JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 5 MAY 1968

# Error Bounds in Equilibrium Statistical Mechanics

Roy G. GORDON*
*Department of Chemistry, Harvard University, Cambridge, Massachusetts*

(Received 20 October 1967)

A new method is presented for the calculation of thermodynamic properties from equilibrium statistical mechanics. Starting from the high-temperature expansion coefficients for the canonical partition function, error bounds are obtained, which are both rigorous and optimal.

## I. INTRODUCTION

The high-temperature expansion method is one of the most widely used techniques in statistical mechanics. It has been used to study the thermodynamic properties of crystalline solids,[1] binary alloys,[2] magnetic properties,[3] pure fluids,[4] fluid mixtures,[5] and condensation from gases.[6] The chief advantage of the method is its wide applicability. The main difficulties of the method are twofold: (1) The series of approximations converges rather slowly, and in some cases the series does not converge at all at low temperatures. (2) The results are of unknown accuracy, when the series are extrapolated.[7]

The purpose of this paper is to develop a new method of high-temperature expansion which helps both of these difficulties. Starting from the coefficients in the usual (truncated) high-temperature series, we derive a new sequence of approximations, which (1) converges much more rapidly than the usual high-temperature expansion, and (2) gives precise upper and lower bounds for the partition function (and some other thermodynamic properties), at each order of approximation. These bounds are optimal in the sense that they are the most precise bounds possible, given the coefficients in the usual high-temperature series.

## II. STATEMENT OF PROBLEM: ASSUMPTIONS

We consider closed equilibrium classical or quantum-statistical systems which are described by canonical distribution law over energy $E$ which we write as

$$e^{-\beta E}\, d\psi(E), \tag{1}$$

where $\beta$ is the reciprocal temperature, and $d\psi(E)$ is a density of states. We always assume that $\psi(E)$ is a nondecreasing function of $E$. We choose the (arbitrary) zero of energy to be the energy of the ground state of the system. Thus $d\psi(E)$ vanishes for $E < 0$.

The problem we pose is to find upper and lower bounds for the canonical partition function defined by the Stieltjes integral

$$Q(\beta) \equiv \int_0^\infty e^{-\beta E}\, d\psi(E), \tag{2}$$

when we are given values for the first $2M$ moments

* Sloan Foundation Fellow.

[1] H. Thirring, Physik Z. **14**, 867 (1913); **15**, 127, 180 (1914); E. W. Montroll, J. Chem. Phys. **10**, 218 (1942); **11**, 481 (1943); E. W. Montroll and D. C. Peaslee, J. Chem. Phys. **12**, 98 (1944); A. A. Maradudin, P. Mazur, E. E. Montroll, and G. H. Weiss, Rev. Mod. Phys. **30**, 175 (1958); C. Domb, A. A. Maradudin, E. W. Montroll, and G. H. Weiss, Phys. Rev. **115**, 18, 24 (1959); J. C. Bradley, Ann. Phys. (New York) **15**, 411 (1961).

[2] J. G. Kirkwood, J. Chem. Phys. **6**, 70 (1938); T. S. Chang, J. Chem. Phys. **9**, 169 (1941); G. S. Rushbrooke, Nuovo Cimento Suppl. **6**, 251 (1949); T. Oguchi, J. Phys. Soc. (Japan) **6**, 499 (1951).

[3] W. Opechowski, Physica **4**, 181 (1937); **6**, 1112 (1938); H. A. Kramers and G. H. Wannier, Phys. Rev. **60**, 252, 263 (1941); E. W. Montroll, J. Chem. Phys. **10**, 61 (1942); L. Onsager, Phys. Rev. **65**, 117 (1944); D. ter Haar, Phys. Rev. **76**, 176 (1949); C. Domb, Proc. Roy. Soc. (London) **A199**, 199 (1949); G. S. Rushbrooke, Nuovo Cimento **6**, Suppl. 2, 251 (1949); V. Zehler, Z. Naturforsch.**A5**, 344 (1950); E. Trefftz, Z. Physik **127**, 371 (1950); J. E. Brooks and C. Domb, Proc. Roy. Soc. (London) **A207**, 343 (1951); A. J. Wakefield, Proc. Cambridge Phil. Soc. **47**, 419 (1951); T. Oguchi, J. Phys. Soc. (Japan) **5**, 75 (1950); **6**, 31 (1951); M. Kurata, R. Kikuchi, and T. Waturi, J. Chem. Phys. **21**, 434 (1953); G. S. Rushbrooke and P. J. Wood, Proc. Phys. Soc. (London) **A68**, 1161 (1955); **A70**, 765 (1956); C. Domb and M. F. Sykes, Proc. Phys. Soc. (London) **B69**, 486 (1956); M. F. Sykes, Phil. Mag. **2**, 733 (1957); C. Domb and M. F. Sykes, Phys. Rev. **108**, 1415 (1957); S. Katsura, Progr. Theoret. Phys. **20**, 192 (1958); G. S. Rushbrooke and P. J. Wood, Mol. Phys. **1**, 257 (1958); C. Domb, Advan. Phys. **9**, 149 (1960); M. F. Sykes, J. Math. Phys. **2**, 52 (1961); C. Domb and M. F. Sykes, J. Math. Phys. **2**, 63 (1961); G. S. Rushbrooke and J. Eve, *ibid.* **3**, 185 (1962); C. Domb and M. F. Sykes, Phys. Rev. **128**, 168 (1962); G. S. Rushbrooke and P. J. Wood, Mol. Phys. **6**, 409 (1963); C. Domb and D. W. Wood, Proc. Phys. Soc. **86**, 1 (1965); G. A. Baker, H. E. Gilbert, J. Eve, and G. S. Rushbrooke, Phys. Letters **20**, 146 (1966); G. A. T. Allan and D. D. Betts, Proc. Phys. Soc. (London) **91**, 341 (1967).

[4] R. W. Zwanzig, J. Chem. Phys. **22**, 1429 (1954); **23**, 1915 (1955); J. S. Rowlinson, Mol. Phys. **7**, 349 (1964); **8**, 107 (1964); D. A. McQuarrie and J. L. Katz, J. Chem. Phys. **44**, 2393 (1966).

[5] J. C. Wheeler, thesis,Cornell University, 1967.

[6] C. N. Yang and T. D. Lee, Phys. Rev. **87**, 404, 410 (1952); J. Wang, Proc. Roy. Soc. (London) **A161**, 127 (1937); J. E. Brooks and C. Domb, Proc. Phys. Soc. (London) **A207**, 343 (1951).

[7] D. Park, Physica **12**, 932 (1956); G. A. Baker, Phys. Rev. **124**, 768 (1961); G. A. Baker and J. L. Gammel, J. Math. Anal. Appl. **2**, 21 (1961); G. A. Baker, J. L. Gammel, and J. G. Wills, *ibid.* **2**, 405 (1961); M. E. Fisher and J. W. Essam, J. Chem. Phys. **38**, 802 (1963); G. A. Baker, Phys. Rev. **129**, 99 (1963); D. S. Gaunt and M. E. Fisher, J. Chem. Phys. **43**, 2840 (1965); G. A. Baker and D. S. Gaunt, Phys. Rev. **155**, 545 (1967).

$\mu_n$ of the distribution function, defined by

$$\mu_n \equiv \int_0^\infty E^n \, d\psi(E) \tag{3}$$

for $n = 0, 1, 2, \cdots, 2M$.

In order for these moments to exist and be finite in a classical mechanical system, it is important to separate out the kinetic energy, and treat (2) as the configurational integral, and the moments $\mu_n$ as averages of powers of the potential energy only:

$$\mu_n = \int_0^\infty V^n \, d\psi(V). \tag{4}$$

The integrals then are just configuration-space averages of powers of the potential energy. Having a potential-energy function of bounded variation is then sufficient to guarantee their existence.

In quantum-mechanical spin systems (such as the Heisenberg model of ferromagnetism) the energy $E$ in (3) may be taken to be the full spin Hamiltonian $H$. The moments may then be evaluated from the quantum-mechanical trace formula

$$\mu_n = \mathrm{Tr}\,[H^n]. \tag{5}$$

The power of this method is that one may calculate the trace in any convenient basis.

The usual form of the high-temperature expansion is now readily obtained by expanding the exponential in (2) in its power series and assuming that one can interchange the order of summation and integration,

$$Q(\beta) = \sum_{n=0}^{\infty} (-\beta)^n \mu_n / n!. \tag{6}$$

By truncating this sum at successively higher (finite) numbers of terms, one obtains a sequence of approximations to the partition function. The higher the temperature, the smaller $\beta$, and hence the series converges most rapidly at high temperatures. Unfortunately, the convergence is often slow at temperatures of interest.

Since the terms in the series (6) alternate in sign, successive partial sums give crude upper and lower bounds to the infinite sum. However, these bounds are not too useful in practice, because for sufficiently low temperatures the lower bounds become negative; hence they furnish no new information, since we knew anyway that the partition function cannot be negative. Similarly, these upper bounds become larger than $Q(0)$, for low temperatures. But we knew already that $Q(\beta) < Q(0)$, so these upper bounds also fail to yield any information at low temperatures. Even when the temperature is high enough so that these bounds furnish some information, the magnitudes of the error limits are so large (for reasonable $n \sim 10$), that one might be lead to the hasty conclusion that a small finite number of terms in the high-temperature expansion tells one little about the partition function.

We hope to demonstrate in the following sections of this paper that such a conclusion is unduly pessimistic. We will construct a new set of error bounds, based on precisely that information contained in the coefficients of the usual high-temperature series. Our new error bounds are far more precise, by more than a factor of a million in an example given in Sec. V.

## III. GENERAL THEORY

We make use of mathematical results from the theories of continued fractions,[8] quasi-orthogonal polynomials and the moment problem,[9] matrix algebra,[10] and Gaussian-type integration[9]; reference should be made to these works for further mathematical background. Where possible, we follow the definitions, terminology and notation of Shohat and Tamarkin[9] (to be referred to as ST in the following).

Consider[11a] the function $I(z)$ defined by the Stieltjes integral

$$I(z) \equiv \int_0^\infty \frac{d\psi(E)}{z + E} \tag{7}$$

over the nondecreasing distribution $d\psi(E)$. The integrand may be expanded according to a finite geometric series, with remainder term

$$\frac{1}{z+E} = \frac{1}{z} - \frac{E}{z^2} + \frac{E^2}{z^3} - \cdots + \frac{(-E)^{n-1}}{z^n} + \frac{(-E)^n}{z^{n+1}(z+E)}. \tag{8}$$

Inserting this series into the integrand in (7) gives

$$I(z) = \frac{1}{z}\int_0^\infty d\psi(E) - \frac{1}{z^2}\int_0^\infty E \, d\psi(E) + \frac{1}{z^3}\int_0^\infty E^2 \, d\psi(E) - \cdots + \frac{1}{z^n}\int_0^\infty (-E)^{n-1} \, d\psi(E) + \frac{1}{z^{n+1}}\int_0^\infty \frac{(-E)^n \, d\psi(E)}{z+E}. \tag{9}$$

The coefficients of the inverse powers of $z$, are recognized to be just the moments $\mu_n$ defined by (3).

[8] H. S. Wall, *Analytic Theory of Continued Fractions* (D. Van Nostrand, Inc., New York, 1948).

[9] J. A. Shohat and J. D. Tamarkin, "The Problem of Moments," Mathematical Surveys **1** (American Mathematical Society, Providence, R.I., 1950) 2nd ed.

[10] J. H. Wilkinson, *The Algebraic Eigenvalue Problem* (Oxford University Press, London, 1965).

[11a] T. J. Stieltjes, *Recherches sur les fractions continues*, Annales dè la Faculté des Sciences de Toulouse **8**, 1 (1894); **9**, 5 (1895).

Thus the formal expansion of the integral (7) has the form

$$I(z) = \frac{\mu_0}{z} - \frac{\mu_1}{z^2} + \frac{\mu_2}{z^3} - \cdots + (-1)^{n-1}\frac{\mu_{n-1}}{z^n} + \cdots + \frac{1}{z^{n+1}}\int_0^\infty \frac{(-E)^n\, d\psi(E)}{z+E}. \quad (10)$$

As long as the remainder term is kept, this is just an identity. However, if we let $n \to \infty$, (10) becomes a formal expansion of the integral in inverse powers of $z$:

$$I(z) \sim \frac{\mu_0}{z} - \frac{\mu_1}{z^2} + \frac{\mu_2}{z^3} - \frac{\mu_3}{z^4} + \cdots. \quad (11)$$

This is ordinarily an asymptotic, but not convergent, series. In order to form a more useful expression, we replace this series by its "corresponding continued fraction"

$$C(z) = \cfrac{\alpha_1}{z + \cfrac{\alpha_2}{1 + \cfrac{\alpha_3}{z + \cfrac{\alpha_4}{1 + \cdots}}}}. \quad (12)$$

The coefficients $\alpha_n$ appearing in this expansion are determined from the moments $\mu_n$, by the requirement that a formal expansion of $C(z)$ in powers of $(1/z)$ have just the same coefficients as those appearing in (11). Values of the first few $\alpha_n$ may be derived directly from this definition by equating coefficients of $(1/z)^n$ in these expansions:

$$\alpha_1 = \mu_0, \quad (13a)$$

$$\alpha_2 = \mu_1/\mu_0, \quad (13b)$$

$$\alpha_3 = \frac{(\mu_2\mu_0 - \mu_1^2)}{\mu_0\mu_1}, \quad (13c)$$

$$\alpha_4 = \frac{(\mu_1\mu_3 - \mu_2^2)\mu_0}{\mu_1(\mu_2\mu_0 - \mu_1^2)}, \quad (13d)$$

$$\alpha_5 = \frac{(\mu_0^2\mu_1\mu_2\mu_4 - \mu_0\mu_1^3\mu_4 - \mu_0\mu_1\mu_2^3 - \mu_0^2\mu_1\mu_3^2 + 2\mu_0\mu_1^2\mu_2\mu_3)}{\mu_0(\mu_1\mu_3 - \mu_2^2)(\mu_0\mu_2 - \mu_1^2)}. \quad (13e)$$

Obviously this direct matching method cannot be practically applied to higher orders. Explicit general expressions for the $\alpha_n$ in terms of the $\mu_n$ can be written down,[9] but they involve determinants of high order, and are not useful in practice. A convenient recursive method for evaluating the $\alpha_n$ is given in Appendix A. In applications, we have always used this recursive method, which we will refer to as the product-difference (PD) algorithm. It follows from the general theory[9,11] that $\alpha_n \geq 0$, provided that $\psi(E)$ is a nondecreasing function which vanishes for $E < 0$.

An infinite continued fraction such as (12) has a mathematical meaning only as a limit. If we consider a sequence of "approximants" $C_n(z)$, which are the finite, truncated fractions obtained by setting $\alpha_{n+1} = \alpha_{n+2} = \cdots 0$, then $C(z)$ is defined to be

$$C(z) = \lim_{n\to\infty} C_n(z). \quad (14)$$

Infinite subsets of $C_n(z)$ may also be considered in this limit. In particular, the subsequence of the *even* approximants $C_{2n}(z)$ and the *odd* approximants $C_{2n-1}(z)$ will play an important role in finding our error bounds.

First we consider the *even* approximants $C_{2n}(z)$. We claim that the following continued fraction $A^e(z)$ has its approximants of order $n$ exactly equal to the even approximants $C_{2n}(z)$:

$$A^e(z) = \cfrac{\alpha_1}{z + \alpha_2 - \cfrac{\alpha_2\alpha_3}{z + \alpha_3 + \alpha_4 - \cfrac{\alpha_4\alpha_5}{z + \alpha_5 + \alpha_6 - \cdots}}}. \quad (15)$$

The equality $A_n^e(z) = C_{2n}(z)$ may be verified by truncating (12) and (15) at appropriate points and rearranging the rational functions thus obtained. [*Definition:* A continued fraction whose approximants match a subset of the approximants of another continued fraction, is called a *contraction* of that other fraction. Thus $A^e(z)$ is a contraction of $C(z)$. In this particular case the contraction $A^e(z)$ has the special name, "the *associated* continued fraction of *even* order, of the corresponding continued fraction $C(z)$."]

We now claim that an exactly equivalent, but more useful, expression for $A_n^e(z)$ is the following:

$$A_n^e(z) = x_1,$$

where $x_1$ is the first component of the solution to the following set of $n$ simultaneous linear equations:

$$\begin{array}{llll}
(z+\alpha_2)x_1 & -(\alpha_2\alpha_3)^{\frac{1}{2}}x_2 & +0x_3 & +0x_4 \quad \cdots = \alpha_1 \\
-(\alpha_2\alpha_3)^{\frac{1}{2}}x_1 & +(z+\alpha_3+\alpha_4)x_2 & -(\alpha_4\alpha_5)^{\frac{1}{2}}x_3 & +0x_4 \quad \cdots = 0 \\
0x_1 & -(\alpha_4\alpha_5)^{\frac{1}{2}}x_2 & +(z+\alpha_5+\alpha_6)x_3 & -(\alpha_6\alpha_7)^{\frac{1}{2}}x_4 \quad \cdots = 0 \\
\cdots & & & \\
\cdots & & & +0x_{n-2} \quad -(\alpha_{2n-2}\alpha_{2n-1})^{\frac{1}{2}}x_{n-1} \quad +(z+\alpha_{2n-1}+\alpha_{2n})x_n = 0.
\end{array} \quad (16)$$

The equivalence of these two rather different looking expressions for $A_n^e(z)$ is demonstrated in Appendix B.[11b]

The linear equations (16) are more conveniently examined in matrix notation:

$$(z\mathsf{I} + \mathsf{M}) \cdot \mathbf{x} = \mathbf{e}_1 \alpha_1, \tag{17}$$

where $\mathsf{I}$ is the $n \times n$ unit matrix, $\mathbf{e}_1$ is the unit vector

$$\mathbf{e}_1 \equiv \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{18}$$

with $n$ rows, and $\mathsf{M}$ is an $n \times n$ symmetric tridiagonal matrix

$$\mathsf{M} \equiv \begin{pmatrix} \alpha_2 & -(\alpha_2\alpha_3)^{\frac{1}{2}} & 0 & 0 & 0 & \cdots \\ -(\alpha_2\alpha_3)^{\frac{1}{2}} & (\alpha_3 + \alpha_4) & -(\alpha_4\alpha_5)^{\frac{1}{2}} & 0 & 0 & \cdots \\ 0 & -(\alpha_4\alpha_5)^{\frac{1}{2}} & (\alpha_5 + \alpha_6) & -(\alpha_6\alpha_7)^{\frac{1}{2}} & 0 & \cdots \\ & & \cdots & & & \\ & & & & -(\alpha_{2n-2}\alpha_{2n-1})^{\frac{1}{2}} & (\alpha_{2n-1} + \alpha_{2n}) \end{pmatrix}. \tag{19}$$

The formal solution to these linear equations is

$$\mathbf{x} = \alpha_1 (z\mathsf{I} + \mathsf{M})^{-1} \mathbf{e}_1 . \tag{20}$$

If we now transform to the basis in which the matrix $\mathsf{M}$ is diagonal, with eigenvalues $\xi_{jn}^e$, and transformation matrix $\mathsf{U}$,

$$\xi_{jn}^e = (\mathsf{U}^{-1}\mathsf{M}\mathsf{U})_{jj}, \tag{21}$$

we have

$$\mathbf{x} = \alpha_1 \mathsf{U}\mathsf{U}^{-1}(z\mathsf{I} + \mathsf{M})^{-1}\mathsf{U}\mathsf{U}^{-1}\mathbf{e}_1 \tag{22}$$

or

$$x_1 = \sum_j \alpha_1 (z + \xi_{jn}^e)^{-1} U_{1j} U_{j1}^{-1}. \tag{23}$$

Since the transformation $\mathsf{U}$ is orthogonal ($\mathsf{M}$ is real symmetric), we have $U_{1j} = U_{j1}^{-1}$ and

$$x_1 = \alpha_1 \sum_j (z + \xi_{jn}^e)^{-1} U_{1j}^2 \tag{24}$$

or

$$A_n^e(z) = \alpha_1 \sum_{j=1}^n \frac{U_{1j}^2}{(z + \xi_{jn}^e)}, \equiv \sum_{j=1}^n \frac{\rho_n^e(j)}{(z + \xi_{jn}^e)}, \tag{25}$$

where we have set

$$\rho_n^e(j) \equiv \alpha_1 U_{1j}^2 . \tag{26}$$

Thus we obtain the relation

$$\int_0^\infty \frac{d\psi(E)}{z + E} = \lim_{n\to\infty} A_n^e(z), = \lim_{n\to\infty} \sum_{j=1}^n \frac{\rho_n^e(j)}{(z + \xi_{jn}^e)}. \tag{27}$$

[11b] An equivalent result is given by Wall, Ref. 8, p. 226.

This is an integration formula for the function $(z + E)^{-1}$ with respect to the weight function $d\psi(E)$. The obvious generalization of (27) to a function $f(E)$ is

$$\int_0^\infty f(E)\, d\psi(E) = \lim_{n\to\infty} \sum_{j=1}^n f(\xi_{jn}^e)\rho_n^e(j). \tag{28}$$

Equation (28) is valid whenever the integral converges, provided $f(E)$ is analytic in a region including the positive real axis. In the integration formula (28), the weights $\rho_n^e(j)$ are obviously positive, from their definition (26), in which $U_{1j}$ is a real element of an orthogonal transformation matrix, and $\alpha_1$ is positive from (13a), (3) and the positive definiteness of the distribution $d\psi(E)$. It is shown in the general theory of the moment problem (ST, p. 109) that all the $\xi_{jn}$ are distinct and positive.

The correspondence between continued fractions and matrix theory which we have just exploited, makes it clear why inverse functions play such an important role. The inverse of a tridiagonal matrix can be calculated directly as a continued fraction, allowing one to evaluate integrals of the Stieltjes transform type (27). General integrals of the type (28) can be calculated only after finding the eigenvalues and eigenvectors of the matrix in (19), whereas for finding its inverse, the eigenvalue calculation is not needed.

What we have done here is to recast the theory of Gaussian-type integration with respect to an arbitrary weight function, into a form in which the positions and weights are determined by the solution to an eigenvalue problem for a real symmetric tridiagonal matrix. The usual form of this integration theory requires one to find roots of high-order polynomials.[9]

The numerical problem of finding roots of high-order polynomials is notoriously ill-conditioned.[10] In contrast, our solution is obtained in terms of the well-conditioned[10] eigenvalue problem of a real symmetric tridiagonal matrix, which can be solved accurately by a variety of procedures, without the catastrophic loss of accuracy associated with the corresponding polynomial problem.

The integration formula (28) was based on the even approximants $C_{2n}(z)$ to the corresponding continued fraction $C(z)$. A second integration formula of this type can be found, starting from the *odd* approximants $C_{2n-1}(z)$. The derivation follows steps analogous to those which have taken us from the even contraction $A^e(z)$ [Eq. (15)] to the even integration formula (28). The only difference is that in Eqs. (16) and (19), the coefficient $\alpha_{2n}$ is omitted. This of course changes the numerical values of the eigenvalues and the weights (in fact one of the eigenvalues now lies at $E = 0$), but the rest of the formal development remains unchanged, giving the odd integration formula

$$\int_0^\infty f(E)\, d\psi(E) = \lim_{n\to\infty} \sum_{j=1}^{n} \rho_n^o(j) f(\xi_{jn}^o). \tag{29}$$

If we now apply these general results to integrating the particular function

$$f(E) = e^{-\beta E}, \tag{30}$$

we obtain two expressions for the canonical partition function (2),

$$Q(\beta) \equiv \int_0^\infty e^{-\beta E}\, d\psi(E),$$

$$= \lim_{n\to\infty} \sum_{j=1}^{n} \rho_n^e(j) \exp(-\beta \xi_{jn}^e) \tag{31}$$

from (28), and

$$Q(\beta) = \lim_{n\to\infty} \sum_{j=1}^{n} \rho_n^o(j) \exp(-\beta \xi_{jn}^o) \tag{32}$$

from (29).

## IV. ERROR BOUNDS

In actual applications of interest in statistical mechanics, one has available only a finite number of moments $\mu_n$, say for $n = 0, 1, 2, \cdots, 2M$. Thus the infinite limiting processes called for in Eqs. (28) and (29) must be stopped at a finite value $n = M$. Thus we must add "remainder" terms to the truncated integration formulas. Expressions for these remainder terms are derived (ST, p. 119), which can be put into the form

$$\int_0^\infty f(E)\, d\psi(E) = \sum_{j=1}^{M} \rho_M^e(j) f(\xi_{jM}^e) + \frac{\alpha_{2M}\alpha_{2M+1} f^{(2M)}(x)}{(2M)!} \tag{33}$$

and

$$\int_0^\infty f(E)\, d\psi(E) = \sum_{j=1}^{M} \rho_M^o(j) f(\xi_{jM}^o) + \frac{\alpha_{2M-1}^1 \alpha_{2M-2}^1 f^{(2M-1)}(x)}{(2M-1)!}, \tag{34}$$

where $\alpha_n^1$ are the continued fraction coefficients for the moment problem in which $\mu_{j+1}$ replaces $\mu_j$; and where $x$ is somewhere in the interval $[0, \infty]$. Since the $\alpha_n$ coefficients are all positive (or zero), the remainder terms for the even or odd approximations take the same signs, respectively, as the even or odd derivatives of $f(E)$.

Applying these results to $f(E) = \exp(-\beta E)$, we see that the even derivatives in this case are positive, and the odd derivatives negative, uniformly on the whole range of integration. Thus we have the following two results, which supply our error bounds: (1) The even approximations to the partition function,

$$Q_M^e(\beta) \equiv \sum_{j=1}^{M} \rho_M^e(j) \exp(-\beta \xi_{jM}^e), \tag{35}$$

furnish a nondecreasing sequence of lower bounds to the partition function. (2) The odd approximations to the partition function,

$$Q_M^o(\beta) \equiv \sum_{j=1}^{M} \rho_M^o(j) \exp(-\beta \xi_{jM}^o), \tag{36}$$

furnish a nonincreasing sequence of upper bounds to the partition function.[11c]

Thus the possible error $\epsilon$ in the approximate partition function $\bar{Q}_M(\beta)$ is

$$\epsilon_M \leq \tfrac{1}{2}(Q_M^o - Q_M^e), \tag{37}$$

where the $M$th approximation to $Q(\beta)$ is

$$\bar{Q}_M(\beta) \equiv \tfrac{1}{2}(Q_M^o + Q_M^e). \tag{38}$$

[11c] The referee points out that error bounds for integrating functions, all of whose derivatives alternate in sign, have recently been obtained independently by G. A. Baker [Phys. Rev. **161**, 434 (1967)], by a completely different method. Our bounds (33) and (34) are considerably more specific, in that the sign of the error is determined by the sign of a *single* derivative, whereas Baker requires *all* of the derivatives to alternate in sign. Our results can thus be applied to many other cases, such as functions, some of whose derivatives have a definite sign on the positive real axis, or are bounded there, etc.

Before presenting quantitative examples of these results in Sec. V, we can sketch here the qualitative behavior of these error bounds as a function of temperature $T$. For the even approximations, all of the eigenvalues $\xi^e_{jn}$ are positive, since they must lie within the range of integration (ST, p. 108), which is here the positive real axis. Thus the lower bounds (35) approach zero as $T \to 0$. However, for finite $T$, they always remain positive, and never give the useless negative values produced by the usual truncated high-temperature expansion (6), at low temperatures.

For the odd approximations, one may show that there is always one and only one eigenvalue ($\xi^o_{jM}$) at the ground-state energy $E = 0$. Thus the upper bounds tend, as $T \to 0$, to a finite positive value (between 0 and 1) equal to the weight $\rho^o_M(1)$. This finite value $Q^o_X(\infty)$ is the maximum possible fraction of the energy spectrum which can be degenerate with the ground state, and still be consistent with the known (first $2M$) moments of the spectrum.

We may remark finally that it can be shown by combining some theorems from the theory of the moment problem,[9] that the bounds we have obtained are in fact the *most precise* which are possible, based on a given finite number of terms in the high temperature expansion. In particular, we can explicitly construct distributions $\psi(E)$ which have precisely the same high-temperature expansion, through $2M$ terms, but whose exact partition functions lie at any point we may choose within the error bounds, including the upper and lower bounds themselves. These lower bounds are *attained* for the distribution

$$d\psi(E) = \sum_{j=1}^{M} \rho^e_M(j)\delta(E - \xi^e_{jM})$$

and the upper bounds are attained by taking

$$d\psi(E) = \sum_{j=1}^{M} \rho^o_M(j)\delta(E - \xi^o_{jM}).$$

Each of these positive distributions has the correct values of the known moments, and thus each is a counterexample against having more precise error bounds from the given information.

## V. TWO EXAMPLES

In this section we apply these results to two examples for which exact results are known. The first example is a classical particle in a V-shaped potential well:

$$V(x) = \begin{cases} x, & |x| \leq 1 \\ \infty, & |x| > 1. \end{cases}$$

TABLE I. Fractional error in configurational integral for a V-well oscillator, using $\mu_0, \mu_1, \cdots, \mu_{12}$.

| $T^*$ | ($\epsilon$) This work | ($\epsilon$) Kramer's series |
|---|---|---|
| 0.02 | $1.9 \times 10^{-1}$ | ··· |
| 0.03 | $4.2 \times 10^{-2}$ | ··· |
| 0.04 | $1.0 \times 10^{-2}$ | ··· |
| 0.05 | $2.6 \times 10^{-3}$ | ··· |
| 0.06 | $7.6 \times 10^{-4}$ | ··· |
| 0.08 | $8.2 \times 10^{-5}$ | ··· |
| 0.10 | $1.2 \times 10^{-5}$ | ··· |
| 0.12 | $2.2 \times 10^{-6}$ | ··· |
| 0.15 | $2.4 \times 10^{-7}$ | ··· |
| 0.20 | $1.2 \times 10^{-8}$ | $9.9 \times 10^{-2}$ |
| 0.30 | $1.3 \times 10^{-10}$ | $5.2 \times 10^{-4}$ |
| 0.40 | $4.7 \times 10^{-12}$ | $1.3 \times 10^{-5}$ |
| 0.50 | $3.5 \times 10^{-13}$ | $7.6 \times 10^{-7}$ |
| 0.60 | $4.0 \times 10^{-14}$ | $7.6 \times 10^{-8}$ |

The configurational density of states in this case is simply

$$d\psi(V) = \begin{cases} dV, & 0 \leq V \leq 1 \\ 0, & 1 < V. \end{cases}$$

This case thus corresponds to integration with respect to a constant weight over a finite interval, and hence the abscissas $\xi^e_{jn}$ and weights $\rho^e_n(j)$ are just those of Gaussian quadrature. We have checked a number of the $\xi^e_{jn}$ and $\rho^e_n(j)$ generated by our method, with those tabulated[12] for Gaussian quadrature, and obtain agreement to 13 of the 14 figures carried in the calculation (which was carried out on a CDC 6400 computer). This not only checks the method, but also verifies the expected numerical stability of the matrix formulation. The diagonalization was carried out using the $Q$–$R$ algorithm.[13]

The moments in this case are given by

$$\mu_n \equiv \langle V^n \rangle = 1/(n + 1).$$

From these moments the approximate configurational integrals $\bar{Q}_M$ (38) were constructed, along with the error bounds $\epsilon$ in (37).

The exact configuration integral for this V-well model is

$$Q(\beta) = [1 - \exp(-\beta)]/\beta.$$

In all cases tried the $\bar{Q}_M$ converged nicely to this exact function, and every stage stayed within the error bounds. The error bounds decreased smoothly and rapidly as the number of moments was increased,

[12] *Handbook of Mathematical Functions*, National Bureau of Standards Applied Mathematics Series 55 (U.S. Government Printing Office, 1964), p. 916.

[13] J. G. F. Francis, Computer J. **4**, 265, 332 (1961); V. N. Kublanovskaya, Zh. Vych. Mat. **1**, 555 (1961).

TABLE II. Number of moments required to obtain an accuracy of better than 1% in the configurational integral, for a V-well oscillator.

| $T^*$ | This work | Kramer's series |
|---|---|---|
| 0.04 | 12 | 68 |
| 0.05 | 11 | 55 |
| 0.06 | 10 | 46 |
| 0.08 | 9 | 35 |
| 0.10 | 8 | 28 |
| 0.12 | 7 | 24 |
| 0.15 | 6 | 19 |
| 0.20 | 6 | 15 |
| 0.30 | 5 | 10 |
| 0.40 | 4 | 8 |

until the fractional error reached about $10^{-14}$, after which it was lost in the round-off error of the 14 figure calculation.

The usual high-temperature series (6) for this model is

$$Q(\beta) = \sum_{n=0}^{\infty} \frac{(-\beta)^n}{(n+1)!}.$$

In Table I our error bounds are compared to those given by this usual high-temperature series, for a typical case of a 12-term series. Our method produces error bounds which are more than a million times smaller than the usual series, for the same number of terms. Furthermore, the relative advantage of our method *increases* as the number of terms kept in both series increases.

If one asks for results of a given accuracy (say 1%), Table II compares the number of terms which are required to obtain that accuracy, for a V-well oscillator. Considering the enormously increasing difficulty of evaluating higher terms in expansions for realistic statistical problems, the smaller number of terms required for a given accuracy is a great advantage for our method.

The second example we consider is the two-dimensional Ising model, for which we can compare our results with the famous exact solution of Onsager.[14] Of particular interest is the partition function for this model, when the number of sites $N$ becomes infinite. In order to obtain finite moments as $N \to \infty$, we may consider the high-temperature expansion of the partition function per site, which is defined by

$$Q = \lim_{N\to\infty} (Q_N)^{1/N},$$

where $Q_N$ is the partition function for a model with $N$ sites. The high-temperature expansion for $Q$ is independent of $N$, and can be obtained by rearranging the hyperbolic tangent expansion of Kramers and Wannier[3] and Onsager.[14] In general, the $N$th root of a partition function need not be expressible as a Laplace integral (2) with a nondecreasing distribution function. Such a nondecreasing distribution does appear to exist for the two-dimensional Ising model in zero field, since the coefficients $(\alpha_{2n}\alpha_{2n+1})$ turn out to be positive (at least through $n = 8$). These are necessary conditions for a nondecreasing distribution, and sufficient conditions if true for all $n$. However, the distribution of states corresponding to this expansion is symmetric about $E = 0$ and extends to $\pm\infty$, rather than starting at 0 as we assumed in deriving the error bounds. Because of this extension to $-\infty$, the error term in (34) is no longer applicable, and hence no upper bound (36) is furnished in this case. This failure of the upper bound might be expected in this case, since the model exhibits a phase transition at which the heat capacity becomes infinite,[14] so that one could hardly expect to find upper bounds to thermodynamic properties in this case. Nevertheless, the even error term (33) continues to be valid, and yields lower bounds which are remarkably accurate. In Fig. 1 we have plotted the error in these lower bounds, as a function of the number of moments ($2M$) used, for the partition function at the critical temperature $T_c$ of the infinite two-dimensional Ising model. For comparison, we have also plotted the error of the high-temperature (hyperbolic tangent) series of Kramers and Wannier,[3] as a function of the number of terms. The greatly improved accuracy and rate of convergence

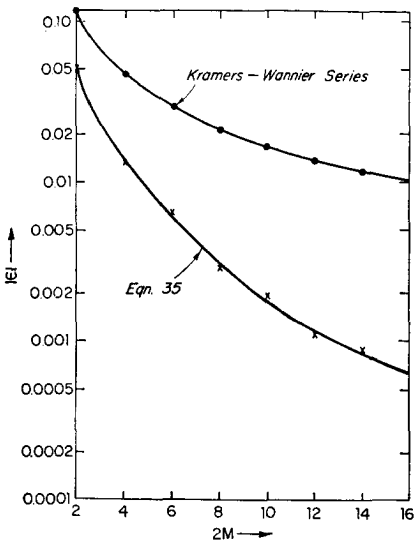



FIG. 1. Fraction error in various approximations to partition function of the 2-dimensional Ising model on a quadratic lattice, at the critical temperature, as a function of the number ($2M$) of moments used in constructing the approximation.

[14] L. Onsager, Phys. Rev. 65, 117 (1944).

TABLE III. Product–difference table for the corresponding continued fraction, including all those terms which can be calculated using $\mu_0$, $\mu_1$, $\mu_2$, $\mu_3$, and $\mu_4$.

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 1 | $\mu_0$ | $\mu_1$ | $(\mu_0\mu_2 - \mu_1^2)$ | $\mu_0(\mu_3\mu_1 - \mu_2^2)$ | $(\mu_0^2\mu_1\mu_3\mu_4 - \mu_0\mu_1^3\mu_4 - \mu_0\mu_1\mu_2^3 - \mu_0^3\mu_1\mu_3^2 + 2\mu_0\mu_1^2\mu_2\mu_3)$ |
| 2 | 0 | $-\mu_1$ | $-\mu_2$ | $-(\mu_0\mu_3 - \mu_2\mu_1)$ | $\mu_0(\mu_2\mu_3 - \mu_1\mu_4)$ | |
| 3 | 0 | $\mu_2$ | $\mu_3$ | $(\mu_0\mu_4 - \mu_1\mu_3)$ | | |
| 4 | 0 | $-\mu_3$ | $-\mu_4$ | | | |
| 5 | 0 | $\mu_4$ | | | | |
| 6 | 0 | | | | | |

of our method is apparent. It is important to emphasize that precisely the same combinational information is required to construct our $2M$th-order approximation, as is required to find $2M$ terms in the usual high-temperature series.


## ACKNOWLEDGMENTS

We wish to thank Walter Nielsen for expert programming of the $Q$–$R$ transform, and John C. Wheeler and David Chandler for discussions of the thermodynamic limit.

This work was supported by the National Science Foundation under Grant GP-6712, and by the Alfred P. Sloan Foundation.


## APPENDIX A: CONSTRUCTION OF THE CORRESPONDING CONTINUED FRACTION

The basic data we use are the moments $\mu_n$, from which we want to construct the coefficients of the corresponding continued fraction (15). We set out the moments as the second column of Table III. The first column is initialized to zero, except for the (1,1) element $P_{11}$ which is taken as unity. The table is then filled up by columns, proceeding to the right from column 3, according to a "product–difference" (PD) recursion relation

$$P_{ij} = P_{1,j-1} \cdot P_{i+1,j-2} - P_{1,j-2} \cdot P_{i+1,j-1}. \quad \text{(A1)}$$

Within each column one starts at the top and works downward. When a triangular portion of the table is complete, then the $\alpha_n$ are given by

$$\alpha_n = P_{1,n+1}/(P_{1,n} \cdot P_{1,n-1}). \quad \text{(A2)}$$

This scheme has many points in common with Rutishauser's quotient–difference (QD) algorithm,[15] which starts from moment *ratios*, rather than the moments themselves, and produces the $\alpha_n$ by alternating quotients and differences. Essentially the same results are obtained by the PD and QD algorithms. The main advantage of the PD method is that it saves all divisions until the end, whereas the QD algorithm may break down during iteration because of trying to divide by zero or a very small number.

Both the PD and QD algorithms are rather sensitive to round-off error, and must be carried out with double precision arithmetic. In this respect, the PD algorithm has an additional advantage, in that the entire recursion (A1) can be carried out in the field of exact integer arithmetic, completely avoiding the round-off error.

The standard formal mathematical method of constructing the $\alpha_n$ is through the Hankel determinants.[9] Since the evaluation of the determinants needed requires $\sim n!$ multiplications and additions, compared to the $\sim n^2$ operations required to construct the PD table, it is clear that the recursion schemes are more suitable for calculation.

