The upper sign corresponds to Bose statistics, the lower sign to Fermi statistics. Obviously the energy spectrum is no affectetd by this prescription. This

* Permanent address: Physics Department, Rome University, Rome, Italy.

[1] H. R. Post, Proc. Phys. Soc. (London) **A69**, 936 (1956).

[2] J. Hurley, J. Math. Phys. **8**, 813 (1967).

[3] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics* (Pergamon Press, Inc., New York, 1958), Sec. 35.

[4] The factor $2^{-\frac{1}{2}}$ in the definition of x is convenient for the comparison with the three-body problem.

[5] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series and Products* (Academic Press Inc., 1965); *Higher Transcendental Functions*, A. Erdélyi, Ed. (McGraw-Hill Book Co., Inc., 1953), Vol. II. Note that the definition of Laguerre polynomials given here is different from that used in some textbooks, for instance in Ref. 3.

[6] A wavefunction $\psi(x)$ is considered physically acceptable if both $|\psi(x)|^2$ and $\psi(x)\psi'(x)$ are continuous. This condition may be interpreted as deriving from the requirement that both the density and the current of probability (that the particle be found at x) vary continuously with x. Moreover, the wavefunction must be normalizable (for closed problems).

happens because the restriction to one dimension together with the singular nature of the centrifugal interaction at short distance implies a complete separation of the configurations with $x > 0$ from those with $x < 0$. Of course if the particles are assumed to satisfy Bose or Fermi statistics, the energy levels (2.6) are nondegenerate, while if they satisfy Boltzmann statistics, each level is two-fold degenerate.[7]

In the $g \to 0$ limit, the eigenfunctions (2.4) and eigenvalues (2.6) go over into the "antisymmetrical" eigenfunctions and eigenvalues of the harmonic-oscillator problem:

$$\psi_n(x) = \exp(-\tfrac{1}{4}\omega x^2) H_{2n+1}[(\tfrac{1}{2}\omega)^{\frac{1}{2}} x], \quad n = 0, 1, 2 \cdots, \tag{2.8}$$

$$E_n = \omega(2n + \tfrac{3}{2}), \quad n = 0, 1, 2 \cdots. \tag{2.9}$$

Here of course H_m is a Hermite polynomial.[5] This happens quite independently of the type of statistics that the particles satisfy, although, of course, to extend these solutions to the whole interval $-\infty < x < \infty$, one should now use the negative sign in Eq. (2.7). The even solutions of the oscillator problem, that do not vanish at $x = 0$, would instead result, in the $g \to 0$ limit, from the functions (2.4) with the negative value of a corresponding to the negative determination of the square root in Eq. (2.5); that is, they result in inserting $a = -\frac{1}{2}$ in Eqs. (2.4) and (2.6) and taking of course the positive sign in Eq. (2.7). But they cannot be obtained as the limit of the eigenfunctions of the problem with $g \neq 0$. This is consistent with the fact that, owing to the singular nature of the interaction, for $g \neq 0$ all eigenfunctions vanish at $x = 0$.

In conclusion we may assert that switching on the "centrifugal" interaction shifts all the "odd" eigenvalues of the harmonic oscillator problem by the constant amount

$$\omega(a - \tfrac{1}{2}) = \tfrac{1}{2}\omega[(1 + 2g)^{\frac{1}{2}} - 1], \tag{2.10}$$

while it eliminates altogether the "even" eigenvalues. This is also consistent with the indication of first-order perturbation theory, which yields a finite answer when applied to an "odd" oscillator state (whose wavefunction, vanishing at the origin, compensates the divergence of the centrifugal potential), but yields a divergent answer if applied to an "even" oscillator state. Of course, the energy shift is positive for repulsive interaction ($g > 0$), negative for attractive interaction ($-\frac{1}{2} < g < 0$).

On the other hand, for $\omega = 0$, the solution of Eq. (2.3) is

$$\psi(x) = x^{\frac{1}{2}} J_a(kx) \tag{2.11}$$

and the corresponding eigenvalues belong to the continuous spectrum

$$E = k^2, \quad 0 \leq k < \infty. \tag{2.12}$$

In Eq. (2.11), J_a is a Bessel function[5] and a is still defined by Eq. (2.5). Note, however, that in the $\omega \to 0$ limit, all the solutions (2.4) go over into the zero-energy solution

$$\psi(x) = Cx^{a+\frac{1}{2}}. \tag{2.13}$$

This case with $\omega = 0$ is of no interest because there is no energy quantization and, on the other hand, the nature of the problem implies that the particles cannot overtake one another [see the discussion after Eq. (2.7) above]. In fact, it is immediately seen that the transmission coefficient, evaluated from Eqs. (2.11) and (2.7), vanishes identically for all values of the energy.

3. THE THREE-BODY PROBLEM

The Schrödinger equation for the three-body problem under consideration is

$$\Big\{-\frac{\partial^2}{\partial x_1^2} - \frac{\partial^2}{\partial x_2^2} - \frac{\partial^2}{\partial x_3^2} + \tfrac{1}{8}\omega^2[(x_1 - x_2)^2 + (x_2 - x_3)^2 + (x_3 - x_1)^2] + g_3(x_1 - x_2)^{-2} + g_1(x_2 - x_3)^{-2} + g_2(x_3 - x_1)^{-2}\Big\}\psi = E\psi. \tag{3.1}$$

We will solve the two cases

$$g_1 = g_2 = 0, \quad g_3 = g, \tag{3.2}$$

and

$$g_1 = g_2 = g_3 = g. \tag{3.3}$$

As in the two-body case, and for the same reason, we assume throughout that $g > -\frac{1}{2}$.

It is first of all convenient to go over to the center-of-mass and "Jacobi" coordinates

$$R^{(3)} = \tfrac{1}{3}(x_1 + x_2 + x_3), \quad x = 2^{-\frac{1}{2}}(x_1 - x_2), \quad y = 6^{-\frac{1}{2}}(x_1 + x_2 - 2x_3). \tag{3.4}$$

In these coordinates, the Schrödinger equation becomes, after elimination of the center-of-mass

[7] In the case of equal but distinguishable particles (Boltzmann statistics), each independent eigenfunction may be chosen to correspond to a definite ordering of the particles, because the singularity of the "centrifugal" pair interaction excludes the possibility that the particles overtake one another. Such an eigenfunction vanishes unless the variable which distinguishes the different orderings of the particles (x in the two-body case, ϕ in the three-body case) lies within the appropriate range.

motion,

$$\left[-\frac{\partial^2}{\partial x^2}-\frac{\partial^2}{\partial y^2}+\tfrac{3}{8}\omega^2(x^2+y^2)+\tfrac{1}{2}g_3x^{-2}+\tfrac{1}{2}g_1[(3)^{\frac{1}{2}}y-x]^{-2}+\tfrac{1}{2}g_2[(3)^{\frac{1}{2}}y+x]^{-2}-E\right]\psi=0, \tag{3.5}$$

where E now indicates the energy in the center-of-mass system.

We introduce next the two-dimensional "spherical" coordinates r and ϕ:

$$\begin{aligned} r^2 &= x^2+y^2=(x_1-R^{(3)})^2+(x_2-R^{(3)})^2+(x_3-R^{(3)})^2 \\ &= \tfrac{1}{3}[(x_1-x_2)^2+(x_2-x_3)^2+(x_3-x_1)^2], \\ \phi &= \tan^{-1}(x/y) \\ &= \tan^{-1}[\sqrt{3}(x_1-x_2)/(x_1+x_2-2x_3)], \end{aligned} \tag{3.6a}$$

$$x=r\sin\phi, \qquad y=r\cos\phi. \tag{3.6b}$$

The range of these variables is

$$0\le r<\infty \tag{3.7a}$$

and

$$0\le\phi<2\pi. \tag{3.7b}$$

Moreover, it is easily seen that the value of ϕ corresponds to the ordering of the three particles as follows:

$$\begin{aligned} &0<\phi<\tfrac{1}{3}\pi, && x_1>x_2>x_3, \\ &\tfrac{1}{3}\pi<\phi<\tfrac{2}{3}\pi, && x_1>x_3>x_2, \\ &\tfrac{2}{3}\pi<\phi<\pi, && x_3>x_1>x_2, \\ &\pi<\phi<\tfrac{4}{3}\pi, && x_3>x_2>x_1, \\ &\tfrac{4}{3}\pi<\phi<\tfrac{5}{3}\pi, && x_2>x_3>x_1, \\ &\tfrac{5}{3}\pi<\phi<2\pi, && x_2>x_1>x_3. \end{aligned} \tag{3.8}$$

In fact, from Eqs. (3.4) and (3.6), we have

$$\begin{aligned} x_1-x_2 &= 2^{\frac{1}{2}}r\sin\phi, \\ x_2-x_3 &= 2^{\frac{1}{2}}r\sin(\phi+\tfrac{2}{3}\pi), \\ x_3-x_1 &= 2^{\frac{1}{2}}r\sin(\phi+\tfrac{4}{3}\pi). \end{aligned} \tag{3.9}$$

In these new variables the Schrödinger equation reads

$$\left(-\frac{\partial^2}{\partial r^2}-\frac{1}{r}\frac{\partial}{\partial r}+\tfrac{3}{8}\omega^2r^2+\frac{M}{r^2}-E\right)\psi=0, \tag{3.10}$$

with

$$M=-\frac{\partial^2}{\partial\phi^2}+\frac{1}{2}\left(\frac{g_3}{\sin^2\phi}+\frac{g_1}{\sin^2(\phi+\frac{2}{3}\pi)}+\frac{g_2}{\sin^2(\phi+\frac{4}{3}\pi)}\right). \tag{3.11}$$

It is, therefore, solvable by separation of variables as

$$\psi=R(r)f(\phi). \tag{3.12}$$

Let us denote the eigenvalues of the operator M by b_l^2, $l=0,1,2\cdots$ (they are certainly positive; see below), i.e.,

$$Mf_l(\phi)=b_l^2f_l(\phi), \quad l=0,1,2\cdots. \tag{3.13}$$

We then get the equation

$$\left(-\frac{d^2}{dr^2}-\frac{1}{r}\frac{d}{dr}+\tfrac{3}{8}\omega^2r^2+\frac{b_l^2}{r^2}-E\right)R=0 \tag{3.14}$$

for the corresponding radial wavefunctions. The normalizable solutions of this equation are

$$R_{nl}(r)=r^{b_l}\exp[-\tfrac{1}{4}(\tfrac{3}{2})^{\frac{1}{2}}\omega r^2]L_n^{b_l}[\tfrac{1}{2}(\tfrac{3}{2})^{\frac{1}{2}}\omega r^2], \quad n=0,1,2\cdots, \quad l=0,1,2\cdots, \tag{3.15}$$

where L_n^b are again the Laguerre polynomials.[5]

The corresponding energy eigenvalues are

$$E_{nl}=(\tfrac{3}{2})^{\frac{1}{2}}\omega(2n+b_l+1), \quad n=0,1,2\cdots, \quad l=0,1,2\cdots. \tag{3.16}$$

There remains now the less trivial part of the exercise, which is to determine the eigenfunctions and eigenvalues of the operator M of Eq. (3.11).

We consider first the case with $g_1=g_2=0$, $g_3=g$.[8] Then the differential equation

$$Mf_l(\phi)=-\frac{d^2}{d\phi^2}+\frac{g}{2\sin^2\phi}f_l(\phi)=b_l^2f_l(\phi) \tag{3.17}$$

may be transformed into the hypergeometric equation by appropriate substitutions, so that we obtain

$$f_l(\phi)=(\sin\phi)^{a+\frac{1}{2}}F(\tfrac{1}{2}(a+\tfrac{1}{2}-b_l),\tfrac{1}{2}(a+\tfrac{1}{2}+b_l),1+a;\sin^2\phi) \tag{3.18a}$$

$$=\cos\phi(\sin\phi)^{a+\frac{1}{2}}F(\tfrac{1}{2}(a+\tfrac{3}{2}-b_l),\tfrac{1}{2}(a+\tfrac{3}{2}+b_l);1+a;\sin^2\phi). \tag{3.18b}$$

Here a is defined by Eq. (2.5), $F(A,B;C;t)$ is the usual hypergeometric function, and Eq. (3.18b) is obtained from Eq. (3.18a) by using the well-known identity

$$F(A,B;C;t)=(1-t)^{C-A-B}F(C-B,C-A;C;t). \tag{3.19}$$

[8] This problem is separable (and easily solvable) also in the x, y variables [see Eq. (3.5)]. (In fact, any problem characterized by equal-strength harmonical potentials acting between every two particles and, in addition, by one arbitrary potential depending only upon the interparticle distance of one pair is completely separable[2]; and this statement remains true for any number of particles and spatial dimensions.) But the solution of the more complete three-body problem with $g_1=g_2=g_3=g$ can be simply obtained only using the spherical variables r, ϕ (see below).

The negative square root in Eq. (2.5) would still yield a solution of Eq. (3.17), but it would not be an acceptable one owing to its behavior at $\phi = 0$ and $\phi = \pi$ (i.e., at $x_1 = x_2$). This situation is identical to its counterpart in the two-body case [see Eq. (2.4) and the discussion following it].

The function $f_l(\phi)$ is, however, physically unacceptable[6] unless the hypergeometric functions in Eqs. (3.18a) or (3.18b) reduce to a polynomial, owing to the discontinuous behavior of its derivative at $\phi = \frac{1}{2}\pi$.[9] Thus, the only acceptable solutions (in the interval $0 \leq \phi \leq \pi$), and the corresponding eigenvalues, obtain from the requirement that the first argument of either one of the hypergeometric functions in Eqs. (3.18) coincides with a nonpositive integer. In this manner one obtains

$$f_l(\phi) = (\sin\phi)^{a+\frac{1}{2}} C_l^{a+\frac{1}{2}}(\cos\phi), \quad l = 0, 1, 2 \cdots, \tag{3.20}$$

$$b_l = l + a + \tfrac{1}{2}, \quad l = 0, 1, 2 \cdots, \tag{3.21}$$

where C_n^λ is a Gegenbauer polynomial.[5]

The extension of these solutions to the whole interval $0 \leq \phi < 2\pi$ is performed, just as in the two-body case, by the prescription

$$f_l(2\pi - \varphi) = \pm f_l(\phi), \quad 0 \leq \phi \leq \pi, \tag{3.22}$$

the positive (negative) sign corresponding to states symmetrical (antisymmetrical) under the exchange of particles 1 and 2. There is, of course, no symmetry under the exchange of particle 1 (or 2) with particle 3, because in the case under consideration particle 3 has an interaction different from that of particles 1 and 2.

We may, therefore, conclude that the (normalizable) eigenfunctions of the three-body problem (with $g_1 = g_2 = 0$, $g_3 = g$) are

$$\psi_{nl}(r, \phi) = r^{l+a+\frac{1}{2}} \exp[-\tfrac{1}{4}(\tfrac{3}{2})^{\frac{1}{2}}\omega r^2] L_n^{l+a+\frac{1}{2}}(\tfrac{1}{2}(\tfrac{3}{2})^{\frac{1}{2}}\omega r^2) \times (\sin\phi)^{a+\frac{1}{2}} C_l^{a+\frac{1}{2}}(\cos\phi), \quad 0 \leq \phi \leq \pi, \tag{3.23a}$$

$$\psi_{nl}(r, \phi + \pi) = \pm(-)^l \psi_{nl}(r, \phi), \quad 0 \leq \phi \leq \pi, \tag{3.23b}$$

and the corresponding eigenvalues are

$$E_{2n+l} = (\tfrac{3}{2})^{\frac{1}{2}}\omega(2n + l + a + \tfrac{3}{2}). \tag{3.24}$$

We recall that a is defined by Eq. (2.5) and that both quantum numbers n and l take all nonnegative integral values. If the two equal particles 1 and 2 satisfy Bose (Fermi) statistics, only the positive (negative) sign should be taken in Eq. (3.23b); if they satisfy Boltzmann statistics, both possibilities are allowed (so that the degeneracy of each eigenvalue is doubled).[7]

Equation (3.24) implies that the spectrum is linear, with the same spacing as in the three-body "oscillator" problem without any "centrifugal" potential.

It is also easily seen that, just as in the two-body case, in the limit $g \to 0$, $a \to \frac{1}{2}$, the eigenfunctions (3.23a) and the eigenvalues (3.24) go over into those eigenfunctions and eigenvalues of the three-body "oscillator" problem that correspond to states antisymmetrical under the exchange of particles 1 and 2. The eigenfunctions and eigenvalues corresponding to "oscillator" states symmetrical under the exchange of particles 1 and 2 result inserting $a = -\frac{1}{2}$ in Eqs. (3.23a) and (3.24), and therefore cannot be obtained in the $g \to 0$ limit from the eigenstates and eigenfunctions of the problem with $g \neq 0$ (compare with the analogous phenomenon in the two-body case, discussed in the preceding section). Of course, to extend the corresponding eigenfunctions to the whole interval $0 \leq \phi < 2\pi$, one must choose in Eq. (3.23b) the negative (positive) sign for antisymmetrical (symmetrical) states.

The degeneracy of each level E_N, $N = 2n + l$, is the integral part of $(N + 2)/2$ (if the two equal particles satisfy Bose or Fermi statistics; twice that if they are distinguishable, namely, if they satisfy Boltzmann statistics).[7] The shift of each energy level from the corresponding level of the oscillator problem is $(\frac{3}{2})^{\frac{1}{2}}(a - \frac{1}{2})\omega$, namely, $(\frac{3}{2})^{\frac{1}{2}}$ times that found in the two-body case.

We proceed now to the equal-particle case ($g_1 = g_2 = g_3 = g$). To do this we must solve the eigenvalue equation

$$MF_l(\phi) = \left[-\frac{d^2}{d\phi^2} + \frac{g}{2}\left(\frac{1}{\sin^2\phi} + \frac{1}{\sin^2(\phi + \frac{2}{3}\pi)} + \frac{1}{\sin^2(\phi + \frac{4}{3}\pi)}\right)\right] F_l(\phi) = B_l^2 F_l(\phi). \tag{3.25}$$

We have used capital letters for the eigenfunctions and eigenvalues of this equation, to distinguish them from those of the analogous equation (3.17).

The solution of this problem is immediately reduced to that of the previous one, Eq. (3.17), by the use of the trigonometric identity

$$(\sin\phi)^{-2} + [\sin(\phi + \tfrac{2}{3}\pi)]^{-2} + [\sin(\phi + \tfrac{4}{3}\pi)]^{-2} = 9(\sin 3\phi)^{-2}. \tag{3.26}$$

This identity,[10] which may be verified by explicit computation, was actually discovered through a

[9] At $\varphi = \frac{1}{2}\pi$ the differential equation (3.17) has no singularity, but the mapping between the variable φ and the argument of the hypergeometric function does.

[10] It is amusing to recognize that the equality

$$[\sin\varphi]^{-p} + [\sin(\varphi + \tfrac{2}{3}\pi)]^{-p} + [\sin(\varphi + \tfrac{4}{3}\pi)]^{-p} = 3^p[\sin 3\varphi]^{-p}$$

holds both for $p = 1$ and $p = 2$. It does not, however, hold for $p = 3$.

solution of the eigenvalue problem, Eq. (3.25), that had been obtained by a painstaking analysis. Such an analysis is, however, completely unnecessary once the existence of this identity is recognized, because its insertion in Eq. (3.25) and a comparison of this equation with Eq. (3.17) immediately implies that

$$F_l(\phi) = f_l(3\phi), \tag{3.27}$$

$$B_l = 3b_l. \tag{3.28}$$

These equations, together with Eqs. (3.20) and (3.21), yield "acceptable" solutions of Eq. (3.25) in the interval $0 \leq \phi \leq \pi/3$. The extension of these solutions to the whole interval $0 \leq \phi < 2\pi$ is easily achieved by the prescriptions

$$F_l(\phi + \tfrac{1}{3}p\pi) = \pm F_l(\phi),$$
$$0 \leq \phi \leq \tfrac{1}{3}\pi, \quad p = 1, 2, 3, 4, 5. \tag{3.29}$$

The justification of these prescriptions and the choice of the sign [depending on the statistics of the particles; see also Eq. (3.8)] should be obvious by now. Therefore, we proceed immediately to the final expressions for the eigenfunctions and eigenvalues of the problem:

$$\psi_{nl}(r, \phi) = r^{3l+3a+\frac{3}{2}} \exp\left[-\tfrac{1}{4}(\tfrac{3}{2})^{\frac{1}{2}}\omega r^2\right] L_n^{3l+3a+\frac{3}{2}}\left[\tfrac{1}{2}(\tfrac{3}{2})^{\frac{1}{2}}\omega r^2\right] \times (\sin 3\phi)^{a+\frac{1}{2}} C_l^{a+\frac{1}{2}}(\cos 3\phi),$$
$$0 \leq \phi \leq \tfrac{1}{3}\pi, \tag{3.30}$$

$$\psi_{nl}(r, \phi + \tfrac{1}{3}p\pi) = (-)^{pl}\psi_{nl}(r, \phi), \quad 0 \leq \phi \leq \tfrac{1}{3}\pi,$$
$$p = 1, 2, 3, 4, 5, \tag{3.31a}$$

$$\psi_{nl}(r, \phi + \tfrac{1}{3}p\pi) = (-)^{p(l+1)}\psi_{nl}(r, \phi), \quad 0 \leq \phi \leq \tfrac{1}{3}\pi,$$
$$p = 1, 2, 3, 4, 5, \tag{3.31b}$$

$$E_{2n+3l} = (\tfrac{3}{2})^{\frac{1}{2}}\omega(2n + 3l + 3a + \tfrac{5}{2}). \tag{3.32}$$

Here a is again defined by Eq. (2.5), L_m^b and C_n^λ indicate again the Laguerre and Gegenbauer polynomials,[5] the quantum numbers n and l take all nonnegative integral values, and r and ϕ are connected to the coordinates of the three particles by Eqs. (3.5). Note that these equations imply that $\sin 3\phi$ and $\cos 3\phi$ depend in a symmetric manner on the coordinates of the three particles:

$$\sin 3\phi = -3(6)^{\frac{1}{2}}(x_1 - x_2)(x_2 - x_3)(x_3 - x_1) \times [(x_1 - x_2)^2 + (x_2 - x_3)^2 + (x_3 - x_1)^2]^{-\frac{3}{2}},$$
$$\cos 3\phi = 2^{\frac{1}{2}}(x_1 + x_2 - 2x_3)(x_2 + x_3 - 2x_1) \times (x_3 + x_1 - 2x_2) \times [(x_1 - x_2)^2 + (x_2 - x_3)^2 + (x_3 - x_1)^2]^{-\frac{3}{2}}. \tag{3.33}$$

Equation (3.31a) is for the case of Bose statistics, Eq. (3.31b) for Fermi statistics; in these cases a unique wavefunction corresponds to each pair n, l. In the Boltzmann case, in each one of the six "angular" sectors $(\frac{1}{3}p\pi \leq \phi \leq \frac{1}{3}(p+1)\pi, \ p = 0, 1, 2, 3, 4, 5)$ the wavefunction can be defined through Eqs. (3.30) and (3.31) and can be set to zero in the remaining five sectors, giving altogether six different states corresponding to each pair n, l.[7] On the other hand, the degeneracy of each level E_N, $N = 2n + 3l$, is, in the case of identical particles (Bose or Fermi statistics), the integral part of $\frac{1}{6}(N + 6)$, and of course six times that in the case of distinguishable particles (Boltzmann statistics).

It should be noted that the energy spectrum is again linear; it reproduces the spectrum of the problem without "centrifugal" forces *and with Fermi statistics*, except for a constant shift of all the energy eigenvalues by the quantity $3(\frac{3}{2})^{\frac{1}{2}}\omega(a - \frac{1}{2})$. In fact it is easily seen that in the $g \to 0$ limit ($a \to \frac{1}{2}$) the eigenfunctions (3.30) and (3.31b) and the eigenvalues (3.32) go over into the eigenfunctions and eigenvalues of the "oscillator" three-body problem with Fermi statistics, in analogy to the previous cases.

In the $\omega \to 0$ limit, the eigenfunctions (3.30) go over into the zero-energy eigenfunctions of the problem with only "centrifugal" forces. On the other hand, the complete set of eigenfunctions of this problem is

$$\psi_{kl}(r, \phi) = J_{3l+3a+\frac{3}{2}}(kr)(\sin 3\phi)^{a+\frac{1}{2}} C_l^{a+\frac{1}{2}}(\cos 3\phi),$$
$$0 \leq \phi \leq \tfrac{1}{3}\pi, \quad l = 0, 1, 2, \cdots, \tag{3.34}$$

and the corresponding energy eigenvalues are

$$E = k^2, \quad 0 \leq k < \infty. \tag{3.35}$$

The complete definition of the wavefunction in the whole range $0 \leq \phi < 2\pi$ may be given with the same prescriptions as in the previous case [see Eqs. (3.31)], depending on the statistics. These eigenfunctions depend now on the discrete index l and the continuous index k; they are, of course, not normalizable.

4. CONCLUSION

In this paper the Schrödinger equation for the three-body problem under consideration has been explicitly solved, and in this manner the eigenvalues and eigenfunctions of the problem have been determined. The simplicity (and degeneracy) of the spectrum obtained suggests that a solution by group-theoretical techniques should also be possible. This has not been attempted here.

The possibility of solving this problem completely suggests attacking similar, but more complicated, ones. The type of generalization that appears natural is an increase in the number of particles and/or

dimensions of the space.[11] It is, of course, well known that, when no "centrifugal" forces are present, the problem is easily solved for an arbitrary number of particles and dimensions; but such a problem is of very limited interest, because it corresponds to a collection of completely decoupled oscillators.[12]

Finally, as regards the possible usefulness of the present model, it is perhaps worthwhile to recall that one of the major difficulties in many-body (or rather, more-than-two-body) physics has to do with the presence of interactions that are singular at short interparticle distance. The present model, which features interactions just of this kind, might therefore, in spite of its extreme simplicity, provide useful tests for approximation schemes and computational techniques or a convenient starting point for the (approximate) solution of more realistic problems.

[11] An interesting question to ask in this connection is whether the energy spectrum will continue to depend linearly on all quantum numbers. It is expected that it will if only the number of particles is increased, but that it will not if the number of space dimensions is increased. Incidentally, using the approach of this paper, one can immediately solve the one-dimensional N-body problem with equal-strength harmonical forces between every two particles and, in addition, either only one or only three (equal-strength) centrifugal potentials depending, respectively, either only upon the interparticle distance between two particles or only upon the three interparticle distances between three particles.

[12] Also the problem considered in this paper is, however, in some sense equivalent to that with decoupled oscillators, as indicated by the structure of its spectrum. But this correspondence does not appear to be a trivial one.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge a useful correspondence with Professor H. R. Post concerning the separability of this three-body problem. Conversations with friends have also been helpful, especially with Dr. P. Camiz, Dr. C. Marchioro, and Dr. Yu. A. Simonov.

The research reported here was completed in Rome, but the final version of this paper was written in London. I would like to thank the Accademia Nazionale dei Lincei and the Royal Society for a joint grant which made my visit to London possible, and the Physics Department at Imperial College for their kind hospitality.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 12 DECEMBER 1969

Ground State of a One-Dimensional N-Body System

F. CALOGERO
Istituto di Fisica, Università di Roma and Istituto Nazionale di Fisica Nucleare Sezione di Roma, Rome, Italy

(Received 14 March 1969)

The problem of N quantum-mechanical equal particles interacting pairwise by inverse-cube forces ("centrifugal potential") in addition to linear forces ("harmonical potential") is considered in a one-dimensional space. An explicit expression for the ground-state energy and for the corresponding wavefunction is exhibited. A class of excited states is similarly displayed.

1. INTRODUCTION

The one-dimensional N-body problem with linear ("harmonical") and inverse-cube ("centrifugal") pair forces has been recently solved for $N = 3$; the complete energy spectrum and all the corresponding eigenfunctions have been exhibited.[1] For arbitrary N, the ground state and a set of excited states are given in this paper. From these results, a remarkably simple expression is obtained for the ground-state wavefunction of the one-dimensional system composed of N oscillators interacting pairwise and obeying Fermi statistics.

Units are chosen so that $\hbar = 1$.

2. THE N-BODY PROBLEM

The Schrödinger equation for the N-body problem under consideration is

$$\left[-\frac{1}{2m}\sum_{i=1}^{N}\frac{\partial^2}{\partial x_i^2} + \tfrac{1}{4}m\omega^2\sum_{i=2}^{N}\sum_{j=1}^{i-1}(x_i - x_j)^2 + g\sum_{i=2}^{N}\sum_{j=1}^{i-1}(x_i - x_j)^{-2} - E\right]\psi = 0. \quad (2.1)$$

We assume throughout that $g > -\frac{1}{2}$, to avoid two-body collapse[1,2]; and we consider only the sector of the N-body phase-space characterized by the inequalities

$$x_1 \leq x_2 \leq \cdots x_i \leq x_{i+1} \leq \cdots x_N. \quad (2.2)$$

As explained in I, for $g \neq 0$ the extension of the wavefunction to the whole phase-space is performed by the simple prescription

$$\psi(Px) = \eta_P\psi(x), \quad (2.3)$$

where x indicates the set $\{x_i; i = 1, 2 \cdots N\}$, P indicates an arbitrary permutation, and η_P equals unity if the particles obey Bose statistics and equals the parity of the permutation if the particles obey Fermi statistics. If the particles are considered distinguishable, i.e., if they obey Boltzmann statistics, each wavefunction $\psi(x)$ gives rise to $N!$ different states, each one of these being characterized by a wavefunction vanishing identically for all but one of the $N!$ particle orderings, and that one is given by Eq. (2.3).[1]

We introduce now the two completely symmetrical variables z and r^2:

$$z = \prod_{i=2}^{N}\prod_{j=1}^{i-1}(x_i - x_j), \quad (2.4)$$

$$r^2 = \frac{1}{N}\sum_{i=2}^{N}\sum_{j=1}^{i-1}(x_i - x_j)^2 = \frac{1}{2N}\sum_{i=1}^{N}\sum_{j=1}^{N}(x_i - x_j)^2; \quad (2.5)$$

and we consider the subset of solutions of Eq. (2.1) of the form

$$\psi = z^{a+\frac{1}{2}}\varphi(r), \quad (2.6)$$

with

$$a = +\tfrac{1}{2}(1 + 4mg)^{\frac{1}{2}}. \quad (2.7)$$

Before proceeding to show that such a subset is not empty, we note that the ansatz (2.6) implies that ψ vanishes proportionally to $(x_i - x_j)^{a+\frac{1}{2}}$ whenever the distance $x_i - x_j$ vanishes. This phenomenon originates from the singularity of the "centrifugal" potential; together with the one-dimensional nature of the model, it implies the impossibility for any particle to overtake any other particle.[1] Thus configurations characterized by different particle orderings are dynamically separated; and the prescription (2.3) is justified, because the wavefunction resulting from its application does indeed possess the required continuity properties.[1] We also note that since both variables z and r are obviously translation invariant, so is the wavefunction (2.6), i.e., it describes states in the center-of-mass frame.

[1] F. Calogero, J. Math. Phys. **10**, 2191 (1969) (preceding paper), hereafter referred to as I.
[2] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics* (Pergamon Press, Inc., New York, 1958), Sec. 35.

We now insert the ansatz (2.6) in the Schrödinger equation (2.1):

$$\Big[-\frac{1}{2m}\sum_{i=1}^{N}\Big\{(a+\tfrac{1}{2})(a-\tfrac{1}{2})z^{-2}\Big(\frac{\partial z}{\partial x_i}\Big)^2 + (a+\tfrac{1}{2})z^{-1}\frac{\partial z}{\partial x_i} + 2(a+\tfrac{1}{2})z^{-1}\frac{\partial z}{\partial x_i}\frac{\partial r}{\partial x_i}\frac{\partial}{\partial r} + \Big(\frac{\partial r}{\partial x_i}\Big)^2\frac{\partial^2}{\partial r^2} + \frac{\partial^2 r}{\partial x_i^2}\frac{\partial}{\partial r}\Big\} + \frac{N}{4}m\omega^2 r^2 + g\sum_{i=2}^{N}\sum_{j=1}^{i-1}(x_i-x_j)^{-2} - E\Big]z^{a+\frac{1}{2}}\varphi(r) = 0. \quad (2.8)$$

In the Appendix it is proved that

$$z^{-2}\sum_{i=1}^{N}\Big(\frac{\partial z}{\partial x_i}\Big)^2 = 2\sum_{i=2}^{N}\sum_{j=1}^{i-1}(x_i-x_j)^{-2} = \sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}(x_i-x_j)^{-2}, \quad (2.9)$$

$$\sum_{i=1}^{N}\frac{\partial^2 z}{\partial x_i^2} = 0, \quad (2.10)$$

$$\sum_{i=1}^{N}\frac{\partial z}{\partial x_i}\frac{\partial r}{\partial x_i} = \tfrac{1}{2}N(N-1)zr^{-1}, \quad (2.11)$$

$$\sum_{i=1}^{N}\Big(\frac{\partial r}{\partial x_i}\Big)^2 = 1, \quad (2.12)$$

$$\sum_{i=1}^{N}\frac{\partial^2 r}{\partial x_i^2} = (N-2)r^{-1}. \quad (2.13)$$

Using these equations and Eq. (2.7), Eq. (2.8) becomes

$$\Big[-\frac{1}{2m}\Big\{\frac{d^2}{dr^2} + [(a+\tfrac{1}{2})N(N-1)+N-2]\frac{1}{r}\frac{d}{dr}\Big\} + \frac{N}{4}m\omega^2 r^2 - E\Big]\varphi(r) = 0. \quad (2.14)$$

The normalizable solutions of this equation are

$$\varphi_n(r) = \exp[-\tfrac{1}{2}m\omega(mN)^{\frac{1}{2}}r^2]L_n^b[m\omega(mN)^{\frac{1}{2}}r^2], \quad n = 0, 1, 2, \cdots, \quad (2.15)$$

where L_n^b is a Laguerre polynomial[3] and

$$b = \tfrac{1}{4}N(N+1) + \frac{a}{2}N(N-1) - \tfrac{3}{2}. \quad (2.16)$$

The corresponding energy eigenvalues are

$$E_n = \tfrac{1}{2}\omega(N/2)^{\frac{1}{2}}[4n + \tfrac{1}{2}N(N+1) + aN(N-1) - 1] \quad (2.17a)$$

$$= \tfrac{1}{2}\omega(N/2)^{\frac{1}{2}}[4n + N^2 - 1 + (a-\tfrac{1}{2})N(N-1)]. \quad (2.17b)$$

For $N = 2$, the eigenfunctions (2.6), (2.15) exhaust the whole set of eigenfunctions of the problem. For $N \geq 3$, they constitute only a subset of all the eigenfunctions; for $N = 3$, they correspond to the $l = 0$ eigenfunctions of I. But for all N and for $g \neq 0$, the $n = 0$ eigenfunction

$$\psi_0 = z^{a+\frac{1}{2}}\exp[-\tfrac{1}{2}m\omega(mN)^{\frac{1}{2}}r^2] \quad (2.18)$$

and the corresponding eigenvalue

$$E_0 = \tfrac{1}{2}\omega(N/2)^{\frac{1}{2}}[N^2 - 1 + (a-\tfrac{1}{2})N(N-1)] \quad (2.19)$$

correspond to the ground state of the system. This is implied by the evident property of ψ_0: to be nodeless within the sector (2.2). Moreover, in the $g \to 0$, $a \to \frac{1}{2}$ limit, ψ_0 and E_0, Eqs. (2.18) and (2.19), become, respectively, the ground-state wavefunction and energy of the N-fermion oscillator problem[4]:

$$\psi_0 = z\exp[-\tfrac{1}{2}m\omega(N/2)^{\frac{1}{2}}z^2], \quad (2.20)$$

$$E_0 = \tfrac{1}{2}\omega(N/2)^{\frac{1}{2}}(N^2-1). \quad (2.21)$$

The eigenvalue E_0 was already well known,[5] but the expression (2.20) of the ground-state wavefunction is considerably simpler than that given in Ref. 5.

In fact, in the $g \to 0$, $a \to \frac{1}{2}$ limit, all the eigenvalues and eigenfunctions, Eqs. (2.17), (2.6), and (2.15), go over into eigenvalues and eigenfunctions of the N-oscillator problem *with Fermi statistics*.[4] This phenomenon has been explained in sufficient detail in I (for $N = 3$) so not to warrant any further discussion here. It appears reasonable to conjecture that the only difference between the complete energy spectrum of the problem considered here and that of the pure oscillator problem *with Fermi statistics* is a shift of each energy level by the constant amount ΔE:

$$\Delta E = (a+\tfrac{1}{2})\omega(N/2)^{\frac{1}{2}}\tfrac{1}{2}N(N-1). \quad (2.22)$$

It is also easily seen that, again as in I, in the $a \to -\frac{1}{2}$ limit the eigenvalues and eigenfunctions, Eqs. (2.17), (2.6), and (2.15), go over into eigenvalues and eigenfunctions of the pure oscillator problem *with Bose statistics*[4]; and in particular in this limit ψ_0 and E_0, Eqs. (2.18) and (2.19), go over into the

[3] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series and Products* (Academic Press Inc., New York, 1965); *Higher Transcendental Functions*, A. Erdélyi, Ed. (McGraw-Hill Book Co., Inc., New York, 1953), Vol. II. Note that the definition of Laguerre polynomial given here is different from that used in some textbooks, for instance in Ref. 2.

[4] Provided of course the appropriate prescription is used in Eq. (2.3) to define the eigenfunctions throughout the whole N-body phase-space; compare with the analogous situation as discussed in I.

[5] H. R. Post, Proc. Phys. Soc. (London) **66A**, 649 (1953). See also, for instance, J. M. Lévy-Leblond, Phys. Letters **26A**, 540 (1968).

ground state of the N-oscillator problem with Bose statistics.[4]

ACKNOWLEDGMENT

It is a pleasure to acknowledge a suggestive conversation with Professor J. G. Valatin.

APPENDIX

In this Appendix Eqs. (2.9)–(2.13) are proved.

Proof of Eqs. (2.12) and (2.13)

Differentiating the definition of r^2, Eq. (2.5), we get

$$r\frac{\partial r}{\partial x_i} = N^{-1}\sum_{j=1}^{N}(x_i - x_j). \tag{A1}$$

Differentiating again with respect to x_i, we get

$$\left(\frac{\partial r}{\partial x_i}\right)^2 + r\frac{\partial^2 r}{\partial x_i^2} = \frac{N-1}{N}. \tag{A2}$$

Thus

$$\sum_{i=1}^{N}\left(\frac{\partial r}{\partial x_i}\right)^2 + r\sum_{i=1}^{N}\frac{\partial^2 r}{\partial x_i^2} = N - 1. \tag{A3}$$

On the other hand, from Eq. (A1),

$$r^2\left(\frac{\partial r}{\partial x_i}\right)^2 = N^{-2}\sum_{j=1}^{N}\sum_{l=1}^{N}(x_i - x_j)(x_i - x_l). \tag{A4}$$

If we now use

$$\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{l=1}^{N}(x_i - x_j)(x_i - x_l) = N^2 r^2, \tag{A5}$$

we get, from Eq. (A4), Eq. (2.12); and from this equation and Eq. (A3), we get Eq. (2.13).

There remains to prove Eq. (A5). This can be done as follows:

$$\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{l=1}^{N}(x_i - x_j)(x_i - x_l)$$

$$= \sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{l=1}^{N}(x_i - x_j)(x_i - x_j + x_j - x_l) \tag{A6a}$$

$$= 2Nr^2 + \sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{l=1}^{N}(x_i - x_j)(x_j - x_l) \tag{A6b}$$

$$= 2N^2 r^2 - \sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{l=1}^{N}(x_i - x_j)(x_i - x_l). \tag{A6c}$$

Equation (A6a) is trivial; Eq. (A6b) obtains using the definition of r^2, Eq. (2.5); Eq. (A6c) follows exchanging the dummy indices i and j, and it immediately implies Eq. (A5). Q.E.D.

Proof of Eq. (2.11)

Differentiating the definition of z, Eq. (2.4), we get

$$\frac{\partial z}{\partial x_i} = z\sum_{j=1}^{N}{}'(x_i - x_j)^{-1}. \tag{A7}$$

Here, and always below, the apex in a sum indicates that the divergent term must be omitted. Equation (2.11) is then immediately implied by the combination of Eqs. (A1) and (A7) with the formula

$$\sum_{i=1}^{N}\sum_{j=1}^{N}{}'\sum_{l=1}^{N}\frac{x_i - x_l}{x_i - x_j} = \tfrac{1}{2}N^2(N-1). \tag{A8}$$

As for this formula, it can be proved as follows:

$$\sum_{i=1}^{N}\sum_{j=1}^{N}{}'\sum_{l=1}^{N}\frac{x_i - x_l}{x_i - x_j}$$

$$= \sum_{i=1}^{N}\sum_{j=1}^{N}{}'\sum_{l=1}^{N}\frac{x_i - x_j + x_j - x_l}{x_i - x_j} \tag{A9a}$$

$$= N^2(N-1) + \sum_{i=1}^{N}\sum_{j=1}^{N}{}'\sum_{l=1}^{N}\frac{x_j - x_l}{x_i - x_j} \tag{A9b}$$

$$= N^2(N-1) - \sum_{i=1}^{N}\sum_{j=1}^{N}{}'\sum_{l=1}^{N}\frac{x_i - x_l}{x_i - x_j}. \tag{A9c}$$

These steps to derive Eq. (A9c) are essentially the same as those for the derivation of Eq. (A6c) above; on the other hand, Eq. (A9c) immediately implies Eq. (A8). Q.E.D.

Proof of Eqs. (2.9) and (2.10)

Differentiating Eq. (A7), we get

$$\frac{\partial^2 z}{\partial x_i^2} = \frac{\partial z}{\partial x_i}\sum_{j=1}^{N}{}'(x_i - x_j)^{-1} - z\sum_{j=1}^{N}(x_i - x_j)^{-2} \tag{A10a}$$

$$= z^{-1}\left(\frac{\partial z}{\partial x_i}\right)^2 - z\sum_{j=1}^{N}{}'(x_i - x_j)^{-2} \tag{A10b}$$

$$= z\left[\sum_{j=1}^{N}{}'\sum_{l=1}^{N}{}'(x_i - x_j)^{-1}(x_i - x_l)^{-1} - \sum_{j=1}^{N}{}'(x_i - x_j)^{-2}\right]. \tag{A10c}$$

Equations (A10b) and (A10c) follow from Eqs. (A10a) and (A7). But the r.h.s. of Eq. (A10c) yields a vanishing result when summed over i, namely,

$$a_N \equiv \sum_{i=1}^{N}\sum_{j=1}^{N}{}'\sum_{\substack{l=1\\ l\neq j}}^{N}{}'(x_i - x_j)^{-1}(x_i - x_l)^{-1} = 0. \tag{A11}$$

This immediately implies, through Eq. (A10c) itself, the validity of Eq. (2.10), and through Eq. (A10b), the validity of Eq. (2.9).

There remains to prove Eq. (A11). First note that

$$a_N \equiv \sum_{i=1}^{N}\sum_{j=1}^{N}{}'\left[\sum_{l=1}^{j-1}{}'(x_i - x_j)^{-1}(x_i - x_l)^{-1} + \sum_{l=j+1}^{N}{}'(x_i - x_j)^{-1}(x_i - x_l)^{-1}\right] \tag{A12a}$$

$$= 2\sum_{i=1}^{N}\sum_{j=2}^{N}{}'\sum_{l=1}^{j-1}{}'(x_i - x_j)^{-1}(x_i - x_l)^{-1}. \tag{A12b}$$

Equation (A12b) is obtained exchanging the dummy indices j and l in the second term in the rhs of (A12a), thereby demonstrating its equality to the first term. Incidentally, in writing Eq. (A12a), and also below, we adopt the convention that any sum vanishes whenever the upper limit of the running index is smaller than its lower limit.

It is now easy to show, by direct computation, that

$$a_3 = 0. \tag{A13}$$

Then, defining b_N through

$$b_N = \tfrac{1}{2}(a_N - a_{N-1}), \tag{A14}$$

Eq. (A11) is proved if one can show that

$$b_N = 0, \quad N \geq 4. \tag{A15}$$

In fact, from Eqs. (A14) and (A12b),

$$b_N = \sum_{j=1}^{N-1}(x_N - x_j)^{-1}\Big[\sum_{l=1}^{j-1}(x_N - x_l)^{-1} - \sum_{l=1}^{N-1}{}'(x_j - x_l)^{-1}\Big]. \tag{A16}$$

But since

$$\sum_{l=1}^{j-1}(x_N - x_l)^{-1} - \sum_{l=1}^{N-1}{}'(x_j - x_l)^{-1} = \sum_{l=1}^{j-1}(x_N - x_l)^{-1} - \sum_{l=1}^{j-1}(x_j - x_l)^{-1} - \sum_{l=j+1}^{N-1}(x_j - x_l)^{-1} \tag{A17a}$$

$$= -(x_N - x_j)\sum_{l=1}^{j-1}(x_N - x_l)^{-1}(x_j - x_l)^{-1} - \sum_{l=j+1}^{N-1}(x_j - x_l)^{-1}, \tag{A17b}$$

we get

$$b_N = -\sum_{j=1}^{N-1}\sum_{l=1}^{j-1}(x_N - x_l)^{-1}(x_j - x_l)^{-1} - \sum_{j=1}^{N-1}\sum_{l=j+1}^{N-1}(x_N - x_j)^{-1}(x_j - x_l)^{-1}. \tag{A18}$$

By exchanging the dummy indices j and l in the second term in the rhs of this equation, it is apparent that this term is just the opposite of the first one, so that b_N vanishes. Q.E.D.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 12 DECEMBER 1969

Self-Consistent Approximations in Many-Body Systems. II

O. SHLIDOR* AND M. REVZEN
Department of Physics, Technion—Israel Institute of Technology, Haifa, Israel

(Received 1 July 1968)

A stationary property of the grand-canonical potential is introduced. This stationary property is used to define a class of self-consistent approximations. The class of self-consistent approximations is shown to be particularly suitable for the random-phase approximation (RPA). The conditions for self-consistency are only sufficient.

I. INTRODUCTION

Luttinger and Ward[1] (L.W. henceforth) give an exact diagrammatic (i.e., perturbative) proof that the grand-canonical potential[2] Ω is stationary with respect to variation of the proper self-energy Σ. This stationary property enables the formulation of a corresponding class of self-consistent approximations (SCA).[3,4]

In this paper the exact grand-canonical potential Ω is shown to be stationary with respect to variation of some quantity W, which is related to the density–density correlation function.[5] The method of proof is diagrammatic (i.e., perturbative) and is parallel to the L.W. proof for the aforementioned stationarity. But, whereas L.W. concentrate on the particle propagator, here the interaction line is the object of our analysis.

The stationarity with respect to W enables the definition of a new class of SCA. The self-consistency appears in the fact that one may calculate any desired thermodynamic property at equilibrium from the approximate grand-canonical potential, as well as from the corresponding approximate propagator, or the approximate density–density correlation, the results being identical.

It will be seen that RPA is a natural approximation in this SCA in the same sense as the Hartree–Fock approximation is natural in the previous[3,4] class of SCA.

II. STATIONARITY OF THE GRAND-CANONICAL POTENTIAL

The grand-canonical potential (GCP) Ω is connected with statistical mechanics through the grand partition function Z_G by

$$Z_G = \exp(-\beta\Omega). \tag{1}$$

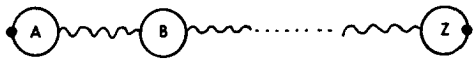


FIG. 1. A representative $F_{k(\zeta_l)}$ diagram. A, B, etc., by their structure are arbitrary vacuum–vacuum diagrams, each one lacking a single wavy line. The wavy lines joining A, B, etc., carry the momentum "k" and "energy" "ζ_l," which go in and out through the heavy dots.

It can be represented in terms of linked diagrams, namely,

$$\Omega - \Omega_0 = -\frac{1}{\beta}\sum_{n=1}^{\infty}\frac{(-1)^n}{n!}\int_0^\beta du_1 \cdots \int_0^\beta du_n \times \langle T_u\{V_{(u_1)} \cdots V_{(u_n)}\}\rangle_{\mathrm{L}}, \tag{2}$$

where Ω_0 is the GCP with no interaction. (We use the notation as in Paper I, apart from replacing A by Ω.)

Our aim is to analyze a functional which emphasizes the role of the interaction line in the formal building of the vacuum–vacuum (GCP) and the propagator diagrams. To this end we consider the density–density correlation function[6]

$$F_{k(u)} = \langle T_u\{\rho_{-k(u)}\rho_{k(0)}\}\rangle. \tag{3}$$

The density operator in momentum space is defined as

$$\rho_{k(u)} = \sum_r a^+_{r(u)}a_{r+k(u)}. \tag{4}$$

The perturbative development of $F_{k(u)}$ is

$$F_{k(u)} = \sum_{n=0}^{\infty}\frac{(-1)^n}{n!}\int_0^\beta du_1 \cdots \int_0^\beta du_n \times \langle T_u\{\rho_{-k(u)}\rho_{k(0)}V_{(u_1)} \cdots V_{(u_n)}\}\rangle_{\mathrm{L}}, \tag{5}$$

where the subscript L signifies linked[1] and the expectation value is now taken with a noninteracting Hamiltonian.

The Fourier transform of $F_{k(u)}$ is defined

$$F_{k(\zeta_l)} = \int_0^\beta e^{-\zeta_l u}F_{k(\zeta_l)}\,du, \tag{6}$$

where

$$\zeta_l = 2l\pi i/\beta + \mu, \quad l = 0, \pm1, \pm2, \cdots.$$

The general shape of an $F_{k(\zeta_l)}$ diagram is drawn in Fig. 1. k and ζ_l are the momentum and "energy,"

* Based on a thesis submitted to the Senate of Technion, Israel Institute of Technology, in partial fulfillment of the requirements for M.Sc. degree.

[1] J. M. Luttinger and J. Ward, Phys. Rev. **118**, 1417 (1960).

[2] H. B. Callen, *Thermodynamics* (John Wiley & Sons, Inc., New York, 1960).

[3] M. Revzen, J. Math. Phys. **6**, 450 (1965). This article is referred to as I.

[4] G. Baym, Phys. Rev. **127**, 1391 (1962).

[5] C. De Dominicis and P. C. Martin, J. Math. Phys. **5**, 14, 31 (1964).

[6] See, e.g., David Pines, *Elementary Excitations in Solids: Electrons, Phonons and Plasmons* (W. A. Benjamin, Inc., New York, 1963).

respectively, which enter and leave the diagram through the heavy dots on the two opposite sides of the diagram. Conservation of momentum and "energy" obviously is valid in the heavy dots. The numerical factor of an nth-order diagram is

$$\sum_{\substack{\text{all free}\\ \text{indices}}} \frac{(-1)^n}{n!}\frac{1}{2^n}\frac{1}{\beta^{n+1}}\prod_{i=1}^{n} v_{k_i}.$$

A, B, etc., are, pictorially, vacuum–vacuum diagrams, which lack a single wavy (interaction) line and cannot be separated into two disconnected parts by merely removing a single wavy line. The momentum and "energy" carried along by the wavy lines which join A, B, etc., are k and ζ_l.

We may sum up all $F_{k(\zeta_l)}$ diagrams of m parts as depicted pictorially in Fig. 2. $W_{k(\zeta_l)}$ indicates the sum of all the vacuum–vacuum diagrams which lack a single wavy line and are not constructed of distinct parts joined by wavy lines only. k and ζ_l denote the momentum and "energy" of the missing interaction line.

The full expression for $F_{k(\zeta_l)}$ is

$$\begin{aligned} F_{k(\zeta_l)} &= W_{k(\zeta_l)} + W_{k(\zeta_l)}v_k W_{k(\zeta_l)} + \cdots \\ &= W_{k(\zeta_l)} + W_{k(\zeta_l)}v_k[W_{k(\zeta_l)} + \cdots \\ &= W_{k(\zeta_l)} + W_{k(\zeta_l)}v_k F_{k(\zeta_l)}; \end{aligned} \tag{7}$$

hence,

$$F_{k(\zeta_l)} = W_{k(\zeta_l)}/(1 - W_{k(\zeta_l)}v_k). \tag{8}$$

By naked diagrams of $F_{k(\zeta_l)}$, we mean those which are not constructed of distinct parts joined by wavy lines. In these lines an appearance of momentum q on a certain wavy line does not imply (by the conservation laws) that some other line should be assigned the same momentum q; e.g., see Fig. 3. (Diagrams which do not fulfill these criteria are termed "dressed.")

In order to construct $W_{k(\zeta_l)}$ diagrammatically, one has to sum up all the naked diagrams of $F_{k(\zeta_l)}$ and replace in them every wavy line by a saw-tooth line.

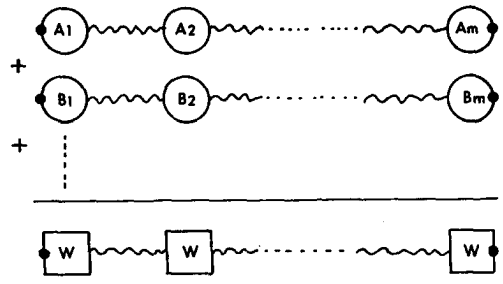


FIG. 2. The mth-order $F_{k(\zeta_l)}$ diagrams are summed up to give a single diagram built up of m equal $W_{k(\zeta_l)}$ blocks joined by wavy lines.

FIG. 3. Examples for $F_{k(\zeta_l)}$ diagrams not constructed by distinct parts; (a) is naked and (b) dressed.

The analytical meaning of a saw-tooth line is

$$U_{k(\zeta_l)} = v_k(1 + F_{k(\zeta_l)}v_k), \tag{9}$$

and its graphical meaning is given in Fig. 4. By (8) we get

$$U_{k(\zeta_l)} = v_k/(1 - W_{k(\zeta_l)}v_k). \tag{10}$$

Let $F_{nk(\zeta_l)}$ represent the sum of all nth-order $F_{k(\zeta_l)}$ diagrams. The vacuum–vacuum diagrams of order $n+1$ may be constructed from $F_{nk(\zeta_l)}$ by the relation

$$(\Omega - \Omega_0)_{n+1} = \frac{-1}{2(n+1)}\frac{1}{\beta}\sum_k\sum_l v_k F_{nk(\zeta_l)}e^{\zeta_l 0^+}. \tag{11}$$

By introducing an interaction-strength parameter λ: $v_k \to \lambda v_k$ and using (8), (10), and (11), we obtain[1]

$$\Omega_{(\lambda)} = \Omega_0 - \frac{1}{2\beta}\sum_k\sum_l\int_0^\lambda \frac{d\lambda'}{\lambda'} W_{k(\zeta_l,\lambda')}U_{k(\zeta_l,\lambda')}e^{\zeta_l 0^+}, \tag{12}$$

so that

$$\lambda\frac{\partial\Omega}{\partial\lambda} = -\frac{1}{2\beta}\sum_k\sum_l W_{k(\zeta_l,\lambda)}U_{k(\zeta_l,\lambda)}e^{\zeta_l 0^+}. \tag{13}$$

We define the functional R' to be the sum of all naked vacuum–vacuum diagrams, in which every wavy line is replaced by a saw-tooth line:

$$R'_{(\lambda)} = -\sum_n\sum_k\frac{1}{2\beta}\sum_l\frac{1}{n+1}U_{k(\zeta_l,\lambda)}W_{nk(\zeta_l,\lambda)}e^{\zeta_l 0^+}. \tag{14}$$

(W_n corresponds to F_n in the natural manner.)

The functional R'' is defined as

$$R''_{(\lambda)} = \frac{1}{2\beta}\sum_k\sum_l e^{\zeta 0^+}\left\{\ln\left[-\frac{1}{\lambda v_k} + W_{k(\zeta_l,\lambda)}\right] + U_{k(\zeta_l,\lambda)}W_{k(\zeta_l,\lambda)}\right\}. \tag{15}$$

It can readily be proved that

$$R = R' + R'' \tag{16}$$

is not sensitive to first-order variations in $W_{k(\zeta_l,\lambda)}$, i.e.,

$$\frac{\delta R_{(\lambda)}}{\delta W_{k(\zeta_l,\lambda)}} = 0. \tag{17}$$

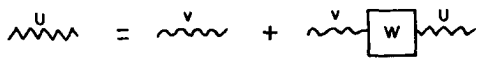


FIG. 4. Graphical meaning for Eq. (9).

On calculating the derivative of $R(\lambda)$ with respect to λ, we obtain

$$\lambda \frac{\partial R}{\partial \lambda} = -\frac{1}{2\beta} \sum_k \sum_l U_{k(\zeta_l,\lambda)} W_{k(\zeta_l,\lambda)} e^{\zeta_l 0^+} \tag{18}$$

and, therefore,

$$\frac{\partial R}{\partial \lambda} = \frac{\partial \Omega}{\partial \lambda}. \tag{19}$$

Since $R_{(\lambda=0)} = 0$, we see that

$$R = \Omega - \Omega_0, \tag{20}$$

and hence, in using (17),

$$\frac{\delta \Omega}{\delta W_{k(\zeta_l)}} = 0. \tag{21}$$

III. SELF-CONSISTENT APPROXIMATIONS

The stationarity of Ω with respect to variation of $W_{k(\zeta_l)}$ is an exact result,

$$\frac{\delta \Omega}{\delta W_{k(\zeta_l)}} = \frac{\delta \Omega}{\delta U_{k(\zeta_l)}} = 0, \tag{22}$$

and is analogous with the stationarity of Ω with respect to the proper self-energy[1]:

$$\frac{\delta \Omega}{\delta \Sigma_{k(\zeta_l)}} = \frac{\delta \Omega}{\delta G_{k(\zeta_l)}} = 0.$$

The basis for the discussion above was the pair of implicit equations

$$F = W(u)[1 + vF], \tag{7'}$$

$$U = v(1 + vF). \tag{9'}$$

Thus the naked diagrams of F were used as building bricks for W, which was the building brick for F, which, in turn was the building brick for U, the latter being used to construct W. The procedure does not require the whole set of naked diagrams of F. It is sufficient to use a partial set, finite or infinite, and still to retain the exact functional interrelations among F_A, W_A, U_A, and Ω_A (subscript A denotes "approximate"). The stationary property, namely

$$\frac{\delta \Omega_A}{\delta W_{Ak(\zeta_l)}} = 0, \tag{23}$$

thus holds.

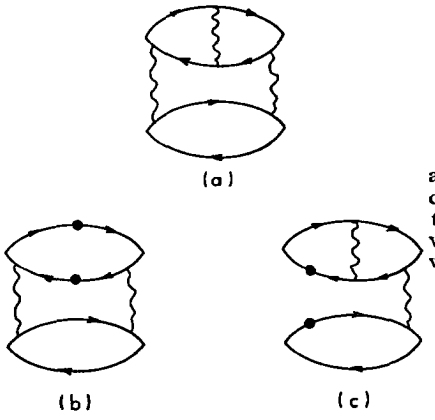


FIG. 5. (b) and (c) are two different $F_{k(\zeta_l)}$ diagrams which yield the same vacuum–vacuum diagram, (a), when closed.

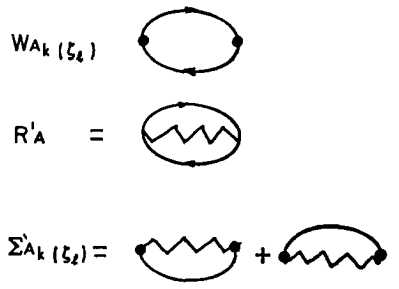


FIG. 6. The diagrammatic base for the RPA.

It follows that any thermodynamic property (at equilibrium) which is calculated with F_A may be obtained from Ω_A as well, with no difference whatsoever.

The set of naked F diagrams forming the approximation cannot, however, be chosen arbitrarily. The inclusion of certain F diagrams requires the inclusion of all other F diagrams, which, together with the original F when closed, yield the same vacuum–vacuum diagram (see Fig. 5). Moreover, all vacuum–vacuum diagrams which are topologically equivalent to the latter should be constructed too.

Given a set of R' diagrams, there is a one-valued correspondence between these diagrams and a set of self-energy diagrams Σ' obtained by removing in turn each one of the free propagator lines of the former. These self-energy diagrams contain sawtooth lines for their interaction lines, so that each diagram represents an infinite number of ordinary, self-energy diagrams. The unique value is due to the unique way of identifying the R' source of a given self-energy diagram.

If our functional R' is an approximate one, say R'_A, then we use this procedure to define the set of self-energy diagrams consistent with this approximation.

As an example, consider the approximation which is sketched graphically in Fig. 6. (These diagrams lead to the RPA.)

In order to illustrate that the correspondence between R' and Σ' is self-consistent, let us calculate the average number of particles in a grand-canonical ensemble.[4] From its definition,

$$\langle N \rangle = -\frac{\partial \Omega}{\partial \mu} = -\frac{\partial R}{\partial \mu} - \frac{\partial \Omega_0}{\partial \mu}. \tag{24}$$

R'' is dependent on μ only through its dependence on $W_{k(\zeta_l)}$; hence $\partial R''/\partial\mu$ will cancel out that part of $\partial R'/\partial\mu$ which corresponds to the dependence of R' on $W_{k(\zeta_l)}$. Thus we have to consider only the dependence of R' on μ through the free propagators:

$$\begin{aligned}\frac{\partial R'}{\partial\mu} &= \sum_{n=0}^{\infty}\sum_{k,l}\frac{\partial R'_{n+1}}{\partial G^0_{k(\zeta_l)}}\frac{\partial G^0_{k(\zeta_l)}}{\partial\mu}\,e^{\zeta_l O^+}\\ &= -\sum_{n=0}^{\infty}\sum_{k,l}\frac{\partial R'_{n+1}}{\partial G^0_{k(\zeta_l)}}(G^0_{k(\zeta_l)})^2 e^{\zeta_l O^+}\\ &= -\frac{1}{\beta}\sum_{n=0}^{\infty}\sum_{k,l}\Sigma'_{n+1\,k(\zeta_l)}(G^0_{k(\zeta_l)})^2 e^{\zeta_l O^+}\\ &= -\frac{1}{\beta}\sum_{k,l}\Sigma'_{k(\zeta_l)}(G^0_{k(\zeta_l)})^2 e^{\zeta_l O^+}. \qquad (25)\end{aligned}$$

Hence,

$$\begin{aligned}\langle N\rangle &= \frac{1}{\beta}\sum_{k,l}G^0_{k(\zeta_l)}\Sigma'_{k(\zeta_l)}e^{\zeta_l O^+} - \frac{\partial\Omega_0}{\partial\mu}\\ &= \frac{1}{\beta}\sum_{k,l}G_{k(\zeta_l)}e^{\zeta_l O^+},\end{aligned}$$

as expected. (In the above calculation, a subindex A is not mentioned, but it should be remarked that the calculation holds for properly approximated functionals as well.)

IV. THE SELF-CONSISTENT HARTREE AND HARTREE–FOCK APPROXIMATIONS AND THE SELF-CONSISTENT RPA

Special attention should be drawn to the case $k = \zeta_l = 0$, where broken F diagrams also appear, such as in Fig. 7(a). (They are, of course, linked!) Such diagrams should not be considered when the interaction line is dressed, but have to be taken into account when the self-consistent self-energy diagrams are formed.

If one desires to construct the self-consistent Hartree–Fock diagrams, one should consider, among others, an infinite set of broken F diagrams [e.g., Fig. 7(a), Parts (1)–(3)]. Examples of other naked F diagrams for Hartree–Fock approximation are given in Fig. 7(a), Parts (4)–(6). It is readily seen that when the interaction line is dressed, undesired diagrams, such as in Fig. 7(b), Part (1), appear.

Similarly, if one wishes to formulate a self-consistent RPA in the propagator formalism, one faces a similar difficulty, namely, the necessity for an infinite number of diagrams as a base and the appearance of undesired diagrams [e.g., Fig. 7(b), Part (2)] when an exact evaluation is made.

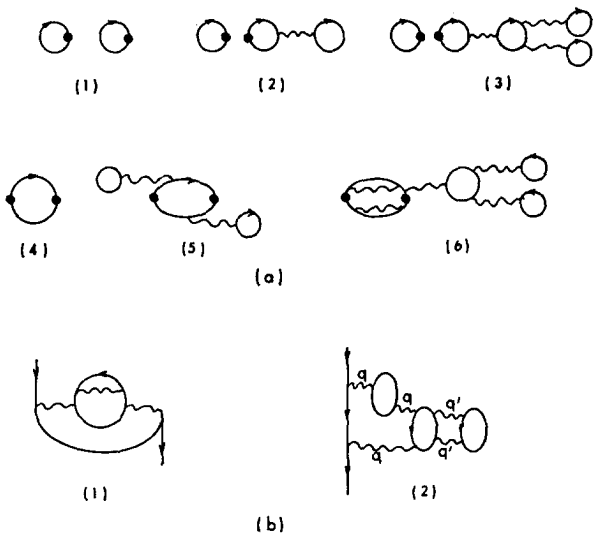


FIG. 7. (a) Several naked $F_{k(\zeta_l)}$ diagrams out of the infinite set required for the Hartree–Fock approximation. (b) (1) An undesired diagram which appears upon evaluating the Hartree–Fock approximation through the interaction-line formalism. (2) An undesired diagram which appears upon evaluating the RPA through the propagator formalism.

On the other hand, the interaction-line formalism sketched above is tailored for the RPA, as only a single R'_A diagram is responsible for it.

V. REMARKS AND CONCLUSIONS

The interaction line and the propagator line play similar roles in the formal construction and evaluation of diagrams. In this paper this similarity is utilized to formulate, in a diagrammatic manner, a variational principle of the vacuum–vacuum diagrams.

It is expected that propagator-based mathematical properties of diagrammatic functionals should correspond to analogous properties based on the interaction line.

If the bare interaction line is considered as a "propagator" of a free boson field (e.g., phonons) through the vertices, then the analogy is natural. In the case where the particles are electrons interacting via the Coulomb interaction, the wavy line may be regarded as the bare plasmon propagator. In those effects where the plasmons rather than the electrons play the significant role, it is suggested that the new approach to self-consistency should be preferred.

ACKNOWLEDGMENTS

We wish gratefully to acknowledge useful discussions with Dr. A. Casher, Professor C. G. Kuper, and Professor D. Lurie.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 12 DECEMBER 1969

Vacancy Annihilation for One-Dimensional Dumbbell Kinetics

R. B. McQuistan
University of Wisconsin–Milwaukee, Milwaukee, Wisconsin

(Received 3 March 1969)

Expressions are developed which describe the ensemble average of the annihilation of groups of contiguous vacant compartments when spatially random attempts are made to place dumbbells on a linear array of N compartments. It is shown that in the limit, as the number of compartments tends to infinity, $\langle\theta_p(t)\rangle$, the ensemble average of the fraction of the compartments which is composed of p contiguous vacant compartments, is given by

$$\langle\theta_p(t)\rangle = pe^{-2}\left(1 - \sum_{n=2}^{\infty} C_n e^{-nvt}\right)e^{-(p-1)vt},$$

where v is the striking frequency of the dumbbells, t is time, and the C_n's are appropriately defined coefficients.

I. INTRODUCTION

It was shown in a previous paper[1] that $\langle\theta(m, N)\rangle$, the ensemble average fraction of the compartments occupied after m spatially random attempts to place dumbbells on a linear array of N compartments, is given by

$$\langle\theta(m, N)\rangle = 1 - \left(\frac{N-1}{N}\right) \times \sum_{k=0}^{N-3} \frac{(-2)^{k-1}}{k!}\left(\frac{N-1-k}{N-1}\right)\left(\frac{N-2-k}{N-1}\right) \times \sum_{j=0}^{k}(-1)^j {}_kC_j\left(\frac{N-1-j}{N-1}\right). \quad (1)$$

In the limit as N tends to infinity the fraction of compartments occupied becomes

$$\theta(t) = 1 - \exp\{-2[1 - \exp(-vt)]\}. \quad (2)$$

The present paper is concerned with another aspect of this problem, namely, the kinetics of the annihilation of the p-tuple contiguous vacant compartments when dumbbells are placed on a linear array of compartments in a spatially random manner.

Our task is to determine $\langle\theta_p(m, N)\rangle$, the average fraction of the array which is composed of p-tuple contiguous compartments after m spatially random attempts to place dumbbells on a one-dimensional array of N compartments. A somewhat similar problem has been treated by Flory.[2]

II. CALCULATION OF $\langle\theta_p(m, N)\rangle$

We showed in a previous article[1] that the appropriate recursion relationship for $N_p(m, N)$, the number of p-tuple contiguous vacant compartments after m spatially random attempts to place indistinguishable dumbbells on a linear array of N compartments, is

$$N_p(m, N) = (N - p - 2)N_p(m - 1, N) + 2\sum_{h=p}^{N-2} N_h(m - 1, N) - 2N_{p+1}(m - 1, N)\sum_{h=1}^{N-3}\delta_{ph}, \quad (3)$$

where δ_{ph} is the Kronecker δ.

To facilitate computations it has been found that $N_p(m, N)$ may also be represented by

$$N_p(m, N) = (N - p)N_p(m - 1, N) + 2\sum_{h=p+2}^{N-2} N_h(m - 1, N), \quad (4)$$

with the restriction that either $p \geq N - 4$ or the summation term is neglected. If Eq. (4) is recurred r times, the result is

$$N_p(m, N) = (N - p)^r N_p(m - r, N) + 2(N - p)^{r-1} \times \sum_{h=p+2}^{N-2}\left[1 + \left(\frac{N-h}{N-p}\right)^1 + \left(\frac{N-h}{N-p}\right)^2 + \cdots + \left(\frac{N-h}{N-p}\right)^r\right] N_p(m - r, N) + 2^2(N - p)^{r-2}\sum_{h=p+2}^{N-4}\sum_{i=h+2}^{N-2}\left[1 + \left(\frac{N-h}{N-p}\right)^1 + \left(\frac{N-i}{N-p}\right)^1 + \left(\frac{N-h}{N-p}\right)^2 + \left(\frac{N-h}{N-p}\right)\left(\frac{N-i}{N-p}\right)\right.$$

[1] R. B. McQuistan and D. Lichtman, J. Math. Phys. **9**, 1680 (1968).
[2] P. J. Flory, J. Am. Chem. Soc. **61**, 1518 (1939).

$$+\left(\frac{N-i}{N-p}\right)^2+\cdots+\left(\frac{N-i}{N-p}\right)^{r-2}\Bigg]N_p(m-r,N)+2^3(N-p)^{r-3}$$
$$\times\sum_{h=p+2}^{N-6}\sum_{i=h+2}^{N-4}\sum_{j=i+2}^{N-2}\Bigg[1+\left(\frac{N-h}{N-p}\right)+\left(\frac{N-i}{N-p}\right)+\left(\frac{N-j}{N-p}\right)+\left(\frac{N-h}{N-p}\right)^2$$
$$+\left(\frac{N-h}{N-p}\right)\left(\frac{N-i}{N-p}\right)+\left(\frac{N-i}{N-p}\right)^2+\left(\frac{N-i}{N-p}\right)\left(\frac{N-j}{N-p}\right)+\left(\frac{N-j}{N-p}\right)^2+\cdots+\left(\frac{N-j}{N-p}\right)^{r-3}\Bigg]$$
$$\times N_p(m-r,N)+2^4(N-p)^{r-4}\cdots, \tag{5}$$

with the restriction that either the summation indices are nonnegative or the summation is neglected.

Since $N_p(1,N)=2$, Eq. (5) may be written, for $r=m-1$, as

$$N_p(m,N)=2(N-p)^{m-1}\Bigg\{1+2\sum_{h=p+2}^{N-2}\frac{1}{h-p}\Bigg[1-\left(\frac{N-h}{N-p}\right)^{m-1}\Bigg]$$
$$+2^2\sum_{h=p+2}^{N-4}\sum_{i=h+2}^{N-2}\frac{1}{(h-p)(i-p)}\Bigg[1-\left(\frac{i-p}{i-h}\right)\left(\frac{N-h}{N-p}\right)^{m-1}+\left(\frac{h-p}{i-h}\right)\left(\frac{N-i}{N-p}\right)^{m-1}\Bigg]$$
$$+2^3\sum_{h=p+2}^{N-6}\sum_{i=h+2}^{N-4}\sum_{j=i+2}^{N-2}\frac{1}{(h-p)(i-p)(j-p)}\Bigg[1-\left(\frac{i-p}{i-h}\right)\left(\frac{j-p}{j-h}\right)\left(\frac{N-h}{N-p}\right)^{m-1}$$
$$+\left(\frac{h-p}{i-h}\right)\left(\frac{j-p}{j-i}\right)\left(\frac{N-i}{N-p}\right)^{m-1}-\left(\frac{h-p}{j-h}\right)\left(\frac{i-p}{j-i}\right)\left(\frac{N-j}{N-p}\right)^{m-1}\Bigg]+2^4\cdots\Bigg\}. \tag{6}$$

To determine $\langle\theta_p(m,N)\rangle$ we multiply Eq. (6) by p and divide by $N(N-1)^m$; the result may be written in the form

$$\langle\theta_p(m,N)\rangle=\frac{2p}{N(N-1)}\left(\frac{N-p}{N-1}\right)^{m-1}\Bigg\{1+2\sum_{h=2}^{N-p-2}\frac{1}{h}\Bigg[1-\left(1-\frac{h}{N-p}\right)^{m-1}\Bigg]$$
$$+2^2\sum_{h=2}^{N-p-4}\sum_{i=h+2}^{N-p-2}\frac{1}{ih}\Bigg[1-\left(\frac{i}{i+h}\right)\left(1-\frac{h}{N-p}\right)^{m-1}+\left(\frac{h}{i+h}\right)\left(1-\frac{i}{N-p}\right)^{m-1}\Bigg]$$
$$+\cdots+2^3\sum_{h=2}^{N-p-6}\sum_{i=h+2}^{N-p-4}\sum_{j=i+2}^{N-p-2}\Bigg[1-\left(\frac{i}{i-h}\right)\left(\frac{j}{j-h}\right)\left(1-\frac{h}{N-p}\right)^{m-1}$$
$$+\left(\frac{h}{i-h}\right)\left(\frac{j}{j-i}\right)\left(1-\frac{i}{N-p}\right)^{m-1}+\left(\frac{h}{j-h}\right)\left(\frac{i}{j-i}\right)\left(1-\frac{j}{N-p}\right)^{m-1}\Bigg]$$
$$+2^4\sum_{h=2}^{N-p-8}\sum_{i=h+2}^{N-p-6}\sum_{j=i+2}^{N-p-4}\sum_{k=j+2}^{N-p-2}\frac{1}{hijk}\Bigg[1-\left(\frac{i}{i-h}\right)\left(\frac{j}{j-h}\right)\left(\frac{k}{k-h}\right)\left(1-\frac{h}{N-p}\right)^{m-1}$$
$$+\left(\frac{j}{j-i}\right)\left(\frac{k}{k-i}\right)\left(\frac{h}{i-h}\right)\left(1-\frac{i}{N-p}\right)^{m-1}-\left(\frac{k}{k-j}\right)\left(\frac{h}{j-h}\right)\left(\frac{i}{j-i}\right)\left(1-\frac{j}{N-p}\right)^{m-1}$$
$$+\left(\frac{h}{k-h}\right)\left(\frac{i}{k-i}\right)\left(\frac{j}{k-j}\right)\left(1-\frac{k}{N-p}\right)^{m-1}\Bigg]+2^5\cdots\Bigg\}. \tag{7}$$

To obtain $\theta_p(t)$, we consider $\langle\theta_p(m,N)\rangle$ to be expanded in the form

$$\langle\theta_p(m,N)\rangle=p\left(\frac{N-p}{N-1}\right)^{m-1}\sum_{n=0}^{\infty}a_n\left(1-\frac{n}{N-p}\right)^{m-1}. \tag{8}$$

Comparing the coefficients of $[1-n/(N-p)]^{m-1}$, $n=0,1,2,\cdots$, in Eqs. (7) and (8), we see that

$$a_0=\frac{2}{N(N-1)}\Bigg(1+2\sum_{h=2}^{N-p-2}\frac{1}{h}+2^2\sum_{h=2}^{N-p-4}\sum_{i=h+2}^{N-p-2}\frac{1}{hi}$$
$$+2^3\sum_{h=2}^{N-p-6}\sum_{i=h+2}^{n-p-4}\sum_{j=i+2}^{N-p-2}\frac{1}{hij}+\cdots\Bigg). \tag{9}$$

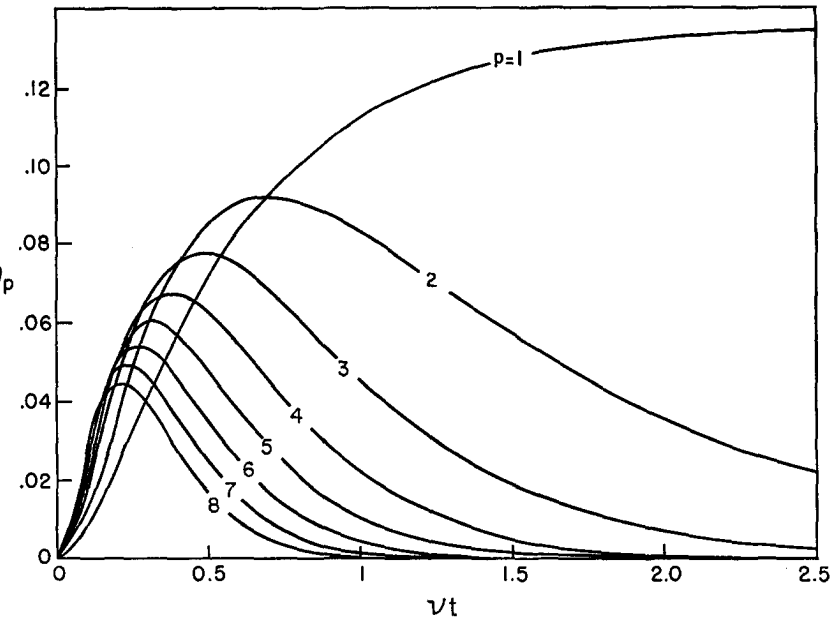


FIG. 1. θ_p (the fraction of the compartments consisting of p-tuple contiguous vacant compartments when dumbbells are placed in a spatially random manner on an infinite array of compartments), as a function of exposure, is given for various values of p.

It is known[1] that, for the isolated vacancies ($p = 1$) when N tends to infinity and $m = Nvt$, we may write

$$\lim_{N\to\infty} \langle \theta_1(\infty, N) \rangle = e^{-2}. \tag{10}$$

Thus, it is seen from Eq. (8) that

$$a_0 = e^{-2}. \tag{11}$$

In general, for $N \to \infty$ one may write, for $n \geq 2$,

$$a_n = \frac{-2a_0}{n}\left(1 - 2\sum_{h=2}^{n-2}\frac{1}{h} + 2^2\sum_{h=2}^{n-4}\sum_{i=h+2}^{n-2}\frac{1}{ih} - 2^3\sum_{h=2}^{n-6}\sum_{i=h+2}^{n-4}\sum_{j=i+2}^{n-2}\frac{1}{ijh} + 2^4\cdots\right), \tag{12}$$

where the first term is found by choosing $h = n$ in Eq. (7); the second term by choosing $i = n$ and $h = 2, 3, \cdots, n-2$; the third term by choosing $j = n$, $i = 4, 5, \cdots, n-4$, and $h = 2, 3, \cdots, n-6$, etc.

Thus, in the limit as $N - p$ tends to infinity, where $m = Nvt$ (v is the dumbbell striking frequency and t is time), Eq. (7) becomes

$$\theta_p(t) = pe^{-2}\left(1 - \sum_{n=2}^{\infty} C_n e^{-nvt}\right)e^{-(p-1)vt}. \tag{13}$$

These C_n in Eq. (13) are related to the a_n by

$$C_n = -a_n e^2, \quad n \geq 2.$$

For example,

$$C_2 = 1,$$
$$C_3 = \tfrac{2}{3},$$
$$C_4 = 0,$$
$$C_5 = -\frac{2^2}{(5)(3)},$$
$$C_6 = -\frac{2^2}{(6)(3)}.$$

In addition, the C_n satisfy the quasinormalization condition

$$\sum_{n=2}^{\infty} C_n = 1.$$

Figure 1 shows $\theta_p(t)$ for several values of p. It is seen that only the curve representing the fraction of compartments which are isolated ($p = 1$) is monotonically increasing, i.e., the isolated vacancies cannot be annihilated. In addition, those groups of vacancies with $p > 1$ initially show an increasing surface density but eventually are annihilated to form groups of vacancies of lower order.

ACKNOWLEDGMENTS

The author wishes to express his appreciation to D. Lichtman, T. Casselman, and F. N. Simon for many helpful suggestions.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 12 DECEMBER 1969

Canonical States in Quantum-Statistical Mechanics*

ROBERT E. KVARDA†
Department of Mathematics, Oregon State University, Corvallis, Oregon

(Received 14 April 1969)

The quantum-mechanical analog of the classical Gibbs canonical density is characterized by considering a large collection Q of noninteracting quantum systems, each in an equilibrium statistical state. The set Q, the Hamiltonian operator for each system, and the statistical states are assumed to have certain properties which are given as axioms. It is shown that these assumptions imply that each member of Q is in a canonical state at a temperature which is the same for all systems. The possibility of zero absolute temperature is included.

1. INTRODUCTION

The purpose of this paper is to present a characterization of the quantum-mechanical analog of the Gibbs canonical state in statistical mechanics. The approach is based on a method which considers composite mechanical systems composed of two mechanically and statistically independent components. We proceed in a way which parallels, as closely as possible, the usual measure—theoretic description of classical statistical mechanics. The development presented here differs from the classical approach in that it is not possible, in the most general case, to consider only a single two-component system. The problem then is to define in a mathematically precise way an appropriate collection Q of quantum systems, with each system in an equilibrium statistical state, such that the statistical state of each system in Q is defined by a Gibbs density operator corresponding to a uniform temperature.

2. FUNDAMENTAL ASSERTIONS OF QUANTUM-STATISTICAL MECHANICS

We denote a quantum-mechanical system by a pair $(\mathcal{H}, H)$, where $\mathcal{H}$ is a Hilbert space and H is the system Hamiltonian operator. Let $\mathcal{M}$ denote the set of all closed linear subspaces of $\mathcal{H}$. Following Mackey,[1] we define a statistical state of $(\mathcal{H}, H)$ as a probability measure on $\mathcal{M}$ in the sense of the following definition:

Definition 1: A probability measure on the closed linear subspaces of $\mathcal{H}$ is a function p which assigns to every closed subspace $M \subset \mathcal{H}$ a nonnegative real number $p(M)$ such that

(a) $p(\mathcal{H}) = 1, p(\{0\}) = 0$;

(b) $0 \leq p(M) \leq 1$ for all closed subspaces $M \subset \mathcal{H}$;

(c) if $\{M_i\}$ is a countable collection of mutually orthogonal subspaces having closed linear span M, then $p(M) = \sum_{i=1}^{\infty} p(M_i)$.

The following theorem, due to Gleason,[2] identifies each statistical state of $(\mathcal{H}, H)$ with a density operator in a one-to-one way:

Theorem 1: Let p be a probability measure on the closed subspaces of a separable (real or complex) Hilbert space of dimension at least three. Then there exists a unique, nonnegative, self-adjoint operator D of the trace class such that for all closed subspaces M of $\mathcal{H}$ we have $p(M) = \mathrm{Tr}\,(DP_M)$, where P_M is the orthogonal projection of $\mathcal{H}$ onto M.

Since D is of trace class, it is also completely continuous.[3] Being completely continuous and self-adjoint, it admits a complete orthonormal sequence of characteristic vectors. Its nonzero (necessarily real) characteristic values have finite multiplicities[4] and form either a finite or countably infinite sequence $\{\lambda_n\}$. A density operator therefore has the spectral representation

$$D = \sum_{n=1}^{\infty} \lambda_n P_n; \quad 0 \leq \lambda_n \leq 1 \quad \text{and} \quad \sum_{n=1}^{\infty} m_n \lambda_n = 1,$$

where P_n is the projection operator onto the characteristic subspace of $\mathcal{H}$ of dimension m_n corresponding to the characteristic value λ_n.

In the case of composite quantum systems, one is led to consider composite probability measures on the closed subspaces of a tensor product of Hilbert spaces. We denote the tensor product of $\mathcal{H}_1$ and $\mathcal{H}_2$ by $\mathcal{H}_1 \otimes \mathcal{H}_2$. If $\varphi \in \mathcal{H}_1$ and $\psi \in \mathcal{H}_2$, then their tensor product is denoted by $\varphi \times \psi$. The space $\mathcal{H}_1 \otimes \mathcal{H}_2$

* The material in this paper is discussed in greater detail in the author's doctoral dissertation. See R. E. Kvarda, "Canonical States in Quantum-Statistical Mechanics," Tech. Report No. 21, Dept. of Math., Oregon State University, Corvallis, Oregon. Work supported in part by a National Science Foundation Grant in Applied Analysis and by the Naval Electronics Laboratory Center, San Diego, Calif.

† Present address: Mathematics Research Group, Naval Weapons Laboratory, Dahlgren, Virginia 22448.

[1] G. W. Mackey, *Mathematical Foundations of Quantum Mechanics* (W. A. Benjamin, Inc., New York, 1963).

[2] A. M. Gleason, J. Math. Mech. 6, 885 (1953).

[3] R. Schatten, *Normed Ideals of Completely Continuous Operators* (Springer-Verlag, Berlin, 1960).

[4] F. Riesz and B. von Sz-Nagy, *Functional Analysis* (Frederick Ungar Publ. Co., New York, 1955).

is the completion of the set $(\mathcal{H}_1 \otimes \mathcal{H}_2)'$ consisting of all finite linear aggregates of the form

$$\Phi = \sum_{m=1}^{p} (\varphi_m \times \psi_m).$$

If $\Psi = \sum_{n=1}^{q} (\xi_n \times \eta_n)$ is in $\mathcal{H}_1 \otimes \mathcal{H}_2$, then the inner product (Φ, Ψ) is defined by

$$(\Phi, \Psi) = \sum_{m=1}^{p} \sum_{n=1}^{q} (\varphi_m, \xi_n)(\psi_m, \eta_n).$$

By defining the functional

$$(\varphi \times \psi)(\theta_1, \theta_2) = (\varphi, \theta_1)(\psi, \theta_2)$$

with $\varphi, \theta_1 \in \mathcal{H}_1$ and $\psi, \theta_2 \in \mathcal{H}_2$, every member of $(\mathcal{H}_1 \otimes \mathcal{H}_2)'$ may be identified with an antibilinear functional on the Cartesian product $\mathcal{H}_1 \times \mathcal{H}_2$. It is not difficult to show[5] that the norm convergence of a sequence from $(\mathcal{H}_1 \otimes \mathcal{H}_2)'$ implies pointwise convergence as antibilinear functionals on $\mathcal{H}_1 \times \mathcal{H}_2$ so that every member of $\mathcal{H}_1 \otimes \mathcal{H}_2$ may be identified with a certain antibilinear functional on $\mathcal{H}_1 \times \mathcal{H}_2$.

If M, N are closed linear subspaces in $\mathcal{H}_1$, $\mathcal{H}_2$, respectively, then the closure of $(M \otimes N)'$ is a closed linear subspace of $\mathcal{H}_1 \otimes \mathcal{H}_2$ and is denoted by $M \otimes N$. We denote by $\mathcal{M}_1 \otimes \mathcal{M}_2$ the set of all closed linear subspaces of $\mathcal{H}_1 \otimes \mathcal{H}_2$. The set of all subspaces of the form $M \otimes N$ is a proper subset of $\mathcal{M}_1 \otimes \mathcal{M}_2$.

Let D_1, D_2 be density operators on $\mathcal{H}_1$, $\mathcal{H}_2$, respectively. The operator $D_1 \times D_2$ which is the *tensor product* of D_1 and D_2 is defined as follows: If $\{\varphi_1, \cdots, \varphi_m\} \subset \mathcal{H}_1$ and $\{\psi_1, \cdots, \psi_m\} \subset \mathcal{H}_2$, then for finite m,

$$(D_1 \times D_2) \sum_{n=1}^{m} (\varphi_n \times \psi_n) = \sum_{n=1}^{m} (D_1\varphi_n \times D_2\psi_n).$$

It is easy to show that $D_1 \times D_2$ is a density operator having a unique extension to all of $\mathcal{H}_1 \otimes \mathcal{H}_2$. In particular, for projection operators, we have

$$P_M \times P_N = P_{M \otimes N}.$$

For brevity, we denote a quantum-mechanical system in a statistical state p by a triple $(\mathcal{H}, \mathcal{M}, p)$ and call it a *Q-space*. A Q-space then is a nondistributive probability structure which plays the same role in quantum statistics as a probability space in the measure—theoretic description of classical statistics. The Q-space $(\mathcal{H}_1 \otimes \mathcal{H}_2, \mathcal{M}_1 \otimes \mathcal{M}_2, p_{12})$ denotes a composite quantum system in a statistical state p_{12} composed of two components which may or may not be independent. In either case, the state p_{12} *induces* certain states p_1 in $(\mathcal{H}_1, H_1)$ and p_2 in $(\mathcal{H}_2, H_2)$ according to the definitions

$$p_1(M) = p_{12}(M \otimes \mathcal{H}_2), \quad p_2(N) = p_{12}(\mathcal{H}_1 \otimes N)$$

for all closed subspaces M, N. In general, the induced states p_1, p_2 do not determine p_{12} uniquely unless the components are *statistically independent*. This is given by the following existence theorem:

Theorem 2: Let $(\mathcal{H}_1, \mathcal{M}_1, p_1)$ and $(\mathcal{H}_2, \mathcal{M}_2, p_2)$ be two Q-spaces. Then there exists one and only one probability measure on $\mathcal{M}_1 \otimes \mathcal{M}_2$, denoted by $p_1 \otimes p_2$, such that

$$(p_1 \otimes p_2)(M \otimes N) = p_1(M) \cdot p_2(N)$$

for all closed subspaces of the form $M \otimes N$.