## APPENDIX B: PROOF OF EQ. (16)

We wish to show that the first component $x_1$ of the solution to the set of linear equations

$$\begin{aligned}
b_1x_1 \quad &-(a_2)^{\frac{1}{2}}x_2 \quad +0x_3 \quad +0x_4 \quad 0 \quad &\cdots \quad &= -a_1 \\
-(a_2)^{\frac{1}{2}}x_1 \quad &+b_2x_2 \quad -(a_3)^{\frac{1}{2}}x_3 \quad +0x_4 \quad 0 \quad &\cdots \quad &= 0 \\
0x_1 \quad &-(a_3)^{\frac{1}{2}}x_2 \quad +b_3x_3 \quad -(a_4)^{\frac{1}{2}}x_4 \quad 0 \quad &\cdots \quad &= 0 \\
&\cdots \quad &\cdots \quad &\\
&\cdots \quad &-(a_{n-1})^{\frac{1}{2}}x_{n-2} \quad +b_nx_{n-1} \quad -(a_n)^{\frac{1}{2}}x_n \quad &= 0 \\
&\cdots \quad &0x_{n-2} \quad -(a_n)^{\frac{1}{2}}x_{n-1} \quad +b_nx_n \quad &= 0
\end{aligned}$$

[15] H. Rutishauser, *Der Quotienten-Differenzen-Algorithmus* (Birkhäuser, Basel/Stuttgart, 1957); P. Henrici, Proc. Symp. Appl. Math. **15**, 159 (1963).

is equivalent to the value of the fraction

$$f \equiv \cfrac{-a_1}{b_1 - \cfrac{a_2}{b_2 - \cfrac{a_3}{b_3 - \cdots \cfrac{}{\cdots \cfrac{-a_{n-1}}{b_{n-1} - \cfrac{a_n}{b_n}}}}}}.$$

We solve the linear equations by Gaussian elimination, starting at the bottom. From the last equation, we have

$$-(a_n)^{\frac{1}{2}}x_n = (-a_n/b_n)x_{n-1}.$$

Substituting this into the second-to-last equation and solving gives

$$-(a_{n-1})^{\frac{1}{2}}x_{n-1} = \left(\cfrac{-a_{n-1}}{b_{n-1} - \cfrac{a_n}{b_n}}\right)x_{n-2}.$$

Similarly substitution of this into the third equation from the bottom gives

$$-(a_{n-2})^{\frac{1}{2}}x_{n-2} = \left(\cfrac{-a_{n-2}}{b_{n-2} - \cfrac{a_{n-1}}{b_{n-1} - \cfrac{a_n}{b_n}}}\right)x_{n-3}.$$

It is by now clear that each of the factors in parenthesis represent larger and larger portions of the bottom of the fraction $f$. At the last stage,

$$-(a_2)^{\frac{1}{2}}x_2 = \left(\cfrac{-a_2}{b_2 - \cfrac{a_3}{b_3 - \cfrac{a_4}{b_4 - \cdot \cdot \cdot \cfrac{-a_n}{b_n}}}}\right)x_1.$$

Substituting this into the first linear equation, and solving finally for $x_1$, gives the desired result

$$x_1 = \cfrac{-a_1}{b_1 - \cfrac{a_2}{b_2 - \cfrac{a_3}{b_3 - \cdots \cfrac{}{\cdots - \cfrac{a_n}{b_n}}}}}.$$

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 5 MAY 1968

# $O(5)$ Polynomial Bases*

R. T. SHARP† AND S. C. PIEPER
*Department of Physics, University of Illinois, Urbana, Illinois*

(Received 6 September 1967)

Polynomial bases are derived for the irreducible representations of the group $O(5)$. The matrix elements of the infinitesimal generators are given.

## 1. INTRODUCTION

In recent years there has been a great deal of interest by physicists in the classical groups; see, for example, an excellent review by Behrends, Dreitlein, Fronsdal, and Lee[1] (we refer to this paper as BDFL). Attention has been centered to some extent on the special unitary groups as describing the symmetries of strongly interacting particles.

Recently the orthogonal group $O(4)$ has been recognized as useful in classifying families of Regge trajectories.[2] These ideas have also been extended to higher orthogonal groups in order to include internal symmetries on the same footing as space-time properties[3]; in this connection the group $O(5)$ plays a special role as the largest subgroup of $O(6) \approx SU(4)$.

The fact that an $O(n)$ symmetric system subjected to a reflection condition can have $SU(n-1)$ symmetry[4] may lend some additional interest to the orthogonal groups.

For physical applications it is very useful to have

* Supported by the U.S. Office of Naval Research under Contract 00014-67-A-0305-0005.

† On leave from McGill University, Montreal, Canada.

[1] R. E. Behrends, J. Dreitlein, C. Fronsdal, and W. Lee, Rev. Mod. Phys. **34**, 1 (1962).

[2] D. Z. Freedman and J. M. Wang, Phys. Rev. Letters **18**, 863 (1967).

[3] A. Salam and J. Strathdee, Phys. Rev. Letters **19**, 339 (1967).

[4] D. G. Ravenhall, R. T. Sharp, and W. J. Pardee, Phys. Rev. **164**, 1950 (1967).

is equivalent to the value of the fraction

$$f \equiv \cfrac{-a_1}{b_1 - \cfrac{a_2}{b_2 - \cfrac{a_3}{b_3 - \cdots \cfrac{}{\cdots \cfrac{-a_{n-1}}{b_{n-1} - \cfrac{a_n}{b_n}}}}}}.$$

We solve the linear equations by Gaussian elimination, starting at the bottom. From the last equation, we have

$$-(a_n)^{\frac{1}{2}}x_n = (-a_n/b_n)x_{n-1}.$$

Substituting this into the second-to-last equation and solving gives

$$-(a_{n-1})^{\frac{1}{2}}x_{n-1} = \left(\cfrac{-a_{n-1}}{b_{n-1} - \cfrac{a_n}{b_n}}\right)x_{n-2}.$$

Similarly substitution of this into the third equation from the bottom gives

$$-(a_{n-2})^{\frac{1}{2}}x_{n-2} = \left(\cfrac{-a_{n-2}}{b_{n-2} - \cfrac{a_{n-1}}{b_{n-1} - \cfrac{a_n}{b_n}}}\right)x_{n-3}.$$

It is by now clear that each of the factors in parenthesis represent larger and larger portions of the bottom of the fraction $f$. At the last stage,

$$-(a_2)^{\frac{1}{2}}x_2 = \left(\cfrac{-a_2}{b_2 - \cfrac{a_3}{b_3 - \cfrac{a_4}{b_4 - \cdot \cdot \cdot \cfrac{-a_n}{b_n}}}}\right)x_1.$$

Substituting this into the first linear equation, and solving finally for $x_1$, gives the desired result

$$x_1 = \cfrac{-a_1}{b_1 - \cfrac{a_2}{b_2 - \cfrac{a_3}{b_3 - \cdots \cfrac{}{\cdots - \cfrac{a_n}{b_n}}}}}.$$

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 5 MAY 1968

# $O(5)$ Polynomial Bases*

R. T. SHARP† AND S. C. PIEPER
*Department of Physics, University of Illinois, Urbana, Illinois*

(Received 6 September 1967)

Polynomial bases are derived for the irreducible representations of the group $O(5)$. The matrix elements of the infinitesimal generators are given.

## 1. INTRODUCTION

In recent years there has been a great deal of interest by physicists in the classical groups; see, for example, an excellent review by Behrends, Dreitlein, Fronsdal, and Lee[1] (we refer to this paper as BDFL). Attention has been centered to some extent on the special unitary groups as describing the symmetries of strongly interacting particles.

Recently the orthogonal group $O(4)$ has been recognized as useful in classifying families of Regge trajectories.[2] These ideas have also been extended to higher orthogonal groups in order to include internal symmetries on the same footing as space-time properties[3]; in this connection the group $O(5)$ plays a special role as the largest subgroup of $O(6) \approx SU(4)$.

The fact that an $O(n)$ symmetric system subjected to a reflection condition can have $SU(n-1)$ symmetry[4] may lend some additional interest to the orthogonal groups.

For physical applications it is very useful to have

* Supported by the U.S. Office of Naval Research under Contract 00014-67-A-0305-0005.

† On leave from McGill University, Montreal, Canada.

[1] R. E. Behrends, J. Dreitlein, C. Fronsdal, and W. Lee, Rev. Mod. Phys. **34**, 1 (1962).

[2] D. Z. Freedman and J. M. Wang, Phys. Rev. Letters **18**, 863 (1967).

[3] A. Salam and J. Strathdee, Phys. Rev. Letters **19**, 339 (1967).

[4] D. G. Ravenhall, R. T. Sharp, and W. J. Pardee, Phys. Rev. **164**, 1950 (1967).

explicit matrix representations of the group generators; Clebsch–Gordan coefficients for reducing products of representations are also often needed. To calculate these in a simple form, polynomial bases for the irreducible representations are a useful tool.

In this paper we derive polynomial bases for $O(5)$; the generator matrix elements appear as part of the derivation. The Clebsch–Gordan coefficients we leave for the future.

## 2. COMMUTATION RULES AND PHASE CONVENTIONS

The ten (Hermitian) generators of $O(5)$, the group of rotations in five dimensions, may be taken as $J_{ij} = -J_{ji}$; $i \neq j$; $i, j = 1, \cdots, 5$. $J_{ij}$ generates rotations in the $ij$ plane.

The commutation rules for the $J$'s are

$$[J_{ij}, J_{kl}] = i(\delta_{ik}J_{jl} + \delta_{jl}J_{ik} - \delta_{il}J_{jk} - \delta_{jk}J_{il}). \quad (2.1)$$

We define

$$\begin{aligned} S_1 &= \tfrac{1}{2}(J_{23} + J_{14}), & S_2 &= \tfrac{1}{2}(J_{31} + J_{24}), \\ S_3 &= \tfrac{1}{2}(J_{12} + J_{34}), & T_1 &= \tfrac{1}{2}(J_{23} - J_{14}), \\ T_2 &= \tfrac{1}{2}(J_{31} - J_{24}), & T_3 &= \tfrac{1}{2}(J_{12} - J_{34}), \\ U_\pm &= J_{15} \pm iJ_{25}, & V_\pm &= J_{35} \pm iJ_{45}. \end{aligned} \quad (2.2)$$

Then $\bar{S}$, $\bar{T}$ are two commuting spins which generate $O(4)$ transformations.[5] We give the nonzero commutators of the generators $\bar{S}$, $\bar{T}$, $U_\pm$, $V_\pm$, omitting the well-known rules for $S_\pm$, $S_3$ (and $T_\pm$, $T_3$) with each other; $S_\pm$, $T_\pm$ mean $S_1 \pm iS_2$, $T_1 \pm iT_2$:

$$\begin{aligned} [S_3, U_\pm] &= \pm\tfrac{1}{2}U_\pm, & [T_3, U_\pm] &= \pm\tfrac{1}{2}U_\pm, \\ [S_3, V_\pm] &= \pm\tfrac{1}{2}V_\pm, & [T_3, V_\pm] &= \mp\tfrac{1}{2}V_\pm, \\ [S_\pm, U_\mp] &= \pm V_\pm, & [T_\pm, U_\mp] &= \pm V_\mp, \\ [S_\pm, V_\mp] &= \mp U_\pm, & [T_\pm, V_\pm] &= \mp U_\pm, \\ [U_\pm, V_\pm] &= \mp 2S_\pm, & [U_\pm, V_\mp] &= \mp 2T_\pm. \end{aligned} \quad (2.3)$$

All but the last two relations are summarized by the remark that $U_\pm$, $V_\pm$ under $O(4)$ form a $(\frac{1}{2}\,\frac{1}{2})$ quartet of operators $\mathbf{U}^{\frac{1}{2}\frac{1}{2}}_{s_3t_3}$ with

$$\begin{aligned} \mathbf{U}^{\frac{1}{2}\frac{1}{2}}_{\frac{1}{2}\frac{1}{2}} &= -U_+, & \mathbf{U}^{\frac{1}{2}\frac{1}{2}}_{\frac{1}{2}-\frac{1}{2}} &= V_+, \\ \mathbf{U}^{\frac{1}{2}\frac{1}{2}}_{-\frac{1}{2}-\frac{1}{2}} &= U_-, & \mathbf{U}^{\frac{1}{2}\frac{1}{2}}_{-\frac{1}{2}\frac{1}{2}} &= V_-. \end{aligned} \quad (2.4)$$

We use states in which $S_3$, $T_3$ are diagonal. If $S_3$, $T_3$ are plotted as Cartesian coordinates, the other generators move states as shown in the "root diagram," Fig. 1.

A systematic derivation of the root diagram and the corresponding generators is described by BDFL; our generators are related to theirs by $S_\pm = (6)^{\frac{1}{2}}E_{\pm 1}$,

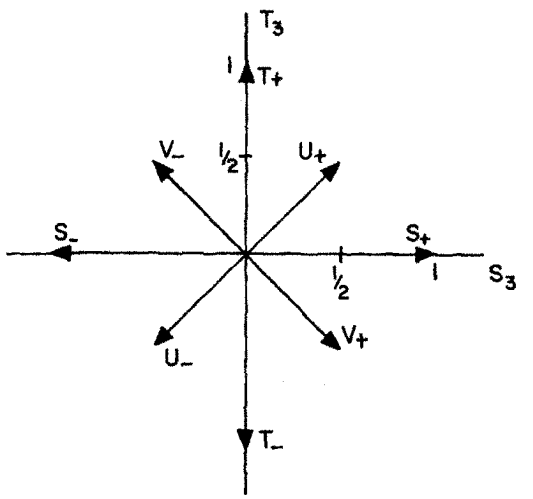



FIG. 1. The 10 generators of $O(5)$; the arrows show the amount by which the eight nondiagonal generators $S_\pm$, $T_\pm$, $U_\pm$, $V_\pm$ move the point $(S_3, T_3)$ whose coordinates are the eigenvalues of the two diagonal generators.

$T_\pm = (6)^{\frac{1}{2}}E_{\pm 3}$, $U_\pm = -2(3)^{\frac{1}{2}}E_{\pm 2}$, $V_\pm = 2(3)^{\frac{1}{2}}E_{\mp 4}$, $S_3 = 2(3)^{\frac{1}{2}}H_1$, $T_3 = 2(3)^{\frac{1}{2}}H_2$.

$O(5)$ representations may be labeled by two nonnegative integers $(p, q)$; they are the $(\lambda_1, \lambda_2)$ of BDFL, or the $(2\mu_2, \mu_1 - \mu_2)$ of Hamermesh.[6] According to BDFL, p. 8, we need $\frac{1}{2}(r - 3l) = 2$ quantum numbers to label states besides the eigenvalues of $S_3$, $T_3$; here $r = 10$ is the number of generators and $l = 2$ is the largest number of mutually commuting generators. We choose to use the pair $S$, $T$ which label $O(4)$ representations. Thus our general state is $|pq; ST; S_3T_3\rangle$; where there is no ambiguity we may suppress the $pq$ labels.

We adopt the Condon–Shortley phase convention for the $S$ and $T$ spin states in each $O(4)$ basis. To define the relative phases of different $O(4)$ bases we invoke the $O(4)$ Wigner–Eckart theorem:

$$\begin{aligned} &\langle S'T'; S_3'T_3'|\, \mathbf{B}^{st}_{s_3t_3} \,|ST; S_3T_3\rangle \\ &\quad = \langle S'T' \| \mathbf{B} \| ST\rangle \langle SS_3ss_3 \,|\, S'S_3'\rangle \langle TT_3tt_3 \,|\, T'T_3'\rangle \\ &\qquad \times [(2S' + 1)(2T' + 1)]^{-\frac{1}{2}}; \end{aligned} \quad (2.5)$$

we use the double bar notation throughout to denote $O(4)$ reduced matrix elements. We now ask that the reduced matrix element $\langle S'T' \| \mathbf{U} \| ST\rangle$ be positive when $\Delta S = \frac{1}{2}$, $\Delta T = \frac{1}{2}$, when $\Delta S = \frac{1}{2}$, $\Delta T = -\frac{1}{2}$, when $\Delta S = -\frac{1}{2}$, $\Delta T = \frac{1}{2}$ and negative when $\Delta S = -\frac{1}{2}$, $\Delta T = -\frac{1}{2}$. Here $\Delta S$, $\Delta T = S' - S$, $T' - T$. It then turns out that all matrix elements of $U_\pm$, $V_\pm$ are positive except those of $U_+$ with $\Delta S = \frac{1}{2}$, $\Delta T = \frac{1}{2}$, of $U_-$ with $\Delta S = -\frac{1}{2}$, $\Delta T = -\frac{1}{2}$, of $V_+$ with $\Delta S = -\frac{1}{2}$, $\Delta T = \frac{1}{2}$, and of $V_-$ with $\Delta S = \frac{1}{2}$, $\Delta T = -\frac{1}{2}$, which are negative.

[5] P. Roman, *Theory of Elementary Particles* (North-Holland Publ. Co., Amsterdam, 1964).

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

### 3. THE REPRESENTATIONS $(p, 0)$ AND $(0, q)$

The basic $(1, 0)$ representation contains two $O(4)$ representations $(\frac{1}{2}\ 0)$ and $(0\ \frac{1}{2})$. The basis states $\alpha = |\frac{1}{2}\ 0;\ \frac{1}{2}\ 0\rangle$, $\beta = |\frac{1}{2}\ 0;\ -\frac{1}{2}\ 0\rangle$, $\gamma = |0\ \frac{1}{2};\ 0\ \frac{1}{2}\rangle$, $\delta = |0\ \frac{1}{2};\ 0\ -\frac{1}{2}\rangle$ are shown in Fig. 2(a). It can be verified that we reproduce the correct commutation rules if we represent the generators by the differential operators

$$\begin{aligned} S_3 &= \tfrac{1}{2}(\alpha\partial_\alpha - \beta\partial_\beta), & T_3 &= \tfrac{1}{2}(\gamma\partial_\gamma - \delta\partial_\delta), \\ S_+ &= \alpha\partial_\beta, & S_- &= \beta\partial_\alpha, \\ T_+ &= \gamma\partial_\delta, & T_- &= \delta\partial_\gamma, \\ U_+ &= \gamma\partial_\beta + \alpha\partial_\delta, & U_- &= \beta\partial_\gamma + \delta\partial_\alpha, \\ V_+ &= \alpha\partial_\gamma - \delta\partial_\beta, & V_- &= \gamma\partial_\alpha - \beta\partial_\delta. \end{aligned} \tag{3.1}$$

The representation $(p, q)$ has dimension

$$\tfrac{1}{6}(p + 1)(q + 1)(p + q + 2)(p + 2q + 3)$$

according to BDFL. For $q = 0$ this is just the number of monomials of degree $p$ in four variables. In fact we see that

$$|p\ 0;\ ST;\ S_3T_3\rangle = \frac{\alpha^{S+S_3}\beta^{S-S_3}\gamma^{T+T_3}\delta^{T-T_3}}{[(S + S_3)!\,(S - S_3)!\,(T + T_3)!\,(T - T_3)!]^{\frac{1}{2}}}. \tag{3.2}$$

For these $(p\ 0)$ states $S$ and $T$ are not independent but

$$T = \tfrac{1}{2}p - S. \tag{3.3}$$

Hence we may without ambiguity suppress $T$ in the state label. $S$ takes integer and half-integer values in the range $0 \leq S \leq p/2$.

The $(0, 1)$ representation contains two $O(4)$ representations $(\frac{1}{2}, \frac{1}{2})$ and $(0, 0)$. The basis states $\eta = |\frac{1}{2}\ \frac{1}{2};\ \frac{1}{2}\ \frac{1}{2}\rangle$, $\xi = |\frac{1}{2}\ \frac{1}{2};\ -\frac{1}{2},\ \frac{1}{2}\rangle$, $\theta = |\frac{1}{2}\ \frac{1}{2};\ \frac{1}{2},\ -\frac{1}{2}\rangle$, $\zeta = |\frac{1}{2}\ \frac{1}{2};\ -\frac{1}{2}\ -\frac{1}{2}\rangle$, $\lambda = |0\ 0,\ 0\ 0\rangle$ are shown in Fig. 2(b). It can be verified that the commutation rules are reproduced if we write

$$\begin{aligned} S_3 &= \tfrac{1}{2}(\eta\partial_\eta + \theta\partial_\theta - \xi\partial_\xi - \zeta\partial_\zeta), \\ T_3 &= \tfrac{1}{2}(\eta\partial_\eta + \xi\partial_\xi - \theta\partial_\theta - \zeta\partial_\zeta), \\ S_+ &= \eta\partial_\xi + \theta\partial_\zeta, & S_- &= \xi\partial_\eta + \zeta\partial_\theta, \\ T_+ &= \eta\partial_\theta + \xi\partial_\zeta, & T_- &= \theta\partial_\eta + \zeta\partial_\xi, \\ U_+ &= (2)^{\frac{1}{2}}(\lambda\partial_\zeta - \eta\partial_\lambda), & U_- &= (2)^{\frac{1}{2}}(\zeta\partial_\lambda - \lambda\partial_\eta), \\ V_+ &= (2)^{\frac{1}{2}}(\theta\partial_\lambda + \lambda\partial_\xi), & V_- &= (2)^{\frac{1}{2}}(\lambda\partial_\theta + \xi\partial_\lambda). \end{aligned} \tag{3.4}$$

Now $\eta\zeta - \xi\theta + \frac{1}{2}\lambda^2$ is an $O(5)$ scalar so to avoid duplication of representations we must discard states proportional to powers of it. The number of monomials of degree $q$ in five variables is

$$\tfrac{1}{24}(q + 1)(q + 2)(q + 3)(q + 4).$$

The number of discarded states is the same, with $q$ replaced by $q - 2$; the remaining number is just the required dimension of the representation $(0, q)$.

For $(0, q)$ states, it turns out that $S = T$, so again we may suppress $T$ from the state label; $S$ takes integer and half-integer values in the range $0 \leq S \leq q/2$. We find

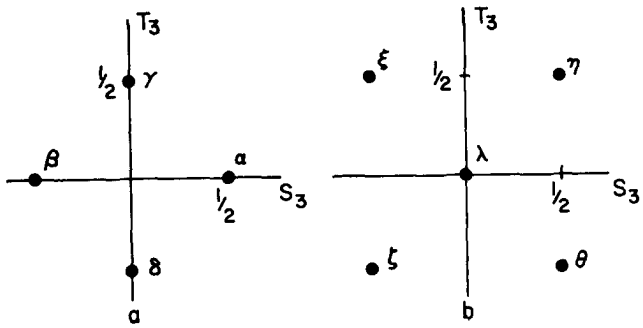



FIG. 2. The two basic representations of $O(5)$; (a) shows the four states of the (1,0) representation; (b) shows the five states of the (0,1) representation.

$$|0\ q;\ S;\ SS\rangle = N_S\eta^{2S}\sum_x \frac{\lambda^{q-2S-2x}(\xi\theta - \eta\zeta)^x}{2^x x!\,(q - 2S - 2x)!\,(x + 2S + 1)!} \tag{3.5}$$

with

$$N_S = \{[(2S + 1)(q - 2S)!\,(q + 2S + 2)! \times (q + 1)!\,2^{q-2S}]/(2q + 2)!\}^{\frac{1}{2}}. \tag{3.6}$$

To derive Eq. (3.5) we first assume $|0\ q;\ S;\ SS\rangle = \eta^{2S}F(\lambda, \eta\zeta - \xi\theta)$; since $\lambda$ and $\eta\zeta - \xi\mu$ are the only $O(4)$ scalars, this is the most general form with the correct $O(4)$ behavior. $F$ is a polynomial of degree $q - 2S$ in $\eta\xi\theta\zeta\lambda$ and may be determined to within a multiplicative constant by the condition

$$(\partial_\eta\partial_\zeta - \partial_\xi\partial_\theta + \tfrac{1}{2}\partial_\lambda^2)\eta^{2S}F = 0. \tag{3.7}$$

The operator in Eq. (3.7) obviously gives zero on the perimeter state $\eta^q/(q!)^{\frac{1}{2}}$; and since it is an $O(5)$ scalar and commutes with all the generators, it must give zero on every state.

The normalization constant (3.6) is fixed by demanding the equality of the matrix elements

$$\langle S + \tfrac{1}{2};\ S + \tfrac{1}{2}S + \tfrac{1}{2}|\ U_+\ |S;\ SS\rangle$$

and

$$\langle S;\ SS|\ U_-\ |S + \tfrac{1}{2};\ S + \tfrac{1}{2}S + \tfrac{1}{2}\rangle.$$

In calculating these matrix elements with the states (3.5), one operates with $U_\pm$ on the right-hand state and identifies the coefficient of the left-hand state; one does not calculate scalar products. In this way all sums are avoided; the calculation is further facilitated by retaining only the $\lambda^{q-2S}$ term on both sides. From the equality of the matrix elements, together with the phase conventions, the ratio $N_S/N_{S+1}$ and the matrix element

$$\langle S;\ SS|\ U_-\ |S + \tfrac{1}{2};\ S + \tfrac{1}{2}S + \tfrac{1}{2}\rangle = -\left[\frac{(2S + 1)(q - 2S)(q + 2S + 3)}{2(S + 1)}\right]^{\frac{1}{2}} \tag{3.8}$$

are determined.

The other states of the $(SS)$ representation of $O(4)$ are easily constructed from the corner state (3.5) by repeated use of the generators $S_-$, $T_-$. They are given [apart from an $O(4)$ scalar factor] by Eq. (2.2) of a paper by one of us.[7]

The states $(0\ q)$ are closely related to the five-dimensional hyperspherical harmonics. In fact, the harmonics are obtained from the states $(0\ q)$, apart from a numerical factor, by the replacement of $\eta\xi\theta\zeta\lambda$ by the corresponding first-degree hyperspherical harmonics.

## 4. THE GENERAL REPRESENTATION $(p, q)$

The $O(4)$ content of the general representation is indicated in Fig. 3. The $ST$ which appear are those in or on the rectangle whose corners are $(\frac{1}{2}p, 0)$, $(\frac{1}{2}p + \frac{1}{2}q, \frac{1}{2}q)$, $(\frac{1}{2}q, \frac{1}{2}p + \frac{1}{2}q)$, $(0, \frac{1}{2}p)$, and for which $2(S - T)$ has the same parity as $p$. Representations with odd $p$ are spinor representations in the sense that the component of angular momentum in any plane in the five-dimensional space (i.e., eigenvalue of any $J_{ij}$) is half-odd; representations with even $p$ have integer angular momentum.

We wish to construct the representation $(p, q)$ as a product of the representations $(p, 0)$ and $(0, q)$:

$$|pq; ST; S_3T_3\rangle = \sum_{S'S''} |pq; S'S''; ST; S_3T_3\rangle\, A_{ST}(S'S''). \tag{4.1}$$

We have adopted the abbreviation

$$\begin{aligned}|pq; S'S''; ST; S_3T_3\rangle &= \sum_{S_3'T_3'} |p\,0; S'; S_3'T_3'\rangle\, |0\,q; S''; S_3 - S_3'T_3 - T_3'\rangle \\ &\quad \times \langle S'S_3'S''S_3 - S_3' \,|\, SS_3\rangle\langle T'T_3'T''T_3 - T_3' \,|\, TT_3\rangle.\end{aligned} \tag{4.2}$$

Here $T' = \frac{1}{2}p - S'$ according to Eq. (3.3) and $T'' = S''$. The coefficients $A(S'S'')$ in Eq. (4.1) are special stretched $O(4)$ scalar factors, i.e., special $O(5)$ Clebsch–Gordan coefficients with the $O(4)$ Clebsch–Gordan coefficients removed. The generators are now the sums of those in Eqs. (3.1) and (3.4).

According to BDFL Table V we can form an $O(5)$ $(1, 0)$ quartet from $(1, 0)$ and $(0, 1)$. In terms of our states it is

$$\begin{aligned}\alpha' &= [\alpha\lambda + (2)^{\frac{1}{2}}(\delta\eta - \lambda\theta)]/(5)^{\frac{1}{2}}, \\ \beta' &= [\beta\lambda + (2)^{\frac{1}{2}}(\delta\xi - \gamma\zeta)]/(5)^{\frac{1}{2}}, \\ \gamma' &= [-\gamma\lambda + (2)^{\frac{1}{2}}(\alpha\xi - \beta\eta)]/(5)^{\frac{1}{2}}, \\ \delta' &= [-\delta\lambda + (2)^{\frac{1}{2}}(\alpha\zeta - \beta\theta)]/(5)^{\frac{1}{2}}.\end{aligned} \tag{4.3}$$

We are concerned with these states because $|pq; ST; S_3T_3\rangle$ must be orthogonal to any state which contains powers of them; such states would belong to representations lower than $(pq)$ in the product of $(p\ 0)$ and $(0\ q)$.

[7] R. T. Sharp, J. Math. & Phys. (to be published).

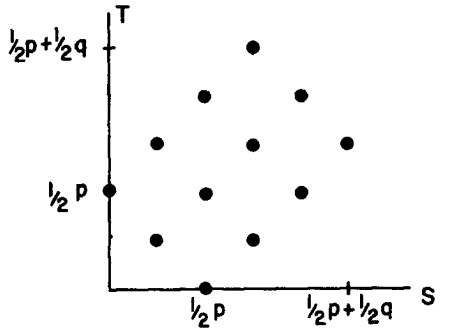



FIG. 3. The $O(4)$ content of the $O(5)$ representation $(pq)$ all $O(4)$ representations $(ST)$ are included (shown by dots) for which the point $(ST)$ lies in or on the rectangle $\frac{1}{2}p \le S + T \le q + \frac{1}{2}p$, $0 \le |S - T| \le \frac{1}{2}p$ and for which $2(S + T)$ has the parity of $p$; the case illustrated is $p = 2$, $q = 3$; the scale is $\frac{1}{2}$ that of Figs. 1 and 2.

Corresponding to the states (4.3) we construct a $(1, 0)$ quartet of operators $\partial_i'$. For example,

$$\partial_\beta' = \partial_\beta\partial_\lambda + (2)^{\frac{1}{2}}(\partial_\delta\partial_\xi - \partial_\gamma\partial_\zeta) \tag{4.4}$$

is the $(\frac{1}{2}\ 0)$ member. We probe by induction that $\partial_i'|x\rangle = 0$ where $|x\rangle$ is any state of the representation $(pq)$. Assume the statement true for $|x\rangle$, and let $G$ be a generator of $O(5)$. Then

$$\partial_i'G|x\rangle = G\partial_i'|x\rangle + [\partial_i', G]|x\rangle = 0$$

since $[\partial_i', G]$ is a linear combination of $\partial_i'$. The statement is true for the heaviest state $\alpha^p\eta^q/(p!\,q!)^{\frac{1}{2}}$ so it is true for all since they can all be reached by repeated application of generators.

We could determine the relative values of the $A(S'S'')$ in the state (4.1) by applying $\partial_\beta'$ and $\partial_\alpha'$ to it and asking that the results vanish. But rather than apply them directly, we notice that $(\partial_\beta', \partial_\alpha')$ is a $(\frac{1}{2}, 0)$ doublet $\partial'$ under $O(4)$. In fact it is the sum of two composite doublets, one $(\partial_\beta\partial_\lambda)$ formed from a $(\frac{1}{2}\ 0)$ doublet in the $p$ variables and a $(0\ 0)$ singlet in the $q$ variables; the other $(\partial_\delta\partial_\xi - \partial_\lambda\partial_\zeta)$ formed from a $(0\ \frac{1}{2})$ doublet in $p$ and a $(\frac{1}{2}\ \frac{1}{2})$ quartet in $q$. The reduced matrix elements can be calculated using the $O(4)$ generalization of Edmonds'[8] Eq. (7.1.5),

$$\begin{aligned}&\langle S_a'T_a'; S_b'T_b'; S'T' \| \mathbf{X}^{st} \| S_aT_a; S_bT_b; ST\rangle \\ &\quad \times \langle S_a'T_a' \| \mathbf{A}^{s_at_a} \| S_aT_a\rangle\langle S_b'T_b' \| \mathbf{B}^{s_bt_b} \| S_bT_b\rangle \\ &\quad \times [(2S + 1)(2S' + 1)(2T + 1) \\ &\quad \times (2T' + 1)(2s + 1)(2t + 1)]^{\frac{1}{2}} \\ &\quad \times X\begin{bmatrix} S_a' & S_a & s_a \\ S_b' & S_b & s_b \\ S' & S & s \end{bmatrix} X\begin{bmatrix} T_a' & T_a & t_a \\ T_b' & T_b & t_b \\ T' & T & t \end{bmatrix},\end{aligned} \tag{4.5}$$

[8] A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton University Press, Princeton, N.J., 1957).

involving $X$ coefficients or $9j$ symbols. We omit Edmonds' $\gamma$'s since in our case the $ST$, etc., labels are sufficient to identify the states. **X** in Eq. (4.5) is a composite $O(4)$ tensor:

$$\mathbf{X}^{st}_{s_3t_3} = \sum_{s_3't_3'} \mathbf{A}^{s_at_a}_{s_3't_3'}\mathbf{B}^{s_bt_b}_{s_3-s_3't_3-t_3'} \times \langle s_a s_3' s_b\, s_3 - s_3' \,|\, s\, s_3\rangle\langle t_a t_3' t_b\, t_3 - t_3' \,|\, t\, t_3\rangle.$$

The condition

$$\begin{aligned} 0 &= \langle p-1\; q-1; S'S''; S \pm \tfrac{1}{2}T \| \partial' \| pq; ST\rangle \\ &= \sum_{S'''S''''} \langle p-1\; q-1; S'S''; S \pm \tfrac{1}{2}T \| \partial' \| pq; S'''S''''; ST\rangle A(S''', S''''), \end{aligned}$$

which holds for all $S'S''$ and for $S \pm \frac{1}{2}$, leads to two independent relations connecting $A(S' + \frac{1}{2}, S'')$, $A(S', S'' - \frac{1}{2})$, $A(S', S'' + \frac{1}{2})$. Solving these gives

$$A(S' + \tfrac{1}{2}, S'' + \tfrac{1}{2}) = \left[\frac{(q - 2S'')(T + T'' - T' + 1)(T - T'' + T')(S'' + 1)2}{(2S'' + 1)(q + 2S'' + 3)(S'' - S + S' + 1)(S'' + S' + S + 2)}\right]^{\frac{1}{2}} A(S', S''). \quad (4.6)$$

and a second equation in which the roles of $S$ and $T$ are interchanged. Iterating these equations gives

$$A_{ST}(S', S'') = N_{ST}\left[\frac{(T + T'' - T')!\,(S + S'' - S')!\,(2S'' + 1)}{(q - 2S'')!\,(T + T' - T'')!\,(S'' + S' - S)!\,(S + S' + S'' + 1)!\,(S + S' - S'')!}\right]^{\frac{1}{2}}$$
$$\times \{[(T + T' + T'' + 1)!\,(T' + T'' - T)!\,(q + 2S'' + 2)!]^{-1}\}^{\frac{1}{2}}, \quad (4.7)$$

where $N_{ST}$ is a phase-normalization factor to be determined. The $S'S''$ in the sum (4.1) are restricted to those for which $0 \le S' \le p/2$, $0 \le S'' \le q/2$ and for which the arguments of the factorials in Eq. (4.7) are non-negative integers, i.e., for which $2(S' \pm S'')$ have the parity of $S$ and satisfy the inequalities $\max(\frac{1}{2}p - T, S) \le S' + S'' \le \frac{1}{2}p + T$, $-S \le S' - S'' \le \min(S, \frac{1}{2}p - T)$.

From the phase conventions and the equality of the matrix elements

$$\langle S + \tfrac{1}{2}\; T + \tfrac{1}{2}; S + \tfrac{1}{2}\; S + \tfrac{1}{2}|\; U_+ \;|ST; ST\rangle = \langle ST; ST|\; U_- \;|S + \tfrac{1}{2}\; S + \tfrac{1}{2}; T + \tfrac{1}{2}\; T + \tfrac{1}{2}\rangle, \quad (4.8)$$

the ratio $N_{ST}/N_{S+1T+1}$ and the matrix elements in question are determined. Again we avoid summations in calculating the matrix elements (4.8) by operating in each case with $U_\pm$ on the right-hand state and picking out the coefficient of the left state, not by taking scalar products. The work is facilitated by working with the term $|\underline{S}' + \frac{1}{2}\; \underline{S}'' + \frac{1}{2}; S + \frac{1}{2}\; T + \frac{1}{2}; S + \frac{1}{2}\; T + \frac{1}{2}\rangle$ in the state $|S + \frac{1}{2}\; T + \frac{1}{2}; S + \frac{1}{2}\; T + \frac{1}{2}\rangle$, where $\underline{S}' = \frac{1}{2}(S - T) + \frac{1}{4}p$, $\underline{S}'' = \frac{1}{2}(S + T) - \frac{1}{4}p$; in this term the $S$ and $T$ spins from the $p$ and $q$ spaces are stretched.

By iteration $N_{ST}$ can be expressed in terms of its value for a boundary state $S + T = q + \frac{1}{2}p$. Its value is

$$N_{ST} = \left[\frac{(q + \frac{1}{2}p - S - T)!\,(q + \frac{1}{2}p + S + T + 2)!\,(\frac{1}{2}p + S + T + 1)!}{(2q + p + 2)!\,(S + T - \frac{1}{2}p)!\,(p + q + 1)!\,(q + 1)!}\right]^{\frac{1}{2}}$$
$$\times [(q + \tfrac{1}{2}p - S + T + 1)!\,(q + \tfrac{1}{2}p + S - T + 1)!\,(\tfrac{1}{2}p - S + T)!\,(\tfrac{1}{2}p + S - T)!\,(2q + 2)!]^{\frac{1}{2}}. \quad (4.9)$$

For the reduced matrix elements of the UV generators we find

$$\begin{aligned}
\langle S + \tfrac{1}{2}\; T + \tfrac{1}{2} \| \mathbf{U} \| ST\rangle &= [(q + \tfrac{1}{2}p + S + T + 3)(\tfrac{1}{2}p + S + T + 2)(q + \tfrac{1}{2}p - S - T)(S + T - \tfrac{1}{2}p + 1)]^{\frac{1}{2}}, \\
\langle S - \tfrac{1}{2}\; T - \tfrac{1}{2} \| \mathbf{U} \| ST\rangle &= -[(q + \tfrac{1}{2}p + S + T + 2)(\tfrac{1}{2}p + S + T + 1)(q + \tfrac{1}{2}p - S - T + 1)(S + T - \tfrac{1}{2}p)]^{\frac{1}{2}}, \\
\langle S + \tfrac{1}{2}\; T - \tfrac{1}{2} \| \mathbf{U} \| ST\rangle &= [(q + \tfrac{1}{2}p + T - S + 1)(q + S - T + \tfrac{1}{2}p + 2)(\tfrac{1}{2}p + T - S)(\tfrac{1}{2}p + S - T + 1)]^{\frac{1}{2}}, \\
\langle S - \tfrac{1}{2}\; T + \tfrac{1}{2} \| \mathbf{U} \| ST\rangle &= [(q + \tfrac{1}{2}p + S - T + 1)(q + \tfrac{1}{2}p - S + T + 2)(\tfrac{1}{2}p + S - T)(\tfrac{1}{2}p - S + T + 1)]^{\frac{1}{2}}.
\end{aligned} \quad (4.10)$$

This completes our results. The basis states are given by Eq. (4.1) with the coefficients $A(S'S'')$ given by Eq. (4.7) and the normalization factor $N_{ST}$ by Eq. (4.9). The matrix element of the generators (other than the trivial ones $\bar{S}$, $\bar{T}$) are given by Eqs. (2.4) and (2.5) in terms of the reduced matrix elements (4.10).

Instead of working with basic (1,0) and (0,1) representations one might use two independent (1,0)'s; then (0, 1) would appear as a composite state. In this way one would need eight independent variables instead of nine. Although we have not pursued this alternative course, we may note an analogy with $SU(3)$ where states may be built from two independent quark (1,0) representations or from a quark and an antiquark (0,1) representation; the second course, to which this paper closely corresponds, involves a simpler isospin structure and leads to simpler formulas when the bases are used, for example, to compute Clebsch–Gordan coefficients,[9] an important consideration in view of the large number of internal summations which arise.

[9] C. K. Chew and R. T. Sharp, Nucl. Phys. **B2**, 697 (1967).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 5 MAY 1968

# Quantum-Mechanical Description of a Brownian Particle

L. DE LA PEÑA-AUERBACH*
*Instituto de Física, Universidad de México, México, D.F.*

E. BRAUN†
*Reactor Centro Nuclear Comisión Nacional de Energía Nuclear, México, D.F.*

AND

L. S. GARCÍA COLÍN
*Facultad de Ciencias, Universidad de México, México, D.F.*

(Received 14 July 1967)

It is shown that the motion of a Brownian particle in the Smoluchowski approximation may be described by a Schrödinger-like equation defining a complex probability amplitude whose norm is the same as the stochastic probability density. Furthermore, the quantum dynamical operators have a physical meaning which arises in a natural way, from the stochastic nature of the process. These operators satisfy the usual commutation relations and thus the uncertainty principle. Here the constant $\hbar$ is replaced by a parameter depending on the characteristics of the system. In particular, the potential-energy operator for a Brownian particle subject to no external forces can be interpreted as a Rayleigh dissipative function.

## I. INTRODUCTION

The idea of giving an alternative interpretation to the quantum theory by essentially assuming, first, that to each of the fundamental particles one may associate a field $\psi$ which is a solution to the ordinary Schrödinger's equation, second, that this field is an average over random fluctuations originating in a subquantum level, and third, that this field exerts a "quantum-mechanical" force on the particle which begins to manifest itself strongly on the atomic level, originated 40 years ago in the work of de Broglie[1] and Madelung[2] and was continued by Bohm *et al.*, about 30 years later.[3] Although some progress has been achieved in clarifying the physical concepts lying behind such ideas, the solution to the problem is far from being satisfactory and, furthermore, complete.

In recent years many authors[4] have revived this field trying to bring up a relationship between stochastic processes and quantum mechanics and as one of us has pointed out,[5] this is feasible in a quite simple way. Following the ideas introduced in paper I,[6] we would like to set forth in a series of papers a different method by means of which one can visualize some of the outstanding features of the general and complex problem posed in the previous paragraph. This paper, which is the second one of the series, is devoted to the study of the possibility of describing the dynamical properties of a stochastic process defined by a simple diffusion or Smoluchowski equation using quantum-mechanical methods. This description is therefore restricted in the sense of the limitations which are intrinsic in the derivation of Smoluchowski's equation,[6] namely, that we shall be considering only time intervals of the particle long compared with its relaxation time. The more general case in which the particle is described by means of a Fokker–Planck equation will be dealt with in a forthcoming paper.

In Sec. II we sketch the model for our discussion and from which stems Schrödinger's equation describing the stochastic process. Section III is devoted, using an adequate language, to a discussion of the physical interpretation of the "potential function" appearing in Schrödinger's equation derived in Sec. II. The most important physical example which can be treated within the context of Smoluchowski's equation, namely, that of a free particle, is given in Sec. IV, and finally in Sec. V a discussion of our results is presented.

## II. DERIVATION OF SCHRÖDINGER'S EQUATION

We describe the motion of a particle in the configuration space through a real single-valued function $\rho(x, t)$ where $x(t)$ is a stochastic process and $\rho$ is the probability density at $x(t)$. Then, we postulate that this probability density is conserved, namely

$$\partial\rho/\partial t + \operatorname{div}(\mathbf{v}\rho) = 0, \tag{1}$$

$\rho$ satisfies a continuity equation. Here, $\mathbf{v}$ is the macroscopic or flow velocity of the particle, which in

*Consultant, Comisión Nacional de Energía Nuclear, México.
† Facultad de Ciencias, Universidad de México.
[1] L. de Broglie, Compt. Rend. **183**, 447 (1926); **184**, 273 (1927); **185**, 380 (1927).
[2] E. Madelung, Z. Physik **40**, 332 (1926).
[3] D. Bohm, *Causality and Chance in Modern Physics* (D. Van Nostrand Co., Inc., Princeton, N.J., 1957), and references there cited.
[4] G. Della Riccia and N. Wiener, J. Math. Phys. **7**, 1372 (1966) and also some other references in Ref. 5.
[5] L. de la Peña-Auerbach, Phys. Letters **24A**, 603 (1967). (Hereafter this paper will be referred to as I.)
[6] N. Wax, Ed., *Selected Papers on Noise and Stochastic Processes* (Dover Publ., Inc., New York, 1954); see especially papers by S. Chandrasekhar and by M. C. Wang and G. E. Uhlenbeck.

general may be written in the following way:

$$v_i = a_i + \frac{1}{\rho}\sum_j \frac{\partial}{\partial x_j}(b_{ij}\rho), \tag{2}$$

where $a_i = K_i/\beta$ is the $i$th component of the external force $\mathbf{K}$, per unit mass, acting on the particle, divided by the friction coefficient $\beta$, and $b_{ij}$ is the diffusion tensor.[6] Because $\rho$ is a real positive single-valued function, we may now introduce a real function $R$ such that

$$\rho = e^{2R}, \tag{3}$$

in terms of which Eq. (1) may be written as follows:

$$\partial R/\partial t = -\tfrac{1}{2}\,\mathrm{div}\,\mathbf{v} - \mathbf{v}\cdot\mathrm{grad}\,R. \tag{4}$$

The question here arises if whether we can obtain an equation of the Schrödinger type defining a field $\psi$ such that the probability density $\psi\psi^* = \rho$ describes the same physical situation as Eq. (1). The answer to this question is dealt with in paper I where it is shown that if we introduce the function

$$\psi = e^{R+iS}, \tag{5}$$

where $S$ is a real, single-valued function, then indeed Eq. (1) may be cast into the form

$$i(\partial\psi/\partial t) = -\tfrac{1}{2}\alpha\nabla^2\psi + V\psi, \tag{6}$$

together with

$$\mathbf{v} = \alpha\,\mathrm{grad}\,S, \tag{7}$$

corresponding to an "irrotational" flow in configuration space. This implies, of course, that the external force is conservative and $\alpha$ is an undetermined parameter characteristic of the system.[7]

Equation (6) is a Schrödinger-like equation with a potential energy function $V$ which in terms of $R$ and $S$ is given by

$$V = -\partial S/\partial t + \tfrac{1}{2}\alpha[\nabla^2 R + (\mathrm{grad}\,R)^2 - (\mathrm{grad}\,S)^2], \tag{8}$$

whose physical interpretation, in the quantum-mechanical sense, is the main subject of this paper, but at this stage remains unknown. Notice should be made, however, that the definition of $\psi$ given by Eq. (5) satisfies the requirement that $\rho = \psi^*\psi$ is a probability density. Therefore, Eq. (6) implies that $\psi$ satisfies a Schrödinger-type equation with a potential-energy function defined by Eq. (8) and its modulus squared gives the probability density for a Markoff process in the Smoluchowski approximation where the flow velocity is irrotational.

[7] A brief discussion concerning the nature of $\alpha$ is given in paper I.

## III. THE PHYSICAL CONTENT OF SCHRÖDINGER'S EQUATION

As was pointed out in the previous section, the core of our discussion is to provide a physical meaning for the potential-energy function $V$ appearing in Eq. (6). To accomplish this task it is convenient to introduce various definitions whose meaning will become clearer as we proceed with the argument.

Let $\hat{f}$ be any operator. We define $\langle\hat{f}\rangle_{\mathrm{av}}$ the mean value of $\hat{f}$ weighted with the probability distribution $\rho$ in the usual way, namely,

$$\langle\hat{f}\rangle_{\mathrm{av}} \equiv \int \hat{f}\rho\, d\mathbf{r}, \tag{9}$$

the integration extending over all the configuration space.

Also, we define the expectation value $\langle\hat{f}\rangle$ of the operator $\hat{f}$ by

$$\langle\hat{f}\rangle \equiv \int \psi^*\hat{f}\psi\, d\mathbf{r}, \tag{10}$$

where, in particular, if $\hat{f}$ is a $c$ function then, trivially, $\langle\hat{f}\rangle = \langle\hat{f}\rangle_{\mathrm{av}} = f$.

From the results of the previous section, it is seen that the probability current $\mathbf{J}$ associated with our Brownian particle is given by

$$\mathbf{J} = \rho\mathbf{v}. \tag{11}$$

On the other hand, if we restrict ourselves, for the sake of simplicity, to consider an isotropic diffusion tensor, $b_{ij} = -\delta_{ij}D$, then Eqs. (2) and (7) yield

$$\mathbf{v} = \mathbf{a} - (D/\rho)\,\mathrm{grad}\,\rho = \alpha\,\mathrm{grad}\,S. \tag{12}$$

Let us now define the operator $\mathbf{v}$ through the equation

$$\hat{\mathbf{v}} \equiv \mathbf{a} - D\,\mathrm{grad} \tag{13}$$

so that

$$\hat{\mathbf{v}}\rho = \mathbf{v}\rho. \tag{14}$$

Finally, in analogy with quantum mechanics the "momentum" operator will be defined by

$$\hat{\mathbf{p}} \equiv -im\alpha\,\mathrm{grad}. \tag{15}$$

Let us now derive some results concerning the mean and expectation values of these operators using the definitions introduced above.

The mean value of the momentum operator $\hat{\mathbf{p}}$ is readily shown to be equal to zero. In fact, substitution of Eq. (15) into Eq. (9) and integration by parts, using the boundary condition that $\rho$ must vanish at infinity, yields immediately the result

$$\langle\hat{\mathbf{p}}\rangle_{\mathrm{av}} = 0. \tag{16}$$

Use of this result in Eq. (13) leads to another interesting equation, namely, that

$$\langle\hat{\mathbf{v}}\rangle_{\mathrm{av}} = \bar{\mathbf{a}}. \tag{17}$$

In other words, the mean value of the flow velocity of the particle is proportional to the mean value of the external force per unit mass acting on it. We stress the fact that this result is a consequence of working with Smoluchowski's equation.

On the other hand, the expectation value of the operator $\hat{\mathbf{p}}$ is given by

$$\langle\hat{\mathbf{p}}\rangle = \int \psi^* \hat{\mathbf{p}} \psi \, d\mathbf{r}. \tag{18}$$

Using Eq. (5), Eq. (18) may be rewritten as follows:

$$\langle\hat{\mathbf{p}}\rangle = -\mathrm{im}\,\alpha \int \psi^* (\mathrm{grad}\, R + i\, \mathrm{grad}\, S)\psi \, d\mathbf{r},$$

and therefore,

$$\langle\hat{\mathbf{p}}\rangle = -im\alpha \langle \mathrm{grad}\, R\rangle + m\alpha \langle \mathrm{grad}\, S\rangle.$$

However, from its definition $\hat{\mathbf{p}}$ is a Hermitian operator, hence

$$\langle \mathrm{grad}\, R\rangle = 0,$$

so that

$$\langle\hat{\mathbf{p}}\rangle = \alpha m \langle \mathrm{grad}\, S\rangle. \tag{19}$$

This result is also quite interesting. Indeed, from Eq. (7) we see that $\langle \mathrm{grad}\, S\rangle = \alpha^{-1}\langle \mathbf{v}\rangle$, so that Eq. (19) reads

$$\langle\hat{\mathbf{p}}\rangle = \langle m\mathbf{v}\rangle = m\bar{\mathbf{v}}, \tag{20}$$

or that the expectation value of the "momentum operator" is equal to the mean value of the flow momentum associated with the particle. It is indeed this result which allows one to interpret $\hat{\mathbf{p}}$ as the "flow momentum" operator associated to the Brownian particle.

Let us now introduce the corresponding statistical deviations from the two average values defined at the beginning of this section. Define two quantities $\delta\hat{f}$ and $\Delta\hat{f}$ through the equations:

$$\delta\hat{f} \equiv \hat{f} - \langle\hat{f}\rangle_{\mathrm{av}}, \tag{21a}$$

$$\Delta\hat{f} \equiv \hat{f} - \langle\hat{f}\rangle, \tag{21b}$$

the two being equal if $\hat{f}$ is a $c$ function. Then, the mean square deviations associated with $\delta\hat{f}$ and $\Delta\hat{f}$ are

$$\langle(\delta\hat{f})^2\rangle_{\mathrm{av}},\ \langle(\delta\hat{f})^2\rangle,\ \langle(\Delta\hat{f})^2\rangle_{\mathrm{av}},\ \text{and}\ \ \langle(\Delta\hat{f})^2\rangle,$$

the four being equal to each other when $\hat{f}$ is a $c$ function. In particular, if we take for $\hat{f}$ the momentum operator we see, using Eqs. (16) and (20), that

$$\delta\hat{\mathbf{p}} = \hat{\mathbf{p}} - \langle\hat{\mathbf{p}}\rangle_{\mathrm{av}} = \hat{\mathbf{p}}$$

and

$$\Delta\hat{\mathbf{p}} = \hat{\mathbf{p}} - \langle m\mathbf{v}\rangle = \hat{\mathbf{p}} - m\langle\mathbf{v}\rangle_{\mathrm{av}}.$$

Then,

$$\langle(\Delta\hat{\mathbf{p}})^2\rangle_{\mathrm{av}} = \langle(\delta\hat{\mathbf{p}})^2\rangle_{\mathrm{av}} + (m\langle\mathbf{v}\rangle_{\mathrm{av}})^2. \tag{22a}$$

Also,

$$\langle(\Delta\hat{\mathbf{p}})^2\rangle = \langle\hat{\mathbf{p}}^2\rangle - (m\langle\mathbf{v}\rangle_{\mathrm{av}})^2. \tag{22b}$$

Moreover, since $\delta\hat{\mathbf{p}} = \hat{\mathbf{p}}$, we see that

$$\langle(\delta\hat{\mathbf{p}})^2\rangle_{\mathrm{av}} = -\alpha^2 m^2 \int \mathrm{div}\,(\mathrm{grad}\,\rho)\, d\mathbf{r},$$

which vanishes due to the boundary conditions obeyed by $\rho$. Hence,

$$\langle(\Delta\hat{\mathbf{p}})^2\rangle_{\mathrm{av}} = (m\langle\mathbf{v}\rangle_{\mathrm{av}})^2. \tag{23}$$

We are now prepared to undertake our task, namely, to arrive at a physical interpretation for the potential function $V$ defined through Eq. (8). Let us then begin by calculating the expectation value of the square of the momentum operator $\hat{\mathbf{p}}$. We have

$$\langle\hat{\mathbf{p}}^2\rangle = -m^2\alpha^2 \int \psi^* \nabla^2 \psi \, d\mathbf{r}$$

$$= -2m^2\alpha \int \psi^* [-i(\partial\psi/\partial t) + V\psi]\, d\mathbf{r},$$

using Eq. (6). The integration is straightforward, yielding the result that

$$\langle\hat{\mathbf{p}}^2\rangle = 2m^2\alpha[\langle i(\partial/\partial t)\rangle - \langle V\rangle] \tag{24a}$$

or

$$\langle m\alpha i(\partial/\partial t)\rangle = \langle\hat{\mathbf{p}}^2/2m\rangle + \langle m\alpha V\rangle. \tag{24b}$$

This equation is quite suggestive itself. Indeed if we define two operators, namely,

$$\hat{E} \equiv im\alpha(\partial/\partial t) \tag{25a}$$

and

$$\hat{U} \equiv m\alpha V, \tag{25b}$$

then Eq. (24b) takes the form of

$$\langle\hat{E}\rangle = \langle\hat{\mathbf{p}}^2/2m\rangle + \langle\hat{U}\rangle, \tag{26}$$

which has the conventional form for the relationship between the expectation values for the total energy $\hat{E}$, the kinetic energy $\hat{\mathbf{p}}^2/2m$, and the potential energy $\hat{U}$. Thus, the operator

$$\hat{H} = (\hat{\mathbf{p}}^2/2m) + \hat{U} \tag{27}$$

could be interpreted as the Hamiltonian operator for the particle. However a certain amount of care has to be exercised when this analogy is carried up to this point. In fact, the term $\hat{\mathbf{p}}^2/2m$ is not the usual kinetic energy, but the kinetic energy associated with the microscopic flow of the particle and thus inherent in it there is a stochastic contribution. Also, $V$ is not the external field but a complicated function whose nature is still unknown, but which of course might contain terms of a purely stochastic character. It is therefore necessary, before assigning any physical meaning to the formal results given by Eqs. (26) and (27), to

study more closely Eq. (26). For this purpose let us retrace our steps to our former equations. From Eqs. (8) and (25b) we see that

$$\hat{U} = -m\alpha(\partial S/\partial t) + \tfrac{1}{2}m\alpha^2[\nabla^2 R + (\text{grad } R)^2 - (\text{grad } S)^2] \quad (28)$$

and let us eliminate the terms containing $S$, expressing them in terms of $R$. Notice first that through Eq. (12) we get that

$$(\text{grad } S)^2 = \mathbf{v}^2/\alpha^2 = \alpha^{-2}[\mathbf{a}^2 + 4D^2(\text{grad } R^2) - 4D\mathbf{a}\cdot(\text{grad } R)],$$

where use has been made of Eq. (3). Substitution of this result back into Eq. (28) yields

$$\hat{U} = -m\alpha(\partial S/\partial t) + \tfrac{1}{2}m\alpha^2\{\nabla^2 R + (4D/\alpha^2)\mathbf{a}\cdot\text{grad } R - (\mathbf{a}^2/\alpha^2) + [1 - (4D^2/\alpha^2)](\text{grad } R)^2\}.$$

At this stage it is convenient to emphasize the fact that $\alpha$ is still an undetermined constant about which nothing has been said. The above equation shows, however, that some simplification results if we set $\alpha = 2D$, so that we shall introduce this value for the constant $\alpha$. The simpler form of the last equation thus reads

$$\hat{U} = -2mD(\partial S/\partial t) + 2mD^2[\nabla^2 R + D^{-1}\mathbf{a}\cdot\text{grad } R - (\mathbf{a}^2/4D^2)]. \quad (29)$$

If we now take the gradient of both terms in this equation and use the fact that

$$\text{grad } (\partial S/\partial t) = (2D)^{-1}\partial\mathbf{a}/\partial t - \text{grad } (\partial R/\partial t), \quad (30)$$

we get for grad $\hat{U}$ the expression

$$\text{grad } \hat{U} = -m(\partial\mathbf{a}/\partial t) + 2mD^2 \text{ grad } [\nabla^2 R + D^{-1}\mathbf{a}\cdot\text{grad } R + D^{-1}(\partial R/\partial t) - (\mathbf{a}^2/4D^2)]. \quad (31)$$

Let us now assume that the external force $\mathbf{K}$ is time independent. Then, since $\mathbf{a} = \mathbf{K}/\beta$, $\partial\mathbf{a}/\partial t = 0$, and Eq. (31) can be readily integrated to give

$$\hat{U} = 2mD[D\nabla^2 R + \mathbf{a}\cdot\text{grad } R + (\partial R/\partial t) - (\mathbf{a}^2/4D) + \varphi(t)], \quad (32)$$

where $\varphi(t)$ is an arbitrary differentiable function of time *only*. Comparison between this equation and Eq. (29) leads immediately to the following relationship, namely

$$-\partial S/\partial t = \partial R/\partial t + \varphi(t). \quad (33)$$

This equation already yields some important results. Indeed, if we calculate the expectation value of the operator $\hat{E}$, we get

$$\langle\hat{E}\rangle = 2mD[i\langle\partial R/\partial t\rangle - \langle\partial S/\partial t\rangle].$$

But $\langle\hat{E}\rangle$, $\partial R/\partial t$, and $\partial S/\partial t$ are all real quantities[8] so that

$$\langle\partial R/\partial t\rangle = \left\langle\frac{\partial R}{\partial t}\right\rangle_{\text{av}} = 0 \quad (34)$$

and therefore,

$$\langle\hat{E}\rangle = -2mD\langle\partial S/\partial t\rangle. \quad (35)$$

Comparison of Eqs. (33) and (35) thus implies that

$$\langle\hat{E}\rangle = 2mD\varphi(t), \quad (36)$$

meaning that $\varphi(t)$ is a function which, multiplied by the constant $2mD$, equals the expectation value of the total energy operator.

Equation (33) together with Eq. (30) leads also to a relationship between $R$ and $S$, namely,

$$R + S = -F(t) + (2D)^{-1}A(\mathbf{r}), \quad (37)$$

where

$$\partial F/\partial t = \varphi(t) \quad \text{and} \quad \text{grad } A(\mathbf{r}) = \mathbf{a}, \quad (38)$$

that is, consistently to what we assumed, the external force is given as the gradient of some scalar function and $F$ is a function whose time rate of change is proportional to the expectation value of the total energy operator.

Returning to our problem of disclosing the nature of $\hat{U}$, let us substitute Eq. (36) back into Eq. (32). This yields

$$\langle\hat{E}\rangle - \hat{U} = -2mD[D\nabla^2 R + \mathbf{a}\cdot\text{grad } R + \partial R/\partial t - (\mathbf{a}^2/4D)]. \quad (39)$$

Making use of Eqs. (3), (4), and (12) to express the first three terms of this equation in terms of $\mathbf{v}$ and $\mathbf{a}$, leads immediately to the following expression:

$$\langle\hat{E}\rangle - \hat{U} = -2mD[(2D)^{-1}(\mathbf{a} - \mathbf{v})^2 - \text{div } \mathbf{v} + \tfrac{1}{2}\text{div } \mathbf{a} - (\mathbf{a}^2/4D)]. \quad (40)$$

If we now take the expectation value of $\langle\hat{E}\rangle - \hat{U}$ and use the fact that

$$\langle\text{div } \mathbf{v}\rangle = -\langle 2\mathbf{v}\cdot\text{grad } R\rangle = -D^{-1}\langle\mathbf{v}\cdot(\mathbf{a} - \mathbf{v})\rangle,$$

where use has been made of Eq. (34), we finally obtain that

$$\langle\hat{E}\rangle = \langle\hat{U}\rangle + \langle\tfrac{1}{2}m\mathbf{v}^2 - \tfrac{1}{2}m(\mathbf{v} - \mathbf{a})^2 + mD \text{ div } (\mathbf{v} - \mathbf{a})\rangle. \quad (41)$$

This equation lends itself to the following interpretation: If we define an effective potential

$$\phi_{\text{eff}} = \hat{U} - \tfrac{1}{2}m(\mathbf{v} - \mathbf{a})^2 + mD \text{ div } (\mathbf{v} - \mathbf{a}), \quad (42)$$

then the expectation value of the total energy equals

[8] $\langle\hat{E}\rangle$ is real because $(i\partial/\partial t)$ is Hermitian.

the expectation value of the flow kinetic energy plus the expectation value of $\phi_{\text{eff}}$. The introduction of this effective potential is, however, somewhat artificial because the last two terms of Eq. (33) correspond to a kinetic energy and not to a potential energy. Indeed from Eqs. (24), (28), and (35) we find that

$$\langle \hat{\mathbf{p}}^2/2m \rangle = -2mD^2\langle \nabla^2 R + (\text{grad } R)^2 - (\text{grad } S)^2 \rangle.$$

Furthermore, if we use Eq. (12) together with the fact that $\alpha = 2D$ to eliminate grad $S$ and the relationship

$$\mathbf{v} = \mathbf{a} - 2D \text{ grad } R$$

to eliminate the $R$-dependent terms from the above expression, we reach the final result that

$$\langle \hat{\mathbf{p}}^2/2m \rangle = \langle \tfrac{1}{2}m\mathbf{v}^2 - \tfrac{1}{2}m(\mathbf{v} - \mathbf{a})^2 + mD \text{ div } (\mathbf{v} - \mathbf{a}) \rangle, \tag{43}$$

so that the last two terms are related to the total kinetic energy of the particle and not to a potential energy. In this context, Eq. (41) has the conventional significance, namely,

$$\langle \hat{E} \rangle = \langle \hat{\mathbf{p}}^2/2m \rangle + \langle \hat{U} \rangle. \tag{44}$$

The discussion leading to Eq. (44) thus shows that the operator $\hat{U}$ can be interpreted as the potential-energy operator for the quantum analog of the Brownian particle although we are not giving the set of rules whereby one can calculate this operator explicitly as a function of $\mathbf{r}$ and $t$, assuming that they exist. A more thorough discussion of this point will be given in a later paper.

## IV. THE BROWNIAN FREE PARTICLE

In this section we would like to illustrate the results obtained in the previous sections by applying them to the free-particle case. By free we mean a Brownian particle subject to no external forces.

Let us study this case starting from Schrödinger's equation given by

$$i\partial\psi/\partial t = -D\nabla^2\psi + V\psi. \tag{45}$$

For the potential $V$ we take its value obtained from $\hat{U}$, dividing by $m\alpha$ after we have set the external force, and therefore $\mathbf{a}$, equal to zero. Hence, from Eq. (28)

$$V = -\partial S/\partial t + D\nabla^2 R, \tag{46}$$

where use has been made of the fact that from Eq. (12)

$$\mathbf{v} = -2D \text{ grad } R = 2D \text{ grad } S \tag{47}$$

and hence the two gradients squared are equal.

Then, Eq. (45) reads

$$i\partial\psi/\partial t = -D\nabla^2\psi + (-\partial S/\partial t + D\nabla^2 R)\psi$$

and using (5) we get

$$\partial R/\partial t = D\nabla^2 R + 2D(\text{grad } R)^2 \tag{48}$$

which is a nonlinear partial differential equation for $R$. This result has been obtained, noticing from Eq. (47) that $\nabla^2 R = -\nabla^2 S$.

Equation (48) yields interesting conclusions regarding the spatial dependence of both $R$ and $S$. In fact, since $\langle \text{grad } R \rangle = \langle \text{grad } R \rangle_{\text{av}} = 0$ we see that $R$ can only be an even function in the variable $r$. But due to the structure of Eq. (48) the highest exponent of $r$ can be two, so that

$$R = lr^2 + n, \tag{49}$$

where $l$ and $n$ are only functions of time. Hence, both $\nabla^2 R$ and $\nabla^2 S$ are functions of time only.

From Eq. (47),

$$\text{div } \mathbf{v} = -2D\nabla^2 R \tag{50}$$

and hence div $\mathbf{v}$ is only a function of time.

These results imply that $V$ is a quadratic function of the velocity. In fact, from Eqs. (4), (32), and (50) we obtain that

$$V = \varphi(t) + 2D\nabla^2 R + 2D(\text{grad } R)^2$$

and since $\nabla^2 R$ is a function of $t$ only, we can choose $l$ in Eq. (49) so that $\varphi(t) + 2D\nabla^2 R = 0$. Hence, using Eq. (47),

$$V = \mathbf{v}^2/2D, \tag{51}$$

which is the assertion we wanted to prove.

Equation (51) leads us to conclude that even in the absence of an external force acting on the particle, its motion in the heat bath is affected by this latter one and that the interaction is proportional to $\mathbf{v}^2$. Hence $V$ may be interpreted as a Rayleigh dissipation function, in this simple case.

Having established Eq. (51), we could follow alternatively one of two roads to accomplish the solution of Schrödinger's equation. First, to solve directly Eq. (48) to find $R$ and then use Eq. (37) with $A = 0$ to find $S$. The second alternative is to use Eq. (51) directly in Eq. (45) and follow the same procedure leading to Eq. (48). This yields a system of two partial differential equations for $R$ and $S$ which is easily solved.

We choose to follow the first method because we already have the structure of the solution to Eq. (48). Also, for simplicity we consider a one-dimensional motion. Substitution of Eq. (49) into Eq. (48) yields

$$l = -(8Dt)^{-1}, \quad n = -\tfrac{1}{4}\ln ct$$

and hence,

$$R = -x^2/8Dt - \tfrac{1}{4}\ln ct. \tag{52}$$

On the other hand, since $2D\nabla^2 R = -\varphi(t)$, we have that

$$\varphi(t) = (2t)^{-1} \tag{53}$$

and therefore,

$$F(t) = \tfrac{1}{2} \ln ct. \tag{54}$$

From Eqs. (52), (54), and the fact that $R + S = -F(t)$, we get

$$S = x^2/8Dt - \tfrac{1}{4} \ln ct. \tag{55}$$

The constant $c$ appearing in Eqs. (52), (54), and (55) must be fixed through the normalization condition which $\rho$ must satisfy, namely,

$$\int_{-\infty}^{+\infty} e^{2R}\, dx = 1.$$

Using Eq. (52) we find that

$$c = (4\pi D)^{-\frac{1}{2}}$$

and thus,

$$\rho = (4\pi Dt)^{-\frac{1}{2}} \exp(-x^2/4Dt). \tag{56}$$

With the explicit form for the probability density $\rho$ we can now calculate all the relevant average values for the case. Indeed, using Eq. (10) we have the well-known Einstein relation, namely,

$$\overline{x^2} = \langle x^2 \rangle = 2Dt. \tag{57}$$

Also, from Eq. (43) and (47) we see that

$$\langle \hat{\mathbf{p}}^2/2m \rangle = mD/2t. \tag{58}$$

These last two equations provide a further interesting analogy with well-known quantum-mechanical results. Indeed, using Eq. (15) defining $\hat{\mathbf{p}}$ we can immediately obtain the commutator

$$[\hat{x}, \hat{p}] = x\hat{p} - \hat{p}x = 2imD \tag{59}$$

and from Eqs. (57) and (58)[9]

$$\langle (\Delta x)^2 \rangle \langle (\Delta \hat{p})^2 \rangle = \langle x^2 \rangle \langle \hat{p}^2 \rangle = 2m^2D^2, \tag{60}$$

which is Heisenberg's uncertainty principle applied to the free Brownian particle. Emphasis should be made on the fact that these results stem out only from the assumptions concerning the stochastic process used to describe the problem and thus are not additional ones put into the theory itself. Also, the results provide a correct description of the motion only for times such that $\beta t \gg 1$, a limitation inherent in the approximate nature of Smoluchowski's equation.[6]

Finally, it is interesting to mention a possible interpretation of the motion of a force-free Brownian particle in terms of the Schrödinger-like equation describing it. Since by Eqs. (47) and (55)

$$V = x/2t,$$

we see that the potential energy $V$ is given by $x^2(8Dt^2)^{-1}$, that is, it corresponds to a harmoniclike motion in the spatial coordinate $x$ but with a force constant dying out to zero as $t^{-2}$ in a way characteristic of the system. Thus, Eq. (6) reads

$$i\partial\psi/\partial t = -D\partial^2\psi/\partial x^2 + (x^2/8Dt^2)\psi, \tag{61}$$

which of course cannot be transformed into an "eigenvalue equation" $\hat{H}\psi = E\psi$ because of the time dependence of $V$. The physical meaning of this fact is that if the particle is initially in a certain state with a given energy, due to the friction forces between it and the heat bath it will lose energy continuously.

Notice also that

$$\langle \hat{U} \rangle = (m/4t^2)\langle x^2 \rangle = mD/2t$$

and

$$\langle \hat{T} \rangle = \langle \hat{p}^2/2m \rangle = mD/2t$$

and hence,

$$\langle \hat{U} \rangle = \langle \hat{T} \rangle = \tfrac{1}{2}\langle \hat{E} \rangle \tag{62}$$

which is the virial theorem for the free Brownian motion.

## V. CONCLUSIONS

Throughout this paper we have shown that a particle of mass $m$ undergoing a Brownian motion described in the Smoluchowski (or static) approximation may be also understood via a Schrödinger-like equation giving a probability amplitude $\psi$ whose norm is equal to the probability density $\rho$ appearing in the corresponding stochastic equation. This picture is valid if the flow velocity $\mathbf{v}$ associated with the particle is irrotational. This restriction merely simplifies our mathematical machinery and could be easily removed, but no further assumptions are introduced. Under these conditions, the following results are obtained: (a) The usual operators which have to be introduced explicitly into the quantum mechanics, such as the momentum of the particle, its energy, etc., appear here in a natural way together with their corresponding physical significance. The usual commutation relations for these operators also follow and are expressed in terms of a parameter depending only on the system. (b) To the potential-energy function appearing in Schrödinger's equation, an operator $\hat{U}$ may be associated which can be interpreted as the "potential-energy" operator arising from the external and frictional forces acting on the particle. However, no prescription is advanced on how to calculate this operator in terms of the coordinates and time. (c) The average kinetic energy of the particle is shown to consist of the ordinary kinetic energy of flow plus an additional term which represents the contribution

[9] The first equality in Eq. (60) follows because $\Delta\hat{p} = \hat{p} - m\langle v \rangle$, but for the free particle, $\langle v \rangle = 0$. Also, $\langle \hat{x} \rangle = 0$.

arising from the stochastic nature of the motion. (d) The Hamiltonian operator $\hat{H}$ defined as the sum of the kinetic-energy operator $\hat{T}$ plus $\hat{U}$ is such that the conventional energy equation is satisfied, namely, the expectation value of $\hat{H}$ equals the sum of the expectation values of $\hat{T}$ and $\hat{U}$. (e) When the method is applied to the case of a Brownian particle subject to no external forces, explicit evaluation of the mean-square deviation for the displacement and momentum shows that they satisfy Heisenberg's uncertainty relationship in terms of a constant characteristic of the system. Furthermore, it is found in this case that the expectation value of $\hat{U}$ can be interpreted as a Rayleigh dissipative function associated with the macroscopic flow. Thus, no eigenvalue equation for $\hat{H}$ can be formulated.

All these conclusions are valid within the approximations inherent in Smoluchowski's equation, namely, that we are considering the particle at times long compared with its relaxation time. The removal of this limitation, as mentioned in the text, shall be dealt with in a forthcoming paper of the series.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 5 MAY 1968

# Green's Theorem and Invariant Transformations*

JOSHUA N. GOLDBERG
*Physics Department, Syracuse University, Syracuse, New York*

(Received 6 October 1967)

Conservation laws are derived with the use of Green's theorem. These are studied specifically for the wave equation. A trivial class of invariant transformations exists which maps solutions into solutions. The mapping consists in the addition of a particular solution to all solutions. Because of the linearity of the wave equation, this sum will again be a solution. It is shown that this class of transformations has the Newman–Penrose constants among its generators. Calculations are carried out explicitly for the scalar wave equation and for Maxwell's equations.

## 1. INTRODUCTION

Newman and Penrose[1,2] have discovered a set of new constants of the motion. In linear field theories an infinite number of such constants of the motion exist, while in the nonlinear theory of general relativity only a finite number occur. An earlier attempt[3] to understand the origin of these constants from the point of view of an invariant transformation ended in failure. That investigation started out from the assumption that the field equations were derivable from a variational principle and hence that Noether's theorem[4] connecting an invariant transformation to a conserved quantity would be valid. The conclusion was that the Newman–Penrose (N-P) constants, as generators of invariant transformations, generate a zero change in the field variables[5] $\bar{\delta}y_A = 0$. This does not give any insight into the N–P constants, except to suggest that they are trivial from the point of view of transformation theory. Therefore, a further effort has been made to understand these constants without relying explicitly on Noether's theorem. It was hoped that an alternate approach to invariant transformations and conservation laws would give additional insight into these unusual constants of the motion.

The principal characteristic of physically interesting differential conservation laws is that there exists a set of quantities $t^\rho$ such that $t^\rho{}_{,\rho} = 0$ whenever a set of field equations are satisfied, but not otherwise. For example, Noether's theorem results in the expression

$$-\bar{\delta}y_A L^A \equiv t^\rho{}_{,\rho}, \tag{1.1}$$

where $\bar{\delta}y_A$ is the invariant transformation and $L^A = 0$ are the field equations for the field variables $y_A$. Clearly, when the field equations are satisfied, i.e., $L^A = 0$, the weak conservation law $t^\rho{}_{,\rho} = 0$ follows.

An alternate approach to obtain an expression similar to (1.1) is suggested by Green's theorem. Green's theorem may be stated generally as follows: If $L^A = 0$ are the field equations to be satisfied by the variables $y_A$, the adjoint system of equation $L^{+A} = 0$ for variables $z_A$ is defined by

$$z_A L^A(y) - y_A L^{+A}(Z) \equiv C^\rho{}_{,\rho}. \tag{1.2}$$

* Research supported in part by the National Science Foundation.

[1] E. T. Newman and R. Penrose, Phys. Rev. Letters 15, 231 (1965).

[2] E. T. Newman and R. Penrose (to be published).

[3] J. N. Goldberg, J. Math. Phys. 8, 2161 (1967).

[4] A. Trautman, in *Gravitation*, L. Witten, Ed. (John Wiley & Sons, Inc., New York, 1962).

[5] In this section the field variables are denoted generally by $y_A$, $A$ ranging from 1–$N$, the number of variables. Greek indices, $\mu, \nu, \rho, \sigma$, etc. range over 0, 1, 2, 3. The signature for the Minkowski metric is chosen to be $-2$. Partial differentiation will be usually denoted by a comma as $\partial\phi/\partial x_\rho = \phi_{,\rho}$. Occasionally we denote partial differentiation as $\partial_r\phi = \partial\phi/\partial r$ or $\partial_\rho\phi = \partial\phi/\partial x^\rho$. The context should make clear what is meant.

arising from the stochastic nature of the motion. (d) The Hamiltonian operator $\hat{H}$ defined as the sum of the kinetic-energy operator $\hat{T}$ plus $\hat{U}$ is such that the conventional energy equation is satisfied, namely, the expectation value of $\hat{H}$ equals the sum of the expectation values of $\hat{T}$ and $\hat{U}$. (e) When the method is applied to the case of a Brownian particle subject to no external forces, explicit evaluation of the mean-square deviation for the displacement and momentum shows that they satisfy Heisenberg's uncertainty relationship in terms of a constant characteristic of the system. Furthermore, it is found in this case that the expectation value of $\hat{U}$ can be interpreted as a Rayleigh dissipative function associated with the macroscopic flow. Thus, no eigenvalue equation for $\hat{H}$ can be formulated.

All these conclusions are valid within the approximations inherent in Smoluchowski's equation, namely, that we are considering the particle at times long compared with its relaxation time. The removal of this limitation, as mentioned in the text, shall be dealt with in a forthcoming paper of the series.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 5 MAY 1968

# Green's Theorem and Invariant Transformations*

JOSHUA N. GOLDBERG
*Physics Department, Syracuse University, Syracuse, New York*

(Received 6 October 1967)

Conservation laws are derived with the use of Green's theorem. These are studied specifically for the wave equation. A trivial class of invariant transformations exists which maps solutions into solutions. The mapping consists in the addition of a particular solution to all solutions. Because of the linearity of the wave equation, this sum will again be a solution. It is shown that this class of transformations has the Newman–Penrose constants among its generators. Calculations are carried out explicitly for the scalar wave equation and for Maxwell's equations.

## 1. INTRODUCTION

Newman and Penrose[1,2] have discovered a set of new constants of the motion. In linear field theories an infinite number of such constants of the motion exist, while in the nonlinear theory of general relativity only a finite number occur. An earlier attempt[3] to understand the origin of these constants from the point of view of an invariant transformation ended in failure. That investigation started out from the assumption that the field equations were derivable from a variational principle and hence that Noether's theorem[4] connecting an invariant transformation to a conserved quantity would be valid. The conclusion was that the Newman–Penrose (N-P) constants, as generators of invariant transformations, generate a zero change in the field variables[5] $\bar{\delta}y_A = 0$. This does not give any insight into the N–P constants, except to suggest that they are trivial from the point of view of transformation theory. Therefore, a further effort has been made to understand these constants without relying explicitly on Noether's theorem. It was hoped that an alternate approach to invariant transformations and conservation laws would give additional insight into these unusual constants of the motion.

The principal characteristic of physically interesting differential conservation laws is that there exists a set of quantities $t^\rho$ such that $t^\rho{}_{,\rho} = 0$ whenever a set of field equations are satisfied, but not otherwise. For example, Noether's theorem results in the expression

$$-\bar{\delta}y_A L^A \equiv t^\rho{}_{,\rho}, \tag{1.1}$$

where $\bar{\delta}y_A$ is the invariant transformation and $L^A = 0$ are the field equations for the field variables $y_A$. Clearly, when the field equations are satisfied, i.e., $L^A = 0$, the weak conservation law $t^\rho{}_{,\rho} = 0$ follows.

An alternate approach to obtain an expression similar to (1.1) is suggested by Green's theorem. Green's theorem may be stated generally as follows: If $L^A = 0$ are the field equations to be satisfied by the variables $y_A$, the adjoint system of equation $L^{+A} = 0$ for variables $z_A$ is defined by

$$z_A L^A(y) - y_A L^{+A}(Z) \equiv C^\rho{}_{,\rho}. \tag{1.2}$$

* Research supported in part by the National Science Foundation.

[1] E. T. Newman and R. Penrose, Phys. Rev. Letters 15, 231 (1965).

[2] E. T. Newman and R. Penrose (to be published).

[3] J. N. Goldberg, J. Math. Phys. 8, 2161 (1967).

[4] A. Trautman, in *Gravitation*, L. Witten, Ed. (John Wiley & Sons, Inc., New York, 1962).

[5] In this section the field variables are denoted generally by $y_A$, $A$ ranging from 1–$N$, the number of variables. Greek indices, $\mu, \nu, \rho, \sigma$, etc. range over 0, 1, 2, 3. The signature for the Minkowski metric is chosen to be $-2$. Partial differentiation will be usually denoted by a comma as $\partial\phi/\partial x_\rho = \phi_{,\rho}$. Occasionally we denote partial differentiation as $\partial_r\phi = \partial\phi/\partial r$ or $\partial_\rho\phi = \partial\phi/\partial x^\rho$. The context should make clear what is meant.

Suppose $z_A$ can be chosen such that

$$y_A L^{+A}(Z) \equiv Q^\rho{}_{,\rho}. \tag{1.3}$$

Then by defining

$$t^\rho \overset{\text{def}}{=} C^\rho + Q^\rho, \tag{1.4}$$

Eq. (1.2) is reduced to an expression of the form of (1.1). If $z_A$ can be identified with $-\bar{\delta}y_A$, a correspondence with an invariant, transformation results. For linear equations this identification is possible only if the system of equations $L^A = 0$ is self-adjoint, i.e., $L^A \equiv L^{+A}$. For nonlinear equations, the adjoint equations can be defined if we allow them to depend on the $y_A$ as well as the $z_A$. However, in that case the significance is not clear.

In Sec. 2 this approach will be applied to the scalar field as an illustration. The usual transformations giving energy–momentum conservation will be shown and finally a transformation giving the N–P constants will be written down. In order to illustrate the technique applicable to spinor equations, Maxwell's equations are treated in Sec. 3. All of this work is carried out in Minkowski space. Section 4 presents a summing up together with a brief discussion of what to expect in a curved space–time.

## 2. THE MASSLESS SCALAR FIELD

### A. Invariant Transformations

The massless scalar field $\phi(x)$ satisfies the linear field equation

$$\eta^{\mu\nu}\phi_{,\mu\nu} = 0, \tag{2.1}$$

where the metric tensor is defined by

$$\begin{aligned} ds^2 &= \eta_{\mu\nu}\, dx^\mu\, dx^\nu \\ &= (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2. \end{aligned}$$

Green's theorem takes the form

$$\psi\eta^{\mu\nu}\phi_{,\mu\nu} - \phi\eta^{\mu\nu}\psi_{,\mu\nu} \equiv (\psi\phi_{,}{}^\nu - \phi\psi_{,}{}^\nu)_{,\nu}\,. \tag{2.2}$$

The scalar wave equation is clearly self-adjoint. Comparison with (1.2) identifies

$$C^\nu \overset{\text{def}}{=} \psi\phi_{,}{}^\nu - \phi\psi_{,}{}^\nu. \tag{2.3}$$

We can divide our further consideration into two cases:

(i) $\phi\psi_{,\mu\nu}\eta^{\mu\nu} \equiv Q^\nu{}_{,\nu}$,

but

$$\phi_{,}{}^\rho{}_{,\rho} = 0 \Rightarrow \psi_{,}{}^\rho{}_{,\rho} = 0;$$

(ii) $\psi^\mu{}_{,\mu} = 0$ independent of $\phi$.