Theorem 2 follows at once from Theorem 1 and the following statement[6]:

Theorem 3: Let $(\mathcal{H}_1 \otimes \mathcal{H}_2, \mathcal{M}_1 \otimes \mathcal{M}_2, p_{12})$ be a composite system composed of the two components $(\mathcal{H}_1, \mathcal{M}_1, p_1)$ and $(\mathcal{H}_2, \mathcal{M}_2, p_2)$, and let D_{12}, D_1, and D_2 be their respective density operators. Then $D_{12} = D_1 \times D_2$ if and only if $p_{12}(M \otimes N) = p_1(M) \cdot p_2(N)$ for all closed subspaces $M \subset \mathcal{H}_1$ and $N \subset \mathcal{H}_2$.

To prove Theorem 3, the following lemma is needed:

Lemma 1: Let A be a self-adjoint operator of the trace class defined for all $\Phi \in \mathcal{H}_1 \otimes \mathcal{H}_2$. If

$$\mathrm{Tr}\,(AP_{M \otimes N}) = 0$$

for all projection operators of the form $P_{M \otimes N}$, where M, N are closed linear subspaces in $\mathcal{H}_1$, $\mathcal{H}_2$, respectively, then $A = 0$.

Proof: Let $[\varphi]$ and $[\psi]$ denote the one-dimensional subspaces spanned by the unit vectors φ and ψ. Then $\mathrm{Tr}\,(AP_{[\varphi] \otimes [\psi]}) = 0$ implies $(A(\varphi \times \psi), \varphi \times \psi) = 0$ for all $\varphi, \psi \in \mathcal{H}_1, \mathcal{H}_2$. Using the distributive property of tensor multiplication, one finds that

$$(A(\varphi_1 \times \psi_1), \varphi_1 \times \psi_2) = 0,$$
$$(A(\varphi_1 \times \psi_1), \varphi_2 \times \psi_2) = 0,$$

for all $\varphi_1, \varphi_2 \in \mathcal{H}_1$ and all $\psi_1, \psi_2 \in \mathcal{H}_2$. It follows that $(A\Phi, \Phi) = 0$ for all $\Phi \in (\mathcal{H}_1 \otimes \mathcal{H}_2)'$. By the continuity of the inner product, we have $(A\Phi, \Phi) = 0$ for all $\Phi \in \mathcal{H}_1 \otimes \mathcal{H}_2$. Hence $A = 0$.

Proof of Theorem 3: If $D_{12} = D_1 \times D_{12}$, then for all subspaces $M \subset \mathcal{H}$ and $N \subset \mathcal{H}_2$ we obtain

$$p_{12}(M \otimes N) = \mathrm{Tr}\,(D_1P_M) \cdot \mathrm{Tr}\,(D_2P_N).$$

[5] J. von Neumann and F. Murray, Ann. Math. 37(2), 116 (1936).

[6] Theorem 3 is presumably well known to the experts; however, we have been unable to find a reference to it in the literature.

Hence $p_{12}(M \otimes N) = p_1(M) \cdot p_2(N)$. On the other hand, if $p_{12}(M \otimes N) = p_1(M) \cdot p_2(N)$ for all M, N, then

$$p_{12}(M \otimes N) = \mathrm{Tr}\,(D_1 P_M) \cdot \mathrm{Tr}\,(D_2 P_N) = \mathrm{Tr}\,((D_1 \times D_2) P_{M \otimes N}).$$

Since p_{12} corresponds to D_{12}, we also have

$$p_{12}(M \otimes N) = \mathrm{Tr}\,(D_{12} P_{M \otimes N})$$

for all M, N. Therefore

$$\mathrm{Tr}\,([(D_1 \times D_2) - D_{12}] P_{M \otimes N}) = 0.$$

Applying the lemma, we obtain

$$D_{12} = D_1 \times D_2. \tag{2.1}$$

Theorem 3 may be extended to the case of a composite system composed of an arbitrary but finite number of components in a straightforward way. In the case of an infinite number of components, the technique of this paper breaks down since an infinite tensor product of Hilbert spaces (of dimension greater than one) is always nonseparable. However, in the next section we shall define a (possibly) nondenumerable collection of quantum systems for which Eq. (2.1) holds under the *pairwise* formation of composite systems. By placing appropriate conditions on the collection, the resulting family of operator equations may be solved to yield the canonical density operator.

3. THE AXIOMS

We shall view Q as a set in the mathematical sense. However, Q may be interpreted physically as a collection of systems, each pair of which can be brought together into equilibrium at a common temperature. We do not treat the mechanism for attaining equilibrium. We deal instead with the limiting case of zero interaction, and postulate that each component is in a limiting state p, independent of the second component of the composite pair. It is not necessary to make an explicit assumption regarding a common temperature for all systems. The temperature simply appears as a free parameter in the class of canonical states.

The axioms fall into three categories. The first axiom restricts the collection Q to systems whose Hamiltonian operators have pure point spectra, and the next three axioms ensure that Q is sufficiently large for our characterization to succeed. The last two axioms place natural restrictions on the equilibrium states.

Axiom 1: For each system in Q, the Hamiltonian operator H has a pure point spectrum $S(H)$ consisting of zero and a sequence of real numbers increasing to infinity:

$$0 = \lambda_0 < \lambda_1 < \lambda_2 < \cdots < \lambda_n < \cdots, \quad \lim_{n \to \infty} \lambda_n = \infty.$$

Since the Hamiltonian operator is arbitrary up to an additive constant, the requirement that the smallest characteristic value be zero merely requires that each Hamiltonian be normalized by a constant energy shift.

Let I_1, I_2 denote the identity operators on $\mathcal{H}_1$, $\mathcal{H}_2$, respectively.

Axiom 2: The set Q is closed under the pairwise formation of noninteracting composite systems. That is, if $(\mathcal{H}_1, H_1)$ and $(\mathcal{H}_2, H_2)$ are any two members of Q, then the composite system $(\mathcal{H}_1 \otimes \mathcal{H}_2, H_1 \times I_2 + I_1 \times H_2)$ also belongs to Q.

This axiom provides the mathematical counterpart for the physical assertion that each pair of systems can be brought into equilibrium with each other in the limit of zero interaction. Since H_1 and H_2 are unbounded operators, $H_1 \times I_2 + I_1 \times H_2$ is an unbounded essentially self-adjoint operator and has the spectrum $\{\lambda_{1m} + \lambda_{2n} : m, n = 1, 2, \cdots\}$, where $\{\lambda_{1m}\} = S(H_1)$ and $\{\lambda_{2n}\} = S(H_2)$. To obtain the self-adjoint operator corresponding to the energy of the composite system, one identifies the symbol $H_1 \times I_2 + I_1 \times H_2$ with its smallest closed extension. It follows that the set $\mathfrak{D}$ consisting of all Hamiltonian spectral values for systems in Q is closed under addition. The next axiom ensures that $\mathfrak{D}$ is closed under positive differences.

Axiom 3: Let $\mathfrak{D}$ be the union of the Hamiltonian spectra for all systems in Q. Then $\mathfrak{D}$ is closed under subtraction in the sense that if $\lambda_1, \lambda_2 \in \mathfrak{D}$ and $\lambda_2 > \lambda_1$, then $\lambda_2 - \lambda_1 \in \mathfrak{D}$.

The next axiom requires that Q contains certain systems which, for convenience, we call "harmonic oscillators."

Axiom 4: For each $\lambda \in \mathfrak{D}$, $\lambda \neq 0$, there exists a system $(\mathcal{H}, H)$ in Q such that $S(H) = \{n\lambda : n = 0, 1, 2, \cdots\}$.

Axiom 5: Let $(\mathcal{H}, H)$ be a system in Q in the equilibrium state p, and let D be the density operator for this state. Then there is a function $f(\lambda)$ defined on $S(H)$ such that $D = f(H)$.

This axiom ensures that D commutes with the dynamical group $\{T_t = \exp(-iHt): t \in R\}$ and therefore qualifies as a true integral of the motion.

In view of Axiom 1, each Hamiltonian has a spectral representation $H = \sum_{n=0}^{\infty} \lambda_n P_n$, where P_n is the projection onto the characteristic subspace of $\mathcal{H}$ corresponding to the characteristic value λ_n. If f is any function defined on $S(H)$, then by the spectral theorem one may write $f(H) = \sum_{n=0}^{\infty} f(\lambda_n) P_n$. Since $\mathrm{Tr}\, D = 1$, every function f defined by Axiom 5 must satisfy

$$0 \leq f(\lambda_n) \leq 1 \quad \text{for} \quad n = 0, 1, 2, \cdots,$$

$$\sum_{n=0}^{\infty} m_n f(\lambda_n) = 1,$$

where m_n is the (necessarily finite) multiplicity of λ_n.

The final axiom presents the main statistical assumption. It asserts that the mechanical independence of any pair of systems in Q implies their statistical independence also.

Axiom 6: For a composite system $(\mathcal{H}_{12}, H_{12})$ composed of two noninteracting components $(\mathcal{H}_1, H_1)$ and $(\mathcal{H}_2, H_2)$ in Q, the component systems are statistically independent; that is, their statistical states satisfy $p_{12}(M \otimes N) = p_1(M) \cdot p_2(N)$ for all closed subspaces $M \subset \mathcal{H}_1$ and $N \subset \mathcal{H}_2$.

It should be stressed that the axioms do not determine the set Q uniquely. For example, Q might consist of only those systems built up by composition from a single harmonic oscillator, in which case all systems in Q would have the same spectrum $\mathfrak{D} = \{0, \lambda, 2\lambda, \cdots\}$. Therefore, it is important to know that any collection of systems Q_0 which we may wish to consider can be enlarged to obtain a collection Q satisfying Axioms 1 through 4, provided only that Q_0 satisfied Axiom 1.

Theorem 4: Let Q_0 be any collection of systems satisfying Axiom 1. Then there exists a second collection Q which contains Q_0 and satisfies Axioms 1 through 4.

Proof: Let $\mathfrak{D}_0$ be the union of all Hamiltonian spectra of systems in Q_0. Enlarge $\mathfrak{D}_0$ to obtain the set

$$\mathfrak{D}_1 = \{n_1\lambda_1 + \cdots + n_k\lambda_k : k = 1, 2, \cdots;\ n_j = 0, \pm 1, \pm 2, \cdots, \text{for all } j = 1, 2, \cdots k;\ \lambda_j \in \mathfrak{D}_0\}.$$

Let $\mathfrak{D}_+ = \{\lambda: \lambda \in \mathfrak{D}_1, \lambda \geq 0\}$. For each $\lambda \in \mathfrak{D}_+$, $\lambda \neq 0$, adjoin to Q_0 the harmonic oscillator whose Hamiltonian spectrum is $\{n\lambda: n = 0, 1, 2, \cdots\}$ and denote the enlarged collection by Q_1. Let Q_1' be the collection obtained by noninteracting pairwise composition of members from Q_1 and define $Q_2 = Q_1 \cup Q_1'$. Now, for each integer $m \geq 1$, let Q_m' be the collection obtained by noninteracting pairwise composition of members from Q_m and define $Q_{m+1} = Q_m \cup Q_m'$. Then $Q = \bigcup_{m=1}^{\infty} Q_m$ satisfies Axioms 1 through 4.

4. THE FUNCTIONAL EQUATION

It follows from Axiom 6 and Theorem 3 that to each pair of systems $(\mathcal{H}_1, H_1)$ and $(\mathcal{H}_2, H_2)$ in Q there corresponds an operator equation $D_{12} = D_1 \times D_2$, or equivalently,

$$f_{12}(H_1 \times I_2 + I_1 \times H_2) = f_1(H_1) \times f_2(H_2) \quad (4.1)$$

everywhere in $\mathcal{H}_1 \otimes \mathcal{H}_2$. From the spectral expansion of both sides of (4.1) one obtains

$$f_{12}(\lambda_1 + \lambda_2) = f_1(\lambda_1) \cdot f_2(\lambda_2) \quad (4.2)$$

for all λ_1, $\lambda_2 \in S(H_1)$, $S(H_2)$. By considering every pairwise combination of systems in Q, one obtains a system of functional equations

$$f_{mn}(\lambda_m + \lambda_n) = f_m(\lambda_m) \cdot f_n(\lambda_n), \quad (4.3)$$

which holds for all $\lambda_m \in S(H_m)$ and all $\lambda_n \in S(H_n)$. The subscripts index the systems in Q and range over a possibly nondenumerable set. Our objective is to solve this system of equations to obtain the equilibrium density operator for each member of Q.

We first note that if f is any function specified by Axiom 5, then $f(0) \neq 0$. Consider first the harmonic oscillators of Axiom 4. For two identical oscillators one obtains

$$f_{12}(n\lambda + m\lambda) = f(n\lambda) \cdot f(m\lambda).$$

Consequently, if $f(0) = 0$, then f_{12} is zero everywhere on $S(H_{12})$, which contradicts the fact that $\mathrm{Tr}\, D_{12} = 1$. If $f(0) = 0$ for an arbitrary system $(\mathcal{H}, H)$ with $D = f(H)$, there exists a $\lambda' \in S(H)$, $\lambda' > 0$, such that $f(\lambda') \neq 0$. By considering the composite system composed of $(\mathcal{H}, H)$ and the harmonic oscillator $(\mathcal{H}', H')$ having $S(H') = \{n\lambda': n = 0, 1, 2, \cdots\}$ and $D' = f'(H')$, one obtains $f_{12}(\lambda') = f'(0) \cdot f(\lambda') = f'(\lambda') \cdot f(0)$. If $f(0) = 0$, then $f'(\lambda') \cdot f(0)$ is zero but $f'(0) \cdot f(\lambda')$ is not, which again is a contradiction.

The fact that $f(0) \neq 0$ for every system in Q allows one to replace the system of equations (4.3) by a single functional equation.

Lemma 2: Consider the set $\mathfrak{D}$ which is the union of all Hamiltonian spectra for all systems in Q. There exists a function F, defined on $\mathfrak{D}$, satisfying the functional equation

$$F(\lambda_1 + \lambda_2) = F(\lambda_1) \cdot F(\lambda_2) \quad (4.4)$$

for all $\lambda_1, \lambda_2 \in \mathfrak{D}$ and the initial condition $F(0) = 1$ such that for each system $(\mathfrak{K}, H)$ in Q, with equilibrium density operator $D = f(H)$, the function f satisfies

$$f(\lambda) = f(0) \cdot F(\lambda) \tag{4.5}$$

for all $\lambda \in S(H)$.

Proof: By considering two systems in Q with $D_1 = f_1(H_1)$ and $D_2 = f_2(H_2)$ which may have a characteristic value λ common to both $S(H_1)$ and $S(H_2)$, one obtains from (4.2) $f_1(\lambda)/f_1(0) = f_2(\lambda)/f_2(0)$. Consequently, (4.5) defines F uniquely on $\mathfrak{D}$. To obtain (4.4), substitute (4.5) into (4.2) and use (4.2) with $\lambda_1 = \lambda_2 = 0$.

Lemma 3: $0 \leq F(\lambda) < 1$ for all positive $\lambda \in \mathfrak{D}$.

Proof: Consider the harmonic oscillator having $S(H) = \{n\lambda : n = 0, 1, 2, \cdots\}$. Then, by induction from (4.4), one obtains $F(n\lambda) = [F(\lambda)]^n$. Since

$$\operatorname{Tr} D = \sum_{n=0}^{\infty} m_n f(n\lambda) \geq f(0) \sum_{n=0}^{\infty} [F(\lambda)]^n,$$

we must have $0 \leq F(\lambda) < 1$.

The next lemma defines the canonical state of zero absolute temperature, or equivalently, the microcanonical state of zero energy. It is convenient to include this limiting state in the family of canonical states. We note that the canonical density operator corresponding to an absolute temperature T is

$$D_T = \frac{e^{-H/kT}}{\operatorname{Tr} e^{-H/kT}} = \frac{\sum_{n=0}^{\infty} e^{-\lambda_n/kT} P_n}{\sum_{n=0}^{\infty} m_n e^{-\lambda_n/kT}}.$$

Since we require $\lambda_0 = 0$, we have

$$\lim_{T \to 0} D_T = D_0 = \frac{1}{m_0} P_0,$$

where P_0 is the projector of $\mathfrak{K}$ onto the null space of H.

Lemma 4: If $F(\lambda) = 0$ for some $\lambda > 0$, then $F(\lambda) = 0$ for all $\lambda > 0$ and each system in Q is in its canonical state of zero absolute temperature.

Proof: The proof depends on the fact that $\mathfrak{D}$ is closed under positive differences. Let $F(\lambda_1) = 0$ and let λ_2 be any positive number in $\mathfrak{D}$.

Case I: If $\lambda_2 > \lambda_1$, then $\lambda_2 - \lambda_1 \in \mathfrak{D}$ and $F(\lambda_2) = F(\lambda_1) \cdot F(\lambda_2 - \lambda_1)$. Hence $F(\lambda_2) = 0$.

Case II: If $\lambda_2 < \lambda_1$, then, for some positive integer n, we have $n\lambda_2 > \lambda_1$ and by Case I we have $F(n\lambda_2) = [F(\lambda_2)]^n = 0$. Hence $F(\lambda_2) = 0$. Consequently, $F(\lambda) = 0$ for all $\lambda > 0$, and if $(\mathfrak{K}, H)$ is any system in Q with equilibrium density operator $D = f(H)$, then

$$f(\lambda) = \frac{1}{m_0}, \quad \text{if} \quad \lambda = \lambda_0 = 0,$$
$$f(\lambda) = 0, \quad \text{if} \quad \lambda \neq 0.$$

This implies $D = D_0 = (1/m_0)P_0$.

The next lemma considers the alternative to Lemma 4. In this case, $F(0) = 1$ and $0 < F(\lambda) < 1$ for all positive $\lambda \in \mathfrak{D}$. It is convenient to define the function $y(\lambda) = \ln F(\lambda)$ and work instead with the functional equation

$$y(\lambda_1 + \lambda_2) = y(\lambda_1) + y(\lambda_2). \tag{4.6}$$

By defining $y(-\lambda) = -y(\lambda)$, it is straightforward to show that every such extended solution satisfies (4.6) for all λ_1, λ_2 in the extended domain $\mathfrak{D} \cup -\mathfrak{D}$ and that

$$y(m\lambda) = my(\lambda) \tag{4.7}$$

for any integer m and all $\lambda \in \mathfrak{D} \cup -\mathfrak{D}$. We may now prove the following:

Lemma 5: Consider the case where $0 < F(\lambda) < 1$ for all positive $\lambda \in \mathfrak{D}$. Then there exists a positive constant θ such that $F(\lambda) = \exp(-\theta\lambda)$ and each system in Q is in the Gibbs canonical state with the density operator

$$D = e^{-\theta H}/\operatorname{Tr} e^{-\theta H}.$$

Proof: We prove that $y(\lambda)/\lambda = -\theta$ for all positive $\lambda \in \mathfrak{D}$. Suppose that $y(\lambda)/\lambda$ is not a constant. Then, for some $\lambda_1, \lambda_2 \in \mathfrak{D}$ and any two positive numbers a and b, the equations

$$\alpha y(\lambda_1) + \beta y(\lambda_2) = a,$$
$$\alpha\lambda_1 + \beta\lambda_2 = b$$

can be solved for the real coefficients α and β. Moreover, we can always find rational numbers r_1/s_1, r_2/s_2 with $s_1, s_2 > 0$, such that

$$(r_1/s_1)y(\lambda_1) + (r_2/s_2)y(\lambda_2) > 0,$$
$$(r_1/s_1)\lambda_1 + (r_2/s_2)\lambda_2 > 0.$$

Multiplying by $s_1 s_2$ and using (4.6) and (4.7), these inequalities become

$$y(m_1\lambda_1 + m_2\lambda_2) > 0,$$
$$m_1\lambda_1 + m_2\lambda_2 > 0,$$

where m_1 and m_2 are integers. This contradicts the fact

that $0 < F(\lambda) < 1$ for all positive $\lambda \in \mathfrak{D}$. Hence $y(\lambda)/\lambda$ is a constant and the result follows.

We summarize the results of Lemmas 4 and 5 with the following theorem:

Theorem 5: Let Q be any collection of quantum systems, each in an equilibrium statistical state, satisfying Axioms 1 through 6. Then each system is in a Gibbs canonical state at some absolute temperature $T \geq 0$. The temperature is arbitrary, but it is the same for all systems in Q.

ACKNOWLEDGMENT

The author wishes to express his gratitude to Professor David S. Carter, who suggested the problem and who directed the dissertation on which this paper is based.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 12 DECEMBER 1969

Infinite Renormalization of the Hamiltonian Is Necessary

JAMES GLIMM*
Courant Institute of Mathematical Sciences, New York University, New York, New York
AND
ARTHUR JAFFE†
Lyman Laboratory of Physics, Harvard University, Cambridge, Massachusetts

(Received 28 April 1969)

We show that the unrenormalized Hamiltonian in quantum field theory is unbounded from below whenever lowest-order perturbation theory indicates that this is true. We conclude that perturbation theory is an accurate guide to the divergence of the vacuum energy in quantum field theory.

In this paper we show that the unrenormalized Hamiltonian is unbounded from below when lowest-order perturbation theory predicts that this is true.

The proof is a simple calculation. The unrenormalized Hamiltonian H is a densely defined bilinear form on Fock space.[1] We choose a sequence Ω_κ of unit vectors in the domain of H. As $\kappa \to \infty$, the expectation values of the Hamiltonian in the vectors Ω_κ tends to $-\infty$. The ground state of H given by first-order perturbation theory motivates our choice of the vectors Ω_κ.

We concentrate on the boson self-interaction Φ^{2n} in $(s+1)$-dimensional space–time. The second-order vacuum energy has a momentum divergence for $s \geq 2$, $n \geq 2$ and for $s \geq 3$, $n = 1$, and these cases will be treated first. Afterwards, the remaining cases (which have a volume divergence) will be treated.

The Hamiltonian that we study is

$$H = H_0 + \lambda \int :\Phi(x)^{2n}: \, dx = H_0 + \lambda H_I. \tag{1}$$

The methods and the results hold equally for the spatially cut-off Hamiltonian:

$$H(g) = H_0 + \lambda \int :\Phi(x)^{2n}: g(x)\, dx, \quad 0 \leq g(x) \leq 1. \tag{2}$$

We work on the Fock space; H_0 is the free-field Hamiltonian for mass $m > 0$, and Φ is the standard boson field.

We study first the cases $s \geq 2$, $n \geq 2$ and $s \geq 4$, $n \geq 1$. The vectors Ω_κ are defined by

$$\Omega_\kappa = c(\psi_0 - \lambda \psi_{2n}),$$

where ψ_0 is the Fock no-particle vector and ψ_{2n} is a $2n$-particle vector. The constant c is chosen so that $\|\Omega_\kappa\| = 1$, and

$$\psi_{2n}(k_1, \cdots, k_{2n}) = \left(\sum_{i=1}^{2n} \mu(k_i)\right)^{-1} ((2n)!)^{\frac{1}{2}} h\left(\sum_{i=1}^{2n} k_i\right) \prod_{i=1}^{2n} [\mu(k_i)^{\tau - \frac{1}{2}} \chi_\kappa(k_i)]. \tag{3}$$

Here $\mu(k) = (k^2 + m^2)^{\frac{1}{2}}$ and h is a smooth, positive, rapidly decreasing function. The function $\chi_\kappa(k)$ equals unity if $|k| \leq \kappa$; it equals zero otherwise. We choose

$$\tau = -(s-1)\left(1 - \frac{1}{2n}\right) + \frac{\epsilon + 2}{2n},$$

* Supported by the New York Science and Technology Foundation, Grant SSF-(8)-8.
† Alfred P. Sloan Foundation Fellow. Supported in part by the Air Force Office of Scientific Research, AF49(638)-1380.
[1] A. Galindo Proc. Natl. Acad. Sci. (U.S.) **48**, 1128 (1962).

that $0 < F(\lambda) < 1$ for all positive $\lambda \in \mathfrak{D}$. Hence $y(\lambda)/\lambda$ is a constant and the result follows.

We summarize the results of Lemmas 4 and 5 with the following theorem:

Theorem 5: Let Q be any collection of quantum systems, each in an equilibrium statistical state, satisfying Axioms 1 through 6. Then each system is in a Gibbs canonical state at some absolute temperature $T \geq 0$. The temperature is arbitrary, but it is the same for all systems in Q.

ACKNOWLEDGMENT

The author wishes to express his gratitude to Professor David S. Carter, who suggested the problem and who directed the dissertation on which this paper is based.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 12 DECEMBER 1969

Infinite Renormalization of the Hamiltonian Is Necessary

JAMES GLIMM*
Courant Institute of Mathematical Sciences, New York University, New York, New York
AND
ARTHUR JAFFE†
Lyman Laboratory of Physics, Harvard University, Cambridge, Massachusetts

(Received 28 April 1969)

We show that the unrenormalized Hamiltonian in quantum field theory is unbounded from below whenever lowest-order perturbation theory indicates that this is true. We conclude that perturbation theory is an accurate guide to the divergence of the vacuum energy in quantum field theory.

In this paper we show that the unrenormalized Hamiltonian is unbounded from below when lowest-order perturbation theory predicts that this is true.

The proof is a simple calculation. The unrenormalized Hamiltonian H is a densely defined bilinear form on Fock space.[1] We choose a sequence Ω_κ of unit vectors in the domain of H. As $\kappa \to \infty$, the expectation values of the Hamiltonian in the vectors Ω_κ tends to $-\infty$. The ground state of H given by first-order perturbation theory motivates our choice of the vectors Ω_κ.

We concentrate on the boson self-interaction Φ^{2n} in $(s+1)$-dimensional space–time. The second-order vacuum energy has a momentum divergence for $s \geq 2$, $n \geq 2$ and for $s \geq 3$, $n = 1$, and these cases will be treated first. Afterwards, the remaining cases (which have a volume divergence) will be treated.

The Hamiltonian that we study is

$$H = H_0 + \lambda \int :\Phi(x)^{2n}: \, dx = H_0 + \lambda H_I. \tag{1}$$

The methods and the results hold equally for the spatially cut-off Hamiltonian:

$$H(g) = H_0 + \lambda \int :\Phi(x)^{2n}: g(x)\, dx, \quad 0 \leq g(x) \leq 1. \tag{2}$$

We work on the Fock space; H_0 is the free-field Hamiltonian for mass $m > 0$, and Φ is the standard boson field.

We study first the cases $s \geq 2$, $n \geq 2$ and $s \geq 4$, $n \geq 1$. The vectors Ω_κ are defined by

$$\Omega_\kappa = c(\psi_0 - \lambda \psi_{2n}),$$

where ψ_0 is the Fock no-particle vector and ψ_{2n} is a $2n$-particle vector. The constant c is chosen so that $\|\Omega_\kappa\| = 1$, and

$$\psi_{2n}(k_1, \cdots, k_{2n}) = \left(\sum_{i=1}^{2n} \mu(k_i)\right)^{-1} ((2n)!)^{\frac{1}{2}} h\left(\sum_{i=1}^{2n} k_i\right) \prod_{i=1}^{2n} [\mu(k_i)^{\tau - \frac{1}{2}} \chi_\kappa(k_i)]. \tag{3}$$

Here $\mu(k) = (k^2 + m^2)^{\frac{1}{2}}$ and h is a smooth, positive, rapidly decreasing function. The function $\chi_\kappa(k)$ equals unity if $|k| \leq \kappa$; it equals zero otherwise. We choose

$$\tau = -(s-1)\left(1 - \frac{1}{2n}\right) + \frac{\epsilon + 2}{2n},$$

* Supported by the New York Science and Technology Foundation, Grant SSF-(8)-8.
† Alfred P. Sloan Foundation Fellow. Supported in part by the Air Force Office of Scientific Research, AF49(638)-1380.
[1] A. Galindo Proc. Natl. Acad. Sci. (U.S.) **48**, 1128 (1962).

where ϵ is in the interval $0 < \epsilon < \frac{1}{2}$. We remark that the choices $h = \delta$, $\kappa = \infty$, $\tau = 0$ would give the ground state in first-order perturbation theory. With the above restrictions on n, s, and ϵ, we have $\tau < -\frac{1}{8}$. Furthermore,

$$\epsilon = 2n\tau + (2n - 1)s - (2n + 1).$$

Theorem 1:

$$\lim_{\kappa \to \infty} (\Omega_\kappa, H\Omega_\kappa) = -\infty.$$

Proof: We compute the inner product as

$$(\Omega_\kappa, H\Omega_\kappa) = c^2\{P_0 + \lambda P_1 + \lambda^2 P_2 + \lambda^3 P_3\}.$$

In this expansion it is easy to see that

$$P_0 = P_1 = 0$$

and that

$$P_2 = (\psi_{2n}, H_0\psi_{2n}) - (\psi_0, H_I\psi_{2n}) - (\psi_{2n}, H_I\psi_0).$$

The proof is completed by showing that, for large κ,

$$c^{-2} = 1 + \lambda^2 \|\psi_{2n}\|^2 = O(1),$$
$$(\psi_{2n}, H_0\psi_{2n}) = O(1),$$
$$P_3 = O(1),$$

and that for some positive constant D,

$$D\kappa^\epsilon \leq (\psi_0, H_I\psi_{2n}) = (\psi_{2n}, H_I\psi_0).$$

These orders of growth are established by standard power-counting arguments.[2] This completes the proof. The remaining case $n = 1$, $s = 3$ is handled by similar methods, modified to deal with a logarithmic divergence.

In the cases not covered by Theorem 1, perturbation theory predicts no momentum divergence. Thus, when g has compact support, perturbation theory predicts that $H(g)$ is bounded from below. This lower bound has been proved rigorously.[3] Perturbation theory predicts that H has a vacuum-energy divergence which is linear in the volume, and thus it predicts a lower bound for $H(g)$ which is linear in the volume. (The "volume" here is the area of the support of g.)

It is known that the true bounds are no worse than this prediction. Thus, for $s = 1$ or $s = 2$, $n = 1$, the lower bound diverges no faster than a constant times the volume.[3] We now show that, for $s = 1$, the H defined in Eq. (1) is unbounded from below. The same proof shows that the lower bound on $H(g)$ in Eq. (2) tends to $-\infty$ as $g \to 1$, and similar results hold for the case $s = 2$, $n = 1$.

Let

$$h_V(k) = Vh(kV).$$

In the definition (3) for ψ_{2n}, we substitute h_V for h and set $\tau = 0$, $\kappa = \infty$. We define

$$\Omega_V = \psi_0 - V^{-\frac{1}{2}}\lambda\psi_{2n}.$$

As before, one proves the following theorem:

Theorem 2:

$$\lim_{V \to \infty} (\Omega_V, H\Omega_V) = -\infty,$$

$$\limsup_{V \to \infty} \|\Omega_V\|^2 < \infty.$$

We conclude that perturbation theory is an accurate guide to the divergence of the vacuum energy in quantum field theory.

[2] S. Weinberg, Phys. Rev. **118**, 839 (1960).

[3] For the case $s = 1$, any n, see E. Nelson in *Mathematical Theory of Elementary Particles* (M.I.T. Press, Cambridge, Mass., 1966); J. Glimm, Commun. Math. Phys. **8**, 12 (1968); J. Glimm and A. Jaffe (to be published). The case $s = 2$, $n = 1$ can be computed explicitly or estimated.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 12 DECEMBER 1969

Relativistic Effects of Strong Binding on Slow Particles*

AMNON KATZ
Physics Department, Weizmann Institute, Rehovoth, Israel

(Received 29 March 1969)

The relativistic effect of strong binding on the equations of motion of slow particles is derived by taking the appropriate limit of classical relativistic equations of motion of interacting particles. The expected effect on the total mass of the system is verified. The relative motion is also affected—in a model-dependent way.

1. THE MASS OF BINDING ENERGY

Einstein derived his famous formula[1]

$$E = mc^2 \tag{1.1}$$

by kinematic arguments showing that the emission of electromagnetic radiation carrying energy E reduces the mass of the emitting system by E/c^2. The conservation of energy then leads to the general relation between energy and mass, Eq. (1.1). This argument cleverly evades the problem of actually computing the energy of a relativistic system of particles. In fact, most theories are not able to predict the total energy of a system of particles without encountering infinite self-energies.

One common application of Eq. (1.1) is to a system of slowly moving particles. One uses the Newtonian equations of motion and separates them nonrelativistically into center-of-mass motion and relative motion. One then uses the energy of the relative motion to correct the total mass via Eq. (1.1) in violation of the equation for the center-of-mass motion. Most calculations of nuclear structure start from this rather arbitrary premise, namely, that the defect in the total nuclear mass is not accompanied by any change in the equations of the relative motion.[2]

Recently, an attempt was made to extend this approach to the quark model,[3] i.e., to the systems of quarks that supposedly made up nucleons and mesons. In this case the effect of the relativistic connection between mass and energy became quite dramatic, since the mass defect was almost equal to the total original mass. A slight change in the parameters (or in the number of quarks grouped together) would send the mass of the system through zero and into the negative range. Still, the equations of relative motion were assumed unchanged.

No experimental verification of the assumptions of the quark model is probable short of the discovery of quarks. Also in the nuclear case, the connection of theory and experiment is not quite direct. A theoretical study thus seems indicated.

The natural way to verify a nonrelativistic approximation is to derive it from the full relativistic theory. This could not be done until recently for want of a suitable relativistic theory of interacting particles. The only theory free of infinite self-energies was the Fokker–Tetrode[4,5] formulation of classical electrodynamics in terms of action at a distance. This theory was considered useless because it led to half-retarded–half-advanced interactions. In any case, neither this form of electrodynamics nor related formulations of meson theory[6] allow the coexistence of slow motion with strong binding by virtue of the virial theorem (see below).

Recently, Van Dam and Wigner[7] proposed a classical dynamics of interacting particles that is a natural generalization of the Fokker–Tetrode theory but features an arbitrary function characterizing the interaction. A corresponding "scalar" theory has been proposed by this author.[8] Either of these theories reduces in the classical limit to Newton's equations with an interparticle potential depending only on distance. A situation of slow motion and strong binding may be arranged by adjusting the arbitrary function in the relativistic theory so as to make the limiting classical potential a deep well with a flat bottom. This may be done in either the Van Dam–Wigner theory or the scalar theory. In this paper we use a linear combination of both to demonstrate that the effect of

* This research has been sponsored in part by the Air Force Office of Scientific Research through the European Office of Aerospace Research, OAR, U.S. Air Force, under Contract F-61052-68-C-0070.

[1] A. Einstein, Ann. Physik **17**, (1905) [English transl.: *The Principle of Relativity* (Dover Publications, Inc., New York, 1923), p. 67].

[2] E.g., K. A. Brueckner in *The Many-Body Problem, Notes of the 1958 Les Houches Summer School* (John Wiley & Sons, New York, 1959).

[3] G. Morpurgo, Physics **2**, 95 (1965).

[4] H. Tetrode, Z. Physik **10**, 317 (1922); A. D. Fokker, Z. Physik **58**, 386 (1929); Physica **9**, 33 (1929); **12**, 145 (1932).

[5] J. A. Wheeler and R. P. Feynman, Rev. Mod. Phys. **21**, 425 (1949).

[6] P. Havas, Phys. Rev. **87**, 309 (1952).

[7] H. Van Dam and E. P. Wigner, Phys. Rev. **138**, B1576 (1965).

[8] A. Katz, J. Math. Phys. **10**, 1929 (1969).

strong binding on the relative motion is model-dependent.

The relativistic theory is introduced in Sec. 2. The notation is that of Ref. 8 and the results of that reference are recalled. Section 3 explains the limit of slow motion and strong binding. Like the ordinary classical limit, it involves letting c tend to infinity. However, certain interaction terms are assumed of order c^2. In Sec. 4 the limiting procedure is applied to the relativistic equation of motion. The limiting equations of motion [Eqs. (5.1), (5.2)] are discussed and separated into center-of-mass and relative equations in Sec. 5. The limits of validity of the limiting procedure are discussed in Sec. 7. It turns out that the case in which the contribution of the binding energy completely cancels the rest mass cannot be discussed within the present scheme. The relativistic dynamics we use is not the most general possible. In Appendix A we refer to a different form of dynamics, which, however, becomes identical with ours in the limit of slow motion but strong binding.

The limiting equations of motion [Eqs. (5.1), (5.2)] are the main result of the present paper.

2. RELATIVISTIC DYNAMICS AND CLASSICAL LIMIT

Consider a system of classical relativistic particles. Let $x_i(\tau_i)$ represent the world line of the ith particle (x_i is a four-vector and τ_i a parameter). We derive our dynamical equations from the action

$$A = m_1c^2\int(\dot{x}_1^2)^{\frac{1}{2}}\,d\tau_1 + m_2c^2\int(\dot{x}_2^2)^{\frac{1}{2}}\,d\tau_2 - \tfrac{1}{2}\int\chi((x_1-x_2)^2)\dot{x}_1\cdot\dot{x}_2\,d\tau_1\,d\tau_2 - \tfrac{1}{2}\int\psi((x_1-x_2)^2)(\dot{x}_1^2)^{\frac{1}{2}}(\dot{x}_2^2)^{\frac{1}{2}}\,d\tau_1\,d\tau_2. \quad (2.1)$$

$\dot{x}_i$ is the derivative of x_i with respect to τ_i. The generalization for more than two particles is obvious. χ and ψ are functions describing the two-particle interaction. They are subject only to the restriction

$$\chi(p) = \psi(p) = 0, \quad \text{when} \quad p > 0. \quad (2.2)$$

Our metric is $g_{00} = c^2$, $g_{0r} = 0$, $g_{rs} = -\delta_{rs}$; the interaction thus occurs at spacelike separations, which is the relativistic generalization of an instantaneous interaction. [Condition (2.2) may be weakened to require only that χ and ψ tend to zero as $p \to \infty$ fast enough for the integrals used to converge.]

The χ and ψ terms correspond to interactions of the Van Dam–Wigner[7] type and of the "scalar" type,[8] respectively. It has been shown in Ref. 8 that, in the ordinary nonrelativistic limit, both interactions reduce to Newtonian potentials depending only on the interparticle distance r given by

$$V(r) = \tfrac{1}{2}\int\chi(\theta^2 - r^2)\,d\theta, \quad (2.3)$$

$$U(r) = \tfrac{1}{2}\int\psi(\theta^2 - r^2)\,d\theta. \quad (2.4)$$

The full relativistic equations of motion for the action (2.1) are

$$\left(m_1 + \frac{1}{2c}\int\psi\,d\tau_2\right)\ddot{x}_1 = \frac{1}{c}\int\chi'\,d\tau_2\{\dot{x}_1\cdot\dot{x}_2(x_1-x_2) - (x_1-x_2)\cdot\dot{x}_1\dot{x}_2\} + \frac{1}{c}\int\psi'\,d\tau_2\{c^2(x_1-x_2) - (x_1-x_2)\cdot\dot{x}_1\dot{x}_1\} \quad (2.5)$$

and the equation obtained by exchanging the particle indices 1 and 2. χ' and ψ' are the derivatives of χ and ψ with respect to their arguments, the arguments still being $(x_1 - x_2)^2$. In Eq. (2.5) the arbitrariness connected with the choice of the parameters τ_1 and τ_2 has been removed by enforcing the requirement

$$\dot{x}_i^2 = c^2, \quad i = 1, 2. \quad (2.6)$$

This turns τ_i into the proper time of particle i.

3. THE LIMIT OF SLOW MOTION BUT STRONG BINDING

The procedure for obtaining the limit of slow motion while retaining the effects of strong binding is to let c tend to infinity while assuming that the classical potential energies $V(r)$ and $U(r)$ of Eqs. (2.3) and (2.4) are large of order c^2. This means that quantities such as $V(r)/c^2$ and $U(r)/c^2$ cannot be neglected. Terms proportional to any power of $\dot{\mathbf{x}}_i/c$ will be discarded.

The above prescription assumes that the potential energy $V(r) + U(r)$ is of order c^2 while the kinetic energy $\frac{1}{2}m_1\dot{\mathbf{x}}_1^2 + \frac{1}{2}m_2\dot{\mathbf{x}}_2^2$ is of order 1. However, the virial theorem[9,10] connects the average kinetic energy to the average of $(\mathbf{x}_1 - \mathbf{x}_2)\cdot(\partial/\partial\mathbf{x}_1)[V(r) + U(r)]$. For consistency we therefore assume that while $V(r)$ and $U(r)$ are of order c^2, their derivatives $dV(r)/dr$ and $dU(r)/dr$ are of order 1. This means that we

[9] H. Goldstein, *Classical Mechanics* (Addison-Wesley Publ. Co., Inc., London, 1959), Secs. 3–5.

[10] The virial theorem in Ref. 9 applies to ordinary Newtonian mechanics. We anticipate that a similar relation obtains also in the case of the equations of motion of strongly bound slow particles that we are about to derive. This assumption may be verified *a posteriori* by inspecting Eqs. (5.1) and (5.2).

consider motions where each particle is near the bottom of a deep potential well where the "floor" of the well is not steep.[11,12]

As the changes in $U(r)$ and $V(r)$ during the motion are of order 1, the limits

$$v \equiv \lim_{c\to\infty} V(r)/c^2, \tag{3.1}$$

$$u \equiv \lim_{c\to\infty} U(r)/c^2 \tag{3.2}$$

for a given solution are constants independent of r.

Our next task is to expand the equation of motion (2.5) in powers of $1/c$. Terms containing χ and ψ must be expanded up to order $(1/c)^2$. In fact, by our assumptions

$$V(r) = \tfrac{1}{2}\int \chi(\theta^2 - r^2)\,d\theta = -\int \chi'(\theta^2 - r^2)\theta^2\,d\theta$$
$$= \tfrac{2}{3}\int \chi''(\theta^2 - r^2)\theta^4\,d\theta = \cdots, \tag{3.3}$$

$$U(r) = \tfrac{1}{2}\int \psi(\theta^2 - r^2)\,d\theta = -\int \psi'(\theta^2 - r^2)\theta^2\,d\theta$$
$$= \tfrac{2}{3}\int \psi''(\theta^2 - r^2)\theta^4\,d\theta = \cdots \tag{3.4}$$

are of order c^2, but

$$\frac{1}{r}\frac{dV(r)}{dr} = -\int \chi'(\theta^2 - r^2)\,d\theta$$
$$= 2\int \chi''(\theta^2 - r^2)\theta^2\,d\theta = \cdots, \tag{3.5}$$

$$\frac{1}{r}\frac{dU(r)}{dr} = -\int \psi'(\theta^2 - r^2)\,d\theta$$
$$= 2\int \psi''(\theta^2 - r^2)\theta^2\,d\theta = \cdots \tag{3.6}$$

are of order 1. Therefore, integrals of the form (3.3) and (3.4), when multiplied by $1/c^2$, will be replaced by v and u, respectively; integrals of the form (3.5) and (3.6), when multiplied by powers of $1/c$, will be discarded.

Up to the order $1/c^2$, the time coordinates of the particles may be expressed as

$$x_i^0(\tau_i) = \tau_i + f_i(\tau_i)/c^2, \quad i = 1, 2. \tag{3.7}$$

The condition (2.6) then requires

$$f_i(\tau_i) = \tfrac{1}{2}\dot{\mathbf{x}}_i(\tau_i)^2. \tag{3.8}$$

The argument of the function χ and ψ, when expanded to the same order, becomes

$$(x_1 - x_2)^2 = c^2(\tau_1 - \tau_2)^2 + 2(\tau_1 - \tau_2)[f_1(\tau_1) - f_2(\tau_2)] - [\mathbf{x}_1(\tau_1) - \mathbf{x}_2(\tau_2)]^2; \tag{3.9}$$

also

$$(x_1 - x_2)\cdot\dot{x}_1 = c^2(\tau_1 - \tau_2) + f_1(\tau_1) - f_2(\tau_2) + \tfrac{1}{2}(\tau_1 - \tau_2)\dot{\mathbf{x}}_1(\tau_1)^2 - [\mathbf{x}_1(\tau_1) - \mathbf{x}_2(\tau_2)]\cdot\dot{\mathbf{x}}_1(\tau_1), \tag{3.10}$$

$$\dot{x}_1\cdot\dot{x}_2 = c^2 + \tfrac{1}{2}[\dot{\mathbf{x}}_1(\tau_1) - \dot{\mathbf{x}}_2(\tau_2)]^2. \tag{3.11}$$

Equations (3.10) and (3.11) may be obtained from Eq. (3.9) by taking its derivatives first with respect to τ_1, then with respect to τ_2.

4. THE LIMIT OF THE EQUATION OF MOTION

In this section we apply the preparations of the last section to the equation of motion (2.5).

Let us first change the variable of integration in all the integrals in Eq. (2.5) from τ_2 to

$$\theta = c(\tau_2 - \tau_1). \tag{4.1}$$

This together with the expanded expressions (3.7) through (3.10) turns the space components of Eq. (2.5) into

$$\left(m_1 + \frac{1}{c^2}\int \psi\,d\theta\right)\ddot{\mathbf{x}}_1$$
$$= \int(\chi' + \psi')\left[\mathbf{x}_1(\tau_1) - \mathbf{x}_2\left(\tau_1 + \frac{\theta}{c}\right)\right]d\theta$$
$$+ \frac{1}{c^2}\int \chi'\,d\theta\,\frac{1}{2}\left[\dot{\mathbf{x}}_1(\tau_1) - \dot{\mathbf{x}}_2\left(\tau_1 + \frac{\theta}{c}\right)\right]^2$$
$$\times\left[\mathbf{x}_1(\tau_1) - \mathbf{x}_2\left(\tau_1 + \frac{\theta}{c}\right)\right]$$
$$- \frac{1}{c^2}\int\left[\dot{\mathbf{x}}_2\left(\tau_1 + \frac{\theta}{c}\right)\chi' + \dot{\mathbf{x}}_1(\tau_1)\psi'\right]$$
$$\times\left[c\theta + f_1(\tau_1) - f_2\left(\tau_1 + \frac{\theta}{c}\right) + \frac{1}{2}\frac{\theta}{c}\dot{\mathbf{x}}_1(\tau_1)^2 - \left[\mathbf{x}_1(\tau_1) - \mathbf{x}_2\left(\tau_1 + \frac{\theta}{c}\right)\right]\cdot\mathbf{x}_1(\tau_1)\right]d\theta, \tag{4.2}$$

where the argument of the functions χ, ψ, χ', and ψ' is

$$(x_1 - x_2)^2 = \theta^2 + 2\frac{\theta}{c}\left[f_1(\tau_1) - f_2\left(\tau_1 + \frac{\theta}{c}\right)\right] - \left[\mathbf{x}_1(\tau_1) - \mathbf{x}_2\left(\tau_1 + \frac{\theta}{c}\right)\right]^2. \tag{4.3}$$

One should now complete the expansion to order $1/c^2$ by applying the Taylor expansion to all functions the arguments of which contain parts proportional to θ/c:

$$\mathbf{x}_2(\tau_1 + \theta/c) = \mathbf{x}_2(\tau_1) + (\theta/c)\dot{\mathbf{x}}_2(\tau_1) + \tfrac{1}{2}(\theta/c)^2\ddot{\mathbf{x}}_2(\tau_1), \tag{4.4}$$

[11] Quantum-mechanical considerations allow slow motion in a deep potential well only if it is wide enough (cf. Ref. 4).

[12] The prescription for obtaining the χ or ψ corresponding to a given $V(r)$ or $U(r)$ is given in Ref. 8.

etc. The functions χ, ψ, χ', ψ' should also be expanded. When this process is finished, all the θ integrals are of the form

$$\int \chi^{(n)}(\theta^2 - r^2)\theta^m \, d\theta \quad \text{or} \quad \int \psi^{(n)}(\theta^2 - r^2)\theta^m \, d\theta, \quad (4.5)$$

where $\chi^{(n)}$, $\psi^{(n)}$ are the nth derivatives of χ, ψ with respect to their argument and

$$r^2 = [\mathbf{x}_1(\tau_1) - \mathbf{x}_2(\tau_1)]^2. \quad (4.6)$$

Of these integrals, the ones with odd m vanish by symmetry. The others are either of the form (3.3), (3.4) or of the form (3.5), (3.6). The latter, when multiplied by $1/c^2$, are discarded; the former are replaced by v or u. This turns Eq. (4.2) into

$$(m_1 + u)\ddot{\mathbf{x}}_1 + \tfrac{1}{2}(v - u)\ddot{\mathbf{x}}_2 = -\frac{\partial}{\partial \mathbf{x}_1}[V(r) + U(r)]. \quad (4.7)$$

In the last equation both $\mathbf{x}_1$ and $\mathbf{x}_2$ depend on τ_1, which becomes the common time. It may be checked that the time component of Eq. (2.5) is satisfied identically.

5. CENTER-OF-MASS AND RELATIVE MOTIONS

Let us reproduce the equations of motion in the limit of slow motion but strong binding:

$$(m_1 + u)\ddot{\mathbf{x}}_1 + \tfrac{1}{2}(v - u)\ddot{\mathbf{x}}_2 = -\frac{\partial}{\partial \mathbf{x}_1}[V(r) + U(r)] + \mathbf{F}_1, \quad (5.1)$$

$$\tfrac{1}{2}(v - u)\ddot{\mathbf{x}}_1 + (m_2 + u)\ddot{\mathbf{x}}_2 = -\frac{\partial}{\partial \mathbf{x}_2}[V(r) + U(r)] + \mathbf{F}_2. \quad (5.2)$$

We have tentatively included external forces $\mathbf{F}_1$, $\mathbf{F}_2$ acting on particles 1 and 2. These forces could be produced by interaction of the two particles with a third particle that is extremely heavy and extremely far away. They are necessary in order fully to appreciate the effect of strong binding on the center-of-mass motion. Otherwise, Eq. (5.1) is identical with Eq. (4.7) and Eq. (5.2) is obtainable from it by exchanging the indices 1 and 2. Both $\mathbf{x}_1$ and $\mathbf{x}_2$ depend on a common time; dots denote derivatives with respect to this common time.

When the two equations (5.1) and (5.2) are added together, the interparticle force cancels and we find that

$$M\ddot{\mathbf{X}} = \mathbf{F}_1 + \mathbf{F}_2, \quad (5.3)$$

where

$$M = m_1 + m_2 + v + u \quad (5.4)$$

and

$$M\mathbf{X} = m_1\mathbf{x}_1 + m_2\mathbf{x}_2 + \tfrac{1}{2}(v + u)(\mathbf{x}_1 + \mathbf{x}_2). \quad (5.5)$$

Obviously, $\mathbf{X}$ plays the role of a center-of-mass coordinate and Eq. (5.3) shows that M is the inertial mass of the two-particle system.

For the purpose of studying the relative motion, let us neglect the external forces $\mathbf{F}_1$ and $\mathbf{F}_2$ and define a relative coordinate

$$\boldsymbol{\xi} \equiv \mathbf{x}_1 - \mathbf{x}_2. \quad (5.6)$$

Equations (5.1) and (5.2) then lead to

$$\mu\ddot{\boldsymbol{\xi}} = -\frac{\partial}{\partial \xi}[V(|\boldsymbol{\xi}|) + U(|\boldsymbol{\xi}|)], \quad (5.7)$$

where

$$\mu = \frac{(m_1 + u)(m_2 + u) - \tfrac{1}{4}(v - u)^2}{m_1 + m_2 + v + u}. \quad (5.8)$$

The last equation replaces the usual expression for the reduced mass.

Note that the total mass (5.4) agrees with Einstein's formula (1.1) regardless of the part of the potential energy contributed by the χ and ψ terms of the interaction. The effect of strong binding on the reduced mass (5.8) is crucially model-dependent. It may altogether disappear for a suitable choice of u and v, e.g., when $m_1 = m_2 = m$, and $v = 3u$, the reduced mass retains its usual value $\frac{1}{2}m$ although the total mass is shifted to $2m + 4u$.

6. TOTAL ENERGY AND GRAVITATIONAL MASS

It might be of value to compare the total mass M of the last section with the zero component of the energy–momentum vector [cf. Eqs. (3.1) and (4.4) of Ref. 8]. One finds that the total energy is equal to Mc^2 in leading order.

Another approach to the total energy is to consider space–time as infinitesimally curved and to evaluate the stress-energy tensor as

$$T^{\mu\nu}(x, t) = \frac{\delta A}{\delta g_{\mu\nu}(x, t)}, \quad (6.1)$$

where A is the action (2.1). The energy calculated as the integral of T^{00} over all of space also agrees with Mc^2 in leading order. The mass obtained in this way may be considered the active gravitational mass of the system, since the $T^{\mu\nu}$ serves as the source in Einstein's equations of gravity.

7. LIMITS OF VALIDITY

The equations of motion (5.1) and (5.2) relate a linear form in the accelerations to the forces. For some values of v and u the linear form may become singular, forcing certain components of the acceleration to become infinite. When this happens, we have exceeded the limits of validity of our assumptions. The moment one of the eigenvalues of the matrix multiplying the accelerations becomes small of order $1/c$, accelerations of order c are indicated and our assumption of slow motion is violated.

The vanishing of the determinant of the linear form in Eqs. (5.1) and (5.2) occurs on a hyperbola in the v, u plane described by

$$(m_1 + u)(m_2 + u) - \tfrac{1}{4}(v - u)^2 = 0 \qquad (7.1)$$

(see Fig. 1). This coincides with the vanishing of the reduced mass μ. The total mass M vanishes on the straight line

$$m_1 + m_2 + v + u = 0. \qquad (7.2)$$

It is not possible to connect continuously the origin $v = u = 0$ with the points of the line (7.2) without cutting the hyperbola (7.1). Therefore, the vanishing of the total mass occurs outside the region of validity of our approximation of slow motion and strong binding.[13] For the case of equal masses, the hyperbola degenerates into its two asymptotes, one of which is (7.2).

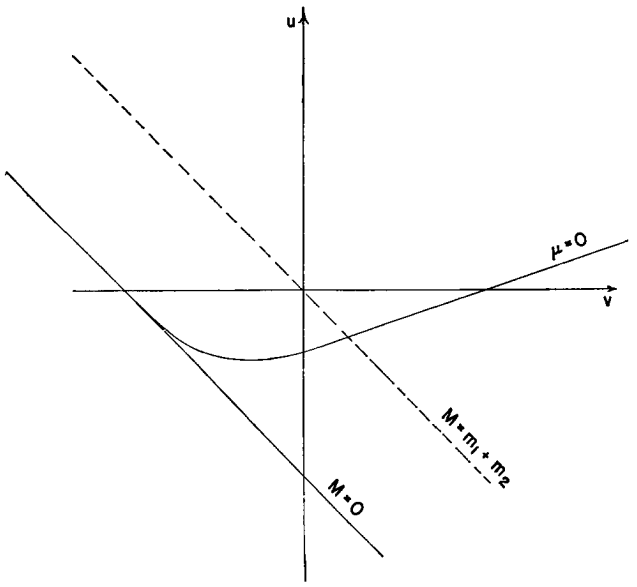


FIG. 1. The straight line $m_1 + m_2 + u + v = 0$, on which the total mass vanishes is an asymptote of the hyperbola

$$(m_1 + u)(m_2 + u) - \tfrac{1}{4}(v - u)^2 = 0,$$

on which the reduced mass vanishes.

APPENDIX A

In this appendix we consider an action of the form

$$A = m_1 c^2 \int (\dot{x}_1^2)\, d\tau_1 + m_2 c^2 \int (\dot{x}_2^2)\, d\tau_2 - \tfrac{1}{2}\iint \sigma((x_1 - x_2)^2)(\dot{x}_1 \cdot \dot{x}_2)^{1-q}(\dot{x}_1^2)^{\frac{1}{2}q}(\dot{x}_2^2)^{\frac{1}{2}q}\, d\tau_1\, d\tau_2. \qquad (A1)$$

The corresponding equation of motion is

$$\left\{\left[m_1 + qc^{2q-3}\tfrac{1}{2}\int \sigma\, d\tau_2 (\dot{x}_1 \cdot \dot{x}_2)^{1-q}\right]\delta_\nu^\mu - q(1-q)c^{2q-1} \times \tfrac{1}{2}\int \sigma\, d\tau_2 (\dot{x}_1 \cdot \dot{x}_2)^{-1-q}\left(\dot{x}_2^\mu - \frac{\dot{x}_1 \cdot \dot{x}_2}{c^2}\dot{x}_1^\mu\right)\dot{x}_{2\nu}\right\} x_1^\nu = c^{2q-1}\int \sigma'\, d\tau_2 (\dot{x}_1 \cdot \dot{x}_2)^{1-q} \times \left[x_1^\mu - x_2^\mu - (x_1 - x_2)\cdot \dot{x}_1\left(\frac{1-q}{\dot{x}_1 \cdot \dot{x}_2}\dot{x}_2^\mu + \frac{q}{c^2}\dot{x}_1^\mu\right)\right]. \qquad (A2)$$

The present theory depends on a continuous parameter q. For $q = 0$ we recover the Van Dam–Wigner theory, for $q = 1$ the scalar theory.

For $q \neq 0, 1$ we have a new theory distinct from the combination of Van Dam–Wigner and scalar theory considered in this paper. However, in the limit of slow motion and strong binding this new theory becomes identical to our mixed theory of Sec. 2 with

$$\chi = (1 - q)\sigma, \qquad (A3)$$

$$\psi = q\sigma. \qquad (A4)$$

[13] There is one case in which the approximation of slow motion involves no error: It is the case of two particles at rest, each at the bottom of the potential well created by the other. This situation is an exact solution of the full relativistic equation of motion (2.5). The energy taken as the zero component of the energy–momentum vector of Ref. 8 is, in leading order, $(m_1 + m_2 + v + u)c^2$, and can be made to vanish by increasing the interaction terms by a factor. This approach neglects questions of stability. Also the case of stationary particles may be exceptional.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 12 DECEMBER 1969

Half-Space Multigroup Transport Theory*

S. PAHOR† AND J. K. SHULTIS‡
Department of Nuclear Engineering, The University of Michigan, Ann Arbor, Michigan

(Received 4 April 1968)

A method for solving various half-space multigroup transport problems for the case of a symmetric transfer matrix is explained. This method is based on the full-range completeness and orthogonality properties of the infinite-medium eigenfunctions. First, the albedo problem is considered. A system of Fredholm integral equations is derived for the emergent distribution of the albedo problem, and it is shown that this system has a unique solution. Then, by using the full-range eigenfunction completeness, the inside angular distribution is obtained from the emergent distribution. Finally, the Milne problem and the half-space Green's function problem are solved in terms of the emergent distribution of the albedo problem and the infinite-medium eigenfunctions.

1. INTRODUCTION

In recent years much effort has been given to solving the energy-dependent Boltzmann equation. Various approximations have been used. The most rewarding approximation to date has been the multigroup technique, and often the diffusion-theory approximation is employed to simplify the calculations further. However, there is a definite need for exact solutions of the multigroup transport equations, since these solutions serve as a standard against which one can compare the approximate results.

Recently, the solution of the infinite-medium Green's function has been obtained explicitly for the two-group[1] and N-group[2,3] cases. Several two-group half-space problems have been investigated,[4–6] and in a paper by Siewert and Zweifel[7] a special N-group Milne problem for radiative transfer was solved. The general case of N-group half-space problems with symmetric transfer matrix was studied by Leonard and Ferziger[3]; they proved full- and half-range completeness of the N-group transport-equation eigenfunctions. In all these works, the solution of a half-space transport problem is expanded in terms of the eigenfunctions and then a set of equations for the expansion coefficients is derived.

In this paper we consider also N-group half-space problems for a symmetric transfer matrix. This form of $\mathbf{C}$ is not so restrictive as may appear at first glance. For instance, all two-group problems (see Appendix B) and the N-group equations for thermal neutrons may be transformed into such a case (see Appendix A and Ref. 3). Symmetric transfer also appears in special astrophysical radiative transfer problems for a medium in local thermodynamic equilibrium.[7]

In our approach, we do not need the half-range completeness property of the eigenfunctions. We solve half-space transport problems in two steps: First the emergent distribution is calculated and then the distribution inside the medium is evaluated by using the full-range completeness and orthogonality properties of the N-group eigenfunctions. These eigensolutions to the N-group isotropic transport equation and their full-range completeness theorem have been known for several years,[8] while their orthogonality relations have recently been obtained by Leonard and Ferziger[3] and Yoshimura.[2]

Section 2 briefly summarizes the N-group eigenfunctions and their full-range orthogonality relations as described by Yoshimura.[2] In addition, it is shown for symmetric $\mathbf{C}$ that the discrete eigenvalues are real or purely imaginary. In Sec. 3, a system of Fredholm equations is obtained which uniquely determines the emergent distribution for the albedo problem. It is shown that the uniqueness of solution of this system of Fredholm equations also implies half-range completeness of the eigenfunctions. Finally, in Sec. IV, the emergent distributions of the Milne's- and Green's-function problems are expressed in terms of the emergent albedo-problem distribution and the complete solutions obtained from the full-range completeness and orthogonality properties.

2. INFINITE-MEDIUM EIGENFUNCTIONS

The linear Boltzmann equation for N energy groups in plane geometry and with isotropic scattering

* Work supported by the National Science Foundation, and based, in part, upon a Ph.D. thesis of one of the authors (J. K. S.).

† On leave from the University of Ljubljana, Yugoslavia.

‡ Present address: Dept. of Nuclear Engineering, Kansas State Univ., Manhattan, Kansas.

[1] C. E. Siewert and P. S. Shieh, J. Nucl. Energy **21**, 383 (1962).

[2] T. Yoshimura and S. Katsuragi, Nucl. Sci. Eng. **33**, 297 (1968).

[3] A. Leonard and J. H. Ferziger, Nucl. Sci. Eng. **26**, 170 (1966).

[4] R. Zelazny and A. Kuszell, Ann. Phys. (N.Y.) **16**, 81 (1961).

[5] D. Metcalf, Ph.D. thesis, University of Michigan.

[6] C. E. Siewert and P. F. Zweifel, Ann. Phys. (N.Y.) **36**, 61 (1966).

[7] C. E. Siewert and P. F. Zweifel, J. Math. Phys. **7**, 2092 (1966).

[8] R. Zelazny and A. Kuszell, *Physics of Fast and Intermediate Reactors* (IAEA, Vienna, 1964), Vol. II, p. 55.

can be written in the form[4]

$$\mu \frac{\partial}{\partial x} \psi(x, \mu) + \mathbf{\Sigma}\psi(x, \mu) = \mathbf{C}\int_{-1}^{1} d\mu' \psi(x, \mu'). \quad (2.1)$$

The vector $\psi(x, \mu)$ is an N-component vector, of which the ith component, $\psi_i(x, \mu)$, is the angular flux of the ith group. The components of the diagonal matrix, the $\mathbf{\Sigma}$, are the $\sigma_i \delta_{ij}$, where σ_i is the total interaction cross section for the ith group. The elements C_{ij} of the transfer matrix $\mathbf{C}$ describe the transfer of neutrons from the jth group to the ith group. In some problems, for instance, thermal neutron-transport theory, $\mathbf{C}$ can be written as a product of diagonal matrices $\mathbf{D}_i$ and a symmetric matrix $\mathbf{A}$ (see Appendix A), as

$$\mathbf{C} = \mathbf{D}_1 \mathbf{A} \mathbf{D}_2. \quad (2.2)$$

Equation (2.1) can then be so transformed that the elements of the transformed $\mathbf{\Sigma}$ matrix are ordered as

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_N, \quad (2.3)$$

and the new $\mathbf{C}$ matrix is symmetric (Appendix C). It will be assumed for the remainder of the paper that the transport equation has this special form of an ordered $\mathbf{\Sigma}$ matrix and symmetric $\mathbf{C}$ matrix. Finally, by measuring distance in units of the smallest mean free path, we can set $\sigma_N = 1$.

Using the analogy of the one-group problem,[9] we seek a set of eigenfunction solutions $\psi(\nu, x, \mu)$ to Eq. (2.1) of the form

$$\psi(\nu, x, \mu) = e^{-x/\nu} \boldsymbol{\phi}(\nu, \mu). \quad (2.4)$$

Substituting Eq. (2.4) into Eq. (2.1), the self-adjoint equation for the eigenvectors $\boldsymbol{\phi}(\nu, \mu)$ is obtained as

$$(\mathbf{\Sigma} - (\mu/\nu)\mathbf{E})\boldsymbol{\varphi}(\nu, \mu) = \mathbf{C}\int_{-1}^{1} d\mu' \boldsymbol{\phi}(\nu, \mu'), \quad (2.5)$$

where $\mathbf{E}$ is the unit matrix. The explicit form of these eigenfunctions has been obtained by several authors.[1–3,8] We will use, with slight changes, the notation of Yoshimura.[2]

The eigenvector spectrum is divided into two regions.

(a) *Region I:* $\nu \notin (-1, 1)$. In this region there may exist an even number, say $2M$, of discrete eigenvectors, which are written in component form as

$$\phi_i(\nu_{0s}, \mu) = \nu_{0s} b_i(\nu_{0s})/(\sigma_i \nu_{0s} - \mu), \quad i = 1, \cdots, N, \quad (2.6)$$

where $\mathbf{b}(\nu_{0s})$ is a well-defined vector.[2] It can be shown that if ν_{0s} is an eigenvalue, then also $-\nu_{0s}$ and ν_{0s}^* (complex conjugate) are eigenvalues with

$$\mathbf{b}(\nu_{0s}) = \mathbf{b}(-\nu_{0s}) = \mathbf{b}^*(\nu_{0s}). \quad (2.7)$$

For our case of symmetric $\mathbf{C}$, the discrete eigenvalues ν_{0s} are either real or imaginary—never complex. To see this, multiply Eq. (2.5) by $\tilde{\boldsymbol{\phi}}^*(\nu, \mu)$ and integrate over μ. (Here the superscript tilde denotes the transpose.) In this way one obtains the equation

$$\frac{1}{\nu_{0s}} \int_{-1}^{1} d\mu \mu \tilde{\boldsymbol{\phi}}^*(\nu_{0s}, \mu) \boldsymbol{\phi}(\nu_{0s}, \mu) = \int_{-1}^{1} d\mu \tilde{\boldsymbol{\phi}}^*(\nu_{0s}, \mu) \mathbf{\Sigma} \boldsymbol{\phi}(\nu_{0s}, \mu) - \int_{-1}^{1} d\mu \tilde{\boldsymbol{\phi}}^*(\nu_{0s}, \mu) \mathbf{C} \int_{-1}^{1} d\mu' \boldsymbol{\phi}(\nu_{0s}, \mu'). \quad (2.8)$$

Since $\mathbf{\Sigma}$ is diagonal and $\mathbf{C}$ is symmetric, the right-hand side of Eq. (2.8) is real since it is a sum of products of complex-conjugate terms. The integral on the left-hand side of Eq. (2.8), which in view of Eq. (2.6) can be written as

$$\int_{-1}^{1} d\mu \mu \tilde{\boldsymbol{\phi}}^*(\nu_{0s}, \mu) \boldsymbol{\phi}(\nu_{0s}, \mu) = \nu_{0s} \nu_{0s}^* \sum_{i=1}^{N} b_i(\nu_{0s}) b_i^*(\nu_{0s}) \int_{-1}^{1} \frac{\mu \, d\mu}{(\nu_{0s}\sigma_i - \mu)(\nu_{0s}^*\sigma_i - \mu)}, \quad (2.9)$$

is also real.

If the above integral (2.9) is not zero, it follows then that the eigenvalue ν_{0s} must be real! It will now be shown that this integral can vanish only for purely imaginary eigenvalues.

Let us assume, for the sake of the argument, that ν_{0s} is complex and $\mathrm{Re}\{\nu_{0s}\} > 0$. It can easily be verified that in this case

$$0 < (\nu_{0s}\sigma_i - \mu)(\nu_{0s}^*\sigma_i - \mu) < (\nu_{0s}\sigma_i + \mu)(\nu_{0s}^*\sigma_i + \mu), \quad \mu > 0, \quad i = 1, \cdots, N. \quad (2.10)$$

Hence, each integral in the sum on the right-hand side of Eq. (2.9) is strictly positive and, since at least one of the terms $b_i(\nu_{0s})b_i^*(\nu_{0s})$ is also strictly positive in view of Eq. (2.7), the sum is strictly positive for $\mathrm{Re}\{\nu_{0s}\} > 0$. Similarly, it can be proved that for $\mathrm{Re}\{\nu_{0s}\} < 0$ the sum is strictly negative. Thus the integral (2.9) never vanishes if $\mathrm{Re}\{\nu_{0s}\} \neq 0$.

However, if ν_{0s} is purely imaginary, we have

$$(\nu_{0s}\sigma_i - \mu)(\nu_{0s}^*\sigma_i - \mu) = (\nu_{0s}\sigma_i + \mu)(\nu_{0s}^*\sigma_i + \mu) \quad (2.11)$$

and each integral in the right-hand side of Eq. (2.9) is zero. Thus, we conclude, the discrete eigenvalues ν_{0s} lie on only the real or imaginary axis.

[9] K. M. Case and P. F. Zweifel, *Linear Transport Theory* (Addison-Wesley, Reading, Mass., 1967).

(b) *Region II:* This region is divided into N subintervals $\nu_j, j = 1, \cdots, N$, such that for $\nu \in \nu_j$, $1/\sigma_{j-1} < |\nu| \leq 1/\sigma_j$, where $\sigma_0 = 1$. For the jth subinterval, there are $(N - j + 1)$ linearly independent eigenvectors $\boldsymbol{\phi}_j^m(\nu, \mu)$, whose ith component has the form

$$[\boldsymbol{\phi}_j^m(\nu, \mu)]_i = P\frac{\nu}{\sigma_i\nu - \mu}[\mathbf{b}_j^m(\nu)]_i + \delta(\sigma_i\nu - \mu)[\boldsymbol{\lambda}_j^m(\nu)]_i,$$
$$m = j, \cdots, N, \quad j = 1, \cdots, N, \quad (2.12)$$

where P indicates the Cauchy principle value is to be used when these functions are integrated. The vectors $\mathbf{b}_j^m(\nu)$ and $\boldsymbol{\lambda}_j^m(\nu)$ are also defined by Yoshimura.[2]

From the eigenvalue equation (2.5), one finds that the eigenvectors are orthogonal in the following sense:

$$\int_{-1}^{1} d\mu\mu\tilde{\boldsymbol{\phi}}(\nu, \mu)\boldsymbol{\phi}(\nu', \mu) = 0 \quad \text{if} \quad \nu' \neq \nu. \quad (2.13)$$

Moreover, it is possible to choose particular linear combinations of eigenvectors for the independent eigenvectors of each subinterval ν_j, such that all the "continuum" eigenvectors are orthogonal in the following sense:

$$\int_{-1}^{1} d\mu\mu\tilde{\boldsymbol{\phi}}_j^m(\pm\nu, \mu)\boldsymbol{\phi}_j^{m'}(\pm\nu', \mu) = \pm N_j^m(\nu)\delta_{mm'}\delta(\nu - \nu'). \quad (2.14)$$

Similarly, for the "discrete" eigenvalues we have

$$\int_{-1}^{1} d\mu\mu\tilde{\boldsymbol{\phi}}(\pm\nu_{0s}, \mu)\boldsymbol{\phi}(\pm\nu_{0s'}, \mu) = \pm N_s\delta_{ss'}, \quad s = 1, \cdots, M. \quad (2.15)$$

The functions N_s and $N_j^m(\nu)$ are given by Yoshimura,[2] and it can be shown that $N_j^m(\nu)$ is positive for $\nu \geq 0$.

Finally, there is one more relationship between the eigenvectors which we will need later. From Yoshimura's work,[2] the functions $\mathbf{b}_j^m(\nu)$ and $\boldsymbol{\lambda}_j^m(\nu)$ are even functions of ν, and it follows that

$$\boldsymbol{\phi}(-\nu, \mu) = \boldsymbol{\phi}(\nu, -\mu). \quad (2.16)$$

3. THE ALBEDO PROBLEM

In this section we will consider the albedo problem for a half-space. This problem will be shown to be important because the solutions of all other half-space problems can be expressed in terms of the albedo solution.

A. Emergent Distribution

Let us now consider the albedo problem for which the incident neutron beam belongs solely to the ith energy group. In this case the angular flux will be denoted by $\boldsymbol{\psi}^i(0, \mu_0; x, \mu)$. It is the solution of Eq. (2.1) with the boundary conditions

$$\boldsymbol{\psi}^i(0, \mu_0; 0, \mu) = \mathbf{e}_i\delta(\mu - \mu_0), \quad \mu > 0, \quad \mu_0 > 0, \quad (3.1)$$

$$\lim_{x\to\infty} \boldsymbol{\psi}^i(0, \mu_0; x, \mu) = \mathbf{0}, \quad (3.2)$$

where $\mathbf{e}_i$ is a vector, all of whose components are zero except the ith, which is unity. Since our eigenfunctions are complete,[2,8] the solution for this albedo problem can be expanded in terms of the eigenfunctions which satisfy the boundary conditions at infinity:

$$\boldsymbol{\psi}^i(0, \mu_0; x, \mu) = \sum_{s=1}^{M} \alpha(\nu_{0s})\boldsymbol{\phi}(\nu_{0s}, \mu)e^{-x/\nu_{0s}} + \sum_{j=1}^{N}\int_{\eta_{j-1}}^{\eta_j} d\nu\left\{\sum_{m=j}^{N} A_j^m(\nu)\boldsymbol{\phi}_j^m(\nu, \mu)e^{-x/\nu}\right\}, \quad i = 1, \cdots, N, \quad (3.3)$$

where $\eta_j = 1/\sigma_j$, $j = 1, \cdots, N$, and $\eta_0 = 0$. We will assume that all the ν_{0s} are real. Clearly, eigenfunctions with imaginary eigenvalues cannot satisfy our infinity boundary condition. Setting $x = 0$ and using the full-range orthogonality relations plus boundary condition (3.1), we obtain the expansion coefficients as

$$\alpha(\nu_{0s}) = \frac{\mu_0}{N_s}\tilde{\boldsymbol{\phi}}(\nu_{0s}, \mu_0)\mathbf{e}_i - \frac{1}{N_s}\int_0^1 d\mu\mu\tilde{\boldsymbol{\phi}}(\nu_{0s}, -\mu)\boldsymbol{\psi}^i(0, \mu_0; 0, -\mu) \quad (3.4)$$

and

$$A_j^m(\nu) = \frac{\mu_0}{N_j^m(\nu)}\tilde{\boldsymbol{\phi}}_j^m(\nu, \mu_0)\mathbf{e}_i - \frac{1}{N_j^m(\nu)}\int_0^1 d\mu\mu\tilde{\boldsymbol{\phi}}_j^m(\nu, -\mu)\boldsymbol{\psi}^i(0, \mu_\eta; 0, -\mu). \quad (3.5)$$

Substituting these coefficients into Eq. (3.3) with $x = 0$, we obtain the following inhomogeneous Fredholm equation for the emergent distribution:

$$\boldsymbol{\psi}^i(0, \mu_0; 0, -\mu) = \mathbf{F}(\mu)\mathbf{e}_i - \int_0^1 d\mu'\mu'\mathbf{K}(\mu', \mu)\boldsymbol{\psi}^i(0, \mu_0; 0, -\mu'), \quad \mu > 0. \quad (3.6)$$

Here we have defined the matrices

$$\mathbf{F}(\mu) = \mu_0\sum_{s=1}^{M}\frac{1}{N_s}\boldsymbol{\phi}(\nu_{0s}, -\mu)\tilde{\boldsymbol{\phi}}(\nu_{0s}, \mu_0) + \mu_0\sum_{j=1}^{N}\int_{\eta_{j-1}}^{\eta_j} d\nu\left\{\sum_{m=j}^{N}\frac{1}{N_j^m(\nu)}\boldsymbol{\phi}_j^m(\nu, -\mu)\tilde{\boldsymbol{\phi}}_j^m(\nu, \mu_0)\right\} \quad (3.7)$$

and

$$\mathbf{K}(\mu', \mu) = \sum_{s=1}^{M} \frac{1}{N_s} \boldsymbol{\phi}(\nu_{0s}, -\mu)\tilde{\boldsymbol{\phi}}(\nu_{0s}, -\mu') + \sum_{j=1}^{N} \int_{\eta_{j-1}}^{\eta_j} d\nu \left\{ \sum_{m=j}^{N} \frac{1}{N_j^m(\nu)} \boldsymbol{\phi}_j^m(\nu, -\mu)\tilde{\boldsymbol{\phi}}_j^m(\nu, -\mu') \right\}. \tag{3.8}$$

It can be verified that $\mathbf{K}(\mu', \mu)$ and $\mathbf{F}(\mu)$ are continuous functions of their arguments for $\mu, \mu' > 0$.

One can also obtain a singular integral equation for $\boldsymbol{\psi}^i(0, \mu_0; 0, -\mu)$ by considering the incident distribution as given by Eqs. (3.3)–(3.5); explicitly, we have

$$\delta(\mu - \mu_0)\mathbf{e}_i = \mathbf{F}(-\mu)\mathbf{e}_i - \int_0^1 d\mu'\mu'\mathbf{K}(\mu', -\mu)\boldsymbol{\psi}^i(0, \mu_0; 0, -\mu'). \tag{3.9}$$

Either Eqs. (3.6) and/or (3.9) may be used to determine the emergent distribution. Case has obtained the same pair of equations, expressed in terms of the infinite-medium Green's function,[10] by using a different approach. When explicit expressions for the Green's functions are substituted into his equations, Eqs. (3.6) and (3.9) are obtained.

In the one-speed case, the singular integral equation (3.9) and the Fredholm equation (3.6) may be solved together in closed form.[10] However, for the multigroup situation, no closed-form solutions have been obtained, and numerical procedures must be used to determine the emergent distribution.

It will be shown that the emergent distribution is uniquely determined by the system of Fredholm integral equations (3.6) alone, and this system of equations can be solved by standard numerical techniques.

Once Eq. (3.6) has been solved for $\boldsymbol{\psi}^i(0, \mu_0; 0, -\mu)$, $\mu > 0$, the expansion coefficients can be completely determined from Eqs. (3.4) and (3.5). Then Eq. (3.3) gives the complete solution for the albedo problem.

B. Uniqueness of Solution of Fredholm Equation

To show that our Fredholm equation has a unique solution, we consider the homogeneous equation

$$\boldsymbol{\psi}'(0, \mu_0; 0, -\mu) = -\int_0^1 d\mu'\mu'\mathbf{K}(\mu', \mu)\boldsymbol{\psi}'(0, \mu_0; 0, -\mu'), \quad \mu > 0. \tag{3.10}$$

Defining

$$\mathbf{X}(\mu) = (\mu)^{\frac{1}{2}}\boldsymbol{\psi}'(0, \mu_0; 0, -\mu), \tag{3.11}$$

$$\mathbf{D}(\mu', \mu) = (\mu\mu')^{\frac{1}{2}}\mathbf{K}(\mu', \mu), \tag{3.12}$$

we have

$$\mathbf{X}(\mu) = -\int_0^1 \mathbf{D}(\mu', \mu)\mathbf{X}(\mu')\, d\mu'. \tag{3.13}$$

Let us assume a nontrivial solution exists. Multiplying Eq. (3.13) by $\tilde{\mathbf{X}}^*(\mu)$, integrating over μ, and substituting explicitly for $\mathbf{D}(\mu', \mu)$ from Eqs. (3.12) and (3.8), one obtains

$$\begin{aligned}\int_0^1 d\mu\tilde{\mathbf{X}}^*(\mu)\mathbf{X}(\mu) &= -\sum_{s=1}^{M} \frac{1}{N_s} \int_0^1 d\mu(\mu)^{\frac{1}{2}}[\tilde{\boldsymbol{\phi}}(\nu_{0s}, -\mu)\mathbf{X}^*(\mu)] \\ &\quad \times \int_0^1 d\mu'(\mu')^{\frac{1}{2}}\tilde{\boldsymbol{\phi}}(\nu_{0s}, -\mu')\mathbf{X}(\mu') \\ &\quad - \sum_{j=1}^{N} \int_{\eta_{j-1}}^{\eta_j} d\nu \left\{ \sum_{m=j}^{N} \frac{1}{N_j^m(\nu)} \int_0^1 d\mu(\mu)^{\frac{1}{2}}[\tilde{\boldsymbol{\phi}}_j^m(\nu, -\mu)\mathbf{X}^*(\mu)] \right. \\ &\quad \left. \times \int_0^1 d\mu'(\mu')^{\frac{1}{2}}\tilde{\boldsymbol{\phi}}_j^m(\nu, -\mu')\mathbf{X}(\mu') \right\}. \end{aligned} \tag{3.14}$$

Since all the eigenvalues are real, $\boldsymbol{\phi}(\nu, \mu)$ is also real, and hence both sides of Eq. (3.14) are composed of terms which are products of complex conjugates. Thus we have a contradiction: the right-hand side of Eq. (3.14) must be real and negative, while the left-hand side is real and strictly positive. Hence $\mathbf{X}(\mu)$ must be identically zero, or equivalently, the homogeneous equation (3.10) has only the null vector as a solution.

Because a system of integral equations may be transformed into a single integral equation,[11] it follows from the known properties of Fredholm integral equations that the solution of Eq. (3.6) exists and is unique since the homogeneous equation has only the trivial zero solution.[12]

An immediate consequence of this result is that the eigenvectors $\boldsymbol{\phi}(\nu, \mu)$, $\nu \geq 0$, $\mu \in (0, 1)$, are half-range complete in the sense of Case.[9] In fact, with $\boldsymbol{\psi}^i(0, \mu_0 \to 0, -\mu)$ being uniquely determined by Eq. (3.6), we see that Eq. (3.9) is just the half-range expansion of the vector $\delta(\mu_0 - \mu)\mathbf{e}_i$.

4. SOLUTIONS OF TYPICAL HALF-SPACE PROBLEMS

By using the results of the previous section, it will be shown how the emergent distributions for various

[10] K. M. Case, in *Transport Theory, SIAM–AMS Proceedings* (Am. Math. Soc., Providence, R.I., 1969).

[11] S. G. Mikhlin, *Integral Equations* (Pergamon Press, Inc., New York, 1964).

[12] W. Pogorzelski, *Integral Equations and Their Applications* (Pergamon Press, Inc., New York, 1964).

half-space problems may be expressed in terms of the emergent distributions of the albedo problems, $\psi^i(0, \mu_0; 0, -\mu)$, $i = 1, \cdots, N$.

A. Generalized Milne Problem

For every positive eigenvalue $\nu \in (0, 1)$ or $\nu = \nu_{0s}$, $s = 1, \cdots, M$, we define a Milne problem $\psi_\nu(x, \mu)$ by Eq. (2.1) and the following boundary conditions:

$$\psi_\nu(0, \mu) = \mathbf{0}, \quad \mu > 0, \tag{4.1}$$

$$\lim_{x\to\infty} \psi_\nu(x, \mu) = \phi(-\nu, \mu)e^{+x/\nu}, \tag{4.2}$$

where $\phi(-\nu, \mu)$ may be any of the eigenvectors—regular or singular.

First let us determine the emergent distribution $\psi_\nu(0, -\mu)$. Consider a solution $\psi(x, \mu)$ of the transport equation defined as

$$\psi(x, \mu) = \psi_\nu(x, \mu) + \psi_a(x, \mu), \tag{4.3}$$

where $\psi_a(x, \mu)$ is also a solution of the transport equation with the boundary conditions

$$\psi_a(0, \mu) = \phi(-\nu, \mu), \quad \mu > 0, \tag{4.4}$$

$$\lim_{x\to\infty} \psi_a(x, \mu) = \mathbf{0}. \tag{4.5}$$

From (4.3), therefore, $\psi(x, \mu)$ must have the boundary conditions

$$\psi(0, -\mu) = \phi(-\nu, \mu), \quad \mu > 0, \tag{4.6}$$

$$\lim_{x\to\infty} \psi(x, \mu) = \phi(-\nu, \mu)e^{x/\nu}. \tag{4.7}$$

Clearly the unique solution for $\psi(x, \mu)$ is

$$\psi(x, \mu) = \phi(-\nu, \mu)e^{x/\nu}. \tag{4.8}$$

Equations (4.3) and (4.8) then yield, for $x = 0$,

$$\phi(-\nu, -\mu) = \psi_\nu(0, -\mu) + \psi_a(0, -\mu), \quad -1 \le \mu \le 1. \tag{4.9}$$

Using the results of the previous section, the reflected distribution $\psi_a(0, -\mu)$, $\mu > 0$, can be expressed in terms of the incident distribution $\phi(-\nu, \mu)$ as

$$\psi_a(0, -\mu) = \sum_{i=1}^{N} \int_0^1 d\mu' [\phi(-\nu, \mu')]_i \psi^i(0, \mu'; 0, -\mu), \quad \mu > 0. \tag{4.10}$$

Thus the emergent distribution for the Milne problem becomes, in view of Eqs. (4.9) and (2.16),

$$\psi_\nu(0, -\mu) = \phi(\nu, \mu) - \sum_{i=1}^{N} \int_0^1 d\mu' [\phi(-\nu, \mu')]_i \psi^i(0, \mu'; 0, -\mu). \tag{4.11}$$

Finally, to obtain the complete solution for the generalized Milne problem, we use the following expansion:

$$\psi_\nu(x, \mu) = \phi(-\nu, \mu)e^{x/\nu} + \sum_{s=1}^{M} \alpha(\nu_{0s})\phi(\nu_{0s}, \mu)e^{-x/\nu_{0s}} + \sum_{j=1}^{N} \int_{\eta_{j-1}}^{\eta_j} d\nu \left\{ \sum_{m=j}^{N} A_j^m(\nu)\phi_j^m(\nu, \mu) \right\} e^{-x/\nu}. \tag{4.12}$$

The expansion coefficients are obtained by applying full-range orthogonality relations and Eq. (4.11). Explicitly they are

$$\alpha(\nu_{0s}) = -\frac{1}{N_s} \int_0^1 d\mu\mu \tilde{\phi}(\nu_{0s}, -\mu) \times \left[\phi(\nu, \mu) - \sum_{i=1}^{N} \int_0^1 d\mu' [\phi(-\nu, \mu')]_i \times \psi^i(0, \mu'; 0, -\mu) \right] \tag{4.13}$$

and

$$A_j^m(\nu) = -\frac{1}{N_j^m(\nu)} \int_0^1 d\mu\mu \tilde{\phi}_j^m(\nu, -\mu) \times \left[\phi(\nu, \mu) - \sum_{i=1}^{N} \int_0^1 d\mu' [\phi(-\nu, \mu')]_i \times \psi^i(0, \mu'; 0, -\mu) \right]. \tag{4.14}$$

B. Half-Space Green's Function

In a manner similar to that used for the Milne problem, the emergent distribution for the half-space Green's function can be expressed in terms of the emergent albedo-problem distributions. The half-space Green's function, with the source neutrons belonging to the ith group $\mathbf{G}^i(x_0, \mu_0; x, \mu)$, is defined by the equation

$$\left(\mu \frac{\partial}{\partial x}\mathbf{E} + \mathbf{\Sigma}\right)\mathbf{G}^i(x_0, \mu_0; 0, \mu) = \mathbf{C}\int_{-1}^{1} d\mu' \mathbf{G}^i(x_0, \mu_0; x, \mu') + \delta(\mu_0 - \mu)\delta(x - x_0)\mathbf{e}_i, \quad x_0 > 0, \tag{4.15}$$

with the boundary conditions

$$\mathbf{G}^i(x_0, \mu_0; 0, \mu) = \mathbf{0}, \quad \mu > 0, \tag{4.16}$$

$$\lim_{x\to\infty} \mathbf{G}^i(x_0, \mu_0; x, \mu) = \mathbf{0}. \tag{4.17}$$

To determine this function, we will assume that the infinite-medium Green's function $\mathbf{G}_\infty^i(x_0, \mu_0; x, \mu)$, which also satisfies Eq. (4.15), is known.[2] This

infinite-medium Green's function can be expressed in terms of the half-space Green's function as

$$\mathbf{G}^i_\infty(x_0, \mu_0; x, \mu) = \mathbf{G}^i(x_0, \mu_0; x, \mu) + \boldsymbol{\psi}_a(x, \mu), \quad x_0 > 0, \quad (4.18)$$

where $\boldsymbol{\psi}_a(x, \mu)$ is an albedo-problem solution satisfying Eq. (2.1) with boundary condition

$$\boldsymbol{\psi}_a(0, \mu) = \mathbf{G}^i_\infty(x_0, \mu_0; 0, \mu), \quad \mu > 0, \quad (4.19)$$

$$\lim_{x \to \infty} \boldsymbol{\psi}_a(x, \mu) = 0. \quad (4.20)$$

Expressing the emergent distribution for this albedo problem in terms of the known incident distribution and the vectors $\boldsymbol{\psi}^i(0, \mu_0; 0, -\mu)$, Eq. (4.18) yields

$$\mathbf{G}^i(x_0, \mu_0; 0, -\mu) = \mathbf{G}^i_\infty(x_0, \mu_0; 0, -\mu) - \sum_{j=1}^{N} \int_0^1 d\mu' [\mathbf{G}^i_\infty(x_0, \mu_0; 0, \mu')]_j \boldsymbol{\psi}^j(0, \mu'; 0, -\mu). \quad (4.21)$$

Since the angular flux for the half-space Green's function is now known at $x = 0$ for all μ, the complete solution can be found by using the full-range completeness and orthogonality theorems. Explicitly,

$$\mathbf{G}^i(x_0, \mu_0; x, \mu) = \mathbf{G}^i_\infty(x_0, \mu_0; x, \mu) + \sum_{s=1}^{M} \alpha(\nu_{0s}) \boldsymbol{\Phi}(\nu_{0s}, \mu) e^{-x/\nu_{0s}} + \sum_{j=1}^{N} \int_{\eta_{j-1}}^{\eta_j} d\nu \left\{ \sum_{m=j}^{N} A^m_j(\nu) \boldsymbol{\Phi}^m_j(\nu, \mu) \right\} e^{-x/\nu}, \quad (4.22)$$

where

$$\alpha(\nu_{0s}) = -\frac{1}{N_s} \int_0^1 d\mu \mu \tilde{\boldsymbol{\Phi}}(\nu_{0s}, \mu) \mathbf{G}^i_\infty(x_0, \mu_0; 0, \mu) + \frac{1}{N_s} \int_0^1 d\mu \mu \tilde{\boldsymbol{\Phi}}(\nu_{0s}, -\mu) \times \sum_{k=1}^{N} \int_0^1 d\mu' [\mathbf{G}^i_\infty(x_0, \mu_0; 0, \mu')]_k \times \boldsymbol{\psi}^k(0, \mu'; 0, -\mu), \quad (4.23)$$

$$A^m_j(\nu) = -\frac{1}{N^m_j(\nu)} \int_0^1 d\mu \mu \tilde{\boldsymbol{\Phi}}^m_j(\nu, \mu) \mathbf{G}^i_\infty(x_0, \mu_0; 0, \mu) + \frac{1}{N^m_j(\nu)} \int_0^1 d\mu \mu \tilde{\boldsymbol{\Phi}}^m_j(\nu, -\mu) \times \sum_{k=1}^{N} \int_0^1 d\mu' [\mathbf{G}^i_\infty(x_0, \mu_0; 0, \mu')]_k \times \boldsymbol{\psi}^k(0, \mu'; 0, -\mu). \quad (4.24)$$

5. SUMMARY

It has been shown that the solutions of all multigroup half-space problems involving a symmetric transfer matrix can be expressed in terms of the emergent albedo-problem distribution and the infinite-medium eigenfunctions. This emergent albedo distribution is uniquely determined by the Fredholm equation (3.6), which can be solved by standard numerical procedures.