This implies that an expression of the form (1.1) results with $Q^\rho = 0$ even when $\phi$ is not a solution of the scalar wave equation.

*Case i:* Under a Lorentz transformation,

$$\begin{aligned} &\delta x^\rho = a^\rho + \omega^\rho_\sigma x^\sigma, \quad (\omega_{\rho\sigma} + \omega_{\sigma\rho} = 0), \\ &\text{and} \\ &\delta\phi = -\phi_{,\rho}\delta x^\rho. \end{aligned} \tag{2.4}$$

Let $\psi = -\bar{\delta}\phi$; then one easily finds

$$\begin{aligned} &\phi\eta^{\mu\nu}\psi_{,\mu\nu} \equiv Q^\nu{}_{,\nu}, \\ &Q^\nu \overset{\text{def}}{=} \phi(\phi_{,\rho}\delta x^\rho)_{,}{}^\nu - \tfrac{1}{2}\phi_{,\rho}\phi_{,}{}^{\rho}\delta x^\nu. \end{aligned} \tag{2.5}$$

With this choice for $\psi$, (2.2) takes the form

$$-\bar{\delta}\phi\eta^{\mu\nu}\phi_{,\mu\nu} = t^\nu{}_{,\nu}; \quad t^\nu \overset{\text{def}}{=} t_\mu{}^\nu a^\mu + \tfrac{1}{2}M^{\nu\rho\sigma}\omega_{\rho\sigma}, \tag{2.6}$$

where $t_\mu{}^\nu$ and $M^{\nu\rho\sigma}$ are the canonical energy–momentum tensor and angular-momentum tensor, respectively,

$$\begin{aligned} t_\mu{}^\nu &\overset{\text{def}}{=} \phi_{,\mu}\phi_{,}{}^\nu - \tfrac{1}{2}\delta_\mu{}^\nu\eta^{\rho\sigma}\phi_{,\rho}\phi_{,\sigma}, \\ M^{\nu\rho\sigma} &\overset{\text{def}}{=} t^{\nu\rho}x^\sigma - t^{\nu\sigma}x^\rho. \end{aligned} \tag{2.7}$$

Equations (2.4)–(2.7) show that in this manner we have recovered the results usually obtained through Noether's theorem.

*Case ii:* Because the scalar wave equation is self-adjoint, $\phi' = \phi + \psi$ will be a solution of

$$\eta^{\mu\nu}\phi'_{,\mu\nu} = 0, \tag{2.8}$$

if and only if $\phi$ is a solution of (2.1). Thus (2.8) represents an invariant transformation in that, solutions are mapped into solutions while the field equations are unchanged in form. This transformation may appear to be trivial, but we shall see shortly that it contains the transformations which generate the N–P constants.

Consider the elementary solution

$$\psi(x) = D(x - x'), \tag{2.9}$$

where $D(x)$ is the free-field propagator which satisfies the initial conditions

$$D(x)|_{x^0=0} = 0, \tag{2.9'}$$

$$\frac{\partial}{\partial x^0} D(x)|_{x^0=0} = -\delta(\mathbf{x}). \tag{2.9''}$$

In (2.9) $D(x - x')$ is to be considered as a generalized function of $x$ while $x'$ is a parameter which identifies a particular one of the set of such functions. Thus,

$$\bar{\delta}\phi(x) = -D(x - x') \tag{2.10}$$

is an invariant transformation in the sense of Case (ii). Note that since $x'$ is merely a parameter

$$\psi(x) = D(x - x')f(x'), \tag{2.11}$$

an arbitrary continuous function $f(x')$ also satisfies the condition of being an invariant transformation of

$\phi(x)$. Similarly

$$\psi(x) = \int_{R'_4} D(x - x')f(x')\, d^4x' \qquad (2.12)$$

is a linear combination of such solutions and therefore is itself an invariant change in $\phi(x)$. In fact, this clearly represents the most general transformation of Case (ii). Therefore, by studying (2.10) we can hope to learn the significance of these transformations.

### B. Newman–Penrose Constants

The generator of an invariant transformation is the constant of the motion resulting from that transformation. To determine the generator when $\psi(x)$ is chosen as in (2.9), consider a region $R_4$ bounded by two spacelike surfaces $\sigma_1$, and $\sigma_2$ (Fig. 1) and let $x'$ lie in $\sigma_2$. By application of Stokes' integral theorem one finds, when $\phi = 0$,

$$\int_{R_4} t^\rho{}_{,\rho}\, d^4x = \int_{\sigma_2} t^\rho n_\rho\, d\sigma - \int_{\sigma_1} t^\rho n_\rho\, d\sigma = 0, \qquad (2.13)$$

and where $n_\rho$ is the future-pointing normal to $\sigma_1$ and $\sigma_2$. The constant of the motion

$$C = \int_{\sigma_1} t^\rho n_\rho\, d\sigma \qquad (2.14)$$

is the generator sought for. Its value can be determined by evaluating the integral over $\sigma_2$:

$$C = \phi(x'). \qquad (2.15)$$

This result is precisely what we should have expected, since from quantum field theory we know

$$\delta\phi = -i[\phi(x), \phi(x')] = D(x - x').$$

Clearly, choosing $\psi(x)$ as in (2.11) merely gives us

$$C = \phi(x')f(x'). \qquad (2.16)$$

The integral (2.14) is not suitable for the study of N–P constants. In the limit of $x'^0 \to +\infty$, the support of the integrand in Eq. (2.14) moves out to spacelike infinity. Actually we are interested in integrals taken in the limit of null infinity. To achieve this result, we

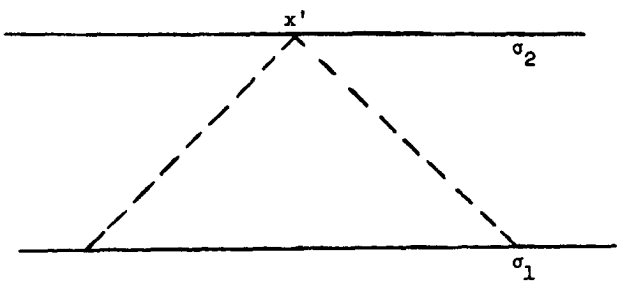



FIG. 1. The region of integration $R_4$ is bounded by two spacelike surfaces extending out to spacelike infinity. The dotted lines indicate the support for the generalized function of $x$, $D(x - x')$.

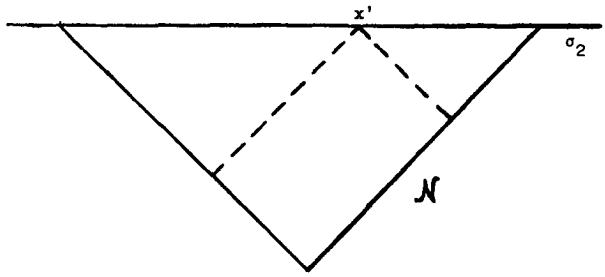



FIG. 2. The region of integration $R_4$ is now bounded by the null cone and a spacelike surface $\sigma_2$. The dotted lines again indicate the support of the generalized function of $x$, $D(x - x')$.

distort $\sigma_1$ into an outgoing null cone $\mathcal{N}$. The region $R_4$ in (2.13) is then bounded by $\mathcal{N}$ and $\sigma_2$ as in Fig. 2. We find

$$\phi(x') = \int_{\mathcal{N}} [2\phi_{,\rho}l^\rho + \phi l^\rho{}_{,\rho}]D(x - x')\, d\tau_{(3)}, \qquad (2.17)$$

where $l^\rho$ is the future-pointing null vector lying in $\mathcal{N}$. Now taking the limit $x'^0 \to +\infty$, the support of $D(x - x')$ on $\mathcal{N}$ moves out to null infinity.

To complete the discussion of the N–P constants, it is convenient to introduce spherical coordinates $(r, \theta, \phi)$ on $\mathcal{N}$ such that the equation for $S$, the intersection of the support of $D(x - x')$ with $\mathcal{N}$, is given by $r = r'$. The parameter $r$ is an affine parameter along the null rays: $l^\rho = \partial x^\rho/\partial r$ and $l^\rho{}_{,\rho} = 2/r$. Equation (2.17) then becomes (the prime is now dropped from $r'$)

$$\phi(x') = \int_S \frac{1}{r}\left[\frac{\partial}{\partial r}\phi + \frac{1}{r}\phi\right] r^2 \sin\theta\, d\theta\, d\phi \qquad (2.18)$$

which can be recognized as the Kirchhoff integral theorem.

In discussing the behavior of free fields at null infinity Newman and Penrose assume[2]

$$\phi = \sum_{j=0}^{l+1} \frac{\phi^j}{r^{j+1}} + O\left(\frac{1}{r^{l+2}}\right). \qquad (2.19)$$

Now, as $x'^0 \to +\infty$, $S \to$ null infinity. Therefore, the integral in (2.18) $\sim 1/r^2$ and

$$\lim_{x'^0 \to \infty} \phi(x') = 0.$$

However, if we had chosen $\psi(x)$ as in (2.11) with $(x'^0)^2 = f(x')$, we would have obtained

$$\lim_{x'^0 \to \infty} (x'^0)^2\phi(x') = \lim_{r \to \infty} 4\int_S r^2\partial_r(r\phi) \sin\theta\, d\theta\, d\varphi. \qquad (2.20)$$

When this limit exists, it is the first of the N–P constants for the massless scalar field. From (2.19) and (2.20) we find

$$C = -4\int \phi^1 \sin\theta\, d\phi.$$

Note that the invariant transformation is

$$\bar{\delta}\phi(x) = -(x'^0)^2 D(x - x'),$$

and in the limit $x'^0 \to \infty$, $\bar{\delta}\phi$ differs from zero only at null infinity. At every point $\bar{\delta}\phi = 0$. This result not only agrees with the earlier calculations,[3] but explains it. The fact is that the N–P constants are surface integrals at infinity not because they come from a strong conservation law, but because the support of the invariant change in $\phi$ is limited to a null cone, in particular, the retrograde null cone at infinity.

To obtain the remaining N–P constants we require generalization of the transformations allowed by (2.12). Consider

$$\psi(x) = f(x')^{\rho_1 \cdots \rho_n} \partial_{\rho_1'} \cdots \partial_{\rho_n'} D(x - x'). \quad (2.21)$$

We can construct $(2n + 1)$ additional constants for each $n \leq l$ where $l$ is defined by the summation in Eq. (2.19). To obtain an infinite number of constants, the solution would have to be analytic at null infinity.

### 3. ELECTRICITY AND MAGNETISM

Maxwell's equations in empty space take the form

$$\begin{aligned} &F^{\mu\nu}{}_{,\nu} = 0; \quad F_{[\mu\nu,\rho]} = 0 \Leftrightarrow F^{*\mu\nu}{}_{,\nu} = 0, \\ &F^{*\mu\nu} = \tfrac{1}{2}\epsilon^{\mu\nu\rho\sigma} F_{\rho\sigma}, \quad \epsilon_{0123} = -\epsilon^{0123} = 1. \end{aligned} \quad (3.1)$$

The square brackets around indices indicates complete antisymmetrization of the indices; $\epsilon^{\mu\nu\rho\sigma}$ is a tensor constructed from the permutation symbol. Operating on the cyclic equations with $\partial^\rho$ and then applying the first of Eqs. (3.1), we obtain

$$F^{\mu\nu}{}_{,\nu}{}^{,\rho} = 0. \quad (3.2)$$

To assure that a solution of (3.2) be a solution of (3.1), it is necessary and sufficient that (3.1) be satisfied on an initial spacelike hypersurface. Green's theorem for (3.2) now reads

$$\psi_{\mu\nu} F^{\mu\nu}{}_{,}{}^{,\rho} - F^{\mu\nu} \psi_{\mu\nu,}{}^{,\rho} \equiv (\psi_{\mu\nu} F^{\mu\nu}{}_{,}{}^{,\rho} - F^{\mu\nu} \psi_{\mu\nu,}{}^{,\rho})_{,\rho}. \quad (3.3)$$

Just as for the scalar field, there are two cases here. The first, arising from the Lorentz transformation is of no further interest to us now. The second follows from the requirement

$$\psi_{\mu\nu,}{}^{,\rho}{}_{,\rho} = 0. \quad (3.4)$$

Choosing $\bar{\delta}F_{\mu\nu} = -\psi_{\mu\nu}$ we see that $\psi_{\mu\nu}$ maps solutions of (3.2) into other solutions of the tensor wave equation. Thus, even if $F_{\mu\nu}$ was originally a solution of Maxwell's equations, $F^{\mu\nu} - \psi^{\mu\nu}$ will, in general, not be a solution. The transformation we are discussing, then, is not an invariant transformation of Maxwell's equations, but of a somewhat more general system of equations, the tensor wave equation. Nonetheless, if the more general system (3.2) possesses N–P constants, we can also expect the more restrictive class, solutions of (3.1), to possess them unless the restrictions are such as to exclude them. Indeed, since the N–P constants are known to exist for the Maxwell field, we know *a priori* that the restrictions of the first-order equations (3.1) do not exclude them entirely.

Because we are dealing with the Maxwell equations we can understand the nature of the restrictions imposed by the first-order equations. The more general system of equations (3.2) not only does not require charge conservation, and therefore may have solutions which exhibit monopole radiation, but each component separately behaves like a scalar field. Hence there are six states of polarization possible which, of course, include monopole radiation. As a result, the solutions of (3.2) will in general have six sets of N–P constants. Of all the solutions of (3.2), those which are solutions of Maxwell's equations certainly do not permit monopole radiation (charge is conserved), but even further, have only two transverse states of polarization. Therefore, we expect at most two independent sets of N–P constants.

To see this point we have to follow the argument of the previous section. Choose

$$\psi_{\mu\nu}(x) = \delta^{\rho'\sigma'}_{\mu\nu} D(x - x'); \quad (3.5)$$

$\delta^{\rho'\sigma'}_{\mu\nu} = \delta^{\rho'}_{\mu}\delta^{\sigma'}_{\nu} - \delta^{\sigma'}_{\mu}\delta^{\rho'}_{\nu}$ as a function of $x$ is to be considered as a constant tensor of rank 2. The indices $\rho'$ and $\sigma'$ are labels which are associated with $x'$. Therefore, $\delta^{\rho'\sigma'}_{\mu\nu}$ may also be considered a constant tensor with respect to $x'$. This identification is reinforced by the primes on the relevant indices. When there is no confusion, the primes will be dropped as on the left-hand side of Eq. (3.6) below. We find then that

$$F^{\rho\sigma}(x') = \int_{\mathcal{N}} \delta^{\rho'\sigma'}_{\mu\nu} \{2F^{\mu\nu}{}_{,\rho} l^\rho + F^{\mu\nu} l^\rho{}_{,\rho}\} D(x - x')\, d\tau_{(3)}, \quad (3.6)$$

where $l^\rho$ is the null vector lying in the surface $\mathcal{N}$ as in Eq. (2.17). Thus $F^{\rho\sigma}(x')$ is the generator of the invariant transformation

$$\bar{\delta}F_{\mu\nu} = -\delta^{\rho'\sigma'}_{\mu\nu} D(x - x'),$$

for the wave equations (3.2).

In order to reduce the integrand in (3.6) to the independent data for solutions of the Maxwell equations, it is convenient to introduce a null tetrad and the corresponding tetrad components of the field. Choose $l^\rho$ and $r$ as in (2.18). Let $n^\rho$ be the corresponding null vector lying in the retrograde null cone from $x'$. Scale $n^\rho$ so that on $S$, $l^\rho n_\rho = 1$. Then define $n_\rho$ on $\mathcal{N}$

by parallel transport along $l^\rho$: $n_{\rho,\sigma}l^\sigma = 0$. Introduce two orthogonal spacelike unit vectors in $S$: $a^\mu a_\mu = b^\mu b_\mu = a^\mu b_\mu - 1 = -1$. Define the complex null vectors

$$\sqrt{2}m^\mu = a^\mu + ib^\mu. \tag{3.7}$$

Define $m^\mu$ over $\mathcal{N}$ by parallel transport: $m^\mu{}_{,\rho}l^\rho = 0$. Thus, at each point of $\mathcal{N}$ we have defined a null tetrad $(l^\mu, n^\mu, m^\mu, \bar{m}^\mu)$ with the properties

$$\begin{gathered} l^\mu l_\mu = n^\mu n_\mu = m^\mu m_\mu = l^\mu m_\mu = n^\mu m_\mu = 0, \\ l^\mu n_\mu = -m^\mu \bar{m}_\mu = 1. \end{gathered} \tag{3.8}$$

Next we can define the bivectors

$$\begin{aligned} V^{\mu\nu} &= l^\mu m^\nu - l^\nu m^\mu, \\ M^{\mu\nu} &= l^\mu n^\nu - l^\nu n^\mu - m^\mu \bar{m}^\nu + m^\nu \bar{m}^\mu, \\ U^{\mu\nu} &= n^\mu \bar{m}^\nu - n^\nu \bar{m}^\mu, \end{aligned} \tag{3.9}$$

and the complex-conjugate bivectors. Since the bivectors written explicitly above satisfy the relations

$$\begin{aligned} V^{*\mu\nu} &= -iV^{\mu\nu}, \\ M^{*\mu\nu} &= -iM^{\mu\nu}, \\ U^{*\mu\nu} &= -iU^{\mu\nu}, \end{aligned} \tag{3.10}$$

they form a basis on $\mathcal{N}$ for the expansion of

$$\begin{aligned} F^{(-)\mu\nu} &\overset{\text{def}}{=} \tfrac{1}{2}(F^{\mu\nu} + iF^{*\mu\nu}) \\ &\overset{\text{def}}{=} \phi_2 V^{\mu\nu} - \phi_1 M^{\mu\nu} - \phi_0 U^{\mu\nu}. \end{aligned} \tag{3.11}$$

Because of the linearity of all of the operations involved, Eq. (3.6) can be rewritten for $F^{(-)\rho\sigma}$. Then using (3.11) we have

$$\begin{aligned} F^{(-)\rho\sigma}(x') = \int_{\mathcal{N}} \delta^{\rho'\sigma'}_{\mu\nu} \{ & V^{\mu\nu}(2\phi_{2,\kappa}l^\kappa + \phi_2 l^\kappa{}_{,\kappa}) \\ & - M^{\mu\nu}(2\phi_{1,\kappa}l^\kappa + \phi_1 l^\kappa{}_{,\kappa}) \\ & - U^{\mu\nu}(2\phi_{0,\rho} + \phi_0 l^\rho{}_{,\rho})\} D(x - x')\, d\tau_{(3)}. \end{aligned} \tag{3.12}$$

On $\mathcal{N}$ we find from Maxwell's equations (3.1)

$$\begin{aligned} V^{\mu\nu}(2\phi_{2,\kappa}l^\kappa + \phi_2 l^\kappa{}_{,\kappa}) &= -2(V^{\mu\nu}\bar{m}^\kappa\phi_1)_{,\kappa} - M^{\mu\nu}\phi_1 l^\kappa{}_{,\kappa}, \\ M^{\mu\nu}(\phi_{1,\kappa}l^\kappa + \phi_1 l^\kappa{}_{,\kappa}) &= -(M^{\mu\nu}\bar{m}^\kappa\phi_0)_{,\kappa} + U^{\mu\nu}\phi_0 l^\kappa{}_{,\kappa}. \end{aligned}$$

Thus, we finally arrive at

$$\begin{aligned} &F^{(-)\rho\sigma}(x') \\ &\quad = \int_{\mathcal{N}} \delta^{\rho'\sigma'}_{\mu\nu} U^{\mu\nu}\{2\phi_{0,\kappa}l^\kappa + 3l^\kappa{}_{,\kappa}\phi_0\} D(x - x')\, d\tau_{(3)}, \end{aligned} \tag{3.13}$$

where we have used the fact that $D(x - x')$ is constant on $S$. From (3.13) we see then that specifying one complex function $\phi_0$ on $\mathcal{N}$ (two real functions) a solution of the homogeneous Maxwell equations is determined. It is clear that suitable assumptions about the asymptotic behavior of $F^{(-)\rho\sigma}(x')$ in the limit of null infinity will lead to nonvanishing weighted limits of (3.13). These limits are linear combinations of the N–P constants.

To derive an explicit expression for the N–P constants it is necessary to define a null tetrad at $x'$, $L^\mu$, $N^\mu$, and $M^\mu$ which satisfy the same algebraic conditions as $l^\mu$, $n^\mu$, $m^\mu$, Eq. (3.8). Then defining $\phi_0(x')$, $\phi_1(x')$, and $\phi_2(x')$ in terms of this tetrad as in (3.12), we have

$$\begin{aligned} \begin{pmatrix} \phi_0 \\ \phi_1 \\ \phi_2 \end{pmatrix} = \int_{\mathcal{N}} & \begin{pmatrix} L_\mu M_\nu \\ L_\mu N_\nu \\ \bar{M}_\mu N_\nu \end{pmatrix} \\ & \times U^{\mu\nu}\{2\phi_{0,\kappa}l^\kappa + 3\phi_0 l^\kappa{}_{,\kappa}\} D(x - x')\, d\tau_{(3)}. \end{aligned} \tag{3.14}$$

Introducing once again spherical coordinates on $\mathcal{N}$, one can choose the origin of polar coordinates on $S$ so that

$$\begin{aligned} L_\mu M_\nu U^{\mu\nu} &= -\frac{1}{r_2}\,{}_{-1}Y_{1,-1}(\theta, \varphi), \\ L_\mu N_\nu U^{\mu\nu} &= \frac{1}{r_2}\,{}_{-1}Y_{1,0}(\theta, \varphi), \\ \bar{M}_\mu N_\nu U^{\mu\nu} &= \frac{1}{r_2}\,{}_{-1}Y_{1,1}(\theta, \varphi), \end{aligned} \tag{3.15}$$

where the ${}_{-1}Y_{1,m}$ are spin-$s$ spherical functions[6,7] of spin weight $-1$. Then

$$\begin{aligned} Q_m &= \lim_{x'^0 \to \infty} \tfrac{1}{16}(x'^0)^4 \phi_{1+m} \\ &= \lim_{r\to\infty} r^2 \int_S {}_{-1}Y_{1,m}\partial_r(r^3\phi_0)\sin\theta\, d\theta\, d\varphi \end{aligned} \tag{3.16}$$

which agrees with the definition given by Newman and Penrose.[2] The higher-order N–P constants can be obtained in a similar manner to that outlined in Sec. 2 and indicated in Eq. (2.21).

## 4. DISCUSSION

The above calculations can be easily extended to what Penrose[8] calls *basic free fields*. These are fields which can be represented by wholly symmetric spinors with either dotted or undotted indices, but not both. Further, the basic free fields satisfy the first-order field equations

$$\nabla^{A\dot{K}}\psi_{ABC\ldots} = 0. \tag{4.1}$$

Linearized general relativity is included in these basic free fields. The conformal tensor $C_{\mu\nu\rho\sigma}$ can be represented by a completely symmetric rank 4 spinor: $\psi_{ABCD}$. When the Einstein equations are satisfied, $R_{\mu\nu} = 0$, the only part of the curvature tensor which may be different from zero is the conformal tensor. Then the linearized Bianchi identities take the form of

[6] E. Newman and R. Penrose, J. Math. Phys. 7, 863 (1966).
[7] J. N. Goldberg, A. J. MacFarlane, E. T. Newman, F. Rohrlich, and E. C. G. Sudarshan, J. Math. Phys. 8, 2155 (1967).
[8] R. Penrose, preprint, reproduced in P. G. Bergman's A.R.L. Tech. Documentary Report 63–65 (1963).

Eq. (4.1) which may be taken to be a propagation equation for the conformal tensor itself. Unfortunately, the extension of these ideas to the nonlinear gravitational theory has not yet been completed. However, preliminary results indicate that the approach given here is essentially correct.

One question which naturally arises is whether one could find an invariant transformation which maps solutions of the first-order equations into solutions of the first-order equations rather than into solutions of the wave equation. This can be done by using the Hertz potentials[9] to construct the fields. The Hertz potentials always satisfy the wave equation. Therefore, by the methods of this paper Hertz potentials may be mapped into Hertz potentials. Fields of spin $s$ are related to the Hertz potentials by $2s$ differentiations. These fields then satisfy the first-order equations (4.1).

The reason we have chosen not to present this approach in the main body of the paper is that we are interested in developing a method which may be applicable in a space–time with curvature. While it is true that a potential exists for Maxwell's field, spin 1, even in a curved space–time, it is not true for any other spin. In particular, it is not true for the gravitational field. Preliminary calculations indicate that the Green's theorem approach to the wave equation for the field itself does show promise of giving an explanation in curved space–time as well as it does in Minkowski space.

It is fairly clear, however, that for linear equations in flat space–time these constants of the motion are trivial. The invariant transformation is simply the fact that the solutions of the wave equation form a linear vector space. Therefore, the constants are not related to an intrinsic property of the field analogous to charge or energy. The situation may be different in a curved space–time or for nonlinear equations.

[9] R. Penrose, Proc. Roy. Soc. (London) **284A**, 159 (1965).


## ACKNOWLEDGMENTS

The author wants to thank E. T. Newman and R. Penrose for several lengthy discussions and for letting me see their manuscript prior to publication. Professor Peter Bergmann and David Robinson independently suggested using the Hertz potentials which are discussed in the last section. Finally, the author wants to acknowledge discussions with and criticisms from E. Couch and A. Staruszkiewicz, R. Torrence, and A. Trautman.


JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 5 MAY 1968

# Frequency Spectra of Harmonic Lattices with Weak Long-Range Interactions*

ROBERT ZWANZIG
*Institute for Fluid Dynamics and Applied Mathematics and*
*Institute for Molecular Physics, University of Maryland, College Park, Maryland*

(Received 8 November 1967)

Vibrational-frequency spectra are calculated analytically for certain harmonic lattice models with weak long-range interactions, in one and two dimensions. The force constants are chosen to decay, for large separations $n$, approximately as $\exp(-\gamma n)$, where $\gamma$ is an inverse range parameter. In the limit of infinite $\gamma$, standard nearest-neighbor results are recovered. In the limit of vanishing $\gamma$, or infinite interaction range, the frequency spectra have entirely different singularities. In the examples studied here, these can be either poles or branch points at the maximum frequency. When $\gamma$ is small but not zero, then the singularities are still qualitatively the same as in the corresponding nearest-neighbor models; but the general shapes of the spectra are dominated by "false" singularities lying outside the allowed frequency domain. Except for a small region near the maximum frequency, spectra obtained in the limit of vanishing $\gamma$ are good approximations to the correct spectra for small $\gamma$.

## INTRODUCTION

In recent years much attention has been given to the theoretical study of many-body systems with weak long-range forces. For the most part, this activity appears to be motivated not by the physical reality of the models, but rather by interest in the mathematical singularities of various properties of such systems, for example, critical points, etc.

The basis for interest in singularities is the commonly held feeling that the singularities of a function are "fingerprints" that determine the essential properties of the function. It is hoped that precise knowledge about the nature and location of singularities, together with a few numerical values to fill in gaps, will

* This research was supported in part by National Science Foundation Grant NSF GP-7652.

Eq. (4.1) which may be taken to be a propagation equation for the conformal tensor itself. Unfortunately, the extension of these ideas to the nonlinear gravitational theory has not yet been completed. However, preliminary results indicate that the approach given here is essentially correct.

One question which naturally arises is whether one could find an invariant transformation which maps solutions of the first-order equations into solutions of the first-order equations rather than into solutions of the wave equation. This can be done by using the Hertz potentials[9] to construct the fields. The Hertz potentials always satisfy the wave equation. Therefore, by the methods of this paper Hertz potentials may be mapped into Hertz potentials. Fields of spin $s$ are related to the Hertz potentials by $2s$ differentiations. These fields then satisfy the first-order equations (4.1).

The reason we have chosen not to present this approach in the main body of the paper is that we are interested in developing a method which may be applicable in a space–time with curvature. While it is true that a potential exists for Maxwell's field, spin 1, even in a curved space–time, it is not true for any other spin. In particular, it is not true for the gravitational field. Preliminary calculations indicate that the Green's theorem approach to the wave equation for the field itself does show promise of giving an explanation in curved space–time as well as it does in Minkowski space.

It is fairly clear, however, that for linear equations in flat space–time these constants of the motion are trivial. The invariant transformation is simply the fact that the solutions of the wave equation form a linear vector space. Therefore, the constants are not related to an intrinsic property of the field analogous to charge or energy. The situation may be different in a curved space–time or for nonlinear equations.

[9] R. Penrose, Proc. Roy. Soc. (London) **284A**, 159 (1965).


## ACKNOWLEDGMENTS

The author wants to thank E. T. Newman and R. Penrose for several lengthy discussions and for letting me see their manuscript prior to publication. Professor Peter Bergmann and David Robinson independently suggested using the Hertz potentials which are discussed in the last section. Finally, the author wants to acknowledge discussions with and criticisms from E. Couch and A. Staruszkiewicz, R. Torrence, and A. Trautman.


JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 5 MAY 1968

# Frequency Spectra of Harmonic Lattices with Weak Long-Range Interactions*

ROBERT ZWANZIG
*Institute for Fluid Dynamics and Applied Mathematics and*
*Institute for Molecular Physics, University of Maryland, College Park, Maryland*

(Received 8 November 1967)

Vibrational-frequency spectra are calculated analytically for certain harmonic lattice models with weak long-range interactions, in one and two dimensions. The force constants are chosen to decay, for large separations $n$, approximately as $\exp(-\gamma n)$, where $\gamma$ is an inverse range parameter. In the limit of infinite $\gamma$, standard nearest-neighbor results are recovered. In the limit of vanishing $\gamma$, or infinite interaction range, the frequency spectra have entirely different singularities. In the examples studied here, these can be either poles or branch points at the maximum frequency. When $\gamma$ is small but not zero, then the singularities are still qualitatively the same as in the corresponding nearest-neighbor models; but the general shapes of the spectra are dominated by "false" singularities lying outside the allowed frequency domain. Except for a small region near the maximum frequency, spectra obtained in the limit of vanishing $\gamma$ are good approximations to the correct spectra for small $\gamma$.

## INTRODUCTION

In recent years much attention has been given to the theoretical study of many-body systems with weak long-range forces. For the most part, this activity appears to be motivated not by the physical reality of the models, but rather by interest in the mathematical singularities of various properties of such systems, for example, critical points, etc.

The basis for interest in singularities is the commonly held feeling that the singularities of a function are "fingerprints" that determine the essential properties of the function. It is hoped that precise knowledge about the nature and location of singularities, together with a few numerical values to fill in gaps, will

* This research was supported in part by National Science Foundation Grant NSF GP-7652.

specify the function sufficiently well for practical purposes.

This may be in fact a valid point of view. It is known to work very well, for example, in calculating the vibrational-frequency spectrum of a harmonic lattice with typical short-range forces.

Whether this point of view is valid with respect to systems with weak long-range forces seems to be more doubtful. In this article an example is presented where knowledge of the precise nature and location of a singularity is irrelevant and even misleading in a practical sense. The frequency spectrum of a harmonic lattice with interactions that are weak and have a long but finite range is the example presented.

From the work of Van Hove[1] and others[2] it is known that the frequency spectrum of a harmonic lattice with interactions of finite range has certain kinds of singularities, depending on dimensionality. For example, in one dimension the typical singularity is inverse square root; in two dimensions the typical singularity is logarithmic.

In the examples to be discussed here, the singularities predicted by Van Hove's arguments are the only ones present in the frequency spectrum and yet the shape of the spectrum is determined almost completely (e.g., over many decades in the distribution function) by a "false" singularity that the spectrum does not possess.

## MOTIVATIONS

The present investigation was suggested by several observations and an analogy. First we make the observations.

As is well known, a one-dimensional gas whose molecules interact with a finite-range potential does not undergo a phase transition, i.e., its thermodynamic properties are analytic functions of temperature, etc. If, however, a weak long-range interaction is added on, and the appropriate limit to infinite range and zero strength is taken, then *in this limit* the gas undergoes a phase transition of the van der Waals type, and the thermodynamic functions have singularities in temperature, etc.[3]

A three-dimensional gas whose molecules interact with a finite-range potential is known experimentally to undergo a phase transition, and some analytic properties of its thermodynamic functions are known with moderate precision (e.g., its critical-point exponents). The corresponding three-dimensional van der Waals gas has been investigated theoretically[4]; in the limit of infinite range and zero strength, the analytic behavior of thermodynamic functions is entirely different than observed for finite-range interactions.

In this connection, attention is called to Van Kampen's treatment of condensation of a van der Waals gas.[5] A remarkable feature of this work is that dimensionality and the limit of infinite range are not invoked explicitly. At first glance, the results appear to be valid for interactions of finite range and strength and they have the same character as those found in the limit. Clearly this theory cannot be entirely correct, because we know that condensation does not occur in one dimension except in the limit. Van Kampen's derivation, however, is so simple and plausible that it should be taken seriously, i.e., we should not immediately jump to the conclusion that his results are correct *only* in the limit of infinite range. They must have some kind of validity even for finite range.

The observations that have just been made suggest two conclusions. First, it is clear that thermodynamic functions are *not continuous* at the limit of infinite range and zero strength. Second, it seems likely, in particular from Van Kampen's work, that the behavior found in the limit is *representative in some approximate sense* of the behavior before the limit is taken.

We are led to analogous conclusions by our investigation of the vibrational-frequency spectrum of a harmonic lattice with weak finite long-range interactions.

Another related observation comes from work by Bowers and Rosenstock,[6] who calculated the frequency spectrum of a two-dimensional harmonic lattice with nearest- and next-nearest-neighbor interactions. They found that the distribution function $G(\omega^2)$ of the spectrum changes substantially as the relative strength of next-nearest-neighbor interaction is varied. With only nearest-neighbor interactions, the function $G(\omega^2)$ has a logarithmic singularity at $\omega^2 = \frac{1}{2}\omega_{\max}^2$. As more and more next-nearest-neighbor interaction is added on, the singularity remains logarithmic (as it must), but it shifts to higher and higher frequencies. When the nearest- and next-nearest-neighbor-interaction strengths are equal, the logarithmic peak falls on the upper limit $\omega_{\max}^2$ of the spectrum.

This example shows that striking quantitative changes in the spectrum can be produced by increasing

[1] L. Van Hove. Phys. Rev. **89**, 1189 (1953).

[2] The basic reference in this field is A. A. Maradudin, E. W. Montroll, and G. H. Weiss, *Theory of Lattice Dynamics in the Harmonic Approximation* (Academic Press Inc., New York, 1963). Chapter III contains a comprehensive review and bibliography on singularities.

[3] M. Kac, G. Uhlenbeck, and P. Hemmer, J. Math. Phys. **4**, 216, 229 (1963); **5**, 60 (1964).

[4] J. Lebowitz and O. Penrose, J. Math. Phys. **7**, 98 (1966).

[5] N. Van Kampen, Phys. Rev. **135**, A362 (1964).

[6] W. A. Bowers and H. B. Rosenstock, J. Chem. Phys. **18**, 1056 (1950).

the range of interaction. However, the qualitative (i.e., analytic) behavior is not changed; Van Hove's theorem still holds.

We bypass substantial literature, reviewed by Maradudin, Montroll, and Weiss,[2] concerned with long-range interactions of ionic type; here, Van Hove's arguments do not apply, and all sorts of singularities are found.

The analogy referred to earlier is concerned with the dual role of the integral

$$\int_{-\pi}^{\pi}\int d\theta_1\, d\theta_2 (z - 2\cos\theta_1 - 2\cos\theta_2)^{-1}, \qquad (1)$$

regarded as a function of the complex variable $z$. This integral can be evaluated exactly in terms of elliptic integrals. It has a logarithmic singularity at $z = 4$.

One place that this integral appears is in the specific heat of a two-dimensional Ising lattice (square, isotropic, nearest-neighbor interactions).[7] In this example, $z$ is a function of temperature; the specific heat has a logarithmic singularity at the temperature determined by the condition $z = 4$.

Another place that this integral appears is in the frequency spectrum of a two-dimensional harmonic lattice (square, isotropic, nearest-neighbor interactions).[8] Here, $z$ is a function of frequency. In consequence, the distribution function $G(\omega^2)$ has a logarithmic singularity at the frequency determined by the condition $z = 4$.

We must, however, not take this analogy too seriously. Cases are known where it is false. For example, the distribution function $G(\omega^2)$ of a three-dimensional harmonic lattice is determined by an integral having the same structure as (1)

$$\int_{-\pi}^{\pi}\int\int d\theta_1\, d\theta_2\, d\theta_3 (z - 3\cos\theta_1 - 3\cos\theta_2 - 3\cos\theta_3)^{-1}. \qquad (2)$$

This integral has square-root singularities, and so does $G(\omega^2)$. As far as is known,[9] the specific heat of the corresponding three-dimensional Ising lattice has entirely different singularities, and cannot be expressed in terms of the same integral.

The validity of this analogy with respect to systems with weak long-range forces in two dimensions is not known at all.

*If* the analogy is good, then investigation of the frequency spectrum of a two-dimensional lattice with weak long-range forces may shed some light on properties of the corresponding Ising lattice. Even if the analogy is not good, such investigations still provide examples of the practical irrelevance of theoretically relevant singularities.

The preceding observations and analogy were motivations for the analysis to be described now.

## ONE-DIMENSIONAL MODEL

Before proceeding to the more interesting two-dimensional model, we discuss a simple example in one dimension which illustrates the main point.

Let $u_n$ be the displacement from equilibrium of the $n$th atom. Then in the harmonic approximation the potential energy may be written as

$$U = \tfrac{1}{2}\sum_m \sum_n A(m-n) u_m u_n. \qquad (3)$$

The force-constant matrix $A(m-n)$ clearly depends on only the separation $|m-n|$. It should be noted that, for reasons of translational invariance, the force constants must obey the condition

$$\sum_m A(m) = 0, \qquad (4)$$

or

$$A(0) = -\sum_{m\neq 0} A(m). \qquad (5)$$

In the nearest-neighbor model, the only nonvanishing matrix elements are $A(1) = A(-1)$ and $A(0) = -2A(1)$.

The long-range model to be analyzed here is defined by the special choice of force constants

$$A(m) = \alpha \exp -\gamma(|m| - 1); \quad m \neq 0; \qquad (6)$$

and $A(0)$ is given by Eq. (5). The parameter $\gamma$ plays the role of an inverse length. Note that only nearest-neighbor interactions remain in the limit of infinite $\gamma$. On the other hand, small $\gamma$ means that interactions extend over many neighbors.

Weiss[10] has analyzed the frequency spectrum of a related system. The essential differences which turn out to be of great importance, are: (1) He treats the nearest neighbors in a different way than the distant neighbors, and (2) his force constants alternate in sign, while ours all have the same sign.

The dispersion relation is given by the standard expression

$$\omega^2(q) = \sum_m A(m) \exp(iqm). \qquad (7)$$

(Throughout this article we will normalize force constants and masses to suit our convenience. Here, for example, we have set the mass equal to unity.)

[7] L. Onsager, Phys. Rev. **65**, 117 (1944).
[8] E. W. Montroll, J. Chem. Phys. **15**, 575 (1947).
[9] The most recently published analysis is by G. A. Baker, Jr., and D. S. Gaunt, Phys. Rev. **155**, 545 (1967).
[10] G. H. Weiss, Bull. Res. Council Israel **7F**, 165 (1958).

In the limit of an infinitely large lattice, the sum in Eq. (7) may be performed explicitly, leading to

$$\omega^2(q) = -2\alpha/(1 - e^{-\gamma}) + 2\alpha \frac{\cos q - e^{-\gamma}}{1 + e^{-2\gamma} - 2e^{-\gamma}\cos q}. \tag{8}$$

It is easy to verify that this reduces to the usual expression

$$\omega^2(q) = -2\alpha(1 - \cos q) \tag{9}$$

in the limit of infinitely large $\gamma$.

The behavior of $\omega^2(q)$ for small $q$ determines the sound velocity,

$$\omega^2(q) \simeq -[\alpha(1 + e^{-\gamma})/(1 - e^{-\gamma})^3]q^2 + \cdots. \tag{10}$$

We choose the coefficient $\alpha$ so that the long-wavelength sound velocity is unity,

$$\alpha = \alpha(\gamma) = -(1 - e^{-\gamma})^3/(1 + e^{-\gamma}). \tag{11}$$

Then the dispersion relation is

$$\omega^2(q) = \frac{(1 - e^{-\gamma})^2 \cdot 2(1 - \cos q)}{(1 - e^{-\gamma})^2 + 2e^{-\gamma}(1 - \cos q)}. \tag{12}$$

The maximum allowed frequency is $\omega_m$,

$$\omega_m = 2(1 - e^{-\gamma})/(1 + e^{-\gamma}) = 2\tanh(\gamma/2). \tag{13}$$

Thus the dispersion relation becomes

$$\omega^2(q) = \frac{2(1 - \cos q)}{1 + (1/\omega_m^2 - \frac{1}{4}) \cdot 2(1 - \cos q)}. \tag{14}$$

This form seems most useful for calculating.

The density of states $g(\omega)$ can be found by differentiation,

$$g(\omega) = dq/d\omega. \tag{15}$$

Note that in this one-dimensional calculation we use the distribution $g(\omega)$ in $\omega$, rather than the distribution $G(\omega^2)$ in $\omega^2$ that is more customary in two and three dimensions.

On inversion, $q$ can be found as a function of $\omega$,

$$q = 2 \arcsin (1 + 4/\omega^2 - 4/\omega_m^2)^{-\frac{1}{2}}, \tag{16}$$

and the differentiation is elementary. The result is

$$g(\omega) = (1 - \omega^2/\omega_m^2)^{-\frac{1}{2}}(1 + \omega^2/4 - \omega^2/\omega_m^2)^{-\frac{1}{2}}. \tag{17}$$

In the nearest-neighbor limit, $\gamma$ is infinite and $\omega_m = 2$. Then (17) reproduces the standard result: an inverse square-root singularity at the maximum frequency.

In fact this is the only singularity possible. By a simple rearrangement, (17) is transformed into

$$g(\omega) = \frac{4}{4 - \omega_m^2}\left(\frac{4}{4 - \omega_m^2} - \frac{\omega^2}{\omega_m^2}\right)^{-1}\left(1 - \frac{\omega^2}{\omega_m^2}\right)^{-\frac{1}{2}}. \tag{18}$$

For any finite $\gamma$, the maximum frequency $\omega_m$ is always smaller than 2, and so

$$4/(4 - \omega_m^2) > 1. \tag{19}$$

The inverse first-power singularity in the distribution function clearly always lies outside the allowed range of frequency

$$\omega^2/\omega_m^2 < 1. \tag{20}$$

Nevertheless, when $\gamma$ is very small, and $\omega_m$ is also very small, the distribution function can be represented to a very good approximation by an inverse three-halves power singularity,

$$g(\omega) \cong (1 - \omega^2/\omega_m^2)^{-\frac{3}{2}}. \tag{21}$$

This approximation is good for frequencies of the order of

$$0 \leq \omega^2/\omega_m^2 \lesssim 1 - \omega_m^2/4. \tag{22}$$

Once the frequency becomes larger than the upper limit in the inequality (22), then the inverse square-root singularity takes over, and from then on the distribution function is approximately

$$g(\omega) \cong 4/\omega_m^2(1 - \omega^2/\omega_m^2)^{-\frac{1}{2}}. \tag{23}$$

As an illustration of orders of magnitude, let us suppose that $\gamma = 0.01$. Then the approximate (21), with the inverse three-halves power singularity, is numerically good for the frequency domain

$$0 \leq \omega^2/\omega_m^2 \lesssim 0.999, \tag{24}$$

and fails only when $\omega^2/\omega_m^2$ becomes greater than 0.999. From a practical point of view, almost all of the spectrum is described by a "false" singularity, lying outside the physically allowed frequency domain.

The distribution function $g(\omega)$ *in the limit* of vanishing $\gamma$ is

$$\lim_{\gamma \to 0} g(\omega) = (1 - \omega^2/\omega_m^2)^{-\frac{3}{2}}. \tag{25}$$

In the limit, the inverse square-root singularity disappears entirely. This illustrates the lack of continuity at the point $\gamma = 0$ which seems to be characteristic of systems with long-range interactions.

[It may be noticed that the maximum frequency itself vanishes in the limit. This is a consequence of imposing the requirement that the long-wavelength sound velocity remain fixed as the limit is taken. By making a different choice of the coefficient $\alpha(\gamma)$, then $\omega_m$ can be kept different from zero in the limit.]

The limiting form (25) is identical with the approximation (21), which is known, for long but finite-range interaction, to be a good approximation over a wide range of frequency. The limiting form (25) fails to

represent the correct distribution for long but finite-range interaction only over a very small range of frequency near the true singularity.

## TWO-DIMENSIONAL MODEL

Two-dimensional models are perhaps more interesting because of the possible analogy with Ising lattice problems. However, they show essentially the same qualitative features as the one-dimensional model just discussed.

What little is known about $G(\omega^2)$ for two-dimensional lattices with long-range forces is summarized in Ref. 2. Attention has been given mainly to interactions of ionic character, where the singularities are entirely different from those found in lattices with finite range interactions.

The procedure used to get the distribution function $g(\omega)$ in one dimension, based on Eq. (15), does not work in two or three dimensions, and another approach is needed. Our procedure is simply to express the distribution function for the system of interest in terms of the known distribution function for the square lattice with isotropic nearest-neighbor interactions. This idea has been used before, by Mazur,[11] to relate properties of diatomic lattices to those of the corresponding monatomic lattices.

Even this procedure is difficult to follow in general; so we work backwards. That is, we guess at a dispersion relation, chosen so that the actual distribution function can be related to the nearest-neighbor distribution function, and then we verify that the guessed dispersion relation arises from long-range interactions. It will be seen that this inverse method works nicely.

The starting point for our calculation is the expression

$$G(\omega^2) = \langle \delta[\omega^2 - \omega^2(q_1, q_2)] \rangle_{\text{av}} \tag{26}$$

for the distribution function $G(\omega^2)$. The average, denoted by angular brackets, means an average over all angles

$$\langle \ \ \rangle_{\text{av}} = \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi}\int dq_1\, dq_2. \tag{27}$$

In a square lattice with isotropic nearest-neighbor interactions, the dispersion relation is given by

$$\omega_0^2(q_1, q_2) = 4 - 2\cos q_1 - 2\cos q_2. \tag{28}$$

This particular spectrum is indicated by the subscript zero. For convenience, we normalize units so that the long-wavelength sound velocity is unity

$$\omega_0^2 \simeq q_1^2 + q_2^2 + \cdots. \tag{29}$$

The distribution function $G_0(\omega^2)$ for this spectrum was found by Montroll[8]; it can be expressed in terms of the complete elliptic integral $K(m)$

$$G_0(\omega^2) = (1/2\pi^2)K[\omega(8 - \omega^2)^{\frac{1}{2}}/4]. \tag{30}$$

We use the notation of Ref. 12, where the variable $m$ replaces the more conventional variable $k$, according to $m = k^2$. This reference gives polynomial approximations for $K(m)$ that are quite useful for numerical calculations.

Now let us *assume* that the actual dispersion relation $\omega^2(q_1, q_2)$ can be expressed as a function $F$ of the nearest-neighbor frequency,

$$\omega^2(q_1, q_2) = F[\omega_0^2(q_1, q_2)]. \tag{31}$$

The actual form of the function $F(x)$ will be discussed later. The inverse function is denoted by $F^{-1}(\omega^2)$

$$\omega_0^2(q_1, q_2) = F^{-1}(\omega^2). \tag{32}$$

The derivative of $F(x)$ is denoted by the standard $F'(x)$.

Next we observe that the delta function *of a function* $f(x)$ can be expressed, according to a familiar formula, in terms of the delta function of the variable $x$

$$\delta[f(x) - a] = |f'(x_0)|^{-1}\delta(x - x_0), \tag{33}$$

where $x_0$ is the root of the equation

$$f(x_0) = a. \tag{34}$$

(If this equation has several roots, then we must sum them all; but in the present instance, only one root occurs.)

On making use of the above observation, we find that the distribution function $G(\omega^2)$ can be transformed to

$$G(\omega^2) = \frac{\langle \delta[F^{-1}(\omega^2) - \omega_0^2(q_1, q_2)] \rangle_{\text{av}}}{|F'[F^{-1}(\omega^2)]|}. \tag{35}$$

But the numerator is just the nearest-neighbor distribution function $G_0$, so that the desired distribution function is related to the known $G_0$ by

$$G(\omega^2) = G_0[F^{-1}(\omega^2)]/|F'[F^{-1}(\omega^2)]|. \tag{36}$$

The rest of the calculation depends on specific choices of $F(x)$.

This function is not arbitrary. We want to choose it so that the interaction force constants fall off asymptotically in the desired way, with a long but finite

[11] P. Mazur, thesis, University of Maryland, 1956 (unpublished); see also, A. A. Maradudin, P. Mazur, E. W. Montroll, and G. H. Weiss, Rev. Mod. Phys. **30**, 175 (1958).

[12] *Handbook of Mathematical Functions*, M. Abramowitz and I. A. Stegun, Eds. (National Bureau of Standards Applied Mathematics Series AMS 55 1964), see especially pp. 590–591.

range. Suppose that the potential energy in the harmonic approximation is

$$U = \tfrac{1}{2} \sum A(m_1 - m_2, n_1 - n_2) u(m_1, n_1) u(m_2, n_2), \tag{37}$$

where $u(m, n)$ is the displacement (in the $z$ direction) of the atom whose equilibrium position is $x = m$, $y = n$. Then the force constants are $A(m_1 - m_2, n_1 - n_2)$. The dispersion relation is given by

$$\omega^2(q_1, q_2) = \sum A(n_1, n_2) \exp i(q_1 n_1 + q_2 n_2). \tag{38}$$

On Fourier inversion, the force constants are evidently

$$A(n_1, n_2) = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int dq_1 \, dq_2 \omega^2(q_1, q_2) \exp -i(q_1 n_1 + q_2 n_2). \tag{39}$$

So a choice of $F(x)$ implies a choice of $A(m, n)$.

The force constants for large separation are determined mainly by the frequency $\omega^2$ for small $q_1$ and $q_2$, and in this neighborhood, Eq. (29) shows that the variable $x = \omega^2$ is essentially just the square of the vector $\mathbf{q} = (q_1, q_2)$. Then, with the corresponding definition of the vector $\mathbf{n} = (n_1, n_2)$, we can write approximately

$$A(n) \simeq \frac{1}{4\pi^2} \int d^2q F(q^2) \exp (-i\mathbf{q} \cdot \mathbf{n}). \tag{40}$$

Because only small $q$ is involved, we may replace the correct finite limits of integration by infinity. The integration over angles is then trivial, and leads to a Bessel function of the magnitudes of the vectors $\mathbf{n}$ and $\mathbf{q}$,

$$A(n) \sim \frac{1}{2\pi} \int_0^{\infty} dq q J_0(nq) F(q^2). \tag{41}$$

By imposing some desired form of asymptotic decay on the force constants $A(\mathbf{n})$, this integral equation can be inverted to yield $F(q^2)$.

This procedure, while feasible, leads to rather complicated formulas. Instead of following it closely, we combine it with other criteria for $F(q^2)$, among which simplicity of form is important.

The actual criteria are as follows. First, we ask that the force constants decay asymptotically as

$$A(n) \sim (n\gamma)^{\alpha - \frac{3}{2}} \exp (-n\gamma), \tag{42}$$

where $\alpha$ is some extra parameter. This includes simple exponential decay as the special case $\alpha = \frac{3}{2}$. Then, we require that $F(x)$ should approach $x$ for sufficiently small $x$, so that the long-wavelength sound velocity is unity. A third criterion is that $F(x)$ should become identical with $x$ in the limit of infinite $\gamma$; this corresponds to the limit of only nearest-neighbor interactions. A final criterion was that the choice should facilitate computation of the inverse function $F^{-1}$ and the derivative, so that Eq. (36) could be used conveniently.

By inspection of tables of integral transforms, it appeared that the choice

$$F(x) = (\gamma^2/\alpha)[1 - \gamma^{2\alpha}/(\gamma^2 + x)^{\alpha}] \tag{43}$$

would lead to useful results. It clearly fits several of our criteria; let us check on the asymptotic form of $A(\mathbf{n})$. When (43) is substituted in (41), the integral can be performed and the result for large $n$ contains the modified Bessel function $K_{\alpha-1}$,

$$A(n) \sim - \frac{\gamma^4}{2\pi\alpha\Gamma(\alpha)} (n\gamma)^{\alpha-1} K_{\alpha-1}(n\gamma). \tag{44}$$

If, furthermore, we put in the asymptotic form of the modified Bessel function, we obtain for large $n\gamma$ the limiting behavior

$$A(\mathbf{n}) \sim - \frac{\gamma^4}{2\pi\alpha\Gamma(\alpha)} \frac{1}{2^{\alpha-1}} (n\gamma)^{\alpha - \frac{3}{2}} \exp (-n\gamma), \tag{45}$$

which is in agreement with the criterion (41).

While other choices of $F(x)$ are no doubt possible, the present one appears to satisfy all of the natural requirements. In particular, it gives the nearest-neighbor spectrum when $\gamma$ becomes infinite, and it gives force constants that fall off approximately exponentially for large separation when $\gamma$ is small. The rest of our discussion is restricted to the particular choice (43).

The maximum allowed frequency in the nearest-neighbor case is $\omega_0^2 = 8$; so the maximum allowed frequency in the present case is $\omega_m^2 = F(8)$. For convenience we normalize the actual frequencies as follows:

$$z = \omega^2/\omega_m^2, \tag{46}$$

and our goal is to find the function $G(z)$. One more bit of notation is useful:

$$Q = 1 - \gamma^{2\alpha}/(\gamma^2 + 8)^{\alpha}. \tag{47}$$

Now we apply Eq. (36) to obtain $G(z)$,

$$G(z) = [(1 - Qz)^{1+1/\alpha}]^{-1} G_0\left[\frac{\gamma^2}{(1 - Qz)^{1/\alpha}} - \gamma^2\right]. \tag{48}$$

This formula, together with Eq. (30) for the function $G_0$, solves the problem of finding the distribution function.

It can be verified easily that $G(z)$ approaches the correct nearest-neighbor result in the limit of infinite $\gamma$.

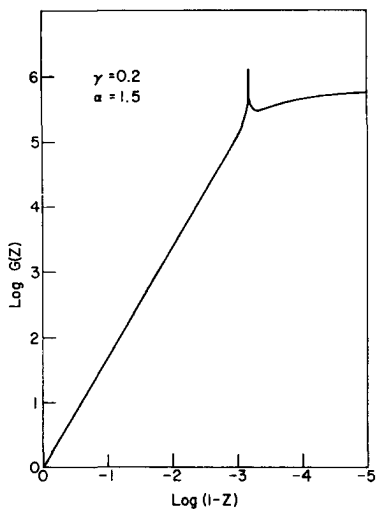



FIG. 1. Logarithm (to base 10) of the distribution function $G(z)$ plotted against logarithm of the deviation $1 - z$ from the maximum frequency.

When $\gamma$ is finite, the quantity $Q$ is always less than unity. The reduced frequency $z$ is by definition always less than unity. Thus the first factor in (48) cannot be singular in the allowed frequency range. The only possible singularity is the logarithmic one contained in $G_0$. This is in accord with Van Hove's argument; when $\gamma$ is finite, the range of interaction also is finite.

Now consider what happens when $\gamma$ is small. Then the quantity $Q$ is only slightly less than unity. For most of the allowed frequency range, $1 - Qz$ is of order unity, and the argument of the function $G_0$ is small, of order $\gamma^2$. This means that we are far from the logarithmic singularity. The actual location of this singularity is

$$z_s = \frac{1 - [\gamma^{2\alpha}/(4 + \gamma^2)^\alpha]}{1 - [\gamma^{2\alpha}/(8 + \gamma^2)^\alpha]}. \tag{49}$$

(If $\gamma = 0.2$ and $\alpha = \frac{3}{2}$, then the singularity is at $z_s \cong 0.994$.) As long as $z$ remains appreciably smaller than $z_s$, the factor $G_0$ changes only slightly, and the distribution function is determined mainly by the first factor in (48).

*In the limit* of vanishing $\gamma$ (or infinite range), the distribution function approaches

$$\lim_{\gamma \to 0} G(z) = [(1 - z)^{1+1/\alpha}]^{-1}. \tag{50}$$

In this limit, the singularity is not logarithmic; it is a pole or branch point, depending on our choice of $\alpha$.

For small but nonvanishing $\gamma$, the distribution function is well approximated by its limit, provided that $z$ remains sufficiently smaller than $z_s$. We see again, as in the one-dimensional model, that most of the frequency range is described, for practical or numerical purposes, by a singularity that lies outside the allowed frequency range.

Figure 1 shows graphically how good the approximation can be. This was calculated numerically for the parameters $\alpha = \frac{3}{2}$ (corresponding to asymptotic exponential decay of force constants), and $\gamma = 0.2$ (or a range of the order of five lattice spacings). The ordinate is the logarithm (to base ten) of the distribution function $G(z)$, and the abscissa is the logarithm (to base ten) of the deviation $1 - z$ from unity. The limiting curve, for $\gamma = 0$, would be a straight line passing through the origin with slope 5/3. The actual curve follows this limiting curve very closely up to the point (5, −3), and then the effect of the true logarithmic singularity begins to be felt. The limiting function (49) is a good approximation for about *five decades* of $G(z)$, or for values of $z$ ranging from zero up to about 0.999. If an experimenter presented data in just this region, we would probably decide by curve fitting that the singularity is an inverse 5/3 power, and not at all logarithmic.

## CONCLUSIONS

What can we learn from the calculations just presented? The first point is that investigation of mathematical singularities, in many-body systems with weak long-range forces, can be misleading. In particular, singularities in the physically allowed range of variables may have very little relation to the general shape of the functions under investigation. Singularities *outside* the physically allowed region may dominate the singularities inside that region, except in very limited circumstances.

The second point is that results obtained in the limit of infinite range and zero strength may be remarkably good approximations to behavior for finite but long range and small strength. Such approximations must fail eventually, but their failure may be insignificant from a practical point of view.

The latter point perhaps explains Van Kampen's results[5] on the van der Waals gas. His calculations, for interactions of long and finite range, probably provide a good approximation to the correct equation of state of the gas, except in a very small region near the correct transition.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 4 MAY 1968

# Electrostatic Interactions of the Configuration $d^{n-1}sp$

CHARLES ROTH
*Department of Mathematics, McGill University, Montreal, Canada*

(Received 6 September 1967)

The matrix elements of the Coulomb interactions of the configuration $d^{n-1}sp$ were obtained for $L$–$S$ coupling in the form of linear combinations of certain radial integrals. The method used can be extended to the configuration $l^{n-1}l'l''$.

## 1. INTRODUCTION

The spectra of the neutral and singly ionized atoms of the iron group show a strong overlapping between the configurations $d^np$ and $d^{n-1}sp$. Furthermore, in some cases the configuration $d^{n-2}s^2p$ is not too far away. Thus in order to fit the experimental levels of the main odd configurations of the first and second spectra of the iron group to theoretically predicted values, it is necessary to consider the configurations $(d + s)^np$, i.e., the three configurations $d^np$, $d^{n-1}sp$, $d^{n-2}s^2p$ and the interactions among these configurations. The purpose of this paper is to use the Racah algebra[1] in order to obtain in closed form the angular parts of the electrostatic-energy matrix elements for the configuration $d^{n-1}sp$ in $L$–$S$ coupling.

Theoretical results are available for all the configurations $(d + s^np$ in the first spectra of the iron group and for the configurations $d^np + d^{n-1}sp$ of Sc II, Ti II, V II, and Cu II. The author hopes to publish these results soon.

According to the exclusion principle, the eigenfunctions of an atomic state are antisymmetric with respect to all electrons. An antisymmetric eigenfunction of the configuration $d^{n-1}sp$ can be written as

$$\begin{aligned}\psi(d^{n-1}sp) = [n(n+1)]^{-\frac{1}{2}}[&\psi(\{d_1\cdots d_{n-1}\}s_np_{n+1}) - \psi(\{d_1\cdots d_{n-1}\}s_{n+1}p_n)\\ &- \sum_{j=1}^{n-1}\psi(\{d_1\cdots d_{j-1}d_{n+1}d_{j+1}\cdots d_{n-1}\}s_np_j) - \sum_{i=1}^{n-1}\psi(\{d_1\cdots d_{i-1}d_nd_{i+1}\cdots d_{n-1}\}s_ip_{n+1})\\ &+ \sum_{i=1}^{n-1}\psi(\{d_1\cdots d_{i-1}d_{n+1}d_{i+1}\cdots d_{n-1}\}s_ip_n) + \sum_{j=1}^{n-1}\psi(\{d_1\cdots d_{j-1}d_nd_{j+1}\cdots d_{n-1}\}s_{n+1}p_j)\\ &+ \sum_{i\neq j=1}^{n-1}\psi(\{d_1\cdots d_{i-1}d_nd_{i+1}\cdots d_{j-1}d_{n+1}d_{j+1}\cdots d_{n-1}\}s_ip_j)].\end{aligned} \tag{1}$$

We then consider the matrix elements

$$\langle d^{n-1}sp|\sum_{i<j=1}^{n+1}e^2/r_{ij}\,|d^{n-1}sp\rangle = \frac{n(n+1)}{2}\langle d^{n-1}sp|\,e^2/r_{n,n+1}\,|d^{n-1}sp\rangle. \tag{2}$$

By using the expansion (1) for $\psi(d^{n-1}sp)$ we obtain contributions representing the $d$–$s$, $p$–$s$, $d$–$p$, and $d$–$d$ interactions, each interaction being characterized by the labeling of the electrons $n$ and $n + 1$.

## 2. THE $d$–$p$ INTERACTION

Since $e^2/r_{n,n+1}$ acts only on the electrons $n$ and $n + 1$, the matrix elements with $i$ different from $j$ vanish. Hence from (1) and (2) the $d$–$p$ interaction is given by

$$\begin{aligned}(n-1)[&\langle\{d_1\cdots d_{n-2}\cdots d_n\}s_{n-1}p_{n+1}|\,e^2/r_{n,n+1}\,|\{d_1\cdots d_{n-2}\cdots d_n\}s_{n-1}p_{n+1}\rangle\\ &- \langle\{d_1\cdots d_{n-2}\cdots d_n\}s_{n-1}p_{n+1}|\,e^2/r_{n,n+1}\,|\{d_1\cdots d_{n-2}\cdots d_{n+1}\}s_{n-1}p_n\rangle].\end{aligned} \tag{3}$$

We now specify

$$|\{d_1\cdots d_{n-2}\cdots d_n\}s_{n-1}p_{n+1}\rangle \equiv |d^{n-1}(\nu_1S_1L_1)s_{n-1}(S_2L_1)p_{n+1}SL\rangle.$$

Expanding by means of coefficients of fractional parentage[2] yields

$$\begin{aligned}|d^{n-1}(\nu_1S_1L_1)s_{n-1}(S_2L_1)p_{n+1}SL\rangle = \sum_{\nu_4S_4L_4}[&|d^{n-2}(\nu_4S_4L_4)d_n(S_1L_1)s_{n-1}(S_2L_1)p_{n+1}SL\rangle\\ &\times\langle d^{n-2}(\nu_4S_4L_4)\,dS_1L_1\,|\}\,d^{n-1}\nu_1S_1L_1\rangle].\end{aligned}$$

[1] U. Fano and G. Racah, *Irreducible Tensorial Sets* (Academic Press Inc., New York, 1958), henceforth referred to as ITS.
[2] G. Racah, Phys. Rev. 63, 367 (1943).

Recoupling the $s_{n-1}$ and $d_n$ electrons and then coupling the electrons $d_n$ and $p_{n+1}$ yields

$$d^{n-1}(v_1S_1L_1)s_{n-1}(S_2L_1)p_{n+1}SL\rangle = \sum_{\substack{v_4S_4L_4\\S_3S_5L_5}} [|d^{n-2}(v_4S_4L_4)s_{n-1}(S_3L_4),\, d_np_{n+1}(S_5L_5),\, SL\rangle\langle S_3L_4,\, d_np_{n+1}(S_5L_5),\, SL \mid S_3L_4d_n(S_2L_1)p_{n+1}SL\rangle \times \langle S_4L_4s_{n-1}(S_3L_4)d_nS_2L_1 \mid S_4L_4d_n(S_1L_1)s_{n-1}S_2L_1\rangle\langle d^{n-2}(v_4S_4L_4)dS_1L_1 \mid\} d^{n-1}v_1S_1L_1\rangle]. \quad (4)$$

Similarly, by using also (7.11) of ITS we obtain

$$|\{d_1\cdots d_{n-2}\cdots d_{n+1}\}s_{n-1}p_n\rangle = \sum_{\substack{v_4S_4L_4\\S_3S_5L_5}} [|d^{n-2}(v_4S_4L_4)s_{n-1}(S_3L_4),\, p_nd_{n+1}(S_5L_5),\, SL\rangle\,[\exp \pi i(S_5+L_5)] \times \langle S_3L_4,\, d_{n+1}p_n(S_5L_5),\, SL \mid S_3L_4d_{n+1}S_2L_1p_nSL\rangle\langle S_4L_4s_{n-1}(S_3L_4)d_{n+1}S_2L_1 \mid S_4L_4d_{n+1}(S_1L_1)s_{n-1}S_2L_1\rangle \times \langle d^{n-2}(v_4S_4L_4)dS_1L_1 \mid\} d^{n-1}v_1S_1L_1\rangle]. \quad (5)$$

### A. The Direct Part of the $d$–$p$ Interaction

From (3) and (4) the direct part of the $d$–$p$ interaction is given by

$$\begin{aligned}\text{D.I.}\ (d\text{–}p) &= (n-1)[\langle\{d_1\cdots d_{n-2}\cdots d_n\}s_{n-1}p_{n+1}|\, e^2/r_{n,n+1}\,|\{d_1\cdots d_{n-2}\cdots d_n\}s_{n-1}p_{n+1}\rangle]\\ &= (n-1)\sum_{\substack{v_4S_4L_4S_3S_5L_5\\v_4'S_4'L_4'S_3'S_5'L_5'}} [\langle d^{n-2}(v_4S_4L_4)s_{n-1}(S_3L_4),\, d_np_{n+1}(S_5L_5),\, SL|\, e^2/r_{n,n+1}\\ &\quad\times |d^{n-2}(v_4'S_4'L_4')s_{n-1}(S_3'L_4'),\, d_np_{n+1}(S_5'L_5'),\, SL\rangle\langle S_3L_4,\, d_np_{n+1}(S_5L_5),\, SL \mid S_3L_4d_n(S_2L_1)p_{n+1}SL\rangle\\ &\quad\times \langle S_3'L_4',\, d_np_{n+1}(S_5'L_5'),\, SL \mid S_3'L_4'd_n(S_2'L_1')p_{n+1}SL\rangle\langle S_4L_4s_{n-1}(S_3L_4)d_nS_2L_1 \mid S_4L_4d_n(S_1L_1)s_{n-1}S_2L_1\rangle\\ &\quad\times \langle S_4'L_4's_{n-1}(S_3'L_4')d_nS_2'L_1' \mid S_4'L_4'd_n(S_1'L_1')s_{n-1}S_2'L_1'\rangle\langle d^{n-1}(v_4'S_4'L_4')dS_1'L_1' \mid\} d^{n-1}v_1'S_1'L_1'\rangle\\ &\quad\times \langle d^{n-1}(v_1S_1L_1)\,\{|\, d^{n-2}(v_4S_4L_4)dS_1L_1\rangle]. \end{aligned} \quad (6)$$

The electrostatic interaction is of course diagonal in $S$ and $L$. Since the interaction is between the electrons $n$ and $n+1$, the quantum numbers of the first $(n-1)$ electrons must be the same on both sides. Thus in the summation of (6) we must insert $\delta(v_4S_4L_4, v_4'S_4'L_4')$ and $\delta(S_3, S_3')$. Since the matrix element of the interaction can be written as

$$\langle d_np_{n+1}S_5L_5|\, e^2/r_{n,n+1}\,|d_np_{n+1}S_5'L_5'\rangle,$$

it is apparent that we must have also $\delta(S_5L_5, S_5'L_5')$ in the summation of (6).

We then have

$$\begin{aligned}\text{D.I.}\ (d\text{–}p) &= (n-1)\sum_{\substack{v_4S_4L_4\\S_3S_5L_5}} [\langle d_np_{n+1}S_5L_5|\, e^2/r_{n,n+1}\,|d_np_{n+1}S_5L_5\rangle\langle S_3L_4,\, d_np_{n+1}(S_5L_5),\, SL \mid S_3L_4d_n(S_2L_1)p_{n+1}SL\rangle\\ &\quad\times \langle S_3L_4,\, d_np_{n+1}(S_5L_5),\, SL \mid S_3L_4d_n(S_2'L_1')p_{n+1}SL\rangle\langle S_4L_4s_{n-1}(S_3L_4)d_nS_2L_1 \mid S_4L_4d_n(S_1L_1)s_{n-1}S_2L_1\rangle\\ &\quad\times \langle S_4L_4s_{n-1}(S_3L_4)d_nS_2'L_1' \mid S_4L_4d_n(S_1'L_1')s_{n-1}S_2'L_1'\rangle\langle d^{n-2}(v_4S_4L_4)dS_1'L_1' \mid\} d^{n-1}v_1'S_1'L_1'\rangle\\ &\quad\times \langle d^{n-1}(v_1S_1L_1)\,\{|\, d^{n-2}(v_4S_4L_4)dS_1L_1\rangle. \end{aligned} \quad (7)$$

Since

$$\langle S_4L_4s_{n-1}(S_3L_4)d_nS_2L_1 \mid S_4L_4d_n(S_1L_1)s_{n-1}S_2L_1\rangle = \langle S_4\tfrac{1}{2}(S_3),\, \tfrac{1}{2}S_2 \mid S_4\tfrac{1}{2}(S_1),\, \tfrac{1}{2}S_2\rangle\langle L_40(L_4)2L_1 \mid L_42(L_1),\, 0L_1\rangle,$$

we obtain from Eqs. (11.10) and (11.12) of ITS

$$\langle S_4L_4s_{n-1}(S_3L_4)d_nS_2L_1 \mid S_4L_4d_n(S_1L_1)s_{n-1}S_2L_1\rangle\langle S_4L_4s_{n-1}(S_3L_4)d_nS_2'L_1' \mid S_4L_4d_n(S_1'L_1')s_{n-1}S_2'L_1'\rangle = [\exp \pi i(2S_3+S_1+S_1')](2S_3+1)[(2S_1+1)(2S_1'+1)]^{\frac{1}{2}}W\begin{pmatrix}\frac{1}{2} & S_4 & S_3\\ \frac{1}{2} & S_2 & S_1\end{pmatrix}W\begin{pmatrix}\frac{1}{2} & S_4 & S_3\\ \frac{1}{2} & S_2' & S_1'\end{pmatrix}. \quad (8)$$

Similarly, using (11.8) of ITS yields

$$\begin{aligned}&\langle S_3L_4,\, d_np_{n+1}(S_5L_5),\, SL \mid S_3L_4d_n(S_2L_1)p_{n+1}SL\rangle\langle S_3L_4,\, d_np_{n+1}(S_5L_5),\, SL \mid S_3L_4d_n(S_2'L_1')p_{n+1}SL\rangle\\ &\quad= [\exp \pi i(2S_3+2S)](2S_5+1)(2L_5+1)[(2S_2+1)(2S_2'+1)(2L_1+1)(2L_1'+1)]\\ &\quad\times W\begin{pmatrix}S_3 & \frac{1}{2} & S_2\\ \frac{1}{2} & S & S_5\end{pmatrix}W\begin{pmatrix}S_3 & \frac{1}{2} & S_2'\\ \frac{1}{2} & S & S_5\end{pmatrix}W\begin{pmatrix}L_4 & 2 & L_1\\ 1 & L & L_5\end{pmatrix}W\begin{pmatrix}L_4 & 2 & L_1'\\ 1 & L & L_5\end{pmatrix}. \end{aligned} \quad (9)$$

As $S_3$ and $S_5$ do not appear in the coefficients of fractional parentage and the interaction matrix element of (7) does not depend on $S_5$, we can sum the net spin contribution, denoted by R.C.$_{\text{spin}}$, over $S_3$ and $S_5$. Then from (8) and (9) we have

$$\sum_{S_3S_5} \text{R.C.}_{\text{spin}} = \sum_{S_3S_5} \{[\exp \pi i(2S + S_1 + S_1')](2S_3 + 1)(2S_5 + 1)[2S_1 + 1)(2S_1' + 1)(2S_2 + 1)(2S_2' + 1)]^{\frac{1}{2}} \times \overline{W}\begin{pmatrix} \frac{1}{2} & S_4 & S_3 \\ \frac{1}{2} & S_2 & S_1 \end{pmatrix} \overline{W}\begin{pmatrix} \frac{1}{2} & S_4 & S_3 \\ \frac{1}{2} & S_2' & S_1' \end{pmatrix} \overline{W}\begin{pmatrix} S_3 & \frac{1}{2} & S_2 \\ \frac{1}{2} & S & S_5 \end{pmatrix} \overline{W}\begin{pmatrix} S_3 & \frac{1}{2} & S_2' \\ \frac{1}{2} & S & S_5 \end{pmatrix}. \quad (10)$$

Use of the symmetry properties of the $\overline{W}$'s and repeated application of (11.15) of ITS yields

$$\sum_{S_3S_5} \text{R.C.}_{\text{spin}} = \delta(S_1S_1')\delta(S_2S_2'). \quad (11)$$

Also from (8) and (9), the net contribution of the $L$-recoupling coefficients is given by

$$\text{R.C.}_{\text{a.m.}} = (2L_5 + 1)[(2L_1 + 1)(2L_1' + 1)]^{\frac{1}{2}} \overline{W}\begin{pmatrix} L_4 & 2 & L_1 \\ 1 & L & L_5 \end{pmatrix} \overline{W}\begin{pmatrix} L_4 & 2 & L_1' \\ 1 & L & L_5 \end{pmatrix}. \quad (12)$$

From Eqs. (16.15) and (16.17) of ITS we have

$$\langle d_n p_{n+1} S_5 L_5 |\, e^2/r_{n,n+1} \,| d_n p_{n+1} S_5 L_5 \rangle = \sum_k R^k(dp, dp)[\exp \pi i(1 + L_5 + k)](2|\,|C^k|\,|2)(1|\,|C^k|\,|1) \overline{W}\begin{pmatrix} 2 & 2 & k \\ 1 & 1 & L_5 \end{pmatrix}. \quad (13)$$

Using (14.12) of ITS, the fact that $(1 + 1 + k)$ in the reduced matrix element must be even, and the triangular condition, (7.4) of ITS, for each triad of the $\overline{W}$ coefficient yields

$$\langle d_n p_{n+1} S_5 L_5 |e^2/r_{n,n+1}| d_n p_{n+1} S_5 L_5 \rangle = R^0(dp, dp)[\exp \pi i(L_5 + 1)](15)^{\frac{1}{2}} \times \overline{W}\begin{pmatrix} 2 & 1 & L_5 \\ 1 & 2 & 0 \end{pmatrix} + R^2(dp, dp)[\exp \pi i(L_5 + 1)](\tfrac{1}{7}\tfrac{2})^{\frac{1}{2}} \overline{W}\begin{pmatrix} 2 & 1 & L_5 \\ 1 & 2 & 2 \end{pmatrix}. \quad (14)$$

Inserting (11), (12), and (14) into (7) gives for the direct part of the $d$–$p$ interaction

$$\text{D.I.} \quad (d\text{–}p) = (n - 1)\delta(S_1S_1')\delta(S_2S_2') \sum_{\nu_4 S_4 L_4 L_5} \{[\exp \pi i(L_5 + 1)](2L_5 + 1)[(2L_1 + 1)(2L_1' + 1)]^{\frac{1}{2}} \times \overline{W}\begin{pmatrix} L_4 & 2 & L_1 \\ 1 & L & L_5 \end{pmatrix} \overline{W}\begin{pmatrix} L_4 & 2 & L_1' \\ 1 & L & L_5 \end{pmatrix} \left[ (15)^{\frac{1}{2}} R^0(dp, dp) \overline{W}\begin{pmatrix} 2 & 1 & L_5 \\ 1 & 2 & 0 \end{pmatrix} + (\tfrac{1}{7}\tfrac{2})^{\frac{1}{2}} R^2(dp, dp) \overline{W}\begin{pmatrix} 2 & 1 & L_5 \\ 1 & 2 & 2 \end{pmatrix} \right] \times [\langle d^{n-1} \nu_1 S_1 L_1 \{ |\, d^{n-2}(\nu_4 S_4 L_4) d S_1 L_1 \rangle \langle d^{n-2}(\nu_4 S_4 L_4) d S_1' L_1' |\, \} \, d^{n-1} \nu_1' S_1' L_1' \rangle]\}. \quad (15)$$

Using Biedenharn's identity (Appendix I 3 of ITS) and then (11.12) of ITS yields

$$\text{D.I.} \quad (d\text{–}p) = (n - 1)\delta(S_1S_1')\delta(S_2S_2') \sum_{\nu_4 S_4 L_4} \{R^0(dp, dp)\delta(L_1L_1') + R^2(dp, dp) \times [\exp \pi i(L_1 + L_1' + L_4 + L + 1)](\tfrac{1}{7}\tfrac{2})^{\frac{1}{2}}[(2L_1 + 1)(2L_1' + 1)]^{\frac{1}{2}} \overline{W}\begin{pmatrix} L_1 & L_1' & 2 \\ 2 & 2 & L_4 \end{pmatrix} \overline{W}\begin{pmatrix} L_1 & L_1' & 2 \\ 1 & 1 & L \end{pmatrix} \times [\langle d^{n-1} \nu_1 S_1 L_1 \{ |\, d^{n-2}(\nu_4 S_4 L_4) d S_1 L_1 \rangle \langle d^{n-2}(\nu_4 S_4 L_4) d S_1' L_1' |\, \} \, d^{n-1} \nu_1' S_1' L_1' \rangle]\}. \quad (16)$$

Due to the normalization of the fractional parentage coefficients we finally obtain for the direct part of the $d$–$p$ interaction

$$\text{D.I.} \quad (d\text{–}p) = (n - 1)F_0(dp)\delta(L_1L_1')\delta(S_1S_1')\delta(S_2S_2') + (n - 1) \sum_{\nu_4 S_4 L_4} \{[\exp \pi i(L_1 + L_1' + L_4 + L + 1)] \times 10[21(2L_1 + 1)(2L_1' + 1)]^{\frac{1}{2}} \overline{W}\begin{pmatrix} L_1 & L_1' & 2 \\ 2 & 2 & L_4 \end{pmatrix} \overline{W}\begin{pmatrix} L_1 & L_1' & 2 \\ 1 & 1 & L \end{pmatrix} \delta(S_1S_1')\delta(S_2S_2')F_2(dp) \times \langle d^{n-1} \nu_1 S_1 L_1 \{ |d^{n-2}(\nu_4 S_4 L_4) d S_1 L_1 \rangle \langle d^{n-2}(\nu_4 S_4 L_4) d S_1' L_1' | \} \, d^{n-1} \nu_1' S_1' L_1' \rangle\}, \quad (17)$$

where we let[3]

$$R^0(dp, dp) = F^0(dp) = F_0(dp), \quad R^2(dp, dp) = F^2(dp) = 35F_2(dp).$$

[3] G. Racah, Phys. Rev. **62**, 438 (1942).

### B. The Exchange Part of the $d$–$p$ Interaction

From (3) and (5) we obtain for the exchange part of the $d$–$p$ interaction

$$\begin{aligned}
\text{E.I.}\ (d\text{–}p) &= -(n-1)[\langle\{d_1\cdots d_{n-2}\cdots d_n\}s_{n-1}p_{n+1}|\, e^2/r_{n,n+1}\,|\{d_1\cdots d_{n-2}\cdots d_{n+1}\}s_{n-1}p_n\rangle] \\
&= (n-1)\sum_{\substack{\nu_4S_4L_4S_3S_5L_5\\ \nu_4'S_4'L_4'S_3'S_5'L_5'}}[\langle d^{n-2}(\nu_4S_4L_4)s_{n-1}(S_3L_4),\, d_np_{n+1}(S_5L_5),\, SL|\, e^2/r_{n,n+1} \\
&\quad\times |d^{n-2}(\nu_4'S_4'L_4')s_{n-1}(S_3'L_4'),\, p_nd_{n+1}(S_5'L_5'),\, SL\rangle[\exp \pi i(1+S_5'+L_5')] \\
&\quad\times \langle S_3L_4,\, d_np_{n+1}(S_5L_5),\, SL\,|\, S_3L_4d_n(S_2L_1)p_{n+1}SL\rangle\langle S_3'L_4',\, d_{n+1}p_n(S_5'L_5'),\, SL\,|\, S_3'L_4'd_{n+1}(S_2'L_1')p_nSL\rangle \\
&\quad\times \langle S_4L_4s_{n-1}(S_3L_4)d_nS_2L_1\,|\,S_4L_4d_n(S_1L_1)s_{n-1}S_2L_1\rangle\langle S_4'L_4's_{n-1}(S_3'L_4')d_{n+1}S_2'L_1'\,|\,S_4'L_4'd_{n+1}(S_1'L_1')s_{n-1}S_2'L_1'\rangle \\
&\quad\times \langle d^{n-2}(\nu_4'S_4'L_4')dS_1'L_1'\,|\}\, d^{n-1}\nu_1'S_1'L_1'\rangle\langle d^{n-1}\nu_1S_1L_1\,\{|\, d^{n-2}(\nu_4S_4L_4)dS_1L_1\rangle]. \qquad (18)
\end{aligned}$$

By similar reasoning, as for the direct part, we must insert into the above sum $\delta(\nu_4S_4L_4, \nu_4'S_4'L_4')$, $\delta(S_5L_5, S_5'L_5')$, and $\delta(S_3, S_3')$.