In this paper, attention was restricted to those cases which could be transformed such that the transfer matrix was symmetric. This assumption was necessary to prove that (i) the eigenvalues of the transport equation are real or imaginary, and (ii) the emergent albedo distribution is uniquely determined by Eq. (3.6). In a future paper, this restriction will be relaxed and the case of a general transfer matrix will be discussed.

ACKNOWLEDGMENTS

We wish to thank Professor P. F. Zweifel, Professor G. C. Summerfield, and Professor C. E. Siewert for stimulating discussions and suggestions. One of us (S. P.) wishes to express his gratitude to the members of the Department of Nuclear Engineering, The University of Michigan, for the hospitality shown him and to the Institute of Science and Technology for financial support.

APPENDIX A. THERMAL REACTOR MODEL

The linear Boltzmann equation for a homogeneous nonmultiplying medium in plane geometry and with isotropic scattering may be written as

$$\left(\mu \frac{\partial}{\partial x} + \Sigma(E)\right) \psi(x, \mu, E) = \int_{-1}^{1} d\mu' \int_0^\infty \Sigma_s(E' \to E) \psi(x, \mu', E'), \quad (A1)$$

where $\psi(x, \mu, E)$ is the angular flux, and $\Sigma(E)$ and $\Sigma_s(E' \to E)$ are the total and differential scattering cross sections, respectively.

Using the usual multigroup technique,[13] the energy variable is split into N regions; integrating Eq. (A1) over the ith region, we obtain the ith multigroup equation

$$\left(\mu \frac{\partial}{\partial x} + \sigma_i\right) \psi_i(x, \mu) = \sum_{j=1}^{N} C_{ij} \int_{-1}^{1} d\mu' \psi_j(x, \mu'), \quad (A2)$$

[13] J. H. Ferziger and P. F. Zweifel, *Theory of Neutron Slowing Down in Nuclear Reactors* (MIT Press, Cambridge, 1967).

where we define

$$\psi_i(x,\mu) = \int_{\Delta E_i} dE\psi(x,\mu,E), \tag{A3}$$

$$\sigma_i = \frac{1}{\psi_i(x,\mu)} \int_{\Delta E_i} dE\Sigma(E)\psi(x,\mu,E), \tag{A4}$$

$$C_{ij} = \frac{1}{\psi_j(x,\mu)} \int_{\Delta E_i} dE \int_{\Delta E_j} dE'\Sigma_s(E' \to E)\psi(x,\mu,E'). \tag{A5}$$

To make the multigroup constants σ_i and C_{ij} independent of x and μ, it is usual to assume that the energy dependence of the angular flux is separable. Further, for a system in thermal equilibrium, a good first approximation is to assume this energy dependence is Maxwellian with some effective temperature T. With these assumptions the multigroup parameters are given by

$$C_{ij} = \alpha_j \int_{\Delta E_i} dE \int_{\Delta E_j} dE'\Sigma_s(E' \to E)M(E', T), \tag{A6}$$

$$\sigma_i = \alpha_i \int_{\Delta E_i} dE\Sigma(E)M(E, T), \tag{A7}$$

$$\frac{1}{\alpha_i} = \int_{\Delta E_i} dEM(E, T). \tag{A8}$$

The cross section $\Sigma_s(E' \to E)$ must obey the detailed balance relation[14]

$$\Sigma_s(E' \to E)M(E', T) = \Sigma_s(E \to E')M(E, T) \tag{A9}$$

or

$$\alpha_i C_{ij} = \alpha_j C_{ji}. \tag{A10}$$

Finally, defining the symmetric matrix $\mathbf{A}$ as

$$[\mathbf{A}]_{ij} = \alpha_j^{-1} C_{ij}, \tag{A11}$$

the transfer matrix can be written in the special form

$$\mathbf{C} = \mathbf{AD}, \tag{A12}$$

where $\mathbf{D}$ is a diagonal matrix with elements $\alpha_i > 0$.

APPENDIX B. THE TRANSFER MATRIX

In certain physical models, the transfer matrix may be written as $\mathbf{C} = \mathbf{D}_1\mathbf{A}\mathbf{D}_2$, where $\mathbf{D}_1$ and $\mathbf{D}_2$ are diagonal matrices with strictly positive diagonal elements and $\mathbf{A}$ is a positive symmetric matrix. The elements of the $\boldsymbol{\Sigma}$ matrix generally will not be ordered, but will be arranged as

$$\sigma_k \geq \sigma_l \geq \cdots \geq \sigma_m > 0, \quad 1 \leq k, l, \cdots, m \leq N. \tag{B1}$$

[14] M. M. R. Williams, *The Slowing Down and Thermalization of Neutrons* (North-Holland Publishing Co., Amsterdam, 1966).

It is possible to transform Eq. (2.1) into a form which has a purely symmetric transfer matrix and an ordered matrix. First, we construct a permutation matrix $\mathbf{P}$, such that

$$\begin{aligned} &[\mathbf{P}]_{1k} = 1, \quad [\mathbf{P}]_{1i} = 0, \quad i \neq k, \\ &[\mathbf{P}]_{2l} = 1, \quad [\mathbf{P}]_{2i} = 0, \quad i \neq l, \\ &\quad\cdot \qquad\qquad\cdot \\ &\quad\cdot \qquad\qquad\cdot \\ &\quad\cdot \qquad\qquad\cdot \\ &[\mathbf{P}]_{Nm} = 1, \quad [\mathbf{P}]_{Ni} = 0, \quad i \neq m. \end{aligned} \tag{B2}$$

By multiplying Eq. (2.1) from the left by $\mathbf{P}$, one obtains

$$\left[\mu \frac{\partial}{\partial x}\mathbf{E} + \boldsymbol{\Sigma}'\right]\boldsymbol{\psi}'(x,\mu) = \mathbf{D}_1'\mathbf{A}'\mathbf{D}_2' \int_{-1}^{1} d\mu\boldsymbol{\psi}'(x,\mu), \tag{B3}$$

where

$$\begin{aligned} \boldsymbol{\psi}'(x,\mu) &= \mathbf{P}\boldsymbol{\psi}(x,\mu), \\ \boldsymbol{\Sigma}' &= \mathbf{P}\boldsymbol{\Sigma}\mathbf{P}^{-1}, \\ \mathbf{A}' &= \mathbf{P}\mathbf{A}\mathbf{P}^{-1}, \\ \mathbf{D}_i' &= \mathbf{P}\mathbf{D}_i\mathbf{P}^{-1}, \quad i = 1, 2. \end{aligned} \tag{B4}$$

Since $\mathbf{P}^{-1} = \tilde{\mathbf{P}}$, it can be shown by inspection that $\boldsymbol{\Sigma}'$ is a diagonal matrix with ordered elements

$$\sigma_1' \geq \sigma_2' \geq \cdots \geq \sigma_N'. \tag{B5}$$

Furthermore, $\mathbf{D}_1'$ and $\mathbf{D}_2'$ are diagonal matrices with positive diagonal elements and $\mathbf{A}'$ is symmetric.

Now we define the diagonal matrices $\mathbf{D}_i^{\frac{1}{2}}$ and $\mathbf{D}_i^{-\frac{1}{2}}$ as

$$[\mathbf{D}_i^{\pm\frac{1}{2}}]_{jk} = \{[\mathbf{D}_i]_{jk}\}^{\pm\frac{1}{2}}, \quad i = 1, 2. \tag{B6}$$

Multiplying Eq. (B3) from the left by $\mathbf{D}_1^{-\frac{1}{2}}\mathbf{D}_2^{\frac{1}{2}}$, we have

$$\left[\mu \frac{\partial}{\partial x}\mathbf{E} + \boldsymbol{\Sigma}'\right]\boldsymbol{\psi}''(x,\mu) = \mathbf{A}'' \int_{-1}^{1} d\mu\boldsymbol{\psi}''(x,\mu), \tag{B7}$$

where

$$\begin{aligned} \boldsymbol{\psi}''(x,\mu) &= \mathbf{D}_1^{-\frac{1}{2}}\mathbf{D}_2^{\frac{1}{2}}\boldsymbol{\psi}''(x,\mu), \\ \mathbf{A}'' &= \mathbf{D}_1^{\frac{1}{2}}\mathbf{D}_2^{\frac{1}{2}}\mathbf{A}'\mathbf{D}_1^{\frac{1}{2}}\mathbf{D}_2^{\frac{1}{2}} = \tilde{\mathbf{A}}''. \end{aligned} \tag{B8}$$

For the two-group model, there exists a transformation $\mathbf{S}$ which will symmetrize any strictly positive $\mathbf{C}$ matrix and leave $\boldsymbol{\Sigma}$ diagonal, namely,

$$\mathbf{S} = \begin{pmatrix} 0 & (C_{12})^{\frac{1}{2}} \\ (C_{21})^{\frac{1}{2}} & 0 \end{pmatrix}. \tag{B9}$$

On the other hand, if one or both off-diagonal elements are zero, the resulting multigroup equations can be solved consecutively by applying one-speed theory.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 12 DECEMBER 1969

Substitution Group and the Stretched Isoscalar Factors for the Group R_5

S. J. ALIŠAUSKAS AND A. P. JUCYS
Institute of Physics and Mathematics of the Academy of Sciences of Lithuanian SSR, Vilnius

(Received 31 August 1968)

The phase relations for basis functions and Clebsch–Gordan coefficients of the representations of the group R_5 under the elements of the substitution group are given. The stretched isoscalar factors as well as the semistretched factors of the first kind are expressed in terms of the quantities of the theory of representations of SU_2.

1. INTRODUCTION

In contemporary theoretical and mathematical physics the theory of representations of semisimple Lie groups assumes ever greater importance. The simplest of these, SU_2, is of the first rank and has been investigated in detail. Much attention has also been paid to the representations of the groups of the second rank, SU_3 and R_5. The Clebsch–Gordan coefficients of the representations of these groups are under intensive examination. For SU_3 they have been examined in a more detailed way than those of R_5.

The representations of R_5 are considered in a review paper by Behrends *et al.*[1] and in a paper of Hecht.[2] In the latter, expressions are given for the isoscalar factors (i.f.) of the Clebsch–Gordan coefficients when one of the representations is a fundamental, or the regular, one. In this paper we are going to find expressions for more general i.f. However, we are limiting ourselves to particular types of them. These will be the stretched ones and one kind of semistretched i.f.

For SU_2 and SU_3 the substitutions of the parameters characterizing the representations were very useful.[3,4,5] These substitutions constitute the group isomorphous to the Weyl group of the same Lie group. In this paper we are going to find such a substitution group, the corresponding phase relations for the basis functions and its connection with the contragredience operation for R_5. We give the phase relations for those Clebsch–Gordan coefficients which are to be used for the simplification of the calculations to be carried out in order to obtain the formulas for some of the i.f.

We use methods similar to those of Sharp and von Baeyers[6] and of Ališauskas and Jucys[7] for obtaining the recurrence formulas for the i.f. of the stretched Clebsch–Gordan coefficients. By the term "stretched" we understand such a coupling in which the highest weight of the resulting representation is equal to the sum of highest weights of the representations being coupled. The i.f. obtained from the stretched ones with the help of the elements of the substitution group are considered as stretched i.f.

In obtaining the expression for semistretched i.f. of the first kind (a concept to be introduced in Sec. 7), we use the formulas for those of the stretched ones. In all cases the i.f. are expressed in terms of quantities of the theory of representations of SU_2, which are examined in detail in Ref. 3.

2. THE SUBSTITUTION GROUP OF THE GROUP R_5

The group R_5 is isomorphous to the symplectic group Sp_4 and has 10 infinitesimal operators. The notations and commutation relations are given in Ref. 2. We denote the representations by the symbol $\langle K\Lambda\rangle$ which corresponds to $(J_m\Lambda_m)$ of Ref. 2; K and Λ may take on integer or half-integer values independently of one another. They are related to integers λ_1 and λ_2 by the expressions

$$K = \tfrac{1}{2}(\lambda_1 + \lambda_2), \quad \Lambda = \tfrac{1}{2}\lambda_2, \tag{1}$$

λ_1 and λ_2 being the coefficients in the expression for the dominant weight in terms of the fundamental dominant weights [formula (III.1) of Ref. 1].

For labeling the basis functions we shall use the reduction scheme

$$R_5 \supset SU_2 \times SU_2$$

rather than

$$R_5 \supset R_4 \supset R_3 \supset R_2,$$

because in the last case we lose one weight component. We label the representations of the subgroups SU_2

[1] R. E. Behrends, J. Dreitlein, C. Fronsdal, and W. Lee, Rev. Mod. Phys. **34**, 1 (1962).
[2] K. T. Hecht, Nucl. Phys. **63**, 177 (1965).
[3] A. P. Jucys and A. A. Bandzaitis, *The Theory of Angular Momentum in Quantum Mechanics* ("Mintis," Vilnius, 1965), in Russian.
[4] S. J. Ališauskas, Z. D. Rudzikas, and A. P. Jucys, Dokl. Akad. Nauk—SSSR **172**, 58 (1967).
[5] S. J. Ališauskas and A. P. Jucys, J. Math. Phys. **8**, 2250 (1967).
[6] R. T. Sharp and H. von Baeyers, J. Math. Phys. **7**, 1105 (1966).
[7] S. J. Ališauskas and A. P. Jucys, Preprint, 1968.

by I and J and their basis functions by M and N, respectively. The basis functions of the representation of R_5 are

$$\left|\begin{matrix}\langle K\Lambda\rangle \\ IMJN\end{matrix}\right\rangle. \tag{2}$$

They are invariant (up to a sign) under the substitutions

$$K\Lambda \to -K-2, \Lambda, \tag{3a}$$
$$\to K, -\Lambda-1, \tag{3b}$$
$$\to \Lambda-\tfrac{1}{2}, K+\tfrac{1}{2}, \tag{3c}$$

and their compositions. All eight substitutions (including identity) constitute the substitution group isomorphous to the Weyl group of R_5. To these substitutions corresponds the similarity transformation of basis functions, i.e., the passage to the equivalent representations. These similarity transformations reduce to the phase relations between the basis functions of equivalent representations. These phase relations may be found by the use of the explicit expressions for the matrix elements of the infinitesimal operators of the group.[2] It is easy to see that the phase factors depend on the parameters I, M, J, and N in the following way:

$$\left|\begin{matrix}\langle K\Lambda\rangle \\ IMJN\end{matrix}\right\rangle = (-1)^{I-J+K-\Lambda}\left|\begin{matrix}\langle \Lambda-\frac{1}{2}, K+\frac{1}{2}\rangle \\ IMJN\end{matrix}\right\rangle \tag{4a}$$
$$= (-1)^{I+J-K-\Lambda}\left|\begin{matrix}\langle -K-2, \Lambda\rangle \\ IMJN\end{matrix}\right\rangle \tag{4b}$$
$$= (-1)^{2J}\left|\begin{matrix}\langle -\Lambda-\frac{3}{2}, K+\frac{1}{2}\rangle \\ IMJN\end{matrix}\right\rangle \tag{4c}$$
$$= (-1)^{I+J-K+\Lambda}\left|\begin{matrix}\langle K, -\Lambda-1\rangle \\ IMJN\end{matrix}\right\rangle \tag{4d}$$
$$= (-1)^{2J}\left|\begin{matrix}\langle \Lambda-\frac{1}{2}, -K-\frac{3}{2}\rangle \\ IMJN\end{matrix}\right\rangle \tag{4e}$$
$$= (-1)^{I+J-K+\Lambda}\left|\begin{matrix}\langle -K-2, -\Lambda-1\rangle \\ IMJN\end{matrix}\right\rangle \tag{4f}$$
$$= (-1)^{2J}\left|\begin{matrix}\langle -\Lambda-\frac{3}{2}, -K-\frac{3}{2}\rangle \\ IMJN\end{matrix}\right\rangle \tag{4g}$$

It must be noted that these phase relations enable us to compare two basis functions, one of which is defined by parameters in the normal region, the parameters of the second one being subjected to the transformation according to the elements of the substitution group. The parameters K and Λ are included in the phase factors in order to make them real. The dependence of the phase factors on K and Λ is connected with the phase system of the representations.

Realizing the mirror-reflection symmetry transformation[3] in SU_2, we obtain the phase relations

$$\left|\begin{matrix}\langle K\Lambda\rangle \\ IMJN\end{matrix}\right\rangle = (-1)^{J+M-K+\Lambda}\left|\begin{matrix}\langle K\Lambda\rangle \\ -I-1, MJN\end{matrix}\right\rangle \tag{5a}$$
$$= (-1)^{I+N-K+\Lambda}\left|\begin{matrix}\langle K\Lambda\rangle \\ IM, -J-1, N\end{matrix}\right\rangle \tag{5b}$$
$$= (-1)^{M+N-K+\Lambda}\left|\begin{matrix}\langle K\Lambda\rangle \\ -I-1, M, -J-1, N\end{matrix}\right\rangle, \tag{5c}$$

where the dependence of the phases on K and Λ is brought into accordance with the phase system of Ref. 2. It is appropriate to mention that, upon performing several substitutions in succession, the resultant factors are not simply the products of the individual phase factors.

Of particular importance is the relation

$$\left|\begin{matrix}\langle -K-2, -\Lambda-1\rangle \\ -I-1, M, -J-1, N\end{matrix}\right\rangle = (-1)^{I+J+M+N}\left|\begin{matrix}\langle K\Lambda\rangle \\ IMJN\end{matrix}\right\rangle. \tag{6}$$

Here the phase factor is equal to the phase factor of the contragredience transformation. For this reason we obtain the relation

$$\left|\begin{matrix}\langle K\Lambda\rangle \\ IMJN\end{matrix}\right\rangle^{\dagger} = \left|\begin{matrix}\langle -K-2, -\Lambda-1\rangle \\ -I-1, -M, -J-1, -N\end{matrix}\right\rangle, \tag{7}$$

in analogy with that for the representations of SU_n (Ref. 5). The substitutions

$$M, N \to -M, -N \tag{8a}$$

mean the inversion of the weight space and

$$K, \Lambda, I, J \to -K-2, -\Lambda-1, -I-1, -J-1 \tag{8b}$$

is to be interpreted as the reflection of the coordinate system of the weight space or that of the coordinates of commuting infinitesimal operators with respect to the rest of the whole space. This constitutes the generalized mirror reflection symmetry.[3]

3. THE STRETCHED ISOSCALAR FACTORS OF SYMMETRIC REPRESENTATIONS

The symmetric representations are $\langle K0\rangle$ and $\langle \Lambda\Lambda\rangle$. The first one we shall call the representation of symmetrized spinors and the second one that of symmetrized vectors. In both cases the parameters I and

J of the basis functions are linearly dependent. In the first case $I + J = K$, and in the second $I = J \leq \Lambda$. The aim of this section is to find explicit expressions for the i.f. giving the coupling of two representations of the maximal highest weight.

We denote the general Clebsch–Gordan coefficient of R_5 by

$$\begin{bmatrix} \langle K_1\Lambda_1\rangle & \langle K_2\Lambda_2\rangle & \langle K\Lambda\rangle_\omega \\ I_1M_1J_1N_1 & I_2M_2J_2N_2 & IMJN \end{bmatrix} = \begin{bmatrix} \langle K_1\Lambda_1\rangle & \langle K_2\Lambda_2\rangle & \langle K\Lambda\rangle_\omega \\ I_1J_1 & I_2J_2 & IJ \end{bmatrix} \begin{bmatrix} I_1I_2I \\ M_1M_2M \end{bmatrix} \begin{bmatrix} J_1J_2J_2 \\ N_1N_2N \end{bmatrix}. \tag{9}$$

The first factor of the right-hand side is the i.f., and the remaining two factors are the Clebsch–Gordan coefficients of the subgroups SU_2. The index ω labels the representations in the case of multiplicity. In this paper this index will not be needed.

We use the method of Ref. 6 to obtain the recurrence formulas; it is based on the fact that the result of the coupling of three basis functions into a maximal one does not depend on the coupling scheme. The generalized Clebsch–Gordan coefficients must be equal to each other, independently of the coupling scheme. Both sides of the equality are to be multiplied by two Clebsch–Gordan coefficients of SU_2 of such a kind that subsequent summation with respect to parameters M and N leaves only isoscalar factors on the one side and the same number of i.f. and two recoupling matrices of SU_2 on the other side.

For representations of symmetrized spinors one obtains the recurrence formula

$$\begin{bmatrix} \langle K_2 - \frac{1}{2}, 0\rangle & \langle \frac{1}{2}0\rangle & \langle K_20\rangle \\ I_2 - \frac{1}{2}, K_2 - I_2 & \frac{1}{2}0 & I_2, K_2 - I_2 \end{bmatrix} \begin{bmatrix} \langle K_10\rangle & \langle K_20\rangle & \langle K_1 + K_2, 0\rangle \\ I_1, K_1 - I_1 & I_2, K_2 - I_2 & I_1 + I_2, K_1 + K_2 - I_1 - I_2 \end{bmatrix}$$
$$= \begin{bmatrix} \langle K_10\rangle & \langle K_2 - \frac{1}{2}, 0\rangle & \langle K_1 + K_2 - \frac{1}{2}, 0\rangle \\ I_1, K_1 - I_1 & I_2 - \frac{1}{2}, K_2 - I_2 & I_1 + I_2 - \frac{1}{2}, K_1 + K_2 - I_1 - I_2 \end{bmatrix}$$
$$\times \begin{bmatrix} \langle K_1 + K_2 - \frac{1}{2}, 0\rangle & \langle \frac{1}{2}0\rangle & \langle K_1 + K_2, 0\rangle \\ I_1 + I_2 - \frac{1}{2}, K_1 + K_2 - I_1 - I_2 & \frac{1}{2}0 & I_1 + I_2, K_1 + K_2 - I_1 - I_2 \end{bmatrix}$$
$$\times \langle I_1, I_2 - \tfrac{1}{2}\tfrac{1}{2}(I_2)I_1 + I_2 \,|\, I_1I_2 - \tfrac{1}{2}(I_1 + I_2 - \tfrac{1}{2})\tfrac{1}{2}I_1 + I_2\rangle. \tag{10}$$

(Here one of two recoupling matrices of SU_2 equals unity.) The parameters of the types I and J satisfy the relations $I = I_1 + I_2$, $J = J_1 + J_2$. Taking the expression for the i.f. containing the representation $\langle \frac{1}{2}, 0\rangle$ from the table of Ref. 2 and then using the recurrence formulas of the type (10), one obtains the expression

$$\begin{bmatrix} \langle K_10\rangle & \langle K_20\rangle & \langle K_1 + K_2, 0\rangle \\ I_1, K_1 - I_1 & I_2, K_2 - I_2 & I_1 + I_2, K_1 + K_2 - I_1 - I_2 \end{bmatrix} = \left[\frac{(2K_1)!\,(2K_2)!\,(2I_1 + 2I_2)!\,(2K_1 + 2K_2 - 2I_1 - 2I_2)!}{(2K_1 + 2K_2)!\,(2I_1)!\,(2K_1 - 2I_1)!\,(2I_2)!\,(2K_2 - 2I_2)!}\right]^{\frac{1}{2}}, \tag{11}$$

which is normalized (equals unity when $I_1 = K_1$ and $I_2 = K_2$).

For the i.f. coupling the representations of symmetrized vectors we obtain

$$\begin{bmatrix} \langle \Lambda_1\Lambda_1\rangle & \langle \Lambda_2\Lambda_2\rangle & \langle \Lambda_1 + \Lambda_2, \Lambda_1 + \Lambda_2\rangle \\ I_1I_1 & I_2I_2 & II \end{bmatrix} = (-1)^{I_1+I_2-I}\left[\frac{2(2I_1 + 1)(2I_2 + 1)(4\Lambda_1 + 1)!\,(4\Lambda_2 + 1)!\,(2\Lambda_1 + 2\Lambda_2 - 2I)!\,(2\Lambda_1 + 2\Lambda_2 + 2I + 2)!}{(2I + 1)(4\Lambda_1 + 4\Lambda_2 + 1)!\,(2\Lambda_1 - 2I_1)!\,(2\Lambda_1 + 2I_1 + 2)!\,(2\Lambda_2 - 2I_2)!\,(2\Lambda_2 + 2I_2 + 2)!}\right]. \tag{12}$$

Using the recurrence formula

$$\begin{bmatrix} \langle \Lambda_1\Lambda_1\rangle & \langle \Lambda_2\Lambda_2\rangle & \langle \Lambda_1 + \Lambda_2, \Lambda_1 + \Lambda_2\rangle \\ I_1I_1 & I_2I_2 & II \end{bmatrix} \begin{bmatrix} \langle \Lambda_2 - \frac{1}{2}, \Lambda_2 - \frac{1}{2}\rangle & \langle \frac{1}{2}\frac{1}{2}\rangle & \langle \Lambda_2\Lambda_2\rangle \\ I_2I_2 & 00 & I_2I_2 \end{bmatrix}$$
$$= \begin{bmatrix} \langle \Lambda_1\Lambda_1\rangle & \langle \Lambda_2 - \frac{1}{2}, \Lambda_2 - \frac{1}{2}\rangle & \langle \Lambda_1 + \Lambda_2 - \frac{1}{2}, \Lambda_1 + \Lambda_2 - \frac{1}{2}\rangle \\ I_1I_1 & I_2I_2 & II \end{bmatrix}$$
$$\times \begin{bmatrix} \langle \Lambda_1 + \Lambda_2 - \frac{1}{2}, \Lambda_1 + \Lambda_2 - \frac{1}{2}\rangle & \langle \frac{1}{2}\frac{1}{2}\rangle & \langle \Lambda_1 + \Lambda_2, \Lambda_1 + \Lambda_2\rangle \\ II & 00 & II \end{bmatrix} \tag{13}$$

(with recoupling matrices of SU_2 equal to unity) and the recurrence formula obtained by transposition of $\Lambda_1, I_1, \Lambda_2, I_2$, expression (12) is brought to the form in which $I_1 = \Lambda_1$ and $I_2 = \Lambda_2$. Afterwards, we take the formula, analogous to (13), in which one of the basis functions of the representations $\langle \frac{1}{2}, \frac{1}{2} \rangle$ has $i_1 = i_2 = \frac{1}{2}$. These operations lead us to expression (12).

It is more difficult to prove the expression

$$\begin{bmatrix} \langle K0\rangle & \langle \Lambda\Lambda\rangle & \langle K+\Lambda, \Lambda\rangle \\ K-J_1, J_1 & I_2 I_2 & IJ \end{bmatrix} = (-1)^{K+2I_2-I-J} \begin{bmatrix} \langle \Lambda\Lambda\rangle & \langle K0\rangle & \langle K+\Lambda, \Lambda\rangle \\ I_2 I_2 & K-J_1, J_1 & IJ \end{bmatrix} \tag{14a}$$

$$= \frac{B[\langle K+\Lambda, \Lambda\rangle IJ]\nabla(KIJ)}{\nabla(K-J_1, I_2, I)\nabla(J_1 I_2 J)} \left[\frac{2(2I_2+1)(2K)!\,(4\Lambda+1)!}{(2\Lambda-2I_2)!\,(2\Lambda+2I_2+2)!} \right]^{\frac{1}{2}}. \tag{14b}$$

Here and in the following text we use the notation

$$B[\langle K\Lambda\rangle IJ] = \left[\frac{(K+\Lambda-I-J)!\,(K+\Lambda-I+J+1)!\,(K+\Lambda+I-J+1)!\,(K+\Lambda+I+J+2)!}{(2\Lambda)!\,(2K+1)!\,(2K-2\Lambda)!\,(2K+2\Lambda+2)!} \right]^{\frac{1}{2}}, \tag{15}$$

$$\nabla(abc) = \left[\frac{(a+b-c)!\,(a-b+c)!\,(a+b+c+1)!}{(-a+b+c)!} \right]^{\frac{1}{2}}. \tag{16}$$

In order to prove expression (14b), we first obtain a formula similar to (13), and with its help we express the i.f. (14b) through the one with $\Lambda = I_2$. Afterwards, in analogy to (10), we obtain the recurrence formula

$$\begin{aligned} &\begin{bmatrix} \langle K0\rangle & \langle \Lambda\Lambda\rangle & \langle K+\Lambda, \Lambda\rangle \\ K-J_1, J_1 & \Lambda\Lambda & IJ \end{bmatrix} (2I_1)(2I+1)[(2K+2\Lambda+1)(2K+4\Lambda+1)]^{\frac{1}{2}} \\ &\quad = [(K-I+J)(K-J_1+\Lambda-I)(I+J-K+1)(I-K+J_1+\Lambda+1) \\ &\qquad \times (K+2\Lambda-I-J)(2\Lambda+K-I+J+1)]^{\frac{1}{2}} \\ &\qquad \times \begin{bmatrix} \langle K-\frac{1}{2}, 0\rangle & \langle \Lambda\Lambda\rangle & \langle K+\Lambda-\frac{1}{2}, \Lambda\rangle \\ K-J_1-\frac{1}{2}, J_1 & \Lambda\Lambda & I+\frac{1}{2}, J \end{bmatrix} \\ &\qquad + [(K+I+J+1)(K+I-J)(K-J_1-\Lambda+I)(K-J_1+\Lambda+I+1) \\ &\qquad \times (2\Lambda+K+I-J+1)(2\Lambda+K+I+J+2)]^{\frac{1}{2}} \\ &\qquad \times \begin{bmatrix} \langle K-\frac{1}{2}, 0\rangle & \langle \Lambda\Lambda\rangle & \langle K+\Lambda-\frac{1}{2}, \Lambda\rangle \\ K-J_1-\frac{1}{2}, J_1 & \Lambda\Lambda & I-\frac{1}{2}, J \end{bmatrix}, \end{aligned} \tag{17}$$

diminishing the parameter K.

Formula (17) resembles the recurrence formula (23.16) of Ref. 3 for the $6j$ coefficient of SU_2. Disregarding the last two factors under the square roots in the terms on the right-hand side, we observe that (17) is satisfied by the stretched $6j$ coefficient

$$\begin{Bmatrix} K-J_1 & J_1 & K \\ J & I & I_2 \end{Bmatrix}. \tag{18}$$

The remaining two factors give no complementary selection rules. This shows that the i.f. (14b) is proportional to (18). To complete the demonstration of formula (14b), several other recurrence formulas must be used until the maximal weights are obtained.

4. THE STRETCHED ISOSCALAR FACTORS OF THE GENERAL FORM

By the methods of Ref. 6 we are going to find the expression for general stretched i.f. of R_5. We couple basis functions of four symmetric representations by two different coupling schemes with the help of the formulas of Sec. 3 and the Clebsch–Gordan coefficients of SU_2. Then we compare the coefficients at the

functions with the same coupling scheme relative to subgroups SU_2. The result is

$$\begin{bmatrix} \langle K_1\Lambda_1\rangle & \langle K_2\Lambda_2\rangle & \langle K_1+K_2,\Lambda_1+\Lambda_2\rangle \\ I_1J_1 & I_2J_2 & IJ \end{bmatrix} \begin{bmatrix} \langle K_1-\Lambda_1,0\rangle & \langle\Lambda_1\Lambda_1\rangle & \langle K_1\Lambda_1\rangle \\ i_1j_1 & i_1'i_1' & I_1J_1 \end{bmatrix} \begin{bmatrix} \langle K_2-\Lambda_2,0\rangle & \langle\Lambda_2\Lambda_2\rangle & \langle K_2\Lambda_2\rangle \\ i_2j_2 & i_2'i_2' & I_2J_2 \end{bmatrix}$$
$$= \begin{bmatrix} \langle K_1-\Lambda_1,0\rangle & \langle K_2-\Lambda_2,0\rangle & \langle K_1+K_2-\Lambda_1-\Lambda_2,0\rangle \\ i_1j_1 & i_2j_2 & i_1+i_2,j_1+j_2 \end{bmatrix}$$
$$\times \sum_{I'} \begin{bmatrix} \langle\Lambda_1\Lambda_1\rangle & \langle\Lambda_2\Lambda_2\rangle & \langle\Lambda_1+\Lambda_2,\Lambda_1+\Lambda_2\rangle \\ i_1'i_1' & i_2'i_2' & I'I' \end{bmatrix}$$
$$\times \begin{bmatrix} \langle K_1+K_2-\Lambda_1-\Lambda_2,0\rangle & \langle\Lambda_1+\Lambda_2,\Lambda_1+\Lambda_2\rangle & \langle K_1+K_2,\Lambda_1+\Lambda_2\rangle \\ i_1+i_2,j_1+j_2 & I'I' & IJ \end{bmatrix}$$
$$\times \langle i_1i_2(i_1+i_2)i_1'i_2'(I')I \mid i_1i_1'(I_1)i_2i_2'(I_2)I\rangle\langle j_1j_2(j_1+j_2)i_1'i_2'(I')J \mid j_1i_1'(J_1)j_2i_2'(J_2)J\rangle. \quad (19)$$

Here $i_1+j_1=K_1-\Lambda_1$ and $i_2+j_2=K_2-\Lambda_2$. Expressing the i.f. according to the formulas of Sec. 3 and using the definition (27.4) of Ref. 3, we obtain the stretched $15j$ coefficient of the type $\{2,2\}$. Bisecting the diagrams of this coefficient by the methods of Ref. 3, we reduce it to a $9j$ coefficient because many of the triads are stretched. In this way we finally obtain the expression

$$\begin{bmatrix} \langle K_1\Lambda_1\rangle & \langle K_2\Lambda_2\rangle & \langle K_1+K_2,\Lambda_1+\Lambda_2\rangle \\ I_1J_1 & I_2J_2 & IJ \end{bmatrix}$$
$$= (-1)^{I_1+I_2-I}\frac{B[\langle K_1+K_2,\Lambda_1+\Lambda_2\rangle IJ]}{B[\langle K_1\Lambda_1\rangle I_1J_1]B[\langle K_2\Lambda_2\rangle I_2J_2]}$$
$$\times \left[\frac{(2I_1+1)(2J_1+1)(2I_2+1)(2J_2+1)(2K_1+2K_2-2\Lambda_1-2\Lambda_2+1)!}{(2K_1-2\Lambda_1)!\,(2K_2-2\Lambda_2)!}\right]^{\frac{1}{2}}$$
$$\times \begin{Bmatrix} K_1-\Lambda_1 & K_2-\Lambda_2 & K_1+K_2-\Lambda_1-\Lambda_2 \\ I_1 & I_2 & I \\ J_1 & J_2 & J \end{Bmatrix}. \quad (20)$$

The stretched $9j$ coefficient may be expressed with the help of formula (25.20) of Ref. 3 or by the corresponding formula given by Sharp.[8] Such an expression contains two summation parameters. In more special cases this $9j$ coefficient turns into a $6j$ coefficient (which contains one summation parameter only) or into a double stretched $9j$ coefficient (which contains no summation parameters at all).

5. THE PHASE RELATIONS FOR THE ISOSCALAR FACTORS UNDER THE SUBSTITUTION GROUP

We confine ourselves to giving the phase relations for the i.f. under three elements of the substitution group and for one substitution of parameters of the subgroups SU_2. These relations are

$$\begin{bmatrix} \langle K\Lambda\rangle & \langle K'\Lambda'\rangle & \langle K+a,\Lambda+b\rangle \\ IJ & I'J' & I+\alpha,J+\beta \end{bmatrix} \quad (21a)$$
$$= \sum_{\omega'}\delta^{b}_{\omega\omega'}(-1)^{\alpha+\beta-|a|-b}\begin{bmatrix} \langle -K-2,\Lambda\rangle & \langle K'\Lambda'\rangle & \langle -K-2-a,\Lambda+b\rangle_{\omega'} \\ IJ & I'J' & I+\alpha,J+\beta \end{bmatrix} \quad (21b)$$
$$= \sum_{\omega'}\delta^{c}_{\omega\omega'}(-1)^{\alpha+\beta-a+|b|}\begin{bmatrix} \langle K_1-\Lambda-1\rangle & \langle K'\Lambda'\rangle & \langle K+a,-\Lambda-1-b\rangle_{\omega'} \\ IJ & I'J' & I+\alpha,J+\beta \end{bmatrix} \quad (21c)$$
$$= \sum_{\omega'}\delta^{d}_{\omega\omega'}(-1)^{\alpha-\beta+a+b+2\Lambda'}\begin{bmatrix} \langle \Lambda-\frac{1}{2},K+\frac{1}{2}\rangle & \langle K'\Lambda'\rangle & \langle \Lambda-\frac{1}{2}+b,K+\frac{1}{2}+a\rangle_{\omega'} \\ IJ & I'J' & I+\alpha,J+\beta \end{bmatrix} \quad (21d)$$
$$= (-1)^{J'-\alpha+a-b}\begin{bmatrix} \langle K\Lambda\rangle & \langle K'\Lambda'\rangle & \langle K+a,\Lambda+b\rangle_{\omega} \\ I,-J-1 & I'J' & I+\alpha,-J-1-\beta \end{bmatrix}. \quad (22)$$

When multiple representations are labeled in a proper way, the $\delta^{x}_{\omega\omega'}$ are diagonal. This will be discussed in our next paper. The dependence of the phase factors on the parameters characterizing the representations of R_5 are coordinated with the phases of i.f. containing fundamental representations, as given in Ref. 2.

[8] R. T. Sharp, Nucl. Phys. **A95**, 222 (1967).

From the expression (20), together with the phase relations (21), we obtain the i.f. having these resultant representations:

$$\langle K_1+\Lambda_2, \Lambda_1+K_2\rangle, \langle K_1-K_2, \Lambda_1+\Lambda_2\rangle,$$
$$\langle K_1-K_2, \Lambda_1-\Lambda_2\rangle, \langle K_1-\Lambda_2, \Lambda_1+K_2\rangle,$$
$$\langle K_1-\Lambda_2, \Lambda_1-K_2\rangle, \langle K_1+K_2, \Lambda_1-\Lambda_2\rangle,$$
$$\langle K_1+\Lambda_2, \Lambda_1-K_2\rangle. \tag{23}$$

Whenever factorials of negative numbers appear, these are to be transformed by the methods given in Sec. 10 of Ref. 3. In some cases the $9j$ coefficient undergoes no change when its parameters are reflected in the ordinary way, according to formulas (24.19)–(24.31) of Ref. 3. For several substitutions two of the parameters exceed in absolute value the sum of two other parameters of corresponding triads. These $9j$ coefficients we call nonstandard. However, they have definite values and may be calculated with the help of corresponding algebraic expressions. It must be emphasized that substitutions can be applied only in those columns of i.f. in which the parameters I, J, K, and Λ are linearly independent.

6. ISOSCALAR FACTORS COUPLING TWO REPRESENTATIONS OF SYMMETRIZED SPINORS

By graphical methods[1] it may be shown that

$$\langle K_1 0\rangle \times \langle K_2 0\rangle = \sum_{\substack{m\geq 0, n\geq 0 \\ m+n\leq 2K_2}} \langle K_1+\tfrac{1}{2}(m-n), K_2-\tfrac{1}{2}(m+n)\rangle \tag{24}$$

when $K_1 \geq K_2$.

Particular cases of (24) will be the representations $\langle K_1+K_2, 0\rangle$, $\langle K_1-K_2, 0\rangle$, and $\langle K_1 K_2\rangle$. The i.f. for the first case is given by formula (11), the second may be obtained easily from formula (A22) of Ref. 2 by permutation of parameters, and the third is obtained by applying relation (21d) to a particular case of (20) after performing a simple transformation according to the methods of Ref. 3.

For a more general case we obtain

$$\begin{bmatrix} \langle K_1 0\rangle & \langle K_2 0\rangle & \langle K_1+K_2-\Lambda, \Lambda\rangle \\ I_1, K_1-I_1 & I_2, K_2-I_2 & IJ \end{bmatrix}$$
$$= (-1)^{I_1+I_2-I}\left[\frac{(2K_1+1)!\,(2K_2+1)!\,(2K_1+2K_2-4\Lambda+1)!}{(2I_1)!\,(2K_1-2I_1)!\,(2I_2)!\,(2K_2-2I_2)!}\right]^{\frac{1}{2}}$$
$$\times B[\langle K_1+K_2-\Lambda, \Lambda\rangle IJ]\begin{Bmatrix} I_1 & I_2 & I \\ K_1-I_1 & K_2-I_2 & J \\ K_1 & K_2 & K_1+K_2-2\Lambda \end{Bmatrix}. \tag{25}$$

In order to derive this formula, the representation $\langle K_2 0\rangle$ is constructed from two representations. The resultant representation will be produced by coupling three representations. This may be done by the stretched Clebsch–Gordan coefficients. Then we take the scalar product of basis functions of initial and resultant representations which is proportional to the Clebsch–Gordan coefficient of R_5. After multiplication by Clebsch–Gordan coefficients of SU_2 and summation with respect to parameters M and N, we obtain a relation, which, after substitution of the expression for the stretched i.f., resembles one of the expressions for the Clebsch–Gordan coefficient of SU_2. This last fact allows us to obtain Eq. (25).

The most general case under consideration is

$$\begin{bmatrix} \langle K_1 0\rangle & \langle K_2 0\rangle & \langle K\Lambda\rangle \\ I_1, K_1-I_1 & I_2, K_2-I_2 & IJ \end{bmatrix} = (-1)^{I_1+I_2+J-K-\Lambda}\begin{bmatrix} I & J & K-\Lambda \\ I_1-I_2 & K_1-K_2-I_1+I_2 & K_1-K_2 \end{bmatrix}$$
$$\times \begin{bmatrix} \tfrac{1}{2}(K_1+K_2-I-J) & \tfrac{1}{2}(K_1+K_2+I+J)+1 & K+\Lambda+1 \\ \tfrac{1}{2}(K_1+K_2+I-J)-I_1-I_2 & I_1+I_2+\tfrac{1}{2}(I-J-K_1-K_2) & I-J \end{bmatrix}. \tag{26}$$

The derivation of Eq. (26) is similar to that for the preceding formula. The resultant representation may be expressed as $\langle K_1+K_2-m-\Lambda, \Lambda\rangle$ and constructed with the help of representations $\langle K_1 0\rangle$, $\langle K_2-\tfrac{1}{2}m, 0\rangle$, and $\langle \tfrac{1}{2}m, 0\rangle$ (where m = integer). Coupling the three representations to the resultant one, we obtain (26) in a way similiar to that described above.

7. SEMISTRETCHED ISOSCALAR FACTORS OF THE FIRST KIND

When the parameters of the representations satisfy the condition

$$K_1+\Lambda_1+K_2+\Lambda_2 = K+\Lambda, \tag{27}$$

we call the corresponding multiplicity-free i.f. semistretched of the first kind. I.f. of the second kind will

be those with

$$K_1 + K_2 = K. \tag{28}$$

This term includes all i.f. obtained with the help of the substitution group.

For the i.f. under consideration we obtain

$$\begin{bmatrix} \langle K_1\Lambda_1\rangle & \langle K_2\Lambda_2\rangle & \langle K_1 + K_2 - \tfrac{1}{2}m, \Lambda_1 + \Lambda_2 + \tfrac{1}{2}m\rangle \\ I_1J_1 & I_2J_2 & IJ \end{bmatrix}$$
$$= (-1)^{I_1+I_2-I} \frac{B[\langle K_1 + K_2 - \tfrac{1}{2}m, \Lambda_1 + \Lambda_2 + \tfrac{1}{2}m\rangle IJ]}{B[\langle K_1\Lambda_1\rangle I_1J_1]B[\langle K_2\Lambda_2\rangle I_2J_2]}$$
$$\times \left[\frac{(2I_1+1)(2J_1+1)(2I_2+1)(2J_2+1)(2K_1+2K_2-2\Lambda_1-2\Lambda_2-2m+1)!}{(2K_1-2\Lambda_1)!\,(2K_2-2\Lambda_2)!}\right]^{\frac{1}{2}}$$
$$\times \begin{Bmatrix} K_1-\Lambda_1 & K_2-\Lambda_2 & K_1+K_2-\Lambda_1-\Lambda_2-m \\ I_1 & I_2 & I \\ J_1 & J_2 & J \end{Bmatrix}. \tag{29}$$

Here m is an integer.

To derive Eq. (29) we use techniques similar to those used in obtaining (19). In the formula obtained in this way, which resembles Eq. (19), we substitute the expressions for the i.f. already obtained. The transformation matrices involved are expressed in terms of $3nj$ coefficients. The resulting sum must be dealt with by the methods of Ref. 3 in order to obtain Eq. (29).

It is easy to see that the preceding expressions for the i.f. are particular cases of Eq. (29). The only exception is formula (26).

The elements of the substitution group bring us to another i.f. of the same kind (with different values of m) or to that obtained by transposition of the first and third columns of (17.22) of Ref. 2 with subsequent renumbering of parameters. Some of the elements of the substitution group transfer the $9j$ coefficient to the nonstandard form, as indicated in Sec. 5.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 12 DECEMBER 1969

Spectral Analysis of Classical Central Force Motion

REESE T. PROSSER
Dartmouth College, Hanover, N.H.

(Received 13 June 1968)

It is shown here that the Liouville operator, which governs the development in time of a classical one-particle system, has an absolutely continuous spectrum for a large class of attractive central force potentials. It follows that every absolutely continuous initial distribution of a monatomic ideal gas enclosed in a spherical container must approach a steady-state distribution in time.

INTRODUCTION

It was first observed by Koopman[1] that the motion of a classical mechanical system of point particles, regarded as a motion of its phase space, preserves the volume element of the phase space, and so induces a motion on the space of square-integrable functions which preserves the norm. Koopman showed that statistical properties of the motion of the phase space were reflected in the spectral properties of the induced motion of the function space. Subsequent developments have established the fundamental importance of this observation. In particular, the study of the ergodic and steady-state properties of the system has found a natural expression within this framework and has yielded a considerable body of results.

If the spectrum of the motion in the function space is absolutely continuous (apart from the zero eigenvalue), then the statistical approach to steady state of

[1] B. O. Koopman, Proc. Natl. Acad. Sci. U.S. **17**, 315 (1931).

be those with

$$K_1 + K_2 = K. \tag{28}$$

This term includes all i.f. obtained with the help of the substitution group.

For the i.f. under consideration we obtain

$$\begin{bmatrix} \langle K_1\Lambda_1\rangle & \langle K_2\Lambda_2\rangle & \langle K_1 + K_2 - \tfrac{1}{2}m, \Lambda_1 + \Lambda_2 + \tfrac{1}{2}m\rangle \\ I_1J_1 & I_2J_2 & IJ \end{bmatrix}$$
$$= (-1)^{I_1+I_2-I} \frac{B[\langle K_1 + K_2 - \tfrac{1}{2}m, \Lambda_1 + \Lambda_2 + \tfrac{1}{2}m\rangle IJ]}{B[\langle K_1\Lambda_1\rangle I_1J_1]B[\langle K_2\Lambda_2\rangle I_2J_2]}$$
$$\times \left[\frac{(2I_1+1)(2J_1+1)(2I_2+1)(2J_2+1)(2K_1+2K_2-2\Lambda_1-2\Lambda_2-2m+1)!}{(2K_1-2\Lambda_1)!\,(2K_2-2\Lambda_2)!}\right]^{\frac{1}{2}}$$
$$\times \begin{Bmatrix} K_1-\Lambda_1 & K_2-\Lambda_2 & K_1+K_2-\Lambda_1-\Lambda_2-m \\ I_1 & I_2 & I \\ J_1 & J_2 & J \end{Bmatrix}. \tag{29}$$

Here m is an integer.

To derive Eq. (29) we use techniques similar to those used in obtaining (19). In the formula obtained in this way, which resembles Eq. (19), we substitute the expressions for the i.f. already obtained. The transformation matrices involved are expressed in terms of $3nj$ coefficients. The resulting sum must be dealt with by the methods of Ref. 3 in order to obtain Eq. (29).

It is easy to see that the preceding expressions for the i.f. are particular cases of Eq. (29). The only exception is formula (26).

The elements of the substitution group bring us to another i.f. of the same kind (with different values of m) or to that obtained by transposition of the first and third columns of (17.22) of Ref. 2 with subsequent renumbering of parameters. Some of the elements of the substitution group transfer the $9j$ coefficient to the nonstandard form, as indicated in Sec. 5.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 12 DECEMBER 1969

Spectral Analysis of Classical Central Force Motion

REESE T. PROSSER
Dartmouth College, Hanover, N.H.

(Received 13 June 1968)

It is shown here that the Liouville operator, which governs the development in time of a classical one-particle system, has an absolutely continuous spectrum for a large class of attractive central force potentials. It follows that every absolutely continuous initial distribution of a monatomic ideal gas enclosed in a spherical container must approach a steady-state distribution in time.

INTRODUCTION

It was first observed by Koopman[1] that the motion of a classical mechanical system of point particles, regarded as a motion of its phase space, preserves the volume element of the phase space, and so induces a motion on the space of square-integrable functions which preserves the norm. Koopman showed that statistical properties of the motion of the phase space were reflected in the spectral properties of the induced motion of the function space. Subsequent developments have established the fundamental importance of this observation. In particular, the study of the ergodic and steady-state properties of the system has found a natural expression within this framework and has yielded a considerable body of results.

If the spectrum of the motion in the function space is absolutely continuous (apart from the zero eigenvalue), then the statistical approach to steady state of

[1] B. O. Koopman, Proc. Natl. Acad. Sci. U.S. **17**, 315 (1931).

the system takes a particularly simple form.[2] It is therefore of interest to determine for which system the spectrum of the motion is absolutely continuous. We have already verified that this is the case for a one-dimensional particle subject to a nonlinear restoring force.[2] In this note we show that this is also the case for a three-dimensional particle subject to a nonlinear central restoring force, by relating the spectrum to the fundamental periods of the motion.

It follows that every absolute continuous initial distribution on the phase space of this system approaches a steady-state distribution in the far reaches of time, and that the entropy of the initial distribution ultimately increases to the entropy of the steady-state distribution. An extension of the same arguments to the analogous n-body problem shows that the same conclusion holds for a monatomic ideal gas enclosed in a spherical container.

1. THE CENTRAL FORCE SYSTEM[3]

For our canonical coordinates of the one-particle central force system, we shall take the spherical coordinates (r, θ, φ) of the particle position and the conjugate momenta $(p_r, p_\theta, p_\varphi)$, with

$$p_r = m\frac{dr}{dt}, \qquad p_\theta = mr^2\frac{d\theta}{dt}, \qquad p_\varphi = mr^2\sin^2\theta\frac{d\varphi}{dt}, \tag{1.1}$$

where m is the particle mass. In these coordinates the total energy of the particle is

$$H = T + V, \tag{1.2}$$

where the kinetic energy T is given by

$$T = (2m)^{-1}(p_r^2 + p_\theta^2/r^2 + p_\varphi^2/r^2\sin^2\theta), \tag{1.3}$$

and the potential energy V is a function of r alone.

We shall assume throughout that $V(r)$ is a differentiable monotonic increasing function of r, defined for all $r \neq 0$, so that

$$V'(r) \geq 0, \quad 0 < r < \infty. \tag{1.4}$$

In this case, the particle is subject to the central restoring force $-V'(r)$ directed toward the origin. It is known that if the total energy E of the particle does not exceed the limiting value

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

of the potential, then the motion of the particle is bounded and multiply periodic, and thus may be compactly described in terms of the associated action and angle variables. We now proceed to exploit this knowledge directly.

We first introduce Hamilton's principle function $S(r, \theta, \varphi, t)$ for the problem. This function generates a canonical transformation of the coordinates such that the transformed coordinates are constants of the motion.[4] It is obtained from the Hamilton–Jacobi equation

$$\left(\frac{\partial S}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial S}{\partial \theta}\right)^2 + \frac{1}{r^2\sin^2\theta}\left(\frac{\partial S}{\partial \varphi}\right)^2 + 2m\left(V(r) + \frac{\partial S}{\partial t}\right) = 0. \tag{1.5}$$

If we assume for S the form

$$S(r, \theta, \varphi, t) = S_r(r) + S_\theta(\theta) + S_\varphi(\varphi) + S_t(t), \tag{1.6}$$

then (1.5) separates into ordinary differential equations

$$\frac{dS_t}{dt} = -E, \tag{1.7}$$

$$\frac{dS_\varphi}{d\varphi} = a_\varphi, \tag{1.8}$$

$$\left(\frac{dS_\theta}{d\theta}\right)^2 + \frac{a_\varphi^2}{\sin^2\theta} = a_\theta^2, \tag{1.9}$$

$$\left(\frac{dS_r}{dr}\right)^2 + \frac{a_\theta^2}{r^2} + 2m[V(r) - E] = 0, \tag{1.10}$$

which can be solved for the derivatives of S, giving

$$H = -\frac{dS_t}{dt} = +E, \tag{1.11}$$

$$p_\varphi = \frac{dS_\varphi}{d\varphi} = a_\varphi, \tag{1.12}$$

$$p_\theta = \frac{dS_\theta}{d\theta} = \left(a_\theta^2 - \frac{a_\varphi^2}{\sin^2\theta}\right)^{\frac{1}{2}}, \tag{1.13}$$

$$p_r = \frac{dS_r}{dr} = \left(2m[E - V(r)] - \frac{a_\theta^2}{r^2}\right)^{\frac{1}{2}}. \tag{1.14}$$

Of the three integration constants, E denotes the energy, a_φ the angular momentum about the polar axis ($\theta = 0$), and a_θ the magnitude of the resultant angular momentum.[4]

An analysis of the motion now shows that, under our assumption (1.4) on V, the variables (r, θ, φ) are all periodic variables whenever $V(0) < E < V(\infty)$.

[2] R. T. Prosser, "An Example of Irreversibility," submitted to Am. Math. Monthly.

[3] The development in this section is orthodox and may be found in any standard text on mechanics; see, e.g., Ref. 4.

[4] H. Goldstein, *Classical Mechanics* (Addison-Wesley Publ. Co., Inc., Reading, Mass., 1950).

It follows that they may be compactly described in terms of the action variables $(j_r, j_\theta, j_\varphi)$ and their conjugate angle variables $(w_r, w_\theta, w_\varphi)$, defined as follows[4]:

$$j_i = \oint p_i\, dq_i = \oint dS_i, \quad i = r, \theta, \varphi, \tag{1.15}$$

$$w_i = \frac{\partial S_i}{\partial j_i}, \qquad i = r, \theta, \varphi. \tag{1.16}$$

Here the integrals defining the j_i are each extended over one period of the running variable, and hence are constants of the motion. The motions of the w_i are determined from

$$\frac{dw_i}{dt} = \frac{\partial H}{\partial j_i} = \nu_i, \tag{1.17}$$

where the ν_i are also constants of the motion; hence

$$w_i = \nu_i t + \beta_i. \tag{1.18}$$

Thus the ν_i are the frequencies of the angle variables and the β_i are the initial values.

Substituting from (1.12)–(1.14) in (1.15), we find that

$$j_\varphi = \oint a_\varphi\, d\varphi, \tag{1.19}$$

$$j_\theta = \oint (a_\theta^2 - a_\varphi^2/\sin^2\theta)^{\frac{1}{2}}\, d\theta, \tag{1.20}$$

$$j_r = \oint \{2m[E - V(r)] - a_\varphi^2/r^2\}^{\frac{1}{2}}\, dr. \tag{1.21}$$

The first and second of these integrals may be performed explicitly (see Ref. 4), yielding

$$j_\varphi = 2\pi a_\varphi, \tag{1.22}$$

$$j_\theta = 2\pi(a_\theta - a_\varphi). \tag{1.23}$$

The third depends on the form of $V(r)$. If we solve (1.22) and (1.23) for a_φ and a_θ and substitute in (1.21), we find

$$j_r = \oint \{2m[E - V(r)] - (j_\theta + j_\varphi)^2/4\pi^2 r^2\}^{\frac{1}{2}}\, dr. \tag{1.24}$$

Finally, we introduce the new variables (j_1, j_2, j_3) and their conjugates (w_1, w_2, w_3) via

$$\begin{aligned} j_1 &= j_r + j_\theta + j_\varphi, \\ j_2 &= j_\theta + j_\varphi, \\ j_3 &= j_\varphi, \end{aligned} \tag{1.25}$$

and

$$\begin{aligned} w_1 &= w_r, \\ w_2 &= w_\theta - w_r, \\ w_3 &= w_\varphi - w_\theta. \end{aligned} \tag{1.26}$$

Then (1.24) becomes

$$j_1 = \oint \{2m[E - V(r)] - j_2^2/4\pi^2 r^2\}^{\frac{1}{2}}\, dr + j_2. \tag{1.27}$$

From this equation we see that the total energy E, and hence the Hamiltonian, depends only on j_1 and j_2 and is independent of j_3 and the w_i. It follows that the j_i and w_3 are constants of the motion, while w_1 and w_2 depend on time through (1.18) with the frequencies ν_i given by (1.17).

Thus we have obtained a set of canonical variables whose dependence on time is particularly simple. It can be shown that each has a simple interpretation in terms of the motion of the original particle. Specifically, $2\pi j_2$ is the magnitude of the total angular momentum, and $2\pi j_3$ is its component along the polar axis. Apart from additive constants, $2\pi w_2$ is the angle of the perihelion ($r_{\min}$) from the line of nodes, and $2\pi w_3$ is the angle from the line of nodes to the axis $\varphi = 0$. Finally, $2\pi j_1$ is proportional to the area enclosed by the orbit in the (r, p_r) plane, and, apart from an additive constant, $2\pi w_1$ is the angular mean anomaly of this orbit.[4,5]

From these identifications, it is easy to see that the variables j_i and w_i are independent and assume the following ranges of values:

$$\begin{aligned} 0 \le j_1 < \infty, \\ 0 \le j_2 < \infty, \\ -j_2 \le j_3 \le +j_2, \end{aligned} \tag{1.28}$$

$$-\infty \le w_i \le +\infty. \tag{1.29}$$

Since the original position and momentum variables are periodic functions of the angle variables, however, it suffices to restrict the ranges of the w_i to

$$0 \le w_i \le 1. \tag{1.30}$$

It is now quite straightforward to verify that every point in the phase space admits a unique description in terms of the action and angle variables lying within the ranges (1.28) and (1.30). Moreover, the motion of this point in time is completely described by the equations of motion

$$j_i = \alpha_i, \tag{1.31}$$

$$w_i = \nu_i t + \beta_i, \quad \text{mod } 1. \tag{1.32}$$

Here, the α_i and β_i are the initial values of the j_i and w_i. According to (1.27), the Hamiltonian H may be expressed in terms of j_1 and j_2 alone. It follows from (1.17) that the same is true of the frequencies ν_i, with $\nu_3 = 0$.

[5] M. Born, *The Mechanics of the Atom* (Frederick Ungar Publ. Co., New York, 1960).

Finally, since the action and angle variables are canonically conjugate, the differential form $d\tau = dj_1\, dj_2\, dj_3\, dw_1\, dw_2\, dw_3$ is invariant under the motion, and hence is related to the volume element dV by the formula

$$dV = \rho\, d\tau, \tag{1.33}$$

where the weight function ρ is a constant of the motion.

The form of ρ may be determined as follows. From (1.15) we see that j_r is simply the area enclosed by the orbit in the (r, p_r) plane, and w_r is the fraction of this area mapped out in time t. It follows that the change in area induced by a change in j_r and w_r is simply $dA = dj_r\, dw_r$, from which we conclude that

$$dp_r\, dr = dj_r\, dw_r. \tag{1.34}$$

A quite similar argument shows that the same relations hold for the (θ, p_θ) plane. In the (φ, p_φ) plane, the orbit is periodic rather than closed, but the same argument applies. Combining these results, we find that

$$dp_r\, dp_\theta\, dp_\varphi\, dr\, d\theta\, d\varphi = dj_r\, dj_\theta\, dj_\varphi\, dw_r\, dw_\theta\, dw_\varphi. \tag{1.35}$$

Since the Jacobian of the transformation (1.25)–(1.26) is identically unity, we see finally that the weight function ρ is identically unity, and

$$d\tau = dV. \tag{1.36}$$

2. THE SPECTRUM OF THE MOTION

We now turn to our analysis of the spectrum of the motion.

If P is any point in the phase space, then the motion of the system in time may be completely described as a motion of P through the phase space. If P is given in terms of the action-angle coordinates, then its dependence on time may be expressed in terms of the transformation $T_t\colon P \to P_t$, where

$$T_t P = P_t = P(\mathbf{j}, \boldsymbol{\nu} t + \boldsymbol{\beta}). \tag{2.1}$$

Now let $\mathcal{H}$ denote the space of all complex-valued functions defined on the phase space and square-integrable with respect to the volume element $dV = d\tau$. Then $\mathcal{H}$ becomes a Hilbert space under the inner product

$$(f, g) = \int\!\!\int f(\mathbf{j}, \mathbf{w}) g(\mathbf{j}, \mathbf{w})\, d\mathbf{w}\, d\mathbf{j}. \tag{2.2}$$

The motion T_t of the phase space induces a motion U_t of the function space according to the formula

$$(U_t f)(P) = f_t(P) = f(P_{-t}) = f(\mathbf{j}, -\boldsymbol{\nu} t + \boldsymbol{\beta}). \tag{2.3}$$

Since T_t preserves the volume element of the phase space, U_t preserves the inner product of $\mathcal{H}$. It follows that U_t is a unitary operator on $\mathcal{H}$ which may be expressed in the form

$$U_t = \exp(-iXt), \tag{2.4}$$

where X is the self-adjoint generator with the property

$$iXf = \frac{df}{dt}. \tag{2.5}$$

To determine X, recall that the motion of f_t in time is given by

$$\frac{df}{dt} = \{H, f\}, \tag{2.6}$$

where $\{\ ,\ \}$ is the Poisson bracket. In terms of the action-angle variables, (2.6) becomes

$$\frac{df}{dt} = \sum_{i=1}^{3} \left(\frac{\partial H}{\partial j_i} \frac{\partial f}{\partial w_i} - \frac{\partial H}{\partial w_i} \frac{\partial f}{\partial j_i} \right). \tag{2.7}$$

Since H depends only on j_1 and j_2, this expression reduces to

$$\begin{aligned} \frac{df}{dt} &= \frac{\partial H}{\partial j_1} \frac{\partial f}{\partial w_1} + \frac{\partial H}{\partial j_2} \frac{\partial f}{\partial w_2} \\ &= \nu_1 \frac{\partial f}{\partial w_1} + \nu_2 \frac{\partial f}{\partial w_2}. \end{aligned} \tag{2.8}$$

Comparing (2.5) and (2.8), we find that X is given by

$$X = \frac{1}{i} \left(\nu_1 \frac{\partial}{\partial w_1} + \nu_2 \frac{\partial f}{\partial w_2} \right), \tag{2.9}$$

for all differentiable functions f in $\mathcal{H}$.

Now let $\mathcal{H}_{m,n}$ be the subspace of $\mathcal{H}$ consisting of all functions of the particular form

$$f(\mathbf{j}, \mathbf{w}) = f(\mathbf{j}, w_3) \exp[2\pi i(mw_1 + nw_2)], \tag{2.10}$$

for fixed integers m and n and arbitrary functions $f(\mathbf{j}, w_3)$. From (2.9) and (2.10) we find

$$Xf = 2\pi(m\nu_1 + n\nu_2) f. \tag{2.11}$$

On the subspace $\mathcal{H}_{m,n}$, then, X acts like multiplication by the function $2\pi(m\nu_1 + n\nu_2)$. Since ν_1 and ν_2 depend only on j_1 and j_2, it follows that X leaves this subspace invariant, and on this subspace the spectrum of X consists of all real numbers of the form $2\pi(m\nu_1 + n\nu_2)$.

A standard Fourier analysis now shows that these subspaces are pairwise orthogonal in $\mathcal{H}$, and together they span $\mathcal{H}$. We conclude that the spectrum of X consists of all real numbers of the form $2\pi(m\nu_1 + n\nu_2)$, for arbitrary integers m and n and all admissible frequencies ν_1 and ν_2.

Now let f and g be any two functions in $\mathcal{H}$, and consider the inner product $(U(t)f, g)$. The asymptotic

behavior of this inner product as $t \to \pm\infty$ is determined by the spectrum of X. In fact, we have

$$(U(t)f, g) = \sum_{m,n} (U(t)f_{mn}, g_{mn}), \tag{2.12}$$

where f_{mn} and g_{mn} are the components of f and g in $\mathcal{H}_{m,n}$. But

$$(U(t)f_{mn}, g_{mn}) = \iint \exp[-2\pi i(mv_1 + nv_2)t] f_{mn}(\mathbf{j}, \mathbf{w}) g_{mn}(\mathbf{j}, \mathbf{w})\, d\mathbf{w}\, d\mathbf{j}. \tag{2.13}$$

Hence the time dependence of $(U(t)f, g)$ devolves upon the behavior of the exponent $2\pi i(mv_1 + nv_2)t$.

If $mv_1 + nv_2 \equiv 0$, then X vanishes identically on $\mathcal{H}_{m,n}$, and (2.13) is independent of t. In this case we have

$$\lim_{t\to\pm\infty} (U(t)f_{mn}, g_{mn}) = (f_{mn}, g_{mn}). \tag{2.14}$$

If $mv_1 + nv_2 \not\equiv 0$, then X does not vanish identically on $\mathcal{H}_{m,n}$. In this case we introduce a change of variables in the (j_1, j_2) plane:

$$\begin{pmatrix} j_1 \\ j_2 \end{pmatrix} \to \begin{pmatrix} \xi \\ j_2 \end{pmatrix} = \begin{pmatrix} mv_1 + nv_2 \\ j_2 \end{pmatrix}. \tag{2.15}$$

With this change of variables, (2.13) becomes

$$(U(t)f_{mn}, g_{mn}) = \iint e^{-2\pi i\xi t} f_{mn}(\mathbf{j}, \mathbf{w}) g_{mn}(\mathbf{j}, \mathbf{w}) J^{-1}(\mathbf{j})\, d\mathbf{w}\, d\xi\, dj_2\, dj_3, \tag{2.16}$$

where J is the Jacobian

$$J(j_1, j_2) = \frac{\partial(\xi, j_2)}{\partial(j_1, j_2)} = \frac{\partial \xi}{\partial j_1}. \tag{2.17}$$

In this form we see that $U(t)$ acts on $\mathcal{H}_{m,n}$ like multiplication by $e^{-2\pi i\xi t}$, with a spectral measure of the form $J^{-1}(\mathbf{j})\, d\mathbf{w}\, d\xi\, dj_2\, dj_3$. If the Jacobian $J(j_1, j_2)$ is continuous and vanishes only on a set of Lebesgue measure zero in the (j_1, j_2) plane, then the spectral measure of $U(t)$ is absolutely continuous with respect to Lebesgue measure. It now follows from the familiar Riemann–Lebesgue lemma that the right-hand side of (2.16) vanishes asymptotically for large times. Thus we have

$$\lim_{t\to\pm\infty} (U(t)f_{mn}, g_{mn}) = 0. \tag{2.18}$$

If the Jacobian $J(j_1, j_2)$ vanishes identically in the (j_1, j_2) plane, then we redefine the change of variables (2.15) so that the roles of j_1 and j_2 are interchanged:

$$\begin{pmatrix} j_1 \\ j_2 \end{pmatrix} \to \begin{pmatrix} j_1 \\ \xi \end{pmatrix} = \begin{pmatrix} j_1 \\ mv_1 + nv_2 \end{pmatrix}. \tag{2.19}$$

With this change of variables, (2.16) and (2.17) are replaced by similar formulas with j_1 and j_2 interchanged. In particular,

$$J(j_1, j_2) = \frac{\partial(j_1, \xi)}{\partial(j_1, j_2)} = \frac{\partial \xi}{\partial j_2}. \tag{2.20}$$

If this Jacobian is continuous and vanishes only on a set of Lebesgue measure zero in the (j_1, j_2) plane, then the preceding analysis leads again to (2.18).

If both Jacobians $\partial\xi/\partial j_1$ and $\partial\xi/\partial j_2$ vanish identically in the (j_1, j_2) plane, then $\xi = mv_1 + nv_2$ is actually independent of j_1 and j_2. Since $v_i = \partial E/\partial j_i$, it follows in this case that

$$\begin{aligned} m\frac{\partial^2 E}{\partial j_1^2} + n\frac{\partial^2 E}{\partial j_1 \partial j_2} &= 0, \\ m\frac{\partial^2 E}{\partial j_2 \partial j_1} + n\frac{\partial^2 E}{\partial j_2^2} &= 0. \end{aligned} \tag{2.21}$$

In this case, then, E as a function of j_1 and j_2 must satisfy (2.21). If $m, n \neq 0$, this means that $E = aj_1 + bj_2 + c$ for suitable constants a, b, and c. If $m = 0$, then $E = \varphi_1(j_1) + bj_2 + c$, while if $n = 0$, $E = aj_1 + \varphi_2(j_2) + c$, for suitable functions φ_1 and φ_2. If $m = n = 0$, then $\xi \equiv 0$, and X vanishes on $\mathcal{H}_{m,n}$.

We conclude from this analysis that if $m \neq 0$ and $\partial\xi/\partial j_1 = \partial\xi/\partial j_2 = 0$, then E depends linearly on j_1, and v_1 is independent of j_1. In particular, this remains true if we put $j_2 = 0$. But if $j_2 = 0$, then the angular momentum vanishes, and the motion reduces to that of a simple oscillator of constant frequency. It has been shown by Levin and Shatz[6] that such an oscillator must be harmonic, i.e., that we must have $V(r) = kr^2$. In this case it is well known that $E = aj_1$, and $v_1 = \text{const}$, $v_2 = 0$.

If $m = 0$ but $n \neq 0$, and $\partial\xi/\partial j_1 = \partial\xi/\partial j_2 = 0$, then E depends linearly on j_2, and v_2 is independent of j_2. In particular, this remains true if we put $j_2 = 0$. But if $j_2 = 0$, then the angular momentum again vanishes, and the motion again reduces to that of a simple oscillator. In this case there can be no precession, and we conclude that $v_2 \equiv 0$. Since $m = 0$, it follows that $\xi \equiv 0$, and X vanishes identically on $\mathcal{H}_{m,n}$.

Summarizing briefly, we have shown that if either $\partial\xi/\partial j_1$ or $\partial\xi/\partial j_2$ vanishes only on a set of Lebesgue measure zero in the (j_1, j_2) plane, then $(U(t)f_{mn}, g_{mn}) \to 0$ as $t \to \pm\infty$. If both $\partial\xi/\partial j_1$ and $\partial\xi/\partial f_2$ vanish identically in the (j_1, j_2) plane, then, with one exception, $\xi \equiv 0$, and $(U(t)f_{mn}, g_{mn}) \to (f_{mn}, g_{mn})$ as $t \to \pm\infty$. The single exception is the case of simple

[6] J. J. Levin and S. S. Shatz, J. Math. Anal. Appl. 7, 284 (1963).

harmonic oscillator, where $V(r) = kr^2$, and $\nu_1 = (2\pi/mk)^{\frac{1}{2}}$, $\nu_2 = 0$. In this case $(U(t)f_{mn}, g_{mn}) \to (f_{mn}, g_{mn})$ as $t \to \pm\infty$ if $m = 0$, but behaves like $\exp(2\pi i m\nu_1 t)$ if $m \neq 0$.

We have not yet exhausted all possibilities. It may happen that both $\partial\xi/\partial j_1$ and $\partial\xi/\partial j_2$ vanish on a set whose Lebesgue measure differs from zero, but do not vanish identically. In this case we must divide the (j_1, j_2) plane up into measurable subsets, such that on each subset the Jacobians $\partial\xi/\partial j_1$ either vanish only on a subset of measure zero or else vanish identically. The preceding analysis can then be applied separately to each subset. In the general case we cannot put $j_2 = 0$ in each subset, and the analysis becomes quite complicated.

We notice, however, that if we assume that the potential $V(r)$ is an *analytic* function of r, then it follows that $E(j_1, j_2)$ is an analytic function of both j_1 and j_2. In fact, the relation (1.27) defining j_1 in terms of E and j_2 can then be rewritten as a contour integral in the complex r plane:

$$j_1 = \oint_\Gamma \{2m[E - V(r)] - j_2^2/4\pi^2 r^2\}^{\frac{1}{2}}\, dr, \quad (2.22)$$

where the closed contour Γ surrounds the interval on the positive r axis on which the integrand is real (see Ref. 4, p. 302). If the potential is analytic, then the integrand is analytic on this contour, and the integral defines j_1 as an analytic function of E and j_2. It follows that E may be expressed as a (locally) analytic function of j_1 and j_2. (It is obvious from the physical interpretation of these variables that E can have no singularities in the quadrant $j_1, j_2 > 0$.)

If E is a (locally) analytic function of j_1 and j_2, then so is ξ, and hence so are $\partial\xi/\partial j_1$ and $\partial\xi/\partial j_2$. It follows that $\partial\xi/\partial j_1$ either vanishes on a set of measure zero, or else vanishes identically, and similarly for $\partial\xi/\partial j_2$. Hence, if the potential $V(r)$ is an analytic function of r, then our analysis is complete, and, except for the harmonic oscillator, either (2.14) or (2.18) holds. Combining (2.12), (2.14), and (2.18), we can now draw the following conclusion:

Theorem: If $V(r)$ is an analytic function of r with $V'(r) \geq 0$ and $V'''(r) \not\equiv 0$, $0 < r < \infty$, then

$$\lim_{t\to\pm\infty} (U(t)f, g) = (Pf, g), \quad (2.23)$$

where P is the projection on the subspace of $\mathcal{H}$ consisting of constants of the motion.

Finally, we note that if $\nu_2 \not\equiv 0$, then the only constants of the motion lie in the subspace $\mathcal{H}_{00}$, and so $Pf = f_{00}$. The components f_{mn} of f are determined from the formula

$$f_{mn}(\mathbf{j}, w_3) = \int_0^1\int_0^1 f(\mathbf{j}, \mathbf{w}) \exp[-2\pi i(mw_1 + nw_2)]\, dw_1\, dw_2. \quad (2.24)$$

In particular, we have

$$Pf = f_{00} = \int_0^1\int_0^1 f(\mathbf{j}, \mathbf{w})\, dw_1\, dw_2, \quad (2.25)$$

i.e., Pf is the average value of f taken over one period of each of the angle variables w_1 and w_2. If $\nu_2 \equiv 0$, however, then the constants of the motion all lie in the subspaces $\mathcal{H}_{0,n}$, and $Pf = \sum_n f_{0n}$. In this case we have

$$Pf = \sum_n f_{0n} = \int_0^1 f(\mathbf{j}, \mathbf{w})\, dw_1. \quad (2.26)$$

From (2.20) it follows easily that every absolutely continuous initial probability distribution ρ on the system approaches a steady-state distribution as $t \to \pm\infty$, in the sense that

$$\lim_{t\to\pm\infty} \int f\rho_t\, dV = \int f\bar{\rho}\, dV, \quad (2.27)$$

for each bounded measurable function f on the phase space. Here $\bar{\rho}$ is given by

$$\bar{\rho} = P_\rho. \quad (2.28)$$

For details and discussion we refer to Ref. 2.

In the special case of the Kepler problem, $V(r) = kr^{-2}$ and $E = -2\pi^2 mk^2/j_1^2$. Hence, $\nu_1 = 4\pi^2 mk^2/j_1^3$ and $\nu_2 \equiv 0$. It follows that the spectrum of the generator X consists of all real numbers, and the spectral measure is absolutely continuous with respect to Lebesgue measure, the weighting factor being a multiple of

$$\frac{\partial\nu_1}{\partial j_1} = \frac{12\pi^2 mk^2}{j_1^4}.$$

In particular, the spectrum of the generator X does not approximate the spectrum of the generator of the associated quantum mechanical problem, which can be cast in the same form, but with a weighting factor which is a multiple of

$$\sum_{n=0}^{\infty} \frac{2\pi^2 mk^2}{j_1^2}\, \delta(j_1 - nk).$$

For this reason, it appears unlikely that the classical generator X can be used to describe the motion of the quantum-mechanical problem.

3. AN APPLICATION

We shall now apply the results of the preceding section to obtain a statistical description of the behavior of a monatomic ideal gas enclosed in a spherical container (see Ref. 2).

We shall take as our model the system consisting of n point particles all lying inside a spherical region of E_3 centered at the origin. We take the spherical coordinates and their conjugate momenta of each particle as canonical coordinates, and for the Hamiltonian of the system we adopt the function

$$H = T + V, \tag{3.1}$$

with the kinetic energy T given by

$$T = \sum_{i=1}^{n} T_i, \tag{3.2}$$

where T_i is the kinetic energy (1.3) of the ith particle, and the potential energy V is given by

$$V = \sum_{i=1}^{n} V(r_i), \tag{3.3}$$

with r_i the radial position of the ith particle. Here $V(r)$ is a potential function describing the effects of the walls of the container. We shall take for $V(r)$ any function which satisfies the conditions of Sec. 2 and is small in the interior of the container but large near the walls. A possible example is

$$V(r) = c(a^2 - r^2)^{-1}, \tag{3.4}$$

with $0 < c \ll 1$.

It follows from (3.3) that the only forces acting on the particles are those due to the presence of the walls of the container. Because of the special form of (3.3), it follows that the Hamiltonian may be expressed as the sum of terms of the form (1.2), and hence that the motion of the phase space E_{6n} may be expressed as a product of the motions of the n single-particle phase spaces E_6, each of which has the form described in Sec. 2.

It follows that the motion induced on the function space $\mathcal{K} = \mathfrak{L}_2(E_{6n})$ may also be expressed as a product of the form

$$U(t) = \exp i\left(\sum_{i=1}^{n} X_i\right)t, \tag{3.5}$$

where each generator X_i has the form (2.11). The analysis of Sec. 2 leads again to the conclusion that

$$\lim_{t\to\pm\infty} (U(t)f, g) = (Pf, g). \tag{3.6}$$

It follows that if ρ is any probability distribution on E_{6n}, then

$$\lim_{t\to\pm\infty} \int g\rho_t \, dV = \int g\bar{\rho} \, dV, \tag{3.7}$$

where g is any bounded measurable function on E_{6n}, and $\bar{\rho} = P_\rho$.

The computation of $\bar{\rho}$ proceeds as follows: First we observe that the volume element in each single-particle phase space may be expressed in terms of the energy and angular momentum variables:

$$\begin{aligned} dV &= dw_1 \, dw_2 \, dw_3 \, dj_1 \, dj_2 \, dj_3 \\ &= dw_1 \, dw_2 \, dw_3 \frac{\partial j_1}{\partial E} dE \frac{\partial j_2}{\partial L} dL \frac{\partial j_3}{\partial L_3} dL_3 \\ &= d\mu \, dE \, dL \, dL_3, \end{aligned} \tag{3.8}$$

where

$$d\mu = (1/4\pi^2\nu_1) \, d\mathbf{w}. \tag{3.9}$$

Hence, in terms of these variables we have

$$\begin{aligned} &\bar{\rho}(E_1, \cdots, E_n, L_1, \cdots, L_n, L_{13}, \cdots, L_{n3}) \\ &\quad = \int \rho \prod_{i=1}^{n} d\mu_i = \int \rho \prod_{i=1}^{n} \frac{1}{4\pi^2\nu_{i1}} \, d\mathbf{w}_i. \end{aligned} \tag{3.10}$$

If the initial distribution depends on the momenta only through the total energy E, then the same is true of the steady-state distribution, and (3.10) reduces to

$$\bar{\rho}(E) = \int_{\Sigma(E)} \rho \frac{d\Sigma}{|\text{grad } E|}, \tag{3.11}$$

where $\Sigma(E)$ is the hypersurface of total energy E in the system phase space and $d\Sigma$ its area element. In this case, then, the time average of the initial distribution of the system is equal to its space average, even though the system is not ergodic (see Ref. 2).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 12 DECEMBER 1969

Feynman Path Integrals and Scattering Theory*†‡

DONALD GELMAN AND LARRY SPRUCH
Department of Physics, New York University, Washington Square, New York, New York

(Received 7 March 1969)

The Feynman path-integral formulation of nonrelativistic quantum mechanics is applied to scattering theory. A particularly appealing result of Feynman's is a variational upper bound (VUB) on the ground-state energy, with the bound expressed in terms of path integrals containing a trial functional $\tilde{S}$. When $\tilde{S}$ is of the standard form $-\int \{(KE) + \tilde{V}\}\, dt/\hbar$, with KE the kinetic energy and $\tilde{V}$ a trial potential, the bound reduces to the Rayleigh–Ritz result. Since $\tilde{S}$ need not be of that form, Feynman's result has interesting new possibilities. For central potentials, we extend Feynman's analysis to the subspace $\mathcal{K}^{(L)}$ of states of a given angular momentum L, obtaining a VUB on the energy of the lowest-energy bound state within $\mathcal{K}^{(L)}$. By exploiting the analogy between the lowest-energy bound state and the zero-energy scattering state in the no-bound state case, we obtain a VUB on A_L, the parameter which characterizes low-energy scattering in $\mathcal{K}^{(L)}$. The bound on A_L is expressed in terms of modified path integrals containing a trial functional $\tilde{S}_L$, and it reduces to the usual result if and only if $\tilde{S}_L$ is of the standard form $-\int \{(KE)_L + \tilde{V}\}\, dt/\hbar$, where $(KE)_L$ is the kinetic energy corresponding to motion with fixed angular momentum $\{L(L+1)\hbar^2\}^{\frac{1}{2}}$. The bound therefore contains new features. Finally, a similar but much simpler bound on A^L is derived by using the regularization $\mathrm{Tr}\,(e^{-\beta H_L} - e^{-\beta H_L{}^0})$ of the divergent function $\mathrm{Tr}\, e^{-\beta H_L}$, with Tr the trace, H_L^0 the (reduced) free-particle Hamiltonian in $\mathcal{K}^{(L)}$, and $H_L = H_L^0 + V$.

1. INTRODUCTION

The Feynman formulation of nonrelativistic quantum mechanics is based on a representation of the propagator K as a path integral,[1] which involves the evaluation of the classical action

$$A = \int \mathfrak{L}\, dt,$$

where $\mathfrak{L}$ is the Lagrangian. Using the notation $\int d(\text{paths})$ to symbolize integration over paths, we have for K the representation

$$K = \int d(\text{paths}) e^{iA/\hbar}.$$

Since the phase of the contribution to K from a given path is proportional to the value of the action on that path, contributions from neighboring paths will cancel as $\hbar$ approaches zero, except in the vicinity of that path—there may be more than one—for which the action is stationary. In this limit, then, the only important path is the one which satisfies Hamilton's principle $\delta A = 0$ and which is therefore the classical path.

Although path integrals provide an exceedingly elegant formulation of quantum mechanics—which is equivalent to the more conventional formulations, but which exhibits most clearly the reduction of quantum to classical mechanics in the limit of vanishing $\hbar$—and although Feynman's original investigations have stimulated much further work with path integrals,[2] their possible application to scattering theory appears to have received little attention. It has been suggested that in problems such as the excitation of hydrogen by a high-energy proton, where a classical treatment of the motion of the protons provides a good approximation, path integrals should be useful in obtaining higher-order corrections.[3] This will not be considered here, however.

For reasons to be noted shortly, we propose, instead, to study path integrals in connection with the development of variational bounds in scattering theory. Since useful bounds are normally much more readily obtainable on real than on complex entities, particularly if the real entities are nonnegative, it is natural to work with the positive-definite density matrix ρ rather than with the complex propagator K. A bound on ρ is an intermediate step which enables us to bound quantities (such as scattering lengths) which can assume negative as well as positive values. The functions K and ρ are intimately related, involving a change from real to imaginary time.[4] For a spinless

* The work reported in this article was supported by the U.S. Office of Naval Research under Contract N00014-67-A-0467-0007, and the U.S. Army Research Office under Contract DA-ARO-D-31-124-G989.

† A very preliminary version of this report appears in *Abstracts of the Proceedings of the Fifth International Conference on the Physics of Electronic and Atomic Collisions, Leningrad, July 1967*, I. P. Flaks, Ed. (Nauka, Leningrad, 1967), p. 31.

‡ Submitted by D. Gelman in partial fulfillment of the requirements for the degree of Doctor of Philosophy at New York University.

[1] R. P. Feynman, Rev. Mod. Phys. **20**, 367 (1948).

[2] See S. G. Brush, Rev. Mod. Phys. **33**, 79 (1961) for many references.

[3] P. Pechukas and J. C. Light, J. Chem. Phys. **44**, 3897 (1966).

[4] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill Book Co., New York, 1965), Chap. 10.

particle in a central potential $V(r)$, ρ is given by

$$\rho = \rho(\mathbf{r}'', \mathbf{r}'; \beta) = \langle \mathbf{r}'' | e^{-\beta H} | \mathbf{r}' \rangle, \tag{1.1}$$

where the Hamiltonian H is

$$H = H^0 + V(r), \quad H^0 \equiv (-\hbar^2/2m)\nabla^2. \tag{1.2}$$

(K is the coordinate representation of the time evolution operator $e^{-iTH/\hbar}$.) The superscript zero, when not in parentheses, is used throughout the paper to identify quantities associated with the free-particle system. The parameter β will normally be taken to be $T/\hbar$, with T a time interval, as this interpretation is more appropriate in connection with the path-integral representation of ρ than that involving the temperature.

For ρ Feynman obtained the path-integral representation[4]

$$\rho(\mathbf{r}'', \mathbf{r}'; T/\hbar) = \int_{x'}^{x''} D\mathbf{r}(t) e^{S}, \tag{1.3}$$

where $\int_{x'}^{x''} D\mathbf{r}(t)$ symbolizes integration over three-dimensional paths $\mathbf{r}(t)$ from the initial space–time point $x' \equiv (\mathbf{r}', t')$ to the final space–time point $x'' \equiv (\mathbf{r}'', t'')$, where S, the analog of the action in the path-integral representation of the propagator, is given by

$$S = -\frac{1}{\hbar}\int_{t'}^{t''} dt \left\{ \frac{m}{2}\left(\frac{d\mathbf{r}}{dt}\right)^2 + V(r) \right\} \tag{1.4}$$

and where

$$t'' - t' \equiv T = \beta\hbar. \tag{1.5}$$

By introducing a trial functional $\tilde{S}$, Feynman was able to obtain a variational lower bound (VLB) $\underset{\cdot}{\rho}^F$ on ρ, which he then used to develop a very interesting variational upper bound (VUB) on the ground-state energy.[5] (The dot beneath a symbol always denotes a VLB on that symbol.) Thus, assuming that at least one bound state exists, it follows from the bilinear eigenfunction expansion of ρ that ρ is dominated by the ground-state contribution when β is very large, i.e.,

$$\rho(\mathbf{r}'', \mathbf{r}'; \beta) \to \psi_0(\mathbf{r}'')\psi_0^*(\mathbf{r}')e^{-\beta E_0}, \quad \beta \to \infty,$$

where ψ_0 and E_0 (< 0) denote, respectively, the wavefunction and energy of the ground state. Expressing E_0 in terms of ρ, we have

$$E_0 = -\lim_{\beta\to\infty} \beta^{-1} \ln \rho \equiv E_0(\rho). \tag{1.6}$$

Since E_0 is evidently a monotonic decreasing function of ρ, we have

$$\rho \geq \underset{\cdot}{\rho}^F \Rightarrow E_0(\rho) \leq E_0(\underset{\cdot}{\rho}^F),$$

i.e., $E_0(\underset{\cdot}{\rho}^F)$ is obviously a VUB on E_0. [Clearly, if any quantity Q is a monotonic decreasing (increasing) function of ρ, then $Q(\underset{\cdot}{\rho}^F)$ is a VUB (VLB) on Q.] When the trial functional $\tilde{S}$ is of the (standard) form

$$\tilde{S} = -\frac{1}{\hbar}\int_{t'}^{t''} dt \left\{ \frac{m}{2}\left(\frac{d\mathbf{r}}{dt}\right)^2 + \tilde{V}(r) \right\}, \tag{1.7}$$

with $\tilde{V}(r)$ a trial potential having at least one bound state, $E_0(\underset{\cdot}{\rho}^F)$ generates the usual Rayleigh–Ritz VUB, i.e., we obtain

$$E_0 \leq \int d\mathbf{r} \tilde{\psi}_0^*(\mathbf{r}) H \tilde{\psi}_0(\mathbf{r}),$$

where $\tilde{\psi}_0$ is the ground-state wavefunction of the trial Hamiltonian $\tilde{H} = H^0 + \tilde{V}(r)$. However, $\tilde{S}$ need not be of the standard form, and Feynman's result, therefore, contains some interesting new possibilities. That this additional freedom is of more than just academic interest can be inferred from the fact that the results achieved in the polaron problem,[5] with $\tilde{S}$ not of the standard form, were for a time superior to those obtained by more conventional methods.

If the particle cannot be bound to the potential, the lowest available energy is at the bottom of the continuous spectrum, at zero energy. The zero-energy scattering state for the no-bound state case therefore plays a role very similar to that of the lowest-energy bound state when at least one bound state exists. This analogy has previously been exploited within the context of the conventional formulation of quantum mechanics to obtain a VUB on the scattering length A_0.[6] It can also be exploited within the context of the Feynman formulation of quantum mechanics to obtain a VUB on A_0 in terms of path integrals containing a trial functional $\tilde{S}$. Since a VUB on A_0 is the scattering theory analog of the Rayleigh–Ritz result, the development of this bound is a natural starting point for the present investigation.