The net effect of the spin-recoupling coefficients is the same as for the direct part. However, here we must include $\exp(\pi iS_5)$ in the sum over $S_5$. Also since $S_3$ is the net spin of $(n-1)$ electrons and $S$ is the net spin of $(n+1)$ electrons, $(2S_3+2S)$ must be an even integer. Thus

$$\exp(2\pi iS) = \exp(2\pi iS_3).$$

Hence, as for (10), we can write

$$\begin{aligned}
&\sum_{S_3S_5}[\exp \pi iS_5]\text{R.C.}_{\text{spin}} \\
&\quad= \sum_{S_3S_5}\{[\exp \pi i(2S_3+S_5+S_1+S_1')](2S_3+1)(2S_5+1)[2S_1+1)(2S_1'+1)(2S_2+1)(2S_2'+1)]^{\frac12} \\
&\qquad\times \overline{W}\begin{pmatrix}\frac12 & S_4 & S_3\\ \frac12 & S_2 & S_1\end{pmatrix}\overline{W}\begin{pmatrix}\frac12 & S_4 & S_3\\ \frac12 & S_2' & S_1'\end{pmatrix}\overline{W}\begin{pmatrix}S_3 & \frac12 & S_2\\ \frac12 & S & S_5\end{pmatrix}\overline{W}\begin{pmatrix}S_3 & \frac12 & S_2'\\ \frac12 & S & S_5\end{pmatrix}.
\end{aligned}$$

Using (11.16) of ITS to perform the sum over $S_5$ and then (12.12) of ITS to sum over $S_3$ yields

$$\begin{aligned}
&\sum_{S_3S_5}[\exp \pi iS_5]\text{R.C.}_{\text{spin}} \\
&\quad= [\exp \pi i(S_1+S_1'+S_2+S_2')][(2S_1+1)(2S_1'+1)(2S_2+1)(2S_2'+1)]^{\frac12}X\begin{pmatrix}\frac12 & S_4 & S_1\\ S_2' & S_1' & \frac12\\ S & \frac12 & S_2\end{pmatrix}. \qquad (19)
\end{aligned}$$

The net contribution of the $L$ recoupling coefficients is given by (12). Now by Eqs. (16.15) and (16.17) of ITS

$$\langle d_np_{n+1}S_5L_5|\, e^2/r_{n,n+1}\,|p_nd_{n+1}S_5L_5\rangle = \sum_k R^k(dp, pd)[\exp \pi i(L_5+k)](1||C^k||2)(2||C^k||1)\overline{W}\begin{pmatrix}2 & 1 & k\\ 2 & 1 & L_5\end{pmatrix}.$$

Using (14.12) of ITS, the fact that $(1+k+2)$ in the reduced matrix elements must be even, and the triangular condition, (7.4) of ITS, for each triad of the $\overline{W}$ coefficient, we obtain

$$\begin{aligned}
&\langle d_np_{n+1}S_5L_5|\, e^2/r_{n,n+1}\,|p_nd_{n+1}S_5L_5\rangle \\
&\qquad= [\exp \pi i(L_5+1)]\left[2R^1(dp, pd)\overline{W}\begin{pmatrix}2 & 1 & 1\\ 2 & 1 & L_5\end{pmatrix} + \tfrac{9}{7}R^3(dp, pd)\overline{W}\begin{pmatrix}2 & 1 & 3\\ 2 & 1 & L_5\end{pmatrix}\right]. \qquad (20)
\end{aligned}$$

Inserting (12), (19), and (20) into (18) yields for the exchange part of the $d$–$p$ interaction

$$\begin{aligned}
\text{E.I.}\ (d\text{–}p) &= (n-1)\sum_{\nu_4S_4L_4L_5}\left\{\left[2R^1(dp, pd)\overline{W}\begin{pmatrix}2 & 1 & 1\\ 2 & 1 & L_5\end{pmatrix} + \tfrac{9}{7}R^3(dp, pd)\overline{W}\begin{pmatrix}2 & 1 & 3\\ 2 & 1 & L_5\end{pmatrix}\right]\right. \\
&\quad\times (2L_5+1)[(2L_1+1)(2L_1'+1)]^{\frac12}\overline{W}\begin{pmatrix}L_4 & 2 & L_1\\ 1 & L & L_5\end{pmatrix}\overline{W}\begin{pmatrix}L_4 & 2 & L_1'\\ 1 & L & L_5\end{pmatrix}[\exp \pi i(S_1+S_1'+S_2+S_2')] \\
&\quad\times [(2S_1+1)(2S_1'+1)(2S_2+1)(2S_2'+1)]^{\frac12}X\begin{pmatrix}\frac12 & S_4 & S_1\\ S_2' & S_1' & \frac12\\ S & \frac12 & S_2\end{pmatrix} \\
&\quad\times \left.\langle d^{n-2}(\nu_4S_4L_4)dS_1'L_1'\,|\}\, d^{n-1}\nu_1'S_1'L_1'\rangle\langle d^{n-1}\nu_1S_1L_1\,\{|\, d^{n-2}(\nu_4S_4L_4)dS_1L_1\rangle\right\}.
\end{aligned}$$

Using (12.12) of ITS to perform the sum over $L_5$ finally yields for the exchange part of the $d$–$p$ interaction

$$\text{E.I.}\quad (d\text{–}p) = (n-1)\sum_{\nu_4 S_4 L_4}\{[\exp \pi i(S_1 + S_1' + S_2 + S_2')][(2S_1+1)(2S_1'+1)(2S_2+1)(2S_2'+1)]^{\frac{1}{2}}$$
$$\times X\begin{pmatrix} \frac{1}{2} & S_4 & S_1 \\ S_2' & S_1' & \frac{1}{2} \\ S & \frac{1}{2} & S_2 \end{pmatrix}\left[30X\begin{pmatrix} 1 & L & L_1 \\ 2 & L_1' & L_4 \\ 1 & 1 & 2 \end{pmatrix}G_1(dp) + 105X\begin{pmatrix} 1 & L & L_1 \\ 2 & L_1' & L_4 \\ 3 & 1 & 2 \end{pmatrix}G_3(dp)\right]$$
$$\times \langle d^{n-1}\nu_1 S_1 L_1\{|d^{n-2}(\nu_4 S_4 L_4)dS_1 L_1\rangle\langle d^{n-2}(\nu_4 S_4 L_4)dS_1' L_1'|\}d^{n-1}\nu_1' S_1' L_1'\rangle\}, \tag{21}$$

where $G_1(dp) = \frac{1}{15}R^1(dp, pd)$ and $G_3(dp) = \frac{3}{245}R^3(dp, pd)$.[3]

The net interaction $d$–$p$ is then the sum of (17) and (21).

### 3. THE DIRECT PART OF THE $s$–$p$ INTERACTION

From (1) and (2) the direct part of the $s$–$p$ interaction is given by

$$\text{D.I.}\quad (s\text{–}p) = \langle\{d_1\cdots d_{n-1}\}s_n p_{n+1}|\, e^2/r_{n,n+1}\,|\{d_1\cdots d_{n-1}\}s_n p_{n+1}\rangle$$
$$= \langle d^{n-1}(\nu_1 S_1 L_1)s_n(S_2 L_1)p_{n+1}SL|\, e^2/r_{n,n+1}\,|d^{n-1}(\nu_1' S_1' L_1')s_n(S_2' L_1')p_{n+1}SL\rangle.$$

Since the interaction is between the electrons $n$ and $n+1$, the quantum numbers of the first $(n-1)$ electrons must be the same on both sides. Hence

$$\nu_1 S_1 L_1 \equiv \nu_1' S_1' L_1'.$$

Now,

$$\text{D.I.}\quad (s\text{–}p) = \langle d^{n-1}(\nu_1 S_1 L_1)s_n(S_2 L_1)p_{n+1}SL|\, e^2/r_{n,n+1}\,|d^{n-1}(\nu_1 S_1 L_1)s_n(S_2' L_1)p_{n+1}SL\rangle$$
$$= \sum_{S_3 S_3'}[\langle d^{n-1}(\nu_1 S_1 L_1),\, s_n p_{n+1}(S_3 L_3),\, SL|\, e^2/r_{n,n+1}\,|d^{n-1}(\nu_1 S_1 L_1),\, s_n p_{n+1}(S_3' L_3),\, SL\rangle$$
$$\times \langle S_1 L_1,\, s_n p_{n+1}(S_3 L_3),\, SL \mid S_1 L_1 s_n(S_2 L_1)p_{n+1}SL\rangle$$
$$\times \langle S_1 L_1,\, s_n p_{n+1}(S_3' L_3),\, SL \mid S_1 L_1 s_n(S_2' L_1)p_{n+1}SL\rangle], \tag{22}$$

where $L_3$ necessarily equals 1.

Then from Eqs. (16.15) and (16.17) of ITS we have for the matrix interaction element

$$\langle s_n p_{n+1}S_3 L_3|\, e^2/r_{n,n+1}\,|s_n p_{n+1}S_3' L_3\rangle$$
$$= \delta(S_3, S_3')\sum_k R^k(sp, sp)[\exp \pi i(1 + L_3 + k)](0||C^k||0)(1||C^k||1)\overline{W}\begin{pmatrix} 0 & 0 & k \\ 1 & 1 & L_3 \end{pmatrix}.$$

Using Eqs. (14.12), (11.12), and the triangle condition, (7.4) of ITS, for the triad $(0\ 0\ k)$ of the $\overline{W}$ coefficient, yields

$$\langle s_n p_{n+1}S_3 L_3|\, e^2/r_{n,n+1}\,|s_n p_{n+1}S_3' L_3\rangle = R^0(sp, sp)\delta(S_3 S_3'). \tag{23}$$

Using the fact that $S_3'$ equals $S_3$, we obtain from (11.8) of ITS that the net contribution of the $S$-recoupling coefficients becomes

$$\text{R.C.}_{\text{spin}} = \langle S_1, \tfrac{1}{2}\,\tfrac{1}{2}(S_3), S \mid S_1\,\tfrac{1}{2}(S_2)\tfrac{1}{2}S\rangle\langle S_1, \tfrac{1}{2}\,\tfrac{1}{2}(S_3), S \mid S_1\,\tfrac{1}{2}(S_2')\tfrac{1}{2}S\rangle$$
$$= [\exp \pi i(2S_1 + 2S)](2S_3+1)[(2S_2 + 1(2S_2'+1)]^{\frac{1}{2}}\overline{W}\begin{pmatrix} S_1 & \frac{1}{2} & S_2 \\ \frac{1}{2} & S & S_3 \end{pmatrix}\overline{W}\begin{pmatrix} S_1 & \frac{1}{2} & S_2' \\ \frac{1}{2} & S & S_3 \end{pmatrix}. \tag{24}$$

Then we need to sum only over $S_3$ and using (11.15) of ITS obtain

$$\sum_{S_3}\text{R.C.}_{\text{spin}} = [\exp \pi i(2S_1 + 2S)]\delta(S_2, S_2') = \delta(S_2, S_2'), \tag{25}$$

since $(2S_1 + 2S)$ is an even integer.

Also from Eqs. (11.8) and (11.12) of ITS we obtain

$$\text{R.C.}_{\text{a.m.}} = \langle L_1, 01(1), L \mid L_1 0(L_1)1L\rangle\langle L_1, 01(1), L \mid L_1 0(L_1)1L\rangle = 1. \tag{26}$$

Thus inserting (23), (25), and (26) into (22), we obtain for the direct part of the $s$–$p$ interaction

$$\text{D.I.}\quad (s\text{–}p) = \delta(\nu_1 S_1 L_1, \nu_1' S_1' L_1')\delta(S_2, S_2')R^0(sp, ps) = \delta(\nu_1 S_1 L_1, \nu_1' S_1' L_1')\delta(S_2, S_2')F_0(sp). \tag{27}$$

**The Exchange Part of the $s$–$p$ Interaction**

From (1) and (2) the exchange part of the $s$–$p$ interaction is given by

$$\begin{aligned}\text{E.I.}\quad (s\text{–}p) &= -\langle\{d_1 \cdots d_{n-1}\}s_n p_{n+1}|\ e^2/r_{n,n+1}\ |\{d_1 \cdots d_{n-1}\}s_{n+1}p_n\rangle \\ &= -\langle d^{n-1}(\nu_1 S_1 L_1)s_n(S_2 L_1)p_{n+1}SL|\ e^2/r_{n,n+1}\ |d^{n-1}(\nu_1 S_1 L_1)s_{n+1}(S_2' L_1')p_n SL\rangle.\end{aligned} \tag{28}$$

Coupling the $s$ and $p$ electrons on both sides of (28) and then using (7.11) of ITS, yields

$$\begin{aligned}\text{E.I.}\quad (s\text{–}p) = \sum_{S_3 S_3'}\{&[\exp \pi i(1 - S_3 - L_3)]\langle s_n p_{n+1} S_3 L_3|\ e^2/r_{n,n+1}\ |p_n s_{n+1} S_3' L_3\rangle \\ &\times \langle S_1 L_1, s_n p_{n+1}(S_3 L_3), SL\ |\ S_1 L_1 s_n(S_2 L_1)p_{n+1}SL\rangle \\ &\times \langle S_1 L_1, s_{n+1}p_n(S_3' L_3), SL\ |\ S_1 L_1 s_{n+1}(S_2' L_1)p_n SL\rangle\}.\end{aligned} \tag{29}$$

Using Eqs. (16.15) and (16.17) of ITS, we obtain

$$\begin{aligned}&\langle s_n p_{n+1} S_3 L_3|\ e^2/r_{n,n+1}\ |p_n s_{n+1} S_3' L_3\rangle \\ &\qquad = \delta(S_3 S_3')\sum_k R^k(sp, ps)[\exp \pi i(1 + k)](0|\ |C^k|\ |1)(1|\ |C^k|\ |0)\overline{W}\begin{pmatrix}0 & 1 & k\\ 0 & 1 & 1\end{pmatrix}.\end{aligned} \tag{30}$$

By the triangle rule, (7.4) of ITS, for the triad (0 $k$ 1) of the $\overline{W}$ coefficient, $k$ can only have the value 1 and thus

$$\langle s_n p_{n+1} S_3 L_3|\ e^2/r_{n,n+1}\ |p_n s_{n+1} S_3' L_3\rangle = \delta(S_3 S_3')[R^1(sp, ps)/3]. \tag{31}$$

Since $S_3$ is a whole number and $L_3$ equals 1,

$$\exp \pi i(1 - S_3 - L_3) = \exp \pi i S_3.$$

The net contribution of the $S$-recoupling coefficients is the same as for the direct part and hence is given by (24).

Then since $(2S_1 + 2S)$ is an even integer, we have from (24)

$$\begin{aligned}\sum_{S_3}[\exp \pi i S_3]\text{R.C.}_{\text{spin}} &= \sum_{S_3}[\exp \pi i S_3](2S_3 + 1)[(2S_2 + 1)(2S_2' + 1)]^{\frac{1}{2}}\overline{W}\begin{pmatrix}S_1 & \frac{1}{2} & S_2\\ \frac{1}{2} & S & S_3\end{pmatrix}\overline{W}\begin{pmatrix}S_1 & \frac{1}{2} & S_2'\\ \frac{1}{2} & S & S_3\end{pmatrix} \\ &= [\exp \pi i(S_2 + S_2')][(2S_2 + 1)(2S_2' + 1)]^{\frac{1}{2}}\overline{W}\begin{pmatrix}S_1 & \frac{1}{2} & S_2\\ S & \frac{1}{2} & S_2'\end{pmatrix},\end{aligned} \tag{32}$$

using (11.16) of ITS for the sum over $S_3$.

Inserting (26) for the net contribution of the $L$-coupling coefficients (31) and (32) into (29), we finally obtain

$$\text{E.I.}\quad (s\text{–}p) = [\exp \pi i(S_2 + S_2')][(2S_2 + 1)(2S_2' + 1)]^{\frac{1}{2}}\overline{W}\begin{pmatrix}S_1 & \frac{1}{2} & S_2\\ S & \frac{1}{2} & S_2'\end{pmatrix}\frac{R^1(sp, ps)}{3}. \tag{33}$$

Then the net interaction $s$–$p$, T.I. ($s$–$p$), is obtained by adding (27) and (33),

$$\begin{aligned}\text{T.I.}\quad (s\text{–}p) &= \delta(\nu_1 S_1 L_1, \nu_1' S_1' L_1') \\ &\times \left\{F_0(sp)\delta(S_2, S_2') + [\exp \pi i(S_2 + S_2')][(2S_2 + 1)(2S_2' + 1)]^{\frac{1}{2}}\overline{W}\begin{pmatrix}S_1 & \frac{1}{2} & S_2\\ S & \frac{1}{2} & S_2'\end{pmatrix}G_{ps}\right\},\end{aligned} \tag{34}$$

where[4] $G_{ps} = [R^1(sp, ps)/3]$.

## 4. THE $d$–$s$ INTERACTION

Although the $d$–$s$ interaction can be obtained in the same fashion as the $d$–$p$ and $s$–$p$ interactions, it is much simpler to use the Dirac vector model.[5]

[4] E. U. Condon and G. H. Shortley, *The Theory of Atomic Spectra* (Cambridge University Press, Cambridge, England, 1935), Chap. VII.
[5] P. A. M. Dirac, *The Principles of Quantum Mechanics* (Oxford University Press, Oxford, England, 1947), Chap. XI.

By Eqs. (16.22) and (16.23) of ITS, the direct interaction for the electrons $d$ and $s$ is given by $R^0(ds, ds)$. Hence the direct part of the interaction $d$–$s$ for the configuration $d^{n-1}sp$ is given by

$$\text{D.I.} \quad (d\text{–}s) = (n-1)R^0(ds, ds) = (n-1)F^0(ds) = (n-1)F_0(ds). \tag{35}$$

From Eqs. (16.22), (16.23), and (16.24) of ITS, the exchange interaction between the two electrons $d$ and $s$ is given by

$$-\tfrac{1}{2}[1 + 4s_d \cdot s_s][R^2(ds, sd)/5] = -\tfrac{1}{2}[1 + 4s_d \cdot s_s]G_{ds},$$

where[4]

$$G_{ds} = [R^2(ds, sd)/5].$$

Hence for $(n-1)$ electrons $d$, the $d$–$s$ interaction will be given by

$$\begin{aligned}\text{E.I.} \quad (d\text{–}s) &= -\tfrac{1}{2}(n-1)G_{ds} - \tfrac{1}{2}[4(s_{d_1} + \cdots + s_{d_{n-1}}) \cdot s_s]G_{ds} = -\tfrac{1}{2}[n - 1 + 4S_1 \cdot s_s]G_{ds} \\ &= -\tfrac{1}{2}[n - 1 + 2S_2(S_2+1) - 2S_1(S_1+1) - \tfrac{3}{2}]G_{ds}\delta(\nu_1 S_1 L_1, \nu_1' S_1' L_1')\delta(S_2, S_2'),\end{aligned} \tag{36}$$

where as usual

$$\langle d^{n-1}sp| \equiv \langle d^{n-1}(\nu_1 S_1 L_1)s(S_2 L_1)pSL|, \quad |d^{n-1}sp\rangle \equiv |d^{n-1}(\nu_1' S_1' L_1')s(S_2' L_1')pSL\rangle.$$

Then the net interaction $d$–$s$, T.I. $(d$–$s)$, is obtained by adding (35) and (36),

$$\begin{aligned}\text{T.I.} \quad (d\text{–}s) &= \delta(\nu_1 S_1 L_1, \nu_1' S_1' L_1')\delta(S_2, S_2') \\ &\quad \times \{(n-1)F_0(ds) - \tfrac{1}{2}[n - 1 + 2S_2(S_2+1) - 2S_1(S_1+1) - \tfrac{3}{2}]G_{ds}\}.\end{aligned} \tag{37}$$

## 5. $d$–$d$ INTERACTION

The $d$–$d$ interaction of the configuration $d^{n-1}sp$ is the same as the $d$–$d$ interaction of the configuration $d^{n-1}$, and hence given by the well-known results of Racah.[3]

Finally, it should be noted that starting from (1) and replacing $d$ by $l$, $s$ by $l'$ and $p$ by $l''$, one can obtain the Coulomb interactions for the configuration $l^{n-1}l'l''$. The latter results are naturally more cumbersome since no explicit use is made of the fact that one electron has zero angular momentum.

# Some Inequalities Involving Traces of Operators*

C. L. MEHTA†
*Department of Physics and Astronomy, University of Rochester, Rochester, New York*

(Received 7 September 1967)

We prove that for arbitrary completely continuous operators $A_1, A_2, \cdots, A_n$ and for positive numbers $p_1, p_2, \cdots, p_n$ with $\sum_{k=1}^{n} p_k^{-1} = 1$, the inequalities,

$$|\mathrm{Tr}\,(A_1 A_2 \cdots A_n)| \leq \prod_{k=1}^{n} [\mathrm{Tr}\,(A_k^\dagger A_k)^{\frac{1}{2}p_k}]^{p_k^{-1}}$$

$$|\mathrm{Tr}\exp(A_1 + A_2 + \cdots + A_n)| \leq \prod_{k=1}^{n} [\mathrm{Tr}\exp\{\tfrac{1}{2}p_k(A_k^\dagger + A_k)\}]^{p_k^{-1}}$$

hold. Further if $a_1, a_2, \cdots, a_n$ are the annihilation operators of an $N$-dimensional harmonic oscillator, $m$ and $n$ are any positive integers, and $\rho$ is a nonnegative definite operator, we prove the inequality

$$|\mathrm{Tr}\,(\rho a_{i_1}^\dagger \cdots a_{i_m}^\dagger a_{j_1} \cdots a_{j_n})| \leq \prod_{k=1}^{m} [\mathrm{Tr}\,\rho(a_{i_k}^\dagger a_{i_k})^m]^{1/2m} \prod_{l=1}^{n} [\mathrm{Tr}\,\rho(a_{j_l}^\dagger a_{j_l})^n]^{1/2n}.$$

Some consequences of these inequalities, related results, and some applications to correlation functions of the quantized electromagnetic field are discussed.

## 1. INTRODUCTION

It is well known that for any two arbitrary operators $A$ and $B$, an inequality analogous to the Cauchy–Schwarz inequality holds, i.e., that

$$|\mathrm{Tr}\,(AB)|^2 \leq \mathrm{Tr}\,(A^\dagger A)\,\mathrm{Tr}\,(B^\dagger B), \tag{1.1}$$

where dagger denotes Hermitian adjoint operation. This result has recently been generalized by Thompson[1] for Hermitian, nonnegative definite operators, and an inequality analogous to Hölder inequality has been proved. Thus if $A$ and $B$ are Hermitian, nonnegative definite operators, then one has

$$\mathrm{Tr}\,(AB) \leq [\mathrm{Tr}\,A^p]^{1/p}[\mathrm{Tr}\,B^q]^{1/q}, \tag{1.2}$$

where $p$ and $q$ are positive numbers and $p^{-1} + q^{-1} = 1$. The proof given in Ref. 1 does not state conditions under which the relation (1.2) reduces to an equality. However, it can readily be deduced from the discussion given there that this happens if and only if $A^p$ is a constant multiple of $B^q$.

In this paper we establish a number of inequalities which are analogous to, and may be considered as, generalizations of (1.2). Throughout this paper it is understood that the operators $A$, $B$, etc. belong to a class such that the traces appearing in (1.1), (1.2), and elsewhere in this paper are all well defined. Our results are summarized in the following theorems.

*Theorem I:* Let $A_1, A_2, \cdots, A_n$ be a set of $n$ ($\geq 1$) completely continuous operators in a Hilbert space and let $p_1, p_2, \cdots, p_n$ be positive numbers with

$$p_1^{-1} + p_2^{-1} + \cdots + p_n^{-1} = 1.$$

Then

$$|\mathrm{Tr}\,(A_1 A_2 \cdots A_n)| \leq \prod_{k=1}^{n} [\mathrm{Tr}\,(A_k^\dagger A_k)^{\frac{1}{2}p_k}]^{p_k^{-1}}, \tag{1.3}$$

where positive roots are to be taken of the operators $A_k^\dagger A_k$ on the right-hand side.

*Theorem II:* Let $A_1, A_2, \cdots, A_n$ and $p_1, p_2, \cdots, p_n$ be as defined in Theorem I. Then

$$|\mathrm{Tr}\exp(A_1 A_2 \cdots A_n)| \leq \prod_{k=1}^{n} [\mathrm{Tr}\exp\{\tfrac{1}{2}p_k(A_k^\dagger + A_k)\}]^{p_k^{-1}}. \tag{1.4}$$

*Theorem III:* Consider a harmonic oscillator with $N$ degrees of freedom. Let $a_1, a_2, \cdots, a_N$ be the annihilation operators and $a_1^\dagger, a_2^\dagger, \cdots, a_N^\dagger$ be the creation operators; they satisfy the commutation relations

$$[a_i; a_j] = 0, \tag{1.5a}$$

$$[a_i; a_j^\dagger] = \delta_{ij}. \tag{1.5b}$$

Further let $\rho$ be a Hermitian nonnegative-definite operator. Then for arbitrary positive integral values

* Research supported in part by the U.S. Army Research Office (Durham).

† Some of the research reported in this paper was carried out during the author's stay as Visiting Professor of Physics at Université Laval, Québec 10, Canada.

[1] C. J. Thompson, J. Math. Phys. **6**, 1812 (1965). The inequality (1.2) is proved in this reference for the case when $A$ and $B$ are $n \times n$ positive-definite Hermitian matrices. The same proof is valid even in the case when $A$ and $B$ are any Hermitian, nonnegative-definite completely continuous operators in a Hilbert space.

of $m$ and $n$ the inequality

$$|\mathrm{Tr}\,(\rho a_{i_1}^\dagger \cdots a_{i_m}^\dagger a_{j_1} \cdots a_{j_n})| \leq \prod_{k=1}^{m} [\mathrm{Tr}\,\{\rho(a_{i_k}^\dagger a_{i_k})^m\}]^{1/2m} \prod_{l=1}^{n} [\mathrm{Tr}\,\{\rho(a_{j_l}^\dagger a_{j_l})^n\}]^{1/2n} \tag{1.6}$$

holds. Here each of the indices $i_1, \cdots, i_m, j_1, \cdots, j_n$ may take on any of the values $1, 2, \cdots, N$. Equality in (1.6) holds either as a trivial identity ($0 = 0$) or only when $m = n = 1$ and $i_1 = j_1$.

In Sec. 2 we prove Theorems I and II and also discuss a generalization of (1.2). Theorem III is proved in Sec. 3. A conjecture stating that an inequality of the form (1.6) also holds for arbitrary linear combinations of the annihilation operators is proposed and is proved in a special case. Some other related inequalities and their application to coherence functions of the quantized electromagnetic field are briefly discussed in Sec. 4.

## 2. PROOF OF THEOREMS I AND II

To prove Theorem I, we make use of a result due to Fan[2]: Let $A_1, A_2, \cdots, A_n$ be a set of $n$ ($\geq 1$) completely continuous operators in a Hilbert space $\mathcal{H}$. For each $k(1 \leq k \leq n)$, let $\alpha_1^{(k)}, \alpha_2^{(k)} \cdots$ be the eigenvalues of the operator $A_k^\dagger A_k$ arranged in a nonincreasing order. Then, for any positive integer $N$,

$$\left| \sum_{i=1}^{N} \langle x_i | A_1 A_2 \cdots A_n | x_i \rangle \right| \leq \sum_{i=1}^{N} [\alpha_i^{(1)} \alpha_i^{(2)} \cdots \alpha_i^{(n)}]^{\frac{1}{2}}, \tag{2.1}$$

where $|x_i\rangle$, $i = 1, 2, \cdots, N$ are orthonormal vectors in $\mathcal{H}$.

Using Hölder inequality[3] for positive numbers in (2.1) we can write

$$\sum_{i=1}^{N} \langle x_i | A_1 A_2 \cdots A_n | x_i \rangle \leq \prod_{k=1}^{n} \left[ \sum_{i=1}^{N} (\alpha_i^{(k)})^{\frac{1}{2} p_k} \right]^{p_k^{-1}}, \tag{2.2}$$

where $p_1, p_2, \cdots, p_n$ are positive numbers and $p_1^{-1} + p_2^{-1} + \cdots + p_n^{-1} = 1$.

If we now extend the summation over $i$ in (2.2) to include all the eigenvalues of the operators $A_k^\dagger A_k$, we obtain the inequality expressed by Theorem I,

$$|\mathrm{Tr}\,(A_1 A_2 \cdots A_n)| \leq \prod_{k=1}^{n} [\mathrm{Tr}\,(A_k^\dagger A_k)^{\frac{1}{2} p_k}]^{p_k^{-1}}. \tag{2.3}$$

It may be noted that when $A_1, A_2, \cdots, A_n$ are Hermitian nonnegative-definite operators, the inequality (2.3) reduces to

$$|\mathrm{Tr}\,(A_1 A_2 \cdots A_n)| \leq \prod_{k=1}^{n} [\mathrm{Tr}\,(A_k)^{p_k}]^{p_k^{-1}}, \tag{2.4}$$

and thus it is a generalization of (1.2). Relation (2.4) reduces to an equality if and only if either all the operators $A_1^{p_1}, A_2^{p_2}, \cdots, A_n^{p_n}$ differ from one another by constant multiplication factors, or one of the operators is identically zero.

We prove Theorem II first in the case when $A_1, A_2, \cdots, A_n$ are Hermitian operators. In this case the proof is based on the inequality (1.2) and the relation[4]

$$\mathrm{Tr}\,[\exp(A + B)] \leq \mathrm{Tr}\,[e^A e^B], \tag{2.5}$$

which holds for arbitrary Hermitian operators $A$ and $B$.

From (2.5) and (1.2) we obtain

$$\begin{aligned} &\mathrm{Tr}\,\{\exp[A_1 + A_2 + \cdots + A_n]\} \\ &\leq \mathrm{Tr}\,[e^{A_1} \exp(A_2 + A_3 + \cdots + A_n)] \\ &\leq [\mathrm{Tr}\, e^{p_1 A_1}]^{p_1^{-1}} \\ &\quad \times [\mathrm{Tr}\,\{\exp \alpha (A_2 + A_3 + \cdots + A_n)\}]^{\alpha^{-1}}, \end{aligned} \tag{2.6}$$

where $p_1$ and $\alpha$ are positive numbers with $p_1^{-1} + \alpha^{-1} = 1$. A repeated use of (2.5) and (1.2) in this manner gives the required result

$$\mathrm{Tr}\,[\exp(A_1 + A_2 + \cdots + A_n)] \leq \prod_{k=1}^{n} [\mathrm{Tr} \exp p_k A_k]^{p_k^{-1}}. \tag{2.7}$$

It may be noted that the equality in (2.7) holds if and only if the difference between any two of the operators $p_1 A_1, p_2 A_2, \cdots, p_n A_n$ is a constant multiple of the identity operator.

Next we consider the general case when $A_1, A_2, \cdots, A_n$ are arbitrary completely continuum operators. In this case we make use of another result due to Fan[5]: Let $A$ be a completely continuous operator in $\mathcal{H}$ and let the eigenvalue $\lambda_i$ and $\rho_i$ of $A$ and $\frac{1}{2}(A^\dagger + A)$, respectively, be so arranged that

$$\Re\lambda_1 \geq \Re\lambda_2 \geq \cdots; \quad \rho_1 \geq \rho_2 \geq \rho_3 \cdots,$$

where $\Re$ denotes the real part. Then, for any positive integer $N$,

$$\sum_{i=1}^{N} \Re\lambda_i \leq \sum_{i=1}^{N} \rho_i. \tag{2.8}$$

Equality in (2.8) holds if and only if $A$ is a normal operator,

From (2.8), it follows that

$$\left| \exp\left[ \sum_{i=1}^{N} \lambda_i \right] \right| = \exp\left[ \sum_{i=1}^{N} \Re\lambda_i \right] \leq \exp\left[ \sum_{i=1}^{N} \rho_i \right], \tag{2.9}$$

[2] K. Fan, Proc. Natl. Acad. Sci. U.S., 37, 760 (1951) [cf. Eq. (4)]. Author wishes to thank Professor G. S. Mudholkar and the referee for bringing this paper to his notice. This made possible the generalization of Theorem I in the present form.

[3] See for example, E. F. Beckenbach and R. Bellman, *Inequalities* (Springer-Verlag, Berlin, 1961), p. 20.

[4] See for example, S. Golden, Phys. Rev. 137, B1127 (1965); C. J. Thompson, Ref. 1.

[5] K. Fan, Proc. Natl. Acad. Sci. U.S. 36, 31 (1950) [cf. inequality (12)].

and hence on extending the summation over $i$ to include all the eigenvalues of $A$, we obtain

$$|\mathrm{Tr}\, e^{A}| \leq \mathrm{Tr} \exp [\tfrac{1}{2}(A^{\dagger} + A)]. \tag{2.10}$$

On setting $A = A_1 + A_2 + \cdots + A_n$ in (2.10), it follows that

$$|\mathrm{Tr}\, [\exp (A_1 + A_2 + \cdots + A_n)]| \leq \mathrm{Tr}\, [\exp \{\tfrac{1}{2}(A_1^{\dagger} + A_1) + \tfrac{1}{2}(A_2^{\dagger} + A_2) + \cdots + \tfrac{1}{2}(A_n^{\dagger} + A_n)\}]. \tag{2.11}$$

Theorem II [inequality (1.4)] now follows from (2.7) and (2.11).

We now wish to make a few remarks concerning the inequality (1.2):

If $\rho$ is a Hermitian nonnegative-definite operator, which for convenience is normaled[6] so that $\mathrm{Tr}\, \rho = 1$, then for arbitrary operators $A$ and $B$ [cf. (1.1)],

$$|\mathrm{Tr}\, (\rho AB)|^2 \leq \mathrm{Tr}\, (\rho A^{\dagger} A)\, \mathrm{Tr}\, (\rho B^{\dagger} B). \tag{2.12}$$

A generalization of (2.12) analogous to (1.2), namely the inequality

$$\mathrm{Tr}\, (\rho AB) \leq [\mathrm{Tr}\, \rho A^{p}]^{1/p} [\mathrm{Tr}\, \rho B^{q}]^{1/q} \tag{2.13}$$

also holds whenever $\rho$, $A$, $B$ are Hermitian nonnegative definite, and $\rho$ commutes with both $A$ and $B$. Here $p$ and $q$ are positive numbers with $p^{-1} + q^{-1} = 1$. Inequality (2.13) follows by replacing $A$ and $B$ in (1.2) by $\rho^{1/p} A$ and $\rho^{1/q} B$, respectively. In a similar manner one can show that if $\rho$ commutes with all the operators $A_1, \cdots, A_n$, of Theorem I, then

$$|\mathrm{Tr}\, (\rho A_1 \cdots A_n)| \leq \prod_{k=1}^{n} [\mathrm{Tr}\, \rho (A_k^{\dagger} A_k)^{\frac{1}{2} p_k}]^{p_k^{-1}}, \tag{2.14}$$

where $p_1, \cdots, p_n$ are positive numbers with

$$p_1^{-1} + \cdots + p_n^{-1} = 1.$$

It should be noted that if $\rho$ does not commute with both $A$ and $B$ then (2.13) is in general not true. [Similarly (2.14) is not true in general.] A counter example is provided by

$$\rho = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad p = \tfrac{3}{2}, \quad q = 3. \tag{2.15}$$

Here $\mathrm{Tr}\, (\rho AB) = 4$, whereas $[\mathrm{Tr}\, \rho A^{\frac{3}{2}}]^{\frac{2}{3}} [\mathrm{Tr}\, \rho B^{3}]^{\frac{1}{3}} = 63^{\frac{1}{3}} < 4$.

Another situation where the inequality similar to (2.13) and in fact similar to the more general result (2.14) holds is discussed in Sec. 3 [cf. (3.1)].

[6] $\rho$, for example, may be taken to be the density operator describing the statistical state of a quantum mechanical system. Then $\mathrm{Tr}\, (\rho \mathcal{O})$ is the expectation value of the operator $\mathcal{O}$ in the state described by the density operator $\rho$.

## 3. PROOF OF THEOREM III AND SOME RELATED RESULTS

In order to prove Theorem III, we first observe that it is sufficient to prove the inequality

$$\mathrm{Tr}\, [\rho a_{i_1}^{\dagger} \cdots a_{i_l}^{\dagger} a_{i_1} \cdots a_{i_l}] \leq \prod_{k=1}^{l} [\mathrm{Tr}\, \{\rho (a_{i_k}^{\dagger} a_{i_k})^{l}\}]^{1/l}. \tag{3.1}$$

Here $l$ is an arbitrary positive integer and $\rho$, $\{a_i\}$, $\{a_i^{\dagger}\}$ $(i = 1, 2, \cdots, N)$ are those as described in Theorem III (Sec. 1).

To see that (3.1) implies (1.6), we make use of the Schwarz inequality (2.12) and obtain

$$|\mathrm{Tr}\, [\rho a_{i_1}^{\dagger} \cdots a_{i_m}^{\dagger} a_{j_1} \cdots a_{j_n}]|^2 \leq \mathrm{Tr}\, [\rho a_{i_1}^{\dagger} \cdots a_{i_m}^{\dagger} a_{i_m} \cdots a_{i_1}] \times \mathrm{Tr}\, [\rho a_{j_1}^{\dagger} \cdots a_{j_n}^{\dagger} a_{j_n} \cdots a_{j_1}]. \tag{3.2}$$

From (3.1), (1.5a), and (3.2) we obtain the required result (1.6).

We now proceed to prove the inequality (3.1). Let us suppose that in the expression $a_{i_1} a_{i_2} \cdots a_{i_l}$ the operator $a_i$ occurs $k_1$ times, $a_2$ occurs $k_2$ times etc., so that $k_1, k_2, \cdots, k_N$ are all nonnegative integers and

$$k_1 + k_2 + \cdots + k_n = l. \tag{3.3}$$

Using the commutation relations (1.5), we can then write

$$a_{i_1}^{\dagger} \cdots a_{i_l}^{\dagger} a_{i_1} \cdots a_{i_l} = (a_1^{\dagger k_1} a_1^{k_1}) \cdots (a_N^{\dagger k_N} a_N^{k_N}). \tag{3.4}$$

Further, if $a$ denotes any of the operators $a_1, \cdots, a_N$, we have the relation

$$a^{\dagger k} a^{k} = a^{\dagger} a (a^{\dagger} a - 1) \cdots (a^{\dagger} a - k + 1), \tag{3.5}$$

which can easily be established by the method of induction, and use of the relation $[a, a^{\dagger}] = 1$. On substituting (3.5) in (3.4), we obtain the relation

$$a_{i_1}^{\dagger} \cdots a_{i_l}^{\dagger} a_{i_1} \cdots a_{i_l} = \prod_{j=1}^{N} a_j^{\dagger} a_j (a_j^{\dagger} a_j - 1) \cdots (a_j^{\dagger} a_j - k_j + 1). \tag{3.6}$$

Next we represent the density operator $\rho$ in the number representation

$$\rho = \sum_{\{n_j\}} \sum_{\{m_j\}} \rho_{\{n_j\},\{m_j\}} |\{n_j\}\rangle\langle\{m_j\}|, \tag{3.7}$$

where

$$|\{n_j\}\rangle = \prod_{j=1}^{N} |n_j\rangle, \tag{3.8}$$

and $|n_j\rangle$ are the orthonormalized eigenstates of the operator $a_j^{\dagger} a_j$:

$$a_j^{\dagger} a_j |n_j\rangle = n_j |n_j\rangle, \tag{3.9}$$

$$\langle n_j | m_j\rangle = \delta_{n_j m_j}. \tag{3.10}$$

The summations on the right-hand side of (3.7) runs

over all nonnegative integral values of $n_1, \cdots, n_N$, $m_1, \cdots, m_N$.

From Eqs. (3.6)–(3.10) we obtain the following expression for the left-hand side of (3.1):

$$\mathrm{Tr}\,[\rho a_{i_1}^\dagger \cdots a_{i_l}^\dagger a_{i_1} \cdots a_{i_l}] = \sum_{\{n_j\}} \left[ \rho_{\{n_j\},\{n_j\}} \prod_{j=1}^{N} n_j(n_j-1)\cdots(n_j-k_j+1) \right]. \tag{3.11}$$

Since the product $n_j(n_j-1)\cdots(n_j-k_j+1)$ is always nonnegative, it is less than or equal to $n_j^{k_j}$ (equality holding only if either $n_j = 0$ or $k_j = 1$). Also since $\rho$ is a nonnegative-definite operator, the diagonal elements $\rho_{\{n_j\},\{n_j\}}$ are all nonnegative numbers. We can therefore write

$$\mathrm{Tr}\,[\rho a_{i_1}^\dagger \cdots a_{i_l}^\dagger a_{i_1} \cdots a_{i_l}] \le \sum_{\{n_j\}} \rho_{\{n_j\},\{n_j\}} n_1^{k_1} n_2^{k_2} \cdots n_N^{k_N} \tag{3.12}$$

$$\le \prod_{j=1}^{N} \left\{ \sum_{\{n_j\}} \rho_{\{n_j\},\{n_j\}} n_j^{\,l} \right\}^{k_j/l}. \tag{3.13}$$

In going from (3.12) to (3.13) we have used the Hölder inequality for positive numbers.[3] Simultaneous equalities in (3.12) and (3.13) hold only either as trivial identities ($0 = 0$), or in the case when $l = 1$.