In the remainder of this section, we sketch the contents of this article. In the absence of bound states, the bilinear form of ρ in the limit of large β is dominated by the contribution from the continuum states in the neighborhood of zero energy. By examining the form of this contribution at large distances, we obtain A_0 as a monotonic decreasing function of ρ. Feynman's VLB on ρ then generates a VUB on A_0. When the trial functional $\tilde{S}$ is of the standard form (1.7), this VUB on A_0 reduces to the usual one; since

[5] R. P. Feynman, Phys. Rev. 97, 660 (1955).

[6] L. Spruch and L. Rosenberg, Phys. Rev. 116, 1034 (1959).

$\tilde{S}$ is not required to be of the standard form, the bound has new features. (These new features are presumably of primarily academic interest when dealing with simple scattering, since the effort required to evaluate the necessary path integrals would undoubtedly be more profitably spent improving the trial function in the usual bound. The path-integral approach may nevertheless provide some new insights into scattering, and, as has already been noted, should be significant in the treatment of some compound scattering problems.)

The density matrix ρ is defined in the full space. The analog of ρ within the subspace $\mathcal{H}^{(L)}$ of states of angular momentum L is the function

$$\rho_L(r'', r'; \beta) = \langle r''| e^{-\beta H_L} |r'\rangle = 2\pi r'' r' \int_{-1}^{1} d\xi P_L(\xi) \rho(\mathbf{r}'', \mathbf{r}'; \beta), \quad (1.8)$$

where

$$\xi \equiv \hat{r}'' \cdot \hat{r}', \quad (1.9)$$

$$H_L = H_L^0 + V(r), \quad (1.10)$$

$$H_L^0 \equiv -\frac{\hbar^2}{2m}\left\{\frac{d^2}{dr^2} - L(L+1)r^{-2}\right\}, \quad (1.11)$$

and where the second expression for ρ_L follows from the angular-momentum decomposition

$$\rho(\mathbf{r}'', \mathbf{r}'; \beta) = \sum_{L=0}^{\infty} \frac{2L+1}{4\pi r'' r'} P_L(\xi) \rho_L(r'', r'; \beta). \quad (1.12)$$

For ρ_L we obtain the (modified) path-integral representation

$$\rho_L(r'', r'; T/\hbar) = \int_{y'}^{y''} dr(t) e^{S_L}, \quad (1.13)$$

where $\int_{y'}^{y''} dr(t)$ symbolizes integration over one-dimensional positive-valued paths $r(t)$ from $y' \equiv (r', t')$ to $y'' \equiv (r'', t'')$ and where S_L, as would be expected, is given by

$$S_L = -\frac{1}{\hbar}\int_{t'}^{t''} dt\left\{\frac{m}{2}\left(\frac{dr}{dt}\right)^2 + \frac{L(L+1)\hbar^2}{2mr^2} + V(r)\right\}. \quad (1.14)$$

This representation has previously been given, for the case $V = 0$, by Edwards and Gulyaev.[7] Introducing a trial functional $\tilde{S}_L$ and essentially duplicating Feynman's derivation of ρ^F, we develop the VLB ρ_L^F on ρ_L. If at least one bound state of angular momentum L exists, then, by arguments identical to those leading to Eq. (1.6), we find that

$$E_0^{(L)} = -\lim_{\beta\to\infty} \beta^{-1} \ln \rho_L = E_0(\rho_L),$$

where $E_0^{(L)}$ (< 0) is the energy of the lowest-energy bound state of angular momentum L [and $E_0^{(0)} \equiv E_0$]. It immediately follows that $E_0(\rho_L^F)$ is a VUB on $E_0^{(L)}$. This represents a slight extension of Feynman's result.[5]

By exploiting within $\mathcal{H}^{(L)}$ the analogy between the lowest-energy bound state when at least one bound state exists and the zero-energy scattering state in the no-bound state case, we obtain a VUB on A_L in terms of (modified) path integrals containing a trial functional $\tilde{S}_L$. [The parameter A_L characterizes very-low-energy scattering in $\mathcal{H}^{(L)}$, the threshold-energy dependence of the Lth-partial cross section σ_L being given by

$$\sigma_L \to 4\pi(2L+1)A_L^2 k^{4L}, \quad k \to 0,$$

where k is the wavenumber.] When $\tilde{S}_L$ is of the (standard) form

$$\tilde{S}_L = -\frac{1}{\hbar}\int_{t'}^{t''} dt\left\{\frac{m}{2}\left(\frac{dr}{dt}\right)^2 + \frac{L(L+1)\hbar^2}{2mr^2} + \tilde{V}(r)\right\}, \quad (1.15)$$

with $\tilde{V}(r)$ a trial potential which cannot support a bound state of angular momentum L, the VUB on A_L reduces to the usual result[6]; since $\tilde{S}_L$ need not be of the standard form, the bound has new features. It should be possible to relax the requirement that the trial potential be unable to support a bound state within $\mathcal{H}^{(L)}$ because the usual result does not impose this requirement, but this will not be considered here.

The partition function

$$\mathrm{Tr}\, e^{-\beta H} \equiv \int d\mathbf{r} \rho(\mathbf{r}, \mathbf{r}; \beta)$$

for a system in which all states are bound states is given by

$$\mathrm{Tr}\, e^{-\beta H} = \sum_n e^{-\beta E_n},$$

where Tr denotes the trace and where the sum extends over all states. Using an approach based on the partition function rather than on ρ, Feynman obtained a slightly different VUB on the ground-state energy.[8] For scattering systems, however, the partition function is infinite, since the scattering wavefunctions are unnormalizable. For the same reason, this is also true for the quantity

$$\mathrm{Tr}\, e^{-\beta H_L} \equiv \int_0^{\infty} dr \rho_L(r, r; \beta), \quad (1.16)$$

[7] S. F. Edwards and Y. V. Gulyaev, Proc. Roy. Soc. (London) A279, 229 (1964).

[8] Ref. 4, Chap. 11, p. 303.

the analog of the partition function within $\mathcal{K}^{(L)}$. By first subtracting from ρ_L its free-particle counterpart ρ_L^0 and then integrating, however, one obtains the finite quantity

$$z_L(\beta) \equiv \mathrm{Tr}\,(e^{-\beta H_L} - e^{-\beta H_L^0}). \tag{1.17}$$

With an approach based on z_L rather than on ρ_L, we obtain a slightly different, but very much simpler, VUB on A_L.

The analysis is confined to the establishment of formal variational (upper) bounds on A_L for the system consisting of a nonrelativistic, spinless particle of mass m moving in a short-range, central potential $V(r)$ which cannot support a bound state of angular momentum L. There is not the slightest pretense of mathematical rigor. The procedure is to develop VLB's on ρ, ρ_L, and z_L, thereby reducing the problem to that of expressing A_L as a monotonic (decreasing) function of these. We begin with this latter aspect, developing in Sec. 2 appropriate expressions for A_0 in terms of ρ and for A_L in terms of ρ_L and z_L. In Sec. 3, after briefly reviewing the salient features of both the path-integral representation of ρ and the procedure whereby Feynman bounded ρ, we derive the analogous representation of, and bound on, ρ_L. We also develop other VLB's $\rho^{(n)}$ on ρ, $n = 0, 1, \cdots$, and their analogs in $\mathcal{K}^{(L)}$, the VLB's $\rho_L^{(n)}$ on ρ_L. All of the VLB's on ρ_L are then used to generate corresponding VLB's on z_L. The demonstration that the bounds on A_L reduce to the usual results when the trial functional is of the standard form is presented in Sec. 4.

2. SCATTERING LENGTHS IN TERMS OF DENSITY MATRICES

A. A_0 in Terms of ρ

The density matrix is the solution of the Bloch equation,

$$\left(\frac{\partial}{\partial\beta} + H''\right)\rho(\mathbf{r}'', \mathbf{r}'; \beta) = 0, \tag{2.1}$$

satisfying the boundary condition

$$\rho(\mathbf{r}'', \mathbf{r}'; 0) = \delta^{(3)}(\mathbf{r}'' - \mathbf{r}'), \tag{2.2}$$

where H'' is given by (1.2),with $\mathbf{r}$ replaced by $\mathbf{r}''$, and where the right-hand side (rhs) of Eq. (2.2) is a three-dimensional delta function. Since we are assuming that there are no bound states, the eigenfunction expansion of ρ is

$$\rho(\mathbf{r}'', \mathbf{r}'; \beta) = \int d\mathbf{k} P(\mathbf{r}'', \mathbf{r}'; \mathbf{k}) e^{-\beta E_k}, \tag{2.3}$$

where $E_k = (\hbar k)^2/2m$ and where

$$P(\mathbf{r}'', \mathbf{r}'; \mathbf{k}) \equiv (2\pi)^{-3}\psi(\mathbf{k}, \mathbf{r}'')\psi^*(\mathbf{k}, \mathbf{r}'). \tag{2.4}$$

The function $\psi(\mathbf{k}, \mathbf{r})$ is the solution of the time-independent Schrödinger equation

$$(E_k - H)\psi(\mathbf{k}, \mathbf{r}) = 0, \tag{2.5}$$

with the asymptotic form

$$\psi(\mathbf{k}, \mathbf{r}) \to e^{i\mathbf{k}\cdot\mathbf{r}} + r^{-1}e^{ikr}f_{\mathbf{k}}(\hat{r}), \quad r \to \infty. \tag{2.6}$$

The zero-energy limit is finite, nonvanishing, and has the asymptotic form

$$\psi(0, \mathbf{r}) \to 1 - A_0/r, \quad r \to \infty. \tag{2.7}$$

To obtain a bound on A_0, we must express it as a monotonic function of ρ. To do so, we note that only states with energies less than or of the order of β^{-1} contribute appreciably to the integral in Eq. (2.3) because of the rapidly decaying exponential. In the limit of large β, the whole contribution comes from the neighborhood of zero energy. We can then extract the smoothly varying wavefunctions at $\mathbf{k} = 0$ and integrate the exponential over $\mathbf{k}$ to obtain

$$\rho(\mathbf{r}'', \mathbf{r}'; \beta) \to (2\pi m/\beta\hbar^2)^{\frac{3}{2}}P(\mathbf{r}'', \mathbf{r}'; 0), \quad \beta \to \infty. \tag{2.8}$$

It follows from (2.4), (2.7), and (2.8) that A_0 is given by

$$A_0 = \lim_{r'', r' \to \infty} \{r''r'/(r'' + r')\} \times \left\{1 - \lim_{\beta\to\infty}(2\pi\hbar^2\beta m^{-1})^{\frac{3}{2}}\rho(\mathbf{r}'', \mathbf{r}'; \beta)\right\}. \tag{2.9}$$

The additional freedom associated with the possibility of having $\mathbf{r}''$ and $\mathbf{r}'$ approach infinity independently may prove helpful in the analysis of more difficult problems. (One such problem is the determination of a VUB on A_0 when the potential can support a bound state. This VUB has been obtained using the Schrödinger formalism,[9] but we have been unable to reproduce it using the path-integral formalism.) We will not utilize this additional freedom here, so we set $\mathbf{r}'' = \mathbf{r}'$ and obtain

$$A_0 = \lim_{r' \to \infty} \tfrac{1}{2}r'\left\{1 - \lim_{\beta\to\infty}(2\pi\hbar^2\beta m^{-1})^{\frac{3}{2}}\rho(\mathbf{r}', \mathbf{r}'; \beta)\right\} \equiv A_0(\rho). \tag{2.10}$$

Since A_0 is a monotonic decreasing function of ρ, $A_0(\rho)$ is a VUB on A_0, where ρ denotes any VLB on ρ. [We note that this bound on A_0 applies also when $V = V(\mathbf{r})$, since the derivation of the bound did not utilize the assumption of spherical symmetry.]

B. A_L in Terms of ρ_L

The $L = 0$ case is somewhat unique in the analysis of scattering lengths, since only the $L = 0$ component

[9] L. Rosenberg, L. Spruch, and T. F. O'Malley, Phys. Rev. **118**, 184 (1960).

of $\psi(\mathbf{k}, \mathbf{r})$ survives as $\mathbf{k}$ approaches zero. To obtain bounds on A_L for any L, including $L = 0$, we proceed as follows.

The function ρ_L is the solution of the equation

$$\left(\frac{\partial}{\partial\beta} + H_L\right)\rho_L(r, r'; \beta) = 0, \tag{2.11}$$

satisfying the boundary condition

$$\rho_L(r'', r'; 0) = \delta^{(1)}(r'' - r'), \tag{2.12}$$

where the (reduced) Hamiltonian H_L is given by Eqs. (1.10) and (1.11). In view of our assumptions concerning the potential, the bilinear form of ρ_L is

$$\rho_L(r'', r'; \beta) = \frac{2}{\pi}\int_0^\infty dk u_L(k, r'')u_L(k, r')e^{-\beta E_k}, \tag{2.13}$$

where $u_L(k, r)$ is the (real) solution of the radial equation

$$(E_k - H_L)u_L(k, r) = 0, \tag{2.14}$$

satisfying the boundary conditions

$$u_L(k, 0) = 0, \tag{2.15}$$

$$u_L(k, r) \to kr \cos\eta_L\{j_L(kr) - \tan\eta_L n_L(kr)\}, \quad r \to \infty; \tag{2.16}$$

$\eta_L = \eta_L(k)$ is the phase shift, and $j_L(n_L)$ denotes the usual spherical Bessel (Neumann) function. We note that the outgoing scattering solution $\psi(\mathbf{k}, \mathbf{r})$ has the partial-wave expansion

$$\psi(\mathbf{k}, \mathbf{r}) = \sum_{L=0}^{\infty} i^L(2L + 1)P_L(\hat{k} \cdot \hat{r})\frac{u_L(k, r)}{kr}e^{i\eta_L}, \tag{2.17}$$

and that the threshold-energy dependence of the Lth-partial wave is

$$u_L(k, r) \to k^{L+1}u_L(r), \quad k \to 0. \tag{2.18}$$

The function $u_L(r)$ is the (real) solution of the zero-energy radial equation

$$H_L u_L(r) = 0, \tag{2.19}$$

satisfying the boundary conditions

$$u_L(0) = 0, \tag{2.20}$$

$$u_L(r) \to \{r^{L+1}/(2L + 1)!!\} - A_L(2L - 1)!!\, r^{-L}, \quad r \to \infty, \tag{2.21}$$

where

$$(-1)!! = 1, \quad (2L + 1)!! = (2L + 1)(2L - 1)!!, \quad L = 0, 1, \cdots. \tag{2.22}$$

In the limit of large β, the whole contribution to the integral in Eq. (2.13) comes from the neighborhood of $k = 0$, and we can therefore replace $u_L(k, r'') \times u_L(k, r')$ by its threshold form $k^{2L+2}u_L(r'')u_L(r')$ and then integrate over k to obtain

$$\rho_L(r'', r'; \beta) \to a(L)\beta^{-(L+\frac{3}{2})}u_L(r'')u_L(r'), \quad \beta \to \infty, \tag{2.23}$$

where

$$a(L) = \pi^{-\frac{1}{2}}2^{-(L+1)}(2L + 1)!!\,(2m/\hbar^2)^{L+\frac{3}{2}}. \tag{2.24}$$

Using (2.21) and (2.23) and setting $r'' = r'$, the resultant loss of generality being irrelevant for our present purposes, we find that

$$A_L = \lim_{r' \to \infty}\frac{2L + 1}{2r'} \times \left\{\frac{(r')^{2L+2}}{\{(2L + 1)!!\}^2} - \lim_{\beta \to \infty}\frac{\beta^{L+\frac{3}{2}}}{a(L)}\rho_L(r', r'; \beta)\right\} \equiv A_L'(\rho_L). \tag{2.25}$$

Since A_L' is a monotonic decreasing function of ρ_L, it follows that $A_L'(\rho_L)$ is a VUB on A_L, where ρ_L is any VLB on ρ_L.

C. A_L in Terms of z_L

We consider now a different approach which again leads to a number of variational upper bounds on A_L for any value of L. As shown in Appendix A, the quantity $z_L(\beta)$ defined in Eq. (1.17) is given by

$$z_L(\beta) = \frac{1}{\pi}\int_0^\infty dk e^{-\beta E_k}\frac{d\eta_L(k)}{dk}. \tag{2.26}$$

Since only the neighborhood of $k = 0$ contributes to this integral in the limit of large β, it follows from the threshold-energy dependence

$$\eta_L(k) \to -A_L k^{2L+1} \pmod{\pi}, \quad k \to 0, \tag{2.27}$$

that

$$z_L(\beta) \to -(\hbar^2/2m)A_L a(L)\beta^{-(L+\frac{1}{2})}, \quad \beta \to \infty, \tag{2.28}$$

with $a(L)$ given by Eq. (2.24), so that

$$A_L = (-2m/\hbar^2)\lim_{\beta \to \infty}\{\beta^{L+\frac{1}{2}}z_L(\beta)/a(L)\} \equiv \bar{A}_L(z_L). \tag{2.29}$$

Since $\bar{A}_L$ is a monotonic decreasing function of z_L, it follows that $\bar{A}_L(z_L)$ is a VUB on A_L, where z_L is any VLB on z_L.

3. DENSITY MATRICES: PATH-INTEGRAL REPRESENTATIONS AND VARIATIONAL BOUNDS

A. Path-Integral Representation of ρ

The path-integral representation of the density matrix can be derived[10] from the relation

$$\int d\mathbf{r}\rho^V(\mathbf{r}'', \mathbf{r}; \epsilon/\hbar)\rho(\mathbf{r}, \mathbf{r}'; t/\hbar) = \rho(\mathbf{r}'', \mathbf{r}'; (t + \epsilon)/\hbar) + O(\epsilon^2), \tag{3.1}$$

[10] The derivation presented here is essentially equivalent to that given by R. Abé, Busseiron Kenkyu 79, 101 (1954). This work is also described in Sec. 8 of Ref. 2.

where,[11] for any coordinates and any time,

$$\rho^V(\mathbf{r}'', \mathbf{r}'; T/\hbar) \equiv \rho^0(\mathbf{r}'', \mathbf{r}'; T/\hbar) \exp\{-TV(r'')/\hbar\}, \tag{3.2}$$

$$\rho^0(\mathbf{r}'', \mathbf{r}'; T/\hbar) = (m/2\pi\hbar T)^{\frac{3}{2}} \exp\{-m(\mathbf{r}'' - \mathbf{r}')^2/2\hbar T\}. \tag{3.3}$$

[Equation (3.1) is proved in Appendix B.] Iterating Eq. (3.1) $N-1$ times, we obtain

$$\int d\mathbf{r}_{N-1} \cdots \int d\mathbf{r}_1 \int d\mathbf{r} \rho^V(\mathbf{r}'', \mathbf{r}_{N-1}; \epsilon/\hbar) \cdots \rho^V(\mathbf{r}_1, \mathbf{r}; \epsilon/\hbar) \times \rho(\mathbf{r}, \mathbf{r}'; t/\hbar) = \rho(\mathbf{r}'', \mathbf{r}'; (t + N\epsilon)/\hbar) + O(N\epsilon^2), \tag{3.4}$$

from which, by setting $\epsilon = T/N$, taking the limit as $N \to \infty$, setting $t = 0$, and using Eq. (2.2), it follows that

$$\rho(\mathbf{r}'', \mathbf{r}'; T/\hbar) = \lim_{N\to\infty} \int d\mathbf{r}_1 \cdots \int d\mathbf{r}_{N-1} \prod_{n=1}^{N} \rho^V(\mathbf{r}_n, \mathbf{r}_{n-1}; T/N\hbar) = \lim_{N\to\infty} C(N) \int d\mathbf{r}_1 \cdots \int d\mathbf{r}_{N-1} \exp\{S^{(N)}\}, \tag{3.5}$$

where

$$\mathbf{r}_0 \equiv \mathbf{r}', \quad \mathbf{r}_N \equiv \mathbf{r}'', \tag{3.6}$$

$$C(N) \equiv (m/2\pi\hbar\epsilon)^{3N/2}, \tag{3.7}$$

$$S^{(N)} \equiv -\sum_{n=1}^{N} \left\{ \frac{m(\mathbf{r}_n - \mathbf{r}_{n-1})^2}{2\hbar\epsilon} + \frac{\epsilon V(r_n)}{\hbar} \right\}. \tag{3.8}$$

Since the left-hand side (lhs) of Eq. (3.1) is given by

$$\left(\frac{m}{2\pi\hbar\epsilon}\right)^{\frac{3}{2}} \exp\left(-\frac{\epsilon V(r'')}{\hbar}\right) \times \int d\mathbf{r} \exp\left(-\frac{m(\mathbf{r}'' - \mathbf{r})^2}{2\hbar\epsilon}\right) \rho(\mathbf{r}, \mathbf{r}'; t/\hbar), \tag{3.9}$$

appreciable contributions to the integral arise only for $m(\mathbf{r}'' - \mathbf{r})^2/2\hbar\epsilon$ less than, or of the order of, unity. Similarly, the rhs of Eq. (3.5) is dominated from the regions defined by

$$|\mathbf{r}_n - \mathbf{r}_{n-1}| \lesssim O(\epsilon^{\frac{1}{2}}) = O(N^{-\frac{1}{2}}), \quad n = 1, 2, \cdots, N. \tag{3.10}$$

We consider now the path integral on the rhs of Eq. (1.3). In connection with the remarks following that equation, we note that a path from x' to x'' is a kinematically possible trajectory for a particle going from $\mathbf{r}'$ at t' to $\mathbf{r}''$ at the later time t'', and is represented mathematically by a single-valued continuous function $\mathbf{r}(t)$ satisfying the boundary conditions

$$\mathbf{r}(t') = \mathbf{r}', \quad \mathbf{r}(t'') = \mathbf{r}''. \tag{3.11}$$

Unless specifically stated otherwise, it will be understood that all paths go from x' to x''.

The path integral is evaluated with the help of the particular path

$$\mathbf{r}^{(N)}(\mathbf{r}_1, \cdots, \mathbf{r}_{N-1}; t) \equiv \mathbf{r}^{(N)}(\mathbf{R}; t) = \{(t - t_{n-1})\mathbf{r}_n + (t_n - t)\mathbf{r}_{n-1}\}/\epsilon, \quad t_{n-1} \le t \le t_n, \quad n = 1, 2, \cdots, N, \quad N = 2, 3, \cdots, \tag{3.12}$$

where

$$t_j \equiv t' + j\epsilon, \quad \epsilon = T/N, \quad j = 0, 1, \cdots, N, \tag{3.13}$$

and where $\mathbf{R}$ denotes the ordered set of arbitrarily chosen vectors $\mathbf{r}_1, \cdots, \mathbf{r}_{N-1}$. Intuitively, it is clear that

$$M^{(N)} \equiv \int d\mathbf{r}_1 \cdots \int d\mathbf{r}_{N-1} \exp\{S[\mathbf{r}^{(N)}]\}, \tag{3.14}$$

where $S[\mathbf{r}^{(N)}]$ denotes the value of the functional S on the path $\mathbf{r}^{(N)}(\mathbf{R}; t)$, is proportional to the integral of e^S over all paths expressible in the form of Eq. (3.12). (The square bracket is used exclusively for the purpose of denoting a functional throughout this article.) Substituting Eq. (3.12) into Eq. (1.4), we obtain

$$S[\mathbf{r}^{(N)}] = S^{(N)} + \frac{\epsilon}{\hbar} \sum_{n=1}^{N} \left\{ V(r_n) - \int_0^1 ds V(q_n(s)) \right\}, \tag{3.15}$$

where the integration variable has been changed from t to $s = (t - t_{n-1})/\epsilon$ in the term involving the potential, and where

$$q_n(s) \equiv |\mathbf{r}_n - (1 - s)(\mathbf{r}_n - \mathbf{r}_{n-1})|. \tag{3.16}$$

Expanding $q_n(s)$ in a Taylor series in $|\mathbf{r}_n - \mathbf{r}_{n-1}|$, we find that

$$q_n(s) = r_n + O(|\mathbf{r}_n - \mathbf{r}_{n-1}|), \tag{3.17}$$

from which it follows that

$$V(q_n(s)) = V(r_n) + O(|\mathbf{r}_n - \mathbf{r}_{n-1}|). \tag{3.18}$$

Substituting into Eq. (3.15), and using (3.10) and $\epsilon = T/N$, we obtain

$$S[\mathbf{r}^{(N)}] - S^{(N)} = \epsilon \sum_{n=1}^{N} O(|\mathbf{r}_n - \mathbf{r}_{n-1}|) = \epsilon N O(\epsilon^{\frac{1}{2}}) = O(N^{-\frac{1}{2}}). \tag{3.19}$$

[11] The functions $\rho^V(\mathbf{r}'', \mathbf{r}'; \epsilon/\hbar)$ and $\rho(\mathbf{r}'', \mathbf{r}'; \epsilon/\hbar)$ are clearly positive definite; they differ by terms of order ϵ^2, as may be verified by setting $t = 0$ in Eq. (3.1) and using Eq. (2.2).

Equation (3.5) can therefore be rewritten as

$$\rho(\mathbf{r}'', \mathbf{r}'; T/\hbar) = \lim_{N\to\infty} C(N)M^{(N)}. \tag{3.20}$$

The rhs of this equation represents the integral of e^S over all paths available to a particle moving from x' to x'' with Brownian motion[12]; $C(N)$ plays the role of a normalization constant, ensuring the existence of and nonzero value for the limit. Although the Brownian motion paths do not include every conceivable path, enough are included to yield the equality[12]

$$\int_{x'}^{x''} D\mathbf{r}(t)e^S = \lim_{N\to\infty} C(N)M^{(N)}, \tag{3.21}$$

which, together with Eq. (3.20), establishes the validity of Eq. (1.3).

B. Path-Integral Representation of ρ_L

The procedure for developing a path-integral representation of ρ_L is analogous to that used in establishing Eq. (1.3). The relation analogous to Eq. (3.1) is

$$\int_0^\infty dr\, \bar{\rho}_L^V(r'', r; \epsilon/\hbar)\rho_L(r, r'; t/\hbar) = \rho_L(r'', r'; (t+\epsilon)/\hbar) + O(\epsilon^2), \tag{3.22}$$

where[13]

$$\bar{\rho}_L^V(r'', r'; T/\hbar) \equiv \bar{\rho}_L^0(r'', r'; T/\hbar)e^{-TV(r'')/\hbar}, \tag{3.23}$$

$$\bar{\rho}_L^0(r'', r'; T/\hbar) \equiv \left(\frac{m}{2\pi\hbar T}\right)^{\frac{1}{2}} \exp\left\{-\frac{m(r''-r')^2}{2\hbar T} - \frac{L(L+1)\hbar T}{2mr''^2}\right\}. \tag{3.24}$$

[See Appendix C for the proof of Eq. (3.22).] The function $\bar{\rho}_L^0$ is closely related to the free-particle function

$$\rho_L^0(r'', r'; T/\hbar) = 2\pi r''r' \int_{-1}^{1} d\xi P_L(\xi)\rho^0(\mathbf{r}'', \mathbf{r}'; T/\hbar), \tag{3.25}$$

this equation being the free-particle counterpart of Eq. (1.8). To obtain the relationship between $\bar{\rho}_L^0$ and ρ_L^0, we note from Eq. (3.3) that the angular dependence of ρ^0 is isolated in the factor $\exp(mr''r'\xi/\hbar T)$. Using Bauer's formula

$$\exp(a\xi) = \sum_{L=0}^{\infty}(2L+1)P_L(\xi)(-i)^L j_L(ia) \tag{3.26}$$

to expand this factor, we obtain for ρ^0 an angular-momentum decomposition having the form of Eq. (1.12), with ρ_L replaced by

$$\rho_L^0\left(r'', r'; \frac{T}{\hbar}\right) = \left(\frac{m}{2\pi\hbar T}\right)^{\frac{1}{2}} \exp\left(-\frac{m(r''-r')^2}{2\hbar T}\right)\zeta_L\left(\frac{mr''r'}{\hbar T}\right), \tag{3.27}$$

where

$$\zeta_L(z) = 2z(-i)^L j_L(iz)e^{-z}. \tag{3.28}$$

From the relation

$$j_L(iz) = \tfrac{1}{2}z^{-1}i^L e^z\{1 - \tfrac{1}{2}L(L+1)z^{-1} + O(z^{-2})\} + O(z^{-1}e^{-z}), \tag{3.29}$$

it follows that

$$\zeta_L(z) = 1 - \tfrac{1}{2}L(L+1)z^{-1} + O(z^{-2}) + O(e^{-2z}), \tag{3.30}$$

whence

$$\bar{\rho}_L^0(r'', r'; \epsilon/\hbar)/\rho_L^0(r'', r'; \epsilon/\hbar) = 1 + \{\epsilon L(L+1)\hbar(r''-r')/2mr''^2r'\} + O(\epsilon^2). \tag{3.31}$$

Iterating Eq. (3.22) $N-1$ times, we obtain

$$\int_0^\infty dr_{N-1}\cdots\int_0^\infty dr_1\int_0^\infty dr\, \bar{\rho}_L^V(r'', r_{N-1}; \epsilon/\hbar)\cdots \times \bar{\rho}_L^V(r_1, r; \epsilon/\hbar)\rho_L(r, r'; t/\hbar) = \rho_L(r'', r'; (t+N\epsilon)/\hbar) + O(N\epsilon^2). \tag{3.32}$$

Setting $\epsilon = T/N$ and taking the limit as $N\to\infty$, and also setting $t = 0$ and using Eq. (2.12), we obtain the analog of Eq. (3.5):

$$\rho_L(r'', r'; T/\hbar) = \lim_{N\to\infty}\int_0^\infty dr_1\cdots\int_0^\infty dr_{N-1}\prod_{n=1}^{N}\bar{\rho}_L^V(r_n, r_{n-1}; T/N\hbar) = \lim_{N\to\infty}\int_0^\infty dr_1\cdots\int_0^\infty dr_{N-1}C(N)^{\frac{1}{3}}\exp\{S_L^{(N)}\}, \tag{3.33}$$

where $r_0 = r'$, $r_N = r''$ and where

$$S_L^{(N)} \equiv -\sum_{n=1}^{N}\left\{\frac{m(r_n - r_{n-1})^2}{2\hbar\epsilon} + \frac{\epsilon L(L+1)\hbar}{2mr_n^2} + \frac{\epsilon V(r_n)}{\hbar}\right\}. \tag{3.34}$$

Storer[14] has obtained a representation of ρ_L having the form of Eq. (3.33), with $\bar{\rho}_L^V$ replaced by

$$\rho_L^0(r_n, r_{n-1}; \epsilon/\hbar)\exp(-\epsilon\{V(r_n) + V(r_{n-1})\}/2\hbar).$$

[12] Ref. 2, p. 80.

[13] The positive-definite functions $\bar{\rho}_L^V(r'', r'; \epsilon/\hbar)$ and $\rho_L(r'', r'; \epsilon/\hbar)$ differ by $O(\epsilon^2)$, as can be seen by setting $t = 0$ in Eq. (3.22) and using Eq. (2.12).

[14] R. G. Storer, J. Math. Phys. 9, 964 (1968).

With the lhs of Eq. (3.22) written in the explicit form

$$(m/2\pi\hbar\epsilon)^{\frac{1}{2}} \exp\{-\epsilon L(L+1)\hbar/(2mr''^2)\} \times \exp\{-\epsilon V(r'')/\hbar\} \times \int_0^\infty dr \exp\{-m(r''-r)^2/2\hbar\epsilon\}\rho_L(r, r'; t/\hbar), \quad (3.35)$$

it is seen that significant contributions to the integral arise only when $m(r''-r)^2/2\hbar\epsilon$ is less than, or of the order of, unity. Similarly, the (multiple) integral in Eq. (3.33) is dominated by the contributions from the regions defined by

$$|r_n - r_{n-1}| \lesssim O(\epsilon^{\frac{1}{2}}), \quad n = 1, 2, \cdots, N. \quad (3.36)$$

In analogy to Eq. (3.12), we now introduce the path

$$r^{(N)}(r_1, \cdots, r_{N-1}; t) \equiv r^{(N)}(R; t) = \{(t - t_{n-1})r_n + (t_n - t)r_{n-1}\}/\epsilon, \quad t_{n-1} \le t \le t_n, \quad n = 1, 2, \cdots, N, \quad N = 2, 3, \cdots, \quad (3.37)$$

where the times t_j are defined in Eq. (3.13), and where R denotes the ordered set of arbitrarily chosen non-negative numbers $r_1, \cdots, r_{N-1}$. Unless specifically stated otherwise, it will be understood that all one-dimensional paths are positive-valued and go from y' to y''. Substituting Eq. (3.37) into Eq. (1.14), we obtain the relation analogous to Eq. (3.15):

$$S_L[r^{(N)}] - S_L^{(N)} = \frac{\epsilon}{\hbar}\sum_{n=1}^{N}(B + B'), \quad (3.38)$$

where

$$B \equiv \frac{L(L+1)\hbar^2}{2m}\left\{\frac{1}{r_n^2} - \int_0^1 ds \frac{1}{\{p_n(s)\}^2}\right\} = \frac{L(L+1)\hbar^2(r_{n-1} - r_n)}{2mr_n^2 r_{n-1}}, \quad (3.39)$$

$$B' \equiv V(r_n) - \int_0^1 ds V(p_n(s)), \quad (3.40)$$

$$p_n(s) \equiv r_n - (1-s)(r_n - r_{n-1}). \quad (3.41)$$

By procedures analogous to those used to obtain (3.19), we find that the rhs of Eq. (3.38) is of order $N^{-\frac{1}{2}}$. It then follows from (3.36) that Eq. (3.33) can be written as

$$\rho_L(r'', r'; T/\hbar) = \lim_{N\to\infty}\int_0^\infty dr_1 \cdots \int_0^\infty dr_{N-1} C(N)^{\frac{1}{2}} \exp S_L[r^{(N)}] \equiv \int_{y'}^{y''} dr(t) e^{S_L}, \quad (3.42)$$

with the motivation for this definition being provided by Eq. (3.21).

C. Variational Bounds on ρ

Feynman developed a VLB on ρ by introducing a real-valued trial functional $\tilde{S}$ and writing[5]

$$\rho(\mathbf{r}'', \mathbf{r}'; T/\hbar) = \left\{\int_{x'}^{x''} D\mathbf{r}(t) e^{\tilde{S}}\right\} \times \int_{x'}^{x''} D\mathbf{r}(t) e^{\tilde{S}} e^{S-\tilde{S}} \Big/ \int_{x'}^{x''} D\mathbf{r}(t) e^{\tilde{S}}. \quad (3.43)$$

In view of the fact that $\tilde{S}$ is arbitrary, the expression in curly brackets is not necessarily a function of t' and t'' only through their difference $t'' - t'$; we therefore use the notation

$$\int_{x'}^{x''} D\mathbf{r}(t) e^{\tilde{S}} \equiv \tilde{\rho}(x''; x'). \quad (3.44)$$

Since $\tilde{S}$ is real-valued, $e^{\tilde{S}}$ is positive-valued, and consequently $\tilde{\rho}$ is a positive-definite function. The quantity

$$\langle F\rangle \equiv \int_{x'}^{x''} D\mathbf{r}(t) e^{\tilde{S}} F \Big/ \int_{x'}^{x''} D\mathbf{r}(t) e^{\tilde{S}} \quad (3.45)$$

represents a normalized, positive-weighted average of an arbitrary functional F, the dependence on x' and x'' being understood. In this notation, Eq. (3.43) is

$$\rho(\mathbf{r}'', \mathbf{r}'; T/\hbar) = \tilde{\rho}(x''; x')\langle e^{S-\tilde{S}}\rangle. \quad (3.46)$$

Since the exponential is concave upwards, Jensen's inequality yields the relation

$$\langle e^{S-\tilde{S}}\rangle \ge e^{\langle S-\tilde{S}\rangle}, \quad (3.47)$$

from which it follows that

$$\rho^F(x''; x') \equiv \tilde{\rho}(x''; x') e^{\langle S-\tilde{S}\rangle} \quad (3.48)$$

is a VLB on $\rho(\mathbf{r}'', \mathbf{r}'; T/\hbar)$. Specializations of Feynman's bound ρ^F are obtained by using the inequality

$$e^Z \ge \sum_{j=0}^{2n+1} \frac{Z^j}{j!}, \quad n = 0, 1, \cdots, \quad \mathrm{Im}\, Z = 0, \quad (3.49)$$

with $Z = \langle S - \tilde{S}\rangle$.

Different bounds on ρ are obtained by writing

$$\rho(\mathbf{r}'', \mathbf{r}'; T/\hbar) = \int_{x'}^{x''} D\mathbf{r}(t) e^{\tilde{S}} e^{S-\tilde{S}} \quad (3.50)$$

and using the inequality (3.49) with $Z = S - \tilde{S}$. In

this way, it is seen that

$$\dot{\rho}^{(n)}(x''; x') \equiv \tilde{\dot{\rho}}(x''; x') \sum_{j=0}^{2n+1} \frac{1}{j!} \langle (S - \tilde{S})^j \rangle, \quad n = 0, 1, \cdots, \tag{3.51}$$

are also VLB's on ρ. If the exact and trial functionals S and $\tilde{S}$ differ by order ν, with $|\nu| \ll 1$, then $\rho - \rho^{(n)}$ is positive and of order ν^{2n+2}. For future use, we note that the VLB's ρ^F and $\rho^{(0)}$ (which is to be distinguished from the free-particle density matrix ρ^0) are given by

$$\dot{\rho}^F = \tilde{\dot{\rho}} e^{\tilde{I}/\tilde{\rho}} \tag{3.52}$$

and[15]

$$\dot{\rho}^{(0)} = \tilde{\dot{\rho}} + \tilde{I}, \tag{3.53}$$

respectively, where

$$\tilde{I}(x''; x') \equiv \int_{x'}^{x''} D\mathbf{r}(t) e^{\tilde{S}} (S - \tilde{S}). \tag{3.54}$$

D. Variational Bounds on ρ_L

Because of the formal similarity between the path-integral representations of ρ and ρ_L, the procedures used to bound ρ can also be used to bound ρ_L. Thus, we introduce the real-valued trial functional $\tilde{S}_L$, which is defined on the one-dimensional paths $r(t)$, and we define, in analogy to Eqs. (3.44) and (3.45), the quantities

$$\tilde{\rho}_L(y''; y') \equiv \int_{y'}^{y''} dr(t) e^{\tilde{S}_L}, \tag{3.55}$$

$$\langle F \rangle_L \equiv \int_{y'}^{y''} dr(t) e^{\tilde{S}_L} F \Big/ \int_{y'}^{y''} dr(t) e^{\tilde{S}_L}, \tag{3.56}$$

where F is a functional of the paths $r(t)$ and where the dependence of $\langle F \rangle_L$ on the endpoints y'' and y' has been suppressed. It then follows that

$$\rho_L^F(y''; y') \equiv \tilde{\rho}_L(y''; y') \exp \langle S_L - \tilde{S}_L \rangle_L \tag{3.57}$$

and

$$\rho_L^{(n)}(y''; y') \equiv \tilde{\rho}_L(y''; y') \sum_{j=0}^{2n+1} \frac{1}{j!} \langle (S_L - \tilde{S}_L)^j \rangle_L, \quad n = 0, 1, \cdots, \tag{3.58}$$

are VLB's on ρ_L. We omit the details. For future use, we note that the VLB's ρ_L^F and $\rho_L^{(0)}$ are given by

$$\rho_L^F = \tilde{\rho}_L e^{\tilde{I}_L/\tilde{\rho}_L}, \tag{3.59}$$

$$\rho_L^{(0)} = \tilde{\rho}_L + \tilde{I}_L, \tag{3.60}$$

where

$$\tilde{I}_L(y''; y') \equiv \int_{y'}^{y''} dr(t) e^{\tilde{S}_L} (S_L - \tilde{S}_L). \tag{3.61}$$

For the sake of completeness, we consider the quantity

$$B_L \equiv 2\pi r'' r' \int_{-1}^{1} d\xi P_L(\xi) \rho(x''; x').$$

Since $P_0(\xi) = 1$, it follows from Eq. (1.8) that B_0 is a VLB on ρ_0. For $L \neq 0$, however, P_L is an oscillatory function of ξ, and B_L is therefore not in general a VLB on ρ_L.

E. Variational Bounds on z_L

Every VLB ρ_L on ρ_L generates a corresponding VLB z_L on z_L through the inequality

$$\rho_L(r'', r'; \beta) - \rho_L^0(r'', r'; \beta) \geq \rho_L(y''; y') - \rho_L^0(r'', r'; \beta). \tag{3.62}$$

Setting $r'' = r' = r$ and integrating both sides of the inequality over all values of r, we see from Eqs. (1.16) and (1.17) that the lhs reduces to $z_L(\beta)$. Since the order of the inequality is preserved, we conclude that

$$z_L(t''; t') \equiv \int_0^\infty dr \{ \rho_L(r, t''; r, t') - \rho_L^0(r, r; \beta) \} \tag{3.63}$$

is a VLB on $z_L(\beta)$, where, as noted previously, $\beta\hbar = t'' - t'$. For future use, we note that $\rho_L^{(0)}$ generates the VLB

$$z_L^{(0)}(t''; t') = \int_0^\infty dr \{ \tilde{\rho}_L(r, t''; r, t') - \rho_L^0(r, r; \beta) + \tilde{I}_L(r, t''; r, t') \} \tag{3.64}$$

on z_L. (We note that all results in Sec. 3 are valid whether or not the potential can support a bound state.)

4. SPECIALIZATION TO THE STANDARD RESULTS

We now show that the standard results are obtained when the trial functionals $\tilde{S}$ and $\tilde{S}_L$ are of the forms given by Eqs. (1.7) and (1.15), respectively, where the trial potential $\tilde{V}$ is short-range, central, and incapable of supporting a bound state. [In connection with this last requirement note the remark following Eq. (1.15).]

Since the properties of $\tilde{V}$ are precisely those assumed for the exact potential V, it is clear that we can use all of the results in this article for the system associated with $\tilde{V}$ by merely inserting a tilde (~) where appropriate. From Eq. (1.3), for example, it follows that the path integral $\tilde{\rho}(x''; x')$ defined by Eq. (3.44) is identical

[15] Using the results of Sec. 4 and Appendix D, one can show that, for $\tilde{S}$ of the form given by Eq. (1.7), the inequality $\rho \geq \tilde{\rho} + \tilde{I}$ is the coordinate representation of the operator inequality

$$e^{-\beta H} \geq e^{-\beta \tilde{H}} + \int_0^\beta d\tau e^{-(\beta-\tau)\tilde{H}} (\tilde{H} - H) e^{-\tau \tilde{H}}.$$

to the density matrix $\tilde{\rho}(\mathbf{r}'', \mathbf{r}'; \beta)$ for the system described by the Hamiltonian $\tilde{H} = H^0 + \tilde{V}(r)$. Similarly, Eq. (1.13) implies that the path integral $\tilde{\rho}_L(y''; y')$ defined by Eq. (3.55) is identical to the function $\tilde{\rho}_L(r'', r'; \beta)$ obtained by solving Eqs. (2.11) and (2.12) with the (reduced) Hamiltonian $\tilde{H}_L = H_L^0 + \tilde{V}(r)$.

We consider first the VUB's $A_0(\rho^{(0)})$ on A_0 and $A_L'(\rho_L^{(0)})$ on A_L. In view of the discussion in the preceding paragraph, it follows from Eqs. (2.10) and (3.53) that

$$A_0(\rho^{(0)}) = \tilde{A}_0 - \lim_{r'\to\infty} \tfrac{1}{2} r' \lim_{\beta\to\infty} (2\pi\hbar^2\beta m^{-1})^{\frac{3}{2}} \tilde{I}, \quad (4.1)$$

and it similarly follows from Eqs. (2.25) and (3.60) that

$$A_L'(\rho_L^{(0)}) = \tilde{A}_L - \lim_{r'\to\infty} \frac{2L+1}{2r'} \lim_{\beta\to\infty} \frac{\beta^{L+\frac{3}{2}}}{a(L)} \tilde{I}_L. \quad (4.2)$$

If we define

$$W(r) \equiv \tilde{V}(r) - V(r) = \tilde{H} - H = \tilde{H}_L - H_L, \quad (4.3)$$

then, from Eqs. (1.4) and (1.7) and from Eqs. (1.14) and (1.15), we see that

$$S - \tilde{S} = \frac{1}{\hbar}\int_{t'}^{t''} dt W(r) = S_L - \tilde{S}_L, \quad (4.4)$$

from which it follows that the path integrals $\tilde{I}$ and $\tilde{I}_L$ defined by Eqs. (3.54) and (3.61), respectively, are given by

$$\tilde{I}(x''; x') = \frac{1}{\hbar}\int_{x'}^{x''} D\mathbf{r}(t) e^{\tilde{S}} \int_{t'}^{t''} dt W(r), \quad (4.5)$$

$$\tilde{I}_L(y''; y') = \frac{1}{\hbar}\int_{y'}^{y''} dr(t) e^{\tilde{S}_L} \int_{t'}^{t''} dt W(r). \quad (4.6)$$

By analogy with Eq. (3.21), we have for $\tilde{I}$ the explicit expression

$$\tilde{I} = \frac{1}{\hbar} \lim_{N\to\infty} \int d\mathbf{r}_1 \cdots \int d\mathbf{r}_{N-1} C(N) \exp\{\tilde{S}[\mathbf{r}^{(N)}]\} \times \int_{t'}^{t''} dt W(|\mathbf{r}^{(N)}(\mathbf{R}; t)|). \quad (4.7)$$

Interchanging orders of integration and changing the dummy variable in the time integral from t to s, we obtain

$$\tilde{I} = \frac{1}{\hbar} \lim_{N\to\infty} \int_{t'}^{t''} ds \int d\mathbf{r}_1 \cdots \int d\mathbf{r}_{N-1} \times C(N) \exp\{\tilde{S}[\mathbf{r}^{(N)}]\} W(|\mathbf{r}^{(N)}(\mathbf{R}; s)|). \quad (4.8)$$

Taking the limit $N \to \infty$ inside the s integral, we see, again by analogy with Eq. (3.21), that $\tilde{I}$ is given by

$$\tilde{I} = \frac{1}{\hbar}\int_{t'}^{t''} ds \tilde{J}(s), \quad (4.9)$$

$$\tilde{J}(s) \equiv \int_{x'}^{x''} D\mathbf{r}(t) e^{\tilde{S}} W[|\mathbf{r}(s)|], \quad (4.10)$$

where the notation $W[|\mathbf{r}(s)|]$ indicates that the functional W depends on the distance from the origin of the path only at the single instant s. Using the exactly analogous procedure, we obtain the relations

$$\tilde{I}_L = \frac{1}{\hbar}\int_{t'}^{t''} ds \tilde{J}_L(s), \quad (4.11)$$

$$\tilde{J}_L(s) \equiv \int_{y'}^{y''} dr(t) e^{\tilde{S}_L} W[r(s)]. \quad (4.12)$$

Path integrals containing a functional which depends on the position of the path at a single instant are considered in Appendix D, and it follows from the results therein that

$$\tilde{I}(x''; x') = \int_0^\beta d\tau \int d\mathbf{r} \tilde{\rho}(\mathbf{r}'', \mathbf{r}; \beta - \tau) W(r) \tilde{\rho}(\mathbf{r}, \mathbf{r}'; \tau) \equiv \tilde{I}(\mathbf{r}'', \mathbf{r}'; \beta), \quad (4.13)$$

$$\tilde{I}_L(y''; y') = \int_0^\beta d\tau \int_0^\infty dr \tilde{\rho}_L(r'', r; \beta - \tau) W(r) \tilde{\rho}_L(r, r'; \tau) \equiv \tilde{I}_L(r'', r'; \beta), \quad (4.14)$$

where

$$\tau \equiv (s - t')/\hbar.$$

Substituting for each $\tilde{\rho}$ in Eq. (4.13) the eigenfunction expansion given by Eq. (2.3) with P replaced by $\tilde{P}$, after interchanging orders of integration, we find that

$$\tilde{I} = \int d\mathbf{r} \int d\mathbf{k} \int d\mathbf{k}' \tilde{P}(\mathbf{r}'', \mathbf{r}; \mathbf{k}) W(r) \tilde{P}(\mathbf{r}, \mathbf{r}'; \mathbf{k}') Y, \quad (4.15)$$

$$Y \equiv e^{-\beta E_k} \int_0^\beta d\tau e^{\tau(E_k - E_{k'})} = (e^{-\beta E_{k'}} - e^{-\beta E_k})(E_k - E_{k'})^{-1}. \quad (4.16)$$

Since Y has a well-defined, finite value for $E_k = E_{k'}$, the insertion of a principal value (denoted by $\mathcal{P}$) in the second line of (4.16) has no effect. The insertion is useful, however, for it makes it possible to split the rhs of Eq. (4.15) into two terms. Inserting the principal value in Eq. (4.16) and substituting the resulting expression for Y into Eq. (4.15), we find after some

manipulating that

$$\tilde{I} = -\int d\mathbf{r}\int d\mathbf{k} e^{-\beta E_k}\tilde{Q}(\mathbf{r}'', \mathbf{r}, \mathbf{r}'; \mathbf{k}), \quad (4.17)$$

$$\tilde{Q}(\mathbf{r}'', \mathbf{r}, \mathbf{r}'; \mathbf{k}) \equiv \{\tilde{G}(\mathbf{r}'', \mathbf{r}; k)W(r)\tilde{P}(\mathbf{r}, \mathbf{r}'; \mathbf{k})\} + \{\mathbf{r}'' \leftrightarrow \mathbf{r}'\}^*, \quad (4.18)$$

where the asterisk denotes complex conjugation and where

$$\tilde{G}(\mathbf{r}'', \mathbf{r}'; k) = \mathcal{P}\int d\mathbf{k}' \frac{\tilde{P}(\mathbf{r}'', \mathbf{r}'; \mathbf{k}')}{E_k - E_{k'}}. \quad (4.19)$$

Since $\tilde{G}$ satisfies the inhomogeneous time-independent Schrödinger equation with the full Hamiltonian $\tilde{H}$,

$$(E_k - \tilde{H})\tilde{G}(\mathbf{r}, \mathbf{r}'; k) = \int d\mathbf{k}'\tilde{P}(\mathbf{r}, \mathbf{r}'; \mathbf{k}') = \delta^{(3)}(\mathbf{r} - \mathbf{r}'), \quad (4.20)$$

and since in Eq. (4.19) the principal value of the integral is taken at the singularity $E_{k'} = E_k$, $\tilde{G}$ is the full Green's function appropriate to the boundary condition of standing waves at infinity. It has the angular-momentum decomposition

$$\tilde{G}(\mathbf{r}'', \mathbf{r}'; k) = \sum_{L=0}^{\infty} \frac{2L+1}{4\pi r'' r'} P_L(\xi)\tilde{G}_L(r'', r'; k), \quad (4.21)$$

where

$$\tilde{G}_L(r'', r'; k) = \frac{2}{\pi}\mathcal{P}\int_0^\infty dk' \frac{\tilde{u}_L(k', r'')\tilde{u}_L(k', r')}{E_k - E_{k'}} = (2m/\hbar^2)\tilde{u}_L(k, r_<)\tilde{v}_L(k, r_>)k^{-1}, \quad (4.22)$$

$r_<$ and $r_>$ denoting the smaller and larger, respectively, of r'' and r'. The function $\tilde{v}_L(k, r)$ is the (real) irregular solution of the radial equation

$$(E_k - \tilde{H}_L)\tilde{v}_L(k, r) = 0, \quad (4.23)$$

with the asymptotic form

$$\tilde{v}_L(k, r) \to kr\cos\tilde{\eta}_L\{n_L(kr) + \tan\tilde{\eta}_L j_L(kr)\}, \quad r \to \infty. \quad (4.24)$$

It has the threshold-energy dependence

$$\tilde{v}_L(k, r) \to k^{-L}\tilde{v}_L(r), \quad k \to 0, \quad (4.25)$$

where $\tilde{v}_L(r)$ is the (real) irregular solution of the zero-energy radial equation

$$\tilde{H}_L\tilde{v}_L(r) = 0, \quad (4.26)$$

with the asymptotic form

$$\tilde{v}_L(r) \to -(2L-1)!!\, r^{-L}, \quad r \to \infty. \quad (4.27)$$

Using the procedure exactly analogous to that leading from Eq. (4.13) to Eq. (4.17), we find that Eq. (4.14) can be expressed in the form

$$\tilde{I}_L = -\int_0^\infty dr\int_0^\infty dk e^{-\beta E_k}\tilde{Q}_L(r'', r, r'; k), \quad (4.28)$$

$$\tilde{Q}_L(r'', r, r'; k) \equiv \{\tilde{G}_L(r'', r; k)W(r)(2/\pi) \times \tilde{u}_L(k, r)\tilde{u}_L(k, r')\} + \{r'' \leftrightarrow r'\}. \quad (4.29)$$

In the limit of large β, only the neighborhood of zero energy contributes to the integral over $\mathbf{k}$ in Eq. (4.17). We can therefore extract the smoothly varying function $\tilde{Q}$ at $\mathbf{k} = 0$ and perform the $\mathbf{k}$ integration to obtain

$$\tilde{I}(\mathbf{r}'', \mathbf{r}'; \beta) \to -(2\pi m/\beta\hbar^2)^{\frac{3}{2}}\int d\mathbf{r}\tilde{Q}(\mathbf{r}'', \mathbf{r}, \mathbf{r}'; 0), \quad \beta \to \infty. \quad (4.30)$$

Substituting (4.30) into Eq. (4.1) and using Eqs. (4.18) and (2.4), we obtain

$$A_0(\rho^{(0)}) = \tilde{A}_0 + \lim_{r'\to\infty} r'\tilde{\psi}(0, \mathbf{r}') \times \int d\mathbf{r}\tilde{G}(\mathbf{r}', \mathbf{r}; 0)W(r)\tilde{\psi}(0, \mathbf{r}), \quad (4.31)$$

where we have made use of the fact that the outgoing scattering solution is real at zero energy. Since both V and $\tilde{V}$ are short-range, $W(r)$ is negligible beyond a finite distance, and $\tilde{G}$ in Eq. (4.31) can be replaced by its asymptotic form[16]

$$\tilde{G}(\mathbf{r}', \mathbf{r}; 0) \to -(2m/\hbar^2)(4\pi r')^{-1}\tilde{\psi}(0, \mathbf{r}), \quad r \ll r', \quad r' \to \infty. \quad (4.32)$$

Substituting (4.32) into Eq. (4.31) and using the relation

$$W(r)\tilde{\psi}(0, \mathbf{r}) = -H\tilde{\psi}(0, \mathbf{r}), \quad (4.33)$$

this relation being a consequence of Eqs. (4.3) and (2.5), we arrive at the inequality

$$A_0 \le A_0(\rho^{(0)}) = \tilde{A}_0 + (m/2\pi\hbar^2)\int d\mathbf{r}\tilde{\psi}(0, \mathbf{r})H\tilde{\psi}(0, \mathbf{r}). \quad (4.34)$$

Since $\psi(0, \mathbf{r}) = \psi(0, r) = u_0(r)/r$, we set $\tilde{\psi}(0, \mathbf{r}) = \tilde{u}_0(r)/r$. The upper bound on A_0 then becomes

$$\tilde{A}_0 + (2m/\hbar^2)\int_0^\infty dr\tilde{u}_0(r)H_0\tilde{u}_0(r),$$

the usual result. We note that $A_0(\rho^F)$ can also be shown to reduce to the rhs of the inequality (4.34).

A similar procedure can be used to connect the VUB in Eq. (4.2) with the standard result. Setting $r'' = r'$ in Eq. (4.28) and using the usual arguments

[16] See, for example, A. Messiah, *Quantum Mechanics* (North-Holland Publ. Co., Amsterdam, 1965), Vol. II, Chap. XIX.

concerning the k integration in the limit of large β, we replace $\tilde{Q}_L$ by

$$k^{2L+2}2\tilde{G}_L(r', r; 0)W(r)(2/\pi)\tilde{u}_L(r)\tilde{u}_L(r');$$

after integrating over k, we find that

$$\tilde{I}_L(r', r'; \beta) \to -\frac{2a(L)}{\beta^{L+\frac{3}{2}}} \times \int_0^\infty dr\tilde{G}_L(r', r; 0)W(r)\tilde{u}_L(r)\tilde{u}_L(r'), \quad \beta \to \infty. \quad (4.35)$$

From (2.18) and (4.25), it follows that

$$\tilde{G}_L(r', r; 0) = (2m/\hbar^2)\tilde{u}_L(r_<)\tilde{v}_L(r_>); \quad (4.36)$$

since $W(r)$ is negligible beyond a finite range, we obtain for the integral in (4.35) the asymptotic form

$$\int_0^\infty dr\tilde{G}_L(r', r; 0)W(r)\tilde{u}_L(r) \to (2m/\hbar^2)\tilde{v}_L(r')(\tilde{u}_L, W\tilde{u}_L), \quad r' \to \infty, \quad (4.37)$$

where

$$(\tilde{u}_L, W\tilde{u}_L) = \int_0^\infty dr\tilde{u}_L(r)W(r)\tilde{u}_L(r). \quad (4.38)$$

Substituting (4.35) into Eq. (4.2) and using (4.37), (2.21), (4.27), and the relation

$$W(r)\tilde{u}_L(r) = -H_L\tilde{u}_L(r), \quad (4.39)$$

we obtain the inequality

$$A_L \leq A'_L(\rho_L^{(0)}) = \tilde{A}_L + (2m/\hbar^2)(\tilde{u}_L, H_L\tilde{u}_L), \quad (4.40)$$

which is the usual result.[6] We note that for the given choice of $\tilde{S}_L$, the VUB $A'_L(\rho_L^F)$ on A_L also reduces to the rhs of (4.40).

We consider next the relation between the VUB $\tilde{A}_L(z_L^{(0)})$ and the standard result. From the discussion in the second paragraph of this section, it follows that Eq. (3.64) can be rewritten as

$$z_L^{(0)}(t''; t') = \int_0^\infty dr\{\tilde{\rho}_L(r, r; \beta) - \rho_L^0(r, r; \beta) + \tilde{I}_L(r, r; \beta)\} = \tilde{z}_L(\beta) + \int_0^\infty dr\tilde{I}_L(r, r; \beta) \equiv z_L^{(0)}(\beta). \quad (4.41)$$

Using Eq. (4.14) and the relation

$$\int_0^\infty dr\tilde{\rho}_L(r', r; \tau)\tilde{\rho}_L(r, r''; \beta - \tau) = \tilde{\rho}_L(r', r''; \beta), \quad (4.42)$$

which is a special case of Eq. (D14), we obtain

$$\int_0^\infty dr\tilde{I}_L(r, r; \beta) = \beta\int_0^\infty dr\tilde{\rho}_L(r, r; \beta)W(r). \quad (4.43)$$

Substituting Eq. (4.43) into Eq. (4.41) and using (2.28) and (2.23), we find that

$$z_L^{(0)}(\beta) \to -(\hbar^2/2m)a(L)\beta^{-(L+\frac{1}{2})} \times \{\tilde{A}_L - (2m/\hbar^2)(\tilde{u}_L, W\tilde{u}_L)\}, \quad \beta \to \infty. \quad (4.44)$$

It then follows from Eqs. (2.29) and (4.39) that

$$A_L \leq \tilde{A}_L(z_L^{(0)}) = \tilde{A}_L + (2m/\hbar^2)(\tilde{u}_L, H_L\tilde{u}_L), \quad (4.45)$$

which is the standard result.

Note added in proof: A VUB on A_L when bound states exist has been found and will be submitted for publication shortly. A path-integral representation of ρ_L has recently been developed by D. Peak and A. Inomata, J. Math. Phys. **10,** 1422 (1969). See also P. Pechukas, Phys. Rev. **181,** 166 and 174 (1969).

ACKNOWLEDGMENTS

The authors would like to thank Dr. K. Symanzik and Dr. M. Lieber for many valuable discussions.

APPENDIX A: PROOF OF EQUATION (2.26)

The well-known relation given by Eq. (2.26) is usually proved[17] by introducing a large but finite radius R (which ultimately goes to infinity) and imposing the boundary conditions

$$u_L(k, R) = u_L^0(kR) = 0,$$

where $u_L^0(z) \equiv zj_L(z)$ denotes the Riccatti–Bessel function. We present here a simple if heuristic proof which does not utilize this procedure.[18]

The function

$$\frac{du_L(k, r)}{dk} \equiv \dot{u}_L \quad (A1)$$

satisfies the inhomogeneous equation

$$(E_k - H_L)\dot{u}_L = (-\hbar^2/2m)2ku_L, \quad (A2)$$

as may be verified by differentiating Eq. (2.14) with respect to k. Multiplying Eq. (A2) by u_L, multiplying Eq. (2.14) by $\dot{u}_L$, and then subtracting, we find that

$$\mathcal{W}' = 2ku_L^2, \quad (A3)$$

where the Wronskian $\mathcal{W}$ is defined by

$$\mathcal{W} \equiv u'_L\dot{u}_L - u_L\dot{u}'_L, \quad (A4)$$

and where the prime denotes differentiation with respect to r. In view of the boundary condition given in Eq. (2.15), $\mathcal{W}$ vanishes at the origin, and the integral

$$J \equiv \int_0^\infty dke^{-\beta E_k}\int_0^R dru_L^2(k, r) \quad (A5)$$

[17] See, for example, K. Huang, *Statistical Mechanics* (John Wiley & Sons, Inc., New York, 1963), Chap. 14, pp. 307–311.

[18] For another such proof, see Appendix C in R. N. Hill, J. Math. Phys. **9,** 1534 (1968).

can therefore be written as

$$J=\int_0^\infty dke^{-\beta E_k}\left(\frac{1}{2k}\,\mathfrak{W}\Big|^{r=R}\right). \tag{A6}$$

As seen from (2.18), the integrand in Eq. (A5) is of order k^{2L+2} near $k=0$. Consequently, the value of the integral remains unchanged if, roughly speaking, the point $k=0$ is excluded from the region of integration. With this point deleted, the form of J in the limit of large R can be determined by using for u_L the asymptotic form

$$u_L(k,r)\to \sin(kr-\tfrac{1}{2}L\pi+\eta_L),\quad r\to\infty,\quad 2kr\gg L(L+1). \tag{A7}$$

Substituting (A7) into Eq. (A4), we obtain

$$\mathfrak{W}\to kr+k\dot{\eta}_L-\tfrac{1}{2}\sin 2(kr-\tfrac{1}{2}L\pi+\eta_L),\quad r\to\infty,\quad 2kr\gg L(L+1), \tag{A8}$$

from which it follows that

$$J\to\tfrac{1}{2}\int_0^\infty dke^{-\beta E_k}\times\left(R+\dot{\eta}_L-\frac{1}{2k}\sin 2(kR-\tfrac{1}{2}L\pi+\eta_L)\right),\quad R\to\infty. \tag{A9}$$

Repeating the entire procedure for the potential $V\equiv 0$, corresponding to which the phase shifts vanish and $u_L(k,r)$ reduces to $u_L^0(kr)$, we find that

$$J^0\equiv\int_0^\infty dke^{-\beta E_k}\int_0^R dr\{u_L^0(kr)\}^2, \tag{A10}$$

the free-particle counterpart of J, has the asymptotic form

$$J^0\to\tfrac{1}{2}\int_0^\infty dke^{-\beta E_k}\left(R-\frac{1}{2k}\sin 2(kR-\tfrac{1}{2}L\pi)\right),\quad R\to\infty. \tag{A11}$$

Subtracting J^0 from J and interchanging the order of integration, we see from Eq. (2.13) that

$$\frac{2}{\pi}(J-J^0)=\int_0^R dr\{\rho_L(r,r;\beta)-\rho_L^0(r,r;\beta)\}\to\left\{\frac{1}{\pi}\int_0^\infty d\dot{\kappa}e^{-\beta E_k}\dot{\eta}_L\right\}+X+X',\quad R\to\infty, \tag{A12}$$

where

$$X\equiv\frac{(-1)^{L+1}}{2\pi}\int_0^\infty dk\,\frac{1}{k}\,e^{-\beta E_k}\sin 2\eta_L\cos 2kR, \tag{A13}$$

$$X'\equiv\frac{(-1)^{L}}{\pi}\int_0^\infty dk\,\frac{1}{k}\,e^{-\beta E_k}\sin^2\eta_L\sin 2kR. \tag{A14}$$

Integrating by parts, we obtain

$$X=(-1)^{L+1}(4\pi kr)^{-1}e^{-\beta E_k}\sin 2\eta_L\sin 2kr\Big|_0^\infty+Y,$$

$$X'=(-1)^{L+1}(2\pi kR)^{-1}e^{-\beta E_k}\sin^2\eta_L\cos 2kR\Big|_0^\infty+Y',$$

where Y and Y' each fall off at least as rapidly as R^{-2}, as can be seen by a further integration by parts. In view of (2.27), X and X' vanish in the limit $R\to\infty$, and it therefore follows from (A12) and from Eqs. (1.16) and (1.17) that

$$\lim_{R\to\infty}\frac{2}{\pi}(J-J^0)=z_L(\beta)=\frac{1}{\pi}\int_0^\infty dke^{-\beta E_k}\dot{\eta}_L.$$

APPENDIX B: PROOF OF EQUATION (3.1)

The proof of Eq. (3.1) is virtually identical to a proof which Feynman has given for an analogous relation involving the propagator.[19]

To begin, we note that the lhs of Eq. (3.1) is given in more detail in (3.9). Changing the variable of integration from $\mathbf{r}$ to $\mathbf{q}\equiv\mathbf{r}-\mathbf{r}''$, we can rewrite the lhs of (3.9) as

$$\exp\{-\epsilon V(r'')/\hbar\}J[\rho], \tag{B1}$$

where

$$J[F]\equiv\left(\frac{m}{2\pi\hbar\epsilon}\right)^{\frac{3}{2}}\int d\mathbf{q}\exp\left(-\frac{mq^2}{2\hbar\epsilon}\right)F(\mathbf{q}). \tag{B2}$$

Expanding $\rho(\mathbf{r}''+\mathbf{q},\mathbf{r}';t/\hbar)$ in a Taylor series in q_i, $i=1,2,3$, about $q_i=0$ and substituting into Eq. (B2), we obtain

$$J[\rho]=\left\{J[1]+\sum_{i=1}^3 J[q_i]\frac{\partial}{\partial r_i''}+\sum_{i,j=1}^3\tfrac{1}{2}J[q_iq_j]\frac{\partial^2}{\partial r_i''\partial r_j''}+\cdots\right\}\rho(\mathbf{r}'',\mathbf{r}';t/\hbar); \tag{B3}$$

Performing the integrations, we find that

$$J[1]=1,\quad J[q_iq_j]=\delta_{ij}(\hbar\epsilon/m); \tag{B4}$$

J with an odd number of factors is equal to zero, and J with an even number ($2n$) of factors is of order ϵ_n; δ_{ij} is the Kronecker delta function.

It is now easily seen that (B1) is given by

$$\left\{1-\frac{\epsilon V(r'')}{\hbar}+O(\epsilon^2)\right\}\left\{1+\frac{\hbar\epsilon}{2m}\nabla''^2+O(\epsilon^2)\right\}\rho(\mathbf{r}'',\mathbf{r}';t/\hbar)=\{1-(\epsilon H''/\hbar)+O(\epsilon^2)\}\rho(\mathbf{r}'',\mathbf{r}';t/\hbar). \tag{B5}$$

[19] Ref. 1, pp. 375–376.

In view of the Bloch equation (2.1), we see that the rhs of Eq. (B5) differs from the function

$$\rho(\mathbf{r}'', \mathbf{r}'; (t + \epsilon)/\hbar)$$

by terms of order ϵ^2.

APPENDIX C: PROOF OF EQUATION (3.22)

The proof of Eq. (3.22) is closely analogous to that of Eq. (3.1) given in Appendix B. To begin, we note that the lhs of Eq. (3.22) can be written as

$$\exp\left(-\frac{\epsilon}{\hbar}\left\{\frac{L(L+1)\hbar^2}{2mr''^2} + V(r'')\right\}\right) I[\rho_L], \quad \text{(C1)}$$

where

$$I[\phi] \equiv \left(\frac{m}{2\pi\hbar\epsilon}\right)^{\frac{1}{2}} \int_0^\infty dr \exp\left\{-\frac{m(r''-r)^2}{2\hbar\epsilon}\right\}\phi(r). \quad \text{(C2)}$$

Expanding $\rho_L(r, r'; t/\hbar)$ in a Taylor series in r about $r = r''$ and then substituting into Eq. (C2), we obtain

$$I[\rho_L] = \left\{I[1] + I[r - r''] \frac{\partial}{\partial r''} + \tfrac{1}{2} I[(r-r'')^2] \frac{\partial^2}{\partial r''^2} + \cdots\right\} \rho_L(r'', r'; t/\hbar). \quad \text{(C3)}$$

To first order in ϵ, we can replace the limits of integration 0 to ∞ by the limits $-\infty$ to ∞, since the contribution from $-\infty$ to 0 approaches zero at least as rapidly as $\exp(-mr''^2/2\hbar\epsilon)$. Integrating between the limits $-\infty$ and ∞, we find that

$$I[(r-r'')^{2n}] = \pi^{-\frac{1}{2}}\Gamma(n+\tfrac{1}{2})(2\hbar\epsilon/m)^n,$$
$$I[(r-r'')^{2n+1}] = 0,$$
$$n = 0, 1, \cdots. \quad \text{(C4)}$$

We then find that (C1) is given by

$$\left(1 - \frac{\epsilon}{\hbar}\left\{\frac{L(L+1)\hbar^2}{2mr''^2} + V(r'')\right\}\right) \times \left\{1 + \frac{\hbar\epsilon}{2m}\frac{\partial^2}{\partial r''^2}\right\} \rho_L(r'', r'; t/\hbar) = (1 - \epsilon H_L''/\hbar)\rho_L(r'', r'; t/\hbar), \quad \text{(C5)}$$

with the neglect of terms of order ϵ^2. From Eq. (2.11), it follows that the rhs of Eq. (C5) and the function

$$\rho_L(r'', r'; (t + \epsilon)/\hbar)$$

are the same to first order in ϵ.

APPENDIX D: PROOF OF EQUATIONS (4.13) AND (4.14)

In view of the discussion in the second paragraph of Sec. 4, it is sufficient for the purpose of proving Eq. (4.13) to show that the path integral

$$J(s) \equiv \int_{x'}^{x''} D\mathbf{r}(t) e^S F[\mathbf{r}(s)] \quad \text{(D1)}$$

and the quantity

$$J'(s) = \int d\bar{\mathbf{r}} \rho(\mathbf{r}'', \bar{\mathbf{r}}; (t'' - s)/\hbar) F(\bar{\mathbf{r}}) \rho(\bar{\mathbf{r}}, \mathbf{r}'; (s - t')/\hbar) \quad \text{(D2)}$$

are equal for $t' \leq s \leq t''$, where the notation for the functional F in Eq. (D1) indicates that F depends upon the position of the path only at the single instant s.

We consider first the case in which

$$t' < s < t''. \quad \text{(D3)}$$

The time instant s then divides the fundamental interval (t', t'') into two subintervals (t', s) and (s, t'') having lengths $T' \equiv s - t'$ and $T'' \equiv t'' - s$, respectively, with both T' and T'' nonzero. Partitioning (t', s) and (s, t'') into N' and N'' subintervals of lengths $\epsilon' \equiv T'/N'$ and $\epsilon'' \equiv T''/N''$, respectively, we see from Eq. (3.5) that

$$\rho(\bar{\mathbf{r}}, \mathbf{r}'; T'/\hbar) = \lim_{N' \to \infty} \int d\mathbf{r}_1 \cdots \int d\mathbf{r}_{N'-1} \times \prod_{j=1}^{N'} \rho^V(\mathbf{r}_j, \mathbf{r}_{j-1}; \epsilon'/\hbar) \quad \text{(D4)}$$

and

$$\rho(\mathbf{r}'', \bar{\mathbf{r}}; T''/\hbar) = \lim_{N'' \to \infty} \int d\mathbf{q}_1 \cdots \int d\mathbf{q}_{N''-1} \times \prod_{n=1}^{N''} \rho^V(\mathbf{q}_n, \mathbf{q}_{n-1}; \epsilon''/\hbar), \quad \text{(D5)}$$

where

$$\mathbf{r}_0 \equiv \mathbf{r}', \quad \mathbf{r}_{N'} \equiv \bar{\mathbf{r}} \equiv \mathbf{q}_0, \quad \mathbf{q}_{N''} \equiv \mathbf{r}''. \quad \text{(D6)}$$

Substituting Eqs. (D4) and (D5) into Eq. (D2) and interchanging the order in which the limits are taken and the integrations are performed, we obtain

$$J'(s) = \lim_{N'', N' \to \infty} \int d\mathbf{r}_1 \cdots \int d\mathbf{r}_{N'-1} \int d\bar{\mathbf{r}} \int d\mathbf{q}_1 \cdots \int d\mathbf{q}_{N''-1} \times \left\{\prod_{j=1}^{N'} \rho^V(\mathbf{r}_j, \mathbf{r}_{j-1}; \epsilon'/\hbar)\right\} F(\bar{\mathbf{r}}) \times \left\{\prod_{n=1}^{N''} \rho^V(\mathbf{q}_n, \mathbf{q}_{n-1}; \epsilon''/\hbar)\right\}. \quad \text{(D7)}$$

Using Eq. (1.1), we find that Eq. (D2) can also be expressed in the form

$$J'(s) = \langle \mathbf{r}'' | \exp\{-(t'' - s)H/\hbar\} F_{\text{op}} \times \exp\{-(s - t')H/\hbar\} | \mathbf{r}' \rangle, \quad \text{(D8)}$$

where

$$F_{\text{op}} \equiv \int d\mathbf{r}\, |\mathbf{r}\rangle\, F(\mathbf{r})\, \langle \mathbf{r}|. \tag{D9}$$

If $F(\mathbf{r}) = 1$ for all $\mathbf{r}$, then F_{op} reduces to the identity operator, the rhs of Eq. (D8) reduces to $\rho(\mathbf{r}'', \mathbf{r}'; T/\hbar)$, and we therefore have the representation

$$\int_{x'}^{x''} D\mathbf{r}(t) e^{S} = \lim_{N'', N' \to \infty} \int d\mathbf{r}_1 \cdots \int d\mathbf{r}_{N'-1} \int d\bar{\mathbf{r}} \int d\mathbf{q}_1 \cdots \int d\mathbf{q}_{N''-1} \times \left\{ \prod_{j=1}^{N'} \rho^{V}(\mathbf{r}_j, \mathbf{r}_{j-1}; \epsilon'/\hbar) \right\} \left\{ \prod_{n=1}^{N''} \rho^{V}(\mathbf{q}_n, \mathbf{q}_{n-1}; \epsilon''/\hbar) \right\}. \tag{D10}$$

The integrand on the rhs of this equation can be shown to be, to a sufficiently good approximation, proportional to the value of the functional e^S on the straight-line segment path $\mathbf{r}^*(t)$ passing sequentially through the space–time points

$$(\mathbf{r}', t'), (\mathbf{r}_1, t' + \epsilon'), \cdots, (\mathbf{r}_{N'-1}, s - \epsilon'), (\bar{\mathbf{r}}, s), (\mathbf{q}_1, s + \epsilon''), \cdots, (\mathbf{q}_{N''-1}, t'' - \epsilon''), (\mathbf{r}'', t'');$$

the demonstration is quite similar to that given in part A of Sec. 3 to justify Eq. (3.20), and we therefore omit the details. Since

$$\mathbf{r}^*(s) = \bar{\mathbf{r}}, \tag{D11}$$

it follows from the preceding sentence that the integrand in Eq. (D7) is proportional to the value of the functional

$$e^S F[\mathbf{r}(s)]$$

on the path $\mathbf{r}^*(t)$; since Eq. (D11) is valid for all values of N'' and N', we have by analogy with Eq. (D10) that the rhs of Eq. (D7) is a representation of the path integral $J(s)$.

We have therefore shown that $J(s)$ and $J'(s)$ are equal for s satisfying the inequality (D3). If $s = t'$, then it follows from Eq. (3.11) and Eq. (2.2) that

$$J(t') = F(\mathbf{r}') \int_{x'}^{x''} D\mathbf{r}(t) e^{S}$$

and

$$J'(t') = \int d\bar{\mathbf{r}}\, \rho(\mathbf{r}'', \bar{\mathbf{r}}; T/\hbar) F(\bar{\mathbf{r}}) \delta^{(3)}(\bar{\mathbf{r}} - \mathbf{r}'),$$

respectively. Obviously, J and J' are equal for $s = t'$. They are also equal for $s = t''$, as can easily be verified by again using Eqs. (3.11) and (2.2), and we conclude that

$$J(s) = J'(s), \quad t' \leq s \leq t''. \tag{D12}$$

Using an exactly analogous procedure, we can show that

$$\int_{y'}^{y''} dr(t) e^{S_L} \phi[r(s)] = \int_0^{\infty} d\bar{r}\, \rho_L(r'', \bar{r}; (t'' - s)/\hbar) \phi(\bar{r}) \rho_L(\bar{r}, r'; (s - t')/\hbar), \tag{D13}$$

for $t' \leq s \leq t''$; we omit the proof. In view of the discussion in the second paragraph of Sec. 4, the validity of Eq. (4.14) can be inferred from that of Eq. (D13).

We note that if $\phi(r) = 1$ for all values of r, Eq. (D13) reduces to

$$\rho_L(r'', r'; \beta) = \int_0^{\infty} dr\, \rho_L(r'', r; \beta - \tau) \rho_L(r, r'; \tau), \tag{D14}$$

where

$$\tau \equiv (s - t')/\hbar, \quad \beta = (t'' - t')/\hbar.$$

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 12 DECEMBER 1969

Kinematic Interaction for the Heisenberg Antiferromagnet at Low Temperature*†‡

D. C. HERBERT
Department of Physics (Solid State Theory), Imperial College, London S.W.7, England

(Received 17 February 1969)

The general theory of antiferromagnetic spin waves is examined from a new viewpoint and the nature of the antiferromagnetic ground state is clarified. It is shown that the Anderson canonical transformation to free spin waves is nonunitary with the magnon state vectors having large nonphysical projections, and a projection technique which restores unitarity at low temperatures is developed. The nonphysical projection of the state vectors is shown to give a large kinematic interaction even at zero temperature.

1. INTRODUCTION

A rigorous theory of spin-wave interactions in ferromagnetic insulators was given by Dyson[1,2] in 1956, and since then numerous papers have been written on this subject; however, the antiferromagnet has received comparatively little attention, and even the ground state is unknown. In the spin-wave approach to the theory of magnetism, the spin operators are mapped on to a subspace of a boson space, and the kinematic interaction, i.e., the restriction on the number of spin deviations possible at any given spin site, is generally neglected by working in the full boson space. Dyson[1,2] and Wortis[3] have shown that the kinematic effects of nonphysical states cancel out at low temperatures for the ferromagnet, but so far, corresponding proofs for the antiferromagnet have not been found, and consequently, the spin-wave picture of antiferromagnetism is suspect.

In view of the difficulties surrounding the kinematic interaction, other representations for the spin operators have been introduced by various authors.[4-8] The aim is to map spin operators into operators with c-number commutators, so that Wick's theorem and the powerful diagrammatic techniques of many-body theory can be applied, and also to avoid the difficulties of spin-wave theory by restricting calculations rigorously to the physical domain. Unfortunately, due to the mixing of kinematic and dynamic interactions, the graphical structure of these theories becomes very complicated, and only limited progress has been made.

It is also possible to work directly with the spin algebra and calculate the spin Green's functions by the equations-of-motion method. The equations of motion lead to an infinite chain of coupled Green's functions, and as there is no standard decoupling technique, the usual procedure is to choose the decoupling approximation to give results which agree with spin-wave theory at low temperatures. This is clearly unsatisfactory, as spin-wave results are not rigorous for the antiferromagnetic case.

In this paper the spin-wave approach is pursued as it clearly separates the kinematic and dynamic effects, which can then be studied independently. Tao Yuin[9] and Kenan[7] have recently found a large kinematic effect, even at zero temperature, which, in view of the fact that the boson Hamiltonian is not positive, casts doubt on the validity of the spin-wave picture of antiferromagnetism. Szaniecki,[10,11] however, claims that the kinematic interaction is exponentially small at low temperatures, though Dembinski[12] has severely criticised his conclusions. In this work it is shown that there are two distinct kinematic effects to be considered, so that it is still possible for the nonphysical eigenvalues to have a negligible effect at low temperature.

2. THE BOSON MAPPING

For simplicity, a two-sublattice model is considered, in which near neighbors of spins on one sublattice all lie on the other, and only near-neighbor interaction is included. The spin Hamiltonian for such a system can be written in the form

$$H = J \sum_{i\delta} \mathbf{S}_i^a \cdot \mathbf{S}_{i+\delta}^b - A\left[\sum_i S_i^{az} - \sum_j S_j^{bz}\right], \quad (1)$$

* Present address: Physics Department, University of Bristol, Bristol, England.

† Science Research Council Postdoctoral Fellow.

‡ Based on a Ph.D. Thesis recently submitted to the University of London.

[1] F. J. Dyson, Phys. Rev. **102**, 1230 (1956).

[2] F. J. Dyson, Phys. Rev. **102**, 1217 (1956).

[3] M. Wortis, Phys. Rev. **138**, A1126 (1964).

[4] H. L. Davis, Phys. Rev. **120**, 789 (1960).

[5] Y. L. Wang, H. Callen, and S. Shtrikman, Phys. Rev. **148**, 419 (1966).

[6] R. P. Kenan, J. Appl. Phys. **37**, 1453 (1966).

[7] R. P. Kenan, Phys. Rev. **159**, 430 (1967).

[8] R. L. Mills, R. P. Kenan, and J. Korringa, Physica **26**, 204 (1959).

[9] Tao Yuin, Chinese J. Phys. **22**, 357 (1966).

[10] J. Szaniecki, Acta Phys. Polon. **31**, 969 (1967).

[11] J. Szaniecki, Acta Phys. Polon. **32**, 271 (1967).

[12] S. T. Dembinski, Physica **35**, 119 (1967).

where J is the positive exchange parameter and A represents a positive, effective anisotropy field; a and b label sublattice spins with positive and negative z-components, respectively; and subscripts i, j, and $i+\delta$ denote the crystal sites of spin operators, with $i+\delta$ denoting a near-neighbor spin.

The spin Hamiltonian is defined in a Hilbert space which is a product of all the orthogonal localized spin spaces, and it is convenient to map these component spaces isometrically into boson spaces. There are two main reasons for doing this, the most important being that the noninteracting spin-wave theories of Anderson,[13] Kubo,[14] Ziman,[15,16] and others seem to be a good first approximation at low temperatures, indicating that the low-lying spectrum may be well described by bosons. The other reason is that the boson operators have c-number commutators, so that Wick's theorem is applicable.

The mathematical technique for mapping Fock spaces has been developed in the context of quantum field theories with indefinite metric; here the discussion given by Nagy[17] is used. Consider a single spin operator having the usual commutation relations, which may conveniently be written in the form

$$[S_j^z, S_k^+] = \delta_{jk}S_j^+, \quad [S_j^z, S_k^-] = -\delta_{j,k}S_j^-,$$
$$[S_j^+, S_k^-] = 2\delta_{jk}S_j^z, \tag{2}$$

where

$$S_j^\pm = S_j^x \pm iS_j^y.$$

The ground state $|0_j\rangle$ is defined by

$$S_j^-|0_j\rangle = 0, \quad S_j^z|0_j\rangle = -S|0_j\rangle, \tag{3}$$

and the excited states can be represented by

$$|u_j\rangle = [(2s)^{u_j}u_j!]^{-\frac{1}{2}}(S_j^+)^{u_j}|0_j\rangle, \tag{4}$$

where the state $|u_j\rangle$ has u spin deviations from the ground state. These vectors are orthogonal but not normalized:

$$\langle u_i \,|\, v_j\rangle = \delta_{ij}\delta_{uv}F_u,$$
$$F_u = 1\cdot\left(1-\frac{1}{2s}\right)\cdots\left(1-\frac{u-1}{2s}\right). \tag{5}$$

The following set of orthonormal boson state-vectors is also defined:

$$|u_j) = (u_j!)^{-\frac{1}{2}}(a_j^\dagger)^{u_j}|0_j),$$
$$(u_i \,|\, v_j) = \delta_{uv}\delta_{ij}, \tag{6}$$
$$[a_j, a_{j'}^\dagger] = \delta_{jj'}.$$

A linear one-to-one correspondence of the states $|u_j\rangle$ and $|u_j)$ is now defined such that

$$\alpha|u_j\rangle + \beta|v_j\rangle \rightarrow \alpha|u_j) + \beta|v_j), \tag{7}$$

where α and β are scalars. The metric structure of a Hilbert space is given in terms of the inner product, so that the inner product of vectors in spin space can be used to define a metric operator in the boson space by the relationship

$$\langle u_i \,|\, v_i\rangle = (u_i|\,\eta_i\,|v_i). \tag{8}$$

Having established a metric, the operators in spin space can be mapped into the boson space by requiring that all corresponding matrix elements are equal, so that the boson operator $\hat{O}_i$ corresponding to the spin operator O_i is given by

$$\langle u_i|\,O_i\,|v_i\rangle = (u_i|\,\eta_i\hat{O}_i\,|v_i). \tag{9}$$

The adjoint of an operator has the same effect on the bra vector as the operator has on the ket when the natural imbedding is used to relate bra and ket vectors, and in elementary quantum mechanics, where the Hilbert space of state vectors has the identity operator for metric, or the metric is trivially reducible to the identity by normalizing state vectors, the adjoint and Hermitian conjugate coincide. However, the boson space considered here contains vectors having no correspondents in the physical spin space, and consequently these vectors have zero norm. Hence, the metric operator cannot be reduced to the identity, and, in general, operators will not commute with the metric, so that it is necessary to consider the relationship between adjoints and Hermitian conjugates. If the Hermitian conjugate is denoted by the symbol "$\dagger$" and the adjoint by "*," then in the boson space Hermitian conjugates are defined in the usual way, by putting the metric operator equal to the identity

$$\{(u_i|\,\hat{O}_i\,|v_i)\}^* = (v_i|\,\hat{O}_i^\dagger\,|u_i), \tag{10}$$

where $\{\cdots\}^*$ denotes complex conjugation. Since $\langle u_i \,|\, u_i\rangle = (u_i|\,\eta_i\,|u_i)$ is real, it follows that the metric operator is Hermitian:

$$\eta_i^\dagger = \eta_i. \tag{11}$$

The adjoint operator $\hat{O}_i^*$ is defined by considering a matrix element in the physical spin space and mapping:

$$\langle u_i|\,O_i^*\,|v_i\rangle = \{\langle v_i|\,O_i\,|u_i\rangle\}^* = \{(v_i|\,\eta_i\hat{O}_i\,|u_i)\}^*$$
$$= (u_i|\,\hat{O}_i^\dagger\eta_i\,|v_i) = (u_i|\,\eta_i\eta_i^{-1}\hat{O}_i^\dagger\eta_i\,|v_i). \tag{12}$$

Since $|u_i\rangle$ and $|v_i\rangle$ are arbitrary, it follows that

$$\hat{O}_i^* = \eta_i^{-1}\hat{O}_i^\dagger\eta_i. \tag{13}$$

[13] P. W. Anderson, Phys. Rev., **86**, 694 (1952).
[14] R. Kubo, Phys. Rev. **87**, 568 (1952).
[15] J. M. Ziman, Proc. Phys. Soc. London **A65**, 540 (1952).
[16] J. M. Ziman, Proc. Phys. Soc. London **A65**, 548 (1952).
[17] K. L. Nagy, Nuovo Cimento Suppl. **17**, 92 (1960).

It is important to notice that, rigorously, η_i^{-1} is singular, since η_i is a projection operator having eigenvalues of zero. However, if it is considered to operate only in the subspace of boson state vectors which have physical correspondents, then η_i cannot have eigenvalues of zero and the inverse can be defined. For the case of operators defined in spin space, this procedure is valid, though great care is needed when making transformations on such operators in the boson space, as it can easily happen that the resulting operators are nonphysical, in the sense that they are only defined on the full boson Hilbert space. It turns out that the canonical transformation to free spin waves is of this type, and the usual noninteracting spin-wave theory arises from a transformation which is nonunitary in the physical subspace.

An observable in spin space is represented by a self-adjoint operator which has real expectation values, and, consequently, the boson equivalent is self-adjoint but not necessarily Hermitian, as can be seen from the following argument:

$$\langle u_i|\, O_i^\dagger\, |u_i\rangle = \langle u_i|\, O_i\, |u_i\rangle = (u_i|\, \eta_i \hat{O}_i\, |u_i) = (u_i|\, \hat{O}_i^\dagger \eta_i\, |u_i);$$

therefore

$$\eta_i \hat{O}_i = \hat{O}_i^\dagger \eta_i, \tag{14}$$

so that the boson operator is self-adjoint with respect to the metric operator and has real expectation values. However, the boson operator $\hat{O}_i$ does not commute with the metric operator in general, so that

$$\hat{O}_i^\dagger \eta_i \neq \eta_i \hat{O}_i^\dagger \tag{15}$$

and, consequently, $\hat{O}_i^\dagger$ is not Hermitian:

$$\hat{O}_i^\dagger \neq \hat{O}_i. \tag{16}$$

So far, all that has been done is to give a new formulation of the mapping process implicitly used by Dyson. Maleev[18] gave the corresponding boson transformations for individual spin operators:

$$S_i^+ \to (2s)^{\frac{1}{2}} a_i^\dagger, \quad S_i^- \to (2s)^{\frac{1}{2}} \left(1 - \frac{a_i^\dagger a_i}{2s}\right) a_i,$$
$$S_i^z \to -S + a_i^\dagger a_i. \tag{17}$$

In a rather different approach, Dembinski[19] has used an explicit representation for the metric operator to derive the Dyson, conjugate Dyson, and Holstein–Primakoff boson mappings. These results can also be obtained with the η-formalism in a slightly different and more illuminating form as follows:

$$\langle u_i|\, O_i\, |v_i\rangle = (u_i|\, \eta_i \hat{O}_i\, |v_i). \tag{9}$$

If the metric operator η_i is commutated to the right, the conjugate Dyson mapping $\hat{O}_i^{\text{C.D.}}$ is obtained:

$$(u_i|\, \eta_i \hat{O}_i\, |v_i) = (u_i|\, (\eta_i \hat{O}_i \eta_i^{-1}) \eta_i\, |v_i). \tag{18}$$

This is related to the Dyson transformation by

$$\hat{O}_i^{\text{C.D.}} = \eta_i \hat{O}_i \eta_i^{-1}. \tag{19}$$

Similarly, the square root of the metric operator can be defined by $\eta_i^{\frac{1}{2}} \eta_i^{\frac{1}{2}} = \eta_i$, and the Holstein–Primakoff mapping obtained:

$$(u_i|\, \eta_i \hat{O}_i\, |v_i) = (u_i|\, \eta_i^{\frac{1}{2}} (\eta_i^{\frac{1}{2}} \hat{O}_i \eta_i^{-\frac{1}{2}}) \eta_i^{\frac{1}{2}}\, |v_i). \tag{20}$$

The metric is Hermitian ($\eta_i^\dagger = \eta_i$), so that the square-root operator is also Hermitian ($\eta_i^{\frac{1}{2}\dagger} = \eta_i^{\frac{1}{2}}$) and, if the operator O is Hermitian, it follows that $\eta_i^{\frac{1}{2}} \hat{O}_i \eta_i^{-\frac{1}{2}}$ is also Hermitian. The Holstein–Primakoff boson transformation is related to the Dyson transformation by

$$\hat{O}_i^{\text{H.P.}} = \eta_i^{\frac{1}{2}} \hat{O}_i \eta_i^{-\frac{1}{2}}. \tag{21}$$

A difficulty now arises in the interpretation of the Holstein–Primakoff boson operators, as the spin Hamiltonian is mapped into a Hermitian boson counterpart, but the boson state-vectors are defined by $\eta_i^{\frac{1}{2}}\, |u_i)$. By commuting the operator $\eta_i^{\frac{1}{2}}$ through to operate on the ground state, this can be written in the form

$$\eta_i^{\frac{1}{2}}\, |u_i) = (u_i!)^{-\frac{1}{2}} (\alpha_i^\dagger)^{u_i}\, |0_i), \tag{22}$$

where

$$\alpha_i^\dagger = \eta^{\frac{1}{2}} a_i^\dagger \eta^{-\frac{1}{2}} = a_i^\dagger \left(1 - \frac{a_i^\dagger a_i}{2s}\right)^{\frac{1}{2}}. \tag{23}$$

[This follows from the definition of the states $|u_i)$ and the relation $\eta_i\, |0_i) = |0_i).$] It is clear that boson operators of the type $a_i^\dagger$ do not create particles described by the state vectors $\eta_i^{\frac{1}{2}}\, |u_i)$, as the latter contain a cutoff which is characteristic of spin operators. This means that the operators $a_i^\dagger$ and a_i cannot be treated as bosons except as an approximation, so that, in a rigorous treatment, the Holstein–Primakoff formalism does not seem any simpler to use than the spin operators themselves. It has also been noted by Dyson that the Holstein–Primakoff boson transformation is highly nonlinear, due to the presence of square roots which have to be expanded.

The Dyson–Maleev transformation does give a simplification, as the metric operator acts to the left of the Hamiltonian and the latter can operate on the full boson Hilbert space.

[18] S. V. Maleev, Zh. Eksp. Teor. Fiz. **33**, 1010 (1957) [Sov. Phys.—JETP **6**, 776 (1958)].
[19] S. T. Dembinski, Physica **30**, 1217 (1964).

3. THE ANTIFERROMAGNETIC GROUND STATE

For the simple antiferromagnet described by the Hamiltonian in (1), the z components of spin on the two sublattices are oppositely directed, and it is convenient to define the ground state of spins differently, depending on the particular sublattice:

$$S_i^{a+}|0_i\rangle = 0, \quad S_j^{b-}|0_j\rangle = 0,$$
$$S_i^{az}|0_i\rangle = +S, \quad S_j^{bz}|0_j\rangle = -S. \tag{24}$$

This corresponds to a rotation of every local system of coordinates on lattice a through 180° about the x axis, i.e.,

$$S_i^{a+} \leftrightarrow S_i^{a-}, \quad S_i^{az} \rightarrow -S_i^{az}. \tag{25}$$

Application of this rotation to the Dyson–Maleev transformation leads to the following results:

$$\begin{aligned} S_i^{a+} &\rightarrow (2s)^{\frac{1}{2}}\left(1 - \frac{a_i^+ a_i}{2s}\right)a_i, \\ S_j^{b+} &\rightarrow (2s)^{\frac{1}{2}} b_j^+, \\ S_i^{a-} &\rightarrow (2s)^{\frac{1}{2}} a_i^+, \\ S_j^{b-} &\rightarrow (2s)^{\frac{1}{2}}\left(1 - \frac{b_j^+ b_j}{2s}\right)b_j, \\ S_i^{az} &\rightarrow S - a_i^+ a_i, \\ S_j^{bz} &\rightarrow -S + b_j^+ b_j. \end{aligned} \tag{26}$$

If these transformations are substituted in the spin Hamiltonian and terms higher than bilinear rejected, the usual free spin-wave Hamiltonian results:

$$H_0 = E_0 + E_1\left(\sum_i a_i^+ a_i + \sum_j b_j^+ b_j\right) + E_2 \sum_{i\delta} (a_i^+ b_{i+\delta}^+ + a_i b_{i+\delta}), \tag{27}$$

where

$$E_0 = -JS^2NZ - 2NAS, \quad E_1 = JSZ + A, \quad E_2 = JS. \tag{28}$$

z is the number of near-neighbor sites and N is the number of spins on one sublattice.

The boson operators can be Fourier-transformed in the usual way:

$$\begin{aligned} a_i^+ &= N^{-\frac{1}{2}} \sum_k e^{ik\cdot R_i} a_k^+, \quad a_i = N^{-\frac{1}{2}} \sum_k e^{-ik\cdot R_i} a_k, \\ b_j^+ &= N^{-\frac{1}{2}} \sum_k e^{ik\cdot R_j} b_k^+, \quad b_j = N^{-\frac{1}{2}} \sum_k e^{-ik\cdot R_j} b_k, \end{aligned} \tag{29}$$

and after substitution of these definitions, the free spin-wave Hamiltonian can be written in the following form:

$$H_0 = E_0 + E_1 \sum_k (a_k^+ a_k + b_k^+ b_k) + E_2 \sum_k \gamma_k (a_k^+ b_{-k}^+ + a_k b_{-k}), \tag{30}$$

where

$$\gamma_k = \sum_\delta e^{ik\cdot\delta}. \tag{31}$$

If the crystal has inversion symmetry, the structure factor is real:

$$\gamma_k = \gamma_{-k} = \gamma_k^*. \tag{32}$$

The bilinear Hamiltonian can be diagonalized with a canonical transformation, leading to the result

$$H_0 = \bar{E}_0 + \sum_k \lambda_k (\alpha_k^+ \alpha_k + \beta_k^+ \beta_k), \tag{33}$$

where magnon operators are defined by

$$\begin{aligned} \alpha_k &= u_k a_k - v_k b_{-k}^+, \\ \beta_k &= u_k b_k - v_k a_{-k}^+, \end{aligned} \tag{34}$$

and the canonical constraints are

$$\begin{aligned} [\alpha_k, \alpha_{k'}^+] &= [\beta_k, \beta_{k'}^+] = \delta_{kk'}, \\ [\alpha_k, \beta_{-k}] &= [\beta_k, \alpha_{-k}] = 0. \end{aligned} \tag{35}$$

These lead to the canonical condition

$$|u_k|^2 - |v_k|^2 = 1. \tag{36}$$

The solutions for the parameters are

$$\lambda_k = (E_1^2 - E_2^2\gamma_k^2)^{\frac{1}{2}}, \quad \bar{E}_0 = E_0 - NE_1 + \sum_k \lambda_k, \tag{37}$$

where

$$u_k = \pm\left(\frac{1+\epsilon_k}{2\epsilon_k}\right)^{\frac{1}{2}}, \quad v_k = \mp\left(\frac{1-\epsilon_k}{2\epsilon_k}\right)^{\frac{1}{2}},$$

where

$$\epsilon_k = \lambda_k/E_1 = [1 - (E_2^2/E_1^2)\gamma_k^2]^{\frac{1}{2}}. \tag{38}$$

It is interesting to note that the ground-state energy $\bar{E}_0$ contains the zero-point energy of the magnon oscillators and, in fact, the corresponding zero-point motion causes the ground state to have a very complicated structure, making rigorous treatments extremely difficult.

It is useful to construct an exponential representation for the canonical transformation to free spin waves. This transformation, which maps spin-deviation operators, e.g., a_k, into magnon operators, e.g., α_k, has the following properties:

$$Ua_kU^{-1} = \alpha_k, \quad Ub_kU^{-1} = \beta_k, \tag{39}$$

$$\begin{aligned} U^{-1}\alpha_k U &= u_k\alpha_k + v_k\beta_{-k}^+, \\ U^{-1}\beta_k U &= u_k\beta_k + v_k\alpha_{-k}^+. \end{aligned} \tag{40}$$

The expressions for spin-wave operators (34) can be inverted to give

$$\begin{aligned} a_k &= u_k\alpha_k + v_k\beta_{-k}^+, \quad b_k = u_k\beta_k + v_k\alpha_{-k}^+, \\ a_k^+ &= u_k\alpha_k^+ + v_k\beta_{-k}, \quad b_k^+ = u_k\beta_k^+ + v_k\alpha_{-k}, \end{aligned} \tag{41}$$

and it can be seen that (40) follows directly from the definitions (34) and (39). The transformation operator can be written in the form

$$U = e^{-iS}, \tag{42}$$

where

$$S^+ = S \tag{43}$$

and the solution for the generator is obtained from the expansion

$$e^{-iS}a_k e^{iS} = a_k + i[a_k, S] + \frac{i^2}{2!}[[a_k, S], S] + \cdots \tag{44}$$

as follows:

$$S = i\sum_k \theta_k(a_k^+ b_{-k}^+ - a_k b_{-k}), \tag{45}$$

where θ_k is related to the canonical coefficients by

$$u_k = \cosh\theta_k, \quad v_k = \sinh\theta_k. \tag{46}$$

It can easily be shown that

$$\alpha_k^+\beta_{-k}^+ - \alpha_k\beta_{-k} = a_k^+ b_{-k}^+ - a_k b_{-k}, \tag{47}$$

so that the generator S can be conveniently expressed in terms of magnon operators. It should be noted that the canonical transformation to free spin waves is closely related to the Bogoliubov transformation in the theory of superfluid bosons, and the exponential representation of the latter transformation is well known.[20,21]

The ground state of the magnon oscillators can now be related to the Néel state. The magnon ground state is defined by

$$\alpha_k|\tilde{0}) = \beta_k|\tilde{0}) = 0 \tag{48}$$

and the Néel state by

$$a_k|0) = b_k|0) = 0. \tag{49}$$

If the explicit transformation operator is inserted, the magnon ground state can be written in the form

$$|\tilde{0}) = U|0) = \exp\left[\sum_k \theta_k(a_k^+ b_{-k}^+ - a_k b_{-k})\right]|0) \tag{50}$$

and, when this expression is expanded in powers of the operators a_k^+ and b_{-k}^+, it is clear that it has a projection onto the nonphysical Hilbert space, since terms like $(a_k^+)^n(b_{-k}^+)^n|0)$ exist and n can tend to infinity. Thus the canonical transformation leads to operators whose state vectors have nonphysical projections and this, in fact, causes a large kinematic interaction, even at zero temperature.

[20] M. Girardeau and R. Arnowitt, Phys. Rev. **113**, 755 (1959).
[21] E. P. Gross, Ann. Phys. (N.Y.) **9**, 292 (1960).

By construction, the mapping procedure described earlier produces operators in the boson space which do not have matrix elements from the nonphysical subspace into the physical subspace, and a selection rule can be written in the general form

$$P\hat{O} = P\hat{O}P, \tag{51}$$

where P denotes a projection operator with eigenvalues of zero on the nonphysical subspace. This selection rule has the very important consequence that the full Hamiltonian can never mix in the nonphysical with the physical state-vectors. Dyson[2] showed that if an eigenstate of $\hat{O}$ has a physical projection, then its eigenvalue is physical and corresponds to the projection onto the physical subspace mapped back into spin space. Hence, it is clear that if an eigenstate is found in the full boson space, then its projection onto the physical subspace is the part which has a correspondent in spin space, and the vector should be projected before estimates of physical quantities are made.

The magnon ground state is an eigenstate of the free spin-wave Hamiltonian, but it is not an eigenstate of the full boson Hamiltonian, as the latter has source terms in the interactions which create magnons when operating on the magnon ground state. In addition, the bilinear Hamiltonian does not obey the selection rule (51). However, it is reasonable to take the magnon ground state as a first approximation to the antiferromagnetic ground state, and use perturbation theory to switch the magnon ground state into the ground state of the full boson Hamiltonian.

Due to the nonphysical projection, the magnon ground state overestimates the number of physical spin deviations, and it should therefore be projected before the reduction in sublattice magnetization from the Néel state is calculated. This reduction is given by the expression

$$\frac{\langle\tilde{0}|(s - S_i^{az})|\tilde{0}\rangle}{\langle\tilde{0}\,|\,\tilde{0}\rangle} = \frac{(\tilde{0}|\eta a_i^+ a_i|\tilde{0})}{(\tilde{0}|\eta|\tilde{0})}, \tag{52}$$

where

$$\eta = \prod_i \eta_i. \tag{53}$$

For simplicity, it is convenient to restrict considerations to spin $\frac{1}{2}$ when the metric operator reduces to a theta function:

$$\eta_i = \theta(2s - a_i^+ a_i), \quad \text{for} \quad s = \tfrac{1}{2}, \tag{54}$$

where

$$\theta(x) = 1, \quad x \geq 0,$$
$$= 0, \quad x < 0. \tag{55}$$

Tao Yuin[9] has given an operator representation for the theta function as follows:

$$\theta(2s - a_i^+ a_i) = \sum_{l=0}^{\infty} B_l (a_i^+)^l (a_i)^l,$$

$$a_i^+ a_i \theta(2s - a_i^+ a_i) = \sum_{l=1}^{\infty} C_l (a_i^+)^l (a_i)^l, \tag{56}$$

where

$$B_l = 1, \quad l = 0, \quad C_l = 1, \quad l = 1,$$

$$B_l = 0, \quad l \leq 2s, \quad C_l = 0, \quad l \leq 2s,$$

$$B_l = \frac{(-1)^{l-2s}(2s+1)(2s+2)\cdots(l-1)}{l!\,(l-2s-1)!}, \quad l > 2s,$$

$$C_l = \frac{(-1)^{l-2s} 2s(2s+1)\cdots(l-2)}{(l-1)!\,(l-2s-1)!}, \quad l > 2s. \tag{57}$$

To estimate the expression for the average number operator (52), it is only necessary to project onto a single site i:

$$\frac{\langle \tilde{0}|\, s - S_i^{az} \,|\tilde{0}\rangle}{\langle \tilde{0} \,|\, \tilde{0}\rangle} \simeq \frac{(\tilde{0}|\, \eta_i a_i^+ a_i \,|\tilde{0})}{(\tilde{0}|\, \eta_i \,|\tilde{0})}. \tag{58}$$

For the purpose of evaluating projected averages, it is convenient to introduce some new notation. From (56),

$$\eta_i = \sum_{l=0}^{\infty} B_l (a_i^+)^l (a_i)^l,$$

so that it is necessary to consider the expression $(\tilde{0}|\, (a_i^+)^l (a_i)^l \,|\tilde{0})$ and, since all the complications of the magnon ground state are contained in the unitary operator U, this can be written as

$$(\tilde{0}|\, (a_i^+)^l (a_i)^l \,|\tilde{0}) = (0|\, U^+ (a_i^+)^l (a_i)^l U \,|0). \tag{59}$$

From (39) it follows that

$$U^+ a_k^+ U = u_k a_k^+ + v_k b_{-k}, \tag{60}$$

and using the definition of Fourier transformation, the following operator can be defined:

$$U^+ a_i^+ U = \mu_i^+ = N^{-\frac{1}{2}} \sum_k e^{-ik\cdot R_i} (u_k a_k^+ + v_k b_{-k}). \tag{61}$$

Hence

$$(\tilde{0}|\, (a_i^+)^l (a_i)^l \,|\tilde{0}) = (0|\, (\mu_i^+)^l (\mu_i)^l \,|0), \tag{62}$$

and, using (60), this becomes

$$(0|\, (\mu_i^+)^l (\mu_i)^l \,|0) = (0|\left[N^{-\frac{1}{2}} \sum_k e^{-ik\cdot R_i} v_k b_{-k} \right]^l \times \left[N^{-\frac{1}{2}} \sum_k e^{ik\cdot R_i} v_k b_{-k}^+ \right]^l |0). \tag{63}$$

Wick's theorem can be used to evaluate this expression with the following result:

$$(\tilde{0}|\, (a_i^+)^l (a_i)^l \,|\tilde{0}) = l!\, v^l, \tag{64}$$

where

$$v = N^{-1} \sum_k v_k^2. \tag{65}$$

The ground-state expectation value can now be written in the form

$$\frac{(\tilde{0}|\, \theta_i a_i^+ a_i \,|\tilde{0})}{(\tilde{0}|\, \theta_i \,|\tilde{0})} = \frac{\sum_{l=1}^{\infty} C_l l!\, v^l}{\sum_{l=0}^{\infty} B_l l!\, v^l}, \tag{66}$$

where

$$\theta_i = \theta(2s - a_i^+ a_i), \tag{67}$$

and for spin $\frac{1}{2}$ the summations are easily evaluated:

$$\sum_{l=1}^{\infty} C_l l!\, v^l = v(1+v)^{-2},$$

$$B_0 + \sum_{l=2}^{\infty} B_l l!\, v^l = (1+2v)(1+v)^{-2}. \tag{68}$$

Hence, the reduction in sublattice magnetization due to the zero-point motion of the magnon oscillators is given by

$$\frac{(\tilde{0}|\, \theta_i a_i^+ a_i \,|\tilde{0})}{(\tilde{0}|\, \theta_i \,|\tilde{0})} = \frac{v}{1+2v}. \tag{69}$$

Anderson[13] was the first to estimate the reduction in sublattice magnetization from the Néel state, but his method is equivalent to using the unprojected magnon ground state:

$$(\tilde{0}|\, a_i^+ a_i \,|\tilde{0}) = (0|\, \mu_i^+ \mu_i \,|0) = v. \tag{70}$$

Anderson calculated $v = 0.078$, and if this value is substituted into the expression obtained here, using the projected ground state, a correction to his value results:

$$v/(1+2v) = 0.0675. \tag{71}$$

It is interesting to compare this result with other work which is not based on spin-wave theory, and the calculation of Davis[4] will be taken as representative. Davis used Schwinger's coupled boson representation of the spin operators to keep the theory rigorously in the physical domain, and applied perturbation theory to switch the Néel state into the antiferromagnetic ground state. Comparison of these results (Table I) seems to show that the spin-wave

TABLE I. Values for $\langle a_i^+ a_i \rangle$ in the antiferromagnetic ground state for spin $\frac{1}{2}$ and zero anisotropy.

unprojected magnon ground state (Anderson)	$\langle a_i^+ a_i \rangle = 0.078$
projected magnon ground state	$\langle a_i^+ a_i \rangle = 0.0675$
perturbation theory (Davis)	$\langle a_i^+ a_i \rangle = 0.0637$

picture of the ground state may be better than Anderson's results indicate. The magnon ground state has been projected for $s = \frac{1}{2}$, but it is possible to expand the metric operator for general spin in terms of theta functions to obtain more general results.[22]

4. KINEMATIC INTERACTION AT FINITE TEMPERATURE

At finite temperatures it is necessary to consider the magnon state vectors. These can be defined in the usual way as follows:

$$|\tilde{\alpha}\tilde{\beta}) = \prod_{kk'} (n_k!)^{-\frac{1}{2}}(\alpha_k^+)^{n_k}(n_{k'}!)^{-\frac{1}{2}}(\beta_{k'}^+)^{n_{k'}} |\tilde{0}). \quad (72)$$

Using the properties of the U operator, this becomes

$$|\tilde{\alpha}\tilde{\beta}) = U\,|ab), \quad (73)$$

where

$$|ab) = \prod_{kk'} (n_k!)^{-\frac{1}{2}}(a_k^+)^{n_k}(n_{k'}!)^{-\frac{1}{2}}(b_{k'}^+)^{n_{k'}}|\,0), \quad (74)$$

and it can be seen that the antiferromagnetic spin-wave state vectors are obtained by a unitary transformation on ferromagnetic-type spin-wave states.

When a transformation is made in a normed linear space, in general, the metric has to be transformed in order to preserve norms.[23] A unitary transformation O satisfies the relation

$$O^\dagger O = 1, \quad (75)$$

and the correspondents in the physical subspace of the boson space satisfy

$$\eta^{-1}\hat{O}^\dagger\eta\hat{O} = 1. \quad (76)$$

If the transformation operator does not commute with the metric operator, then

$$\hat{O}^\dagger\hat{O} \neq 1 \quad (77)$$

and the transformation operator is no longer unitary. In this work a unitary transformation U has been defined on the full boson space, but it is not unitary in the physical subspace, i.e.,

$$\eta^{-1}U^\dagger\eta U \neq 1, \quad (78)$$

so that the transformation to free spin-wave operators corresponds to a nonunitary transformation in the physical spin space. The nonunitary nature of the Anderson canonical transformation can also be seen from the fact that magnon state vectors are normalized on the full boson Hilbert space and so cannot be simultaneously normalized on the physical subspace. It is precisely this nonphysical projection which leads to the large kinematic interaction at low temperatures found by recent authors.[7,9] It is interesting to note that the process of Fourier transformation is also nonunitary on the physical subspace, and this causes ferromagnetic spin-wave state vectors to have a nonphysical projection. However, in this case, the nonphysical projection is small and vanishes in the ground state.

In order to allow for the kinematic interaction at finite temperature, it is necessary to evaluate expressions of the type

$$\frac{(\tilde{\alpha}\tilde{\beta}|\, a_i^+a_i\theta(2s - a_i^+a_i)\,|\tilde{\alpha}\tilde{\beta})}{(\tilde{\alpha}\tilde{\beta}|\,\theta(2s - a_i^+a_i)\,|\tilde{\alpha}\tilde{\beta})}. \quad (79)$$

The operator representation for the theta function (56) can be used, and it is convenient to consider the expression

$$(\tilde{\alpha}\tilde{\beta}|\,(a_i^+)^l(a_i)^l\,|\tilde{\alpha}\tilde{\beta}) = (ab|\,(\mu_i^+)^l(\mu_i)^l\,|ab). \quad (80)$$

To evaluate this, the operators a_k, b_k used in the definition of $(ab|$ can be commuted to the right, when the expression $(ab|\,(\mu_i^+)^l(\mu_i)^l\,|ab)$ can be written as a power series in the magnon occupation numbers. For simplicity, terms higher than linear in the number operators are neglected here, and the following expression is obtained:

$$(\tilde{\alpha}\tilde{\beta}|\, a_i^+a_i\theta(2s - a_i^+a_i)\,|\tilde{\alpha}\tilde{\beta}) = v(1+v)^{-2} + \sum_{l=1}^{\infty} C_l v^{l-1} l\cdot l!\,\frac{1}{N} \times \sum_k (n_k^\alpha u_k^2 + n_k^\beta v_k^2) + \cdots. \quad (81)$$

In the perturbation approach, the free-magnon Hamiltonian is taken to describe an unperturbed system, and the remainder of the boson Hamiltonian is regarded as interaction and is used to renormalize the energy. If this procedure is adopted, then the magnon states are regarded as the best eigenstates of the full Hamiltonian, so that for calculating thermal averages of spin deviations the following expression should be used:

$$\overline{a_i^+a_i} = \left[\sum_{\{\alpha\beta\}} e^{-\beta\lambda_{\alpha\beta}}\right]^{-1} \sum_{\{\alpha\beta\}} e^{-\beta\lambda_{\alpha\beta}} \frac{(\tilde{\alpha}\tilde{\beta}|\,\eta a_i^+a_i\,|\tilde{\alpha}\tilde{\beta})}{(\tilde{\alpha}\tilde{\beta}|\,\eta\,|\tilde{\alpha}\tilde{\beta})}. \quad (82)$$

The summation is over all magnon state vectors corresponding to eigenstates having a nonzero physical projection and hence a physical correspondent in spin space, and $\lambda_{\alpha\beta}$ is the renormalized magnon energy corresponding to the state $|\alpha\beta)$. When terms depending on the square of the magnon number-operators are

[22] D. C. Herbert, Ph.D. thesis, University of London, 1968.
[23] R. L. Mills and R. P. Kenan, Ann. Phys. (N.Y.) 37, 104 (1966).

neglected, the magnon state vectors can be projected as follows:

$$\frac{(\tilde{\alpha}\tilde{\beta}|\,\eta a_i^+ a_i\,|\tilde{\alpha}\tilde{\beta})}{(\tilde{\alpha}\tilde{\beta}|\,\eta\,|\tilde{\alpha}\tilde{\beta})} \simeq \frac{\sum_{l=1}^{\infty} C_l v^l l! + \sum_{l=1}^{\infty} C_l v^{l-1} l \cdot l! \frac{1}{N}\sum_k (n_k^\alpha u_k^2 + n_k^\beta v_k^2)}{\sum_{l=0}^{\infty} B_l v^l l! + \sum_{l=1}^{\infty} B_l v^{l-1} l \cdot l! \frac{1}{N}\sum_k (n_k^\alpha u_k^2 + n_k^\beta v_k^2)}. \quad (83)$$

The summations are easily performed:

$$\begin{aligned} \sum_{l=1}^{\infty} C_l v^l l! &= v(1+v)^{-2}, \\ \sum_{l=1}^{\infty} C_l v^{l-1} l \cdot l! &= (1-v)(1+v)^{-3}, \\ B_0 + \sum_{l=1}^{\infty} B_l v^l l! &= (1+2v)(1+v)^{-2}, \\ \sum_{l=1}^{\infty} B_l v^{l-1} l \cdot l! &= -2v(1+v)^{-3}. \end{aligned} \quad (84)$$

The term $(1/N)\sum_k (n_k^\alpha u_k^2 + n_k^\beta v_k^2)$ is just the correction to the ground-state occupation number obtained from the unprojected state vector, and it is convenient to introduce the definitions

$$\overline{\langle a_i^+ a_i\rangle}^{\alpha\beta} = N^{-1}\sum_k (n_k^\alpha u_k^2 + n_k^\beta v_k^2), \quad (85)$$

$$\langle a_i^+ a_i\rangle^{\alpha\beta} = \frac{(\tilde{\alpha}\tilde{\beta}|\,\eta a_i^+ a_i\,|\tilde{\alpha}\tilde{\beta})}{(\tilde{\alpha}\tilde{\beta}|\,\eta\,|\tilde{\alpha}\tilde{\beta})}. \quad (86)$$

The projected expectation value can then be written as

$$\langle a_i^+ a_i\rangle^{\alpha\beta} \simeq \frac{v(1+v)^{-2} + (1-v)(1+v)^{-3}\overline{\langle a_i^+ a_i\rangle}^{\alpha\beta}}{(1+2v)(1+v)^{-2} - 2v(1+v)^{-3}\overline{\langle a_i^+ a_i\rangle}^{\alpha\beta}}. \quad (87)$$

This expression can be expanded into a power series in $\overline{\langle a_i^+ a_i\rangle}^{\alpha\beta}$:

$$\langle a_i^+ a_i\rangle^{\alpha\beta} \simeq v(1+2v)^{-1} + (1+2v)^{-2}\overline{\langle a_i^+ a_i\rangle}^{\alpha\beta} + \cdots, \quad (88)$$

and the result for the thermal average becomes

$$\overline{a_i^+ a_i} = v(1+2v)^{-1} + (1+2v)^{-2}\overline{\langle a_i^+ a_i\rangle} + \cdots, \quad (89)$$

where $\overline{\langle a_i^+ a_i\rangle}$ is a correction to the ground-state values obtained using unprojected state vectors in the thermal trace, but still retaining a cutoff, i.e., nonphysical states having zero physical projection are excluded from the thermal trace.

It is interesting to compare this result with that of Tao Yuin, who gives the expression

$$\langle S_f^z\rangle = \tfrac{1}{2} - \langle b_f^+ b_f\rangle_0/(1 + 2\langle b_f^+ b_f\rangle_0), \quad (90)$$

where $b_f^+ b_f$ is the spin deviation number operator for lattice site "f" and $\langle\cdots\rangle_0$ denotes a thermal average over the full boson space using the non-interacting spin-wave Hamiltonian and no cutoff on the trace. If $\langle\cdots\rangle_0$ is expanded into a series of temperature-dependent terms,

$$\langle b_f^+ b_f\rangle_0 = v + a(T) + \cdots, \quad (91)$$

then

$$\frac{\langle b_f^+ b_f\rangle_0}{1 + 2\langle b_f^+ b_f\rangle_0} = v(1+2v)^{-1} + (1+2v)^{-2}a(T) + \cdots. \quad (92)$$

This is very similar to the expansion obtained in this work, but $a(T)$ is obtained from a trace without cutoff, whereas $\overline{\langle a_i^+ a_i\rangle}$ is obtained from a trace with cutoff. Tao Yuin claims that his method is rigorous, but it is not clear that this is so, as he does not normalize the state vectors on the physical subspace. However, at zero temperature the thermal trace reduces to a ground-state average and normalizations cancel. In this case, Tao Yuin's result reduces to the expression obtained in this work.

It must be emphasized that the nonphysical projection of the state vectors is only a part of the total kinematic interaction; the other part comes from the effect of nonphysical states, i.e., states with zero physical projection. The results of this work show that the kinematic interaction at low temperature obtained by Tao Yuin and Kenan can be explained in terms of a nonphysical projection of the magnon state vectors. The states with zero physical projection introduce a cutoff on the thermal trace (82), and also affect the renormalized energies $\lambda_{\alpha\beta}$.

5. CONCLUSION

The η-formalism of quantum field theory has been used to clarify the mathematical nature of the various boson mappings used in the theory of magnetism, and it is shown that the Holstein–Primakoff transformation has some formal disadvantages which make it unsuitable for mathematical investigation. By constructing an exponential representation for the Anderson canonical transformation to magnon operators, it was possible to relate the magnon ground state to the Néel state, and it was shown that the magnon ground state has a large nonphysical projection. After projecting the magnon ground state on to the physical subspace, close agreement with the results of perturbation theories for the reduction in sublattice magnetization from the Néel state was obtained, indicating that spin-wave theory is better than might

be expected from Anderson's work.[13] It is possible to improve spin-wave theory still further, and investigations are in progress in which the coefficients in the canonical transformation are chosen to minimize the number of magnons in the true ground state. The main conclusion of this paper is that the large kinematic interaction found by recent authors at low temperatures can be understood as a nonphysical projection of the state vectors. This effect can be allowed for, and is relatively innocuous, as the eigenvalues of states with nonzero physical projection are physical. It is, therefore, still possible that the nonphysical eigenvalues may cancel out in the antiferromagnet, as they do in the ferromagnet at low temperature.

ACKNOWLEDGMENTS

The author wishes to thank Dr. S. Doniach for supervising this work. He is also indebted to the Science Research Council for the award of a research studentship.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 12 DECEMBER 1969

Statistical Average of Product of Phase Sums Arising in the Study of Disordered Lattices. I

PRIYAMVADA SAH
Department of Physics, Banaras Hindu University, Varanasi, India

(Received 27 January 1969)

A recurrence formula is established to evaluate the statistical average $\langle S(k_1)S(k_2)\cdots S(k_n)\rangle$ in the limit of $N\to\infty$, where $S(k)=\sum_{j=1}^{N}e^{ikx_j}$, $x_1, x_2, \cdots, x_N$, denotes the position of atoms in a disordered lattice, under the condition that interatomic distances are statistically independent and have the same probability distribution $s(r)$.

1. INTRODUCTION

If we consider the one-electron levels in a disordered array of atoms, the electronic energy levels depend on the position of atoms as parameters. The perturbation expansion of energy in certain approximations would involve the product of phase sums[1] and hence the formulas derived here are useful in evaluating the terms and may be of use in other connections also. Let $x_1, x_2, \cdots, x_N$ denote the positions of atoms from one end and $r_1, r_2, \cdots, r_N$ the interatomic distances. We define

$$S(k)=\sum_{j=1}^{N}e^{ikx_j}$$
$$=e^{ikr_1}+e^{ik(r_1+r_2)}+\cdots+e^{ik(r_1+r_2+\cdots+r_N)}. \quad (1)$$

For any n,

$$S(k_1)S(k_2)\cdots S(k_n)$$
$$=\sum_{K_1,K_2,\cdots,K_n}\exp\left[i(K_1r_1+\cdots+K_nr_n)\right], \quad (2)$$

where

$$K_1=k_1+k_2+\cdots+k_n, \quad (3)$$

and $K_2, \cdots, K_n$ are sums of any subset of $k_1, k_2, \cdots, k_n$ including the full set and empty set in such a way that the terms in $K_{\nu+1}$ are subset of terms in K_ν ($\nu=1, 2, \cdots, N-1$). This restriction in forming the sum appearing in Eq. (2) will be called condition I. Denoting the left-hand side of (2) by $\Xi(N)$, we have

$$\Xi(N)=S(k_1)S(k_2)\cdots S(k_n). \quad (4)$$

The problem is to determine

$$\langle\Xi\rangle=\lim_{N\to\infty}\langle\Xi(N)\rangle.$$

We shall first derive $\langle\Xi(N)\rangle$, where N can be any number greater than n, and finally proceed to the limit as $N\to\infty$. This result is established under the assumption that none of $k_1, k_2, \cdots, k_n$ or their partial sums are zero. If any partial sum vanishes, the procedure has to be modified and will be considered later. These are needed in the complete discussion of perturbation terms.

In all terms of the right-hand side of Eq. (1) K_1 is fixed and given by Eq. (3), but $K_2, \cdots, K_n$ vary subject to condition I. Let us write

$$\Xi(N)=\Xi^{(1)}(N)+\Xi^{(2)}(N)+\cdots+\Xi^{n}(N), \quad (5)$$

so that, in every term of $\Xi^{(1)}(N)$, K_ν is either equal to K_1 or 0. In $\Xi^{(2)}(N)$, at least one of the $K_\nu=K_1^{(1)}$, where $0\subset K_1^{(1)}\subset K_1$. In $\Xi^{(j)}(N)$, K_ν assumes j distinct values $K_1, K_1^{(1)}, \cdots, K_1^{(j-1)}$. Every term in the rhs of Eq. (5)

[1] P. Sah, Ph.D. thesis, University of London, 1959.

be expected from Anderson's work.[13] It is possible to improve spin-wave theory still further, and investigations are in progress in which the coefficients in the canonical transformation are chosen to minimize the number of magnons in the true ground state. The main conclusion of this paper is that the large kinematic interaction found by recent authors at low temperatures can be understood as a nonphysical projection of the state vectors. This effect can be allowed for, and is relatively innocuous, as the eigenvalues of states with nonzero physical projection are physical. It is, therefore, still possible that the nonphysical eigenvalues may cancel out in the antiferromagnet, as they do in the ferromagnet at low temperature.

ACKNOWLEDGMENTS

The author wishes to thank Dr. S. Doniach for supervising this work. He is also indebted to the Science Research Council for the award of a research studentship.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 12 DECEMBER 1969

Statistical Average of Product of Phase Sums Arising in the Study of Disordered Lattices. I

PRIYAMVADA SAH
Department of Physics, Banaras Hindu University, Varanasi, India

(Received 27 January 1969)

A recurrence formula is established to evaluate the statistical average $\langle S(k_1)S(k_2)\cdots S(k_n)\rangle$ in the limit of $N\to\infty$, where $S(k)=\sum_{j=1}^{N}e^{ikx_j}$, $x_1, x_2, \cdots, x_N$, denotes the position of atoms in a disordered lattice, under the condition that interatomic distances are statistically independent and have the same probability distribution $s(r)$.

1. INTRODUCTION

If we consider the one-electron levels in a disordered array of atoms, the electronic energy levels depend on the position of atoms as parameters. The perturbation expansion of energy in certain approximations would involve the product of phase sums[1] and hence the formulas derived here are useful in evaluating the terms and may be of use in other connections also. Let $x_1, x_2, \cdots, x_N$ denote the positions of atoms from one end and $r_1, r_2, \cdots, r_N$ the interatomic distances. We define

$$S(k)=\sum_{j=1}^{N}e^{ikx_j}$$
$$=e^{ikr_1}+e^{ik(r_1+r_2)}+\cdots+e^{ik(r_1+r_2+\cdots+r_N)}. \quad (1)$$

For any n,

$$S(k_1)S(k_2)\cdots S(k_n)$$
$$=\sum_{K_1,K_2,\cdots,K_n}\exp\left[i(K_1r_1+\cdots+K_nr_n)\right], \quad (2)$$

where

$$K_1=k_1+k_2+\cdots+k_n, \quad (3)$$

and $K_2, \cdots, K_n$ are sums of any subset of $k_1, k_2, \cdots, k_n$ including the full set and empty set in such a way that the terms in $K_{\nu+1}$ are subset of terms in K_ν ($\nu=1, 2, \cdots, N-1$). This restriction in forming the sum appearing in Eq. (2) will be called condition I. Denoting the left-hand side of (2) by $\Xi(N)$, we have

$$\Xi(N)=S(k_1)S(k_2)\cdots S(k_n). \quad (4)$$

The problem is to determine

$$\langle\Xi\rangle=\lim_{N\to\infty}\langle\Xi(N)\rangle.$$

We shall first derive $\langle\Xi(N)\rangle$, where N can be any number greater than n, and finally proceed to the limit as $N\to\infty$. This result is established under the assumption that none of $k_1, k_2, \cdots, k_n$ or their partial sums are zero. If any partial sum vanishes, the procedure has to be modified and will be considered later. These are needed in the complete discussion of perturbation terms.

In all terms of the right-hand side of Eq. (1) K_1 is fixed and given by Eq. (3), but $K_2, \cdots, K_n$ vary subject to condition I. Let us write

$$\Xi(N)=\Xi^{(1)}(N)+\Xi^{(2)}(N)+\cdots+\Xi^{n}(N), \quad (5)$$

so that, in every term of $\Xi^{(1)}(N)$, K_ν is either equal to K_1 or 0. In $\Xi^{(2)}(N)$, at least one of the $K_\nu=K_1^{(1)}$, where $0\subset K_1^{(1)}\subset K_1$. In $\Xi^{(j)}(N)$, K_ν assumes j distinct values $K_1, K_1^{(1)}, \cdots, K_1^{(j-1)}$. Every term in the rhs of Eq. (5)

[1] P. Sah, Ph.D. thesis, University of London, 1959.

is further subdivided:

$$\Xi^{(j)}(N) = \Xi_1^{(j)}(N) + \cdots + \Xi_{N-j+1}^{(j)}(N), \quad j = 1, 2, \cdots, n, \quad (6)$$

where the subscript i in $\Xi_i^{(j)}(N)$ gives the number of K's which are equal to K_1. The highest possible subscript is $N - j + 1$ because, if the term has j distinct K's in it, the number of K's equal to K_1 cannot exceed $N - j + 1$.

2. DETERMINATION OF $\langle \Xi^{(j)}(N) \rangle$

$\langle \Xi^{(j)}(N) \rangle$ is now determined by induction. We assert that, for any $N \geq j$,

$$\langle \Xi^{(j)}(N) \rangle = P_j(K_1, K_1^{(1)}, \cdots, K_1^{(j-1)}) + Q_j(K_1, K_1^{(1)}, \cdots, K_1^{(j-1)}, N), \quad (7)$$

where P_j and Q_j are defined as follows:

$$P_j = \sum_{K_1, K_1^{(1)}, \cdots, K_1^{j-1}} \frac{f(K_1)}{1 - f(K_1)} \times \frac{f(K_1^{(1)})}{1 - f(K_1^{(1)})} \cdots \frac{f(K_1^{(j-1)})}{1 - f(K_1^{(j-1)})}, \quad (8)$$

where K_1 is given by Eq. (3) and $K_1^{(1)}, \cdots, K_1^{(j-1)}$ are all distinct. $K_1^{(i)}$, $i = 1, \cdots, (j-1)$, are partial sums of $k_1, k_2, \cdots, k_n$ in such a way that the terms of $K_1^{(i)} \subset K_1^{(i-1)}$, $i = 2, 3, \cdots, j-1$. This will be referred to as condition II. $f(K_1)$ is the Fourier transform of the probability distribution $s(r)$, i.e.,

$$f(K_1) = \int e^{iK_1 r} s(r)\, dr. \quad (9)$$

$s(r)$ is normalized and, therefore, $|f(K_1)| < 1$. The functions

$$Q_j(K_1, K_1^{(1)}, \cdots, K_1^{(j-1)}, N) = \sum_{K_1, \cdots, K_1^{(j-1)}} C_0^j f^N(K_1) + C_1^{(j)} f^N(K_1^{(1)}) + \cdots + C_{j-1}^{(j)} f^N(K_1^{(j-1)}), \quad (10)$$

where $C_i^{(j)}$, $i = (0, j-1)$, are some functions of $f(K_1), f(K_1^{(1)}), \cdots, f(K_1^{(j-1)})$, but do not depend on N. In writing $\langle \Xi^{(j)}(N) \rangle$ as $P_j + Q_j$, we have separated the parts which are dependent and independent of N.

To establish the basis for induction we show that the assertion is true for $j = 1$ and any $N \geq 1$. By definition,

$$\Xi^{(1)}(N) = e^{iK_1 r_1} + e^{iK_1(r_1+r_2)} + \cdots + e^{iK_1(r_1+r_2+\cdots+r_N)}. \quad (11)$$

The terms on the right are $\Xi_1^{(1)}, \Xi_2^{(1)}, \cdots, \Xi_N^{(1)}(N)$, respectively.

Since $r_1, r_2, \cdots, r_N$ are statistically independent, on averaging both sides of Eq. (11) we have, therefore,

$$\begin{aligned} \langle \Xi^{(1)}(N) \rangle &= f(K_1) + f^2(K_1) + \cdots + f^N(K_1) \\ &= \frac{f(K_1)}{1 - f(K_1)} - \frac{f(K_1)}{1 - f(K_1)} \cdot f^N(K_1) \\ &= P_1(K_1) + Q_1(K_1, N). \end{aligned}$$

P_1 and Q_1 have the postulated forms. We now show that the Eqs. (7), (8), and (10), which have been assumed for j, also hold for $j + 1$.

3. CORRESPONDENCE BETWEEN $\Xi^{(j)}(N - \mu)$ AND $\Xi_\mu^{(j+1)}(N)$

The nature of this correspondence can be brought out as follows. The terms of $\Xi^{(j)}(N)$ can be considered to arise from different arrangement of N objects $r_1, r_2, \cdots, r_N$ in j cells $K_1, K_1^{(1)}, \cdots, K_1^{(j-1)}$ which are the values of $K_1, K_2, \cdots, K_N$ appearing in Eq. (2). The cells are numbered from 1 to j and placed as shown in diagrams 1 and 2.

Objects	$r_1, r_2, \cdots, r_{\lambda_1}$	$r_{\lambda_1+1}, \cdots, r_{\lambda_2}$	$\cdots$	$r_{\lambda_{j-1}+1}, \cdots, r_{\lambda_j}$
Cell	K_1	$K_1^{(1)}$	$\cdots$	$K_1^{(j-1)}$

Diagram 1.

If $\lambda_j < N$, then $r_{\lambda_{j+1}} - r_N$ are undistributed.

The above diagram is replaced by a simpler diagram.

Object	$1, \cdots, \lambda_1$	$\lambda_1 + 1, \cdots, \lambda_2$	$\cdots$	$\lambda_{j-1} + 1, \cdots, \lambda_j$
Cell	1	2	$\cdots$	j

Diagram 2.

If $\lambda_j < N$, then $\lambda_j + 1, \cdots, N$ are undistributed. This diagram corresponds to a term

$$\exp [iK_1(r_1 + r_2 + \cdots + r_{\lambda_1}) + K_1^{(1)}(r_{\lambda_1+1} + \cdots + r_{\lambda_2}) + \cdots + K_1^{(j-1)}(r_{\lambda_{j-1}+1} + \cdots + r_{\lambda_j})]$$

in Eq. (2). There is at least one object in each cell, hence their number is $\leq N - (j-1)$ in any cell. Some may be left undistributed. All the terms of $\Xi^{(j)}(N)$ are obtained by assigning to $\lambda_1, \lambda_2, \cdots, \lambda_j$ all values compatible with the above conditions and $K_1^{(1)}, K_1^{(2)}, \cdots, K_1^{(j-1)}$ taking all values in accordance

with condition II. In order to obtain arrangements which give rise to the terms of $\Xi^{(j+1)}(N)$, the cells are renumbered from 2 to $j+1$ and a new cell 1 is added in the front. The objects are rearranged in the new set of cells. This rearrangement is shown in diagram 3. r_1 always has the coefficient K_1 so it must be placed in cell 1, therefore it is taken from cell 2 and placed in cell 1.

The number of objects in cells 2 to $j+1$ is restored by shifting them from cell p to $p-1$, $p = 3, \cdots, j$ and in cell $j+1$ by drawing on undistributed objects as shown in diagram 3. Thus, to every arrangement in $\Xi^{(j)}(N)$ except those in which all objects are distributed, there corresponds one arrangement in $\Xi_1^{(j+1)}(N)$. It is clear that there is one-to-one correspondence between the terms in $\Xi^{(j)}(N-1)$ and $\Xi_1^{(j+1)}(N)$.

Objects	$1, \cdots, \lambda_1$	$\lambda_1+1, \cdots, \lambda_2$	$\lambda_2+1, \cdots, \lambda_3$	$\cdots$	$\lambda_{j-1}+1, \cdots, \lambda_j$	$: \Xi^j(N)$
Cells	1	2	3	$\cdots$	j	

$\lambda_j+1, \cdots, N$ undistributed

Objects	1	$2, \cdots, \lambda_1+1$	$\lambda_1+2, \cdots, \lambda_2+1$	$\lambda_2+2, \cdots, \lambda_3+1$	$\cdots$	$\lambda_{j-1}+2, \cdots, \lambda_j+1$	$: \Xi_1^{(j+1)}(N)$
Cells	1	2	3	4	$\cdots$	$j+1$	

$\lambda_j+2, \cdots, N$ undistributed.

Diagram 3.

Similarly, a one-to-one-correspondence exists between the terms of $\Xi^{(j)}(N-\mu)$ and $\Xi_\mu^{(j+1)}(N)$.

4. EVALUATION OF $\langle \Xi^{(j+1)}(N) \rangle$

Replacing j by $j+1$ in Eq. (6) gives

$$\langle \Xi^{(j+1)}(N) \rangle = \langle \Xi_1^{(j+1)}(N) \rangle + \cdots + \langle \Xi_{N-j}^{(j+1)}(N) \rangle,$$

and $\langle \Xi_1^{(j+1)}(N) \rangle$ can be obtained from $\langle \Xi^{(j)}(N-1) \rangle$ by replacing $K_1, K_1^{(1)}, \cdots, K_1^{(j+1)}$ by $K_1^{(1)}, \cdots, K^{(j)}$ (equivalent to relabeling) and multiplying the entire expression by $f(K_1)$, which is equivalent to putting object 1 in the newly added cell 1 in $\langle \Xi_1^{(j+1)}(N) \rangle$. Therefore, by (7), we have

$$\langle \Xi_1^{(j+1)}(N) \rangle = f(K_1)\{P_j(K_1^{(1)}, K_1^{(2)}, \cdots, K_1^{(j)} + Q_j(K_1^{(1)}, K_1^{(2)}, \cdots, K_1^{(j)}, N-1)\}. \quad (12)$$

Similarly,

$$\langle \Xi_2^{(j+1)}(N) \rangle = f^2(K_1)\{P_j(K_1^{(1)}, \cdots, K_1^{(j)}) + Q_j(K_1^{(1)}, \cdots, K_1^{(j)}, N-2)\} \quad (13)$$

and, finally,

$$\langle \Xi_{N-j}^{(j+1)}(N) \rangle = f^{N-j}(K_1)\{P_j(K_1^{(1)}, K_1^{(2)}, \cdots, K_1^{(j)}) + Q_j(K_1^{(1)}, K_1^{(2)}, \cdots, K_1^{(j)}, j)\}. \quad (14)$$

It follows that the minimum value of N in $\langle \Xi^{(j+1)}(N) \rangle$ is $j+1$.

Using Eqs. (6), (12), (13), and (14), we have

$$\begin{aligned}\langle \Xi^{(j+1)}(N) \rangle &= P_j(K_1^{(1)}, K_1^{(2)}, \cdots, K_1^{(j)}) \\ &\times \{f(K_1) + f^2(K_1) + \cdots + f^{N-j}(K_1)\} \\ &+ f(K_1)Q_j(K_1^{(1)}, K_1^{(2)}, \cdots, K_1^{(j)}, N-1) \\ &+ f^2(K_1)Q_j(K_1^{(1)}, K_1^{(2)}, \cdots, K_1^{(j)}, N-2) \\ &+ \cdots + f^{N-j}(K_1)Q_j(K_1^{(1)}, K_1^{(2)}, \cdots, K_1^{(j)}, j). \quad (15)\end{aligned}$$

One term $C_1^{(j)} f^N(K_1^{(1)})$ of Q_j in Eq. (10) will give the sum

$$\begin{aligned}C_1^{(j)}\{f(K_1)f^{N-1}(K_1^{(1)}) + f^2(K_1)f^{N-2}(K_1^{(1)}) \\ + \cdots + f^{N-j}(K_1)f^j(K_1^{(1)})\} \\ = e_1 f^N(K_1^{(1)}) + e_2 f^N(K_1),\end{aligned}$$

where e_1 and e_2 do not depend on N. Hence the sum in Eq. (10) will give terms which can be put in the form

$$\sum_{K_1^{(1)}, \cdots, K_1^{(j)}} \{C_0'^{(j+1)} f^N(K_1) + C_1^{(j+1)} f^N(K_1^{(1)}) + \cdots + C_j^{(j+1)} f^N(K_1^{(j)}, j)\}, \quad (16)$$

where $C_0'^{(j+1)}, C_1^{(j+1)}, \cdots, C_j^{(j+1)}$ are independent of N. Their explicit form is not needed for the present purpose because in the limit of large N this sum will vanish:

$$\begin{aligned}\langle \Xi^{(j+1)}(N) \rangle &= P_j(K_1^{(1)}, K_1^{(2)}, \cdots, K_1^{(j)}) \\ &\times [f(K_1) - f^{N-j+1}(K_1)]/[1 - f(K_1)] \\ &+ \sum_{K_1, \cdots, K_1^{(j)}} \{C_0^{(j+1)} f^N(K_1) \\ &+ C_1^{(j+1)} f^N(K_1^{(1)}) + \cdots + C_j^{(j+1)} f^N(K_1^{(j)})\}.\end{aligned}$$

Therefore,

$$\begin{aligned}&P_{j+1}(K_1, K_1^{(1)}, \cdots, K_1^{(j)}) \\ &= P_j(K_1^{(1)}, K_1^{(2)}, \cdots, K_1^{(j)}) f(K_1)/[1 - f(K_1)] \\ &= \sum_{K_1, K_1^{(1)}, \cdots, K^{(j)}} \frac{f(K_1)}{1 - f(K_1)} \frac{f(K_1^{(1)})}{1 - f(K_1^{(1)})} \cdots \frac{f(K_1^{(j)})}{1 - f(K_1^{(j)})} \quad (17)\end{aligned}$$

and $Q_{j+1}(N) \to 0$ as $N \to \infty$.

5. CALCULATION OF $\langle \Xi(N) \rangle$

$$\langle \Xi(N) \rangle = \{P_1(K_1) + Q_1(K_1, N)\} + \{P_2(K_1, K_1^{(1)}) + Q_2(K_1, K_1^{(1)}, N)\} + \cdots + \{P_n(K_1, K_1^{(1)}, \cdots, K_1^{(n-1)}) + Q_n(K_1, K_1^{(1)}, \cdots, K_1^{(n-1)}, N)\}. \quad (18)$$

In the limit $N \to \infty$,

$$\lim_{N\to\infty} \langle \Xi(N) \rangle = \langle \Xi \rangle = \langle S(k_1)S(k_2)\cdots S(k_n) \rangle = \sum_{K_1} \frac{f(K_1)}{1-f(K_1)} + \sum_{K_1, K_1^{(1)}} \frac{f(K_1)}{1-f(K_1)} \frac{f(K_1^{(1)})}{1-f(K_1^{(1)})} + \cdots + \sum_{K_1, \cdots, K_1^{(n-1)}} \frac{f(K_1)}{1-f(K_1)} \cdots \frac{f(K_1^{(n-1)})}{1-f(K_1^{(n-1)})}. \quad (19)$$

This formula gives the average $\langle S(k_1)S(k_2)\cdots S(k_n) \rangle$ in the limit $N \to \infty$ in terms of the Fourier transform of the probability distribution of interatomic distances.

We will now express Eq. (19) as a recurrence formula. This is also more suitable for the actual computation of $\langle S(k_1)S(k_2)\cdots S(k_n) \rangle$ as a function of $k_1, k_2, \cdots, k_n$. Since K_1 is fixed,

$$\langle S(k_1)S(k_2)\cdots S(k_n) \rangle = \frac{f(K_1)}{1-f(K_1)} \left(1 + \sum_{K_1^{(1)}} \frac{f(K_1^{(1)})}{1-f(K_1^{(1)})} + \sum_{K_1^{(1)}, K_1^{(2)}} \frac{f(K_1^{(1)})}{1-f(K_1^{(1)})} \frac{f(K_1^{(2)})}{1-f(K_1^{(2)})} + \cdots + \sum_{K_1^{(1)}, \cdots, K_1^{(n-1)}} \frac{f(K_1^{(1)})}{1-f(K_1^{(1)})} \times \frac{f(K_1^{(2)})}{1-f(K_1^{(2)})} \cdots \frac{f(K_1^{(n-1)})}{1-f(K_1^{(n-1)})} \right). \quad (20)$$

From condition II, $K_1^{(1)}$ in the first sum can be the sum of $1, 2, \cdots, n-1$ elements among $k_1, k_2, \cdots, k_n$. Those $K_1^{(1)}$ which are equal to one element occur only in this term since only those $K_1^{(1)}$ which have nonnull subsets can appear in higher sums. From Eq. (20), their sum is $\sum_{i=1}^{n} \langle S(k_i) \rangle$. Those $K_1^{(1)}$ which are a sum of two terms appear in this sum and the second sum. These terms can be combined giving

$$\sum \frac{f(K_1^{(1)})}{1-f(K_1^{(1)})} \left(1 + \sum_{K_1^{(2)}} \frac{f(K_1^{(2)})}{1-f(K_1^{(2)})} \right),$$

where $\sum$ means that $K_1^{(1)}$ is the sum of any two k's from $k_1, k_2, \cdots, k_n$ and $\sum_{K_1^{(2)}}$ means that sum over $K_1^{(2)}$ subject to condition II, i.e., it is different from $K_1^{(1)}$ and 0, and it is a sum of subset of terms of $K_1^{(1)}$. These give rise to $\sum'_{i,j=1}^{n} \langle S(k_i)S(k_j) \rangle$. Those $K_1^{(1)}$ which are sum of $(n-1)$ k's occur in the second, third, $\cdots$, nth terms and the collection of all these is

$$\sum'_{i,j,\cdots,l} \langle S(k_i)S(k_j)\cdots S(k_l) \rangle.$$

We thus obtain the recurrence formula

$$\langle S(k_1)S(k_2)\cdots S(k_n) \rangle = \frac{f(k_1 + \cdots + k_n)}{1-f(k_1 + k_2 + \cdots + k_n)} \left[1 + \sum_{i=1}^{n} \langle S(k_i) \rangle + \sum'_{i,j=1}^{n} \langle S(k_i)S(k_j) \rangle + \cdots + \sum'_{i,j,\cdots,l=1}^{n} \langle S(k_i)S(k_j)\cdots S(k_l) \rangle \right]. \quad (21)$$

There are n terms within the bracket which involve the mean value of S and its twofold, threefold, $\cdots$, $(n-1)$-fold products. The mean value of S can be found by putting $n = 1$. This result is then used to obtain the mean of twofold products which is given in terms of mean of S. To calculate the mean of the n-fold product of S, the mean of S, and its twofold, threefold, $\cdots$, $(n-1)$-fold products will have to be calculated successively by using Eq. (21). An extension of this result when k's are such that some partial sums vanish will be given in a later publication.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 12 DECEMBER 1969

Continuous Representation Theory Using the Affine Group*

ERIK W. ASLAKSEN
Bell Telephone Laboratories, Inc., Holmdel, New Jersey
AND
JOHN R. KLAUDER†
Bell Telephone Laboratories, Inc., Murray Hill, New Jersey

(Received 9 May 1969)

We present a continuous representation theory based on the affine group. This theory is applicable to a mechanical system which has one or more of its classical canonical coordinates restricted to a smaller range than $-\infty$ to ∞. Such systems are especially troublesome in the usual quantization approach since, as is well known from von Neumann's work, the relation $[P, Q] = -iI$ implies that P and Q must have a spectrum from $-\infty$ to ∞ if they are to be self-adjoint. Consequently, if the spectrum of either P or Q is restricted, at least one of the operators, say Q, is not self-adjoint and does not have a spectral resolution. Thus Q cannot generate a coordinate representation. This leads us to consider a different pair of operators, P and B, both of which are self-adjoint and which obey $[P, B] = -iP$. The Lie group corresponding to this latter algebra is the affine group, which has two unitarily inequivalent, irreducible representations, one in which the spectrum of P is positive. Using the affine group as our kinematical group, we have developed continuous representations analogous to those Klauder and McKenna developed for the canonical group, and have shown that the former representations have almost all the desirable properties of the latter.

1. INTRODUCTION

The general concepts and properties of continuous representation theory (CRT) have been developed by Klauder[1]; for convenience we briefly recall them here: Let $\mathcal{H}$ denote abstract Hilbert space and let $U[l]$ be a family of unitary operators on $\mathcal{H}$. If we now choose an arbitrary but fixed unit vector $\Phi_0 \in \mathcal{H}$, called the *fiducial vector*, then we can generate a subset of $\mathcal{H}$ by operating on Φ_0 with $U[l]$. Denote this subset by $\mathfrak{S}$; then

$$\mathfrak{S} = \{U[l]\Phi_0 : l \in \mathfrak{L}\},$$

where $\mathfrak{L}$ is some label space. With any vector $\Psi \in \mathcal{H}$ we can now associate the complex, bounded, continuous function

$$\psi(l) = (U[l]\Phi_0, \Psi),$$

and the set $\mathfrak{C} \equiv \{\psi(l) : \Psi \in \mathcal{H}\}$ is called a *continuous representation* of $\mathcal{H}$.

For the further development it is convenient to let the $U[l]$ be the elements of a kinematical group, and to interpret the labels l as the classical canonical coordinates p and q for a system with one degree of freedom, as we shall be considering here. Without going into any details at this point, we just mention that use of the classical canonical coordinates as the labels leads to a particularly simple physical interpretation of the theory.[2] For the common case, which we shall refer to as the *canonical case*, when the classical Cartesian coordinates p and q can take on any value on the real line, the CRT has been developed in detail by Klauder and McKenna.[3] In this development, the unitary operators of interest are the Weyl operators $U[p, q] = \exp[i(pQ - qP)]$, where Q and P are the familiar self-adjoint operators satisfying $[Q, P] = iI$.

In this paper we develop a CRT appropriate to a different group and suitable for different dynamical systems. Suppose, for example, that the range of the classical variable p is restricted to be positive, $p > 0$. Such restricted coordinates are not unknown; in particular, we were motivated to undertake the present investigation by the case of the gravitational field. There the metric has to satisfy certain positivity requirements,[4] which lead to restrictions on the range of the components $g_{\mu\nu}$. Such restrictions must be reflected in the quantum theory; in our example this requires that the spectrum of the operator P be positive, i.e., $P > 0$. According to a theorem of von Neumann,[5] such a restriction is not compatible with having Q and P both be self-adjoint, and thus the appropriate unitary operators cannot be the familiar Weyl operators of the usual canonical theory.

Elsewhere[6] we have argued that the affine group is

* This paper is based on a thesis submitted by E. W. Aslaksen to Lehigh University in partial fulfillment of the requirement for the Ph.D. degree.

† Part of this work was done while the author was at Syracuse University, with National Science Foundation support.

[1] J. R. Klauder, J. Math. Phys. **4**, 1055 (1963); **5**, 177 (1964).

[2] J. R. Klauder, Talk given at Seminar on Unified Theories of Elementary Particles, Munich, 1965 (unpublished); J. R. Klauder, J. Math. Phys. **8**, 2392 (1967).

[3] J. R. Klauder and J. McKenna, J. Math. Phys. **5**, 878 (1964); **6**, 68 (1965).

[4] C. Moller, *The Theory of Relativity* (Oxford University Press, London, 1962), p. 235.

[5] J. von Neumann, Math. Ann. **104**, 570 (1931).

[6] I. M. Gel'fand and M. A. Naimark, Dokl. Akad. Nauk SSSR, **55**, 570 (1947); E. W. Aslaksen and J. R. Klauder, J. Math. Phys. **9**, 206 (1968).

pertinent for this problem. This group is abstractly defined as the group of linear transformations without reflections on the real line: $x \to (P_0/p)x - q$. With this parameterization the unitary group elements may be given as

$$U[p, q] = e^{-iqP}e^{i\ln(p/p_0)B}, \tag{1.1}$$

where $p_0 > 0$, and where P and B are self-adjoint generators which fulfill

$$[B, P] = iP. \tag{1.2}$$

Although this group is formally "close" to the canonical group, as demonstrated by multiplying both sides of the canonical commutation relation

$$[Q, P] = iI$$

by P and making the identification

$$B = \tfrac{1}{2}(PQ + QP),$$

the actual unitary representations of the affine group are sufficiently different from those in the canonical case to necessitate a reexamination of the associated continuous representations. It is the purpose of this paper to carry out that reexamination for a finite number of degrees of freedom.

2. THE OVERCOMPLETE FAMILY OF STATES

The unitary representations of the affine group have been studied,[6] and it is known that there exist two and only two unitarily inequivalent, irreducible representations, one for which P is positive and one for which it is negative. In particular, if we take our representation space $\mathcal{R}$ to be $L^2(R)$ and denote by R^+ and R^- the positive and negative half of the real line, respectively, then $\mathcal{R}$ can be written as the direct sum of two subspaces which are invariant under $U[p, q]$:

$$\mathcal{R} = \mathcal{R}_+ \oplus \mathcal{R}_-,$$

where

$$\mathcal{R}_+ \equiv \{\phi(k): \phi \in L^2(R),\ \phi(k) = 0 \text{ if } k < 0\},$$
$$\mathcal{R}_- \equiv \{\phi(k): \phi \in L^2(R),\ \phi(k) = 0 \text{ if } k > 0\}.$$

If $\phi(k) \in \mathcal{R}$ and we choose the particular representation where P is just multiplication by k, then

$$U[p, q]\phi(k) = \left(\frac{p_0}{p}\right)^{\frac{1}{2}} e^{-iqk}\phi\left(\frac{p_0}{p}k\right). \tag{2.1}$$

We can evidently treat both inequivalent representations in $L^2(R^+)$ by writing

$$U[p, q]\phi(k) = \left(\frac{p_0}{p}\right)^{\frac{1}{2}} e^{\mp iqk}\phi\left(\frac{p_0}{p}k\right), \tag{2.2}$$

where $\phi(k) \in L^2(R^+)$, and the minus sign corresponds to the representation where P has positive spectrum, the plus sign to the representation where P has negative spectrum. Of course, since we have chosen $P > 0$, we shall always use the corresponding irreducible representation.

From the commutation relation (1.2) we can immediately deduce the following relations, which are frequently used in this paper:

$$U[p, q]U[p', q'] = U[pp'/p_0,\ q + (p_0/p)q'], \tag{2.3}$$

$$U^\dagger[p, q] = U[p_0^2/p,\ -qp/p_0], \tag{2.4}$$

$$PU[p, q] = \frac{p}{p_0}U[p, q]P, \tag{2.5}$$

$$BU[p, q] = U[p, q]\left(B + \frac{pq}{p_0}P\right), \tag{2.6}$$

$$QU[p, q] = U[p, q]\left(\frac{p_0}{p}Q + qI\right). \tag{2.7}$$

In order for the set $\mathfrak{S}$ to be suitable for constructing a continuous representation, we shall demand that it have the following three properties:

(1) For each $\Phi \in \mathfrak{S}$ and every $\delta > 0$, there exists a vector $\Phi' \in \mathfrak{S}$, $\Phi' \neq \Phi$, such that $\|\Phi - \Phi'\| < \delta$.

(2) The mapping $l \to \Phi[l]$ is a many–one continuous map of a separated topological space $\mathfrak{L}$ onto $\mathfrak{S}$. By continuity in $\mathfrak{S}$, we mean the usual weak continuity in $\mathcal{H}$. Thus, if $l_n \to l$, then $(\Phi[l_n], \Psi) \to (\Phi[l], \Psi)$ for all $\Psi \in \mathcal{H}$.

(3) The span of $\mathfrak{S}$ is dense in $\mathcal{H}$.

Such a subset $\mathfrak{S}$ is called an overcomplete family of states (OFS). Following Klauder and McKenna,[3] we show that the subset $\mathfrak{S}$ generated by an irreducible representation of the $U[p, q]$ in Eq. (1.1) is indeed an OFS.

Lemma 2.1: The function $V[q]$, defined by $V[q] \equiv e^{-iqP}$, is a strongly continuous function of q, i.e., $q_\alpha \to q$ implies that $\|(V[q_\alpha] - V[q])\Psi\| \to 0$ for each $\Psi \in \mathcal{H}$. The same is true for $W[p] = \exp[i \ln (p/p_0)B]$.

Proof: Let $\epsilon > 0$ be given, and $\delta = |q_\alpha - q| > 0$. Then

$$\|(V[q_\alpha] - V[q])\Psi\| = \|e^{-iqP}(e^{\pm i\delta P} - I)\Psi\| \leq \|(e^{\pm i\delta P} - I)\Psi\|.$$

By assumption $e^{\pm i\delta P}$ is weakly continuous, hence strongly continuous, so there exists a δ_0 such that $\delta < \delta_0$ implies that the last expression is less than ϵ.

Lemma 2.2: The mapping, $R^+ \times R \to \mathfrak{S}$, defined by $\Phi[p, q]$, is continuous when $R^+ \times R$ has the product topology and $\mathcal{H}$ the strong topology.

Proof: Let (p_0, q_0) and $\epsilon > 0$ be given. Then

$$\begin{aligned}\|\Phi[p_0, q_0] - \Phi[p, q]\| &= \|V[q_0]W[p_0]\Phi_0 - V[q]W[p]\Phi_0\| \\ &\leq \|V[q_0]W[p_0]\Phi_0 - V[p]W[p_0]\Phi_0\| \\ &\quad + \|V[q]W[p_0]\Phi_0 - V[q]W[p]\Phi_0\| \\ &= \|V[q_0]W[p_0]\Phi_0 - V[q]W[p_0]\Phi_0\| \\ &\quad + \|W[p_0]\Phi_0 - W[p]\Phi_0\| \equiv A + D.\end{aligned}$$

Because of Lemma 2.1, there exists a $\delta > 0$ such that $|p_0 - p| < \delta$, $|q_0 - q| < \delta$ implies that $A < \epsilon/2$, $D < \epsilon/2$.

Consider the functions $\psi(p, q) \equiv (\Phi[p, q], \Psi)$, $\Psi \in \mathcal{H}$. Since $\Phi[p, q]$ is strongly continuous, it is certainly weakly continuous; thus $\psi(p, q)$ is a continuous function. Schwartz's inequality, $|(\Psi, \Phi)| \leq \|\Psi\| \cdot \|\Phi\|$, gives $|\psi(p, q)| \leq \|\Psi\|$. Furthermore, $\psi(p, q)$ is square integrable. To show this, denote the unitary map from $\mathcal{H}$ to $\mathfrak{R}_+$ by T_+, and let $\phi_0(k) \in \mathfrak{R}$ and $\psi(k) \in \mathfrak{R}$ be the functions corresponding to Φ_0 and Ψ, respectively, under T_+. Then

$$T_+\Phi[p, q] = \left(\frac{p_0}{p}\right)^{\frac{1}{2}} e^{\mp iqk}\phi_0\left(\frac{p_0}{p}k\right)$$

and

$$\begin{aligned}\psi(p, q) &= (\Phi[p, q], \Psi) \\ &= (T_+\Phi[p, q], T_+\Psi) \\ &= \int e^{\pm iqk}\left(\frac{p_0}{p}\right)^{\frac{1}{2}}\phi_0^*\left(\frac{p_0}{p}k\right)\psi(k)\,dk \\ &= (2\pi)^{-\frac{1}{2}}\int e^{\pm iqk}\left[\left(\frac{2\pi p_0}{p}\right)^{\frac{1}{2}}\phi_0^*\left(\frac{p_0}{p}k\right)\psi(k)\right]dk \\ &= (2\pi)^{-\frac{1}{2}}\int e^{\pm iqk}h(p, k)\,dk,\end{aligned}$$

where

$$h(p, k) = \left(\frac{2\pi p_0}{p}\right)^{\frac{1}{2}}\phi_0^*\left(\frac{p_0}{p}k\right)\psi(k).$$

The functions $\phi_0(k)$ and $\psi(k)$ are measurable on R^+, the positive half of the real line. Since $p > 0$, p^{-1} is a measurable function; therefore $\phi_0(p_0/pk)$ is also measurable[7] as a function of p, so $h(p, k)$ is measurable on $R^+ \times R$. However,

$$\int |h(p, k)|^2\frac{dp}{p} = 2\pi\,|\psi(k)|^2\int\left|\phi_0\left(\frac{p_0}{p}k\right)\right|^2\frac{dp}{p},$$

with $p_0/pk = a$, $dp = -p^2/p_0k\,da$, and

$$\int\left|\phi_0\left(\frac{p_0}{p}k\right)\right|^2\frac{dp}{p} = \int|\phi_0(a)|^2\frac{da}{a} \equiv \langle P^{-1}\rangle.$$

The last integral is not finite for all $\phi_0 \in \mathfrak{R}$, so let $\mathfrak{R}_0$ be the subset of $\mathfrak{R}$ for which the integral is finite, and let the value of the integral be M. Then

$$\int |h(p, k)|^2\,dp = 2\pi M\,|\psi(k)|^2$$

and

$$\int dk\int |h(p, k)|^2\,dp = 2\pi M\,\|\Psi\|^2.$$

By the theorem of Tonelli,[8] $|h(p, k)|^2$ is integrable over $R^+ \times R^+$, and then Fubini's theorem[8] shows that the integration can be performed in any order; in particular, $\int |h(p, k)|^2\,dk$ exists for all p except possibly for a set of measure zero.

Further, $h(p, k)$ is integrable in k since it is the product of two functions which are both square integrable in k. So, for almost all fixed p, $\psi(p, q)$ is the Fourier transform of a function which is both integrable and square integrable on R^+ and, as a function of q, square integrable for almost all p. By Parseval's theorem,

$$\int |\psi(p, q)|^2\,dq = \int |h(p, k)|^2\,dk,$$

and since we have shown that the right-hand side is an integrable function of p, Tonelli's theorem finally gives that $\psi(p, q) \in L^2(R^+ \times R)$.

If we set $\lambda(p, q) = (\Phi[p, q], \Lambda)$, then the same arguments give

$$\iint \psi^*(p, q)\lambda(p, q)\,dp\,dq = 2\pi M(\Psi, \Lambda).$$

In the foregoing we chose the measure with respect to which the integrals over $\mathfrak{S}$ exists as simply $dp\,dq$ times an arbitrary constant. Klauder[1] has shown that, so far as the result of the integration being proportional to the inner product goes, there is no loss of generality in taking the measure to be the left-invariant group measure. For the affine group the left-invariant group measure is (in the usual notation for Lie groups[9])

$$\mu_\sigma^\rho(a) = \left[\frac{\partial\phi^\rho(t, a)}{\partial t^\sigma}\right]_{t=e},$$

[7] P. R. Halmos, *Measure Theory* (D. Van Nostrand Co., Princeton, N.J., 1950), p. 81.

[8] E. J. McShane, *Integration* (Princeton University Press, Princeton, N.J., 1944), pp. 137, 145.

[9] P. M. Cohn, *Lie Groups* (Cambridge University Press, London, 1965).

with $a = (p', q')$, $t = (p, q)$, and $e = (p_0, 0)$. Then

$$\mu_l = \begin{bmatrix} \frac{p}{p_0} & 0 \\ 0 & \frac{p_0}{p} \end{bmatrix},$$

and $|\mu_e| = \Delta_l = 1$. Thus, the measure is just a constant times $dp\,dq$ and, in view of the previous calculations, we choose

$$d\mu(p, q) \equiv (2\pi M)^{-1}\,dp\,dq.$$

We mention that the right-invariant group measure is not equal to the left-invariant group measure; we find

$$\mu_r = \begin{vmatrix} p & 0 \\ q & 1 \end{vmatrix}$$

and $\Delta_r = p$; thus the right-invariant measure is proportional to $p^{-1}\,dp\,dq$.

Having obtained these results, we can simply refer to the arguments of Klauder and McKenna[3]—in particular, those arguments leading to their Theorem 3.2 and Lemmas 3.3 and 3.4, from which our first theorem immediately follows:

Theorem 2.1:
(a) Let $\Psi \in \mathcal{H}$; then

$$\Psi = \iint (\Phi[p, q], \Psi)\Phi[p, q]\,d\mu(p, q).$$

(b) The span of $\mathfrak{S}$ is dense in $\mathcal{H}$.
(c) The identity operator may be written as

$$I = \iint \Phi[p, q]\Phi^{\dagger}[p, q]\,d\mu(p, q).$$

Consequently, $\mathfrak{S}$ is an overcomplete family of states.

3. THE CONTINUOUS REPRESENTATION

Let the map $C: \mathcal{H} \to \mathfrak{C}$ be defined by $C\Psi = \psi(p, q) \equiv (\Phi[p, q], \Psi)$ for each $\Psi \in \mathcal{H}$. Thus $\mathfrak{C} \equiv \{\psi(p, q)\}$. Since $\Phi[p, q] = U[p, q]\Phi_0$, it is clear that $\mathfrak{C}$ will depend on Φ_0; this dependence will be discussed later. Also, depending upon which of the two irreducible representations of the affine group we use, we get two spaces, $\mathfrak{C}_+$ and $\mathfrak{C}_-$. To see what these spaces consist of, consider

$$\psi(p, q) = \int \left(\frac{p_0}{p}\right)^{\frac{1}{2}} e^{iqk}\phi_0^*\left(\frac{p_0}{p}\,k\right)\psi(k)\,dk.$$

Let $q = q' + iq''$; then the space $\mathfrak{C}_+$ arising from $\mathcal{R}_+$ will consist of those functions which are the limit as $q'' \to 0$ of functions analytic in the lower half of the complex q plane, whereas $\mathfrak{C}_-$ will consist of those functions which are the limit as $q'' \to 0$ of functions analytic in the upper half plane. The two spaces have, except for the function which is identically zero, no elements in common, i.e., $\mathfrak{C}_+ \cap \mathfrak{C}_- = \{0\}$. By $\mathfrak{C}$ we mean either $\mathfrak{C}_+$ or $\mathfrak{C}_-$.

We define the inner product in $\mathfrak{C}$ by

$$(\psi', \psi) = \iint \psi'^*(p, q)\psi(p, q)\,d\mu(p, q). \quad (3.1)$$

Using the results of the previous section, we can show that the following theorem is true:

Theorem 3.1: The set $\mathfrak{C}$, given by $\psi(p, q) = (\Phi[p, q], \Psi)$ for all $\Psi \in \mathcal{H}$, is a family of bounded, continuous, and square-integrable functions. When supplied with the inner product displayed in Eq. (3.1), the set $\mathfrak{C}$ is a Hilbert space which is unitarily equivalent to the original space $\mathcal{H}$ under the unitary mapping

$$C\Psi = (\Phi[p, q], \Psi) = \psi(p, q),$$

$$C^{-1}\psi(p, q) = \iint \psi(p, q)\Phi[p, q]\,d\mu(p, q) = \Psi.$$

The set $\mathfrak{C}$ is called a continuous representation of $\mathcal{H}$.

We now investigate the existence of the derivatives of $\psi(p, q)$. Let $\mathfrak{D}_P$ and $\mathfrak{D}_B$ be the domains of P and B on $\mathcal{H}$. We want to see if $U[p, q]\mathfrak{D}_P \subset \mathfrak{D}_P$ and $U[p, q]\mathfrak{D}_B \subset \mathfrak{D}_B$. Let $\Phi \in \mathfrak{D}_P$. Using Stone's theorem,[8] we write P as

$$\lim_{q' \to 0} \frac{i}{q'}(V[q'] - I).$$

Then

$$\begin{aligned} &\frac{i}{q'}(V[q'] - I)U[p, q]\Phi \\ &\quad = \frac{i}{q'}\left(\exp\left(-iq'\left(1 - \frac{p_0}{p}\right)P\right)\right. \\ &\qquad \left. \times\, U[p, q]V[q'] - U[p, q]\right)\Phi \\ &\quad = \frac{i}{q'}\exp\left(-iq'\left(1 - \frac{p_0}{p}\right)\right)U[p, q](V[q'] - I)\Phi \\ &\qquad + \frac{i}{q'}\,U[p, q]\left(\exp\left(-iq'\left(\frac{p}{p_0} - 1\right)P\right) - I\right)\Phi, \end{aligned}$$

and in the limit $q' \to 0$, $PU[p, q]\Phi = U[p, q](p/p_0)P\Phi$; thus $U[p, q]\mathfrak{D}_P \subset \mathfrak{D}_P$.

Let $\Phi \in \mathfrak{D}_B$. Then

$$\frac{i}{\ln(p'/p_0)}(W[p'] - I)U[p, q]\Phi$$
$$= \frac{i}{\ln(p'/p_0)}\left(\exp\left(iq\left(1 - \frac{p_0}{p}\right)P\right)\right.$$
$$\left. \times\, U[p, q]W[p'] - U[p, q]\right)\Phi$$
$$= \frac{i}{\ln(p'/p_0)}\exp\left(iq\left(1 - \frac{p_0}{p}\right)P\right)U[p, q](W[p'] - I)\Phi$$
$$+ \frac{i}{\ln(p'/p_0)}U[p, q]\left(\exp\left(iqp\left(1 - \frac{p_0}{p}\right)P\right) - I\right)\Phi,$$

and in the limit $p'/p_0 \to 1$,

$$BU[p, q]\Phi = U[p, q](B + pqp_0^{-1}P)\Phi.$$

So $U[p, q]\Phi \in \mathfrak{D}_B$ if and only if $\Phi \in \mathfrak{D}_B \cap \mathfrak{D}_P$, or $U[p, q]\mathfrak{D}_B \cap \mathfrak{D}_P \subset \mathfrak{D}_B \cap \mathfrak{D}_P$. We define $\mathfrak{R}_B$ and $\mathfrak{R}_P$ by $T{:}\mathfrak{D}_B \to \mathfrak{R}_B$, $T{:}\mathfrak{D}_P \to \mathfrak{R}_P$, where $T{:}\mathfrak{H} \to \mathfrak{R}$.

Assume $\phi_0 \in \mathfrak{R}_B \cap \mathfrak{R}_P \cap \mathfrak{R}_0$, and let $\ln(p/p_0) \equiv \alpha$; then

$$\frac{\partial\psi}{\partial\alpha} = \lim_{\Delta\alpha\to 0}\frac{\psi(\alpha + \Delta\alpha, q) - \psi(\alpha, q)}{\Delta\alpha}$$
$$= \lim_{\Delta\alpha\to 0}\left(U[\alpha, q]\left\{\frac{W[\Delta\alpha] - I}{\Delta\alpha}\right\}\Phi_0, \Psi\right).$$

From Stone's theorem

$$\lim_{\Delta\alpha\to 0}\frac{W[\Delta\alpha] - I}{\Delta\alpha} = iB;$$

thus

$$\frac{\partial\psi}{\partial\alpha} = -i(U[\alpha, q]B\Phi_0, \Psi).$$

Further, $d\alpha/dp = p^{-1}$, so $\partial\psi/\partial\alpha = p\partial\psi/\partial p$, and

$$\frac{\partial\psi}{\partial p} = -ip^{-1}(U[p, q]B\Phi_0, \Psi),$$
$$\left|\frac{\partial\psi}{\partial p}\right| \leq p^{-1}\,\|B\Phi_0\| \cdot \|\Psi\|.$$

Similarly,

$$\frac{\partial\psi}{\partial q} = \lim_{\Delta q\to 0}\left(\left\{\frac{V[\Delta q] - I}{\Delta q}\right\}U[p, q]\Phi_0, \Psi\right)$$
$$= i(PU[p, q]\Phi_0, \Psi)$$
$$= ipp_0^{-1}(U[p, q]P\Phi_0, \Psi).$$

Thus,

$$\left|\frac{\partial\psi}{\partial q}\right| \leq pp_0^{-1}\,\|P\Phi_0\| \cdot \|\Psi\|.$$

The continuity of the derivatives follows directly from the strong continuity of the family of operators $U[p, q]$.

4. DIAGONAL EXPECTATION VALUES

As we have seen, p and q cannot be the eigenvalues of P and Q. But we can introduce a connection between the operator formalism and c numbers by requiring that p and q be the expectation values of P and Q, respectively, and also that the expectation value of B be pq, with respect to the states $\Phi[p, q]$. We have

$$(\Phi[p, q], P\Phi[p, q]) = (U[p, q]\Phi_0, PU[p, q]\Phi_0)$$
$$= (\Phi_0, U^\dagger[p, q]PU[p, q]\Phi_0)$$
$$= pp_0^{-1}(\Phi_0, P\Phi_0) \equiv p.$$

Thus, the fiducial vector Φ_0 must correspond to a state of the system in which the expectation value of P is p_0. In the representation space $\mathfrak{R}$, this means

$$\int \phi_0^*(k)k\phi_0(k)\,dk = p_0. \tag{4.1}$$

In addition, we always have the normalization

$$\int |\phi_0(k)|^2\,dk = 1.$$

This shows the reason for introducing the constant p_0; its value is determined by Eq. (4.1).

The expectation value of B is

$$(\Phi[p, q], B\Phi[p, q])$$
$$= \left(\Phi_0, U^\dagger[p, q]U[p, q]\left(B + \frac{pq}{p_0}P\right)\Phi_0\right)$$
$$= \left(\Phi_0, \left(B + \frac{qp}{p_0}P\right)\Phi_0\right) = (\Phi_0, B\Phi_0) + qp.$$

We can secure that the expectation value of B equals pq by demanding that $(\Phi_0, B\Phi_0)$ vanish, for which it is sufficient that Φ_0 be real.

The expectation value of Q is determined as follows:

$$QU = Q\exp[-iqP]\exp\left[i\ln\left(\frac{p}{p_0}\right)B\right]$$
$$= \exp[-iqP](Q + qI)\exp\left[i\ln\left(\frac{p}{p_0}\right)B\right]$$
$$= \exp[-iqP]\exp\left[i\ln\left(\frac{p}{p_0}\right)B\right]\left(\frac{p_0}{p}Q + qI\right)$$
$$= U\left(\frac{p_0}{p}Q + qI\right),$$

so

$$U^\dagger QU = U^\dagger U\left(\frac{p_0}{p}Q + qI\right) = \frac{p_0}{p}Q + qI.$$

Then

$$(\Phi[p, q], Q\Phi[p, q]) = \frac{p_0}{p}(\Phi_0, Q\Phi_0) + q.$$

We restrict Φ_0 to satisfy

$$(\Phi_0, Q\Phi_0) = 0; \tag{4.2}$$

this is always satisfied if $\phi_0(k)$ is real; thus $(\Phi[p,q], Q\Phi[p,q]) = q$.

In summary, we list the requirements on the fiducial vector Φ_0:

(a) The unit vector Φ_0 lies in the domain of P, P^{-1}, Q, and B.

(b) The expectation values of these operators are respectively chosen as

$$(\Phi_0, P\Phi_0) = \int |\phi_0(k)|^2 \, k \, dk \equiv p_0,$$

$$(\Phi_0, P^{-1}\Phi_0) = \int |\phi_0(k)|^2 \, k^{-1} \, dk \equiv M,$$

$$(\Phi_0, Q\Phi_0) = (\Phi_0, B\Phi_0) = 0.$$

These restrictions are not severe and there is a large class of allowed vectors Φ_0. For example, it suffices that $\phi_0(k)$ be real, several times differentiable, and vanish sufficiently fast at 0 and ∞.

5. A WEYL-LIKE REPRESENTATION OF GENERAL OPERATORS

Throughout CRT we make use of the operators $V[q]$ and $W[p]$. Analogous operators were introduced in the canonical case by Weyl[10] as a means of going from classical to quantum mechanics; in particular, he asserted that the quantum operator corresponding to the classical quantity

$$f(p,q) = \iint e^{i(\sigma p + \tau q)} \xi(\sigma, \tau) \, d\sigma \, d\tau$$

should be

$$F(P,Q) = \iint e^{i(\sigma P + \tau Q)} \xi(\sigma, \tau) \, d\sigma \, d\tau.$$

We now investigate the use of the analog of these operators in the affine case, and spell out more clearly under what conditions Weyl's representation holds. We are always assuming an irreducible representation of $U[p,q]$ and that the spectrum of P is positive, such that P^{-1} exists and $P^{\frac{1}{2}}$ is uniquely defined.

First of all, consider the following calculation. We know from the development in Sec. 2, particularly Theorem 2.1, that

$$(\Phi_0, P^{-1}\Phi_0)(\Lambda, \Psi) = \iint (\Lambda, \Phi[p,q])(\Phi[p,q], \Psi) \frac{dp\,dq}{2\pi}.$$

Then

$$(X, P^{-1}\Phi)(\Lambda, \Psi) = \iint (\Lambda, U[p,q]\Phi)(U[p,q]X, \Psi) \frac{dp\,dq}{2\pi}, \tag{5.1}$$

so that

$$|\Psi\rangle\langle X| \, P^{-1} = \iint (X, U^\dagger[p,q]\Psi) U[p,q] \frac{dp\,dq}{2\pi}$$

and

$$|\Psi\rangle\langle X| = \iint (X, PU^\dagger[p,q]\Psi) U[p,q] \frac{dp\,dq}{2\pi}.$$

The latter equation can be generalized to

$$A = \iint \mathrm{Tr}\,\{APU^\dagger[p,q]\} U[p,q] \frac{dp\,dq}{2\pi}, \tag{5.2}$$

and we define the *kernel* of A to be

$$\tilde{a}(p,q) \equiv \mathrm{Tr}\,\{APU^\dagger[p,q]\}.$$

In the canonical case the kernel does not contain the operator P, and it is straightforward to show that the appearance of this P is connected with the fact that the right- and left-invariant measures are not the same for the affine group. In Eq. (5.1), the left-hand side remains invariant if we operate on both Λ and Ψ with $U[r,s]$, the right-hand side must be invariant under left multiplication, i.e., we must use the left-invariant measure. However, we can rewrite the equation by letting $\Phi = P\Theta$; using the relation

$$U[p,q]P = p_0 p^{-1} P U[p,q],$$

we obtain

$$(X, \Theta)(\Lambda, \Psi) = \iint (\Lambda, PU[p,q]\Theta)(U[p,q]X, \Psi) \frac{p_0}{p} \frac{dp\,dq}{2\pi},$$

and if we now operate on X and Θ with $U[r,s]$, the left-hand side remains invariant. But this corresponds to multiplication on the right under the integral, so that we must now have the right-invariant measure, which indeed we do. If we commute P through $U[p,q]$, the measure becomes left invariant again, and we are back to the first case. If the P were not present, the right- and left-invariant measures would have to be identical.

Leaving aside the question of the relation between the type of operator and type of kernel for the moment, we first show that the closure of the set $U[p,q]$ (in the weak topology) equals the set of all linear operators on $\mathcal{H}$. Since the representation of the Weyl-like operators is irreducible, Schur's lemma says that if $[B, U[p,q]] = 0$ for all p, q, then $B \sim I$. Let a prime on an operator algebra denote the set of all bounded operators which commute with all the

[10] H. Weyl, *The Theory of Groups and Quantum Mechanics* (Dover Publications, Inc., New York, 1931).

operators in the algebra. Then we have

$$\{U[p,q]\}' = \{I\}.$$

But further we clearly have

$$\{U[p,q]\}'' = \{I\}' = \mathfrak{B}(\mathcal{H}),$$

where $\mathfrak{B}(\mathcal{H})$ is the set of all bounded operators on Hilbert space.

We need the following theorem, due to von Neumann:

Theorem 5.1: Let $\mathcal{A}$ be a $*$-algebra of operators on $\mathcal{H}$ such that $I \in \mathcal{A}$. Then $\mathcal{A}''$ is the weak closure of $\mathcal{A}$. A proof of this theorem can be found in Ref. 11, p. 44.

Applied to our case, this theorem says that any bounded operator B is given by the limit of sequence like

$$B_N = \sum_{n=1}^{N} c_n U[p_n, q_n].$$

But such a limit is given by the expression in Eq. (5.2) if we interpret the kernel as a "distribution." In particular, a matrix element has the form

$$\langle \Lambda | B | \Phi \rangle = \iint \tilde{b}(p,q) \langle \Lambda | U[p,q] | \Phi \rangle \frac{dp\,dq}{2\pi}.$$

Here $\langle \Lambda | U[p,q] | \Phi \rangle$ is the "test function" on which the "distribution" $\tilde{b}(p,q)$ operates. Since the set of all bounded linear operators is dense (in the weak topology) in the set of all linear operators, we have arrived at the desired result.