Finally, it is easy to verify that the right-hand side of (3.1) is identical to that of (3.13). Inequality (3.1) and hence also Theorem III is thus proved.

It is of interest to note a few related inequalities. Using the commutation relations (1.5) and the representation (3.7) of the operator $\rho$, one can easily show that for arbitrary positive integral values of $l$,

$$\mathrm{Tr}\,[\rho a_{i_1}^\dagger \cdots a_{i_l}^\dagger a_{i_1} \cdots a_{i_l}] \le \mathrm{Tr}\,[\rho a_{i_1} \cdots a_{i_l} a_{i_1}^\dagger \cdots a_{i_l}^\dagger]. \tag{3.14}$$

Further, if $b$ is any linear combination of $a_1, a_2, \cdots, a_N$, i.e., if

$$b = \sum_{i=1}^{N} \alpha_i a_i, \tag{3.15}$$

where $\alpha_i$ are arbitrary complex numbers, we show below that the following inequalities hold:

$$\mathrm{Tr}\,[\rho b^{\dagger l} b^l] \le \mathrm{Tr}\,[\rho (b^\dagger b)^l] \le \mathrm{Tr}\,[\rho (bb^\dagger)^l] \le \mathrm{Tr}\,[\rho b^l b^{\dagger l}]. \tag{3.16}$$

Here $l$ is an arbitrary positive integer, and $\rho$, as before, is a Hermitian nonnegative-definite operator.

From Eqs. (1.5b) and (3.15) we obtain the relation

$$[b, b^\dagger] = \lambda \equiv \sum_{i=1}^{N} |\alpha_i|^2 > 0. \tag{3.17}$$

From Eq. (3.17) we find that the eigenvalues of the operator $b^\dagger b$ are $0, \lambda, 2\lambda, \cdots$, and that the relations

$$b^{\dagger l} b^l = b^\dagger b (b^\dagger b - \lambda) \cdots (b^\dagger b - l\lambda + \lambda), \tag{3.18}$$

$$b^l b^{\dagger l} = bb^\dagger (bb^\dagger + \lambda) \cdots (bb^\dagger + l\lambda - \lambda), \tag{3.19}$$

$$= (b^\dagger b + \lambda)(b^\dagger b + 2\lambda) \cdots (b^\dagger b + l\lambda), \tag{3.20}$$

hold. Relations (3.18)–(3.20) can easily be established by induction. If we now represent $\rho$ in the basis formed by the orthonormalized eigenstates of the operator $b^\dagger b$ and also make use of (3.18) and the fact that $\lambda$ is a positive number, we obtain

$$\mathrm{Tr}\,(\rho b^{\dagger l} b^l) = \sum_{n=l}^{\infty} \rho_n \lambda^l n(n-1)\cdots(n-l+1) \le \sum_{n=0}^{\infty} \rho_n (\lambda n)^l = \mathrm{Tr}\,[\rho (b^\dagger b)^l]. \tag{3.21}$$

Here $\rho_n \ge 0$ is the sum of all the diagonal matrix elements of $\rho$ in the states which are eigenstate of $b^\dagger b$, with eigenvalue $n\lambda$. (Note that each of the eigenvalues of $b^\dagger b$ is $N$-fold degenerate.) Other inequalities in (3.16) follow in a similar manner from (3.19) and (3.20).

We now propose the following.

*Conjecture:* Let $b_1, b_2, \cdots, b_M$ be $M$ linearly-independent linear combinations of the annihilation operators $a_1, a_2, \cdots, a_N$, i.e., let

$$b_i = \sum_{j=1}^{N} T_{ij} a_j, \quad i = 1, 2, \cdots, M, \quad M \le N, \tag{3.22}$$

where $T_{ij}$ are some arbitrary complex numbers. Then inequalities of the form (1.6) and (3.14) also hold for the $b$ operators:

$$|\mathrm{Tr}\,(\rho b_{i_1}^\dagger \cdots b_{i_m}^\dagger b_{j_1} \cdots b_{j_n})| \le \prod_{k=1}^{m} [\mathrm{Tr}\,\{\rho (b_{i_k}^\dagger b_{i_k})^m\}]^{1/2m} \times \prod_{l=1}^{n} [\mathrm{Tr}\,\{\rho (b_{j_l}^\dagger b_{j_l})^n\}]^{1/2n}, \tag{3.23}$$

$$\mathrm{Tr}\,(\rho b_{i_1}^\dagger \cdots b_{i_m}^\dagger b_{i_1} \cdots b_{i_m}) \le \mathrm{Tr}\,(\rho b_{i_1} \cdots b_{i_m} b_{i_1}^\dagger \cdots b_{i_m}^\dagger). \tag{3.24}$$

In Eq. (3.23), it is again sufficient to consider only the case when $m = n$ and $i_1 = j_1$, $i_2 = j_2, \cdots, i_m = j_m$.

We prove Eq. (3.23) below in a special case when the transformation matrix $T$ is such that $TT^\dagger$ is diagonal, i.e., when

$$(TT^\dagger)_{ij} = \Lambda_i \delta_{ij}; \quad i, j = 1, 2, \cdots, M \tag{3.25}$$

($\Lambda_i$ are nonnegative numbers). Proof of (3.24) in this special case is similar.

Let us introduce a set of operators $c_1, c_2, \cdots, c_M$, defined by

$$c_i = \Lambda_i^{-\frac{1}{2}} b_i \quad \text{(no summation over } i\text{)}. \tag{3.26}$$

From (1.5), (3.22), (3.25), and (3.26) it follows that these operators satisfy the commutation relations

$$[c_i, c_j] = 0, \tag{3.27a}$$

$$[c_i, c_j^\dagger] = \delta_{ij}. \tag{3.27b}$$

These relations are identical to those satisfied by the $a$ operators [Eqs. (1.5)]. If we employ an argument similar to that used in proving Theorem III, we find that

$$|\mathrm{Tr}\,(\rho c_{i_1}^\dagger \cdots c_{i_m}^\dagger c_{j_1} \cdots c_{j_n})| \leq \prod_{k=1}^{m} [\mathrm{Tr}\,\{\rho (c_{i_k}^\dagger c_{i_k})^m\}]^{1/2m} \prod_{l=1}^{n} [\mathrm{Tr}\,\{\rho (c_{j_l}^\dagger c_{j_l})^n\}]^{1/2n}. \tag{3.28}$$

On substituting for $c_i$'s from (3.26) in (3.28), we obtain the required result (3.23).

## 4. SOME APPLICATIONS TO CORRELATION FUNCTIONS OF THE QUANTIZED ELECTROMAGNETIC FIELD

The inequalities stated in Sec. 3 find applications in the coherence theory of the electromagnetic field. If $A(x)$ denotes the operator corresponding to the field variable (e.g., a typical Cartesian component of either the electric field, the magnetic field, or the vector potential operator) at the space–time point $x \equiv \mathbf{r}, t$ of quantized electromagnetic field, we can write

$$A(x) = A^{(+)}(x) + A^{(-)}(x), \tag{4.1}$$

where

$$A^{(+)}(x) = \{A^{(-)}(x)\}^\dagger = \sum u_\lambda(x) a_\lambda, \tag{4.2}$$

$u_\lambda(x)$ are the mode expansion functions and $a_\lambda$ are the corresponding annihilation operators, with

$$[a_\lambda, a_{\lambda'}] = 0, \quad [a_\lambda, a_{\lambda'}^\dagger] = \delta_{\lambda\lambda'}. \tag{4.3}$$

It is now obvious from (3.16) that for an arbitrary nonnegative integer $n$, the relations

$$\langle :\{A^{(-)}(x)A^{(+)}(x)\}^n: \rangle \leq \langle \{A^{(-)}(x)A^{(+)}(x)\}^n \rangle \leq \langle \{A^{(+)}(x)A^{(-)}(x)\}^n \rangle \leq \langle ``\{A^{(-)}(x)A^{(+)}(x)\}^n" \rangle \tag{4.4}$$

hold. Here : : denotes the normal ordering operation, " " denotes the antinormal ordering operation, and the sharp brackets denote the quantum expectation values, i.e., if the statistical state of the field is described by the density operator $\rho$, then $\langle \mathcal{O} \rangle \equiv \mathrm{Tr}\,\rho\mathcal{O}$.

If the conjecture stated above [inequalities (3.23) and (3.24)] is true, we have in addition the following inequalities[7] valid for arbitrary space–time points $x_1, \cdots, x_m, x_1', \cdots, x_n'$, and for arbitrary positive integral values of $m$ and $n$:

$$|\langle A^{(-)}(x_1) \cdots A^{(-)}(x_m) A^{(+)}(x_1') \cdots A^{(+)}(x_n) \rangle| \leq \prod_{k=1}^{m} \langle [A^{(-)}(x_k)A^{(+)}(x_k)]^m \rangle^{1/2m} \times \prod_{l=1}^{n} \langle [A^{(-)}(x_l')A^{(+)}(x_l')]^n \rangle^{1/2n}, \tag{4.5}$$

$$\langle A^{(-)}(x_1) \cdots A^{(-)}(x_m) A^{(+)}(x_1) \cdots A^{(+)}(x_m) \rangle \leq \langle A^{(+)}(x_1) \cdots A^{(+)}(x_m) A^{(-)}(x_1) \cdots A^{(-)}(x_m) \rangle. \tag{4.6}$$

It is to be noted that

$$\langle A^{(-)}(x_1) \cdots A^{(-)}(x_m) A^{(+)}(x_1') \cdots A^{(+)}(x_n') \rangle$$

is the $(m, n)$th-order coherence function,[8] and $\langle :\{A^{(-)}(x)A^{(+)}(x)\}^n: \rangle$, $\langle \{A^{(-)}(x)A^{(+)}(x)\}^n \rangle$, and $\langle ``\{A^{(-)}(x)A^{(+)}(x)\}^n" \rangle$ are, respectively, proportional to the $n$th normally ordered correlation of the intensity, $n$th moment of the intensity, and $n$th antinormally ordered correlation of the intensity at the space–time point $x$.[9]

We conclude this paper by stating two more inequalities, which in slightly different forms, have been proved elsewhere.[10]

Let $b_1, b_2, \cdots, b_M$ be $M$ linearly independent linear combinations of $a_1, a_2, \cdots, a_N$ as in (3.22), and let $\rho$ be a Hermitian nonnegative-definite operator. Then the inequality

$$|\mathrm{Tr}\,(\rho b_{i_1} \cdots b_{i_m} b_{j_1}^\dagger \cdots b_{j_n}^\dagger)| \leq \prod_{k} [\mathrm{Tr}\,\{\rho b_{i_k}^m b_{i_k}^{\dagger m}\}]^{1/2m} \times \prod_{l=1}^{n} [\mathrm{Tr}\,\{\rho b_{j_l}^n b_{j_l}^{\dagger n}\}]^{1/2n} \tag{4.7}$$

holds.

Let us express the operator $\rho$ in a diagonal coherent state representation[11]

$$\rho = \int \phi(\{v\})\, |\{v\}\rangle\langle\{v\}|\, d^2\{v\}. \tag{4.8}$$

If the functional $\phi(\{v\})$ is nonnegative and well behaved everywhere, then the inequality

$$|\mathrm{Tr}\,(\rho b_{i_1}^\dagger \cdots b_{i_m}^\dagger b_{j_1} \cdots b_{j_n})| \leq \prod_{k=1}^{m} [\mathrm{Tr}\,\{\rho b_{i_k}^{\dagger m} b_{i_k}^m\}]^{1/2m} \times \prod_{l=1}^{n} [\mathrm{Tr}\,\{\rho b_{j_l}^{\dagger n} b_{j_l}^n\}]^{1/2n} \tag{4.9}$$

also holds.


## ACKNOWLEDGMENTS

The author wishes to thank Professor F. Bonenfant and Professor A. Boivin for the hospitality provided to him during his stay at Laval University, Quebec.


[7] The possible validity of the inequality (3.23) has also been suggested in a different paper: C. L. Mehta, J. Math. Phys. **8**, 1798 (1967), inequality (4.29). The inequality (4.30) of this reference stated as a conjecture is now obviously true as can be seen from the discussion given in the present paper [cf. inequalities (4.4) above].

[8] Compare with R. J. Glauber, Phys. Rev. **130**, 2529 (1963); L Mandel and E. Wolf, Rev. Mod. Phys. **37**, 245 (1965).

[9] Compare with L. Mandel, Phys. Rev. **136**, B1221 (1964); **152**, 438 (1966).

[10] Compare with relations (4.4), (6.3), and (6.4) of the paper quoted in Ref. 7.

[11] See footnote 16 of the paper quoted in Ref. 7.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 5 MAY 1968

# Stability of Matter. II

A. LENARD*
*Department of Mathematics, Indiana University, Bloomington, Indiana*
AND
FREEMAN J. DYSON
*The Institute for Advanced Study, Princeton, New Jersey*

(Received 9 October 1967)

The stability of a system of charged point particles is proved under the assumption that all negatively charged particles are fermions. A lower bound on the energy is found to be $-Aq^{\frac{2}{3}}Nme^4\hbar^{-2}$, where $q$ is the number of distinct negative species, $N$ the total number of negative particles, $m$ an upper bound for their mass, $e$ an upper bound for the absolute value of the charge on both negative and positive particles, and $A$ is a numerical constant.

## 1. INTRODUCTION

In a previous paper[1] we presented an analysis of the stability problem of matter, a problem posed by Fisher and Ruelle.[2] The mathematical model for "matter" is a system of point particles, finite in number, obeying the laws of nonrelativistic quantum mechanics, and interacting with each other solely by electrostatic (Coulomb) forces. We use the word "stability" to mean that there exists a lower bound for the energy per particle which is independent of the state of the system and of its size. In I we proved stability under the hypothesis that all particles, both positive and negative, are fermions belonging to some fixed number of different species, and that all masses, as well as all charges, have common bounds. Under these conditions we found[3]

$$E > -Aq^{\frac{2}{3}}N \text{ Ry}, \tag{1.1}$$

with

$$\text{Ry} = me^4/2\hbar^2. \tag{1.2}$$

Here $E$ is the energy, $N$ the total number of particles, $q$ the number of species, $m$ the upper bound on the masses, $e$ the upper bound on the absolute value of the charges, and $A$ is a numerical constant. We have also stated[4] but not proved that (1.1) holds under much weaker assumptions. In the present paper we give the detailed proof of this theorem.

The hypotheses we adopt are as follows:

(a) All negatively charged particles are fermions. Their total number is $N$, and they belong to no more than $q$ distinct species. Their masses do not exceed $m$, and their charges do not exceed $e$ in absolute value.

(b) The number and kind of positively charged particles is arbitrary. They may be bosons or fermions, or they may all belong to different species. Their masses, too, are arbitrary. However, their charges are bounded from above by $e$.[5]

Theorem 5 asserts that under hypotheses (a) and (b) the inequality (1.1) still holds.

Our hypotheses are just those which apply to physics. The assumption that there be no bosons with both signs of charge is essential. That a system with an indefinitely large number of positive and negative bosons is unstable was recently shown by one of us.[6]

Unfortunately, our proof of Theorem 5 is lengthy and complicated. It is constructive in the sense that not only is the existence of the constant $A$ in (1.1) shown, but a definite numerical value for it is obtained. Nevertheless, the result is scarcely better than an existence theorem because we find $A = 1.3 \times 10^{14}$, whereas it is clear from physics that a good value must be of the order of unity. This lack of precision is inherent in our method, which depends on successive use of many inequalities. Each of these is relatively good by itself, yet their total effect is to build up large constants by repeatedly multiplying small ones. We have not been careful to find the best $A$ allowed by our machinery, since in any case, the result would remain absurdly large. It is likely that any significant improvement will come from a stability proof which, at least in part, depends on new ideas.

Another outstanding unsolved problem is the one posed by Fisher and Ruelle,[2] namely to prove the existence of thermodynamic quantities for the Coulomb system in the usual bulk limit of statistical mechanics. It is a simple consequence of our Theorem

* This research was done at the Institute for Advanced Study, Princeton, New Jersey, and was supported by a National Science Foundation grant.

[1] F. J. Dyson and A. Lenard, J. Math. Phys. 8, 423 (1967). We shall refer to this paper as I.

[2] M. E. Fisher and D. Ruelle, J. Math. Phys. 7, 260 (1966).

[3] Theorem 4.

[4] Theorem 5.

[5] We depart from conventional notation in that $e$ is not the "electronic charge" but the "maximum nuclear charge."

[6] F. J. Dyson, J. Math. Phys. 8, 1538 (1967).

5 that the free energy $F_N(v, T)$ of a system of $N$ point charges in a volume $Nv$ at temperature $T$ satisfies an inequality

$$F_N(v, T) > -Nf(v, T), \tag{1.3}$$

where $f$ is independent of $N$. This is a necessary but not sufficient[6(a)] condition for the existence of

$$\lim [F_N(v, T)/N] \quad \text{as } N \to \infty.$$

## 2. THE PLAN OF ATTACK

In Theorem 9 of Paper I we have a result which almost solves the stability problem. Theorem 9 states that a system of negative fermions in an arbitrary continuous distribution of classical charge, with the self-energy of the classical charge included in the total energy, is stable. We want to prove that the fermions remain stable in an arbitrary distribution of positive classical point charges, with the mutual interactions of the positive charges included, but with their individual self-energies excluded. Our whole task consists in sharpening Theorem 9 so as to deal with this apparently trivial detail of the positive particle self-energies. Unfortunately, we have not found a way to sharpen the theorem directly for a many-fermion system.

Our plan of attack is based on the fact that we succeeded in sharpening Theorem 9 only for one negative particle at a time. We are consequently driven to an elaborate and unphysical scheme of chopping up our space into cells, each containing one negative particle (cf. Sec. 4). We then prove a sharpened form of Theorem 9 for each cell separately, with its one negative and an arbitrary number of positive particles [Eq. (5.29)]. Finally, we reassemble the fragments and show that stability in the individual cells implies stability for the whole space (cf. Sec. 5).

To make this complicated argument clearer, we carry it through in this section for the case in which there is only one negative particle in the whole space. In this case there is no need to chop the space into cells, and still the argument is not quite trivial. The results of this section will not be used in the main argument which follows, but they may have some independent interest.

When there is only one negative particle, Theorem 9 may be stated as follows (the numbering of theorems and lemmas runs consecutively from Paper I).

[6(a)] Dr. Robert B. Griffiths (private communication) has shown that Eq. (1.3) implies the existence of the thermodynamic limit in the case when the positive and negative charges are antiparticles of each other. His argument makes essential use of charge-conjugation symmetry and does not work for matter composed of nuclei and electrons.

*Theorem 10:* Let a single nonrelativistic quantum-mechanical particle with mass $m$ and charge $(-e)$ be placed in an arbitrary classical electrostatic potential. Then the total energy $E$, including the self-energy of the classical charge distribution, satisfies the inequality

$$E \geq -\tfrac{1}{4}\,\mathrm{Ry} = -(me^4/8\hbar^2). \tag{2.1}$$

To prove Theorem 10, observe that the proof of Theorem 9 in I demonstrated that

$$E \geq \tfrac{1}{2}E_2, \tag{2.2}$$

where $E_2$ is the energy of a system consisting of a pair of charges $(-e)$ and $(+e)$ with no external potential. Here $E_2$ is just the energy of a positronium atom, which is never less than the ground-state energy $[-\frac{1}{2}\,\mathrm{Ry}]$.

Written out in explicit analytic form, the statement (2.1) becomes

$$\begin{aligned}(\hbar^2/2m)\int |\nabla\psi(\mathbf{r})|^2\,d^3r - e\int U(\mathbf{r})\,|\psi(\mathbf{r})|^2\,d^3r \\ + [(1/8\pi)\int |\nabla U(\mathbf{r})|^2\,d^3r \\ + (me^4/8\hbar^2)]\int |\psi(\mathbf{r})|^2\,d^3r \geq 0,\end{aligned} \tag{2.3}$$

where $\psi(\mathbf{r})$ is the wavefunction and $U(\mathbf{r})$ is the potential. Since $e$ and $m$ are arbitrary parameters, Eq. (2.3) is equivalent to the following inequality.

*Theorem 11:* Let $U(\mathbf{r})$ be any real differentiable function and $\psi(\mathbf{r})$ any complex differentiable function. Then

$$\begin{aligned}\left|\int U(\mathbf{r})\,|\psi(\mathbf{r})|^2\,d^3r\right| \leq \left[(1/4\pi)\int |\nabla U(\mathbf{r})|^2\,d^3r\right]^{\frac{1}{2}} \\ \times \left[\int |\nabla\psi(\mathbf{r})|^2\,d^3r\right]^{\frac{1}{4}}\left[\int |\psi(\mathbf{r})|^2\,d^3r\right]^{\frac{3}{4}},\end{aligned} \tag{2.4}$$

provided that the integrals on the right side converge.

This inequality, which does not appear in the standard textbooks, is the basic tool for our work. The constant $(1/4\pi)$ is not the best possible. In our proof of the main stability theorem we shall need a version of the same inequality (Lemma 9 of Sec. 6) with the integrals confined to a cell instead of extending over all space. The reader can see, by comparing the above one-paragraph proof of Theorem 11 with the brute-force proof of Lemma 9 in Sec. 7, how great is the price we have paid for chopping up the space.

We continue with a statement of the main stability Theorem 5 for the case of a single negative particle.

*Theorem 12:* Let a single quantum-mechanical particle of mass $m$ and charge $(-e)$ interact with an

arbitrary number of fixed classical charges of magnitude $(+Ze)$. The total energy is

$$E = (\hbar^2/2m)t - Ze^2u + Z^2e^2q, \tag{2.5}$$

with

$$t = \int |\nabla\psi|^2 \, d^3r, \tag{2.6}$$

$$u = \int u(\mathbf{r}) \, |\psi|^2 \, d^3r, \tag{2.7}$$

$$u(\mathbf{r}) = \sum_k \frac{1}{|\mathbf{r} - \mathbf{r}_k'|}, \tag{2.8}$$

$$q = \sum_{j<k}\sum \frac{1}{|\mathbf{r}_j' - \mathbf{r}_k'|}, \tag{2.9}$$

the $\mathbf{r}_j'$ being the positions of the fixed charges. Then

$$E > -(Z + \tfrac{2}{8}\tfrac{7})^2 \text{ Ry}. \tag{2.10}$$

We prove Theorem 12 by applying Theorem 11 to a truncated form of the potential $u(\mathbf{r})$. Let $2\rho_k$ be the distance from $\mathbf{r}_k'$ to its nearest neighbor among the $\mathbf{r}_j'$ $(j \neq k)$ (we assume that there are at least two positive particles, since the theorem is manifestly true for one), and let $S_k$ be the sphere of radius $\rho_k$ with center $\mathbf{r}_k'$, so that $S_1, S_2, \cdots$ do not intersect. Let

$$U(\mathbf{r}) = \sum_k \min\left\{\frac{1}{|\mathbf{r} - \mathbf{r}_k'|}, \frac{1}{\rho_k}\right\}. \tag{2.11}$$

$U(\mathbf{r})$ is the potential produced by unit positive charges distributed uniformly over the surfaces of the spheres $S_k$. We write

$$u_1(\mathbf{r}) = u(\mathbf{r}) - U(\mathbf{r}) = \begin{cases} \dfrac{1}{|\mathbf{r} - \mathbf{r}_k'|} - \dfrac{1}{\rho_k} & (\mathbf{r} \text{ in } S_k) \\ 0 & (\text{otherwise}). \end{cases} \tag{2.12}$$

Since each distance $2\rho_k$ occurs at least once but no more than twice among the $|\mathbf{r}_j' - \mathbf{r}_k'|$, we have

$$\sum_k \frac{1}{\rho_k} \leq 4q, \tag{2.13}$$

and, therefore,

$$\frac{1}{8\pi}\int |\nabla U(\mathbf{r})|^2 \, d^3r = q + \sum_k \frac{1}{2\rho_k} \leq 3q. \tag{2.14}$$

Next, we apply Lemma 2 of I to integrals extended over the interiors of the spheres $S_k$,

$$\begin{aligned} &\int_{S_k} u_1 \, |\psi|^2 \, d^3r \\ &\quad \leq \frac{\lambda}{4}\int_{S_k} |\nabla\psi|^2 \, d^3r + \left(\frac{1}{\lambda} + \frac{1}{2\rho_k}\right)\int_{S_k} |\psi|^2 \, d^3r \\ &\quad \leq \frac{\lambda}{4}\int_{S_k} |\nabla\psi|^2 \, d^3r + \int_{S_k}\left(\frac{1}{\lambda} + \tfrac{1}{2}U\right)|\psi|^2 \, d^3r, \end{aligned} \tag{2.15}$$

where $\lambda$ is an arbitrary positive parameter. Summing Eq. (2.15) over the $S_k$ we find

$$u \leq \frac{\lambda}{4}t + \frac{1}{\lambda} + \frac{3}{2}\int U \, |\psi|^2 \, d^3r, \tag{2.16}$$

provided $\psi$ is normalized to 1. By Theorem 11 and Eq. (2.14), this implies

$$u \leq t^{\frac{1}{2}} + \tfrac{3}{2}(6q)^{\frac{1}{2}}t^{\frac{1}{4}}. \tag{2.17}$$

Finally, Eq. (2.5) gives

$$\begin{aligned} E &\geq \frac{\hbar^2}{2m}t - Ze^2t^{\frac{1}{2}} - \tfrac{3}{2}Ze^2(6q)^{\frac{1}{2}}t^{\frac{1}{4}} + Z^2e^2q \\ &\geq \frac{\hbar^2}{2m}t - Ze^2t^{\frac{1}{2}} - \tfrac{2}{8}\tfrac{7}{}e^2t^{\frac{1}{2}} \\ &\geq -(Z + \tfrac{2}{8}\tfrac{7}{})^2\frac{e^4m}{2\hbar^2}. \end{aligned} \tag{2.18}$$

This completes the proof of Theorem 12.

The argument in Sec. 6 is analogous, and differs in detail only in that all integrals are there extended over a cube instead of over all space.

Theorem 12 goes beyond Theorem 5 in one respect, namely by allowing the negative and positive charges to be *different*. In addition, Theorem 12 is asymptotically exact as $Z$ becomes very large, though for $Z$ small it is numerically poor. We conjecture that a similarly strengthened version of Theorem 5 holds, namely the inequality

$$E > -(Z + Aq^{\frac{1}{3}})^2N \text{ Ry} \tag{2.19}$$

for the energy $E$ of $N$ negative fermions of charge $-e$ in the field of positive point charges $Ze$, the other symbols having the same meanings as in Theorem 5. As we now have it, Theorem 5 has a factor $Z^4$ on the right hand side,[5] a result which is obviously too weak for large $Z$. In trying to prove Eq. (2.19) we encountered technical difficulties. This remains a problem for the future.

## 3. PRELIMINARY SIMPLIFICATIONS

At this point the formal proof of Theorem 5 begins.

We assume that the state of the system is given by a wavefunction

$$\psi = \psi(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N, \mathbf{r}_1', \mathbf{r}_2', \cdots, \mathbf{r}_{N'}'), \tag{3.1}$$

which is smooth in all variables and tends to zero sufficiently fast at infinity. The $\mathbf{r}_j$ and the $\mathbf{r}_k'$ are position coordinates of the $N$ negative and $N'$ positive particles. We write

$$\begin{aligned} &\Omega = \int d\tau' \int d\tau \, |\psi|^2, \\ &d\tau = \prod_{j=1}^{N} d^3r_j, \quad d\tau' = \prod_{k=1}^{N'} d^3r_k', \end{aligned} \tag{3.2}$$

and for reasons which will become evident below prefer not to set $\Omega = 1$ but leave the normalization arbitrary.

The total energy is

$$E = (T + W)/\Omega$$
$$= (T_+ + T_- + W_{++} + W_{+-} + W_{--})/\Omega. \quad (3.3)$$

Here

$$T_+ = \sum_{k=1}^{N'} \frac{\hbar^2}{2m_k'} \int d\tau' \int d\tau \, |\nabla_k' \psi|^2, \quad (3.4)$$

$$T_- = \sum_{j=1}^{N} \frac{\hbar^2}{2m_j} \int d\tau' \int d\tau \, |\nabla_j \psi|^2, \quad (3.5)$$

$$W_{++} = \sum_{1 \le k < l \le N'} \sum e_k' e_l' \int d\tau' \int d\tau \frac{|\psi|^2}{|\mathbf{r}_k' - \mathbf{r}_l'|}, \quad (3.6)$$

$$W_{+-} = \sum_{j=1}^{N} \sum_{k=1}^{N'} e_j e_k' \int d\tau' \int d\tau \frac{|\psi|^2}{|\mathbf{r}_j - \mathbf{r}_k'|}, \quad (3.7)$$

$$W_{--} = \sum_{1 \le j < i \le N} \sum e_j e_i \int d\tau' \int d\tau \frac{|\psi|^2}{|\mathbf{r}_j - \mathbf{r}_i|}. \quad (3.8)$$

The charges and masses are subject to the inequalities

$$\begin{aligned} -e \le e_j \le 0 \qquad & (1 \le j \le N), \\ 0 \le e_k' \le e \qquad & (1 \le k \le N'), \end{aligned} \quad (3.9)$$

and

$$0 < m_j \le m \qquad (1 \le j \le N). \quad (3.10)$$

The inequality which we want to prove is

$$T + W > -Aq^{\frac{2}{3}} N\Omega \text{ Ry}. \quad (3.11)$$

Our first remark is that it is both necessary and sufficient to prove

$$T_- + W > -Aq^{\frac{2}{3}} N\Omega \text{ Ry}. \quad (3.12)$$

Necessary, because $T_+$ can be made arbitrarily small by choosing the $m_k'$ large enough and the $m_k'$ do not enter the right-hand side of (3.11). Clearly sufficient also, since $T_+ > 0$. Furthermore, to prove (3.12), it is both necessary and sufficient to prove

$$T_-(\mathbf{r}') + W(\mathbf{r}') > -Aq^{\frac{2}{3}} N\Omega(\mathbf{r}') \text{ Ry}, \quad (3.13)$$

where, for instance,

$$\Omega(\mathbf{r}') = \Omega(\mathbf{r}_1', \mathbf{r}_2', \cdots, \mathbf{r}_{N'}') = \int d\tau \, |\psi|^2 \quad (3.14)$$

and $T_-(\mathbf{r}')$, $W(\mathbf{r}')$ are defined similarly. It is necessary because in the absence of $T_+$ we may choose a wavefunction whose support in the $\mathbf{r}_k'$ variables is entirely near some arbitrarily chosen points. Sufficient also, because Eq. (3.12) follows from Eq. (3.13) by integration over the $N'$ variable points $\mathbf{r}_k'$. The quantity

$$E(\mathbf{r}') = [T_-(\mathbf{r}') + W(\mathbf{r}')]/\Omega(\mathbf{r}') \quad (3.15)$$

is just the energy of the $N$ negative fermions in the field of $N'$ positive point charges fixed at the points $\mathbf{r}_k'$, this energy including the contribution

$$W_{++}(\mathbf{r}')/\Omega(\mathbf{r}') = \sum_{1 \le k < l \le N'} \sum \frac{e_k' e_l'}{|\mathbf{r}_k' - \mathbf{r}_l'|} \quad (3.16)$$

which is a fixed "classical" quantity independent of the wavefunction. We keep this interpretation in mind and revert to the simpler notation of writing $\Omega$, $T$, $W$ instead of $\Omega(\mathbf{r}')$, $T_-(\mathbf{r}')$, $W(\mathbf{r}')$. The wavefunction $\psi$ is thought of as depending on the $\mathbf{r}_j$ only, and all integrals are over these $N$ position variables, while the $\mathbf{r}_k'$ are regarded as arbitrary but fixed parameters.

Since $T_-$ is minimized by choosing the masses $m_j$ as large as possible, it is no restriction if we replace Eq. (3.10) by

$$m_j = m \quad (1 \le j \le N). \quad (3.17)$$

We now take a fixed $\psi$ and consider the dependence of the energy on the charges $e_j$ and $e_k'$. The Coulomb energy $W$, when regarded as a function of a single charge, is a linear function. Suppose that $W$ assumes its minimum value $\hat{W}$ in the $(N + N')$-dimensional cube Eq. (3.9) at some point $(\hat{e}_1, \hat{e}_2, \cdots, \hat{e}_{N'}')$. Suppose $j$ is such that $-e < \hat{e}_j < 0$. A linear function can possess a minimum in the inside of an interval only if, in fact, it is a constant. Thus we still get $W = \hat{W}$ if we replace $\hat{e}_j$ by one of the values 0 or $-e$. This argument shows that *the minimum of $W$ in the cube* (3.9) *is assumed at some vertex of this cube.* At such a vertex a certain number of the charges vanish. If we omit the corresponding terms from the kinetic energy $T$ the energy is further diminished. Thus we may write

$$T + W \ge \int d\tau_1 (T_1 + W_1), \quad (3.18)$$

where $d\tau_1 = \prod_j d^3 r_j$ over those $j$ for which the minimization of $W$ yielded $e_j = 0$, and the subscript 1 signifies the replacement of all remaining $e_j$ by $-e$ and $e_k'$ by $e$, as well as the omission of the unwanted kinetic-energy terms (if any). The integrations defining $T_1$ and $W_1$ are only over the $N_1$ ($\le N$) variables $\mathbf{r}_j$ which actually enter the Coulomb energy. Now

$$\Omega = \int d\tau_1 \Omega_1, \quad (3.19)$$

so that it is sufficient to prove

$$T_1 + W_1 > -Aq^{\frac{2}{3}} N_1 \Omega_1 \text{ Ry}, \quad (3.20)$$

and then Eq. (3.11) follows from Eqs. (3.20) and (3.18) by integration over the "remaining" variables. But Eq. (3.20) is an inequality of the precise form (3.11); only the charges have all been taken with

their maximum allowed absolute value. This shows that it is no restriction of our original hypotheses if we replace the inequalities (3.9) by

$$\begin{aligned} e_j &= -e \quad (1 \leq j \leq N), \\ e_k' &= e \quad\quad (1 \leq k \leq N') \end{aligned} \tag{3.21}$$

from the beginning.

Once this is done, it is natural to adopt a system of atomic units[7] such that $e = 1$ and $\hbar^2/2m = 1$. In these units Ry $= \frac{1}{2}$.

## 4. CONFIGURATIONAL DOMAINS WITH UNIFORM NEAREST-NEIGHBOR SEPARATION

In this section we construct certain domains $S$ in the $3N$-dimensional configuration space ($3N$-space, for short). A point $(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N)$ in $3N$-space will be abbreviated as $(\mathbf{r})$. The $\mathbf{r}_j$ are the components of $(\mathbf{r})$.

For any point $(\mathbf{r})$ in $3N$-space we write

$$R_j = \min_{\substack{1 \leq i \leq N \\ i \neq j}} |\mathbf{r}_i - \mathbf{r}_j| \quad (j = 1, 2, \cdots, N). \tag{4.1}$$

$R_j$ is the distance between the component $\mathbf{r}_j$ and its "nearest neighbor" among the other components.

Let $S$ be an $N$-fold direct product of 3-dimensional cubes. This means that $(\mathbf{r})$ is in $S$ if and only if $\mathbf{r}_j$ is in $C_j$ $(j = 1, 2, \cdots, N)$, where the $C_j$ are cubes in 3-space. We use the letter $S$ always to denote such a domain, and write $L_j = L_j(S)$ for the length of the edges of $C_j$. Suppose then that we have a collection $\{S\}$ of such domains $S$. We shall say that the $S$ in the collection have *uniform nearest-neighbor separation* if there are two positive constants $\alpha$ and $\alpha' > \alpha$ such that, given any $S$ in the collection, and any $(\mathbf{r})$ in $S$, we have

$$\alpha L_j \leq R_j < \alpha' L_j \quad (j = 1, 2, \cdots, N). \tag{4.2}$$

This condition means, roughly speaking, that if we pick any cube $C_j$ among those defining $S$, and then pick its nearest neighbor $C_i$, the distance between $C_j$ and $C_i$ is of the order of magnitude $L_j$, the size of $C_j$ itself. The nearest neighbor of a large cube cannot be too close. The nearest neighbor of a small cube cannot be too far.

It is clear that if we have a collection $\{S\}$ of domains with uniform nearest-neighbor separation, then no point $(\mathbf{r})$, some of whose components coincide, can be in any of the $S$. For if $\mathbf{r}_i = \mathbf{r}_j$ for some $i \neq j$, then $R_j = R_i = 0$ and so the first inequality of Eq. (4.2) cannot hold for $\alpha > 0$. We shall make later essential use of a proposition which in some sense is the converse of this.

[7] See, however, Ref. 5.

*Lemma 6:* Given any positive constants $\alpha$ and $\alpha' \geq 2\alpha + 4\sqrt{3}$ there exists a countable collection $\{S\}$ of domains $S$ in $3N$-space having uniform nearest-neighbor separation, such that (i) no two $S$ in $\{S\}$ intersect, and (ii) if $(\mathbf{r})$ is any point in $3N$-space with distinct components then $(\mathbf{r})$ belongs to one of the $S$ in the collection. That is to say, the whole $3N$-space, with the exception of the points having some identical components, is partitioned into disjoint domains $S$.

The proof consists of the following steps. We take an arbitrary point $(\mathbf{r})$ in $3N$-space, with distinct components, and construct a certain uniquely determined domain $S$ which is a direct product of cubes and contains $(\mathbf{r})$. Second, we show that if two such domains contain a single point in common, they are necessarily identical. Third, we show that the resulting collection of disjoint domains $S$ has uniform nearest-neighbor separation. Finally, we show that only countably many distinct $S$ arise from our construction.

We use the following notation: If $U$ and $U'$ are two sets in 3-space, we write

$$d\{U, U'\} = \inf |\mathbf{r} - \mathbf{r}'|, \quad (\mathbf{r} \text{ in } U, \quad \mathbf{r}' \text{ in } U'). \tag{4.3}$$

By a cube in 3-space we mean in the following always a set of those $\mathbf{r} = (x, y, z)$ for which

$$\begin{aligned} a \leq x < a + L, \\ b \leq y < b + L, \\ c \leq z < c + L \end{aligned} \tag{4.4}$$

for some $a$, $b$, $c$, and $L$.[8] If $C$ and $C'$ are two cubes of edges $L$ and $L'$, respectively, we have

$$d\{C, C'\} \leq |\mathbf{r} - \mathbf{r}'| < d\{C, C'\} + \sqrt{3}(L + L') \quad (\mathbf{r} \text{ in } C, \quad \mathbf{r}' \text{ in } C'). \tag{4.5}$$

We call a cube $C$ a *binary cube* if $L = 2^\nu$ where $\nu$ is an integer, and $a$, $b$, $c$ are multiples of $L$. Binary cubes of given size $L$ are disjoint and form a cubic lattice in 3-space so that every point $\mathbf{r}$ belongs to exactly one of them. We denote by $C_\nu(\mathbf{r})$ the binary cube of edge length $L = 2^\nu$ which contains $\mathbf{r}$. Binary cubes of different sizes either do not intersect or else the larger one wholly contains the smaller one. The nested sequence $C_\nu(\mathbf{r})$ $(\nu = \cdots, -2, -1, 0, 1, 2, \cdots)$ is uniquely determined by and also uniquely determines $\mathbf{r}$.

Let $(\mathbf{r})$ be some point of $3N$-space such that

$$\mathbf{r}_i \neq \mathbf{r}_j \quad (1 \leq i < j \leq N). \tag{4.6}$$

Let $\alpha$ be a fixed positive number. Let $\mu_{ij} = \mu_{ji}$ be the

[8] Thus we consider only cubes whose orientation is the same. The convention about boundary points is made so that even when two cubes are adjoining they have no common points.

largest integer $\nu$ for which

$$d\{C_\nu(\mathbf{r}_i),\ C_\nu(\mathbf{r}_j)\} \geq 2^\nu\alpha \tag{4.7}$$

holds. Evidently such an integer exists. For if (4.7) is satisfied for some $\nu$ it is satisfied for all smaller $\nu$. Further, it is not satisfied for large enough $\nu$, because then $d = 0$. On the other hand it is satisfied for small enough $\nu$, because

$$\lim_{\nu\to-\infty} d\{C_\nu(\mathbf{r}_i), C_\nu(\mathbf{r}_j)\} = |\mathbf{r}_i - \mathbf{r}_j| > 0 \tag{4.8}$$

by our assumption (4.6). We now put

$$\lambda_j = \min_{\substack{1\leq i\leq N\\ i\neq j}} \mu_{ij}. \tag{4.9}$$

The integers $\mu_{ij}$ and $\lambda_j$ are well-defined functions of the point $(\mathbf{r})$.