If we now try to find more general relationships between an operator and its kernel, e.g., $\tilde{a} \in L^2$ if A is bounded, we immediately run into difficulties due to the appearance of the operator P in the definition of the kernel, and it has not been possible to find any such general relationships for the expansion displayed in Eq. (5.2). But if we study the effect of P in the expansion, it becomes evident that we can define a slightly different expansion which has some nice properties. Going back to Eq. (5.1), we let the P^{-1} on the left-hand side operate on X instead of Φ,

$$(P^{-1}X, \Phi)(\Lambda, \Psi) = \iint (\Lambda, U[p,q]P\Phi)(U[p,q]P^{-1}X, \Psi) \frac{dp\,dq}{2\pi},$$

so that Eq. (5.2) now reads

$$A = \iint \mathrm{Tr}\,\{AU^\dagger[p,q]\}U[p,q]P \frac{dp\,dq}{2\pi}.$$

By writing $(X, P^{-1}\Phi) = (P^{-\frac{1}{2}}X, P^{-\frac{1}{2}}\Phi)$, we find that

$$A = \iint \mathrm{Tr}\,\{AP^{\frac{1}{2}}U^\dagger[p,q]\}U[p,q]P^{\frac{1}{2}} \frac{dp\,dq}{2\pi}. \quad (5.3)$$

It then follows immediately that

$$\mathrm{Tr}\,(B^\dagger A) = \iint \mathrm{Tr}\,\{BP^{\frac{1}{2}}U^\dagger[p,q]\}^* \times \mathrm{Tr}\,\{AP^{\frac{1}{2}}U^\dagger[p,q]\} \frac{dp\,dq}{2\pi} \equiv \iint b^*(p,q)a(p,q) \frac{dp\,dq}{2\pi}. \quad (5.4)$$

Since Hilbert–Schmidt operators are those for which $\mathrm{Tr}\,(A^\dagger A) < \infty$, we evidently have the following result:

Theorem 5.2: Every Hilbert–Schmidt operator A can be written in the form

$$A = \iint a(p,q)U[p,q]P^{\frac{1}{2}} \frac{dp\,dq}{2\pi},$$

with

$$a(p,q) \equiv \mathrm{Tr}\,\{AP^{\frac{1}{2}}U^\dagger[p,q]\} \in L^2(R^+ \times R).$$

To proceed further, we introduce a space of test functions similar to the space $\mathcal{S}$ consisting of infinitely differentiable functions of fast decrease, introduced by Schwartz.[12] We define $\mathcal{E}$ to be the linear space of all real-valued functions of two variables, $a(p,q)$, which are infinitely differentiable in p and q, fall off faster than any power of q for $q \to \pm\infty$, and faster than any power of p for $p \to 0$ and $p \to \infty$. It would be straightforward to introduce a topology which would make $\mathcal{E}$ a locally convex Frechet space, but since we shall not go into questions of continuity or any others which require a knowledge of the topology, we content ourselves with letting $\mathcal{E}$ be a linear space.

We now want to show that $a(p,q) \in \mathcal{E}$ implies that A is trace class, and to this end we represent the operator A by a sequence, in the following manner:

$$A = \int a(p,q) \sum_{m,n} |m\rangle\langle m|\, U[p,q]P^{\frac{1}{2}}\, |n\rangle\langle n| \frac{dp\,dq}{2\pi} = \sum_{m,n} A_{mn}\, |m\rangle\langle n|, \quad (5.5)$$

with

$$A_{mn} = \int a(p,q) \langle m|\, U[p,q]P^{\frac{1}{2}}\, |n\rangle \frac{dp\,dq}{2\pi}, \quad (5.6)$$

and where $\{|n\rangle, n = 0, 1, \cdots\}$ is a complete orthonormal sequence of vectors.

[11] J. Dixmier, *Les algebres d'operateurs dans l'espace hilbertien* (Gauthier-Villars, Paris, 1957).

[12] L. Schwartz, *Théorie des distributions* (Hermann & Cie., Paris, 1957), Vol. 2.

For an arbitrary vector $|\psi\rangle \in \mathcal{H}$, we certainly have

$$\begin{aligned}\langle\psi| A^\dagger A |\psi\rangle &= \sum_{n,q} \langle\psi \mid n\rangle\langle q \mid \psi\rangle \sum_m A^*_{qm}A_{nm} \\ &\leq \tfrac{1}{2}\sum_{n,q} [\langle\psi \mid n\rangle\langle n \mid \psi\rangle \\ &\quad + \langle\psi \mid q\rangle\langle q \mid \psi\rangle] \sum_m A^*_{qm}A_{nm}.\end{aligned}$$

Since $A^\dagger A$ is symmetric and positive, we have

$$\begin{aligned}\langle\psi| A^\dagger A |\psi\rangle &\leq \tfrac{1}{2}\sum_{n,q} [\langle\psi \mid n\rangle\langle n \mid \psi\rangle + \langle\psi \mid q\rangle\langle q \mid \psi\rangle] \\ &\quad \times \sum_m |A^*_{qm}A_{nm}| \\ &= \langle\psi| \sum_n E_n |n\rangle\langle n \mid \psi\rangle \\ &\equiv \langle\psi| E |\psi\rangle,\end{aligned}$$

where

$$E_n \equiv \sum_{q,m} |A^*_{qm}A_{nm}|.$$

The operator A is trace class if and only if $\mathrm{Tr}\,(A^\dagger A)^{\frac{1}{2}}$ is finite, but we just showed that $(A^\dagger A)^{\frac{1}{2}} \leq E^{\frac{1}{2}}$, so A is trace class if

$$\sum_n \left\{\sum_{q,m} |A^*_{qm}A_{nm}|\right\}^{\frac{1}{2}} < \infty. \tag{5.7}$$

It is not difficult to show that the inequality in Eq. (5.7) will be satisfied if there exist a constant C, two positive integers n_0 and m_0, and an $\epsilon > 0$ such that, for $m > m_0$ and $n > n_0$,

$$|A_{mn}| \leq C[m + n]^{-4-\epsilon}. \tag{5.8}$$

The sequence $\{A_{mn}\}$ is related to the kernel $a(p, q)$ by Eq. (5.6), and let us in particular assume that the sequence $\{|n\rangle\}$ is generated by an operator H,

$$H |n\rangle = n |n\rangle.$$

From Eq. (5.6) we then obtain

$$n^\beta m^\alpha A_{mn} = \int a(p, q) \langle m| H^\alpha U[p, q] P^{\frac{1}{2}} H^\beta |n\rangle \frac{dp\, dq}{2\pi}.$$

But the operator H to the left of $U[p, q]$ can be written in terms of a differential operator which operates on $U[p, q]$, say

$$R\left[p, q, \frac{\partial}{\partial p}, \frac{\partial}{\partial q}, p^{-1}\right],$$

i.e.,

$$RU[p, q] = HU[p, q].$$

Similarly, the operator H to the right of $P^{\frac{1}{2}}$ can also be written in terms of a differential operator on $U[p, q]$, say

$$S\left[p, q, \frac{\partial}{\partial p}, \frac{\partial}{\partial q}, p^{-1}\right],$$

i.e.,

$$SU[p, q]P^{\frac{1}{2}} = U[p, q]P^{\frac{1}{2}}H. \tag{5.9}$$

We thus find

$$n^\beta m^\alpha A_{mn} = \int a(p, q) \langle m| S^\beta R^\alpha U[p, q] P^{\frac{1}{2}} |n\rangle \frac{dp\, dq}{2\pi}. \tag{5.10}$$

Let us now further assume that H is a polynomial in P and B. It follows that R is a polynomial in p, q, $\partial/\partial p$, $\partial/\partial q$, and p^{-1}. Also, commuting H through $U[p, q]P^{\frac{1}{2}}$ in Eq. (5.9) will yield a polynomial in P, Q, p, q, and p^{-1} [see Eqs. (2.5)–(2.7)], so that S is also a polynomial in $p, q, \partial/\partial p, \partial/\partial q$, and p^{-1}. The operator $S^\beta R^\alpha$ is still a polynomial in $p, q, \partial/\partial p, \partial/\partial q$, and p^{-1}, and we can then, by repeated integrations by parts on each individual term in the polynomial, bring the operator over to operate on $a(p, q)$ in Eq. (5.10). The resulting operator, say $E_{\alpha\beta}$, must also be a polynomial in p, q, $\partial/\partial p$, $\partial/\partial q$, and p^{-1}, and we finally obtain

$$n^\beta m^\alpha A_{mn} \leq \left\{\int |E_{\alpha\beta} a(p, q)|^2 \frac{dp\, dq}{2\pi}\right\}^{\frac{1}{2}} \tag{5.11}$$

by use of Schwarz's inequality.

Since $a(p, q) \in \mathcal{E}$, the right-hand side of Eq. (5.11) is finite, and so by choosing $\alpha > 4$, $\beta > 4$, we have shown that $a(p, q) \in \mathcal{E}$ implies that A is trace class.

Let A be a trace class operator such that its kernel $a(p, q)$ is an element of $\mathcal{E}$, and let B be an arbitrary bounded operator with kernel $b(p, q)$. Then

$$\mathrm{Tr}\,(AB) = \int a(p, q)b(p, q) \frac{dp\, dq}{2\pi} < \infty,$$

since AB is also a trace class operator. This relation defines a valid linear functional on the space $\mathcal{E}$ and leads to our next theorem.

Theorem 5.3: Every bounded operator B can be written in the form

$$B = \iint b(p, q)U[p, q]P^{\frac{1}{2}} \frac{dp\, dq}{2\pi},$$

with

$$b(p, q) = \mathrm{Tr}\,\{BP^{\frac{1}{2}}U^\dagger[p, q]\} \in \mathcal{E}'.$$

The space $\mathcal{E}'$, which is the dual of $\mathcal{E}$, is the analog of the tempered distributions in this case where $p > 0$. In the canonical case it is known that the Weyl kernel of an arbitrary bounded operator is a tempered distribution.[13]

We conclude this section by remarking that in the expansion

$$A = \iint a(p, q)U[p, q]P^n \frac{dp\, dq}{2\pi}, \tag{5.12}$$

[13] G. Loupias, Compt. Rend., Ser. A **262**, 799 (1966); S. Miracle-Solé, *ibid.*, 1478 (1966).

with

$$a(p, q) = \mathrm{Tr}\,\{AP^{1-n}U^{\dagger}[p, q]\},$$

any value of n is admissible so far as forming a legitimate expansion is concerned. It is only necessary to treat the kernel $a(p, q)$ as a "distribution," which can always be found by the following limiting operation:

$$a(p, q) = \lim_{M} \sum_{n=1}^{M} \langle m|\, AP^{1-n}U^{\dagger}[p, q]\, |m\rangle,$$

where $\{|m\rangle, m = 1, 2, \cdots\}$ is a complete sequence of orthonormal vectors. However, the properties of the expansion, i.e., the particular relationships between classes of operators and their corresponding classes of kernels, do depend on the value of n in Eq. (5.12), and we have seen that the choice $n = 0$, which is the closest analog to an expansion in Weyl operators, is not the best choice from this viewpoint.

6. UNIQUENESS OF THE DIAGONAL MATRIX ELEMENTS

To what extent do the diagonal matrix elements $(\Phi[p, q]\, A\Phi[p, q])$ of an operator A determine the complete matrix $(\Phi[p, q], A\Phi[p', q'])$? Assume that two operators A_1 and A_2 lead to the same diagonal elements, and let $D = A_1 - A_2$. Then

$$(\Phi[p, q],\, D\Phi[p, q]) = 0 \tag{6.1}$$

for all p and q. If, in the fashion of (5.2), we write D as

$$D = \iint d(p', q')U[p', q']\frac{dp'\,dq'}{2\pi},$$

then Eq. (6.1) becomes

$$\iint (\Phi_0,\, U^{\dagger}[p, q]U[p', q']U[p, q]\Phi_0) \times d(p', q')\frac{dp'\,dq'}{2\pi} = 0.$$

Now,

$$U^{\dagger}[p, q]U[p', q']U[p, q] = U\left[p', \frac{p}{p_0}(q' - q) + \frac{p}{p'}q\right],$$

so Eq. (6.1) finally can be written

$$\iint \mathcal{K}\left(p_0, 0; p', \frac{p}{p_0}(q' - q) + \frac{p}{p'}q\right) \times d(p', q')\frac{dp'\,dq'}{2\pi} = 0, \tag{6.2}$$

where

$$\mathcal{K}(p, q; p', q') \equiv (\Phi[p, q],\, \Phi[p', q']). \tag{6.3}$$

The function $\mathcal{K}(p, q; p', q')$ is, of course, dependent on our choice of the fiducial vector Φ_0. For each choice of Φ_0 there will be a class of operators which are uniquely determined by their diagonal matrix elements; for some choices of Φ_0 it may even be that all operators are uniquely defined by their diagonal matrix elements. A complete analysis of this uniqueness problem has not been carried out, but we demonstrate that operators of the general form

$$A = \sum_{m=-M}^{M} \sum_{n=0}^{N} a_{mn}P^{m}B^{n}, \tag{6.4}$$

i.e., polynomials in P, P^{-1}, and B, are uniquely determined by their diagonal matrix elements. Let

$$J(p, q) \equiv \sum_{m,n} a_{mn}(\Phi[p, q],\, P^{m}B^{n}\Phi[p, q]) = \sum_{m,n} a_{mn}\left(\Phi_0,\, \left(\frac{p}{p_0}P\right)^{m}\left(B + \frac{pq}{p_0}P\right)^{n}\Phi_0\right);$$

then, if $b \equiv pq/p_0$, we have

$$J\left(p, \frac{p_0 b}{p}\right) = \sum_{m,n} a_{mn}\left(\frac{p}{p_0}\right)^{m}(\Phi_0,\, P^{m}(B + bP)^{n}\Phi_0). \tag{6.5}$$

Fix n at its maximum value, $n = N$, and choose b large enough so that we only have to consider the leading term $(bP)^N$ in $(B + bP)^N$. Then, if the left-hand side of Eq. (6.5) is identically zero, we have

$$\sum_{m=-M}^{M} a_{mN}\left(\frac{p}{p_0}\right)^{m} b^{N}(\Phi_0,\, P^{m+N}\Phi_0) = 0.$$

But P^{m+N} is a positive operator, so this implies that $a_{mN} = 0$ for all m. Then we set $n = N - 1$ and go through the same argument, and so by induction we see that $a_{mn} = 0$ for all m and n. Since the operators of the form displayed in Eq. (6.4) form a linear space and the mapping $A \to (\Phi[p, q],\, A\Phi[p, q])$ is linear, we have proved the following theorem:

Theorem 6.1: Any polynomial in P, P^{-1}, and B of the form

$$A = \sum_{m=-M}^{M} \sum_{n=0}^{N} a_{mn}P^{m}B^{n}$$

is uniquely determined by its diagonal matrix elements in any affine phase-space continuous representation.

Although we have treated only one degree of freedom in this paper, the extension to finitely many degrees of freedom is relatively straightforward and consists mainly of introducing a sufficient number of indices and a convenient notation for handling these. We shall omit going into this matter here, but in a forthcoming paper where we extend the present results to field theories, the necessary notation will be displayed.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 12 DECEMBER 1969

Dirac Formalism and Symmetry Problems in Quantum Mechanics. II. Symmetry Problems*

J.-P. ANTOINE†
Palmer Physical Laboratory, Princeton University, Princeton, N.J.

(Received 10 March 1969)

The quantum-mechanical formalism developed in a previous article and based on the use of a rigged Hilbert space $\Phi \subset \mathcal{H} \subset \Phi'$ is here enlarged by taking into account the symmetry properties of the system. First, the compatibility of a particular symmetry with this structure is obtained by requiring Φ to be invariant under the corresponding representation U of the symmetry group in $\mathcal{H}$. The symmetry is then realized by the restriction of U to Φ and its contragradient representation $\check{U}$ in Φ'. This double manifestation of the symmetry is related to the so-called *active* and *passive* points of view commonly used for interpreting symmetry operations. Next, a general procedure is given for constructing a suitable space Φ out of the labeled observables of the system *and* the representation U describing its symmetry properties. This general method is then applied to the case where U is a semidirect product $G = T \boxtimes \Delta$, with T Abelian. Finally, the examples of the Euclidean, the Galilei, and the Poincaré groups are briefly studied.

I. INTRODUCTION

In a previous article[1] (denoted by Part I in the following), a formalism for quantum mechanics was developed which reproduces the basic features of the standard Dirac bra and ket formalism. Now we turn to the problem of symmetries in that context. For the sake of completeness, however, we first recall the main points of the formalism.

(i) The system is *defined* by a particular family of observables, called *labeled observables* (l.o.) by Roberts.[2] They have both a direct physical definition (i.e., in experimental terms) and a mathematical expression as self-adjoint operators in a Hilbert space $\mathcal{H}$. All l.o. must have a common, invariant, dense domain $D \subset \mathcal{H}$.

(ii) This domain D is endowed with a new topology, finer than that induced by the scalar product of $\mathcal{H}$, and such as to make all l.o. continuous. There exists a *coarsest* topology satisfying these requirements, which is the canonical solution of Roberts.[2] Moreover, we require that this topology be *nuclear*.[3–5] Thus, we obtain a complete nuclear topological vector space, which we denote by $\Phi: \Phi \subset \mathcal{H}$. Φ is dense in $\mathcal{H}$ and the embedding $\Phi \to \mathcal{H}$ is continuous (thus nuclear).

(iii) If Φ' denotes the strong dual of Φ, we obtain finally a rigged Hilbert space[2,4] (RHS)

$$\Phi \subset \mathcal{H} \subset \Phi'. \tag{1}$$

(iv) With respect to Φ', a general spectral theorem holds; any l.o. has in Φ' a complete orthonormal (in the sense of Parseval's relation) system of eigenfunctionals.

(v) The following physical interpretation was proposed:
(1) The vectors of Φ, the *ket vectors*, denoted $|\cdot\rangle$, represent the truly realizable states of the system (preparable states);
(2) the vectors of Φ', the *bra vectors*, denoted $\langle\cdot|$, represent potential experiments or "elementary instruments" that can be applied to the prepared system;
(3) the bracket $\langle\cdot\,|\,\cdot\rangle$, the bilinear form expressing duality between Φ and Φ', generalizes the usual scalar product in $\mathcal{H}$ with the same interpretation as a probability amplitude.

So far, we have not touched upon the symmetry properties of the system. These, however, may play an important role in the theory. For one thing, some of them express fundamental properties of nature (such as relativity, either Galilean or Lorentzian), with which conservation laws are associated. No complete physical theory can ignore those symmetries and, clearly, the corresponding conserved physical quantities *must* be included among the l.o. On the other hand, the use of symmetries facilitates and clarifies greatly the investigation of a problem. Therefore, it is desirable to include explicitly into the theory the symmetry properties of the system, *from the start.*

The problem is thus: How can one describe the invariance of the system under some group G in the framework of an RHS? In conventional Hilbert-space quantum mechanics, this question was answered

---

* This work is based partly on a doctoral dissertation submitted by the author to the University of Louvain, 1966.

† On leave of absence from University of Louvain, Belgium. Present address: Department of Physics, University of Pittsburgh, Pittsburgh, Pa. 15213.

[1] J.-P. Antoine, J. Math. Phys., **10**, 53 (1969).

[2] J. E. Roberts, J. Math. Phys. **7**, 1097 (1966); Commun. Math. Phys. **3**, 98 (1966).

[3] F. Treves, *Topological Vector Spaces, Distributions and Kernels* (Academic Press Inc., New York, 1967); A. Pietsch, *Nukleare lokalkonvexe Räume* (Akademie-Verlag, Berlin, 1965).

[4] I. M. Gel'fand and N. Y. Wilenkin, *Verallgemeinerte Funktionen (Distributionen)*, (VEB Deutsche Verlag der Wissenschaften, Berlin, 1965), Bd. IV [English transl.: *Generalized Functions* (Academic Press Inc., New York, 1964), Vol. 4].

[5] This rather strong requirement might be relaxed, but it makes the theory easier.

long ago by Wigner[6] and Bargmann.[7] The symmetry is simply realized by a unitary representation $g \to U(g)$ of G in $\mathcal{H}$. Thus, the problem becomes: How can one extend such a representation to an RHS? We note first that group representations in RHS have already been considered in the literature,[8] but from a very different point of view. One starts usually with a continuous representation V in some topological vector space E, and seeks a suitable, continuous scalar product on E such that the representation V has a unitary extension into the resulting Hilbert space. Here, however, the problem is just the inverse: The probabilistic axioms give both $\mathcal{H}$ and a unitary representation U of G in $\mathcal{H}$; now the space Φ is the element to be built suitably.

More precisely, the system is defined by its l.o., the representation U, and the transformation properties of the l.o. under U. With these elements, we have to construct a suitable space Φ. It is important to notice that some l.o. are usually derived from the symmetry itself (e.g., infinitesimal generators of the symmetry operations obtained through Stone's theorem), but *not all of them*. We must therefore keep track simultaneously of both the l.o. and the representation U. This raises the question of *compatibility* between the symmetry, i.e., the representation U, and a particular RHS used for describing the system. This problem is examined and solved in Sec. II by requiring that Φ be invariant under U. As a consequence, we have to consider two new representations: the restriction of U to Φ, and its contragredient representation $\check{U}$ in Φ'. This double manifestation of the symmetry is then seen to coincide with the distinction between the so-called active and passive points of view commonly used for interpreting the symmetry operations.

Section III is devoted to the construction problem: Given $\mathcal{H}$ and U, build a suitable RHS. We treat first the representation alone, thus considering only those l.o. which are derived from the symmetry [elements of the Lie algebra (G is supposed to be a Lie group) and of the universal enveloping algebra of G]. Since these elements can be identified with particular distributions on G with support consisting of the neutral element only, we extend the problem to the whole algebra of distributions on G with compact support. Then we construct a space Φ in which all these distributions, including the relevant l.o., are correctly represented. Finally, the solution is extended to take care of those l.o. which are not derived from the symmetry, but have simple transformation properties under U.

In Sec. IV, the general solution is applied to the case where U is an induced representation in the sense of Mackey[9] and Bruhat.[10] Particular attention is given to the physically important case of a semidirect product $G = T \boxtimes \Delta$, with T an invariant Abelian subgroup. Examples of such a situation are studied in Sec. V; they include the Euclidean, the Galilei, and the Poincaré groups. Finally, Appendix A illustrates the duality $U \leftrightarrow \check{U}$ in the case of a parity-violating process, while Appendix B summarizes relevant notions from the theory of distributions over Lie groups.

II. COMPATIBILITY BETWEEN SYMMETRY AND RHS DESCRIPTION

Let us consider an irreducible system, i.e., a system without superselection rules,[11] and suppose that it possesses a group of symmetry G. We will assume that G is a connected and simply connected Lie group (more general situations can be treated in a similar way). In the conventional Hilbert-space formalism, a complete analysis of such a symmetry has been given in the fundamental works of Wigner[6] and Bargmann,[7] namely:

(i) The symmetry of the system under G is represented by a *strongly continuous unitary projective representation* U_ω of G in $\mathcal{H}$, the Hilbert space of the state vectors of the system [i.e., a projective representation, or representation up to a factor, continuous in the strong (norm) topology of $\mathcal{H}$]:

$$U_\omega(g_1)U_\omega(g_2) = \omega(g_1, g_2)U_\omega(g_1g_2), \quad |\omega(g_1, g_2)| = 1.$$

(ii) The projective representation U_ω can be obtained from a *true* (i.e., vector) unitary representation of a group G_ω which is an extension of G by the 1-dimensional Abelian group of the reals.

Henceforth, we shall suppose that this extension has been accomplished; thus, the symmetry is realized by a *true* continuous unitary representation $g \to U(g)$ of G in $\mathcal{H}$. This representation may then, as usual, operate either on the state vectors of the system or on its observables.

On the other hand, we may also associate to the system a RHS $\Phi \subset \mathcal{H} \subset \Phi'$ built with help of its l.o.,

[6] E. P. Wigner, *Gruppentheorie* (Frederick Vieweg und Sohn, Braunschweig, Germany, 1931) [English transl.: *Group Theory* (Academic Press Inc., New York, 1959)]; Ann. Math. **40**, 149 (1939).

[7] V. Bargmann, J. Math. Phys. **5**, 862 (1964); Ann. Math. **59**, 1 (1954).

[8] M. Mayer, in *Proceedings of Seminar on Unified Theories of Elementary Particles*, H. Rechenberg, Ed. (Univ. of München, München, 1965).

[9] G. W. Mackey, *The Theory of Group Representations* (The University of Chicago Press, 1955); Ann. Math. **55**, 101 (1952); **56**, 193 (1953).

[10] F. Bruhat, Bull. Soc. Math. France **84**, 97 (1956).

[11] It is well known [A. S. Wightman, Suppl. Nuovo Cim., **14**, 81 (1959)] that symmetry operations can at most induce a permutation of superselection sectors; therefore, there is no loss of generality in considering only a single coherent subspace or, equivalently, an irreducible system.

according to Part I. The question then arises whether or not this description is *compatible with the symmetry.* What does this mean? The argument is the same one already used in connection with the temporal evolution: We want to space Φ to be as well adapted to the system as possible. Therefore, if Φ is not stable under U, we call it incompatible with the symmetry, in the sense that the choice of Φ distinguishes a particular element, or a class of elements, of G; this contradicts the very notion of symmetry.[6] We must therefore impose the following condition:

Postulate of compatibility: The space Φ is stable under the representation U:

$$U(g)\Phi = \Phi, \quad \text{for all } g \in G. \tag{2}$$

[The equality sign follows from the group property: $U(g^{-1}) = U^{-1}(g)$ must leave Φ invariant as well as $U(g)$.]

Clearly, this postulate imposes a new restriction to the topology of Φ. We already had a lower bound for it, namely, the coarsest topology (Roberts' canonical topology[2]) which makes all l.o. continuous. But we have now also an upper bound. Indeed, from the postulate it follows that G is represented by operators $U(g)$ mapping Φ into itself. Of course, $U(g)$ is continuous when its range, i.e., Φ, is endowed with the topology induced by $\mathcal{H}$, but this is not true *a priori* when Φ has its own, finer topology. The same difficulty appears with the function $g \to U(g)\varphi$ (for any fixed $\varphi \in \Phi$), which maps G into Φ. In other words, the representation U is not necessarily continuous in Φ, in the sense of Bruhat[10] and Mayer.[8] Thus, the upper bound referred to is the finest topology such that the representation $g \to U(g)$ is continuous in Φ. With any topology finer than this one, one may expect some catastrophe: The representation of G will not be continuous, the corresponding representation of the Lie algebra will not be integrable, etc.

We shall show in Sec. III that it is always possible to construct a space Φ such that $g \to U(g)$ is a *continuous* representation of G in Φ. From this and the results of Bruhat,[10] it follows that U can be extended by transposition to a *continuous* representation $g \to \check{U}(g)$ of G in Φ', called the *contragredient representation*; the latter is defined by any one of the equivalent relations:

$$\langle \check{U}(g)\varphi' \mid \varphi \rangle = \langle \varphi' \mid U(g^{-1})\varphi \rangle, \quad \varphi \in \Phi, \varphi' \in \Phi', g \in G, \tag{3a}$$

or

$$\langle \check{U}(g)\varphi' \mid U(g)\varphi \rangle = \langle \varphi' \mid \varphi \rangle. \tag{3b}$$

A sufficient condition for the continuity of $\check{U}$ is that Φ be a *semireflexive space*, i.e., $(\Phi')' = \Phi$ as a vector space, without considering the topology. With the physical interpretation we have adopted in Part I, this condition is obviously an essential consistency requirement of the formalism. Indeed, a physical experiment requires only two kinds of entities: the system, with all its accessible states associated with the vectors of Φ, and the apparatus, i.e., the collection of instruments which determine these states and are represented by the vectors of Φ'. The physical relation between these two classes of entities is mathematically described by the duality between Φ and Φ', and there is no room for a third class associated with $(\Phi')'$.[12] The physical interpretation thus leads us to require Φ to be semireflexive. We could as well ask Φ to be *reflexive*, i.e., $(\Phi')' = \Phi$ both algebraically *and* topologically. The latter requirement is simpler, and practically equivalent for the present theory, since we did not give a physical meaning to the topology of Φ itself. Semireflexivity follows, also, from the nuclearity of Φ, but it is, of course, a much weaker condition: Even if Φ would not be nuclear (e.g., if only the embedding operator $\Phi \to \mathcal{H}$ were nuclear), we must ask that Φ be semireflexive. Moreover, in the canonical construction of Roberts,[2] Φ is always semireflexive, and reflexive if the set of l.o. is countable, i.e., for a system with a finite number of degrees of freedom.

So we have now a double manifestation of the symmetry: U in Φ, and $\check{U}$ in Φ'. This double manifestation corresponds exactly to the two dual points of view of interpreting the symmetry operations: the active and the passive interpretations.[13] Let us go through the details:

(1) In the *active point of view*, the symmetry operations act on the states of the system; since we admit the physically accessible states to be represented by vectors of Φ, clearly the symmetry must be described by the representation U; particularly, the transformation

$$\langle \varphi' \mid \varphi \rangle \to \langle \varphi' \mid U(g)\varphi \rangle$$

means that the state φ and the transformed state $U(g)\varphi$ are compared in the same frame of reference (i.e., apparatus) φ', as requested in the active interpretation.

(2) On the other hand, in the *passive point of view* the same state is observed in two different frames; but

[12] Note, however, that due to the construction, one always would have

$$\Phi \subseteq (\Phi')' \subset \mathcal{H} \subset \Phi',$$

but this manifestly complicates the mathematical situation!

[13] A. S. Wightman, Ref. 11.

in any experiment the only relevant external frame for the system under observation is the one provided by the apparatus (we consider only the ideal situation where the pair system-and-apparatus is truly isolated from the rest of the world; this can be justified by. invoking the classical properties of the apparatus[14]). In other words, a frame here is an observing device or apparatus, identified in our language with a vector of Φ'. Therefore, in the passive interpretation, the symmetry must be described by the representation $\check{U}$ in Φ'. Indeed, the transformation

$$\langle \varphi' \mid \varphi \rangle \to \langle \check{U}(g)\varphi' \mid \varphi \rangle$$

means that the state φ is observed by the two different apparatus φ' and $\check{U}(g)\varphi'$.

(3) The relations (3) express the *equivalence of the two points of view*. This fundamental equivalence, in fact, reflects the invariance of space–time under the group G, as contrasted to the invariance of a particular process (the latter notion has a meaning only when the space–time background is itself invariant; otherwise it would be impossible to define what is meant by an invariant process). Relations (3) therefore hold for *any* particular process, whether invariant or not under G. This point is illustrated in Appendix A for the case of a parity-violating decay.

Remark: If G is not connected, the postulate (2) must be extended to all operators, unitary or antiunitary, which represent the "reflections" leading from the component of the identity to the other components. The whole discussion goes through without any modification.

III. CONSTRUCTION OF A RHS COMPATIBLE WITH THE SYMMETRY

A. Representation of the Observables Derived from the Symmetry

Let G be a connected Lie group (see the remark at the end of Sec. II). The Lie algebra of G, which we denote by $\mathfrak{g}$, and its universal enveloping algebra[15] $\mathfrak{U}(\mathfrak{g})$ are naturally associated with some observables of the system. Indeed, let $\tau \to a(\tau)$ be a 1-parameter subgroup of G; it follows that $\tau \to U(a(\tau))$ is a continuous 1-parameter group of unitary operators. According to Stone's theorem,[16] this group is generated by a self-adjoint operator, the map of an element of $\mathfrak{g}$. In general, this operator is unbounded and can be identified with an l.o. of the system. In the case of rotational invariance, for instance, $G = SO(3)$, the infinitesimal generators L_1, L_2, L_3 ($\in \mathfrak{g}$) of the group and its Casimir operator $L^2[\in \mathfrak{U}(\mathfrak{g})]$ are identified with the components and the norm of the total angular momentum of the system, respectively. Thus, the problem is to know which elements of $\mathfrak{U}(\mathfrak{g})$, besides those of $\mathfrak{g}$, can be considered as observables, and to find then a common, invariant, dense domain D for all l.o. obtained in this way. With such a domain at hand, it remains to endow it with a suitable topology: We will then have constructed a space Φ associated with the symmetry alone. A second step will be to take into account those l.o. which are not given by the symmetry.

A preliminary solution of the problem is given by the *Gårding domain* $\mathcal{K}^G$ of the representation U.[17] (Some useful results from the theory of distributions over Lie groups are collected in Appendix B.) Let $\mathfrak{D}(G)$ be the set of all infinitely differentiable functions on G with compact support. To each $f \in \mathfrak{D}(G)$, one associates the operator $\tilde{U}(f)$, defined by the following relation:

$$\tilde{U}(f)h = \int_G U(g)hf(g)\,dg, \quad \text{for all } h \in \mathcal{H}, \qquad (4)$$

where dg is the (left-invariant) Haar measure on G. $\mathcal{K}^G$ is then defined as the set of all finite linear combinations of vectors of the form $\tilde{U}(f)h, f \in \mathfrak{D}(g)$, $h \in \mathcal{H}$. It is easy to verify that this domain $\mathcal{K}^G$ is dense in $\mathcal{H}$, stable under $U(g)$, $\forall g \in G$, contained in the domain of the infinitesimal generators of *all* 1-parameter subgroups $\tau \to U(a(\tau))$, and stable under all of them. It follows that the enveloping algebra $\mathfrak{U}(\mathfrak{g})$ is (homomorphically) represented by operators in $\mathcal{K}^G$; the element T of $\mathfrak{U}(\mathfrak{g})$ is represented in $\mathcal{K}^G$ by the operator $\tilde{U}(T)$ defined by the following relation (which is extended to the whole $\mathcal{K}^G$ by linearity):

$$\tilde{U}(T)(\tilde{U}(f)h) = \tilde{U}(Tf)h, \quad f \in \mathfrak{D}(G), \quad h \in \mathcal{H}.$$

[elements of $\mathfrak{U}(\mathfrak{g})$ act on $\mathfrak{D}(G)$ as differential operators; see Appendix B.] Clearly, $\mathcal{K}^G$ is stable under all the operators $\tilde{U}(T)$: The Gårding domain is thus a natural domain for representing $\mathfrak{U}(\mathfrak{g})$.

Furthermore, because of theorems due to Segal and Nelson and Stinespring,[18] the following elements of

[14] G. Ludwig, *Die Grundlagen der Quantenmechanik* (Springer-Verlag, Berlin, 1954); A. Daneri, A. Loinger, and G. M. Prosperi, Nucl. Phys. **33**, 297 (1962).

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

[16] F. Riesz and B. Sz-Nagy, *Leçons d'analyse fonctionnelle* (Acad. Sci. Hongrie, Budapest, 1955), 2nd ed.

[17] L. Gårding, Proc. Natl. Acad. Sci. U.S.A., **33**, 331 (1947).

[18] I. E. Segal, Duke Math. J. **18**, 221 (1951); Proc. Am. Math. Soc. **3**, 13 (1952); E. Nelson and W. F. Stinespring, Am. J. Math. **81**, 547 (1959).

$\mathfrak{U}(\mathfrak{g})$ are represented on $\mathcal{K}^G$ by *essentially self-adjoint* operators and may thus define physical observables[19]:

(i) the elements of $\mathfrak{g}$;

(ii) the *symmetric central* elements of $\mathfrak{U}(\mathfrak{g})$ (i.e., the Casimir operators and the symmetric operators constructed from them);

(iii) the *symmetric elliptic*[20] elements of $\mathfrak{U}(\mathfrak{g})$ and those elements M of $\mathfrak{U}(\mathfrak{g})$ for which there exists an element $L \in \mathfrak{U}(\mathfrak{g})$, symmetric and elliptic, such that $\tilde{U}(M^+M)$ commutes with $\tilde{U}(L)$.

Moreover, all these properties still hold on the Gårding domain of a representation of a group G_1 that contains G as a subgroup. These results are broad enough to cover all the physically interesting cases. Therefore, $\mathcal{K}^G$ is a natural candidate for the domain D mentioned above. It is indeed a simple matter to endow the elements of $\mathcal{K}^G$ with a nuclear topology leading to a suitable space Φ. Example of such topologies have been given by Maurin.[21]

However, the space Φ constructed in this way would only be sufficient for describing the system completely if all l.o. were elements of $\mathfrak{g}$ or $\mathfrak{U}(\mathfrak{g})$. But the only system where this is true is the "rigid rotator." In other cases, this space is inadequate, for nothing proves that the other l.o.—let alone the position operators!—are continuous on Φ and leave it invariant. We need thus a more general solution.

Remarks:

(1) The same is true (*a priori*) for any space Φ built solely from the representation U; such are, for instance, spaces associated with the Gårding domain of a representation of a larger group G_1, or the set of all *regular vectors*, or the set of all *analytic vectors* [$h \in \mathcal{K}$ is called *regular* (*analytic*) if the function $g \to U(g)h$ is infinitely differentiable (analytic) on G].

(2) In the construction of the Gårding domain, U must be a *true* (vector) representation. In the case of a projective representation U_ω, $\mathcal{K}^G$ is stable under $U_\omega(g)$ if and only if the function $g \to \omega(g_1, g_1^{-1}g)$ is infinitely differentiable, which is not always true. Moreover, such a representation has no natural extension to $\mathfrak{U}(\mathfrak{g})$. For this reason, we deal only with true representations, supposing that the extension $G \to G_\omega$ has been accomplished (see the beginning of Sec. II).

[19] J.-C. Guillot, Helv. Phys. Acta **41**, 5 (1968).

[20] An abstract element of $\mathfrak{U}(\mathfrak{g})$ is called *elliptic* if it acts on $\mathfrak{D}(G)$ as an elliptic partial differential operator (Appendix B); such is, for instance, the second-order Casimir operator of the maximal compact subgroup of G.

[21] K. Maurin, Bull. Acad. Polon. Sci., Sér. Sci. Math., Astron., Phys. **7**, 471 (1959); **11**, 525 (1963); K. Maurin and L. Maurin, *ibid.* **13**, 199 (1965).

B. Extension of the Problem

In Sec. IIIA, we extended the representation U to the elements of $\mathfrak{g}$ and $\mathfrak{U}(\mathfrak{g})$. We shall now reformulate this extension in somewhat different terms, essentially following Maurin.[22]

The first step consists in identifying the Lie algebra with the set of all right-invariant[23] first-order differential operators on G (see Appendix B). Similarly, according to results of Harish-Chandra,[24] the enveloping algebra $\mathfrak{U}(\mathfrak{g})$ is isomorphic to the algebra of all right-invariant differential operators on G.

As a second step, we may identify invariant[10,25] differential operators on G with distributions on G with support $\{e\}$ (e is the unit element of G). Let us denote by $\mathcal{E}'_e$ the set of all distributions on G with support $\{e\}$. Then, given any $T \in \mathcal{E}'_e$, the convolution mapping $f \to T * f$ [$f \in \mathcal{E}(G)$],

$$[T * f](g) = \int_G f(g_1^{-1}g)\, dT(g_1),$$

is indeed a right-invariant differential operator on G. Conversely, any such operator can be obtained in that way.[10] To summarize, we have the canonical isomorphism

$$\mathfrak{U}(\mathfrak{g}) \simeq \mathcal{E}'_e. \tag{5}$$

This suggests that we consider the whole algebra $\mathcal{E}'(G)$ of distributions on G with compact support (it is an algebra with respect to convolution), of which $\mathcal{E}'_e$ is a subalgebra. Thus, we will extend to $\mathcal{E}'(G)$ the representation U in three successive steps.

(1) With respect to convolution and the involution defined by the relation

$$f^+(g) = \overline{f(g^{-1})},$$

$\mathfrak{D}(G)$ is a $*$-algebra. Then the mapping $f \to \tilde{U}(f)$ defined by Eq. (4) is a $*$-representation of $\mathfrak{D}(G)$ by bounded operators in $\mathcal{K}^G$:

$$\tilde{U}(f_1)\tilde{U}(f_2) = \tilde{U}(f_1 * f_2), \quad \text{for all } f_1, f_2 \in \mathfrak{D}(G),$$
$$\tilde{U}(f^+) = [\tilde{U}(f)]^+, \quad \text{for all } f \in \mathfrak{D}(G).$$

(2) Since the function (with values in $\mathcal{K}$) $g \to U(g)h$ is only continuous, the above definition can be extended (by duality) only to the space of measures with compact support on G, $\mathcal{M}(G)$:

$$\tilde{U}(\mu)h = \int_G U(g)h\, d\mu(g), \quad \mu \in \mathcal{M}(G).$$

[22] K. Maurin, Math. Ann. **165**, 204 (1966).

[23] We could as well use the *right*-invariant Haar measure on G and identify $\mathfrak{g}$ with *left*-invariant operators on G.

[24] See, for instance, S. Helgason, *Differential Geometry and Symmetric Spaces* (Academic Press Inc., New York, 1962).

[25] R. Godement, Trans. Am. Math. Soc. **73**, 496 (1952).

(3) For going over to $\mathcal{E}'(G)$, the function $g \to U(g)h$ must be infinitely differentiable, i.e., h must be a *regular vector* of U. Let $\mathcal{H}^\infty$ denote the (dense) manifold of all the regular vectors of U, ($\mathcal{H}^G \subset \mathcal{H}^\infty$). For any $T \in \mathcal{E}'(G)$, $a \in \mathcal{H}^\infty$, one defines

$$\tilde{U}(T)a = \int_G U(g)a \, dT(g) \tag{6a}$$

or, equivalently,[26] for any $h \in \mathcal{H}$,

$$(h, \tilde{U}(T)a) = \int_G (h, U(g)a) \, dT(g). \tag{6b}$$

Then, the properties of the operators $\tilde{U}(T)$ can be summarized in the following proposition:

Proposition 1 (Maurin):

(i) The mapping $T \to \tilde{U}(T)$ is a $*$-representation of the algebra $\mathcal{E}'(G)$ by operators in $\mathcal{H}^\infty$:

$$\tilde{U}(T)\tilde{U}(S) = \tilde{U}(T * S),$$
$$\tilde{U}(T^+) \subseteq [\tilde{U}(T)]^+, \quad \text{for all } T, S \in \mathcal{E}'(G), \tag{7}$$

where T^+ is defined by the relation

$$\langle T^+, f \rangle = \overline{\langle T, f^+ \rangle}, \quad \text{for all } f \in \mathcal{E}(G).$$

(ii) If $T = T^+$, $\tilde{U}(T)$ is a *Hermitian* operator.

(iii) If $T \in \mathcal{E}'_e$ satisfies any of the conditions of Segal–Nelson–Stinespring stated above in Sec. IIIA, $\tilde{U}(T)$ is *essentially self-adjoint.*

(iv) If this is the case, the self-adjoint operators $\overline{\tilde{U}(T_1)}$ and $\overline{\tilde{U}(T_2)}$ commute strongly (i.e., their spectral projections commute) if and only if $T_1 * T_2 = T_2 * T_1$.

(v) This representation is an extension of the representation U of G in $\mathcal{H}$, for

$$\tilde{U}(\delta_g) = U(g), \quad \text{for all } g \in G. \tag{8}$$

The proof of most of these assertions is immediate.[27]

C. General Solution

The solution given above with the Gårding domain of U was unsatisfactory because nothing could be said about those observables which do not belong to $\mathfrak{U}(\mathfrak{g})$ [or $\mathcal{E}'(G)$]. A better construction can be obtained as follows.[22]

Let Ψ be a dense subset of $\mathcal{H}$ on which a topology is imposed such as to make it a complete nuclear topological vector space continuously embedded into $\mathcal{H}$. The space Φ will then be defined as the set of finite linear combinations of vectors of the form $\tilde{U}(f)\psi$, with $f \in \mathcal{D}(G)$, $\psi \in \Psi$. More precisely, we consider the projective tensor product[3] $\mathcal{D}(G) \otimes_\pi \Psi$ and the mapping $u: \mathcal{D}(G) \otimes_\pi \Psi \to \mathcal{H}$ defined by

$$u(f \otimes \psi) = \tilde{U}(f)\psi.$$

Let N be the kernel of u, i.e., the linear span of

$$N_0 = \{f \otimes \psi : \tilde{U}(f)\psi = 0\}.$$

Since u is continuous, N is a closed subspace of $\mathcal{D}(G) \otimes_\pi \Psi$. Let Θ be the range of u, provided with the topology of the space $\mathcal{D}(G) \otimes_\pi \Psi/N$; we have thus a topological isomorphism $[u]$:

$$[u]: \mathcal{D}(G) \otimes_\pi \Psi/N \to \Theta \subset \mathcal{H}. \tag{9}$$

The space Φ is then defined as the completion of Θ in this topology. The operators $\tilde{U}(T)$, $T \in \mathcal{E}'(G)$ are defined on Θ as before and extended to Φ by continuity. Their properties are summarized in the following propositions:

Proposition 2:

(i) Φ is a complete nulcear topological vector space, dense in $\mathcal{H}$ and continuously embedded into $\mathcal{H}$; $\Phi \subset \mathcal{H} \subset \Phi'$ is thus a rigged Hilbert space.

(ii) For any $T \in \mathcal{E}'(G)$, $\tilde{U}(T)$ is a continuous mapping of Φ into itself; $T \to \tilde{U}(T)$ is thus a $*$-representation of $\mathcal{E}'(G)$ by continuous operators, satisfying assertions (i), (ii), and (v) of Proposition 1.

(iii) In particular, for any $g \in G$, $U(g)$ is a continuous mapping from Φ into itself, and $g \to U(g)$ is a continuous representation of G in Φ.

Proof:

(i) $\mathcal{D}(G)$ and Ψ being nuclear, so are their tensor product $\mathcal{D}(G) \otimes_\pi \Psi$ and its quotient by the closed subspace N. With the topology defined by the isomorphism $[u]$, Θ is thus a nuclear space, and the same holds for its completion Φ. Since, by construction, Θ is a dense subset of $\mathcal{H}^\infty$, it is dense in $\mathcal{H}$ and, thus, so is Φ. Since the embedding $\Psi \to \mathcal{H}$ is continuous, u and $[u]$ are continuous, and, therefore, so is the embedding $\Phi \to \mathcal{H}$.

(ii) Let $\varphi = \sum_i a_i \tilde{U}(f_i)\psi_i$ with $f_i \in \mathcal{D}(G)$, $\psi_i \in \Psi$. Then

$$\tilde{U}(T)\varphi = \sum_i a_i \tilde{U}(T)\tilde{U}(f_i)\psi_i$$
$$= \sum_i a_i \tilde{U}(T * f_i)\psi_i.$$

Since the convolution $(T, f_i) \to T * f_i$ is a bilinear *separately continuous* mapping from $\mathcal{E}'(G) \times \mathcal{D}(G)$ into $\mathcal{D}(G)$,[28] $\tilde{U}(T)$ is a continuous mapping from Θ into itself and thus, by continuity, from Φ into itself. The other properties follow immediately.

[26] L. Schwartz, J. Anal. Math. (Jérusalem) **4**, 88 (1954–55).

[27] We note that $\mathcal{H}^\infty$ has a natural topology (see Ref. 10) under which it is a Fréchet space with all the required invariance properties; unfortunately, this topology is not nuclear, so that we discard it.

[28] L. Schwartz, *Théorie des distributions* (Hermann & Cie., Paris, 1957–1959), Vols. I and II.

(iii) Since $U(g) = \tilde{U}(\delta_g)$, $U(g)$ is a continuous mapping from Φ into itself, and the function $g \to U(g)\varphi$ is continuous from G into Φ for any $\varphi \in \Phi$. Indeed, by construction, this reduces to the continuity of the mapping $g \to \delta_g * \varphi$ for all $\varphi \in \Phi$, which follows as above from the separate continuity of the convolution.

Remark: As stated in Proposition 1(iii), the operator $\tilde{U}(T)$, where $T \in \mathcal{E}'_e$ satisfies any of the conditions of Segal–Nelson–Stinespring,[18] is essentially self-adjoint on $\mathcal{H}^\infty$. But the restriction of $\tilde{U}(T)$ to Φ will not necessarily be essentially self-adjoint on Φ. Maurin's proof for $\mathcal{H}^\infty$ breaks down in the present case because of the arbitrariness of Ψ ($\mathcal{H}^\infty$ is canonical, Φ is not!). However, we do not need $\tilde{U}(T)$ to be essentially self-adjoint on Φ: In order to apply the generalized spectral theorem for an operator A, all we need is A to be self-adjoint in $\mathcal{H}$ (so that a unique spectral decomposition exists) and that its restriction to Φ maps Φ continuously into itself; this *is* true for $\tilde{U}(T)$ on Φ.

Proposition 3: Let $U = \oplus_i U_i$ be a decomposition of U in a direct sum of irreducible representations, $\mathcal{H} = \oplus_i \mathcal{H}_i$ the corresponding decomposition of $\mathcal{H}$ into closed invariant subspaces. Let $\Phi_i = \Phi \cap \mathcal{H}_i$. Then, for each i, Φ_i is a dense subset of $\mathcal{H}_i$, stable under U_i and closed in Φ; with the topology induced by Φ, Φ_i is nuclear.

Proof: $\Phi_i = \Phi \cap \mathcal{H}_i$ is a vector subspace of Φ, invariant under U_i and closed in the topology induced by $\mathcal{H}$ on Φ, thus *a fortiori* closed in the (finer) topology of Φ. For any $f \in \mathcal{D}(G)$ and any $\psi \in \Psi \cap \mathcal{H}_i$, $\tilde{U}(f)\psi$ belongs to Φ_i, since $\mathcal{H}_i$ is invariant under U, thus under $\tilde{U}(f)$; Φ_i is therefore a dense subset of $\mathcal{H}_i$. Finally, since Φ_i is a closed subspace of the nuclear space Φ, it is nuclear in the induced topology.

Remark: Bruhat[10] proved that the restriction of U to $\mathcal{H}^\infty$ (with the topology mentioned in Ref. 27) is topologically irreducible[29] *if and only if* U is. Proposition 3 says that the "only if" part of this theorem is true also for the restriction of U to Φ. But here, the converse is not necessarily true: Topological irreducibility of U does not imply topological irreducibility of $U \mid \Phi$, without supplementary information on the space Ψ. This again stems from the fact that Φ is not canonical.

[29] A representation of G in a space E is called *topologically irreducible* if there is no *closed* invariant subspace in E.

Proposition 4: Let $\{A_i, i = 1, 2, \cdots, N\}$ be a *finite* set of observables, such that

(i) the operators A_i transform irreducibly under U:

$$U^+(g)A_iU(g) = \sum_{j=1}^{N} d_{ij}(g)A_j, \quad i = 1, 2, \cdots, N; \quad (10)$$

(ii) each A_i maps Ψ continuously into itself.

Then, each of the A_i also maps Φ continuously into itself.

Proof: Let us note first that (10) is to be understood as a relation between operators in $\mathcal{H}$, $U(g)$ being bounded and A_i self-adjoint. This means that there exists in $\mathcal{H}$ a dense domain D such that (i) each A_i, $i = 1, \cdots, N$, is essentially self-adjoint on D; (ii) D is invariant under $U(g)$: $U(g)D \subset D$, $\forall g \in G$ (clearly $\Psi \subseteq D$). We need only consider the elements of Φ of the form $\varphi = \tilde{U}(f)\psi, f \in \mathcal{D}(G), \psi \in \Psi$. It follows then from (10) that φ is in the domain of each A_i:

$$\begin{aligned} A_i\varphi &= A_i\tilde{U}(f)\psi \\ &= \int_G dg f(g) A_i U(g)\psi \\ &= \sum_{j=1}^{N} \int_G dg f(g)\, d_{ij}(g) U(g) A_j \psi. \end{aligned}$$

We have interchanged A_i and the integral sign in the second equality; this is permissible since A_i is a closed operator.[30] Now the observables A_i transform under a *finite* N-dimensional representation of G (thus nonunitary if G is noncompact, unless $N = 1$); the mapping $g \to \|d_{ij}(g)\|$ is a continuous homomorphism of the analytic group G into the analytic group $GL(N, \mathbf{C})$ of complex regular $N \times N$ matrices; this homomorphism is therefore analytic,[31] i.e., the functions d_{ij} are analytic on G. Thus, we may write

$$A_i\varphi = \sum_{j=1}^{N} \tilde{U}(d_{ij}f)A_j\psi,$$

which proves the proposition, because multiplication by $d_{ij} \in \mathcal{E}(G)$ is continuous on $\mathcal{D}(G)$, and A_i is continuous from Ψ into Ψ.

Corollary: The theorem is also true for those observables which undergo an *inhomogeneous* transformation under U:

$$U^+(g)A_iU(g) = \sum_{j=1}^{N} d_{ij}(g)A_j + c_i(g), \quad i = 1, 2, \cdots, N.$$

[30] N. Dunford and J. T. Schwartz, *Linear Operators* (Interscience Publishers, Inc., New York, 1958), Vol. I, p. 153.

[31] C. Chevalley, *Theory of Lie Groups* (Princeton Univ. Press, Princeton, N.J., 1946).

This occurs, for instance, with the position variables under both the Lorentz and Poincaré groups (the variables which define the action of G as a transformation group).

Remark: The key word in Proposition 4 is "finite." The result may not be true for observables which have more general transformation properties, e.g.,

$$U^+(g)AU(g) = \int_X^{\oplus} R(g, x)A(x)\, d\mu(x),$$

where

$$A = \int_X^{\oplus} A(x)\, d\mu(x)$$

is a decomposition of A into a direct integral over some measure space $(X, d\mu)$. Such are, for example, the spin and helicity operators for the Poincaré group.[19]

With the preceding results, it now becomes fairly easy to build an appropriate space Φ for a system defined by its l.o. and its symmetry properties. First we build the space Ψ with those l.o. which do not belong to $\mathfrak{U}(\mathfrak{g})$, but transform under G according to (10). In general, there are only few of these observables, and in many cases the position operators will be the only ones.[32] All that remains then is to build Φ according to Proposition 2. Once this is done, the whole analysis of Part I can be applied and $\mathcal{H}$ can be decomposed in a direct integral with respect to any complete system of commuting l.o.

Of particular physical significance are, of course, those decompositions of $\mathcal{H}$ in which U is decomposed into irreducible subrepresentations.[9,33] But, obviously, the physical situation requires that this decompositon be unique up to equivalence, i.e., that G be a *type I group*. This class contains compact groups, Abelian groups, connected semisimple Lie groups, and also the universal covering groups of both the Poincaré and Galilei groups. (This is not trivial since it is not true, in general, that the universal covering group of a type I group is also of type I.)

Such a decomposition may be obtained in the standard way. One considers the von Neumann algebra $\mathfrak{U}$ generated by U: $\mathfrak{U} = \{U(g), g \in G\}''$ and its commutant $\mathfrak{U}'$.[34] Choosing any maximal Abelian subalgebra $\mathfrak{W}$ of $\mathfrak{U}'$, one decomposes $\mathcal{H}$ with respect to $\mathfrak{W}$ into irreducible components, corresponding to irreducible subrepresentations of U. More generally, if we consider only the Abelian subalgebra of $\mathfrak{U}'$ generated by the Casimir operators (which belong to the center of $\mathcal{E}'_e$ and are self-adjoint according to Segal's results[18]), we have a decomposition of $\mathcal{H}$ and U into *primary* components. But a primary representation of G (which is of type I) is a direct sum of copies of the same irreducible representation, labeled by the values of the Casimir operators corresponding to that component (Schur's lemma). If G is compact, the Casimir operators have a discrete spectrum, and U decomposes into a direct sum of irreducible representations (each of them might appear several times). If G is *noncompact*, each component $U(\lambda)$ in the direct integral is a unitary representation in the corresponding space $\mathcal{H}(\lambda)$, and a direct sum of copies of a single irreducible representation.[35] Similarly, the elements of $\mathfrak{g}$ are decomposed and the commutation relations defining $\mathfrak{g}$ hold in each $\mathcal{H}(\lambda)$ separately.

IV. APPLICATION TO INDUCED REPRESENTATIONS

A. General Induced Representations

The method described in the preceding section was designed to cope with those l.o. which do not stem from the symmetry group. But it may also be used for studying the representations of $\mathfrak{U}(\mathfrak{g})$ or $\mathcal{E}'(G)$ themselves. In particular, we shall now consider in more detail the case of an induced representation in the sense of Wigner,[6] Mackey,[9] and Bruhat.[10] It is indeed well known that most of the physically interesting representations are of that kind.

We shall not reproduce here the well-known theory of induced representations,[9] but only recall the basic definitions following the notation of a recent paper by Guillot and Petit.[36] Let G be a Lie group which we will suppose *unimodular*[37] for simplicity. (The general case can be treated in exactly the same way, but the equations are much heavier! This includes already all compact or semisimple Lie groups, plus the Euclidean, the Galilei, and the Poincaré groups.) Let K be a closed subgroup of G, L a unitary representation of K in a Hilbert space $\mathcal{H}_L$. The representation $U = U^L$ of G induced by L is then defined as follows. The representation space $\mathcal{H} \equiv \mathcal{H}_{UL}$ is the set of functions $h: G \to \mathcal{H}_L$ which satisfy the following three conditions:

[32] Clearly, the present method is not applicable if there exist l.o. which neither belong to $\mathfrak{U}(\mathfrak{g})$ nor satisfy (10). The work of Guillot (see Ref. 19) shows, however, that the difficulty can be circumvented in the case of induced representations. This is developed in Sec. IV.

[33] G. A. Pozzi, Suppl. Nuovo Cimento **4**, 37 (1966).

[34] J. Dixmier, *Les algèbres d'opérateurs dans l'espace hilbertien (Algèbres de von Neumann)* (Gauthier-Villars, Paris, 1957). See also Part I.

[35] F. Mautner, Ann. Math. **52**, 528 (1950); Proc. Am. Math. Soc. **2**, 490 (1951).

[36] J.-C. Guillot and J.-L. Petit, Helv. Phys. Acta **39**, 281 (1966).

[37] A locally compact group is called *unimodular* if its Haar measure is both left- and right-invariant.

For any $v \in \mathcal{H}_L$, the function

$$g \to (h(g), v) \text{ is measurable;} \tag{11a}$$

covariance:

$$h(gk) = L^+(k)h(g), \quad \text{for all} \quad k \in K, \quad g \in G; \tag{11b}$$

$$\|h\|^2 = \int_{G/K} d\mu(x)\, \|h(g)\|^2_{\mathcal{H}_L}. \tag{11c}$$

In condition (11c), μ is the invariant measure on the coset space G/K (it is unique up to a constant), and the integrand is the squared norm of h in $\mathcal{H}_L$. The covariance condition (11b) ensures that this squared norm is constant on the cosets gK, i.e., is a function on G/K.

The space $\mathcal{H}$ is then a separable Hilbert space, with the scalar product

$$(h_1, h_2) = \int_{G/K} d\mu(x)(h_1(g), h_2(g))_{\mathcal{H}_L}.$$

In this space, the induced representation $U^L \equiv U$ is defined by left translations and is obviously unitary:

$$(U^L(g_0)h)(g) = h(g_0^{-1}g). \tag{12}$$

A second (and more common) form of this representation can be given. $\mathcal{H}$ is indeed isomorphic to the Hilbert space $L^2_\mu(G/K, \mathcal{H}_L)$ of μ-square-integrable functions on G/K with values in $\mathcal{H}_L$. The isomorphism is given by the following correspondence:

$$h(g) = L^+(\Lambda_x^{-1}g)\hat{h}(x), \quad \hat{h} \in L^2_\mu(G/K, \mathcal{H}_L), \quad h \in \mathcal{H}; \tag{13}$$

and the representation U in the new space is given by

$$(U^L(g)\hat{h})(x) = L(\Lambda_x^{-1}g\Lambda_{g^{-1}x})\hat{h}(g^{-1}x). \tag{14}$$

In these equations, $x = gK$ is the class of g in G/K and Λ_x is an *arbitrary* element of that class, such that Λ_x transforms $x_0 \equiv eK$ into x:

$$\Lambda_x : x_0 \to x;$$

thus, by definition,

$$\Lambda_x^{-1}g \in K.$$

According to the general scheme outlined in Sec. III, we need a nuclear space Ψ dense in the representation space of U^L. Let us take first the space $\mathcal{H}$ defined by (11). A natural candidate for Ψ would be the set of infinitely differentiable functions on G with compact support, satisfying (11). However, such functions will satisfy condition (11b) only if

(i) L is a differentiable representation, i.e., the function $k \to L(k)h$ is a C^∞ mapping of K into $\mathcal{H}_L$ for any $h \in \mathcal{H}_L$;

(ii) K is compact, for the support of any function satisfying (11b) must contain at least one coset (if $L \equiv 1$, for instance, a function can be constant on the cosets only if its support is the union of a certain number of cosets).

Since K is not necessarily compact (indeed, in the most important examples of Sec. V, it is not), we cannot limit ourselves to functions of compact support. Instead, we will require that $\operatorname{supp} f$ be contained in a set of the form AK, where A is a compact subset of G. More precisely, Ψ will be the space $\mathcal{D}^{(L)}(G, \mathcal{H}_L)$ of C^∞ functions $h: G \to \mathcal{H}_L$, whose support has a compact canonical image in G/K and which satisfy condition (11b). This space has been introduced by Bruhat[10] and can be endowed with a topology in the usual way (see Appendix B). It is nuclear if and only if $\mathcal{H}_L$ is nuclear.

Passing now to the second form (14) of U^L, in $L^2_\mu(G/K; \mathcal{H}_L)$, we see immediately that (13) is a topological isomorphism of $\mathcal{D}^{(L)}(G, \mathcal{H}_L)$ onto $\mathcal{D}(G/K, \mathcal{H}_L)$, the space of C^∞ functions $f: G/K \to \mathcal{H}_L$ with compact support, provided that

(i) L is differentiable;

(ii) the function $x \to \Lambda_x$ is also C^∞.

From now on we shall suppose that L is a differentiable representation (we will discuss the implications of this later). But condition (ii) on Λ_x is nontrivial and will bring difficulties in the case of the Poincaré and the Galilei groups. When the condition is fulfilled, we can use indifferently either of the two forms of U^L and take $\Psi = \mathcal{D}^{(L)}(G, \mathcal{H}_L)$ or $\Psi = \mathcal{D}(G/K, \mathcal{H}_L)$. If it is not, only the first form is allowed. However, $\mathcal{D}^{(L)}(G, \mathcal{H}_L)$ is *always* isomorphic to the quotient of $\mathcal{D}(G, \mathcal{H}_L)$ by a closed subspace[10] (see Appendix B).

In order to go further, we must now distinguish two cases, according to whether L is finite-dimensional or not.

Case 1: $\dim L < \infty$. This case is straightforward: If L is finite-dimensional and continuous, it is obviously differentiable (even analytic![31]) and $\mathcal{H}_L$ is trivially nuclear. Therefore, both $\mathcal{D}^{(L)}(G, \mathcal{H}_L)$ and $\mathcal{D}(G/K, \mathcal{H}_L)$ are nuclear spaces. We may now construct the corresponding space Φ as indicated in Sec. III.

Taking first the general form of U^L, we find immediately, for any $f \in \mathcal{D}(G)$,

$$[\tilde{U}(f)\psi](g) = [f * \psi](g), \quad \text{for all} \quad \psi \in \mathcal{D}^{(L)}(G, \mathcal{H}_L). \tag{15}$$

It follows that $\tilde{U}(f)\psi \in \mathcal{D}^{(L)}(G, \mathcal{H}_L)$ and furthermore, $\tilde{U}(f)\psi$ can vanish identically only if either f or ψ does; in other words, Φ is a subset of $\mathcal{D}^{(L)}(G, \mathcal{H}_L)$.[38] It will

[38] More precisely, Φ is the (dense) subset of $\mathcal{D}^{(L)}(G, \mathcal{H}_L)$ spanned by the elements $f * \psi$, $f \in \mathcal{D}(G)$, $\psi \in \mathcal{D}^{(L)}(G, \mathcal{H}_L)$, with the topology of $\mathcal{D}(G) \otimes_\pi \mathcal{D}^{(L)}(G, \mathcal{H}_L)$.

therefore be simpler to take for Φ the whole space $\mathcal{D}^{(L)}(G, \mathcal{H}_L)$ itself. On the latter space, we have indeed the usual representation of $\mathcal{E}'(G)$ by continuous operators:

$$[\tilde{U}(T)\varphi](g) = [T * \varphi](g),$$
$$T \in \mathcal{E}'(G), \quad \varphi \in \mathcal{D}^{(L)}(G, \mathcal{H}_L), \quad (16)$$

and all the results of Propositions 1–4 are valid, *including* the essentially self-adjoint character of the relevant operators. In particular, for any $T \in \mathcal{E}'_e$, $\tilde{U}(T)$ is a right-invariant differential operator on G.

If the function $x \to \Lambda_x$ is C^∞, the second form of U^L is available; $\mathcal{D}^{(L)}(G, \mathcal{H}_L)$ is then topologically isomorphic to $\mathcal{D}(G/K, \mathcal{H}_L)$, which we will take as our space Φ as before. In that case, for any $T \in \mathcal{E}'_e$, $\tilde{U}(T)$ will be a right-invariant differential operator on G/K.

In either case, the corresponding dual space Φ' will be a space of vector-valued distributions.[10,39] More precisely, as $\mathcal{D}^{(L)}(G, \mathcal{H}_L)$ is the quotient of $\mathcal{D}(G, \mathcal{H}_L)$ by a closed subspace, say $\mathcal{F}_L$, its dual is the closed subspace of $[\mathcal{D}(G, \mathcal{H}_L)]'$ orthogonal to $\mathcal{F}_L$. Furthermore, $[\mathcal{D}(G, \mathcal{H}_L)]'$ can be identified[10] with

$$\mathcal{L}(\mathcal{D}(G), \mathcal{H}_L) \equiv \mathcal{D}'(G, \mathcal{H}_L),$$

the space of distributions on G with values in $\mathcal{H}_L$. Thus, Φ' is in general a closed subspace of $\mathcal{D}'(G, \mathcal{H}_L)$. Similarly, if $\Phi = \mathcal{D}(G/K, \mathcal{H}_L)$, then $\Phi' = \mathcal{D}'(G/K, \mathcal{H}_L)$, the space of distributions on G/K with values in $\mathcal{H}_L$. We shall see in the examples that the latter space, despite its complicated appearance, is nothing but the vector space spanned by the "improper basis vectors" commonly used in representation theory.

Case 2: dim $L = \infty$.

(a) If L is an infinite *direct sum* of finite-dimensional representations, then, because of Proposition 3, we are in the same situation as in Case 1.

(b) In any other case, e.g., L irreducible, or L equal to a *direct integral* of (in)finite-dimensional representations, we are in trouble. Indeed, the representation L will usually not be differentiable, and the space $\mathcal{D}^{(L)}(G, \mathcal{H}_L)$ is no longer nuclear. This is, in fact, exactly the situation we started with; we will therefore proceed in the same way. We restrict ourselves to a dense subspace Ψ_L of $\mathcal{H}_L$, consisting of differentiable vectors for L (i.e., $\Psi_L \subseteq \mathcal{H}_L^\infty$) and provided with a nuclear topology (such a space Ψ_L can be obtained just as before if L is itself an induced representation). Then we take $\Psi = \mathcal{D}^{(L)}(G, \Psi_L)$ and once more, $\Phi = \Psi$, the operators $\tilde{U}(f)$, $\tilde{U}(T)$ of Eqs. (15), (16) acting again by convolution. Similarly, if a C^∞ function $x \to \Lambda_x$ is available, we may also take $\Phi = \Psi = \mathcal{D}(G/K, \Psi_L)$. In fact, the whole analysis of Case 1 can be repeated here, replacing $\mathcal{H}_L$ by Ψ_L. In particular, the dual space Φ' will consist again of vector-valued distributions, this time distributions defined on G or G/K, with values in Ψ'_L, the dual of Ψ_L. Further details may be found in the work of Bruhat.[10]

For the sake of completeness, we note finally that L and U^L have always the *same* decomposition into primary constituents[9]: the decomposition

$$L = \int_\Sigma^\oplus d\nu(\sigma) L^{(\sigma)} \quad (17a)$$

implies

$$U^L = \int_\Sigma^\oplus d\nu(\sigma) U^{L^{(\sigma)}}; \quad (17b)$$

in these formulas, Σ is a set of primary representations of K (i.e., a subset of $\hat{K}$, the dual of K); ν is an arbitrary measure on Σ, necessarily discrete if K is compact (since $\hat{K}$ and thus also Σ are then discrete sets). The preceding discussion can be adapted to this case, and Eqs. (11)–(16) are now valid in each component $\{\sigma\}$ separately [convolution commutes with the decomposition (17)].

B. The Case of a Semidirect Product $G = T \boxtimes \Delta$, T Abelian

This case has a particular importance for physical applications; all the examples of Sec. V are of this form. When $G = T \boxtimes \Delta$, a semidirect product of an Abelian invariant subgroup T and a subgroup $\Delta \simeq G/T$, the original method of Wigner[6] can be applied, and it yields[9] all the unitary irreducible representations of G as induced representations (provided G satisfies a mild regularity condition, which is known to be true in most interesting cases; we take it for granted here). Equation (17) then gives the most general unitary representation of G.

Thus, we consider first unitary irreducible representations (UIR) of G. T is an Abelian group; therefore it has only 1-dimensional UIR's (characters). Then, under the action of G (defined by duality), the space of characters $\hat{T}$ decomposes into disjoint *orbits* Ω_m (labeled by m), on each of which G acts transitively. Following Wigner,[6] we pick a character $p_0 \in \Omega_m$ and consider the subgroup K of G which leaves this vector invariant; K has the following form:

$$K = T \boxtimes \Delta_{p_0},$$

where Δ_{p_0}, the *little group* of p_0, is a closed subgroup

[39] L. Schwartz, Ann. Inst. Fourier (Grenoble) **7**, 1 (1957); **8**, 1 (1959).

of Δ. We have then

$$\frac{G}{K} = \frac{T \boxtimes \Delta}{T \boxtimes \Delta_{p_0}} \simeq \frac{\Delta}{\Delta_{p_0}} \simeq \Omega_m .$$

If we pick now an UIR $D^{(\lambda)}$ of Δ_{p_0}, we obtain the most general UIR of G, denoted by $U^{(m,\lambda)}$, in the following standard form, which corresponds to Eq. (14). The representation space is $L^2_\mu(\Omega_m, \mathcal{H}_{D^{(\lambda)}})$, where μ is the unique invariant measure on Ω_m, and the action of $U^{(m,\lambda)}$ is given by

$$[U^{(m,\lambda)}(a, A)h](p) = \langle p, a\rangle D^{(\lambda)}(\Lambda_p^{-1} A \Lambda_{A^{-1}p}) h(A^{-1}p),$$
$$\text{where} \quad a \in T, \quad A \in \Delta, \quad p \in \Omega_m \subseteq \hat{T}. \quad (18)$$

This representation $U^{(m,\lambda)}$ is thus induced by the UIR $L^{(m,\lambda)} \equiv \langle p, \cdot\rangle D^{(\lambda)}$ of K, and the latter has the same dimension as the UIR $D^{(\lambda)}$ of $\Delta_{p_0} \equiv \Delta_m$ (the little groups Δ_p are isomorphic for all $p \in \Omega_m$).

From the UIR $U^{(m,\lambda)}$, the most general unitary representation of G is then obtained according to Eq. (17):

$$U = \int^{\oplus} d\rho(m) U^{(m)}, \quad (19a)$$

with

$$U^{(m)} = \int_{\hat{\Delta}_m}^{\oplus} d\nu_m(\lambda) U^{(m,\lambda)}. \quad (19b)$$

In these equations, ν_m is an arbitrary measure on the set $\hat{\Delta}_m$ of all UIR's of the little group Δ_m. The most general unitary representation corresponding to (or carried by) the orbit Ω_m thus is $U^{(m)} \equiv U^{L^{(m)}}$, where

$$L^{(m)} = \int_{\hat{\Delta}_m}^{\oplus} d\nu_m(\lambda) L^{(m,\lambda)}$$

is the most general unitary representation of Δ_m. ρ is an arbitrary measure on the set of all orbits. Equations (19) allow us to sharpen the discussion of Sec. IV.A, about the dimension of L.

1. A Single Orbit Ω_{m_0}

(i) Δ_{m_0} *compact:* Then ν_{m_0} is a discrete measure and all the UIR $D^{(\lambda)}$ are finite-dimensional; we are thus in the favorable cases 1 or 2a;

(ii) Δ_{m_0} *noncompact:* All the $D^{(\lambda)}$ are infinite-dimensional, except the identity representation $D^{(0)}$. Then:

a. If $\nu_{m_0}(\lambda) = n\delta(\lambda)$ (i.e., $D^{(\lambda)}$ is the direct sum of n copies of $D^{(0)}$: $D^{(\lambda)} \equiv nD^{(0)}$), we are again in cases 1 or 2a;

b. For any other ν_{m_0}, we are in case 2b.

2. A Discrete Set of Orbits Ω_{m_i}, $i = 1, 2, \cdots$, (ρ Discrete)

Obviously, $U \equiv \oplus_i U^{(m_i)}$ will be in the favorable case if and only if each $U^{(m_i)}$ is.

3. A Continuum of Orbits (ρ Continuous)

Then we are always in case 2b.

V. SOME PHYSICAL EXAMPLES

A. The Euclidean Group $E(3)$

By $E(3)$ we mean the universal covering group of the group of all rotations and translations in $\mathbf{R}^3$, which has the following structure:

$$E(3) = T_3 \boxtimes SU(2), \quad T_3 \equiv \mathbf{R}^3.$$

This group has been studied extensively by Wightman[40] in his celebrated paper on localizability. We have only to see how his results fit into the general scheme.

The representations of $E(3)$ may be induced by two different subgroups, $SU(2)$ and $T_3 \boxtimes SO(2)$.

1. $K = SU(2)$

This corresponds to the $\{x\}$ representation, with

$$\frac{G}{K} = \frac{T_3 \boxtimes SU(2)}{SU(2)} \simeq T_3 \equiv \mathbf{R}^3,$$

and $E(3)$ acts transitively on $\mathbf{R}^3$ (unique orbit). The unitary representations of $E(3)$ have the standard form:

$$[U(\mathbf{a}, A)f]_m(\mathbf{x}) = \sum_{m'=-j}^{j} D^{(j)}_{mm'}(A) f_{m'}(A^{-1}(\mathbf{x} - \mathbf{a})),$$
$$\mathbf{a} \in T_3, \quad A \in SU(2), \quad f \in L^2(\mathbf{R}^3, \mathcal{H}_{2j+1}), \quad (20)$$

where $D^{(j)}$ is the $(2j+1)$-dimensional representation of $SU(2)$, acting in $\mathcal{H}_{2j+1}$. The general scheme of Sec. IV would yield $\Phi = \mathcal{D}(\mathbf{R}^3, \mathcal{H}_{2j+1})$. But here we see that other solutions are possible. We could start, for example, with $\Psi = \mathcal{S}(\mathbf{R}^3, \mathcal{H}_{2j+1})$ and obtain $\Phi = \Psi$ as an acceptable solution[41] (the space of fast-decreasing C^∞ functions on $\mathbf{R}^3$ with values in $\mathcal{H}_{2j+1}$). In this case, of course, we would just be studying rotation and translation properties of tempered distributions!

2. $K = T_3 \boxtimes SO(2)$

This corresponds to the general method outlined above, since $SO(2)$ is indeed the little group of any nonzero vector of $\mathbf{R}^3 \equiv \hat{T}_3$. Thus, we have

$$\frac{G}{K} = \frac{T_3 \boxtimes SU(2)}{T_3 \boxtimes SO(2)} \simeq S_2, \quad \text{the unit sphere in } \mathbf{R}^3.$$

This choice of K corresponds (as in the general case) to the $\{p\}$ representation (the translation subgroup is diagonalized). The orbits are all the spheres centered

[40] A. S. Wightman, Rev. Mod. Phys. **34**, 845 (1962).

[41] This is a general feature: For any admissible space Ψ which is invariant under convolution with $\mathcal{D}(G)$, we may take $\Phi = \Psi'$, as we did above with $\mathcal{D}^{(L)}(G, \mathcal{H}_L)$ and $\mathcal{D}(G/K, \mathcal{H}_L)$.

at the origin, and they are compact, i.e., $\mathfrak{D}(\Omega_m) \equiv \mathcal{E}(\Omega_m)$ is a Fréchet space. For further details, we refer to Wightman's article.[40]

B. The Galilei Group

The representations of the Galilei group have been studied by numerous authors[7,42]; we must recall only the main points of the analysis.

(i) G, as well as its extensions G_M (see Sec. II),[43] has the structure of a semidirect product:

$$G = T \boxtimes \Delta,$$

where T is the invariant Abelian subgroup of space–time translations, Δ the subgroup of rotations and pure Galilei transformations.

(ii) In the dual space $\hat{T}$ (with variables E, $\mathbf{p}$), the orbits are the paraboloids Ω_{E_0}:

$$E = E_0 + \mathbf{p}^2/2M,$$

labeled by the parameter E_0; the invariant measure on Ω_{E_0} is

$$d\sigma_{E_0}(E, \mathbf{p}) = dE\, d\mathbf{p}\delta(E - E_0 - \mathbf{p}^2/2M).$$

(iii) For any orbit, the little group is $SU(2)$; thus, we take $K = T \boxtimes SU(2)$, and the general theory of Mackey ensures that any UIR of G ($M \neq 0$), denoted $[M \mid E_0, s]$, is fixed by the choice of the orbit Ω_{E_0} and the $(2s + 1)$-dimensional UIR $D^{(s)}$ of $SU(2)$:

$$[U(a, \Gamma)f]_\sigma(p) = e^{i(Eb - \mathbf{p}\cdot\mathbf{a})} \sum_{\sigma'=-s}^{s} D^{(s)}_{\sigma\sigma'}(R) f_{\sigma'}(\Gamma^{-1}p), \quad (21)$$

where

$p \equiv (E, \mathbf{p}) \in \Omega_{E_0}$,
$a \equiv (b, \mathbf{a}) \in T$, space–time translation,
$\Gamma \equiv (\mathbf{v}, R) \in \Delta$, Galilei acceleration and rotation.

(iv) The most general (projective) representation of G corresponding to the value M of the mass is given by

$$U = \int^{\oplus} d\rho(E_0) \bigoplus_s \nu_{E_0}(s)[M \mid E_0, s], \quad (22)$$

where ρ is an arbitrary measure on $\mathbf{R}$ and $\nu_{E_0}(s)$ is the multiplicity of the representation $D^{(s)}$. In Eqs. (21) and (22), the variables have the obvious physical significance when those equations are taken to describe a single, isolated, Galilei invariant system:

M = total mass,
$\mathbf{p}$ = total momentum (or momentum of the center of mass),
E_0 = total rest energy (or internal energy),
E = total energy,
s = total intrinsic (or internal) angular momentum, i.e., spin.

The corresponding observables[44] are denoted M, $\mathbf{P}$, H_0, H and $\mathbf{S}$, respectively.

Let us go back to our construction: The orbits Ω_{E_0} are noncompact, but the little groups Δ_{E_0} are all isomorphic to $SU(2)$ and are thus compact. From this, it follows that no complication arises if we consider only a single orbit or a discrete set of orbits: The general theory gives us $\Phi = \mathfrak{D}^{(s)}(G_M, \mathcal{K}_{2s+1})$ or $\mathfrak{D}(\Omega_{E_0}, \mathcal{K}_{2s+1})$, or a direct sum of such spaces, which are all nuclear. Here again we see that other solutions are available, such as the space $\mathcal{E}(\Omega_{E_0}, \mathcal{K}_{2s+1})$ considered by Guillot.[19] If, however, we want to include a continuum of orbits, then we need a more refined space: Its elements are functions of E_0, and we must add some requirement on this dependence (such as compact support and infinite differentiability; this may depend on the problem at hand).

In any case, Eq. (22) is the (unique) decomposition of U into primary components, i.e., its decomposition with respect to the Casimir operators H_0 and $\mathbf{S}^2$. According to the general discussion at the end of Sec. III, the complete (spectral) decomposition of $\mathcal{K}$ implies the choice of a complete system of commuting labeled observables, including H_0 and $\mathbf{S}^2$. The other l.o. are of two types:

(1) those l.o. which fix the particular realization of the UIR $[M \mid E_0, s]$; this includes:

(i) the generators $\mathbf{P}$, H of the Abelian translation subgroup, whose diagonalization is the essence of the general method outlined in Sec. IVB;

(ii) the l.o. (call it B), implied by the choice of a particular function $p \to \Lambda_p$ defined in Sec. IVA. We will come back to this problem in a moment.

(2) the supplementary observables A_i which lift the degeneracy in the case $\nu_{E_0}(s) > 1$; we can say at once that these operators must have a *purely discrete spectrum*. Indeed, as G is a type I group, the primary representation $U^{(E_0, s)}$ can be at most a direct *sum* of copies of the UIR $[M \mid E_0, s]$. In most situations, as

[42] J.-M. Lévy-Leblond, J. Math. Phys. **4**, 776 (1963); J. Voisin, J. Math. Phys. **6**, 1519, 1822 (1965).

[43] G_M is an extension of G by the one-parameter group of the reals (see Ref. 7); the parameter M, corresponding to the factor ω_M of G, is identified with the total mass of the system and gives rise to Bargmann's superselection rule.

[44] We do not study here the implications of the symmetry on the choice of the labeled observables of the system. Clearly, all the observables mentioned here are l.o.—all contained in the Lie algebra or its enveloping algebra. We shall come back to this very important point in another publication.

for instance in a 3-body problem, those l.o. A_i are relative angular momenta.[45]

To sum up, the decomposition performed so far is the spectral decomposition of $\mathcal{K}$ with respect to the complete set of commuting l.o. $\{H_0, \mathbf{S}^2; \mathbf{P}, H; B; A_i\}$. According to the general theory of Part I, there exists in the dual space Φ' a complete orthonormal set of eigenfunctionals $\{\langle E_0, s; \mathbf{p}, E; b; a_i|\}$: this is the so-called "improper basis" used in many instances (see for instance the work of Voisin[42]). The theory developed so far tells us that it is the same to study the representation U in Φ or the contragredient representation $\check{U}$ in Φ', more specifically, the transformation properties of those "improper basis vectors" under $\check{U}$. This latter procedure, which is most commonly used, is especially convenient also for discussing tensor products of representations and their Clebsch–Gordan decomposition[42]; clearly a CG coefficient is a particular functional over the product space $\Phi_1 \otimes \Phi_2$ (with an adequate topology); this point of view has been considered by Rideau[46] and Guillot.[19]

Finally, let us discuss the choices of the function $p \to \Lambda_p$; this problem has been studied in considerable detail by Voisin[42] and Guillot and Petit.[19,36] These authors consider the following cases:

(i) *The canonical formalism:* Λ_p is the pure Galilei transformation $(0, 0, \mathbf{p}/M, 1)$; clearly this Λ_p depends analytically on p and all the conditions discussed in Sec. IVA are fulfilled; in the canonical formalism, we may thus use either of the two forms (12) or (14) for the representation. In this case, $B \equiv S_Z$. [Equation (21) is written in this formalism.]

(ii) *The helicity formalism:* Λ_p is the composite transformation $(0, 0, \mathbf{p}/M, R_{\mathbf{p}})$, where $R_{\mathbf{p}}$ is the smallest rotation bringing $\mathbf{p}$ onto a fixed vector $\mathbf{k}$; but this rotation is undefined for $\mathbf{p} = 0$ or $\mathbf{p} = -\mathbf{k}$. Consequently, with this choice, Λ_p does not even depend continuously on p, so that the second form of the representation breaks down. This is linked to the difficulties encountered by Guillot[19] in defining correctly the helicity operator $\Sigma = \mathbf{p} \cdot \mathbf{S}/|\mathbf{p}|$. The name of the formalism stems from the fact that $B \equiv \Sigma$ in this case.

C. The Poincaré Group

This last group, with its well-known structure $\mathcal{P} = T \boxtimes L$ (T, translation group; L, Lorentz group), has been studied in so many places that it is hardly necessary to add any comment. We will, therefore, be content to mention the following few points:

(i) The structure of $\mathcal{P}$ and of its UIR's is exactly the same as that of the Galilei group; we note, however, that only the orbits Ω_m $(m^2 > 0)$ have a compact little group $SU(2)$; all other orbits have noncompact little groups and so the difficulties outlined in Sec. VB are to be expected. We might also mention that the systematic use of "a continuum of orbits" has been suggested recently by Lurçat[47]: This author describes unstable particles by a *multiplicity-free* representation of $\mathcal{P}$, i.e., a representation with a continuous measure $\rho(m)$ and no multiplicity, $\nu_m(s) = \delta_{ss_0}$. [Here, m plays the role of E_0 in G_M, in Eq. (22).]

(ii) The "improper bases" are used in the case of $\mathcal{P}$ exactly as for G_M; they can be justified in the present context in the same way, by constructing a suitable space Φ and taking its dual Φ'. The same is true for the CG coefficients, as studied for instance by Moussa and Stora.[48]

(iii) The same choices of the function $p \to \Lambda_p$ may be done here, leading again to the *canonical* and *helicity formalisms*, respectively, and the same difficulty appears with the latter. Another difficulty already appears in this case with the *spin operators*,[19] because here they have more complicated transformation properties (see the Remark after Proposition 4); the *helicity operator* behaves as badly in $\mathcal{P}$ as it does in G_M.

ACKNOWLEDGMENTS

It is a great pleasure to acknowledge the continuous help and guidance of Professor D. Speiser during the realization of this work; his constructive criticisms have much improved the manuscript. Warm thanks are due also to Professor V. Bargmann, who read the manuscript, to Dr. M. Reed for linguistic advice, and to Professor F. Cerulus, Professor A. Grossmann, Professor E. P. Wigner, and Dr. J.-C. Guillot for enlightening discussions. Finally, we gratefully acknowledge the hospitality of Professor R. H. Dicke at Palmer Laboratory, where the work was completed, as well as a Fulbright–Hays Award.

APPENDIX A: THE DUALITY $U \leftrightarrow \check{U}$ AND PARITY NONCONSERVATION

Given the representation U in Φ, the contragredient representation $\check{U}$ in Φ' was defined in Sec. II by the relation

$$\langle \check{U}(g)\varphi' \mid U(g)\varphi \rangle = \langle \varphi' \mid \varphi \rangle, \quad \text{for all} \quad \varphi \in \Phi,\ \varphi' \in \Phi'. \tag{A1}$$

[45] A. Dragt, J. Math. Phys. **6**, 533 (1965); J.-M. Lévy-Leblond and F. Lurçat, *ibid.* **6**, 1564 (1965).

[46] G. Rideau, Ann. Inst. Henri Poincaré **3**, 339 (1965).

[47] F. Lurçat, Phys. Rev. **173**, 1461 (1968).

[48] P. Moussa and R. Stora, in *Lectures in Theoretical Physics*, W. E. Brittin and A. O. Barut, Eds. (University of Colorado Press, Boulder, Colo., 1965), Vol. VII.A.

It was emphasized that this relation is a mere *definition*, and it is therefore independent of the invariance or noninvariance of a particular process under G. Equation (A1) *always* holds.

As an illustration, we will consider here the β decay of a polarized nucleus, $\mathcal{N} \to \mathcal{N}' + e^- + \bar{\nu}$. For simplicity, we take the nucleus $\mathcal{N}$ at rest at the origin, with its spin completely aligned on the negative z axis. Let $\varphi_- \in \Phi$ be the state vector of the nucleus. The final state, after the decay, is then $S\varphi_-$, where S is the appropriate "S matrix" describing the decay. Clearly, this vector must belong to Φ, since the final state is actually realized; in other words, S must map Φ continuously onto itself in order to have a consistent theory.

Let θ be the angle between the positive z axis and the momentum of the emitted electron. Then the angular distribution of the electron is given by the following amplitude:

$$f_-(\theta) = \langle C_\theta \mid S\varphi_- \rangle,$$

where $C_\theta \in \Phi'$ represents an "elementary" (or idealized) counter placed at an angle θ.

Consider now the operation of space reflection, represented in $\mathcal{H}$ by the unitary operator U_P. According to the general scheme, Φ is invariant under U_P and (A1) defines the contragredient operator $\check{U}_P$ in Φ'. Thus, we have

$$\begin{aligned} f_-(\theta) &= \langle C_\theta \mid S\varphi_- \rangle \\ &= \langle \check{U}_P C_\theta \mid U_P S\varphi_- \rangle \\ &= \langle C_{\pi-\theta} \mid U_P S\varphi_- \rangle. \end{aligned}$$

If parity is conserved in the decay, U_P commutes with S, and since $U_P\varphi_- = \varphi_-$, we obtain, as usual,

$$f_-(\theta) = f_-(\pi - \theta).$$

But if parity is not conserved, $[U_P, S] \neq 0$, then we get

$$f_-(\theta) \neq f_-(\pi - \theta),$$

i.e., the familiar up-down asymmetry characteristic of parity-nonconserving decays. Thus we see that relation (A1) does not contradict the violation of parity in this particular process.

APPENDIX B: DISTRIBUTIONS OVER A LIE GROUP

1. Functional Spaces

Let G be a unimodular Lie group, dg its Haar measure, which is both left- and right-invariant, and unique up to normalization. Thus, G is an analytic manifold, which we suppose *countable at infinity*, i.e., G is the union of a *countable* number of compact subsets. This property allows us to build a theory of distributions over G with all usual properties.[28]

The following spaces can be defined:

$\mathfrak{D}(G)$: space of C^∞ functions on G with compact support;

$\mathcal{E}(G)$: space of C^∞ functions on G;

$\mathfrak{D}'(G)$: space of distributions on G:

$\mathcal{E}'(G)$: space of distributions on G with compact support.

These spaces are endowed with their usual Schwartz topologies[28]: for $\mathcal{E}(G)$, uniform convergence of the functions and all their derivatives on any compact subset; $\mathfrak{D}(G)$ is the (countable) inductive limit of the $\mathfrak{D}_A(G)$, where $\mathfrak{D}_A(G)$ is the space of C^∞ functions on G, with support in a fixed compact set A, provided with the topology induced by $\mathcal{E}(G)$; $\mathfrak{D}'(G)$ and $\mathcal{E}'(G)$, the strong duals of $\mathfrak{D}(G)$ and $\mathcal{E}(G)$, respectively. These four spaces are *nuclear*, because of the countability of G at ∞; $\mathcal{E}(G)$ is a Fréchet space, but $\mathfrak{D}(G)$ is not metrizable if G is noncompact. Identifying the function $f \in \mathfrak{D}(G)$ with the distribution $f(g)\,dg$, one has the usual inclusions (algebraically and topologically):

$$\mathfrak{D}(G) \subset \mathcal{E}'(G) \subset \mathfrak{D}'(G),$$

$$\mathfrak{D}(G) \subset L^2(G, dg) \subset \mathfrak{D}'(G).$$

In a similar way, for any locally convex, complete, Hausdorff topological vector space E,[3] the following spaces can be defined:

$\mathfrak{D}(G, E)$: space of C^∞ functions with compact support on G, with values in E;

$\mathcal{E}(G, E)$: space of C^∞ functions on G, with values in E.

These two spaces are endowed with topologies analogous to those of $\mathfrak{D}(G)$ and $\mathcal{E}(G)$; $\mathfrak{D}(G, E)$ is again the inductive limit of the spaces $\mathfrak{D}_A(G, E)$.[49] One has the following result[10]:

$$\mathcal{E}(G, E) = \mathcal{E}(G) \hat{\otimes}_\pi E,$$

where the right-hand side is the completed projective tensor product.[3] It follows from this that both $\mathcal{E}(G, E)$ and $\mathfrak{D}(G, E)$ are nuclear if and only if E is nuclear (in particular, finite dimensional); this excludes the case of an infinite-dimensional Banach or Hilbert space.

Two kinds of dual spaces can be introduced:

(i) the space of distributions on G, with values in E:

$$\mathfrak{L}(\mathfrak{D}(G), E) \equiv \mathfrak{D}'(G, E)$$

(these are the *vector-valued distributions* of Schwartz[39]);

(ii) the space of E-distributions: $[\mathfrak{D}(G, E)]'$.

[49] This topology, introduced by Grothendieck [*Produits tensoriels topologiques et espaces nucléaires*, Memoirs Am. Math. Soc. **16**, (1955)], and used by Bruhat (see Ref. 10), is *not* the one adopted by Schwartz in Ref. 39.

One has in general the following inclusion:

$$[\mathcal{D}(G, E)]' \subseteq \mathcal{D}'(G, E'), \tag{B1}$$

where E' is the strong dual of E. The equality sign holds if E is a Fréchet space. In particular, for a Hilbert space H, one may identify $[\mathcal{D}(G, H)]'$ and $\mathcal{D}'(G, H)$.

Convolution may be defined as usual. The mapping $(T, f) \to T * f$, defined by the relation

$$(T * f)(g_0) = \int_G f(g^{-1}g_0)\, dT(g), \tag{B2}$$

is a *separately continuous mapping* of $\mathcal{D}'(G) \times \mathcal{D}(G, E)$ into $\mathcal{E}(G, E)$; the same is true for the convolutions $\mathcal{E}(G) \times \mathcal{E}(G, E) \to \mathcal{E}(G, E)$ and $\mathcal{E}'(G) \times \mathcal{D}(G, E) \to \mathcal{D}(G, E)$.

Similar results are valid with G replaced by any C^∞ manifold, countable at infinity—in particular, the quotient G/K of G by a closed subgroup.

Closely related to the preceding ones is the space $\mathcal{D}^{(L)}(G, E)$ introduced by Bruhat[10] (L is a unitary differentiable representation in E of the subgroup K of G). $\mathcal{D}^{(L)}(G, E)$ is the space of C^∞ functions $f: G \to E$, such that:

(i) the canonical image in G/K of $\operatorname{supp} f$ is compact;
(ii) $f(gk) = L^+(k)f(g)$, for all $g \in G$, $k \in K$.

The topology of $\mathcal{D}^{(L)}$ is defined as follows: Let A be a compact subset of G and $\mathcal{D}_A^{(L)}$ the space of those $f \in \mathcal{D}^{(L)}$ which have their support in AK, with the topology induced by $\mathcal{E}(G, E)$. $\mathcal{D}^{(L)}$ is then the inductive limit of all the spaces $\mathcal{D}_A^{(L)}$. If $L \equiv 1$, $\mathcal{D}^{(L)}$ reduces to $\mathcal{D}(G/K)$. Let $\hat{f} \in \mathcal{D}(G, E)$ and define

$$f(g) = \int_K L(k)\hat{f}(gk)\, dk. \tag{B3}$$

Then[10] the mapping $\hat{f} \to f$ is a topological homomorphism from $\mathcal{D}(G, E)$ *onto* $\mathcal{D}^{(L)}(G, E)$. Let $\mathcal{F}_L$ be its kernel, which is closed in $\mathcal{D}(G, E)$. Thus, we have

$$\mathcal{D}^{(L)}(G, E) = \mathcal{D}(G, E)/\mathcal{F}_L. \tag{B4}$$

It follows from this that the induced representation U^L, defined in $\mathcal{D}^{(L)}$ by left translations (see Sec. IV), is the quotient by the closed subspace $\mathcal{F}_L$ of the left regular representation of G in $\mathcal{D}(G, E)$, since the mapping $\hat{f} \to f$ commutes with left translations. And this implies that U^L restricted to $\mathcal{D}^{(L)}$ is a differentiable representation. Because of (B4), the *dual* of $\mathcal{D}^{(L)}$ is the orthogonal complement of $\mathcal{F}_L$, i.e., the closed subspace of $[\mathcal{D}(G, E)]'$ consisting of those functionals which vanish on $\mathcal{F}_L$. [The latter space was discussed before, in connection with Eq. (B1).] Finally, the convolution (B2) is again a separately continuous mapping from $\mathcal{E}'(G) \times \mathcal{D}^{(L)}(G, E)$ into $\mathcal{D}^{(L)}(G, E)$ [and *a fortiori* with $\mathcal{E}'(G)$ replaced by $\mathcal{D}(G)$].

2. Invariant Differential Operators[10,24]

A *differential operator* on a C^∞ manifold V is a continuous mapping D of $\mathcal{D}(V)$ into itself, such that

$$\operatorname{supp} Df \subseteq \operatorname{supp} f, \quad \text{for all } f \in \mathcal{D}(V). \tag{B5}$$

This (purely local) operator can be extended by continuity to the whole space $\mathcal{E}(V)$. A condition equivalent to (B5) is the following: In any local coordinate system $\{x_1 \cdots x_n\}$, D can be expressed as a *finite* sum of partial differential operators:

$$(Df)(x) = \sum_p a_p(x) \frac{\partial^{|p|} f(x)}{\partial x_1^{p_1} \cdots \partial x^{p_n}}, \tag{B6}$$

where $|p| = p_1 + \cdots + p_n$ and $a_p \in \mathcal{E}(V)$; the order of D is the highest integer $|p|$ occurring in the sum.

A differential operator D on the Lie group G is *right-invariant* (*left-invariant*) under G if D commutes with right (left) translations under G. Thus a differential operator on G/K (the space of *left* cosets gK) can be identified with a differential operator on G, right-invariant under K. Left-invariant differential operators on G/K (or more generally on any C^∞ manifold on which G acts by left translations) can be defined similarly: They are restrictions to G/K of differential operators on G which are both left-invariant under G and right-invariant under K.

The *Lie algebra* $\mathfrak{g}$ of G can be identified with the vector space of all first-order right-invariant differential operators on G. The *universal enveloping algebra* $\mathfrak{U}(\mathfrak{g})$ is the quotient $\mathfrak{T}(\mathfrak{g})/\mathfrak{J}$, where $\mathfrak{T}(\mathfrak{g})$ is the tensor algebra on $\mathfrak{g}$ and $\mathfrak{J}$ is the ideal consisting of all the elements of the form $T \otimes S - S \otimes T - [T, S]$, $T, S \in \mathfrak{g}$. It is an associative algebra composed of polynomials in the elements of $\mathfrak{g}$; $\mathfrak{g}$ itself is a subalgebra of $\mathfrak{U}(\mathfrak{g})$ (first-order monomials). The following theorems then hold:

a. $\mathfrak{U}(\mathfrak{g})$ is isomorphic to the algebra of all right-invariant differential operators on G;

b. $\mathfrak{U}(\mathfrak{g})$ is isomorphic to the algebra $\mathcal{E}'_e$ of distributions with support $\{e\}$ (with respect to convolution).

Indeed, for any $T \in \mathcal{E}'_e$, $f \in \mathcal{D}(G)$, the mapping $f \to T * f$ is a right-invariant differential operator ($f \to f * T$ would be left-invariant); this follows from (B6) and the fact that any $T \in \mathcal{E}'_e$ is a *finite* linear combination of derivatives of the distribution δ, with C^∞ coefficients.[3] Conversely, all right-invariant differential operators can be obtained in this way.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 12 DECEMBER 1969

Uniform Asymptotic Theory of Edge Diffraction*

ROBERT M. LEWIS† AND JOHANNES BOERSMA
Courant Institute of Mathematical Sciences, New York University
and
Department of Mathematics, Technological University, Eindhoven, The Netherlands

(Received 28 September 1966)

Geometrical optics fails to account for the phenomenon of diffraction, i.e., the existence of nonzero fields in the geometrical shadow. Keller's geometrical theory of diffraction accounts for this phenomenon by providing correction terms to the geometrical optics field, in the form of a high-frequency asymptotic expansion. In problems involving screens with apertures, this asymptotic expansion fails at the edge of the screen and on shadow boundaries where the expansion has singularities. The uniform asymptotic theory presented here provides a new asymptotic solution of the diffraction problem which is uniformly valid near edges and shadow boundaries. Away from these regions the solution reduces to that of Keller's theory. However, singularities at any caustics other than the edge are not corrected.

1. INTRODUCTION

Geometrical optics fails to account for the phenomenon of diffraction, i.e., the existence of nonzero fields in the geometrical shadow. It is now known that the geometrical-optics field corresponds to the leading term of a high-frequency asymptotic expansion of the solution of a boundary-value problem for the reduced wave equation or Maxwell's equations, and that higher-order terms account for diffraction. Keller's "geometrical theory of diffraction" [1,2] provides a systematic means for computing these terms.

In this paper we consider problems of diffraction by screens. The screens may be portions of planes or other smooth surfaces bounded by smooth curves, and the prescribed incident wave may be arbitrary. We consider here only scalar problems for the reduced wave equation with boundary conditions of the first or second kind ($u = 0$ or $\partial u/\partial n = 0$) on the screen. In Sec. 2 we present a brief but self-contained treatment of Keller's geometrical theory for such problems. This theory depends on a "diffraction coefficient" the value of which is obtained from a special ("canonical") problem, the problem of diffraction of a plane wave by a half plane. Sommerfeld's solution of this problem is discussed in Sec. 3 and there the diffraction coefficient is evaluated.

The geometrical theory has several shortcomings. It fails at the shadow boundaries of the incident and reflected waves as well as at the edge of the screen where the "diffracted wave" becomes infinite. Furthermore, it is difficult to justify the determination of the diffraction coefficient by comparison with the solution of the canonical problem, and this procedure cannot be generalized to yield higher-order terms in the diffracted field. These shortcomings are overcome by the method presented in Secs. 4 and 5 of this paper. Other shortcomings of the geometrical theory (the failure at caustics of the problem) remain. Like Keller's theory, ours is formal in the sense that we do not rigorously prove the asymptotic nature of the solution obtained.

Our approach is motivated by a new representation of the solution of the half-plane problem. By using simple concepts of the geometrical theory such as incident-, reflected-, and diffracted-phase functions, we show in Sec. 3 that Sommerfeld's solution can be expressed in a remarkably simple and suggestive form. This representation involves a special function f which is discussed briefly in Appendix A. It is closely related to the Fresnel integral functions.

The geometrical theory of diffraction is based on an "ansatz" in the form of an asymptotic series involving certain "phase" and "amplitude" functions. By inserting the series into the reduced wave equation, one obtains the eikonal equation for the phase function $s(\mathbf{x})$ and a sequence of transport equations for the amplitude functions $z_m(\mathbf{x})$. These equations can be solved by introducing lines in $\mathbf{x}$-space called "rays." Our approach is based on a new ansatz that involves the function f. Away from the edge of the screen and the shadow boundaries, the new expression reduces to one of the same form as Keller used. Therefore the phase and amplitude functions which appear in the new ansatz also satisfy the eikonal and transport equations. In Keller's theory there is an undetermined "initial condition" for the transport equation of order zero. This leads to the diffraction coefficient. In our approach the initial condition is

* The research in this paper was supported by the Air Force Cambridge Research Laboratories, Office of Aerospace Research, under Contract No. AF 19(628)3868. Reproduction in whole or in part is permitted for any purpose of the U.S. Government.

† R. M. Lewis died on 7 November 1968.

[1] J. B. Keller, J. Opt. Soc. Am. **52**, 116 (1962).

[2] R. M. Lewis and J. B. Keller, New York University Research Report EM-194, 1964.

uniquely determined by imposing the "edge condition," which is a part of the rigorous formulation of the boundary-value problem. Away from the edge and shadow boundaries, the leading term of our result reduces to Keller's, and we verify his expression for the diffraction coefficient.

By construction our solution is continuous and finite at the edge of the screen because the edge condition demands this. It is not immediately obvious that it is also continuous at the shadow boundaries. However, in Sec. 4 we compute the leading term of our expansion and prove that it is continuous at the shadow boundaries as well as at the edge. (For this reason we call our asymptotic solution "uniform.") The generalization of this theorem to higher-order terms has not yet been proved. In Sec. 5 we compute the next term of our expansion. In order to simplify the calculations we restrict the problem at this point to screens which are portions of planes. The computation requires an expression for the Laplacian in "ray coordinates" which are not orthogonal. This expression is derived in Appendix C. When our result is evaluated away from the edge and shadow boundaries, it again reduces to an expression of the form used in the geometrical theory, but now the first two terms of the diffracted wave are given. The second term can be expressed in a form that involves Keller's diffraction coefficient and a new coefficient. There is a special problem (grazing incidence with boundary condition $\partial u/\partial n = 0$) in which Keller's diffraction coefficient vanishes and the second term becomes important. For this case Keller has obtained a special diffraction coefficient by using a special canonical problem. In this case our new coefficient reduces to his.

In several respects our theory is incomplete. We have already mentioned the unproved conjecture that all terms are continuous at the shadow boundary. There is a second unproved conjecture: We have obtained the first two terms of the expansion, at least for plane screens. (This is probably not an essential restriction.) It seems likely that the procedure can be continued to yield higher-order terms. But this too is not obvious and has not yet been proved. [*Note added in proof:* Both conjectures were proved recently; see D. S. Ahluwalia, R. M. Lewis, and J. Boersma, SIAM J. Appl. Math. **16**, 703 (1968).] Furthermore, as we have mentioned, our theory also fails at caustics of the incident and reflected waves and caustic points of the diffracted wave other than those on the edge.

Our theory is also incomplete in another sense. For nonplanar screens, diffracted rays emanating from the edge may strike another portion of the screen giving rise to secondary reflected waves or creeping waves. Such waves are not included in our theory. (See the remarks at the end of Sec. 4.)

Uniform expansions which are valid at caustics have recently been obtained by Kravtsov[3] and Ludwig.[4] In fact, their work partially motivated our approach to the problem of diffraction by screens. A second motivation came from the work of Lewis[5] on the uniform transition from the "forerunner" to the "main signal" of a transient wave propagating into a dispersive medium.

The main motivation, however, came from the recent work of Wolfe.[6] Wolfe considered some special cases of the problems treated here, involving plane and spherical waves incident on a screen which is a portion of a plane. For these problems he obtained uniform asymptotic solutions by means of an ansatz involving Fresnel integrals. This ansatz, which was given in terms of ray coordinates, was substituted into the reduced wave equation which had to be transformed to these same coordinates. This obscures several important features of the method. For example, one does not see that the ansatz involves functions that are identical to the phase and amplitude functions of the geometrical theory. As a consequence Wolfe's method is more complicated than ours. In addition, Wolfe relies on the use of the canonical half-plane problem, since the Fresnel-integral part of his ansatz is derived from the uniform asymptotic expansion of the solution of the half-plane problem for the same incident wave. Since this problem has been solved only for special incident waves (plane, cylindrical, and spherical), this restricts the generality of his method. Nevertheless the essential features of our approach are contained in Wolfe's work and we are very much indebted to him. We are of course also greatly indebted to Keller, not only for his geometrical theory of diffraction, but also for his continuing interest and advice in the course of Wolfe's work and our own.

In closing this introduction we wish to mention some problems closely related to the one treated here. The problem of diffraction by a screen is a special case of diffraction by objects which are locally wedge-shaped. (Along the edge the screen is locally a zero-angled wedge.) Such problems can be treated by Keller's theory. The generalization of our method to these problems is currently under consideration. There

[3] Yu. A. Kravtsov, Radiofiz. **7**, 664 (1964).
[4] D. Ludwig, Commun. Pure Appl. Math. **19**, 215 (1966).
[5] R. M. Lewis, *Proceedings of the U.R.S.I. Symposium on Electromagnetic Wave Theory* (Delft, The Netherlands, 1965).
[6] P. Wolfe, "Diffraction of a Scalar Wave by a Plane Screen," Ph.D. thesis, New York University, 1965.

is also a fairly obvious generalization of our approach to problems of diffraction by screens in inhomogeneous media. We have not included a treatment of such problems because the added complications are not justified by the practical importance of the generalization. In addition it is fairly clear that the method presented here can be applied to Maxwell's equations and other linear partial-differential equations, but this has not yet been done. Many diffraction problems (e.g., diffraction by a slit or by a circular aperture in a plane screen) involve "multiple diffraction" (waves produced at one edge are incident on another). Such problems have been treated by Keller and will be treated by our method in a forthcoming sequel to this paper. Finally there is a whole class of problems of diffraction by smooth objects that can be treated by another part of Keller's theory. Recently uniform asymptotic solutions of these problems have been obtained.[7,8] These solutions improve on Keller's theory in much the same way as the method presented here improves on his theory of edge diffraction.

2. KELLER'S GEOMETRICAL THEORY OF DIFFRACTION

In this section we present a summary of that part of Keller's theory which relates to diffraction by an edge of a screen. Further details are given in References 1 and 2. It is important for us to summarize Keller's theory not only because our work was motivated by it, but because we make heavy use of his results. In Secs. 4 and 5, we use almost all the equations derived here.

We consider asymptotic solutions of the reduced wave equation

$$\Delta u + k^2 u = 0 \tag{2.1}$$

of the form

$$u \sim e^{iks(\mathbf{x})} \sum_{m=0}^{\infty} (ik)^{-m} z_m(\mathbf{x}), \quad k \to \infty. \tag{2.2}$$

By inserting (2.2) into (2.1), we find that the *phase function* $s(\mathbf{x})$ satisfies the *eikonal equation* of geometrical optics

$$(\nabla s)^2 = 1, \tag{2.3}$$

while the *amplitude functions* $z_m(\mathbf{x})$ satisfy the recursive system of *transport equations*

$$2\nabla s \cdot \nabla z_m + z_m \Delta s = -\Delta z_{m-1}; \quad m = 0, 1, 2, \cdots, \quad z_{-1} \equiv 0. \tag{2.4}$$

Solution of (2.3) may be described as follows: Given a surface (*wavefront*) on which s has the constant value s_0, we introduce the two-parameter family of straight lines (*rays*) orthogonal to the surface. If σ denotes distance along the rays from the wavefront (measured positively in the direction of increasing s), then on each ray

$$s = s_0 + \sigma. \tag{2.5}$$

It is then clear that (2.5) satisfies (2.3).

Let σ_2 and σ_3 be the two parameters that label the rays and let us describe a ray parametrically in the form

$$\mathbf{x} = \mathbf{x}(\sigma) = \mathbf{x}(\sigma, \sigma_2, \sigma_3). \tag{2.6}$$

If we set $\sigma = \sigma_1$, then (2.6) defines a transformation from $(\sigma_1, \sigma_2, \sigma_3)$-space to (x_1, x_2, x_3)-space and the Jacobian of the transformation is

$$j = j(\sigma) = j(\sigma, \sigma_2, \sigma_3) = \det\left(\frac{\partial x_i}{\partial \sigma_j}\right), \quad i, j = 1, 2, 3. \tag{2.7}$$

For given z_{m-1} it is easy to see that (2.4) is an ordinary differential equation for z_m along a ray. The solution can be expressed in the form

$$z_m(\sigma) = \left|\frac{j(\sigma_0)}{j(\sigma)}\right|^{\frac{1}{2}} z_m(\sigma_0) - \frac{1}{2}\int_{\sigma_0}^{\sigma} \left|\frac{j(\sigma')}{j(\sigma)}\right|^{\frac{1}{2}} \Delta z_{m-1}(\sigma')\, d\sigma', \quad m = 0, 1, 2, \cdots. \tag{2.8}$$

Here $z_m(\sigma) = z_m[\mathbf{x}(\sigma, \sigma_2, \sigma_3)]$ is the value of z_m at a point σ on a given ray. The solution (2.8) is given in terms of an "initial value" $z_m(\sigma_0)$ at some fixed point on each ray. For $m = 0$ we note that the second term of (2.8) is absent because $z_{-1} \equiv 0$. Two alternative expressions for the ratio of Jacobians are sometimes useful:

$$\frac{j(\sigma_0)}{j(\sigma)} = \frac{da(\sigma_0)}{da(\sigma)} = \frac{(\rho_2 + \sigma_0)(\rho_3 + \sigma_0)}{(\rho_2 + \sigma)(\rho_3 + \sigma)}. \tag{2.9}$$

Here $da(\sigma)$ is the cross-sectional area of an infinitesimal tube of rays, while $\rho_2(\sigma_2, \sigma_3)$ and $\rho_3(\sigma_2, \sigma_3)$ are the principal radii of curvature of the wavefront $\sigma = 0$.

At the two points $\sigma = -\rho_2$ and $\sigma = -\rho_3$ on each ray, we see from (2.9) that z_m becomes infinite and the integral in (2.8) will, in general, diverge. Such points are called *caustic points*. They lie on the *caustic*, which is, in general, a two-sheeted surface forming the envelope of the family of rays (the rays are tangent to the caustic). We shall require an alternative form of (2.8) which remains valid when $\sigma_0 = 0$ is a caustic point. First we rewrite (2.8) in the form

$$|j(\sigma)|^{\frac{1}{2}} z_m(\sigma) = |j(\sigma_0)|^{\frac{1}{2}} z_m(\sigma_0) - \tfrac{1}{2}\int_{\sigma_0}^{\sigma} |j(\sigma')|^{\frac{1}{2}} \Delta z_{m-1}(\sigma')\, d\sigma'. \tag{2.10}$$

[7] R. M. Lewis, N. Bleistein, and D. Ludwig, Commun. Pure Appl. Math. **20**, 295 (1967).
[8] D. Ludwig, Commun. Pure Appl. Math. **20**, 103 (1967).

Then we express the integral in (2.10) in the form

$$\int_{\sigma_0}^{\sigma} = \text{⨍}_0^{\sigma} - \text{⨍}_0^{\sigma_0}. \tag{2.11}$$

Here the dash denotes the "finite part" [9] of a divergent integral. (The ordinary integrals would diverge at $\sigma' = 0$.) Now (2.10) becomes

$$|j(\sigma)|^{\frac{1}{2}} z_m(\sigma) + \tfrac{1}{2}\text{⨍}_0^{\sigma} = |j(\sigma_0)|^{\frac{1}{2}} z_m(\sigma_0) + \tfrac{1}{2}\text{⨍}_0^{\sigma_0}. \tag{2.12}$$

The right side of (2.12) is independent of σ. If we denote its value by ζ_m, then we obtain

$$z_m(\sigma) = \frac{\zeta_m}{|j(\sigma)|^{\frac{1}{2}}} - \frac{1}{2}\text{⨍}_0^{\sigma} \left|\frac{j(\sigma')}{j(\sigma)}\right|^{\frac{1}{2}} \Delta z_{m-1}(\sigma')\, d\sigma', \quad m = 0, 1, 2, \cdots. \tag{2.13}$$

This is the required modification of (2.8). The "initial value" $\zeta_m(\sigma_2, \sigma_3)$ has first to be determined before (2.13) is useful. We will see in Sec. 5 that the finite-part integrals are a useful computational tool. For $m = 0$, the integral in (2.13) is again absent.

We now consider the problem of diffraction by a screen S. The screen is a portion of a smooth surface. It is bounded by an *edge* E consisting of a smooth curve

$$\mathbf{x} = \mathbf{x}_0(\eta). \tag{2.14}$$

Here η is an arclength parameter. For example, S might be an infinite plane with a circular aperture or it could be the complementary disk. Alternatively the aperture may have any smooth shape. In general, S need not be a portion of a plane. We consider an incident wave

$$u_0^{\mathrm{i}} \sim e^{iks^{\mathrm{i}}} \sum_{m=0}^{\infty} (ik)^{-m} z_m^{\mathrm{i}}, \tag{2.15}$$

which is an asymptotic solution of (2.1). Then s^{i} and the z_m^{i} satisfy the equations derived above. The total field u is a solution of (2.1) and satisfies a boundary condition on the screen. We shall consider simultaneously the two conditions

$$u = 0 \quad \text{on } S \tag{2.16a}$$

and

$$\frac{\partial u}{\partial n} = \mathbf{N} \cdot \nabla u = 0 \quad \text{on } S. \tag{2.16b}$$

Here $\mathbf{N}$ is a unit normal vector on S. In addition, $u - u_0^{\mathrm{i}}$ is required to be "outgoing."

To solve the diffraction problem, we first set $u = u_0^{\mathrm{i}} + u_0^{\mathrm{r}}$. We assume that the *reflected wave* u_0^{r} has an asymptotic expansion

$$u_0^{\mathrm{r}} \sim e^{iks^{\mathrm{r}}} \sum_{m=0}^{\infty} (ik)^{-m} z_m^{\mathrm{r}}. \tag{2.17}$$

Then (2.16) will be satisfied, provided

$$s^{\mathrm{r}} = s^{\mathrm{i}} \text{ on } S \tag{2.18}$$

and, for the boundary condition $u = 0$,

$$z_m^{\mathrm{r}} = -z_m^{\mathrm{i}} \quad \text{on } S; \quad m = 0, 1, 2, \cdots. \tag{2.19}$$

For the case $\partial u/\partial n = 0$, (2.19) is replaced by

$$z_m^{\mathrm{i}} \frac{\partial s^{\mathrm{i}}}{\partial n} + z_m^{\mathrm{r}} \frac{\partial s^{\mathrm{r}}}{\partial n} + \frac{\partial z_{m-1}^{\mathrm{i}}}{\partial n} + \frac{\partial z_{m-1}^{\mathrm{r}}}{\partial n} = 0 \quad \text{on } S, \quad m = 0, 1, 2, \cdots. \tag{2.20}$$

It can be shown that (2.18) implies that the incident and reflected rays (which have the direction ∇s^{i} and ∇s^{r}, respectively) satisfy the law of reflection of geometrical optics. If ψ is the angle of incidence (= angle of reflection), then $\partial s^{\mathrm{r}}/\partial n = \cos\psi = -\partial s^{\mathrm{i}}/\partial n$ and (2.20) becomes

$$z_m^{\mathrm{r}} = z_m^{\mathrm{i}} - \frac{1}{\cos\psi}\left(\frac{\partial z_{m-1}^{\mathrm{i}}}{\partial n} + \frac{\partial z_{m-1}^{\mathrm{r}}}{\partial n}\right) \quad \text{on } S, \quad m = 0, 1, 2, \cdots. \tag{2.21}$$

Thus s^{r} is determined on the reflected rays by (2.18) and (2.5), while the functions z_m^{r} are given by (2.8) with $z_m(\sigma_0)$ determined by (2.19) or (2.21). We note that both u_0^{i} and u_0^{r} are zero in their respective "shadow regions," i.e., where there are no incident or reflected rays. Thus each has an "illuminated region" separated from the corresponding shadow region by a *shadow boundary* surface.

The leading term $u = u_0^{\mathrm{i}} + u_0^{\mathrm{r}} \sim z_0^{\mathrm{i}} \exp(iks^{\mathrm{i}}) + z_0^{\mathrm{r}} \exp(iks^{\mathrm{r}})$ is the geometrical-optics solution of the problem, which of course fails to account for diffraction phenomena (nonzero fields in the shadows). The full solution (2.15) + (2.17) is correct only to first order because, according to Keller's theory, there is an additional *diffracted wave* $\hat{u}$. Then

$$u = u_0^{\mathrm{i}} + u_0^{\mathrm{r}} + \hat{u}, \tag{2.22}$$

where

$$\hat{u} \sim k^{-\frac{1}{2}} e^{ik\hat{s}} \sum_{m=0}^{\infty} (ik)^{-m} \hat{z}_m. \tag{2.23}$$

Of course $\hat{s}$ and the $\hat{z}_m$ satisfy the equations derived earlier for phase and amplitude functions. The *diffracted rays* associated with $\hat{s}$ emanate from the edge E of the screen and

$$\hat{s} = s^{\mathrm{i}} \text{ on } E. \tag{2.24}$$

[9] Let $f(\epsilon) = \int_{\epsilon}^{b} g(x)\, dx$ have an asymptotic expansion in (perhaps fractional) powers of ϵ for $\epsilon \to 0$. The coefficient of $\epsilon^0 = 1$ in the expansion is called the *finite part* of the integral and will be denoted by $\text{⨍}_0^b g(x)\, dx$.

Let us introduce the unit tangent, normal, and binormal vectors $\mathbf{t}$, $\mathbf{n}$, and $\mathbf{b}$ of E. Then $\mathbf{t} = \dot{\mathbf{x}}_0(\eta)$, $\mathbf{b} = \mathbf{t} \times \mathbf{n}$, and the equations of Frenet,

$$\dot{\mathbf{t}} = \kappa\mathbf{n}, \quad \dot{\mathbf{n}} = -\kappa\mathbf{t} + \tau\mathbf{b}, \quad \dot{\mathbf{b}} = -\tau\mathbf{n}, \tag{2.25}$$

are satisfied. Here κ is the curvature and τ is the torsion of E. If we differentiate (2.24) with respect to η, we obtain

$$\nabla\hat{s} \cdot \mathbf{t} = \nabla s^{\mathrm{i}} \cdot \mathbf{t}. \tag{2.26}$$

This implies the *law of edge diffraction*: The diffracted rays make the same angle with the tangent $\mathbf{t}$ to the edge as the incident ray at the point of diffraction. Let $\beta = \beta(\eta)$ be this angle. Then, from each point $\mathbf{x}_0(\eta)$ of the edge, the diffracted rays emanating from the point generate a cone of semiangle β. Thus we have a two-parameter family of diffracted rays

$$\mathbf{x} = \mathbf{x}_0(\eta) + \sigma\mathbf{U}(\eta, \alpha), \tag{2.27}$$

where $\mathbf{U}$ is a unit vector given by

$$\mathbf{U} = \cos\beta\mathbf{t} + \sin\beta\cos\alpha\mathbf{n} + \sin\beta\sin\alpha\mathbf{b}. \tag{2.28}$$

We may calculate the Jacobian

$$j(\sigma) = j(\sigma, \alpha, \eta) = \frac{\partial(x_1, x_2, x_3)}{\partial(\sigma, \alpha, \eta)} \tag{2.29}$$

of the transformation defined by (2.27), using (2.25). We find that

$$j = \frac{\partial\mathbf{x}}{\partial\sigma} \cdot \frac{\partial\mathbf{x}}{\partial\alpha} \times \frac{\partial\mathbf{x}}{\partial\eta} = (\sin^2\beta)\sigma\left(1 + \frac{\sigma}{\rho}\right), \tag{2.30}$$

where

$$\rho = -\frac{\sin\beta}{\dot{\beta}(\eta) + \kappa\cos\alpha} = -\frac{\sin^2\beta}{\dot{\beta}\sin\beta + \kappa\cos\delta}, \tag{2.31}$$

and

$$\cos\delta = \sin\beta\cos\alpha = \mathbf{U} \cdot \mathbf{n}. \tag{2.32}$$

Then, if we set $\delta_m(\alpha, \eta) = \zeta_m/\sin\beta$, (2.13) yields

$$\hat{z}_m(\sigma) = \left|\sigma\left(1 + \frac{\sigma}{\rho}\right)\right|^{-\frac{1}{2}}\delta_m - \frac{1}{2}\int_0^\sigma \left|\frac{\sigma'(\rho + \sigma')}{\sigma(\rho + \sigma)}\right|^{\frac{1}{2}}\Delta\hat{z}_{m-1}(\sigma')\,d\sigma'; \quad m = 0, 1, 2, \cdots. \tag{2.33}$$

In particular,

$$\hat{z}_0 = \left|\sigma\left(1 + \frac{\sigma}{\rho}\right)\right|^{-\frac{1}{2}}\delta_0. \tag{2.34}$$

The undetermined factor δ_0 is assumed to be proportional to the amplitude z_0^{i} of the incident wave at the point of diffraction $\mathbf{x}_0(\eta)$. Then $\delta_0 = Dz_0^{\mathrm{i}}[\mathbf{x}_0(\eta)]$ and

$$\hat{z}_0(\sigma) = Dz_0^{\mathrm{i}}\left|\sigma\left(1 + \frac{\sigma}{\rho}\right)\right|^{-\frac{1}{2}}. \tag{2.35}$$

The factor of proportionality D is called a *diffraction coefficient.* In Keller's theory it is determined by comparison with the solution of the problem of diffraction by a half-plane; the motivation is that diffraction is a local phenomenon and locally the screen can be approximated by a half-plane. In the next section we discuss the solution of the half-plane problem and derive the diffraction coefficient. The value of D is given by (3.16). It depends on the angles β, φ, and φ_0. The angles φ and φ_0 are illustrated in Fig. 1.

In closing this section we state and prove two lemmas which will be useful in Sec. 4.

Lemma 1: Let $s^{\mathrm{i,r}}$ be the phase function of the incident (reflected) wave and $\hat{s}$ the phase function of the diffracted wave. Then

$$s^{\mathrm{i,r}}(\mathbf{x}) \leq \hat{s}(\mathbf{x}), \tag{2.36}$$

and $s^{\mathrm{i,r}}(\mathbf{x}) = \hat{s}(\mathbf{x})$ if and only if $\mathbf{x}$ is a point on the shadow boundary of the incident (reflected) wave.

Proof: Let s_0 be the common value of s^{i}, s^{r}, and $\hat{s}$ at a point $\mathbf{Q}$ on the edge. Let $\mathbf{P}$ be any point on a diffracted ray emanating from $\mathbf{Q}$ in the direction of the unit vector $\mathbf{U}$; and let σ_0 be the distance from $\mathbf{Q}$ to $\mathbf{P}$. Since $|\nabla s^{\mathrm{i,r}}| = 1$,

$$s^{\mathrm{i,r}}(\mathbf{P}) = s_0 + \int_{\mathbf{Q}}^{\mathbf{P}} \nabla s^{\mathrm{i,r}} \cdot \mathbf{U}\,d\sigma \leq s_0 + \sigma_0 = \hat{s}(\mathbf{P}). \tag{2.37}$$

Equality holds in (2.37) if and only if $\nabla s^{\mathrm{i,r}} \equiv \mathbf{U}$, i.e., if and only if the diffracted ray coincides with an incident (reflected) ray. But this occurs if and only if

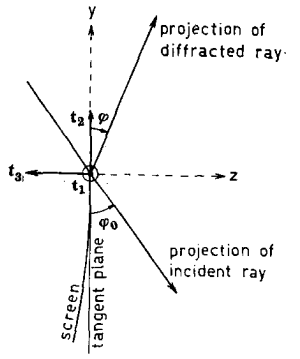


FIG. 1. Angles at the edge of a screen. The unit vectors $\mathbf{t}_2$ and $\mathbf{t}_3$ are orthogonal to the unit vector $\mathbf{t}_1$, which is tangent to the edge. $\mathbf{t}_2$ lies in the tangent plane and points away from the screen. $\mathbf{t}_3 = \mathbf{t}_1 \times \mathbf{t}_2$ is orthogonal to the tangent plane and points toward the illuminated side. Projections of the incident ray and the diffracted ray into the plane orthogonal to the edge are shown. The directions of the projections are determined by the angles φ_0 ($0 \leq \varphi_0 \leq \pi$) and φ ($-\pi \leq \varphi \leq \pi$). For purposes of Sec. 3 the y and z axes are shown. For purposes of Secs. 3 and 5 the screen coincides with the tangent plane.

P is on the shadow boundary of the incident (reflected) wave.

Lemma 2: (See Fig. 2.) Let $\mathbf{t}$ be a unit vector tangent to the edge at $\mathbf{Q}$, and $\mathbf{U}_1$ a unit vector in the direction of the incident (reflected) ray at $\mathbf{Q}$. Let $\mathbf{U}_3$ be a unit vector in the plane T spanned by $\mathbf{t}$ and $\mathbf{U}_1$. $\mathbf{U}_3$ is chosen so that it is perpendicular to $\mathbf{U}_1$ and

$$\mathbf{t} = \cos\beta\mathbf{U}_1 + \sin\beta\mathbf{U}_3. \tag{2.38}$$

Let ρ_0 be the radius of curvature of the normal section of the incident (reflected) wavefront at $\mathbf{Q}$ in the direction of $\mathbf{U}_3$. Then for the diffracted ray emanating from $\mathbf{Q}$ which lies on the shadow boundary of the incident (reflected) wave, the quantity ρ defined by (2.31) has the value

$$\rho = \rho_0. \tag{2.39}$$

Proof: Let S be the shadow-boundary surface of the incident (reflected) wave, $W(\mathbf{x})$ the incident (reflected) wavefront and $\hat{W}(\mathbf{x})$ the diffracted wavefront that passes through the point $\mathbf{x}$. On S, the incident (reflected) and diffracted rays coincide. (See Fig. 2.) Let $\mathbf{Q}$ denote a point on the edge $\mathbf{x} = \mathbf{x}_0(\eta)$ and $\mathbf{P}$ a point on the ray through $\mathbf{Q}$ at a distance σ from $\mathbf{Q}$. From (2.9) and (2.30), we see that the principal radii at $\mathbf{P}$ of $\hat{W}(\mathbf{P})$ are σ and $\rho + \sigma$. The cone of diffracted rays emanating from $\mathbf{Q}$ intersects $\hat{W}(\mathbf{P})$ in a circle. The axis of the cone has the direction of the vector $\mathbf{t}$. From the "formula of Rodrigues" it can easily be shown that the circle is a line of curvature on $\hat{W}(\mathbf{P})$ corresponding to the principal radius of curvature σ. Therefore one of the principal directions at $\mathbf{P}$ is tangent to the circle, hence perpendicular to $\mathbf{t}$. Thus it is perpendicular to the plane T spanned by $\mathbf{t}$ and $\mathbf{U}_1$. The other principal direction corresponding to the principal radius of curvature $\rho + \sigma$ is given by the vector $\mathbf{U}_3$ which lies in the plane T and is perpendicular to $\mathbf{U}_1$. Since $\hat{W}(\mathbf{Q})$ and $\hat{W}(\mathbf{P})$ are parallel surfaces, their principal directions are the same. If we take $\sigma = 0$, we see that the principal radius of curvature of $\hat{W}(\mathbf{Q})$ in the direction $\mathbf{U}_3$ is ρ.

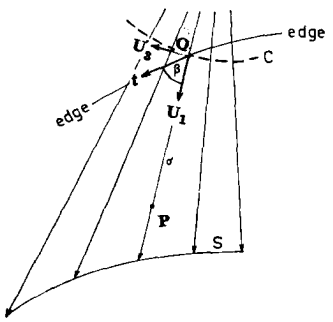


FIG. 2 The shadow boundary surface S (proof of Lemma 2).

Let C be the curve of intersection of S and $W(\mathbf{Q})$. According to Lemma 1, $\hat{s} = s$ on S; hence C also lies on $\hat{W}(\mathbf{Q})$. The rays of S are orthogonal to both $W(\mathbf{Q})$ and $\hat{W}(\mathbf{Q})$ along C; hence $W(\mathbf{Q})$ and $\hat{W}(\mathbf{Q})$ are tangent along C. From this it can easily be shown that the radii of curvature of the normal section of $W(\mathbf{Q})$ and $\hat{W}(\mathbf{Q})$ at any point of C in the direction of the tangent to C are equal. But at $\mathbf{Q}$ the tangent to C has the direction $\mathbf{U}_3$. Hence $\rho_0 = \rho$.

3. SOMMERFELD'S SOLUTION OF THE PROBLEM OF DIFFRACTION BY A HALF-PLANE

In this section we shall express Sommerfeld's well-known solution of the half-plane diffraction problem in a new form. This expression partially motivated our work in this paper. From Sommerfeld's solution we shall also derive Keller's diffraction coefficient.

Let x, y, and z be rectangular coordinates and let a half-plane be given by $z = 0$, $y \leq 0$, as in Fig. 1. We introduce polar coordinates ρ, φ ($0 \leq \rho$; $-\pi \leq \varphi \leq \pi$) defined by the equations $y = \rho\cos\varphi$, $z = \rho\sin\varphi$; and an incident plane wave $u_0^i = \exp(iks^i)$, where

$$\begin{aligned} s^i &= x\cos\beta - y\sin\beta\cos\varphi_0 + z\sin\beta\sin\varphi_0 \\ &= x\cos\beta - \rho\sin\beta\cos(\varphi + \varphi_0). \end{aligned} \tag{3.1}$$

For $\beta = \pi/2$, the solution of the diffraction problem with boundary conditions (2.16) was first derived by Sommerfeld. A simple derivation appears in Ref. 10.[11] For arbitrary β the transformation

$$u = \exp(ikx\cos\beta)u'(y, z)$$

reduces the general problem to the special one ($\beta = \pi/2$), with k replaced by $k\sin\beta$. Thus it is not difficult to obtain the solution of the diffraction problem for the incident plane wave given above. We find that the total field is given by

$$\begin{aligned} u = {} & \exp\{ik[-\rho\sin\beta\cos(\varphi + \varphi_0) + x\cos\beta]\} \\ & \times h[(2k\rho\sin\beta)^{\frac{1}{2}}\cos\tfrac{1}{2}(\varphi + \varphi_0)] \\ & \mp \exp\{ik[-\rho\sin\beta\cos(\varphi - \varphi_0) + x\cos\beta]\} \\ & \times h[-(2k\rho\sin\beta)^{\frac{1}{2}}\cos\tfrac{1}{2}(\varphi - \varphi_0)], \end{aligned} \tag{3.2}$$

where

$$h(x) = \pi^{-\frac{1}{2}}e^{-i\pi/4}\int_{-\infty}^{x} e^{it^2}\,dt. \tag{3.3}$$

In (3.2) the upper (lower) sign is valid for the boundary condition $u = 0$ ($\partial u/\partial n = 0$) on the half-plane.

[10] C. J. Bouwkamp, New York University Research Report EM-50, 1953.

[11] To obtain our notation from Ref. 10, it is necessary to replace by $\pi - \varphi$, θ_0 by φ_0, and y by $-y$.

The above result can be greatly simplified by introducing the reflected and diffracted phase functions s^r and $\hat{s}$ discussed in Sec. 2. It is easily seen that

$$\begin{aligned} s^r &= x\cos\beta - y\sin\beta\cos\varphi_0 - z\sin\beta\sin\varphi_0 \\ &= x\cos\beta - \rho\sin\beta\cos(\varphi-\varphi_0). \end{aligned} \tag{3.4}$$

The diffracted rays may be expressed in the form

$$\begin{aligned} \mathbf{x} = (x,y,z) &= (\eta,0,0) \\ &\quad + \sigma(\cos\beta, \sin\beta\cos\varphi, \sin\beta\sin\varphi) \end{aligned} \tag{3.5}$$

and on each ray

$$\hat{s} = s^i(\eta,0,0) + \sigma = \eta\cos\beta + \sigma. \tag{3.6}$$

Since $\eta = x - \sigma\cos\beta$ and $\rho = \sigma\sin\beta$, we see that

$$\hat{s} = x\cos\beta + \rho\sin\beta. \tag{3.7}$$

We note that

$$\hat{s} - s^i = 2\rho\sin\beta\cos^2\tfrac{1}{2}(\varphi+\varphi_0) \tag{3.8}$$

and

$$\hat{s} - s^r = 2\rho\sin\beta\cos^2\tfrac{1}{2}(\varphi-\varphi_0).$$

We set

$$\epsilon^i = \operatorname{sgn}[\cos\tfrac{1}{2}(\varphi+\varphi_0)], \tag{3.9}$$

$$\epsilon^r = -\operatorname{sgn}[\cos\tfrac{1}{2}(\varphi-\varphi_0)], \tag{3.10}$$

and

$$f(x) = e^{-ix^2}h(x) = \pi^{-\frac{1}{2}}e^{-i\pi/4}e^{-ix^2}\int_{-\infty}^{x} e^{it^2}\,dt. \tag{3.11}$$

Then (3.2) can be written as

$$u = u^i + u^r, \tag{3.12}$$

where

$$u^{i,r} = e^{ik\hat{s}}f\{\epsilon^{i,r}[k(\hat{s}-s^{i,r})]^{\frac{1}{2}}\}z_0^{i,r}, \tag{3.13}$$

and z_0^i and z_0^r are the geometrical incident and reflected amplitudes. Thus

$$z_0^i = 1, \quad z_0^r = \mp 1. \tag{3.14}$$

The geometric structure of the exact solution [(3.12) and (3.13)] becomes even more striking when we recognize that $\epsilon^{i,r} = +1$ in the region illuminated by the incident (reflected) field and $\epsilon^{i,r} = -1$ in the shadow region of the incident (reflected) field.

On the shadow boundary of the incident (reflected) field, $\hat{s} = s^{i,r}$ and the argument of the function f in (3.13) is zero. For points not on a shadow boundary, and for $k \to \infty$, we may introduce the asymptotic expansion of f given in Appendix A. Then we find that

$$u \sim \eta(\epsilon^i)e^{iks^i}z_0^i + \eta(\epsilon^r)e^{iks^r}z_0^r + e^{ik\hat{s}}(k\sigma)^{-\frac{1}{2}}D + O(k^{-\frac{3}{2}}), \tag{3.15}$$

where $\eta(x)$ is the unit-step function (see Appendix A) and

$$D = -\frac{e^{i\pi/4}}{2(2\pi)^{\frac{1}{2}}\sin\beta}[\sec\tfrac{1}{2}(\varphi+\varphi_0) \pm \sec\tfrac{1}{2}(\varphi-\varphi_0)]. \tag{3.16}$$

The factor $\eta(\epsilon^{i,r})$ is one in the illuminated region of the incident (reflected) field and zero in the shadow region. Thus the first two terms in (3.15) are just the geometrical-optics solution. If we remember that the angle β is constant on the edge and the curvature κ of the edge is zero, we see that (2.31) yields $\rho = \infty$ and (2.35) becomes

$$\hat{z}_0 = \sigma^{-\frac{1}{2}}Dz_0^i = \sigma^{-\frac{1}{2}}D. \tag{3.17}$$

Thus we recognize that the third term in (3.15) is the leading term of Keller's diffracted wave (2.23) and D is the diffraction coefficient. Now, however, (3.16) provides a formula[12] for D. In Keller's theory it is assumed that D is given by (3.16) for an arbitrary screen, with the angle β, φ, and φ_0 defined at each point on the edge as in Fig. 1.

4. DIFFRACTION BY A SCREEN

In this section we reconsider the diffraction problem discussed in Sec. 2: The wave u_0^i given by (2.15) is incident on an arbitrary smooth screen with either boundary condition (2.16). The total field u is a solution of the reduced wave equation (2.1) and the boundary condition; and the scattered field $u - u_0^i$ must be outgoing. In addition u must satisfy an "edge condition,"[13] which we shall introduce shortly.

As we have seen in Sec. 2, Keller's approach to this problem is based on the ansatz (2.2). Motivated by the representation (3.12) of the exact solution of the Sommerfeld problem, we shall introduce a new ansatz for the general diffraction problem. We will find that, away from the shadow boundaries, our results will reduce to those of Sec. 2, but we shall be able to obtain more than just the leading term of the diffracted wave. (In this section we obtain the leading term. In Sec. 5 we obtain the next term, and presumably the process can be continued.) Moreover, we shall not resort to the Sommerfeld solution for the determination of the diffraction coefficient. Instead, we shall find that the coefficient and its generalization for higher-order terms arise as a consequence of

[12] If we introduce Keller's angles, $\alpha = \pi/2 - \varphi_0$, $\theta = \pi/2 + \varphi$, we see that (3.16) agrees with the result given in Ref. 1 or Ref. 2. In these references the factor $k^{-\frac{1}{2}}$ is included in the diffraction coefficient. We prefer to define D so that it is dimensionless.

[13] It is well known that an edge condition is necessary for the solution to be unique. [See C. J. Bouwkamp, Rept. Progr. Phys. 17, 35 (1954).] The Sommerfeld solution in Sec. 3 satisfies this condition. The condition also enters indirectly into the method of Sec. 2 because Keller's diffraction coefficient is obtained from the Sommerfeld solution.

imposing the edge condition on our asymptotic solution. We will also see that the asymptotic solution obtained in this section will be continuous at the shadow boundaries as well as at the edge. For this reason we call it "uniform."

Our asymptotic solution is based on the ansatz

$$u = u^{\mathrm{i}} + u^{\mathrm{r}}, \tag{4.1}$$

where

$$u^{\mathrm{i,r}} \sim e^{ik\hat{s}}\Big(f\{\epsilon^{\mathrm{i,r}}[k(\hat{s} - s^{\mathrm{i,r}})]^{\frac{1}{2}}\} \sum_{m=0}^{\infty} (ik)^{-m} z_m^{\mathrm{i,r}} + k^{-\frac{1}{2}} \sum_{m=0}^{\infty} (ik)^{-m} w_m^{\mathrm{i,r}} \Big). \tag{4.2}$$

The function f is given by (3.11). The second term in (4.2) is not present in (3.13), but such a term appears in the uniform asymptotic expansion of the solution of the problem of diffraction of a cylindrical wave by a half-plane. It is also suggested by the work of Wolfe.[6] In (4.2) we take $\epsilon^{\mathrm{i,r}} = 1$ in the illuminated region of the incident (reflected) wave and $\epsilon^{\mathrm{i,r}} = -1$ in the shadow region. For $\hat{s} - s^{\mathrm{i,r}} > 0$, we may introduce the asymptotic expansion of f which is given in Appendix A. Then we find that

$$u^{\mathrm{i,r}} \sim \eta(\epsilon^{\mathrm{i,r}}) e^{iks^{\mathrm{i,r}}} \sum_{m=0}^{\infty} (ik)^{-m} z_m^{\mathrm{i,r}} + k^{-\frac{1}{2}} e^{ik\hat{s}} \sum_{m=0}^{\infty} (ik)^{-m} \hat{z}_m^{\mathrm{i,r}}, \tag{4.3}$$

where $\hat{z}_m^{\mathrm{i,r}}$ is determined by

$$w_m^{\mathrm{i,r}} = \hat{z}_m^{\mathrm{i,r}} + \frac{\epsilon^{\mathrm{i,r}}}{2\pi^{\frac{1}{2}}} e^{i\pi/4} \sum_{n=0}^{m} (\tfrac{1}{2})_n \frac{z_{m-n}^{\mathrm{i,r}}}{(\hat{s} - s^{\mathrm{i,r}})^{n+\frac{1}{2}}}. \tag{4.4}$$

In (4.3), $\eta(\epsilon^{\mathrm{i,r}})$ is one in the illuminated region of the incident (reflected) wave and is zero in the shadow. By inserting (4.2) into the reduced wave equation, one can obtain equations for the determination of $\hat{s}$, s^{i}, s^{r}, z_m^{i}, z_m^{r}, w_m^{i}, w_m^{r}. It is, however, much simpler to use (4.3) and then use (4.4) to obtain $w_m^{\mathrm{i,r}}$.

It is clear from our work in Sec. 2 that (4.3) will satisfy the reduced wave equation, provided $s^{\mathrm{i,r}}$ and $\hat{s}$ satisfy the eikonal equation; and both $z_m^{\mathrm{i,r}}$, $s^{\mathrm{i,r}}$ together, and $\hat{z}_m^{\mathrm{i,r}}$, $\hat{s}$ together satisfy the transport equations (2.4).

We now impose the boundary conditions (2.16). If we insert (4.1, 4.3) into the first boundary condition ($u = 0$), we obtain (2.18), (2.19), and the additional equations

$$\hat{z}_m^{\mathrm{r}} = -\hat{z}_m^{\mathrm{i}} \quad \text{on } S, \quad m = 0, 1, 2, \cdots. \tag{4.5a}$$

If we use the second boundary condition, we again obtain (2.18), (2.20), and the additional equations

$$\frac{\partial \hat{s}}{\partial n}[\hat{z}_m^{\mathrm{i}} + \hat{z}_m^{\mathrm{r}}] + \frac{\partial \hat{z}_{m-1}^{\mathrm{i}}}{\partial n} + \frac{\partial \hat{z}_{m-1}^{\mathrm{r}}}{\partial n} = 0 \quad \text{on } S, \quad m = 0, 1, 2, \cdots. \tag{4.5b}$$

Since the functions s^{r} and z_m^{r} satisfy exactly the same equations as those of Sec. 2, these functions are identical to those of Keller's theory and may be computed by the method of Sec. 2. The functions s^{i} and z_m^{i} are, of course, given. As in Sec. 2, we assume that the diffracted rays emanate from the edge and that $\hat{s} = s^{\mathrm{i}}$ on the edge. It follows that $\hat{s}$ is identical to the diffracted-phase function of Sec. 2. Furthermore, since $\hat{z}_m^{\mathrm{i,r}}$ and $\hat{s}$ together satisfy the transport equations, $\hat{z}_m^{\mathrm{i,r}}$ is given along the diffracted rays by (2.33) with δ_m replaced by $\delta_m^{\mathrm{i,r}}$. These coefficients will be determined shortly by the *edge condition*: u must have a finite limit at the edge.[14]

Once all of the functions in (4.3) are determined, (4.2) follows from (4.4). Since $\hat{s}$ and $s^{\mathrm{i,r}}$ have been identified as the phase functions of Sec. 2, it follows from Lemma 1 that $\hat{s} - s^{\mathrm{i,r}} \geq 0$ and $\hat{s} = s^{\mathrm{i,r}}$ only on the shadow boundary. We take the radical in (4.2) to mean the nonnegative square root.

We may assume that the functions s^{i} and z_m^{i} are defined everywhere. However, the functions s^{r} and z_m^{r} are so far defined only in the illuminated region of the reflected wave. In order for (4.2) to be defined everywhere, we must continue the reflected wave into its shadow region, i.e., we must continue the functions s^{r} and z_m^{r}. We require that these continuations be smooth (i.e., the functions must have sufficiently many derivatives). This smooth continuation can easily be constructed by continuing the screen smoothly past the edge and extending the reflected rays backward through the screen. Then s^{r} and z_m^{r} are defined along these extended rays by the formulas of Sec. 2. As a consequence they will satisfy the eikonal and transport equations in the shadow region. (If the screen has an analytic representation, we can use its analytic continuation.) It can be shown that, to any given order in k^{-1}, the asymptotic expansion (4.2) for u^{r} is independent of the continuation of s^{r} and z_m^{r}, provided it is sufficiently smooth.[15] In fact, if $k(\hat{s} - s^{\mathrm{r}}) \gg 1$, it is clear from (4.3) that u^{r} does not depend on values of s^{r} and z_m^{r} in the shadow, for there $\eta(\epsilon^{\mathrm{r}}) = 0$. For points so close to the shadow boundary that $k(\hat{s} - s^{\mathrm{r}})$ is finite, it can be shown that the distance to the shadow boundary is of order $k^{-\frac{1}{2}}$. Then, if (4.2) is sufficiently smooth, its values in the shadow are given by a Taylor expansion in the distance, hence

[14] This condition is sufficient to ensure uniqueness of the solution of the diffraction problem. See C. J. Bouwkamp, Rept. Progr. Phys. **17**, 35 (1954); L. M. Levine, Commun. Pure Appl. Math. **17**, 147 (1964).

[15] This assertion depends on the assumption that the smoothness of s^{r} and z_m^{r} implies the smoothness of w_m^{r} at the shadow boundary. We will prove (Lemma 3) that w_0^{r} is continuous at the shadow boundary. Further smoothness properties have not yet been proved.

an expansion in powers of $k^{-\frac{1}{2}}$. The coefficients of the expansion up to any order depend only on derivatives at the shadow boundary, therefore are independent of the continuation provided it is sufficiently smooth. Thus any two continuations differ by a term of arbitrarily high order in $k^{-\frac{1}{2}}$, provided they are sufficiently smooth.

Since $f(x)$ is regular at $x = 0$ and the functions $z_m^{i,r}$ are regular at the edge, we see from (4.2) that the edge condition is satisfied provided the *equivalent edge condition*, $|\lim_{\sigma\to 0} w_m^{i,r}| < \infty$, $m = 0, 1, 2, \cdots$, on each diffracted ray, is satisfied. For $m = 0$ we see that

$$w_0^{i,r} = \hat{z}_0^{i,r} + \frac{\epsilon^{i,r}}{2}\pi^{-\frac{1}{2}}e^{i\pi/4}\frac{z_0^{i,r}}{(\hat{s} - s^{i,r})^{\frac{1}{2}}}. \quad (4.6)$$

In order to apply the edge condition we now expand the functions that appear in (4.6) for small σ. From (2.34)

$$\hat{z}_0^{i,r} = \delta_0^{i,r}\sigma^{-\frac{1}{2}} + O(\sigma^{\frac{1}{2}}). \quad (4.7)$$

Since the functions $z_0^{i,r}$ are regular at the edge $\mathbf{x} = \mathbf{x}_0(\eta)$,

$$z_0^{i,r} = z_0^{i,r}[\mathbf{x}_0(\eta)] + O(\sigma). \quad (4.8)$$

In order to expand $s^{i,r}$ near the edge, we introduce the unit vectors $\mathbf{t}_1 = \mathbf{t} = \dot{\mathbf{x}}_0(\eta)$, $\mathbf{t}_2$, and $\mathbf{t}_3 = \mathbf{t}_1 \times \mathbf{t}_2$ as illustrated in Fig. 1. Since ∇s^i is a unit vector, we may set [at $\mathbf{x} = \mathbf{x}_0(\eta)$]

$$\nabla s^i = \cos\beta\mathbf{t}_1 - \sin\beta\cos\varphi_0\mathbf{t}_2 - \sin\beta\sin\varphi_0\mathbf{t}_3. \quad (4.9)$$

This equation defines the angles β ($0 \le \beta \le \pi$) and φ_0 ($0 \le \varphi_0 \le \pi$). From the law of reflection it follows that

$$\nabla s^r = \cos\beta\mathbf{t}_1 - \sin\beta\cos\varphi_0\mathbf{t}_2 + \sin\beta\sin\varphi_0\mathbf{t}_3, \quad (4.10)$$

and by the law of edge diffraction, the diffracted rays are given by

$$\mathbf{x} = \mathbf{x}_0(\eta) + \sigma\mathbf{U}, \quad (4.11)$$

where

$$\mathbf{U} = \cos\beta\mathbf{t}_1 + \sin\beta\cos\varphi\mathbf{t}_2 - \sin\beta\sin\varphi\mathbf{t}_3 \quad (-\pi \le \varphi \le \pi). \quad (4.12)$$

On a diffracted ray,

$$\begin{aligned} s^{i,r} &= s^{i,r}[\mathbf{x}_0(\eta)] + \sigma\mathbf{U}\cdot\nabla s^{i,r} + O(\sigma^2) \\ &= s^i[\mathbf{x}_0(\eta)] + \sigma[\cos^2\beta - \sin^2\beta\cos(\varphi \pm \varphi_0)] + O(\sigma^2). \end{aligned} \quad (4.13)$$

Hence, since $\hat{s} = s^i[\mathbf{x}_0(\eta)] + \sigma$,

$$\epsilon^{i,r}(\hat{s} - s^{i,r})^{\frac{1}{2}} = \pm(2\sigma)^{\frac{1}{2}}\sin\beta\cos\tfrac{1}{2}(\varphi \pm \varphi_0) + O(\sigma^{\frac{3}{2}}). \quad (4.14)$$

Here we have set

$$\epsilon^{i,r} = \pm\operatorname{sgn}[\cos\tfrac{1}{2}(\varphi \pm \varphi_0)], \quad (4.15)$$

and it can be easily verified that $\epsilon^{i,r}$ is indeed $+1$ in the illuminated region of the incident (reflected) wave and -1 in the shadow. By inserting (4.7), (4.8), and (4.14) into (4.6), we find that (for $m = 0$) the edge condition is satisfied provided

$$\delta_0^{i,r} = \mp\frac{e^{i\pi/4}}{2(2\pi)^{\frac{1}{2}}}\frac{z_0^{i,r}[\mathbf{x}_0(\eta)]}{\sin\beta\cos\tfrac{1}{2}(\varphi \pm \varphi_0)}. \quad (4.16)$$

We may now express $z_0^r[\mathbf{x}_0(\eta)]$ in terms of $z_0^i[x_0(\eta)]$ by using (2.19) or (2.21). (For $m = 0$, the latter becomes simply $z_0^r = z_0^i$.) Then from (2.34) we have, for the boundary condition $u = 0$ ($\partial u/\partial n = 0$) on S,

$$\hat{z}_0^i(\sigma) = -\frac{e^{i\pi/4}}{2(2\pi)^{\frac{1}{2}}}\frac{z_0^i[\mathbf{x}_0(\eta)]}{\sin\beta\cos\tfrac{1}{2}(\varphi + \varphi_0)}\left|\sigma\left(1 + \frac{\sigma}{\rho}\right)\right|^{-\frac{1}{2}} \quad (4.17)$$

and

$$\hat{z}_0^r(\sigma) = \frac{e^{i\pi/4}}{2(2\pi)^{\frac{1}{2}}}\frac{z_0^r[\mathbf{x}_0(\eta)]}{\sin\beta\cos\tfrac{1}{2}(\varphi - \varphi_0)}\left|\sigma\left(1 + \frac{\sigma}{\rho}\right)\right|^{-\frac{1}{2}}; \quad z_0^r[\mathbf{x}_0(\eta)] = \mp z_0^i[\mathbf{x}_0(\eta)]. \quad (4.18)$$

Thus we have found the terms of (4.3) and (4.2) for which $m = 0$. The leading term of the nonuniform expansion (4.1, 4.3) is just the geometrical-optics (incident and reflected) field. The second term is given by

$$\hat{u} \sim k^{-\frac{1}{2}}e^{ik\hat{s}}\hat{z}_0, \quad (4.19)$$

where

$$\hat{z}_0 = \hat{z}_0^i + \hat{z}_0^r = Dz_0^i[\mathbf{x}_0(\eta)]\left|\sigma\left(1 + \frac{\sigma}{\rho}\right)\right|^{-\frac{1}{2}}, \quad (4.20)$$

and, for the boundary condition $u = 0$ ($\partial u/\partial n = 0$),

$$D = -\frac{e^{i\pi/4}}{2(2\pi)^{\frac{1}{2}}\sin\beta}[\sec\tfrac{1}{2}(\varphi + \varphi_0) \pm \sec\tfrac{1}{2}(\varphi - \varphi_0)]. \quad (4.21)$$

From (3.16) we see that D is Keller's diffraction coefficient, and, by comparing (4.20) with (2.35) and (4.19) with (2.23), we see that away from the shadow boundary our result reduces to the solution obtained by the geometrical theory of diffraction.

The terms of the uniform expansion [(4.1) and (4.2)] obtained so far are given by (4.1) and

$$u^{i,r} \sim e^{ik\hat{s}}(f\{\epsilon^{i,r}[k(\hat{s} - s^{i,r})]^{\frac{1}{2}}\}z_0^{i,r} + k^{-\frac{1}{2}}w_0^{i,r}). \quad (4.22)$$

Here $w_0^{i,r}$ is given by (4.6), (4.17), and (4.18). Let us compare the present solution (4.22) with Keller's solution. By construction, the solution (4.22) is *finite at the edge* because we have satisfied the edge condition, whereas Keller's geometrical theory fails at the

edge. However, at all other caustic points of the diffracted wave, as well as caustic points of the incident and reflected waves, both our solution and Keller's solution will fail. At the shadow boundary, the geometrical theory breaks down because the geometrical-optics field is discontinuous and the diffraction coefficient becomes infinite. We will now verify that the uniform solution (4.22) is *continuous at the shadow boundary*. Since $f(x)$ is regular at $x = 0$ and $z_0^{i,r}$ is smooth, it is sufficient to examine only the second term of (4.22) which is given by (4.6). At the shadow boundary of the incident (reflected) wave, $\hat{s} - s^{i,r} = 0$ and $\cos \frac{1}{2}(\varphi \pm \varphi_0) = 0$; hence both terms in (4.6) become infinite and we must evaluate the limit of the sum carefully. This is done in Appendix B, where we prove the following:

Lemma 3: $w_0^{i,r}$ is continuous at the shadow boundary of the incident (reflected) wave.

It follows that *the leading term* [(4.1), (4.2)] *of our solution* [(4.1), (4.2)] *is continuous everywhere except at caustics of the incident and reflected waves and any caustic of the diffracted wave other than the edge.*

We have not yet examined Eqs. (4.5a) and (4.5b), which must be satisfied if (4.1) is to satisfy the boundary condition (2.16a) and (2.16b). We find in the next section that these equations are indeed satisfied if the screen is a portion of a plane, provided none of the diffracted rays strike another part of the edge. If that occurs, it is necessary to introduce additional terms into (4.1). Examples of such problems of *multiple diffraction* will be treated by our method in a forthcoming paper. They are treated by Keller's method in Refs. 1 and 2. If the screen is curved, still more terms must be introduced into (4.1) if the boundary condition is to be satisfied. These terms correspond to *secondary reflected waves* which are produced when reflected or diffracted rays strike the screen at an angle of incidence less than $\pi/2$, and *creeping waves*[7] which are excited when incident, reflected, or diffracted rays are tangent to a convex portion of the screen. The existence of such terms was suggested by earlier experience with asymptotic methods in other problems but has not yet been verified for our problem.

5. DIFFRACTION BY A PLANE SCREEN

In order to compute higher-order terms in our expansion (4.1) and (4.2) or (4.1) and (4.3), we may use (2.33) for $\hat{z}_m^{i,r}$. Thus to determine $\hat{z}_1^{i,r}$ it is necessary to find $\Delta\hat{z}_0^{i,r}$. But $\hat{z}_0^{i,r}$ is given by (4.17) and (4.18) as a function of σ, φ, η. Thus it is necessary to transform the Laplacian to these coordinates which are not orthogonal. In order to simplify this and other parts of our work, we specialize our problem and consider only plane screens in this section.

The screen lies in the plane $x_3 = 0$ and is bounded by a smooth curve $\mathbf{x} = \mathbf{x}_0(\eta) = (x_0^1, x_0^2, 0)$. η denotes arclength on the edge; hence $\mathbf{t}_1 = \dot{\mathbf{x}}_0$ is a unit tangent vector. A dot denotes differentiation with respect to η. We assume that the incident wave comes from the region $x_3 < 0$. Then the x_3 axis coincides with the z axis of Fig. 1 and $\mathbf{t}_3 = (0, 0, -1)$. The parameter η is chosen in such a way that the vector $\mathbf{t}_2 = \mathbf{t}_3 \times \mathbf{t}_1$ points away from the screen as in Fig. 1. If $\mathbf{n}$ denotes the unit normal to the edge, then $\mathbf{t}_2 = \pm\mathbf{n}$ and the upper (lower) sign holds when the screen is locally concave (convex). In either case the curvature κ is given by

$$\kappa = \mathbf{n} \cdot \dot{\mathbf{t}}_1 = \mp\tilde{\kappa}, \tag{5.1}$$

where

$$\tilde{\kappa} = \dot{x}_0^1\ddot{x}_0^2 - \dot{x}_0^2\ddot{x}_0^1 \tag{5.2}$$

is the "signed curvature." Then

$$\dot{\mathbf{t}}_1 = \kappa\mathbf{n} = -\tilde{\kappa}\mathbf{t}_2, \quad \dot{\mathbf{t}}_2 = \tilde{\kappa}\mathbf{t}_1, \quad \dot{\mathbf{t}}_3 = 0. \tag{5.3}$$

From (4.11) and (4.12) the diffracted rays are given by

$$\mathbf{x} = \mathbf{x}_0(\eta) + \sigma\mathbf{U}(\eta, \varphi), \tag{5.4}$$

where

$$\mathbf{U} = \cos\beta\mathbf{t}_1 + \sin\beta\cos\varphi\mathbf{t}_2 - \sin\beta\sin\varphi\mathbf{t}_3, \tag{5.5}$$

and $\beta = \beta(\eta)$. By comparing (5.5) with (2.29), we see that $\cos\alpha = \pm\cos\varphi$; hence (2.31) yields

$$\rho = \sin\beta/(\tilde{\kappa}\cos\varphi - \dot{\beta}). \tag{5.6}$$

By using (5.3), (5.6), and standard formulas,[16] the Laplacian can be computed in σ, φ, η coordinates. This is done in Appendix C. We will soon make use of the final result of that appendix, which is an expansion of the Laplacian for small σ.

For the case of a plane screen considered here, we can give a simplified representation of the reflected wave. It is easy to show that now

$$s^r(x_1, x_2, x_3) = s^i(x_1, x_2, -x_3), \tag{5.7}$$

and, for the two boundary conditions $u = 0$ and $\partial u/\partial n = \partial u/\partial x_3 = 0$,

$$z_m^r(x_1, x_2, x_3) = \mp z_m^i(x_1, x_2, -x_3). \tag{5.8}$$

To verify (5.7) and (5.8) we note that s^r satisfies the eikonal equation (2.3) because s^i does. Furthermore, s^r and z_m^r satisfy the system of transport equations (2.4) because s^i and z_m^i satisfy the same system.

[16] J. A. Stratton, *Electromagnetic Theory* (McGraw-Hill Book Co., New York, 1964).

Finally, for the boundary condition $u = 0$, (2.19) is clearly satisfied, and for the boundary condition $\partial u/\partial n = 0$, (2.21) is satisfied because $(\partial/\partial x_3)z^{\mathrm{r}}_m + (\partial/\partial x_3)z^{\mathrm{i}}_m = 0$ on the screen. For both boundary conditions (2.18) is satisfied.

The functions $\hat{z}^{\mathrm{i,r}}_1$ are given by (2.33) for $m = 1$ once the coefficients $\delta^{\mathrm{i,r}}_1$ are determined. As in the evaluation of $\delta^{\mathrm{i,r}}_0$, the values of $\delta^{\mathrm{i,r}}_1$ are uniquely determined by the edge condition. To apply this condition it is necessary to compute [cf. (4.4)]

$$w^{\mathrm{i,r}}_1 = \hat{z}^{\mathrm{i,r}}_1 + \frac{\epsilon^{\mathrm{i,r}}}{2\pi^{\frac{1}{2}}} e^{i\pi/4}\left[\frac{z^{\mathrm{i,r}}_1}{(\hat{s} - s^{\mathrm{i,r}})^{\frac{1}{2}}} + \frac{1}{2}\frac{z^{\mathrm{i,r}}_0}{(\hat{s} - s^{\mathrm{i,r}})^{\frac{3}{2}}}\right] \tag{5.9}$$

for small values of σ. Then the edge condition, $|\lim_{\sigma\to 0} w^{\mathrm{i,r}}_1| < \infty$, will determine the value of $\delta^{\mathrm{i,r}}_1$. The first step is to find $\hat{z}^{\mathrm{i,r}}_1$ for small σ and this requires the determination of $\Delta\hat{z}^{\mathrm{i,r}}_0$ for small σ. Therefore we apply (C16) to the first two terms of the expansion of (4.17) and (4.18) for small values of σ. The computation, although somewhat long, is straightforward and the surprisingly simple result is

$$\Delta\hat{z}^{\mathrm{i,r}}_0 = \frac{\mp 1}{4(2\pi)^{\frac{1}{2}}} e^{i\pi/4} \frac{z^{\mathrm{i,r}}_0[\mathbf{x}_0(\eta)]}{\sin^3\beta\cos^3\frac{1}{2}(\varphi \pm \varphi_0)} \times \left[\sigma^{-\frac{5}{2}} - \frac{1}{2\rho}\sigma^{-\frac{3}{2}} + O(\sigma^{-\frac{1}{2}})\right]. \tag{5.10}$$

The functions $\hat{z}^{\mathrm{i,r}}_1$ are now given by [cf. (2.33)]

$$\hat{z}^{\mathrm{i,r}}_1 = \left[\sigma\left(1 + \frac{\sigma}{\rho}\right)\right]^{-\frac{1}{2}}\left\{\delta^{\mathrm{i,r}}_1 - \tfrac{1}{2}\int_0^{\sigma} g^{\mathrm{i,r}}(t)\,dt\right\}, \tag{5.11}$$

where, from (5.10),

$$g^{\mathrm{i,r}}(t) = \left[t\left(1 + \frac{t}{\rho}\right)\right]^{\frac{1}{2}}\Delta\hat{z}^{\mathrm{i,r}}_0(t) = \mp\frac{1}{4(2\pi)^{\frac{1}{2}}} e^{i\pi/4}\frac{z^{\mathrm{i,r}}_0[\mathbf{x}_0(\eta)]}{\sin^3\beta\cos^3\frac{1}{2}(\varphi \pm \varphi_0)} \times [t^{-2} + O(1)]. \tag{5.12}$$

The finite part integral in (5.11) is now easily evaluated. (We need only ignore the infinite contribution from the lower limit of integration.) Thus we obtain

$$\hat{z}^{\mathrm{i,r}}_1 = \delta^{\mathrm{i,r}}_1\sigma^{-\frac{1}{2}} \mp \frac{1}{8(2\pi)^{\frac{1}{2}}} e^{i\pi/4}\frac{z^{\mathrm{i,r}}_0[\mathbf{x}_0(\eta)]}{\sin^3\beta\cos^3\frac{1}{2}(\varphi \pm \varphi_0)} \times \left[\sigma^{-\frac{3}{2}} - \frac{\sigma^{-\frac{1}{2}}}{2\rho}\right] + O(\sigma^{\frac{1}{2}}). \tag{5.13}$$

It is interesting to note that the term of order t^{-1} in (5.12) is missing. Such a term would have led to a term involving $\log\sigma$ in (5.13) and then the edge condition could not have been satisfied.

In order to compute the remaining terms in (5.9) for small σ, we note that $z^{\mathrm{i,r}}_0$ and $z^{\mathrm{i,r}}_1$ are regular in a neighborhood of the edge; hence

$$z^{\mathrm{i,r}}_1 = z^{\mathrm{i,r}}_1[\mathbf{x}_0(\eta)] + O(\sigma) \tag{5.14}$$

and

$$z^{\mathrm{i,r}}_0 = z^{\mathrm{i,r}}_0[\mathbf{x}_0(\eta)] + a^{\mathrm{i,r}}\sigma + O(\sigma^2), \tag{5.15}$$

where

$$a^{\mathrm{i,r}} = \mathbf{U}(\eta, \varphi)\cdot\nabla z^{\mathrm{i,r}}_0[\mathbf{x}_0(\eta)]. \tag{5.16}$$

Then, by extending (4.13) to one more term, we obtain

$$s^{\mathrm{i,r}} = s^{\mathrm{i}}[\mathbf{x}_0(\eta)] + \sigma[\cos^2\beta - \sin^2\beta\cos(\varphi \pm \varphi_0)] + b^{\mathrm{i,r}}\sigma^2 + O(\sigma^3), \tag{5.17}$$

where

$$2b^{\mathrm{i,r}} = (\mathbf{U}\cdot\nabla)(\mathbf{U}\cdot\nabla)s^{\mathrm{i,r}} = (\mathbf{U}\cdot\nabla)^2 s^{\mathrm{i,r}}. \tag{5.18}$$

Since $s = s^{\mathrm{i}}[\mathbf{x}_0(\eta)] + \sigma$, we find that

$$\hat{s} - s^{\mathrm{i,r}} = 2\sigma\sin^2\beta\cos^2\tfrac{1}{2}(\varphi \pm \varphi_0) - b^{\mathrm{i,r}}\sigma^2 + O(\sigma^3). \tag{5.19}$$

It is now an easy matter to compute $(\hat{s} - s^{\mathrm{i,r}})^{-\frac{1}{2}}$ and $(\hat{s} - s^{\mathrm{i,r}})^{-\frac{3}{2}}$. Then we may evaluate (5.9) for small σ. In so doing we must use (4.15). It is clear that the result will be of the form

$$w^{\mathrm{i,r}}_1 = p^{\mathrm{i,r}}_1\sigma^{-\frac{3}{2}} + p^{\mathrm{i,r}}_2\sigma^{-\frac{1}{2}} + O(\sigma^{\frac{1}{2}}), \tag{5.20}$$

but it is remarkable that the calculation yields

$$p^{\mathrm{i,r}}_1 = 0. \tag{5.21}$$

Then the edge condition will be satisfied if and only if $p^{\mathrm{i,r}}_2 = 0$, and this condition uniquely determines $\delta^{\mathrm{i,r}}_1$. The result is

$$\delta^{\mathrm{i,r}}_1 = \frac{\mp e^{i\pi/4}}{2(2\pi)^{\frac{1}{2}}\sin\beta}\left\{\frac{z^{\mathrm{i,r}}_1(\mathbf{x}_0)}{\cos\frac{1}{2}(\varphi \pm \varphi_0)} + \frac{1}{4\sin^2\beta\cos^3\frac{1}{2}(\varphi \pm \varphi_0)}\left[a^{\mathrm{i,r}} + \frac{1}{2\rho}z^{\mathrm{i,r}}_0(\mathbf{x}_0)\right] + \frac{3b^{\mathrm{i,r}}z^{\mathrm{i,r}}_0(\mathbf{x}_0)}{16\sin^4\beta\cos^5\frac{1}{2}(\varphi \pm \varphi_0)}\right\}. \tag{5.22}$$

With this value of $\delta^{\mathrm{i,r}}_1$, $\hat{z}^{\mathrm{i,r}}_1$ is given by (2.33) for $m = 1$.

We note from (5.20) that the expansion for small σ of $w^{\mathrm{i,r}}_1$ has two terms which become infinite at $\sigma = 0$. One of them automatically vanishes, and the vanishing of the other term, which is required by the edge condition, uniquely determines $\delta^{\mathrm{i,r}}_1$. In general it can be seen that the expansion of $w^{\mathrm{i,r}}_m$ for small σ will be of the form

$$w^{\mathrm{i,r}}_m = p^{\mathrm{i,r}}_1\sigma^{-m-\frac{1}{2}} + p^{\mathrm{i,r}}_2\sigma^{-m+\frac{1}{2}} + \cdots + p^{\mathrm{i,r}}_{m+1}\sigma^{-\frac{1}{2}} + O(\sigma^{\frac{1}{2}}). \tag{5.20'}$$

We conjecture that the first m terms will vanish automatically, i.e., $p_j^{i,r} = 0$ for $j = 1, 2, \cdots, m$. Then the vanishing of the remaining term, i.e., the requirement $p_{m+1}^{i,r} = 0$, will uniquely determine $\delta_m^{i,r}$. But this conjecture has not yet been proved.

While in the realm of conjecture, we should also consider the behavior of our asymptotic solution (4.1, 4.2) in the neighborhood of the shadow boundaries. At the end of Sec. 4 we proved that the leading term (4.1, 4.22) is continuous at the shadow boundaries as well as at the edge. We also conjecture that to every order (4.1, 4.2) is not only continuous but smooth at the shadow boundaries. This conjecture also has not yet been verified. (*Note added in proof:* As mentioned in Sec. 1, both conjectures have recently been proven.) (At the edge, *derivatives* of the asymptotic solution may become singular, but the exact solution has the same property.)

Our nonuniform expansion [(4.1) and (4.3)] may be written in the form

$$u = \eta(\epsilon^i)u_0^i + \eta(\epsilon^r)u_0^r + \hat{u}, \tag{5.23}$$

where

$$u_0^{i,r} \sim e^{iks^{i,r}} \sum_{m=0}^{\infty} (ik)^{-m} z_m^{i,r}, \tag{5.24}$$

$$\hat{u} \sim k^{-\frac{1}{2}} e^{ik\hat{s}} \sum_{m=0}^{\infty} (ik)^{-m} \hat{z}_m, \tag{5.25}$$

and

$$\hat{z}_m = \hat{z}_m^i + \hat{z}_m^r. \tag{5.26}$$

By comparing (5.23)–(5.25) with the results of Sec. 2, we find that they are identical. In Sec. 4 we showed that the leading term of the diffracted wave (5.25) agrees exactly with Keller's formula. The geometrical theory of diffraction (Sec. 2) is not capable of determining higher-order terms in (5.25), but our theory yields these terms as well.[17]

Since $\hat{z}_m^i$ and $\hat{z}_m^r$ both satisfy (2.33) with δ_m replaced by δ_m^i and δ_m^r, it is clear from (5.26) that $\hat{z}_m$ satisfies (2.33) with

$$\delta_m = \delta_m^i + \delta_m^r. \tag{5.27}$$

Since we have determined δ_1^i and δ_1^r, we can give the value of δ_1. First, however, we simplify the terms $a^{i,r}$ and $b^{i,r}$ that appear in (5.22). We introduce the tangential- and normal-gradient operators defined by

$$\nabla_n = \mathbf{N}\mathbf{N}\cdot\nabla, \quad \nabla_t = \nabla - \nabla_n,$$
$$\mathbf{N} = (0, 0, 1) = -\mathbf{t}_3. \tag{5.28}$$

Then (5.7) and (5.8) yield

$$\nabla_n s^r = -\nabla_n s^i, \quad \nabla_t s^r = \nabla_t s^i,$$
$$\nabla_n z_m^r = \pm\nabla_n z_m^i, \quad \nabla_t z_m^r = \mp\nabla_t z_m^i, \tag{5.29}$$

and from (5.5), (5.16), and (5.18) we obtain

$$a^i = \mathbf{U}\cdot\nabla z_0^i, \quad a^r = \mp\mathbf{U}\cdot\nabla z_0^i \pm 2\sin\beta\sin\varphi\frac{\partial z_0^i}{\partial n}, \tag{5.30}$$

and

$$b^i = \tfrac{1}{2}(\mathbf{U}\cdot\nabla)^2 s^i,$$
$$b^r = \tfrac{1}{2}(\mathbf{U}\cdot\nabla)^2 s^i - 2(\mathbf{U}\cdot\nabla_t)(\mathbf{U}\cdot\nabla_n)s^i. \tag{5.31}$$

It now follows from (5.22) and (5.27) that

$$\begin{aligned}\delta_1 = {} & \frac{-e^{i\pi/4}}{2(2\pi)^{\frac{1}{2}}\sin\beta} \\ & \times\Big\{[\sec\tfrac{1}{2}(\varphi+\varphi_0) \pm \sec\tfrac{1}{2}(\varphi-\varphi_0)]z_1^i(\mathbf{x}_0) \\ & + \tfrac{1}{4}\csc^2\beta[\sec^3\tfrac{1}{2}(\varphi+\varphi_0) \pm \sec^3\tfrac{1}{2}(\varphi-\varphi_0)] \\ & \quad\times\left[\mathbf{U}\cdot\nabla z_0^i(\mathbf{x}_0) + \frac{1}{2\rho}z_0^i(\mathbf{x}_0)\right] \\ & + \tfrac{3}{32}\csc^4\beta[\sec^5\tfrac{1}{2}(\varphi+\varphi_0) \pm \sec^5\tfrac{1}{2}(\varphi-\varphi_0)] \\ & \quad\times[(\mathbf{U}\cdot\nabla)^2 s^i]z_0^i(\mathbf{x}_0) \\ & \mp \tfrac{1}{2}\csc\beta\sin\varphi\sec^3\tfrac{1}{2}(\varphi-\varphi_0)\frac{\partial z_0^i}{\partial n}(\mathbf{x}_0) \\ & \mp \tfrac{3}{8}\csc^4\beta\sec^5\tfrac{1}{2}(\varphi-\varphi_0) \\ & \quad\times[(\mathbf{U}\cdot\nabla_t)(\mathbf{U}\cdot\nabla_n)s^i]z_0^i(\mathbf{x}_0)\Big\}. \end{aligned} \tag{5.32}$$

From (4.21) we see that (5.32) can be written in the form

$$\delta_1 = Dz_1^i(\mathbf{x}_0) + D_1 z_0^i(\mathbf{x}_0), \tag{5.33}$$

where D is the "zero-order diffraction coefficient" (Keller's diffraction coefficient) and the "first-order diffraction coefficient" D_1 is a linear differential operator defined by (5.32).

It may happen that the incident rays are tangent to the screen. In this case of "grazing incidence," our results have some special features of interest. There are two cases to consider, depending on whether $\varphi_0 = 0$ or $\varphi_0 = \pi$. (See Fig. 1.) For $\varphi_0 = \pi$, the diffraction problem is pathological. This case will be treated in a sequel to the present paper.

If $\varphi_0 = 0$, the whole region is in the shadow of the reflected wave and illuminated by the incident wave, i.e., $\epsilon^i \equiv 1$ and $\epsilon^r \equiv -1$, and the shadow boundary coincides with the screen. As pointed out in Sec. 4, our results are independent of the values of the functions z_m^r and s^r, provided those functions are sufficiently smooth. Therefore we may continue to define z_m^r and s^r by (5.7) and (5.8). For the boundary condition $\partial u/\partial n = 0$, we see from (4.21) that $D = 0$ and from (4.20) that $\hat{z}_0 \equiv 0$. In this case, the leading term of the diffracted field (5.25) is given by

$$\hat{u} \sim k^{-\frac{1}{2}} e^{ik\hat{s}}(ik)^{-1}\hat{z}_1, \tag{5.34}$$

[17] The determination of the terms z_m for $m = 2, 3, \cdots$ depends on the validity of our first conjecture.

and it is important to compute $\hat{z}_1$. This can easily be done now because $\nabla_n s^i \equiv 0$ on S, and therefore (5.32) becomes

$$\delta_1 = \frac{-e^{i\pi/4}}{4(2\pi)^{\frac{1}{2}} \sin^2 \beta} \sin \varphi \sec^3 \frac{\varphi}{2} \frac{\partial z_0^i}{\partial n}(\mathbf{x}_0). \quad (5.35)$$

Furthermore, since $\hat{z}_0 \equiv 0$, the integral term in (2.33) is absent for $m = 1$. It follows that

$$\hat{u} \sim k^{-\frac{3}{2}} e^{ik\hat{s}} D' \frac{\partial z_0^i}{\partial n}(\mathbf{x}_0) \left| \sigma\left(1 + \frac{\sigma}{\rho}\right) \right|^{-\frac{1}{2}}, \quad (5.36)$$

where D' is a special diffraction coefficient given by

$$D' = -\frac{e^{-i\pi/4}}{2(2\pi)^{\frac{1}{2}} \sin^2 \beta \cos^2 (\varphi/2)} \frac{\sin (\varphi/2)}{}. \quad (5.37)$$

This result was also obtained by Keller by expanding the exact solution of a special half-plane diffraction problem. It is easily seen that our results agree with those given by Keller in Ref. 1.[18]

In order to complete our general treatment of diffraction by a plane screen, it is necessary to verify that the conditions (4.5a) and (4.5b) are satisfied. These conditions result from imposing the boundary condition (2.16a) and (2.16b) on the asymptotic solution [(4.1) and (4.3)]. First we show that

$$\hat{z}_m^r(x_1, x_2, x_3) = \pm \hat{z}_m^i(x_1, x_2, -x_3), \quad m = 0, 1, 2, \cdots. \quad (5.38)$$

From (5.4) and (5.5) it is clear that the points with rectangular coordinates $(x_1, x_2, \pm x_3)$ will have ray coordinates $(\sigma, \pm\varphi, \eta)$. Hence it follows from (4.17) and (4.18) that (5.38) is valid for $m = 0$. The validity of (5.38) for arbitrary m can then be established by an induction argument. Secondly we state that

$$\hat{z}_m^i(x_1, x_2, +0) = -\hat{z}_m^i(x_1, x_2, -0),$$
$$\frac{\partial \hat{z}_m^i}{\partial n}(x_1, x_2, +0) = -\frac{\partial \hat{z}_m^i}{\partial n}(x_1, x_2, -0). \quad (5.39)$$

To prove this we extend the domain of definition of the functions $\hat{z}_m^i(\sigma, \varphi, \eta)$ to arbitrary values of φ. (Only the interval $-\pi \leq \varphi \leq \pi$ corresponds to physical space.) Then it follows from (4.17) that

$$\hat{z}_m^i(\sigma, \varphi + 2\pi, \eta) = -\hat{z}_m^i(\sigma, \varphi, \eta) \quad (5.40)$$

for $m = 0$, and (5.40) can be established for $m = 0, 1, 2, \cdots$, by another induction argument. If we now take $\varphi = -\pi$, we find that

$$\hat{z}_m^i(\sigma, \pi, \eta) = -\hat{z}_m^i(\sigma, -\pi, \eta),$$
$$\frac{\partial \hat{z}_m^i}{\partial \varphi}(\sigma, \pi, \eta) = -\frac{\partial \hat{z}_m^i}{\partial \varphi}(\sigma, -\pi, \eta), \quad (5.41)$$

which is equivalent to (5.39). Finally (5.38) and (5.39) imply (4.5a) and (4.5b) because $\partial\hat{s}/\partial n$ vanishes on the screen.

[18] We must first correct an error in the last part of Eq. (12) of Ref. 1 which has the wrong sign. Then the results agree because $\varphi = \theta - \pi/2$. Note that here D' has been defined so that it is dimensionless.

APPENDIX A: A SPECIAL FUNCTION

Let

$$f(x) = \pi^{-\frac{1}{2}} e^{-i\pi/4} e^{-ix^2} \int_{-\infty}^{x} e^{it^2} dt. \quad (A1)$$

This is an entire function. It is closely related to the Fresnel integral functions. For large real values of x, its asymptotic expansion is

$$f(x) \sim e^{-ix^2}\eta(x) - \tfrac{1}{2}\pi^{-\frac{1}{2}} e^{i\pi/4} x^{-1} \sum_{n=0}^{\infty} (\tfrac{1}{2})_n (ix^2)^{-n}, \quad x \to \pm\infty, \quad (A2)$$

where

$$(\tfrac{1}{2})_0 = 1, \quad (\tfrac{1}{2})_n = \tfrac{1}{2}(\tfrac{1}{2} + 1) \cdots (\tfrac{1}{2} + n - 1), \quad n = 1, 2, 3, \cdots, \quad (A3)$$

and $\eta(x)$ is the unit step function. Thus $\eta(x) = 1$ for $x > 0$ and $\eta(x) = 0$ for $x < 0$.

APPENDIX B: PROOF OF LEMMA 3

From (4.9) and (4.10) we see that the unit vector in the direction of the incident (reflected) ray is

$$\mathbf{U}_1 = \nabla s^{i,r} = \cos \beta \mathbf{t}_1 - \sin \beta \cos \varphi_0 \mathbf{t}_2 \mp \sin \beta \sin \varphi_0 \mathbf{t}_3. \quad (B1)$$

Let

$$\mathbf{U}_2 = \pm \sin \varphi_0 \mathbf{t}_2 - \cos \varphi_0 \mathbf{t}_3 \quad (B2)$$

and

$$\mathbf{U}_3 = \mathbf{U}_1 \times \mathbf{U}_2 = \sin \beta \mathbf{t}_1 + \cos \beta \cos \varphi_0 \mathbf{t}_2 \pm \cos \beta \sin \varphi_0 \mathbf{t}_3. \quad (B3)$$

From (4.12) we have the unit vector in the direction of the diffracted ray:

$$\mathbf{U} = \cos \beta \mathbf{t}_1 + \sin \beta \cos \varphi \mathbf{t}_2 - \sin \beta \sin \varphi \mathbf{t}_3. \quad (B4)$$

Let $\varphi = \mp\varphi_0 \pm \pi \mp \zeta$. Then, for small ζ,

$$\cos \varphi = -\cos \varphi_0 + \zeta \sin \varphi_0 + \tfrac{1}{2}\zeta^2 \cos \varphi_0 + O(\zeta^3), \quad (B5)$$

$$\sin \varphi = \pm(\sin \varphi_0 + \zeta \cos \varphi_0 - \tfrac{1}{2}\zeta^2 \sin \varphi_0) + O(\zeta^3), \quad (B6)$$

and

$$\mathbf{U} = \mathbf{U}_1 \pm \zeta \sin \beta \mathbf{U}_2 + \tfrac{1}{2}\zeta^2 \sin \beta[-\sin \beta \mathbf{U}_1 + \cos \beta \mathbf{U}_3] + O(\zeta^3). \quad (B7)$$

Let $\mathbf{Q}$ be a point on the edge and let $\mathbf{P}$ and $\hat{\mathbf{P}}$ be points on the incident (reflected) and diffracted rays emanating from $\mathbf{Q}$ at a distance σ from $\mathbf{Q}$. Then

$$\mathbf{P} = \mathbf{Q} + \sigma \mathbf{U}_1, \quad \hat{\mathbf{P}} = \mathbf{Q} + \sigma \mathbf{U}. \quad (B8)$$

In a Cartesian coordinate system with basis vectors $\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3$,

$$\mathbf{h} = \hat{\mathbf{P}} - \mathbf{P} = \sigma(\mathbf{U} - \mathbf{U}_1)$$
$$= \sigma(-\tfrac{1}{2}\zeta^2 \sin^2\beta,\ \pm\zeta \sin\beta,\ +\tfrac{1}{2}\zeta^2 \sin\beta\cos\beta) + O(\zeta^3). \quad \text{(B9)}$$

Hence, for $s = s^{i,r}$,

$$s(\hat{\mathbf{P}}) = s(\mathbf{P} + \mathbf{h}) = s(\mathbf{P}) + \mathbf{h}\cdot\nabla s(\mathbf{P}) + \tfrac{1}{2}\sum_{i,j} h_i h_j s_{ij}(\mathbf{P}) + \cdots, \quad \text{(B10)}$$

where $s_{ij} = \partial^2 s/\partial x_i \partial x_j$. But $s(\mathbf{P}) = s^i(\mathbf{Q}) + \sigma = \hat{s}(\hat{\mathbf{P}})$. Therefore, since $\nabla s = (1, 0, 0)$,

$$\hat{s}(\hat{\mathbf{P}}) - s(\hat{\mathbf{P}}) = \tfrac{1}{2}\zeta^2\sigma \sin^2\beta\,[1 - \sigma s_{22}(\mathbf{P})] + O(\zeta^3). \quad \text{(B11)}$$

Let $x_1 = f(x_2, x_3)$ be the equation of the incident (reflected) wavefront passing through $\mathbf{P}$. Then $s(f, x_2, x_3) = \text{const}$. Therefore

$$s_1 f_\nu + s_\nu = 0, \quad \nu = 2, 3, \quad \text{(B12)}$$

and

$$s_{11} f_\mu f_\nu + s_{1\mu} f_\nu + s_1 f_{\nu\mu} + s_{1\nu} f_\mu + s_{\nu\mu} = 0, \quad \nu, \mu = 2, 3. \quad \text{(B13)}$$

Since $(\nabla s)^2 = s_1^2 + s_2^2 + s_3^2 = 1$,

$$s_1 s_{1\nu} + s_2 s_{2\nu} + s_3 s_{3\nu} = 0, \quad \nu = 1, 2, 3. \quad \text{(B14)}$$

The wavefront can be represented parametrically with parameters x_2, x_3 in the form $\mathbf{x} = [f(x_2, x_3), x_2, x_3]$. Then $\mathbf{x}_2 = (f_2, 1, 0)$, $\mathbf{x}_3 = (f_3, 0, 1)$, and $\mathbf{x}_{\nu\mu} = (f_{\nu\mu}, 0, 0)$. At $\mathbf{P}$, since $\nabla s = (s_1, s_2, s_3) = (1, 0, 0)$, we see that $f_2 = f_3 = 0, s_{11} = s_{12} = s_{13} = 0, s_{\nu\mu} = -f_{\nu\mu}$ $(\nu, \mu = 2, 3)$, $\mathbf{x}_2 = (0, 1, 0) = \mathbf{U}_2$, $\mathbf{x}_3 = (0, 0, 1) = \mathbf{U}_3, g_{\nu\mu} = \mathbf{x}_\nu \cdot \mathbf{x}_\mu = \delta_{\nu\mu}, \mathbf{x}_1 = \mathbf{x}_3 \times \mathbf{x}_2 = (-1, 0, 0) = -\mathbf{U}_1$, and $L_{\nu\mu} = \mathbf{x}_1 \cdot \mathbf{x}_{\nu\mu} = -f_{\nu\mu} = s_{\nu\mu}$. It follows that $\rho^* = 1/s_{22} = 1/L_{22}$ is the radius of curvature of the normal section in the direction of $\mathbf{U}_2$ of the incident (reflected) wavefront at $\mathbf{P}$.

In Fig. 3 we illustrate the vectors $\mathbf{U}_2$ and $\mathbf{U}_3$ which are tangent to the incident (reflected) wavefront at $\mathbf{Q}$. We also show the angle θ between these vectors and the principal directions 2 and 3 corresponding to the principal radii of curvature ρ_2 and ρ_3 of the wavefront at $\mathbf{Q}$. Since the wavefronts are parallel (i.e., orthogonal to the same 2-parameter family of rays), the principal radii of curvature of the incident (reflected) wavefront at $\mathbf{P}$ are $(\rho_2 + \sigma)$ and $(\rho_3 + \sigma)$. Furthermore, the principal directions are the same as those at $\mathbf{Q}$. Therefore, according to Euler's formula,

$$s_{22} = \frac{1}{\rho^*} = \frac{\cos^2\theta}{\rho_2 + \sigma} + \frac{\sin^2\theta}{\rho_3 + \sigma}. \quad \text{(B15)}$$

From (B1) and (B3) we see that the unit tangent vector to the edge is given by

$$\mathbf{t} = \mathbf{t}_1 = \sin\beta\mathbf{U}_3 + \cos\beta\mathbf{U}_1. \quad \text{(B16)}$$

In Lemma 2 (Sec. 2) we introduced the radius of curvature ρ_0 of the normal section of the incident (reflected) wavefront at $\mathbf{Q}$ in the direction of $\mathbf{U}_3$ and found that it was equal to ρ. Now we see that

$$\frac{1}{\rho} = \frac{1}{\rho_0} = \frac{\sin^2\theta}{\rho_2} + \frac{\cos^2\theta}{\rho_3}. \quad \text{(B17)}$$

If we eliminate θ from (B15) and (B17), we find that

$$(\rho_2 + \sigma)(\rho_3 + \sigma)s_{22} = (\rho_3 + \sigma)\cos^2\theta + (\rho_2 + \sigma)\sin^2\theta$$
$$= \sigma + \rho_2 + \rho_3 - (\rho_2\cos^2\theta + \rho_3\sin^2\theta)$$
$$= \sigma + \rho_2 + \rho_3 - (\rho_2\rho_3/\rho). \quad \text{(B18)}$$

Hence (B11) becomes

$$\hat{s}(\hat{\mathbf{P}}) - s(\mathbf{P}) = \tfrac{1}{2}\zeta^2 \sin^2\beta\,\frac{\sigma\rho_2\rho_3(\rho + \sigma)}{\rho(\rho_2 + \sigma)(\rho_3 + \sigma)} + O(\zeta^3). \quad \text{(B19)}$$

Since $z_0^{i,r}$ is regular at $\mathbf{P}$, we see from (2.8) and (2.9) that

$$z_0^{i,r}(\hat{\mathbf{P}}) = z_0^{i,r}(\mathbf{P}) + O(\zeta)$$
$$= z_0^{i,r}(\mathbf{Q})\left|\frac{\rho_2\rho_3}{(\rho_2 + \sigma)(\rho_3 + \sigma)}\right|^{\frac{1}{2}} + O(\zeta); \quad \text{(B20)}$$

hence

$$\frac{\epsilon^{i,r}}{2}\pi^{-\frac{1}{2}}e^{i\pi/4}\frac{z_0^{i,r}(\hat{\mathbf{P}})}{\sqrt{\hat{s}(\hat{\mathbf{P}}) - s(\hat{\mathbf{P}})}}$$
$$= \frac{e^{i\pi/4}}{(2\pi)^{\frac{1}{2}}}\frac{\epsilon^{i,r}z_0^{i,r}(\mathbf{Q})}{|\zeta|\sin\beta}\left|\frac{\rho}{\sigma(\rho + \sigma)}\right|^{\frac{1}{2}} + O(1). \quad \text{(B21)}$$

Since $\varphi = \mp\varphi_0 \pm \pi \mp \zeta$,

$$\cos\tfrac{1}{2}(\varphi \pm \varphi_0) = \tfrac{1}{2}\zeta + O(\zeta^3). \quad \text{(B22)}$$

Therefore (4.17) and (4.18) yield

$$\hat{z}_0^{i,r}(\hat{\mathbf{P}}) = \mp\frac{e^{i\pi/4}}{(2\pi)^{\frac{1}{2}}}\frac{z_0^{i,r}(\mathbf{Q})}{\zeta\sin\beta}\left|\frac{\rho}{\sigma(\rho + \sigma)}\right|^{\frac{1}{2}} + O(\zeta). \quad \text{(B23)}$$

Now $\epsilon^{i,r} = \pm\text{sgn}\,\zeta$. Therefore $\epsilon^{i,r}/|\zeta| = \pm 1/\zeta$. According to (4.6), $w_0^{i,r}(\hat{\mathbf{P}})$ is the sum of (B23) and

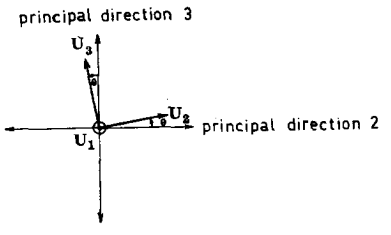


FIG. 3. Principal directions of the incident (reflected) wavefront at Q (proof of Lemma 3).

(B21); hence $w_0^{i,r}(\hat{\mathbf{P}})$ has a finite limit as $\zeta \to 0$, i.e., as $\hat{\mathbf{P}} \to \mathbf{P}$. It follows that $w_0^{i,r}(\hat{\mathbf{P}})$ is continuous at the shadow boundary of the incident (reflected) wave.

APPENDIX C: THE LAPLACIAN IN RAY COORDINATES

We shall transform the Laplacian to the coordinates

$$(y_1, y_2, y_3) = (\sigma, \varphi, \eta), \tag{C1}$$

which are defined by the transformation (5.4) and (5.5). According to Ref. 16, p. 47,

$$\Delta f = \frac{1}{(g)^{\frac{1}{2}}} \sum_{i=1}^{3} \sum_{j=1}^{3} \frac{\partial}{\partial y_i}\left((g)^{\frac{1}{2}} g^{ij} \frac{\partial f}{\partial y_j}\right), \tag{C2}$$

where

$$g_{ij} = \frac{\partial \mathbf{x}}{\partial y_i} \cdot \frac{\partial \mathbf{x}}{\partial y_j}, \tag{C3}$$

$$(g^{ij}) = (g_{ij})^{-1}, \tag{C4}$$

and

$$g = \det (g_{ij}). \tag{C5}$$

From (5.4),

$$\frac{\partial \mathbf{x}}{\partial y_1} = \mathbf{U}, \quad \frac{\partial \mathbf{x}}{\partial y_2} = \sigma \mathbf{U}_\varphi, \quad \frac{\partial \mathbf{x}}{\partial y_3} = \mathbf{t}_1 + \sigma \dot{\mathbf{U}}. \tag{C6}$$

Hence

$$\begin{aligned} &g_{11} = 1, \quad g_{12} = g_{21} = 0, \quad g_{13} = g_{31} = \cos\beta, \\ &g_{22} = \sigma^2 \mathbf{U}_\varphi^2, \quad g_{23} = g_{32} = \sigma \mathbf{U}_\varphi \cdot \mathbf{t}_1 + \sigma^2 \mathbf{U}_\varphi \cdot \dot{\mathbf{U}}, \\ &g_{33} = 1 + 2\sigma \mathbf{t}_1 \cdot \dot{\mathbf{U}} + \sigma^2 \dot{\mathbf{U}}^2. \end{aligned} \tag{C7}$$

By using (5.5) and (5.3), we may easily obtain $\dot{\mathbf{U}}$ and $\mathbf{U}_\varphi$. Then, by using (5.6), we find that

$$\begin{aligned} &g_{22} = \sigma^2 \sin^2\beta, \quad g_{23} = \sigma^2 \tilde{\kappa} \sin\beta \cos\beta \sin\varphi, \\ &g_{33} = 1 + \frac{2\sigma \sin^2\beta}{\rho} + \sigma^2\left[\left(\frac{\sin\beta}{\rho}\right)^2 + (\tilde{\kappa} \cos\beta \sin\varphi)^2\right]. \end{aligned} \tag{C8}$$

The determinant (C5) and inverse matrix (C4) may now be computed by standard methods. We find that

$$(g)^{\frac{1}{2}} = \sigma \sin^2\beta\left(1 + \frac{\sigma}{\rho}\right), \tag{C9}$$

and

$$\begin{aligned} gg^{11} &= \sigma^2 \sin^2\beta\left(1 + \frac{2\sigma \sin^2\beta}{\rho} + \frac{\sigma^2 \sin^2\beta}{\rho^2}\right), \\ gg^{12} &= \sigma^2 \tilde{\kappa} \sin\beta \cos^2\beta \sin\varphi, \\ gg^{13} &= -\sigma^2 \sin^2\beta \cos\beta, \\ gg^{22} &= \sin^2\beta\left(1 + \frac{\sigma}{\rho}\right)^2 + \sigma^2 \tilde{\kappa}^2 \cos^2\beta \sin^2\varphi, \\ gg^{23} &= -\sigma^2 \tilde{\kappa} \sin\beta \cos\beta \sin\varphi, \quad gg^{33} = \sigma^2 \sin^2\beta. \end{aligned} \tag{C10}$$

If we set

$$a^{ij} = (g)^{\frac{1}{2}} g^{ij}, \tag{C11}$$

then (C2) becomes

$$\Delta f = \frac{1}{(g)^{\frac{1}{2}}} \sum_{i=1}^{3} \sum_{j=1}^{3}\left[a^{ij} \frac{\partial^2 f}{\partial y_i \partial y_j} + \frac{\partial a^{ij}}{\partial y_i} \frac{\partial f}{\partial y_j}\right]. \tag{C12}$$

Since the a^{ij} are given by (C9)–(C11), (C12) provides a formula for the Laplacian in the coordinates (C1).

In Sec. 5 we require an expansion of (C12) for small σ. This is easily obtained from our results. We find that

$$\frac{1}{(g)^{\frac{1}{2}}} = \frac{1}{\sigma \sin^2\beta}\left[1 - \frac{\sigma}{\rho} + O(\sigma^2)\right]; \tag{C13}$$

and

$$\begin{aligned} a^{11} &= \sigma\left[1 + \frac{\sigma}{\rho}(2\sin^2\beta - 1)\right] + O(\sigma^3), \\ a^{12} &= \frac{\sigma \tilde{\kappa} \cos^2\beta \sin\varphi}{\sin\beta} + O(\sigma^2), \\ a^{13} &= -\sigma \cos\beta + O(\sigma^2), \\ a^{22} &= \frac{1}{\sigma}\left(1 + \frac{\sigma}{\rho}\right) + O(\sigma), \quad a^{23} = O(\sigma), \quad a^{33} = O(\sigma); \end{aligned} \tag{C14}$$

and

$$\begin{aligned} \sum_i \left(\frac{\partial a^{i1}}{\partial y_i}\right) &= 1 + \sigma\left[\frac{3\sin^2\beta - 1}{\rho} + \frac{\dot{\beta}}{\sin\beta}\right] + O(\sigma^2), \\ \sum_i \left(\frac{\partial a^{i2}}{\partial y_i}\right) &= -\tilde{\kappa} \sin\beta \sin\varphi + O(\sigma), \\ \sum_i \left(\frac{\partial a^{i3}}{\partial y_i}\right) &= -\cos\beta + O(\sigma). \end{aligned} \tag{C15}$$

By inserting (C13)–(C15) in (C12), we find that

$$\begin{aligned} \Delta f = &\frac{1}{\sin^2\beta}\left[1 - \frac{2\cos^2\beta}{\rho}\sigma + O(\sigma^2)\right] f_{\sigma\sigma} \\ &+ \left[\frac{2\tilde{\kappa} \cos^2\beta \sin\varphi}{\sin^3\beta} + O(\sigma)\right] f_{\sigma\varphi} \\ &+ \left[\frac{1}{\sigma^2 \sin^2\beta} + O(1)\right] f_{\varphi\varphi} \\ &+ \left[-\frac{2\cos\beta}{\sin^2\beta} + O(\sigma)\right] f_{\sigma\eta} \\ &+ [O(1)] f_{\varphi\eta} + [O(1)] f_{\eta\eta} \\ &+ \left[\frac{1}{\sigma \sin^2\beta} + \frac{\tilde{\kappa} \cos\varphi}{\sin^3\beta} - \frac{3\cos^2\beta}{\rho \sin^2\beta} + O(\sigma)\right] f_\sigma \\ &+ \left[-\frac{\tilde{\kappa} \sin\varphi}{\sigma \sin\beta} + O(1)\right] f_\varphi \\ &+ \left[-\frac{\cos\beta}{\sigma \sin^2\beta} + O(1)\right] f_\eta. \end{aligned} \tag{C16}$$

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 10, NUMBER 12 DECEMBER 1969

Analytic Properties of a Class of Nonlocal Interactions. II*

D. GUTKOWSKI AND A. SCALIA
Istituto di Fisica dell'Università, Catania, Italy and Sezione Siciliana dell'INFN, Gruppo di Catania, Italy

(Received 25 November 1968)

A definition is given for the function $S_l(k)$ in the complex l plane for a class of nonlocal interactions, called F, which has been considered in a previous paper [J. Math. Phys. **9**, 588 (1968)]. The definition is obtained by means of a new class of nonlocal interactions, called G, for which the definition of $S_l(k)$ in the l plane can be determined by a "dynamical interpolation." The analytic properties of potentials of class G are studied. Then a suitable approximation is defined which allows us to approximate any potential of class F by means of potentials of class G. Comparing the analytic properties of potentials of class G which sufficiently approximate any given potential of class F, it is shown that their total scattering amplitudes can be made as near as we please to each other and that there exist Regge trajectories which can be made as near as we please to each other. With the given definition, $S_l(k)$ turns out to be an analytic function in the complex l plane, but for a finite number of poles. Some general properties of the Regge trajectories are discussed and some examples are given.

1. INTRODUCTION

The Watson–Regge method, which has been such a useful tool in the quantum scattering theory, is far from having reached, for nonlocal potentials, the extension of results it has reached for local potentials.

In fact, among the several recent papers[1–9] devoted to the study of nonlocal potentials, there are only a few[2,3] of those in which the analytic properties of $S_l(k)$ in the complex l plane are studied.

The present paper deals with such a study, for a class of potentials indicated as F, for which in a preceding paper,[10] hereafter referred to as I, the analytic properties in the k plane were studied.

Let V be a potential of class F defined by

$$\langle \mathbf{p}| V |\mathbf{p}'\rangle = -4\pi \frac{\Gamma}{M} \sum_{l=0}^{L} \sum_{i=1}^{I_l} \sum_{m=-l}^{l} g_{il}(p) \times g_{il}(p') Y_l^m(\hat{\mathbf{p}}) Y_l^{m*}(\hat{\mathbf{p}}'). \quad (1.1)$$

If we want to study the analytic properties of $S_l(k)$ in the complex l plane, Eq. (1.1) alone does not allow us to give any sense to an interpolation of a dynamical type analogous to that made for local potentials.

This depends on the fact that while the Schrödinger equation for local potentials,

$$\Psi''_\lambda(x) + \left[k^2 - \frac{\lambda^2 - \frac{1}{4}}{x^2} - V(x)\right]\Psi_\lambda(x) = 0, \qquad \lambda = l + \tfrac{1}{2},$$

makes sense also for complex λ, the integral Schrödinger equation, which for potentials of class F is given by

$$\Psi_l(p) = 4\pi\Gamma \int_0^\infty dp' pp' \sum_{i=1}^{I_l} \frac{g_{il}(p) g_{il}(p')}{p^2 - k^2} \Psi_l(p'),$$

makes sense only for natural l, because its kernel is defined only for natural l.

One might think of getting an analytic interpolation starting from the knowledge of the partial amplitudes which are known for any natural l (in particular, $a_l = (1/2ik)[S_l(k) - 1] = 0$ identically in k for $l > L$). The problem of an analytic interpolation, which is constructed starting from the knowledge of partial amplitudes for any natural l, has been studied by several authors.[11,12]

If we attempt to solve the problem in this way, however, we are faced with great difficulties. The following difficulty is given as an example. Let us suppose that there exists an interpolation of $a_l(k)$ which is analytic in $\mathrm{Re}\, l \geq 0$ apart from a finite number of poles. By multiplying $a_l(k)$ by a convenient polynomial in l, and having fixed k arbitrarily, one obtains a function $f(l)$ analytic in $\mathrm{Re}\, l \geq 0$ and vanishing for any natural $l > L$. Since this function does not vanish identically, from the Carlson theorem[13] one has that $|f(l)| \not< \exp(hl)$ for any $h < \pi$ and sufficiently large $|l|$. It would in consequence be impossible, in the previous conditions for $a_l(k)$, to perform the Sommerfeld–Watson transform.[14]

We have chosen a different formulation of the problem. First, instead of a potential of class F defined

* This work has been supported in part by CSFN e SM.

[1] Y. Yamaguchi, Phys. Rev. **95**, 1628 (1954).
[2] J. T. Cushing, Nuovo Cimento **28**, 819 (1963).
[3] A. N. Mitra and J. D. Anand, Phys. Rev. **130**, 2117 (1963).
[4] G. C. Ghirardi and A. Rimini, J. Math. Phys. **5**, 722 (1964).
[5] S. Tani, Ann. Phys. **37**, 411 (1966).
[6] S. Tani, Ann. Phys. **37**, 451 (1966).
[7] F. Catara and M. Di Toro, J. Math. Phys. **6**, 1720 (1965).
[8] M. Bertero, G. Talenti, and G. A. Viano, Nuovo Cimento **46**, 337 (1966).
[9] M. Bertero, G. Talenti, and G. A. Viano, Commun. Math. Phys. **6**, 128 (1967).
[10] D. Gutkowski and A. Scalia, J. Math. Phys. **9**, 588 (1968).
[11] R. G. Newton, *The Complex j-Plane* (W. A. Benjamin, Inc., New York, 1964), Chap. XV.
[12] A. Gersten, Ann. Phys. **44**, 112 (1967).
[13] E. C. Titchmarsh, *The Theory of Functions* (Clarendon Press, Oxford, 1960), 2nd ed., p. 185.
[14] Ref. 11, p. 4.

by Eq. (1.1), let us consider the following potential:

$$\langle \mathbf{p}| V |\mathbf{p}'\rangle = -4\pi \frac{\Gamma}{M} \sum_{l=0}^{\infty} \sum_{i=1}^{I} \sum_{m=-l}^{l} c_i^2(l) \times g_i(p) g_i(p') Y_l^m(\hat{\mathbf{p}}) Y_l^{m*}(\hat{\mathbf{p}}'), \quad (1.2)$$

where the $g_i(p)$ $(1 \leq i \leq I)$ satisfy all the conditions imposed for potentials of class F to the $g_{il}(p)$ of Eq. (1.1), and the $c_i(l)$ are functions of l $(1 \leq i \leq I)$ defined for complex l apart from a finite number of exceptional points, having real values for any natural l and satisfying further conditions we shall specify later.

Let G be the class of these potentials. Repeating step by step for a potential of class G the proofs given in I for a potential of class F, one has to replace Eq. (3.11) of I with the following equation:

$$f_{i,j;l}(-k) = 4\pi c_i(l) c_j(l) \int_0^{\infty} dk_2 \frac{k_2^2 g_i(k_2) g_j(k_2)}{k^2 - k_2^2}, \qquad \text{Im } k > 0 ;$$

$f_{i,j;l}(-k)$ is defined, as in I, in the whole complex k plane, by analytic continuation. Equations (3.14) and (3.16) of I retain the same form.

So one obtains that $S_l(k)$, $a_l(k)$ for arbitrarily fixed k, apart from exceptional points in the complex k plane, are functions defined in the complex l plane, apart from exceptional points.

We impose on the $c_i(l)$ $(1 \leq i \leq I)$ as functions of l such conditions that it is possible to perform the Sommerfeld–Watson transform on the total scattering amplitude:

$$A(k, \cos\theta) = \frac{1}{2ik} \sum_{l=0}^{\infty} (2l+1)[S_l(k) - 1] P_l(\cos\theta) \quad (1.3)$$

for $0 \leq \theta \leq \pi$.

Consequently,[14] the expression

$$A(k, \cos\theta) = (2k)^{-1} \int_{-\infty}^{\infty} d\lambda' \frac{\lambda'[S(i\lambda', k) - 1]}{\cosh \pi\lambda'} \times P_{i\lambda' - \frac{1}{2}}(-\cos\theta) + i\pi k^{-1} \times \sum_n (\alpha_n + \tfrac{1}{2}) \beta_n P_{\alpha_n}(-\cos\theta)/\sin \pi\alpha_n \quad (1.4)$$

[where $S(\lambda, k) = S_l(k)$; $\alpha_n + \frac{1}{2}$ is a pole of $S(\lambda, k)$ in the λ plane, β_n is the residue relative to the pole $\alpha_n + \frac{1}{2}$], for arbitrarily fixed k, apart from exceptional points, is an analytic function in the $\cos\theta$ plane apart from a branch point.

As a second step, we have to find some connection between potentials of class F and those of class G. This connection is established by showing that for any potential V in F (in G) having total scattering amplitude $A(k, \cos\theta)$ and for any $\epsilon > 0$ there exists a (not unique) potential V_1 in G (in F) having total scattering amplitude $A_1(k, \cos\theta)$ such that, for physical values of k and $\cos\theta$,

$$|A(k, \cos\theta) - A_1(k, \cos\theta)| < \epsilon.$$

Thus we can find potentials $V_1 \in G$ whose total scattering amplitudes approximate, as we please, the total scattering amplitude of any potential $V \in F$ for physical values of k and $\cos\theta$.

Let us remark that the choice of V_1 can be made in such a way as to make

$$\|V_1 - V\| = \left\{ \iint |V_1(\mathbf{p}, \mathbf{p}') - V(\mathbf{p}, \mathbf{p}')|\, d\mathbf{p}, d\mathbf{p}' \right\}^{\frac{1}{2}}$$

as small as we please (see Sec. 4, Theorem 4.1).

What about objects which are defined for a potential in G but not for a potential in F? Let, e.g., such an object be a Regge trajectory. Let

$$\{V_{1\alpha\epsilon}\}_{\alpha \in I_\epsilon}$$

be the subset of G such that $\|V - V_{1\alpha\epsilon}\| < \epsilon$. We can assign by definition a Regge trajectory $l(k)$ to a potential $V \in F$, in such a way that $l(k)$, for sufficiently small ϵ, approximates as we please a trajectory $l_{1\alpha\epsilon}(k)$ relative to any $V_{1\alpha\epsilon}$, for $\alpha \in I_\epsilon$, iff for any $\delta > 0$ there exists $\epsilon_1 > 0$ such that

$$|l_{1\alpha'\epsilon}(k) - l_{1\alpha''\epsilon}(k)| < \delta, \text{ for any } \epsilon < \epsilon_1; \alpha', \alpha'' \in I_\epsilon.$$

We shall see that for any $V \in F$ the above condition can be satisfied for at least one trajectory.

In Sec. 2 of this paper the choice of conditions on the $c_i(l)$ is explained.

In Sec. 3 we derive some general properties and discuss some examples relative to potentials of class G.

In Sec. 4, the "approximation problem" is studied. Under this name we understand roughly the following:

(i) For any potential $V \in F$ and any $\epsilon > 0$ does there exist a potential $V_1 \in G$ such that the modulus of the difference between the total scattering amplitudes relative to V and V_1 for physical values of k and $\cos\theta$ is less than ϵ?

The answer is affirmative, Theorem 4.1; there are infinite potentials V_1 satisfying the above conditions.

(ii) Given two potentials V' and V'' of class G whose total scattering amplitudes, for physical values of k and $\cos\theta$, are sufficiently near to each other, are

"corresponding objects" that one can define starting from them, like the analytic continuation of the total scattering amplitudes in k and $\cos\theta$, the Regge trajectories and so on, to be found near to each other?

The answer is affirmative for m Regge trajectories (see Theorem 4.2) where $m = \min(m', m'')$, and m' (m'') is the number of trajectories relative to V' (V''). For the total scattering amplitudes we have been able to give an affirmative answer only if $m = m' = m''$.

In Appendix A it is shown that results of I can be extended to the case of interactions having both an attractive and a repulsive part.

In Appendix B we consider some mathematical features connected with our interactions and the Sommerfeld–Watson transform.

2. PROPERTIES OF THE $c_i(l)$

In this section we present the reasons for selecting the following hypotheses for the $c_i(l)$ of Eq. (1.2). The starting point is given by the conditions which the $S_l(k)$ have to fulfil as functions of l. We want the total scattering amplitude $A(k, z)$ to exist for physical values of k and z; that is, for real and positive k and for $-1 \leq z \leq 1$,

$$(2ik)^{-1} \sum_{l=0}^{\infty} (2l+1)[S_l(k) - 1]P_l(z)$$

must exist and be finite.

If $A(k, z)$ is a continuous function for physical values of its arguments, the convergence of the above series is equivalent by the Riesz–Fischer[15] theorem to the condition that, for real positive k,

$$(4k^2)^{-1} \sum_{l=0}^{\infty} (2l+1)\,|S_l(k) - 1|^2 < \infty. \qquad (2.1)$$

We therefore impose, as the first condition, that Eq. (2.1) hold true.

Moreover, we want to perform the Sommerfeld–Watson transform. Therefore we impose the asymptotic properties of $S(\lambda, k)$ so that the integral appearing on the rhs of Eq. (1.4) exists. To achieve this result it will be enough to suppose (see Appendix B) that there are a natural number n and a real positive number λ_0, such that, for $|\lambda| > \lambda_0$ ($\operatorname{Re}\lambda \geq 0$) and for any complex k, save at most a finite number of values, the equation

$$|S(\lambda, k) - 1| < |\lambda|^n \qquad (2.2)$$

holds true.

We could impose more general asymptotic conditions than that given by Eq. (2.2) (see Appendix B). One can, however, see that in so doing we shall not attain more general results.[16] We assume therefore, as the second condition, that Eq. (2.2) holds true.

We impose, as a third condition, that $S(\lambda, k)$ is analytic apart from a finite number of poles[17] in the complex half λ plane ($\operatorname{Re}\lambda \geq 0$), for any complex k, apart from a finite number of values.

It is well known that under these hypotheses the right-hand sides of (1.3) and (1.4) are equal to each other. We recall, however, that this equality does not necessarily hold if the series (1.3) diverges. In this case it could indeed happen (see Appendix B) that the rhs of Eq. (1.4) exists as a finite function of k and z. For that reason we have imposed the first condition quite independently from the second and the third ones.

We still want, for Hermitian Hamiltonians with local interactions, the following fourth condition to hold[18]:

$$[S(\lambda, k)]^* = S^{-1}(\lambda^*, k^*). \qquad (2.3)$$

Taking into account that $S(\lambda, k)$ has to satisfy the four conditions we have postulated, we can analyze the conditions to impose on $c_i(l)$.

From (1.1) and (3.13), and (3.15) of I, one gets

$$\begin{aligned} S_i(k) - 1 = \Bigg\{ & 4\pi^2 i\Gamma k \sum_{m,j=1}^{I} c_m(l)c_j(l)g_m(k)g_j(k)\, d\binom{m}{j}_l \Bigg\} \\ & \times \Bigg[\sum_{m,j=1}^{I} \Bigg(\frac{1}{I} \Bigg\{ \delta_{mj} + 2\pi^2 i\Gamma c_m(l)c_j(l) \\ & \times \Bigg[2\sum_{\alpha=1}^{r} \operatorname*{Res}_{k_2=a_{m\alpha}} + 2\sum_{\beta=1}^{s} \operatorname*{Res}_{k_2=a_{j\beta}} - I g_m(k) \\ & \times k g_j(k) \Bigg] \Bigg\} \Bigg) d\binom{m}{j}_l \Bigg]^{-1}. \end{aligned} \qquad (2.4)$$

Let us now suppose that $c_i(l)$ satisfy the following conditions:

(a) They are analytic functions in the right half plane l ($\operatorname{Re} l \geq 0$), apart from isolated not essential singularities;

(b) they do not have an essential singularity at infinity;

(c) they are real for real l.

The possibility of essential singularities in the finite right half plane or at infinity will be considered later.

[15] F. Smithies, *Integral Equations* (Cambridge University Press, Cambridge, 1962), p. 59.

[16] This depends on the generality of our conditions. We could not state that "... we shall not attain more general results" if the possibility of an essential singularity for $S(\lambda, k)$ in $\lambda = \infty$ were allowed. But then, in order for it to be possible to perform the Sommerfeld–Watson transform, we had to impose further conditions.

[17] We make this hypothesis in order to warrant the unicity of the $S(\lambda, k)$ taking given values for physical values of λ. The argument allowing us to prove the unicity is that of Ref. 11, p. 116.

[18] V. De Alfaro and T. Regge, *Potential Scattering* (North-Holland Publ. Co., Amsterdam, 1965), p. 39.

From hypothesis (b) it follows that

$$\lim_{|l|\to\infty} c_i(l), \quad 1 \leq i \leq I,$$

must exist finite or not.

Then the following possibilities may happen:

(i) For any i, $1 \leq i \leq I$, $\lim_{|l|\to\infty} c_i(l) = 0$;

(ii) there exist some $c_i(l)$ for which $\lim_{|l|\to\infty} c_i(l)$ is not vanishing.

In case (i), since the numerator of (2.4) converges to zero, while the denominator converges to one, as l goes to infinity, we can have consistency with Eqs. (2.1) and (2.2). In case (ii), we cannot exclude, under the hypotheses made in I for the $g_i(p)$, that

$$\lim_{|l|\to\infty} S_i(k) - 1 \neq 0.$$

This would lead to a contradiction with Eq. (2.1). Therefore as the fourth condition on the $c_i(l)$ we assume the following:

(d) for any i, $1 \leq i \leq I$, $\lim_{|l|\to\infty} c_i(l) = 0$.

It will be seen that condition (d) together with (a) and (b) implies Eq. (2.1).

From hypothesis (a) and Eq. (2.4), it is easy to deduce the third condition on $S_i(k)$. The singularities of the latter would be in general the zeros of the denominator of the rhs of Eq. (2.4) (which depend on k), because, apart from exceptional values of k, the poles of the numerator due to poles of the $c_i(l)$ also appear at the denominator as poles of the same order.

From hypothesis (c), by applying the Riemann–Schwarz reflection principle,[19] one gets Eq. (2.3).

Let us now consider the case where, keeping hypotheses (a) and (c) and suppressing hypothesis (d), hypothesis (b) is substituted by the following:

(b′) There exist some $c_i(l)$ having an essential singularity at infinity.

If $c_m(l)$ is one of these, from the Picard theorem[20] for any complex a_m (save at most one value) and for any neighborhood of the infinity ($|l| > H$, H arbitrary real positive number), it is possible to satisfy the equation

$$c_m(l) = a_m.$$

From Eq. (2.4) it follows that $S_i(k)$, for $|l| > H$, takes values whose modulus of the difference cannot be made arbitrarily small for a convenient choice of H. Since from hypotheses (a), (b′), (c), and Eq. (2.4), $S_i(k)$ can have in the l plane only isolated singularities, for the fact that

$$\lim_{|l|\to\infty} S_i(k)$$

does not exist, one deduces that $l = \infty$ is for $S_i(k)$ an essential singularity.

Therefore $S_i(k)$ is not bounded in a neighborhood of the infinity, and, in general, conditions for performing the Sommerfeld–Watson transform are not satisfied.

Obviously, under further hypotheses for $c_i(l)$ one could manage that $S_i(k)$ would be bounded on every contour of a convenient class, in such a way that the Sommerfeld–Watson transform could be performed, but, at the level of generality we propose, we shall not deal with this case any further.

Let us now consider the case where we substitute for (a):

(a′) The $c_i(l)$ are analytic functions in the right half l plane ($\mathrm{Re}\, l \geq 0$), apart from isolated singularities, one of which at least is an essential one; we keep conditions (b), (c), and (d).

Repeating the reasoning of the preceding case, one can deduce that also $S_i(k)$ has, in the finite plane, an essential singularity, against the third condition.

As a result of the analysis we have made we assume for the $c_i(l)$ conditions (a), (b), (c), and (d). The class of potentials described by Eq. (1.2) with these conditions will be named G.

Let us remark that from conditions (a), (b), and (d) one can deduce[21] the possibility of expressing the $c_i(l)$ ($1 \leq i \leq I$) in partial fractions, and precisely in a form like

$$c_i(l) = \sum_{j=1}^{n_i} \left[\frac{A^{ji}_{-\beta_{ji}}}{(l - l_{ji})^{\beta_{ji}}} + \cdots + \frac{A^{ji}_{-1}}{l - l_{ji}} \right],$$

where n_i, β_{ji} ($1 \leq j \leq n_i$) are positive integers, $A^{ji}_{-\beta_{ji}}$, l_{ji} are complex numbers.

By substituting this expression into (2.4) it is easy to see that Eq. (2.1) is satisfied.

3. REGGE TRAJECTORIES—GENERAL PROPERTIES AND EXAMPLES

In this section we deal with some general properties of the Regge trajectories for potentials of class G. As is well known, Regge trajectories are given by the functions $\lambda(k)$, for real positive or positive imaginary k such that $D[\lambda(k), -k] = 0$, where $D(\lambda, -k) = D_{\lambda-\frac{1}{2}}(-k)$ ($\lambda = l + \frac{1}{2}$). We shall deal also with some examples with specified potentials of class G.

[19] Ref. 13, p. 155.
[20] Ref. 13, p. 183.
[21] A. I. Markuschevich, *The Theory of Analytic Functions* (Hindustan Publ. Co., Delhi 6, 1963), p. 239.

A. Zeros of $D_l(-k)$

From Eq. (3.14) of I, one gets

$$D(\lambda, -k) = \det |\delta_{ij} + \Gamma c_i(l)c_j(l)f_{ij}(-k)| \quad (i,j = 1, \cdots, I). \quad (3.1)$$

Then $D(\lambda, -k)$ can be expressed as the sum of determinants of the matrices obtained from the matrices $|\delta_{ij}|$ and $|\Gamma c_i(l)c_j(l)f_{ij}(-k)|$ by combining in all the possible ways the columns of the one with the columns of the other; after some easy calculations one gets

$$D(\lambda, -k) = 1 + \sum_{i=1}^{I} c_i^2(l)F_{00\cdots i\cdots 00}(-k) + \sum_{i=1}^{I}\sum_{j=i+1}^{I} c_i^2(l)c_j^2(l)F_{00\cdots i\cdots j\cdots 00}(-k) + \cdots + c_1^2(l)c_2^2(l)\cdots c_I^2(l)F_{1\,2\ldots I}(-k), \quad (3.2)$$

where the functions $F_{\ldots}(-k)$ are determinants of convenient minors of order $\leq I$, having all the elements in $\Gamma f_{ij}(-k)$, extracted from $I \times I$ matrices having a number of columns (from 0 to $I-1$) in δ_{ij} and the other in $\Gamma c_i(l)c_j(l)f_{ij}(-k)$.

Let us recall (see Sec. 2) that from the conditions (a), (b), and (d) on the $c_i(l)$, it is possible to express these functions in the form

$$c_i(l) = \sum_{r=1}^{n_i} \left[\frac{A_{-\beta ri}^{ri}}{(l - l_{ri})^{\beta_{ri}}} + \cdots + \frac{A_{-1}^{ri}}{l - l_{ri}} \right]. \quad (3.3)$$

From (3.2) and (3.3) one can write

$$D_l(-k) = \frac{P_k^m(l)}{Q_k^m(l)}, \quad (3.4)$$

with $P_k^m(l)$ and $Q_k^m(l)$ polynomials of degree m in l, with complex coefficients depending on k. The integer m is given by

$$m = 2\sum_{i=1}^{I}\sum_{r=1}^{n_i} \beta_{ri}. \quad (3.5)$$

Zeros in l of $P_k^m(l)$ are functions of k. The dependence of $Q_k^m(l)$ on k can be expressed by means of a multiplicative factor, viz.,

$$Q_k^m(l) = Q^m(l) \cdot L(k) \quad (3.6)$$

with $Q^m(l)$ polynomial of degree m in l and $L(k)$ function of the only k. The statement that, in general, $P_k^m(l)$ and $Q_k^m(l)$ have no common zeros in l follows from the fact that zeros of $Q_k^m(l)$ in l do not depend on k. Zeros of $D_l(-k)$ in l, which are zeros of $P_k^m(l)$ in l, are thus functions of k, their number m is the number of trajectories $\lambda = \lambda(k)$ in the complex λ plane, provided they are not zeros of the same or greatest order of $D_l(k)$.

B. Hermiticity's Consequences

From the Hermiticity relationship

$$D_l^*(-k^*) = D_l^*(k) \quad (3.7)$$

that we proved in Sec. 2, one sees that for purely imaginary k ($k = ik'$, k' Re), if $D_{l_1}(-ik') = 0$, then $D_{l_1}^*(-ik') = 0$. That is, for purely imaginary k, trajectories are either real or complex conjugate.

C. Allowed Regions in the l Plane

Let us remark: As will be demonstrated in the examples, for potentials of class G there are no limitations in the region of the l plane where poles can be found, contrary to what happens for local potentials.[22]

Besides, it is not necessary, as it is for local potentials,[23] that if a trajectory leaves the real axis at a point $\lambda_0 > \frac{1}{2}$ (or $l_0 > 0$), it does so towards the right.

In fact the vanishing of the centrifugal potential term in the Schrödinger equation cannot have any special significance for potentials of class G, since they explicitly depend on l.

D. Resonance Conditions in l Plane

For potentials of class G, as well as for local potentials, it is possible, from the behavior of the trajectories in the l plane, to get information on poles of $S_l(k)$ in the k plane for physical l.

In fact, if for purely imaginary k the trajectory passes through physical values of l, for those values there are bound states whose energies correspond to values of k for which $l(k)$ has physical values.

From the behavior of the trajectories for real k, one can deduce the existence of poles in the k plane for physical l, having a small imaginary part with respect to the real part.

Let the function $l = l(k)$ indeed be analytic in a region Q of the k plane, including on the positive real semiaxis, and let $dl(k)/dk \neq 0$ in Q. Then $l(k)$ gives a conformal mapping of Q into a region Q' of the l plane. Let $A = k_1 - \Delta k$ and $B = k_1 + \Delta k$ (see Fig. 1) be two real points in the region Q, and A' and B' their images in Q'; let $l_4 = l(k_4)$ be an integer belonging to Q', included between Re B' and Re A'. Let Δk be small enough in order that l_4 be the only integer satisfying the above condition and such that $\Delta k \ll k_1$. If for such a Δk it is possible to satisfy the condition

$$\text{Re } A' - \text{Re } B' \gg \text{Im } l(k_1), \quad (3.8)$$

then, for a small rotation of the point B' around A', the segment of trajectory $A'B'$ will meet the real axis at the point l_4. Then, owing to the conformal mapping,

[22] Ref. 18, p. 79.
[23] Ref. 11, Chap. IX.

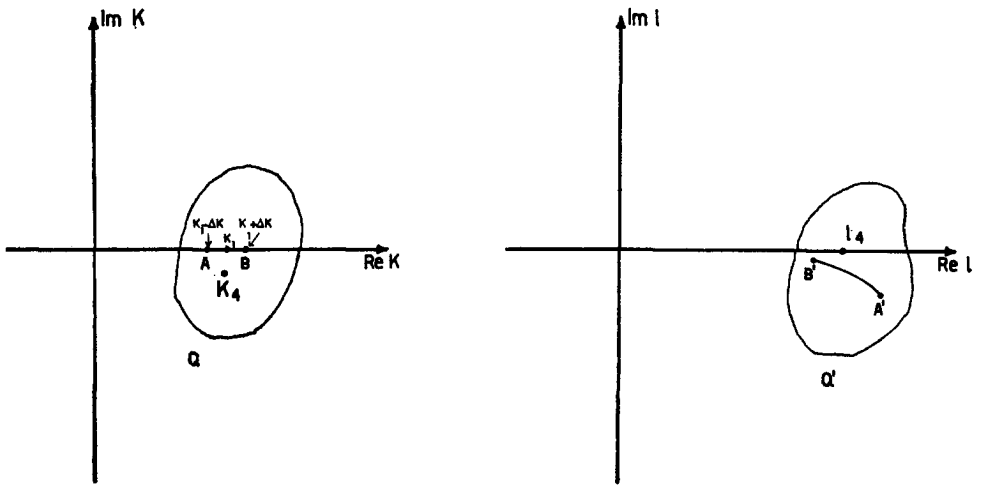


FIG. 1. Mapping of the region Q into the region Q' by means of the analytic functions $l(k)$. A', B', l_4 are, respectively, the corresponding points of A, B, k_4.

k_4 lies on the segment AB rounded by the same small angle.

From (3.8) one gets

$$\left| -2\left[\frac{d}{dk}\,\mathrm{Re}\, l(k)\right]_{k=k_1} \Delta k \right| \gg |I_m l(k_1)|. \qquad (3.9)$$

In order to produce resonances in l in the examples, we have chosen parameters in such a way as to satisfy Eq. (3.9).

E. Examples

In the two following examples we study two potentials satisfying

$$\langle \mathbf{p}|\, V\, |\mathbf{p}'\rangle = -4\pi \frac{\Gamma}{M} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} c^2(l) \times g(p)g(p')Y_l^m(\hat{\mathbf{p}})Y_l^{m*}(\hat{\mathbf{p}}'), \qquad (3.10)$$

where $c(l)$ and $g(p)$ are given by

$$c(l) = \frac{1}{1+bl}, \quad g(p) = \frac{1}{p^2+\beta^2}; \qquad (3.11)$$

β and b are real parameters. The strength Γ is positive (attractive potential) in the first example and negative (repulsive potential) in the second one.

First example: We have imposed a bound state of energy $-\alpha^2/M$ at $l = 0$. With this choice we obtain for the strength

$$\Gamma = \beta(\alpha+\beta)^2/\pi^2, \qquad (3.12)$$

where α and β are the positive square roots of α^2 and β^2.

From (3.1), (3.10), (3.11), and (3.12), we get

$$l_1(k) = \frac{1}{b}\left(\frac{i(\alpha+\beta)}{k+i\beta} - 1\right),$$

$$l_2(k) = \frac{1}{b}\left(\frac{-i(\alpha+\beta)}{k+i\beta} - 1\right), \quad \left(k \neq -i\beta, l \neq -\frac{1}{b}\right). \qquad (3.13)$$

For positive (or negative) b we get trajectories as in Fig. 2(a) and 2(b).

It is possible to verify statements of 3A, 3B, 3C, and 3D. In what concerns 3B, let us remark that for purely

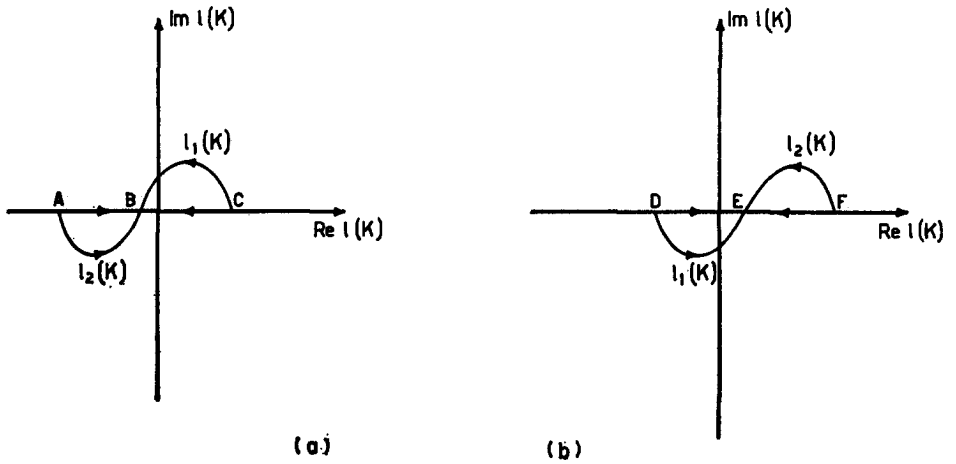


FIG. 2. Example of the behavior of the Regge trajectories for an attractive potential of class G, with different choice of the conditions on parameters.

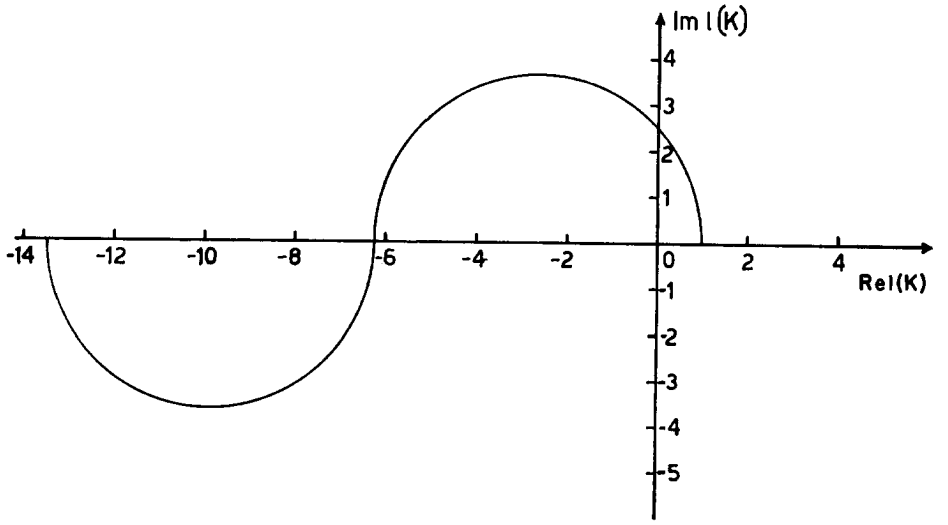


FIG. 3. Regge trajectories for an attractive potential of class G, with specified parameters. There are bound states for $l = 0, 1$.

imaginary k, $l_1(k)$ and $l_2(k)$, are real. In what concerns 3D, only conditions for bound states are verified.

In Fig. 3 trajectories $l_1(k)$ and $l_2(k)$ are drawn for the choice of parameters we have made. There are bound states for $l = 0, 1$. The behavior of the partial cross section $\sigma_0(k)$ given in Fig. 4 confirms, as can be seen from the trajectory, that there are no resonances.

Second Example: We have put $\Gamma = -\Gamma_0$ into (3.10). One gets

$$l_3(k) = -\frac{1}{b} + \frac{1}{b}\frac{\pi(\Gamma_0)^{\frac{1}{2}}(k - i\beta)}{(k^2 + \beta^2)(\beta)^{\frac{1}{2}}},$$

$$l_4(k) = -\frac{1}{b} - \frac{1}{b}\frac{\pi(\Gamma_0)^{\frac{1}{2}}(k - i\beta)}{(k^2 + \beta^2)(\beta)^{\frac{1}{2}}}, \qquad (3.14)$$

$$k \neq -i\beta, \quad l \neq -\frac{1}{b}.$$

For positive (negative) b we get trajectories as in Fig. 5(a) and 5(b).

In Fig. 6, $l_3(k)$ is given for a certain choice of Γ_0, b. From its form one foresees the existence of resonances. This fact is confirmed by Fig. 7, where $\sigma_l(k)$ and $\delta_l(k)$ are given for various l.

4. THE "APPROXIMATION PROBLEM"

Let us now consider the "approximation problem" we have sketched in the Introduction. We introduce the following notations. Let V_1 and V_2 be two potentials of class G defined by equations

$$\langle \mathbf{p}| V_1 |\mathbf{p}'\rangle = -4\pi\frac{\Gamma}{M}\sum_{l=0}^{\infty}\sum_{i=1}^{I}\sum_{m=-l}^{l} c_{1i}^2(l)g_i(p)g_i(p')Y_l^m(\hat{\mathbf{p}})Y_l^{m*}(\hat{\mathbf{p}}'),$$

$$\langle \mathbf{p}| V_2 |\mathbf{p}'\rangle = -4\pi\frac{\Gamma}{M}\sum_{l=0}^{\infty}\sum_{i=1}^{I}\sum_{m=-l}^{l} c_{2i}^2(l)g_i(p)g_i(p')Y_l^m(\hat{\mathbf{p}})Y_l^{m*}(\hat{\mathbf{p}}'). \qquad (4.1)$$

Let us suppose that for $l = 0, 1, \cdots, L$ the following condition holds:

$$c_{1i}(l) = c_{2i}(l). \qquad (4.2)$$

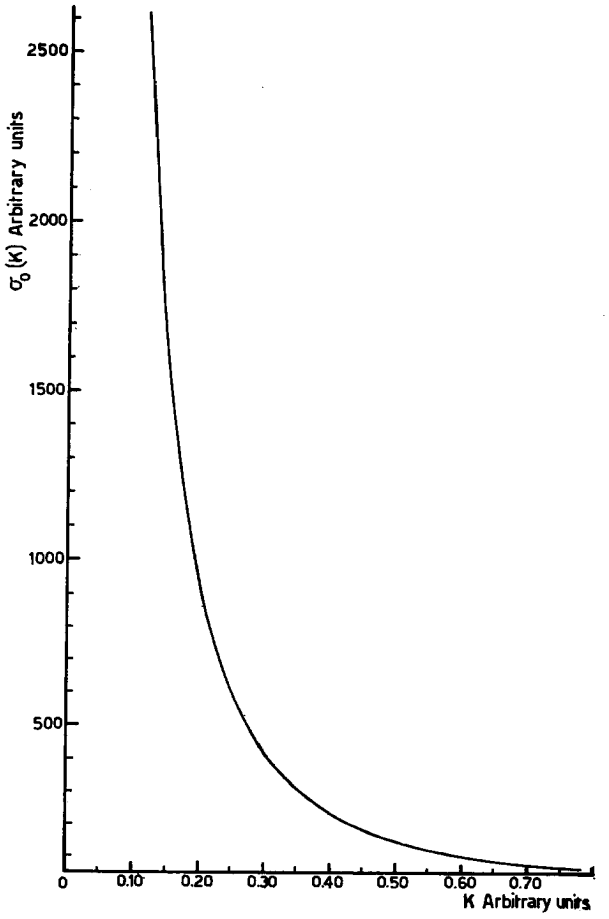


FIG. 4. Partial cross section $\sigma_l(k)$ for $l = 0$ vs the wavenumber k for attractive potential of class G.

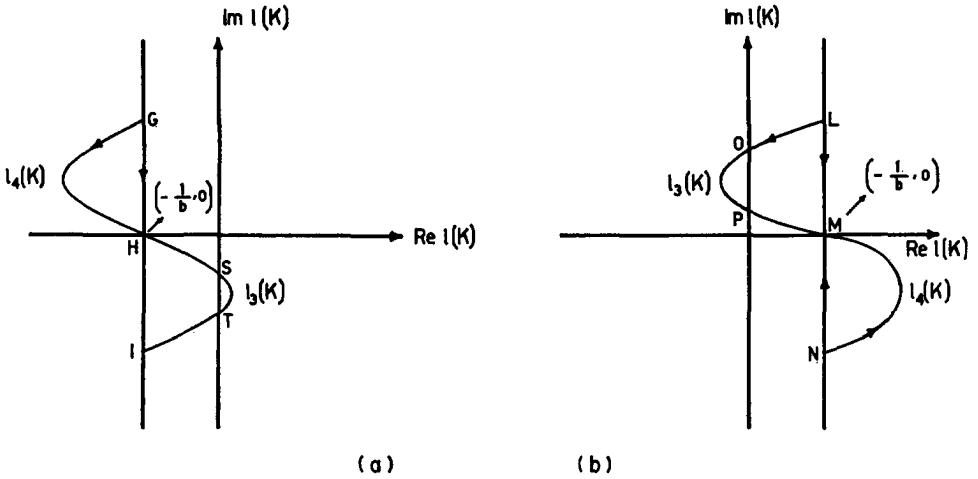


FIG. 5. Example of the behavior of the Regge trajectories for a repulsive potential of class G, with different choice of the conditions on parameters.

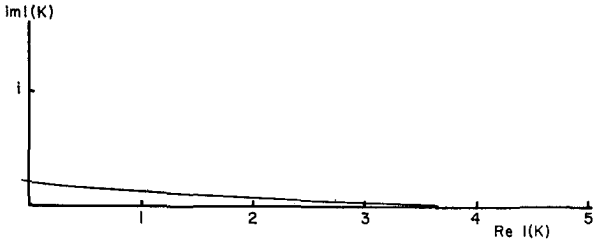


FIG. 6. Part of a Regge trajectory for a repulsive potential of class G, with specified parameters.

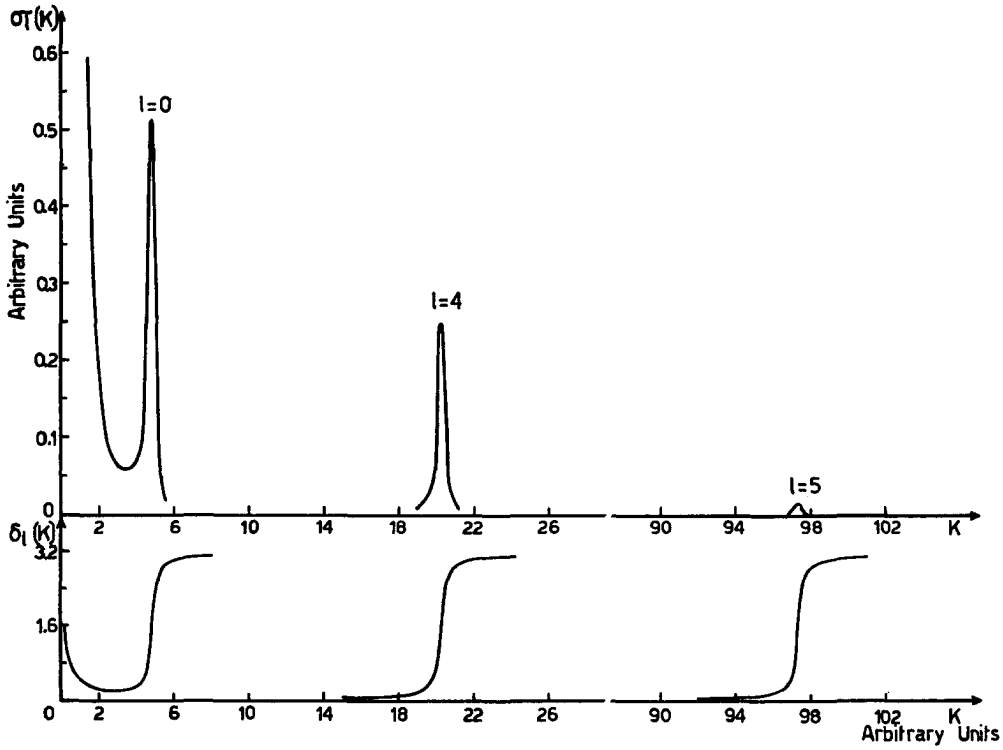


FIG. 7. Partial phase shifts in radians and partial cross sections vs the wavenumber k, for various l, relative to the same potential referred to in Fig. 6.

L is a natural number, eventually depending on one or more parameters.

All the objects related to V_1 and V_2 will be marked by a lower index 1 and 2, respectively, which will be at the first place in the case of several indices.

$V_{1L} = V_{2L}$ is the potential of class F obtained by cutting the sum over l at L in (4.1). All the objects related to it will be marked by a second lower index L.

The total scattering amplitudes will be named $A\ldots$, the Regge trajectories and their analytic continuation will be named $\lambda\ldots(k)$, their number will be named $m.$, where the dots indicate possible indices.

T is a natural number, $\{k_t, T\}$ is a set of T points k_t $(1 \leq t \leq T)$ of the k plane, $\{\overline{k_t, \rho_t, T}\}$ is a set of points which is the complement in the k plane of the union of T open circles having centers at the points k_t and radius ρ_t. Now let us prove the following existence theorem:

Theorem 4.1: For any potential V of class F defined by (1.1) and for any $\epsilon > 0$, there exists a potential V_1 of class G defined by (4.1) and such that

$$V_{1L} = V, \tag{4.3}$$

and for any physical k and z,

$$|A_1(k, z) - A_{1L}(k, z)| < \epsilon. \tag{4.4}$$

Proof: We shall first construct a potential $\bar{V}_1$ of class G satisfying (4.3) [not necessarily (4.4)].

Let us order the linearly independent $g_{ii}(p)$ appearing in the definition of V and let I be their number. From our assumptions, $\bar{c}_{1i}(l)$ can be written as follows:

$$\bar{c}_{1i}(l) = \frac{\bar{p}_{1i}(l)}{\bar{q}_{1i}(l)}, \tag{4.5}$$

where $\bar{p}_{1i}(l)$, $\bar{q}_{1i}(l)$ are polynomials in l.

Taking into account Eq. (4.5), we can set a linear system in the generally complex[24] coefficients of polynomials $\bar{p}_{1i}(l)$, $\bar{q}_{1i}(l)$ by equating the corresponding factors in the equations defining V and $\bar{V}_{1L}$ for $l = 0, 1, \cdots, L$.

If the degrees of $\bar{p}_{1i}(l)$, $\bar{q}_{1i}(l)$ are $\bar{r}_{1i}$, $\bar{s}_{1i}$, respectively, for the system to have a solution we must choose

$$\bar{r}_{1i} + \bar{s}_{1i} \geq L \tag{4.6}$$

for $i = 1, 2, \cdots, L$. If the equality holds in (4.6), the solution is unique for given $\bar{r}_{1i}$ and $\bar{s}_{1i}$, apart from a multiplicative constant.

We must have

$$\bar{s}_{1i} > \bar{r}_{1i} \tag{4.7}$$

by conditions (a), (b), and (d) of Sec. 2.

As the total scattering amplitude

$$\bar{A}_1(k, z) = \frac{1}{2ik} \sum_{l=0}^{\infty} (2l + 1)[S_{1i}(k) - 1]P_l(z)$$

relative to $\bar{V}_1$ converges for physical values of k and z, for any $\epsilon/2$ there exists $\bar{L}(\epsilon/2, k)$ such that, for physical k and z,

$$|\bar{A}_1(k, z) - \bar{A}_{1\bar{L}}(k, z)| < \epsilon/2.$$

We can make $\bar{L}$ independent on k since $\bar{A}_1(k, z)$ and $\bar{A}_{1\bar{L}}(k, z)$ are bounded and continuous functions for $-1 \leq z \leq 1$, $0 \leq k \leq \infty$ and tend to zero as $k \to \infty$ (see Ref. 25).

Suppose $\bar{L} > L$ (otherwise the theorem would be proved). Let us put

$$c_{1i}(l) = \bar{c}_{1i}(l) \frac{1}{1 + b\{l(l-1)\cdots(l-L)\}}$$

with $b > 0$. We get

$$c_{1i}(l) = \bar{c}_{1i}(l) \quad \text{for natural} \quad l \leq L$$

and

$$|c_{1i}(l)| < |\bar{c}_{1i}(l)| \quad \text{for natural} \quad l > L.$$

By choosing b conveniently, we can satisfy the condition

$$\left| \frac{1}{2ik} \sum_{l=L+1}^{\bar{L}} (2l + 1)[S_{1i}(k) - 1]P_l(z) \right| < \frac{\epsilon}{2},$$

so that potential V_1, corresponding to the $c_{1i}(l)$, has a total scattering amplitude satisfying both (4.3) and (4.4). Q.E.D.

The main result concerning the "approximation problem" for the Regge trajectories is given by the following theorem:

Theorem 4.2: If for physical values of k and z the following conditions hold:

(i) $|A_1(k, z) - A_{1L}(k, z)| < \epsilon$,
$|A_1(k, z) - A_2(k, z)| < \epsilon$, for any $\epsilon > \epsilon_1$; (4.8)

(ii) $m_1 \leq m_2$,

[24] Hypothesis (c) of Sec. 2 does not imply the reality of coefficients of polynomials $\bar{p}_{1i}(l)$, $\bar{q}_{1i}(l)$ as the following example shows:

$$\bar{p}_{1m}(l) = (1 + i) + [(1 + \sqrt{2})(1 + i) - 1/(1 + \sqrt{2})]l,$$
$$\bar{q}_{1m}(l) = (1 + i) + [(1 + i)\sqrt{8} - 1/(1 + \sqrt{2})]l + [(1 + i) - \sqrt{8}/(1 + \sqrt{2}) + 1]l^2.$$

[25] The possible existence of a bound state for $k^2 = k_1^2 > 0$ is not a difficulty since $\lim_{k \to k_1} S_l(k) = 1$ from the Hermiticity relationship (3.7).

then there exists a set $\{k_t, T\}$ such that for any choice of $\rho_1, \rho_2, \cdots, \rho_T$, for any $k \in \{\overline{k_t, \rho_t, T}\}$, for any $\Delta > 0$, it is possible to satisfy the condition

$$|\lambda_{1i}(k) - \lambda_{2i}(k)| < \Delta, \tag{4.9}$$

for $i = 1, 2, \cdots, m_1$, provided we choose a small enough ϵ_1.

Proof: The existence of $A_1(k, z)$, $A_2(k, z)$ satisfying (i) for arbitrary $\epsilon_1 > 0$ is insured by Theorem 4.1. From our assumptions, the $c_{1i}(l)$ and $c_{2i}(l)$ can be written as follows:

$$c_{1i}(l) = \frac{p_{1i}(l)}{q_{1i}(l)},$$

$$c_{2i}(l) = \delta \frac{l(l-1)(l-2)\cdots(l-L)p_{2i}(l)}{q_{2i}(l)} + \frac{p_{1i}(l)}{q_{1i}(l)}, \tag{4.10}$$

where $p_{1i}(l)$, $q_{1i}(l)$, $p_{2i}(l)$, $q_{2i}(l)$ are polynomials in l, the degree of $q_{1i}(l)$ is greater than the degree of $p_{1i}(l)$, the degree of $q_{2i}(l)$ is greater than $(L + 1)$ + degree of $p_{2i}(l)$, and δ is a real parameter we have introduced for convenience.

Taking into account Eqs. (3.2), (3.4), and (4.1), we get the equations of the Regge trajectories for potentials V_1 and V_2 by solving respectively the following algebraic equations in l:

$$p_{1k}^{m_1}(l) = 0,$$

$$p_{2k}^{m_2}(l) = \prod_{i=1}^{I} (q_{2i}(l))^2 p_{1k}^{m_1}(l) + \delta \tilde{p}_{2k}^{m_2}(l) = 0, \tag{4.11}$$

where $\tilde{p}_{2k}^{m_2}(l)$ is a polynomial of degree m_2 in l, whose coefficients are of degree from zero to $2I - 1$ in δ. For any k, apart from T exceptional points k_t, given by the zeros of $L(k)$ of Eq. (3.6), and for any $\Delta > 0$, there exists $\eta(k, \Delta) > 0$, such that for $\delta < \eta(k, \Delta)$,

$$|\lambda_{1i}(k) - \lambda_{2i}(k)| < \Delta \quad (i = 1, 2, \cdots, m_1).$$

The above equation follows from the properties of roots of polynomials satisfying an equation like (4.11).

We can make η independent on k by choosing arbitrarily T real positive numbers ρ_t $(t = 1, 2, \cdots, T)$ and letting k take values in $\{\overline{k_t, \rho_t, T}\}$. This follows from the fact that, by hypothesis (c) of $g_{il}(p)$ of I, the $F...(-k)$ of Eq. (3.2) tend to zero as $|k| \to \infty$, and then the coefficients of polynomials $p_{1k}^{m_1}(l)$, $p_{2k}^{m_2}(l)$ divided by $L(k)$ are continuous and bounded functions of k in the closed set $\{\overline{k_t, \rho_t, T}\}$.

In order to achieve the proof we want to show that, by choosing ϵ_1 conveniently in (i), we can make δ as small as we please. As for any $\epsilon > 0$ satisfying the first Eq. (4.8), there exists $\delta > 0$ satisfying the second one, it will be enough to prove that, for any fixed $\delta > 0$, one can choose ϵ in such a way that if the first Eq. (4.8) holds, the second one does not hold.

Indeed, if $A_1(k, z) = A_2(k, z)$ identically in z for physical values of z, it will be, for any natural l, $S_{1l} = S_{2l}$.

But, according to analytic and asymptotic properties in l of S_{1l} and S_{2l}, this would imply, by the Carlson theorem, $S_{1l} = S_{2l}$ identically for complex l. But this identity for $\delta > 0$ is absurd. Then there exist physical z and k, $\bar{\epsilon} > 0$, such that $|A_1(k, z) - A_2(k, z)| > \bar{\epsilon}$. If the first Eq. (4.8) holds for $\epsilon < \bar{\epsilon}$, it is thus proved that the second one does not hold.

The analysis of the construction of $\bar{V}_1$ in Theorem 4.1 shows that the minimum number $\bar{m}_{1\,\min}$ of trajectories, relative to $\bar{V}_1$ as $\bar{r}_{1i}$ and $\bar{s}_{1i}$ vary, satisfying (4.6) have the upper bound $I(L + 2)$. In fact, $I(L + 2)$ is greater or equal to the number of trajectories of a potential $\bar{V}_1$ for which $\bar{r}_{1i} + 1 = \bar{s}_{1i} = (L + 1)/2$ for odd L, $\bar{s}_{1i} = \frac{1}{2}L + 1$ for even L.

As the potential V_1 constructed in Theorem 4.1 has at most $2I(L + 1)$ trajectories more than those of $\bar{V}_1$, we can state that the minimum number $m_{1\,\min}$ of trajectories of potential V_1, as V_1 varies in the set of potentials satisfying (4.3) and (4.4), satisfies the condition

$$m_{1\,\min} \leq I(3L + 4).$$

By Theorem 4.2, every potential V_1 satisfying (4.3) and (4.4) for small enough ϵ must have $m_{1\,\min}$ Regge trajectories as near as we please to certain Regge trajectories, save the exception in k that we have seen earlier.

If we want to associate to potentials of class F, which does not possess any analytic property in l, all the "common" properties ("common" in the sense of the approximation problem), it is thus reasonable to choose the properties of the class of potentials V_1 satisfying (4.3) and (4.4) and having, for small enough ϵ, $m_{1\,\min}$ trajectories.

It turns out that, by restricting oneself to the above class of potentials, we can prove the following statement for the total scattering amplitudes: Given a potential V of class F, there exists a set $\{k_t, T\}$ such that for any choice of $\rho_1, \rho_2, \cdots, \rho_T$, for any $R > 0$, for any $k \in \{\overline{k_t, \rho_t, T}\}$, for any z such that $|z| < R$, for any $\Delta > 0$, we can satisfy the equation

$$|A_1(k, z) - A_2(k, z)| < \Delta, \tag{4.12}$$

provided V_1 and V_2 are chosen so that:

(a) $V_{1L} = V_{2L} = V$;

(b) V_1 and V_2 have only $m_{1\,\min}$ Regge trajectories;

(c) for small enough ϵ and for physical values of k and z,

$$|A_1(k, z) - A_{1L}(k, z)| < \epsilon,$$

$$|A_1(k, z) - A_2(k, z)| < \epsilon.$$

This result is achieved by remembering Theorems 4.1 and 4.2 and Eq. (1.4).

ACKNOWLEDGMENTS

The authors are grateful to Professor A. Agodi for stimulating discussions and advice. They also thank Dr. M. Baldo for interesting discussions.

APPENDIX A

It is easy to remove the restriction we have made in Ref. 11 of I, by proving that all the results we have obtained there hold also for a potential having both attractive and repulsive parts.

In this case we may substitute the following equation for Eq. (2.1) of I:

$$\langle \mathbf{p}| V |\mathbf{p}'\rangle = -4\pi \frac{\Gamma}{M} \left\{ \sum_{l=0}^{L} \sum_{j=1}^{I_l} \sum_{m=-l}^{l} g_{jl}(p) g_{jl}(p') Y_l^m(\hat{\mathbf{p}}) Y_l^{m*}(\hat{\mathbf{p}}') + \sum_{j=I_l+1}^{I_l+J_l} i g_{jl}(p) i g_{jl}(p') Y_l^m(\hat{\mathbf{p}}) Y_l^{m*}(\hat{\mathbf{p}}') \right\}, \quad \text{(A1)}$$

where the $g_{jl}(p)$ $(j = 1, 2, \cdots, I_l + J_l)$ satisfy the conditions we have imposed in I.

One gets

$$D_l(-k) = \det |\delta_{mn} + \Gamma f_{m,n;l}(-k)|\,;$$

$m, n = 1, 2, \cdots, I_l + J_l$, where if $m, n = 1, 2, \cdots, I_l$ or

$$m, n = I_l + 1, I_l + 2, \cdots, I_l + J_l,$$

then

$$f_{m,n;l}(-k) = \pm 4\pi \int_0^\infty dk_2 \frac{k_2^2 g_{ml}(k_2) g_{nl}(k_2)}{k^2 - k_2^2}, \qquad \text{Im } k > 0,$$

and it is given by analytic continuation in the whole complex k plane. Therefore, in this case, $f_{m,n;l}(-k)$ is real for purely imaginary k.

If $m = 1, 2, \cdots, I_l; n = I_l + 1, I_l + 2, \cdots, I_l + J_l$ or $m = I_l + 1, I_l + 2, \cdots, I_l + J_l; n = 1, 2, \cdots, I_l$, then

$$f_{m,n;l}(-k) = 4\pi i \int_0^\infty dk_2 \frac{k_2^2 g_{ml}(k_2) g_{nl}(k_2)}{k^2 - k_2^2}, \quad \text{Im } k > 0,$$

and it is given by analytic continuation in the whole complex k plane. Therefore, in this case, $f_{m,n;l}(-k)$ is purely imaginary for purely imaginary k.

We want to show that for purely imaginary k, $D_l(-k)$ is real, since starting from this point we can repeat what we said in I.

The determinant of the matrix A', obtained by multiplying by i all the elements of the matrix $(\delta_{mn} + \Gamma f_{m,n;l}(-k))$ whose row indices are greater than I_l, and then all the elements whose column indices are greater than I_l is given by

$$\det A' = i^{2J_l} D_l(-k) = \pm D_l(-k).$$

A' is a real matrix for purely imaginary k; therefore $D_l(-k)$ would also be real for purely imaginary k.

Q.E.D.

APPENDIX B

In the choice of the conditions on the $c_i(l)$ of Eq. (1.2), we have to take into account some differences for the Sommerfeld–Watson transform with respect to the case of local potentials.

In order to make the comparison we briefly recall first how the Sommerfeld–Watson transform works for a Yukawa potential, and then we show the main differences we have for a potential described by Eq. (1.2).

Let us consider in the complex $\lambda = l + \frac{1}{2}$ plane the contour shown in Fig. 8; it is composed of the following:

(i) of two arcs of circle

$$C_{ab}(\lambda_0, \epsilon): \quad \lambda = \lambda_0 e^{i\theta}, \quad 0 < \epsilon \leq \theta \leq \pi/2,$$

$$C_{cd}(\lambda_0, \epsilon): \quad \lambda = \lambda_0 e^{i\theta}, \quad -\pi/2 \leq \theta \leq -\epsilon;$$

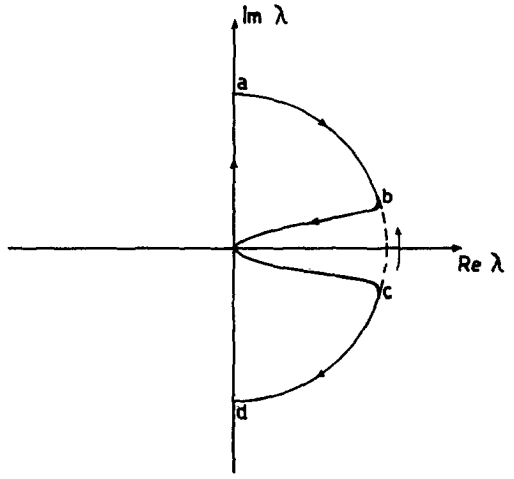


FIG. 8. Contour in the λ plane, relative to the Sommerfeld–Watson transform.

(ii) of the curve $C_{bc}(\lambda_0, \epsilon, \alpha)$, where α denotes a set of parameters, which in addition to λ_0 and ϵ determine such a curve;

(iii) of the segment $C_{da}(\lambda_0)$ on the imaginary axis.

We shall call such a curve $C(\lambda_0, \epsilon, \alpha)$.

The curve which is dashed in Fig. 8 will be called

$$C_1(\lambda_0, \epsilon): \quad \lambda = \lambda_0 e^{i\theta}, \quad \epsilon \geq \theta \geq -\epsilon.$$

For a generalized Yukawa potential

$$V(r) = \int_{\mu_0}^{\infty} d\mu\alpha(\mu)e^{-\mu r}/r, \quad -1 \leq z \leq 1,$$

the sum of the series

$$(2ik)^{-1} \sum_{l=0}^{\infty} (2l + 1)[S_l(k) - 1]P_l(z)$$

converges to the function $A(k, z)$. For $S_l(k)$ (defined for natural l) there exists an analytic interpolation in the complex l plane such that $S_l(k)$ is, for fixed real k, an analytic function in the right half plane, but for a finite number of poles, none of which lies on the real axis.[26]

λ_0, ϵ, and α can be determined in such a way that in the inner region of the curve $C_{bc}(\lambda_0, \epsilon, \alpha)$ there are no poles of $S_l(k)$. Taking into account the analytic properties of $P_\lambda(-z)$, $\cos \pi\lambda$, if $n_1 < \lambda_0 < n_1 + 1$, where n_1 is a natural number, one gets

$$(2ik)^{-1} \sum_{l=0}^{n_1} (2l + 1)[S_l(k) - 1]P_l(z)$$

$$= (2k)^{-1} \int_{C_{bc}(\lambda_0,\epsilon,\alpha)+C_1(\lambda_0,\epsilon)} d\lambda\lambda[S(\lambda, k) - 1] \frac{P_{\lambda-\frac{1}{2}}(-z)}{\cos \pi\lambda},$$

$$A(k, z) = \lim_{\lambda_0 \to \infty} (2k)^{-1} \int_{C_{bc}(\lambda_0,\epsilon,\alpha)+C_1(\lambda_0,\epsilon)} d\lambda \times \lambda[S(\lambda, k) - 1] \frac{P_{\lambda-\frac{1}{2}}(-z)}{\cos \pi\lambda}.$$

Let us now remember that[27]

$$S(\lambda, k) - 1 = O(\lambda^{-\frac{1}{2}} e^{-2\lambda a}) \quad \text{for} \quad \operatorname{Re} \lambda \to \infty, \quad \text{(B1)}$$

where

$$a = \ln\left[\frac{\mu_0}{2k} + \left(1 + \frac{\mu_0^2}{4k^2}\right)^{\frac{1}{2}}\right];$$

$$|P_{\lambda-\frac{1}{2}}(-z)/\cos \pi\lambda| < |\lambda|^{-\frac{1}{2}} \exp(-\operatorname{Re} \theta\, |\operatorname{Im} \lambda|) f(z) \quad \text{(B2)}$$

for $z^2 < 1$, $\operatorname{Re} \lambda > 0$, $|\lambda| \to \infty$, where $f(z)$ is a function of the only $z = \cos \theta$ (Ref. 28).

From (B1) and (B2), one deduces

$$\lim_{\lambda_0 \to \infty} \int_{C_1(\lambda_0,\epsilon)} d\lambda\lambda[S(\lambda, k) - 1] \frac{P_{\lambda-\frac{1}{2}}(-z)}{\cos \pi\lambda} = 0. \quad \text{(B3)}$$

Moreover, if the Yukawian potential satisfies the condition $|V(iy)| < N_\alpha/y^\alpha$, with $\frac{3}{2} < \alpha < 2$ (see Ref. 29), then

$$S(\lambda, k) - 1 = O(\lambda^{\frac{3}{2}-\alpha}) \quad \text{for} \quad |\lambda| \to \infty. \quad \text{(B4)}$$

From (B1), (B2), and (B4), one gets

$$\lim_{\lambda_0 \to \infty} \int_{C_{ab}(\lambda_0,\epsilon)} d\lambda\lambda[S(\lambda, k) - 1] \frac{P_{\lambda-\frac{1}{2}}(-z)}{\cos \pi\lambda}$$

$$= \lim_{\lambda_0 \to \infty} \int_{C_{cd}(\lambda_0,\epsilon)} d\lambda\lambda[S(\lambda, k) - 1] \frac{P_{\lambda-\frac{1}{2}}(-z)}{\cos \pi\lambda} = 0. \quad \text{(B5)}$$

Moreover, for $z^2 < 1$, $\lambda' \to \pm\infty$ (see Ref. 30),

$$P_{i\lambda'-\frac{1}{2}}(-z) \simeq 2(i\lambda' - \tfrac{1}{2})^{-\frac{1}{2}} \exp\{2|\lambda'|\, g(z)\} f(z), \quad \text{(B6)}$$

where $g(z) = \arctan(1 + z/1 - z)^{\frac{1}{2}}$, and then $g(z) < \pi/2$.

By applying the theorem of residue to the contour $C(\lambda_0, \epsilon, \alpha)$ and taking into account (B3) and (B5), one gets

$$A(k, z) = (2k)^{-1} \int_{-\infty}^{\infty} d\lambda\lambda \frac{S(i\lambda, k) - 1}{\cosh \pi\lambda} P_{i\lambda-\frac{1}{2}}(-z) + i\pi k^{-1} \sum_n (\alpha_n + \tfrac{1}{2})\beta_n P_{\alpha_n}(-z)/\sin \pi\alpha_n, \quad \text{(B7)}$$

where α_n ($n = 1, 2, \cdots, N$) are the N poles of $S_l(k)$ in the right half l plane and β_n are their residues.

The integral in (B7) exists by (B4) and (B6).

The following equation holds true in the z plane cut from 1 to ∞ for $\lambda \to \pm\infty$ (Ref. 30):

$$|P_{i\lambda-\frac{1}{2}}(-z)| \simeq 2(2\pi)^{-\frac{1}{2}} |i\lambda - \tfrac{1}{2}|^{-} \, |z^2 - 1|^{-\frac{1}{4}} \times |\exp(2i\lambda \ln |A|) \exp(-2\lambda \arg A) + i \exp(2i\lambda \ln |B|) \exp(-2\lambda \arg B)|, \quad \text{(B8)}$$

where

$$A = (-z + 1)^{\frac{1}{2}} + (-z - 1)^{\frac{1}{2}},$$
$$B = (-z + 1)^{\frac{1}{2}} - (-z - 1)^{\frac{1}{2}}.$$

From (B8) and (B4) one deduces that the rhs of Eq. (B7) has a meaning also for complex z, and is an analytic function of z apart from a branch point.

Let us now consider how a method of this kind can be carried over a potential defined by Eq. (1.2). We get rid of the conditions (B1) and (B4) which

[26] Ref. 18, Chap. VIII.
[27] Ref. 11, p. 44.
[28] Ref. 18, p. 99.
[29] Ref. 18, p. 90.
[30] R. G. Newton, *Scattering Theory of Waves and Particles* (McGraw-Hill Book Co., New York, 1966), p. 402.

belong to Yukawian potentials. From (B2) and (B6), it is easy to deduce that if for $z^2 < 1$, $|\lambda| \to \infty$,

$$\lambda[S(\lambda, k) - 1]P_{\lambda-\frac{1}{2}}(-z)/\cos \pi\lambda = O(\lambda^\beta), \quad \text{(B9)}$$

with $\beta = -(1 + \alpha)$ and arbitrary $\alpha > 0$, then a necessary and sufficient condition for (B3), (B5), and (B7) to hold is that for $|\lambda| \to \infty$, $\operatorname{Re} \lambda \geq 0$, there exists a positive number C such that

$$|S(\lambda, k) - 1| \leq C\,|\lambda|^{-(\frac{3}{2}+\alpha)} \exp(\theta\,|\operatorname{Im} \lambda|) f(z).$$

From Eq. (B8) it is easy to deduce that if, for z belonging to the complex plane cut from 1 to ∞, Eq. (B9) holds, then a necessary and sufficient condition for the convergence of the integral in (B7) is that there exists a positive number D such that

$$|S(\lambda, k) - 1| \leq D\,|\lambda|^{-(\frac{3}{2}+\alpha)}\,|z^2 - 1|^{-\frac{1}{4}}\,[\exp \pi\lambda + \exp(-\pi\lambda)]/|\exp(2i\lambda \ln |A|) \exp(-2\lambda \arg A) + i \exp(2i\lambda \ln |B|) \exp(-2\lambda \arg B)|.$$

Under these conditions, (B7) gives the analytic continuation in z of $A(k, z)$ defined as the sum of the series.

We shall see that it is possible to choose the $c_i(l)$ in such a way that (B5) holds and the integral appearing in (B7) exists and is finite, while

$$A(k, z) = (2ik)^{-1} \sum_{l=0}^{\infty} (2l + 1)[S_l(k) - 1]P_l(z)$$

diverges.

If we choose $c_i(l) = p_i(l)/q_i(l)$, where $p_i(l)$ and $q_i(l)$ are polynomials in l and the degree of $p_i(l)$ is greater or equal to the degree of $q_i(l)$, then this fact is immediately verified. There is no contradiction since the rhs of Eq. (B7) does not give the sum of the series because Eq. (B3) does not hold.

It is also easy to verify that hypotheses (a), (b), (c), and (d) of Sec. 2 are sufficient (but not necessary) conditions for Eqs. (B3), (B5), and (B7) to hold for physical values of z and the rhs of Eq. (B7) gives the analytic continuation in the z plane cut from 1 to ∞.

APPENDIX C

Let A be the statement "$A(k, z)$ is a continuous function for physical values of its arguments"; let B be the statement "the series

$$(2ik)^{-1} \sum_{l=0}^{\infty} (2l + 1)[S_l(k) - 1]P_l(z)$$

is convergent" (punctually); let C be the statement

$$\text{"}(4k^2)^{-1} \sum_{l=0}^{\infty} (2l + 1)\,|S_l(k) - 1|^2 < \infty.\text{"}$$

We have to prove that $A \Rightarrow (B \Leftrightarrow C)$. Let us first prove that $A \Rightarrow (B \Rightarrow C)$. Indeed if A and B are true, $A(k, z)$ is a continuous function of z; by the properties of the Legendre polynomials we get

$$(4k^2)^{-1} \sum_{l=0}^{\infty} (2l + 1)\,|S_l(k) - 1|^2 = \int_{-1}^{1} |A(k, z)|^2\,dz < \infty.$$

Now we have to prove that $A \Rightarrow (C \Rightarrow B)$. Indeed, if A and C are true, the series

$$(2ik)^{-1} \sum_{l=0}^{\infty} (2l + 1)[S_l(k) - 1]P_l(z)$$

converges in the mean to the continuous function $A(k, z)$ by the theorem of Riesz and Fischer.[15] That is, having defined

$$f_L(k, z) = \left| A(k, z) - (2ik)^{-1} \sum_{l=0}^{L} (2l + 1) \times [S_l(k) - 1]P_l(z) \right|^2$$

and

$$\delta_L^2 = \int_{-1}^{1} f_L(k, z)\,dz,$$

and having arbitrarily fixed $\epsilon' > 0$, there exists a positive integer $L'(\epsilon')$ such that

$$\delta_L^2 < \epsilon' \quad \text{(C1)}$$

for any $L > L'(\epsilon')$. The proof follows from the following argument.

Let us suppose that B is false; then there exists $\bar{z} \in (-1, 1)$ such that, having fixed $\epsilon > 0$ and a corresponding $\bar{L}(\epsilon)$,

$$f_L(k, \bar{z}) \geq \epsilon, \text{ for any } \quad L > \bar{L}(\epsilon).$$

But $f_L(k, z)$ is a continuous function of z, because it is the sum of continuous functions; therefore, having fixed ϵ_1 such that $0 < \epsilon_1 < \epsilon$, a corresponding $\Delta(\epsilon_1) > 0$ can be determined for which

$$f_L(k, z) > \epsilon_1, \text{ for any } \quad z \in (\bar{z} - \Delta, \bar{z} + \Delta), \quad L > \bar{L}(\epsilon), \quad -1 \leq \bar{z} - \Delta < \bar{z} + \Delta \leq 1.$$

But it is

$$\delta_L^2 \geq \int_{\bar{z}-\Delta}^{\bar{z}+\Delta} f_L(k, z)\,dz \geq 2\Delta\epsilon_1,$$

and for $\epsilon' = 2\Delta\epsilon_1$ the last relation contradicts Eq. (C1). Q.E.D.