Let now $S$ be the direct product of the cubes $C_{\lambda_1}(\mathbf{r}_1),\ C_{\lambda_2}(\mathbf{r}_2), \cdots, C_{\lambda_N}(\mathbf{r}_N)$. Thus $(\mathbf{r}')$ is in $S$ if and only if $\mathbf{r}'_j$ is in $C_{\lambda_j}(\mathbf{r}_j)$ for $j = 1, 2, \cdots, N$. Clearly $(\mathbf{r})$ itself is in $S$. Thus every point $(\mathbf{r})$ in $3N$-space which has distinct components belongs to a well determined $S$.

We now show that no two *distinct* $S$ have a point in common. It is sufficient to prove that if $S$ is assigned to the point $(\mathbf{r})$, and $(\mathbf{r}')$ is any other point of $S$, then $S'$ which is assigned to $(\mathbf{r}')$ is identical to $S$. Let then $(\mathbf{r}')$ be in $S$ so that the $\mathbf{r}'_j$ are in the $C_{\lambda_j}(\mathbf{r}_j)$, respectively. Then $C_\nu(\mathbf{r}'_j) = C_\nu(\mathbf{r}_j)$ for all $j$ and all $\nu \geq \lambda_j$ because the binary cubes of increasing size containing a given binary cube are uniquely determined. Moreover, because $\mu_{ij} \geq \lambda_i$ and $\mu_{ij} \geq \lambda_j$, we have

$$\begin{cases} C_\nu(\mathbf{r}'_i) = C_\nu(\mathbf{r}_i) \\ C_\nu(\mathbf{r}'_j) = C_\nu(\mathbf{r}_j) \end{cases} \quad (\text{for all } \nu \geq \mu_{ij}). \tag{4.10}$$

Thus

$$d\{C_{\mu_{ij}}(\mathbf{r}'_i),\ C_{\mu_{ij}}(\mathbf{r}'_j)\} \geq 2^{\mu_{ij}}\alpha, \tag{4.11}$$

but

$$d\{C_\nu(\mathbf{r}'_i),\ C_\nu(\mathbf{r}'_j)\} < 2^\nu\alpha \quad (\text{for all } \nu > \mu_{ij}) \tag{4.12}$$

by definition of $\mu_{ij}$. But (4.11) and (4.12) show that $\mu'_{ij} = \mu_{ij}$ and therefore also $\lambda'_j = \lambda_j$. Finally

$$C_{\lambda_{j'}}(\mathbf{r}'_j) = C_{\lambda_j}(\mathbf{r}'_j) = C_{\lambda_j}(\mathbf{r}_j) \tag{4.13}$$

where in the last step we used the fact that $\mathbf{r}'_j$ is in $C_{\lambda_j}(\mathbf{r}_j)$ and that a binary cube of given size containing a point is unique. Thus $S'$, which is a direct product of the $C_{\lambda_{j'}}(\mathbf{r}'_j)$, is identical to $S$.

We next show that all $S$ so constructed have a uniform nearest-neighbor separation. Let $(\mathbf{r})$ be a point of $S$ and let the $\mu_{ij}$ and $\lambda_j$ be given as above. We shall verify the inequalities (4.2) with $L_j = 2^{\lambda_j}$. Take any $j$ and then pick an $i$ so that $R_j = |\mathbf{r}_j - \mathbf{r}_i|$ [cf. Eq. (4.1) above]. Then

$$\begin{aligned} R_j = |\mathbf{r}_j - \mathbf{r}_i| &\geq d\{C_{\lambda_j}(\mathbf{r}_j), C_{\lambda_i}(\mathbf{r}_i)\} \\ &\geq d\{C_{\mu_{ij}}(\mathbf{r}_j), C_{\mu_{ij}}(\mathbf{r}_i)\} \geq 2^{\mu_{ij}}\alpha \\ &\geq 2^{\lambda_j}\alpha = L_j\alpha. \end{aligned} \tag{4.14}$$

This proves the first part of (4.2). Next, take again any $j$ but now pick $i$ so that $\mu_{ij} = \lambda_j$ [cf. Eq. (4.9) above]. Then

$$\begin{aligned} R_j \leq |\mathbf{r}_i - \mathbf{r}_j| &\leq d\{C_{\mu_{ij}+1}(\mathbf{r}_i), C_{\mu_{ij}+1}(\mathbf{r}_j)\} + 2^{\mu_{ij}+2}\sqrt{3} \\ &< 2^{\mu_{ij}+1}\alpha + 2^{\mu_{ij}+2}\sqrt{3} \\ &= 2^{\lambda_j}(2\alpha + 4\sqrt{3}) \leq L_j\alpha' \end{aligned} \tag{4.15}$$

[The second inequality is (4.5) above.] This proves the second part of (4.2).

Finally, we see easily that in the collection $\{S\}$ there are only countably many distinct $S$. For each $S$ has a positive $3N$-dimensional volume, namely

$$\int d^{3N}r = 2^{3(\lambda_1+\lambda_2+\cdots+\lambda_N)}. \tag{4.16}$$

Therefore the number of those $S$ which are inside some bounded domain of $3N$-space, and for which $\lambda_1 + \lambda_2 + \cdots + \lambda_N$ exceeds some given lower bound, is necessarily finite.

This concludes the proof of Lemma 6.

## 5. REDUCTION TO A ONE-PARTICLE PROBLEM

In accordance with Lemma 6 we choose two positive constants $\alpha$ and $\alpha' \geq 2\alpha + 4\sqrt{3}$ and decompose $3N$-space into a collection of nonintersecting domains $S$ with uniform nearest-neighbor separation. Since the exceptional points—those that do not belong to any $S$—form a set of measure zero, we can represent any integral over $3N$-space as an infinite sum of integrals, each term giving the contribution of a particular $S$. Thus we write

$$\Omega = \sum_S \Omega(S), \tag{5.1}$$

where

$$\Omega(S) = \int_S |\psi|^2\, d\tau, \tag{5.2}$$

and we define $T(S)$, $W(S)$ similarly. Integrals over $S$ are of the form

$$\int_S d\tau(\cdots) = \int_{C_1} d^3r_1 \int_{C_2} d^3r_2 \cdots \int_{C_N} d^3r_N(\cdots), \tag{5.3}$$

where the $C_j$ are cubes in 3-space, their edges being of lengths $L_j$, respectively, and such that for any $\mathbf{r}_1$ in $C_1$, $\mathbf{r}_2$ in $C_2, \cdots, \mathbf{r}_N$ in $C_N$ the inequalities (4.2) hold.[9]

[9] From here on it is irrelevant that $L_j = 2^{\lambda_j}$, with $\lambda_j$ an integer. Only the universal validity of (4.2) with the same $\alpha$ and $\alpha'$ is essential.

We now have the following.

*Theorem 13:* For an appropriate choice of $\alpha$ in Lemma 6, there are two constants $A_1$ and $A_2$ such that for all $S$,

$$T(S) + W(S) \geq -\left(A_1 N + A_2 \sum_{j=1}^{N} \frac{1}{L_j}\right)\Omega(S). \quad (5.4)$$

Before proving this theorem, we show that by combining it with the results of our first paper we can deduce Theorem 5 from it.

Indeed, from (4.2) we obtain

$$\frac{1}{L_j}\,\Omega(S) \leq \alpha' \int_S d\tau \frac{|\psi|^2}{R_j}, \quad (5.5)$$

where $R_j = R_j(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N)$ is defined by (4.1). Substituting into Eq. (5.4) and summing over all $S$ we then obtain

$$T + W \geq -(A_1 + A_2\alpha' K)N\Omega, \quad (5.6)$$

where

$$K = \frac{1}{N\Omega} \sum_{j=1}^{N} \int d\tau \frac{|\psi|^2}{R_j}. \quad (5.7)$$

It is at this point, and at this point only, that we use the hypothesis that all negatively charged particles (i.e., all particles that we still treat quantum mechanically) are fermions. In I we proved the following inequality under that assumption[10]:

$$K \leq A_3 q^{\frac{1}{3}} (T/N\Omega)^{\frac{1}{2}}, \quad (5.8)$$

where $q$ is the number of fermion species and $A_3$ is a constant (about 22.2). Thus

$$T \geq A_3^{-2} q^{-\frac{2}{3}} K^2 N\Omega. \quad (5.9)$$

Adding (5.9) to (5.6) and minimizing the right-hand side with respect to $K$, we obtain

$$2T + W \geq -[A_1 + \tfrac{1}{4}(\alpha' A_2 A_3)^2 q^{\frac{2}{3}}]N\Omega. \quad (5.10)$$

$2T$ may be replaced by $T$ since the particle mass $m$ is arbitrary (this produces a factor 2 on the right). Finally, noting that $q \geq 1$, we have

$$T + W \geq -Aq^{\frac{2}{3}}N\Omega \quad (5.11)$$

with

$$A = 2A_1 + \tfrac{1}{2}(\alpha' A_2 A_3)^2. \quad (5.12)$$

This completes the derivation of Theorem 5 from Theorem 13.

The rest of the paper is devoted to the proof of Theorem 13.

Let $S$ be an arbitrary, but from now on fixed, domain of the decomposition given in Lemma 6.

[10] Equation (7.14).

We now consider $W(S)$ a function of the $N'$ parameters $\mathbf{r}'_k$ $(1 \leq k \leq N')$ which are the positions of the positive charges. As a function of $\mathbf{r}'_1$, say, $W(S)$ has the form

$$W(S) = W_1(S) + \Omega(S) \sum_{k=2}^{N'} \frac{1}{|\mathbf{r}'_1 - \mathbf{r}'_k|} - \sum_{j=1}^{N} \int_{C_j} d^3r \frac{\rho(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'_1|}, \quad (5.13)$$

where $W_1(S)$ is independent of $\mathbf{r}'_1$, $\rho(\mathbf{r})$ is nonnegative, and the $C_j$ are the cubes defining $S$. Suppose that $\mathbf{r}'_1$ is in none of the $C_j$. Then we regard $W(S)$ as a function of $\mathbf{r}'_1$ defined in the open domain exterior to the cubes $C_j$ and excluding the points $\mathbf{r}'_k$ $(2 \leq k \leq N')$. It is evident from Eq. (5.13) that this function satisfies Laplace's equation in the coordinates of $\mathbf{r}'_1$. Such a function assumes no minimum in its domain of definition. Thus either a minimum occurs when $\mathbf{r}'_1$ is a point on the boundary of one of the $C_j$, or else there is no minimum at all and $W(S) > W_1(S)$. We see from this argument that for any given $\mathbf{r}'_1, \mathbf{r}'_2, \cdots, \mathbf{r}'_{N'}$ we can find another set $\mathbf{r}''_1, \mathbf{r}''_2, \cdots, \mathbf{r}''_{N''}$ with $N'' \leq N'$ such that by replacing the positive charges at the $\mathbf{r}'_k$ with charges at the $\mathbf{r}''_k$ we do not increase $W(S)$ and, moreover, *every* $\mathbf{r}''_k$ *is inside or on the boundary of one of the* $C_j$. Thus, in attempting to prove the inequality (5.4) for any particular $S$, it is no restriction if we assume the $\mathbf{r}'_k$ to satisfy this condition from the beginning.

The next step is the application of Theorem 6 proved in I. Accordingly,

$$W(S) > U(S), \quad (5.14)$$

where

$$U(S) = \int_S d\tau\, |\psi|^2 \Big\{ -\sum_{j=1}^{N} \frac{1}{2a_j} - \sum_{k=1}^{N'} \frac{1}{2a'_k} + \sum_{1 \leq i < j \leq N} \sum \Delta(|\mathbf{r}_i - \mathbf{r}_j|, a_i, a_j) + \sum_{1 \leq k < l \leq N} \sum \Delta(|\mathbf{r}'_k - \mathbf{r}'_l|, a'_k, a'_l) - \sum_{j=1}^{N} \sum_{k=1}^{N'} \Delta(|\mathbf{r}_j - \mathbf{r}'_k|, a_j, a'_k) \Big\}. \quad (5.15)$$

The positive quantities $a_j$ and $a'_k$ are arbitrary and may be functions of the $\mathbf{r}_j$ and $\mathbf{r}'_k$. The function $\Delta(r, a, b)$ is defined for positive values of its arguments by

$$\Delta(r, a, b) = \begin{cases} \dfrac{1}{r} - \min\left(\dfrac{1}{a}, \dfrac{1}{b}\right) & (0 < r \leq |a - b|) \\ \dfrac{(a + b - r)^2}{4abr} & (|a - b| \leq r \leq a + b) \\ 0 & (a + b \leq r). \end{cases} \quad (5.16)$$

It is positive and monotone decreasing in $r$, and it satisfies the inequality

$$\frac{1}{r} - \frac{1}{a} < \Delta(r, a, a) < \frac{1}{r}. \tag{5.17}$$

We set

$$a_j = \frac{\alpha}{2} L_j \quad (1 \leq j \leq N) \tag{5.18}$$

and

$$a'_k = \frac{\alpha}{2} L_j \quad (\text{for } k \text{ such that } \mathbf{r}'_k \text{ is in } C_j). \tag{5.19}$$

(At this point we use the fact that all $\mathbf{r}'_k$ belong to one of the cubes $C_j$.) We now verify the following fact: *In the double sums of* (5.15) *all terms vanish in which the two particle positions involved belong to two different cubes.* Indeed,

$$|\mathbf{r}_i - \mathbf{r}_j| \geq \max(R_i, R_j) \geq \alpha \max(L_i, L_j), \tag{5.20}$$

where the last inequality follows from (4.2), and hence

$$|\mathbf{r}_i - \mathbf{r}_j| \geq 2 \max(a_i, a_j) \geq a_i + a_j \tag{5.21}$$

so that

$$\Delta(|\mathbf{r}_i - \mathbf{r}_j|, a_i, a_j) = 0. \tag{5.22}$$

Similarly,

$$\Delta(|\mathbf{r}_j - \mathbf{r}'_k|, a_j, a'_k) = 0 \quad (\text{for } \mathbf{r}'_k \text{ not in } C_j) \tag{5.23}$$

and

$$\Delta(|\mathbf{r}'_k - \mathbf{r}'_l|, a'_k, a'_l) = 0 \quad (\text{for } \mathbf{r}'_k \text{ in } C_j, \mathbf{r}'_l \text{ in } C_i, i \neq j). \tag{5.24}$$

Let $\mu_j$ be the number of $\mathbf{r}'_k$ in $C_j$. We may then rewrite (5.15) in the form

$$U(S) = \sum_{j=1}^{N} U_j(S), \tag{5.25}$$

where

$$U_j(S) = \left[ -\frac{1+\mu_j}{\alpha L_j} + \sum_{(k<l)_j} \sum \Delta\left(|\mathbf{r}'_k - \mathbf{r}'_l|, \frac{\alpha}{2} L_j, \frac{\alpha}{2} L_j\right) \right] \Omega(S) - \int_S d\tau\, |\psi|^2 \sum_{(k)} \Delta\left(|\mathbf{r}_j - \mathbf{r}'_k|, \frac{\alpha}{2} L_j, \frac{\alpha}{2} L_j\right). \tag{5.26}$$

In view of (5.17),

$$U_j(S) > \left[ -\frac{1+\mu_j^2}{\alpha L_j} + \sum_{(k<l)_j} \sum \frac{1}{|\mathbf{r}'_k - \mathbf{r}'_l|} \right] \Omega(S) - \int_S d\tau\, |\psi|^2 \sum_{(k)_j} \frac{1}{|\mathbf{r}_j - \mathbf{r}'_k|}. \tag{5.27}$$

In the last two formulas we used the notation $(k)_j$ and $(k < l)_j$ to indicate the restriction of the sums to those subscripts for which $\mathbf{r}'_k$ and $\mathbf{r}'_l$ are in $C_j$. In view of the inequalities (5.14) and (5.27), if we succeed in showing

$$\left[ -\frac{1}{\alpha L_j} (1 + \mu_j^2) + \sum_{(k<l)_j} \sum \frac{1}{|\mathbf{r}'_k - \mathbf{r}'_l|} \right] \Omega(S) + \int_S d\tau \left[ |\nabla_j \psi|^2 - |\psi|^2 \sum_{(k)_j} \frac{1}{|\mathbf{r}_j - \mathbf{r}'_k|} \right] \geq -\left( A_1 + \frac{A_2}{L_j} \right) \Omega(S) \tag{5.28}$$

for appropriate constants $A_1$ and $A_2$, then (5.4) follows by summing over $j = 1, 2, \cdots, N$.

Note that although in (5.28) the integrals are still over the $3N$-dimensional domain $S$, it is only the dependence of $\psi$ on $\mathbf{r}_j$ which is relevant. Thus, it is sufficient to prove (5.28) in a modified form in which the integrals $\int_S d\tau$ are replaced by $\int_{C_j} d^3 r_j$, for then (5.28) follows by integration with respect to the rest of the variables $\mathbf{r}_1, \cdots, \mathbf{r}_{j-1}, \mathbf{r}_{j+1}, \cdots, \mathbf{r}_N$. But then we may ignore the dependence of $\psi$ on these other variables and, since $j$ was *any one* of the $N$ subscripts, we may drop it too. We have thus shown that Theorem 13 is a consequence of the following inequality:

$$\int_C d^3 r \left\{ |\nabla \psi|^2 + |\psi|^2 \left[ \sum_{k=1}^{\mu} \frac{-1}{|\mathbf{r} - \mathbf{r}'_k|} + \sum_{1 \leq k < l \leq \mu} \sum \frac{1}{|\mathbf{r}'_k - \mathbf{r}'_l|} \right] \right\} \geq \left\{ -A_1 + \left( \frac{1+\mu^2}{\alpha} - A_2 \right) \frac{1}{L} \right\} \int_C d^3 r\, |\psi|^2, \tag{5.29}$$

where $C$ is any cube whose edges are of length $L$, $\mathbf{r}'_k$ $(k = 1, 2, \cdots, \mu)$ are arbitrary points in or on the boundary of $C$, and $\psi = \psi(\mathbf{r})$ is an arbitrary smooth function defined in $C$.

The case[11] when $\psi = 0$ identically in $C$ may be ignored for then (5.29) holds trivially. If we then normalize $\psi$

$$\int_C |\psi|^2 d^3 r = 1, \tag{5.30}$$

the left side of (5.29) can be interpreted as the energy of a single negatively charged particle, confined to a box $C$, and attracted to $\mu$ fixed point charges in $C$. The energy includes the static repulsive contribution of these fixed charges. We are then looking for a lower bound for the ground-state energy of this system, and this lower bound must be shown to have the particular dependence on the parameters $\mu$ and $L$ shown on the right-hand side. The many-body aspect of the original problem has been hereby eliminated.

There is one essential remark. Our problem is not the usual "particle in a box," because we have no

[11] This may occur because $\psi(\mathbf{r})$ is the restriction to $C$ of a function defined over all space.

boundary conditions on $\psi$. The customary requirements are either that $\psi = 0$ on the boundary, or that $\psi$ is periodic, modulo $L$, when extended in a natural manner to all space. In contrast our only requirement on $\psi$ is that it remain bounded and differentiable at the boundary of $C$, since it is a restriction to $C$ of a smooth function defined over all space. This lack of boundary conditions on $\psi$ leads to some complication of detail in Sec. 7.

## 6. SOLUTION OF THE ONE-PARTICLE PROBLEM

In considering the proof of (5.29) we have to distinguish three cases, $\mu = 0$, $\mu = 1$, and $\mu \geq 2$.

Let $\mu = 0$. Then there is only the kinetic energy left and the inequality reads[12]

$$\int_C d^3r\, |\nabla\psi|^2 \geq -A_1 + \left(\frac{1}{\alpha} - A_2\right)\frac{1}{L}.$$

For this to be satisfied it is sufficient that

$$A_2 \geq \frac{1}{\alpha}. \tag{6.1}$$

Let now $\mu = 1$. In this case we need to show

$$\int_C d^3r \left\{|\nabla\psi|^2 - \frac{|\psi|^2}{|\mathbf{r} - \mathbf{r}_1'|}\right\} \geq -A_1 + \left(\frac{2}{\alpha} - A_2\right)\frac{1}{L}. \tag{6.2}$$

*Lemma 7:* For any positive $\lambda$, and any complex $\Psi(\mathbf{r})$ with continuous derivatives, and any cube $C$ of size $L$,

$$\int_C d^3r \frac{|\Psi|^2}{|\mathbf{r}|} < \left(\frac{8}{\lambda} + \frac{24}{L}\right)\int_C d^3r\, |\Psi|^2 + 2\lambda \int_C |\nabla\Psi|^2\, d^3r. \tag{6.3}$$

This proposition is closely related to Lemma 2 which we proved in I. There we had a sphere $\Omega$ of radius $b$ centered at the origin; here we have a cube $C$ of edge length $L$ in an arbitrary position. The inequality is of the same form, and only the numerical coefficients are different. In contrast with Lemma 2, here we do not have the best possible constants. The proof of Lemma 7 will be given in Sec. 7. To show (6.2) we take origin of coordinates at $\mathbf{r}_1'$, set $2\lambda = 1$, and take $\psi = \Psi$. It is seen that (6.2) follows from (6.3) if

$$A_1 \geq 16 \tag{6.4}$$

and

$$A_2 \geq \frac{2}{\alpha} + 24. \tag{6.5}$$

[12] From here on we assume $\psi$ to be normalized to unity in the cube $C$, according to (5.30).

Finally, suppose $\mu \geq 2$. We use the notation

$$t = \int_C |\nabla\psi|^2\, d^3r, \tag{6.6}$$

$$w = \sum_{k=1}^{\mu} \int_C \frac{|\psi|^2}{|\mathbf{r} - \mathbf{r}_k'|}\, d^3r, \tag{6.7}$$

and

$$w' = \sum_{1 \leq k < l \leq \mu} \sum \frac{1}{|\mathbf{r}_k' - \mathbf{r}_l'|}, \tag{6.8}$$

so that the inequality to be proved reads

$$t - w + w' \geq -A_1 + \left(\frac{1 + \mu^2}{\alpha} - A_2\right)\frac{1}{L}. \tag{6.9}$$

Let

$$R_k' = \min_{\substack{1 \leq l \leq \mu \\ l \neq k}} |\mathbf{r}_l' - \mathbf{r}_k'| \tag{6.10}$$

and let $\sigma_k$ denote the sphere whose radius is $R_k'/2$ and whose center is $\mathbf{r}_k'$. These spheres do not intersect. They all have a nonempty intersection with the cube $C$, but they need not lie entirely inside $C$. We denote by $\hat{\sigma}_k$ the part of $\sigma_k$ which is inside $C$. The radii of the $\sigma_k$ satisfy the inequality

$$\frac{R_k'}{2} \leq \frac{\sqrt{3}L}{2}. \tag{6.11}$$

We define a function $V(\mathbf{r})$ as follows:

$$\begin{aligned} V(\mathbf{r}) &= \sum_{l=1}^{\mu} \frac{1}{|\mathbf{r} - \mathbf{r}_l'|} \quad (\mathbf{r} \text{ in none of the } \sigma_k), \\ V(\mathbf{r}) &= \sum_{\substack{l=1 \\ l \neq k}}^{\mu} \frac{1}{|\mathbf{r} - \mathbf{r}_l'|} + \frac{2}{R_k'} \quad (\mathbf{r} \text{ in } \sigma_k). \end{aligned} \tag{6.12}$$

$V(\mathbf{r})$ is bounded, continuous, and has piecewise-continuous first derivatives. It is the Coulomb potential of a surface charge distribution, namely a unit positive charge distributed uniformly over the surface of each sphere $\sigma_k$. By elementary electrostatics

$$\frac{1}{8\pi}\int |\nabla V(\mathbf{r})|^2\, d^3r = \sum_{1 \leq k < l \leq \mu} \sum \frac{1}{|\mathbf{r}_k' - \mathbf{r}_l'|} + \sum_{k=1}^{\mu} \frac{1}{R_k'} = w' + \sum_{k=1}^{\mu} \frac{1}{R_k'}, \tag{6.13}$$

where the integral is over all space. Let

$$V(\mathbf{r}) = V_1(\mathbf{r}) + \bar{V}, \tag{6.14}$$

where

$$\bar{V} = \frac{1}{L^3}\int_C d^3r V(\mathbf{r}), \tag{6.15}$$

so that

$$\int_C d^3r V_1(\mathbf{r}) = 0. \tag{6.16}$$

We have

$$V(\mathbf{r}) \leq \sum_{k=1}^{\mu} \frac{1}{|\mathbf{r}-\mathbf{r}_k'|}, \tag{6.17}$$

and

$$\int_C \frac{d^3r}{|\mathbf{r}-\mathbf{r}_k'|} \leq \max_U \int_U \frac{d^3r}{|\mathbf{r}-\mathbf{r}_k'|}, \tag{6.18}$$

where the maximum is taken among all domains $U$ whose volume is $L^3$. It is obvious that this domain is a sphere centered on $\mathbf{r}_k'$. Thus, (6.15), (6.17), and (6.18) give

$$\bar{V} < \frac{3}{2}\left(\frac{4\pi}{3}\right)^{\frac{1}{3}} \mu/L < 3\mu/L. \tag{6.19}$$

We now write

$$w = \int_C V|\psi|^2 \, d^3r + \sum_{k=1}^{\mu} \int_{\hat{\sigma}_k} d^3r \left(\frac{1}{|\mathbf{r}-\mathbf{r}_k'|} - \frac{2}{R_k'}\right)|\psi|^2 \tag{6.20}$$

and estimate the integrals over the domains $\hat{\sigma}_k$ by using a proposition similar to Lemma 7 above.

*Lemma 8:* Let $D$ be a domain common to a sphere of radius $b$ around the origin and some cube whose edges are larger than $b$, and let $D$ contain the origin. For any positive $\lambda$ and any $\Psi(\mathbf{r})$ with piecewise continuous derivatives,

$$\int_D d^3r \frac{|\Psi|^2}{|\mathbf{r}|} \leq \left(\frac{2}{\lambda} + \frac{12}{b}\right)\int_D d^3r\,|\Psi|^2 + 8\lambda \int_D d^3r\,|\nabla\Psi|^2. \tag{6.21}$$

The proof is given in Sec. 7. Making use of this, we have

$$\int_{\hat{\sigma}_k} \frac{|\psi|^2}{|\mathbf{r}-\mathbf{r}_k'|} d^3r \leq \left(\frac{2}{\lambda} + \frac{24}{R_k'}\right)\int_{\hat{\sigma}_k} |\psi|^2 \, d^3r + 8\lambda \int_{\hat{\sigma}_k} |\nabla\psi|^2 \, d^3r. \tag{6.22}$$

But, when $\mathbf{r}$ is in $\hat{\sigma}_k$,

$$V(\mathbf{r}) \geq \frac{2}{R_k'}. \tag{6.23}$$

Hence, by summing (6.22) over $k$, and noting that the disjoint $\hat{\sigma}_k$ are inside $C$, we get

$$w \leq \frac{2}{\lambda} + 8\lambda t + 12\int_C V|\psi|^2 \, d^3r, \tag{6.24}$$

or, minimizing with respect to $\lambda$, and making use of (6.14) and (6.19),

$$w < 8t^{\frac{1}{2}} + \frac{36\mu}{L} + 12\int_C V_1|\psi|^2 \, d^3r. \tag{6.25}$$

It remains to estimate the last term of (6.25). We do this with the help of the following inequality whose proof is given in Sec. 7.

*Lemma 9:* Let $f(\mathbf{r})$ be a complex-valued and $g(\mathbf{r})$ be a real-valued function, defined in a cube $C$, both functions continuous with piecewise continuous derivatives. Let

$$\int_C |f|^2 \, d^3r = 1 \tag{6.26}$$

and

$$\int_C g \, d^3r = 0. \tag{6.27}$$

Then

$$\left|\int_C g|f|^2 \, d^3r\right|^2 \leq \frac{8}{\pi}\left(\int_C |\nabla g|^2 \, d^3r\right)\left(\int_C |\nabla f|^2 \, d^3r\right)^{\frac{1}{2}}. \tag{6.28}$$

We take $f = \psi$ and $g = V_1$ and then our conditions are fulfilled [cf. (6.16) above]. Thus,

$$\int_C V_1|\psi|^2 \, d^3r \leq \left(\frac{8}{\pi}\right)^{\frac{1}{2}} t^{\frac{1}{4}}\left(\int_C |\nabla V_1|^2 \, d^3r\right)^{\frac{1}{2}}. \tag{6.29}$$

But

$$\int_C |\nabla V_1|^2 \, d^3r = \int_C |\nabla V|^2 \, d^3r < \int |\nabla V|^2 \, d^3r = 8\pi\left(w' + \sum_{k=1}^{\mu} \frac{1}{R_k'}\right) \leq 24\pi w', \tag{6.30}$$

by (6.13) above. The last inequality follows from the fact that each term in the sum occurs at least once but no more than twice in the sum (6.8) defining $w'$. Thus we obtain from (6.25), (6.29), and (6.30)

$$t - w + w' > t - 8t^{\frac{1}{2}} - 96t^{\frac{1}{4}}(3w')^{\frac{1}{2}} - \frac{36\mu}{L} + w'. \tag{6.31}$$

We now note that

$$|\mathbf{r}_k' - \mathbf{r}_l'| \leq \sqrt{3}\,L, \tag{6.32}$$

and, therefore,

$$w' \geq \frac{\mu(\mu-1)}{2}\frac{1}{\sqrt{3}\,L} \geq \frac{\mu^2}{4\sqrt{3}\,L}, \tag{6.33}$$

because $\mu \geq 2$. It follows that

$$\frac{w'}{2} - \frac{36\mu}{L} \geq \frac{1}{L}\left(\frac{\mu^2}{8\sqrt{3}} - 36\mu\right) \geq \frac{1}{L}\left(\frac{1+\mu^2}{\alpha} - A_2\right), \tag{6.34}$$

provided that

$$\alpha > 8\sqrt{3} \tag{6.35}$$

and

$$A_2 \geq \frac{1}{\alpha} + \max_{\mu}\left\{36\mu - \mu^2\left(\frac{1}{8\sqrt{3}} - \frac{1}{\alpha}\right)\right\} = \frac{1}{\alpha} + \frac{324}{(1/8\sqrt{3}) - (1/\alpha)}. \quad (6.36)$$

We choose $\alpha$ to satisfy (6.35). Then we choose $A_2$ to satisfy (6.36). Then (6.34) is satisfied, and (6.9) follows from (6.31) if we can choose $A_1$ so that

$$A_1 \geq -t + 8t^{\frac{1}{2}} + 96t^{\frac{1}{4}}(3w')^{\frac{1}{2}} - \tfrac{1}{2}w'. \quad (6.37)$$

On the right-hand side we have a function of the two variables $t$ and $w'$ which is bounded above and so the choice is possible. Numerically, one finds the condition

$$A_1 \geq \max_{t,w'>0}\{-t + 8t^{\frac{1}{2}} + 96t^{\frac{1}{4}}(3w')^{\frac{1}{2}} - \tfrac{1}{2}w'\} = (3.48^2 + 4)^2. \quad (6.38)$$

With (6.35), (6.36), and (6.38) we have proved (6.9). The conditions (6.1), (6.4), and (6.5) now automatically hold and so (5.29) has been proved for arbitrary $\mu \geq 0$. As we have shown in Sec. 4, this implies the truth of Theorem 13. And Theorem 13 implies Theorem 5. So the proof of Theorem 5 is complete except that we still have to prove Lemmas 7, 8, and 9.

It may be of some interest to present our final result in numerical terms. Choosing the reasonable values $\alpha = 30$, $\alpha' = 70$, $A_2 = 10^4$, $A_3^2 = 500$, and $A_1 = 10^8$, we get from (5.12) $A = 1.3 \times 10^{14}$, and this is the constant which appears in Theorem 5.

## 7. PROOF OF LEMMAS 7, 8, AND 9

We begin with a useful extension of Lemma 2 which was proved in I.[1]

*Lemma 2a:* Let $\Omega$ be a domain which contains the sphere $\Omega_0$ of radius $b$ and center at the origin. For any positive $\lambda$ and any complex valued function $\Psi(r)$ with continuous derivatives

$$\int_\Omega d^3r \frac{|\Psi|^2}{|\mathbf{r}|} \leq \left(\frac{1}{\lambda} + \frac{3}{2b}\right)\int_\Omega d^3r\,|\Psi|^2 + \frac{\lambda}{4}\int_\Omega d^3r\,|\nabla\Psi|^2. \quad (7.1)$$

Let $\mathcal{F}(\Psi, \Omega)$ denote the functional

$$\mathcal{F}(\Psi, \Omega) = \left(\frac{\int_\Omega |\nabla\Psi|^2\,d^3r}{\int_\Omega |\Psi|^2\,d^3r}\right)^{\frac{1}{2}} - \frac{\int_\Omega \frac{|\Psi|^2}{|\mathbf{r}|}\,d^3r}{\int_\Omega |\Psi|^2\,d^3r}, \quad (7.2)$$

and let

$$\mu(\Omega) = \inf_{\Psi} \mathcal{F}(\Psi, \Omega). \quad (7.3)$$

By choosing the $\lambda$ in (7.1) optimally, one sees that Lemma 2a is equivalent to the inequality

$$\mu(\Omega) \geq -(3/2b). \quad (7.4)$$

To prove this we take the infimum of (7.3) in two steps.

$$\hat{\mu}(\Omega, \alpha, \beta, \gamma) = \inf_{\Psi\in\mathcal{K}(\alpha,\beta,\gamma)} \mathcal{F}(\Psi, \Omega), \quad (7.5)$$

where $\mathcal{K}(\alpha, \beta, \gamma)$ denotes the class of $\Psi$ which satisfy

$$\int_{\Omega_0} |\Psi|^2\,d^3r = \alpha, \quad (7.6)$$

$$\int_{\Omega} |\Psi|^2\,d^3r = 1, \quad (7.7)$$

$$\int_{\Omega_0} |\nabla\Psi|^2\,d^3r = \beta, \quad (7.8)$$

$$\int_{\Omega} |\nabla\Psi|^2\,d^3r = \gamma, \quad (7.9)$$

with the real numbers $\alpha, \beta, \gamma$ subject to

$$0 \leq \alpha \leq 1 \quad (7.10)$$

and

$$0 \leq \beta \leq \gamma. \quad (7.11)$$

Thus,

$$\mu(\Omega) = \inf_{\alpha,\beta,\gamma} \hat{\mu}(\Omega, \alpha, \beta, \gamma), \quad (7.12)$$

the infimum being over all sets $(\alpha, \beta, \gamma)$ subject to (7.10) and (7.11). If $\Psi$ belongs to $\mathcal{K}(\alpha, \beta, \gamma)$ we have

$$\mathcal{F}(\Psi, \Omega) = \gamma^{\frac{1}{2}} - \int_{\Omega_0} \frac{|\Psi|^2}{|\mathbf{r}|}\,d^3r - \int_{\Omega-\Omega_0} \frac{|\Psi|^2}{|\mathbf{r}|}\,d^3r \geq \gamma^{\frac{1}{2}} - \int_{\Omega_0} \frac{|\Psi|^2}{|\mathbf{r}|}\,d^3r - \frac{1-\alpha}{b}, \quad (7.13)$$

because, for $\mathbf{r} \in \Omega - \Omega_0$, $|\mathbf{r}| \geq b$. By Lemma 2[1]

$$\int_{\Omega_0} \frac{|\Psi|^2}{|\mathbf{r}|}\,d^3r \leq \frac{3\alpha}{2b} + (\alpha\beta)^{\frac{1}{2}}, \quad (7.14)$$

so that for $\Psi \in \mathcal{K}(\alpha, \beta, \gamma)$,

$$\mathcal{F}(\Psi, \Omega) \geq \gamma^{\frac{1}{2}} - (\alpha\beta)^{\frac{1}{2}} - \frac{2+\alpha}{2b}. \quad (7.15)$$

Therefore

$$\hat{\mu}(\Omega, \alpha, \beta, \gamma) \geq \gamma^{\frac{1}{2}} - (\alpha\beta)^{\frac{1}{2}} - \frac{2+\alpha}{2b}. \quad (7.16)$$

The minimization of (7.12) then yields the desired inequality (7.4).

Let $C$ be a cube of edge length $L$, and let $C_1$ be a cube with the same center, with faces parallel to those of $C$ and whose edge length is $2L$. If $\Psi_0(\mathbf{r})$ is a continuously differentiable function in $C$ whose normal derivative vanishes on the faces of $C$, we can extend it

by reflection in the faces of $C$ to a continuously differentiable function $\Psi_1(\mathbf{r})$ in $C_1$. Then

$$\int_{C_1} |\Psi_1|^2 \, d^3r = 8 \int_C |\Psi|^2 \, d^3r \tag{7.17}$$

and

$$\int_{C_1} |\nabla \Psi_1|^2 \, d^3r = 8 \int_C |\nabla \Psi|^2 \, d^3r. \tag{7.18}$$

On the other hand, if $\mathbf{y}_0$ is any point in space,

$$\int_{C_1} \frac{|\Psi_1|^2}{|\mathbf{r} - \mathbf{y}_0|} \, d^3r \geq \int_C \frac{|\Psi|^2}{|\mathbf{r} - \mathbf{y}_0|} \, d^3r. \tag{7.19}$$

Consider now the quantity

$$\left( \frac{\int_C |\nabla \Psi|^2 \, d^3r}{\int_C |\Psi|^2 \, d^3r} \right)^{\frac{1}{2}} - \frac{1}{8} \frac{\int_C \frac{|\Psi|^2}{|\mathbf{r} - \mathbf{y}|} d^3r}{\int_C |\Psi|^2 \, d^3r} = \hat{\mathcal{F}}(\Psi, C, \mathbf{y}). \tag{7.20}$$

If $\mathbf{y}$ is not in $C$ then $\hat{\mathcal{F}}$ is a harmonic function of $\mathbf{y}$ and so it is minimized for some $\mathbf{y}$ on the boundary of $C$. Thus we may assume $\mathbf{y}$ to be in or on the boundary of $C$. Let $\Psi_0(\mathbf{r})$ and $\mathbf{y}_0$ be such that they minimize $\hat{\mathcal{F}}$. Because no boundary condition is imposed on $\Psi$ in the minimization, the variational problem automatically gives the "natural boundary condition" of vanishing normal derivative for $\Psi_0$ on the faces of $C$. Thus we may extend $\Psi_0$ as explained above, and from (7.17), (7.18), and (7.19) it follows that

$$\hat{\mathcal{F}}(\Psi, C, \mathbf{y}) \geq \hat{\mathcal{F}}(\Psi_0, C, \mathbf{y}_0) > \mathcal{F}(\Psi_1, C_1), \tag{7.21}$$

provided we shift the origin of the coordinate system to $\mathbf{y}_0$. But the distance of $\mathbf{y}_0$ from the faces of $C_1$ is at least $\frac{1}{2}L$ so that we may apply Lemma 2a with $b = \frac{1}{2}L$,

$$\mathcal{F}(\Psi_1, C_1) \geq -3/L. \tag{7.22}$$

The resulting inequality

$$\hat{\mathcal{F}}(\Psi, C, \mathbf{y}) > -\frac{3}{L} \tag{7.23}$$

is equivalent to (6.3). This completes the proof of Lemma 7.

The proof of Lemma 8 is similar. We have

$$\hat{\mathcal{F}}(\Psi, D, 0) \geq \hat{\mathcal{F}}(\Psi_0, D, 0), \tag{7.24}$$

where $\Psi_0$ is the minimizing function and therefore has vanishing normal derivative on the boundary of $D$. We now extend $\Psi_0$ by reflection on the faces of the given cube to a function $\Psi_1$ defined in the complete sphere $D_1$: $|\mathbf{r}| \leq b$. Because the origin $\mathbf{r} = 0$ is in $D$ we see that

$$\int_{D_1} |\Psi_1|^2 \, d^3r \leq 8 \int_D |\Psi_0|^2 \, d^3r \tag{7.25}$$

and

$$\int_{D_1} |\nabla \Psi_1|^2 \, d^3r \leq 8 \int_D |\nabla \Psi_0|^2 \, d^3r. \tag{7.26}$$

Also

$$\int_{D_1} \frac{|\Psi_1|^2}{|\mathbf{r}|} > \int_D \frac{|\Psi_0|^2}{|\mathbf{r}|}. \tag{7.27}$$

Note that $\hat{\mathcal{F}}(\Psi_0, D, 0) < 0$, since the left-hand side of (7.24) is negative for $\Psi$ = constant. This, together with the last three inequalities implies

$$\hat{\mathcal{F}}(\Psi_0, D, 0) > \mathcal{F}(\Psi_1, D_1) > -\frac{3}{2b} \tag{7.28}$$

by Lemma 2.[1] The resulting inequality

$$\hat{\mathcal{F}}(\Psi, D, 0) > -\frac{3}{2b} \tag{7.29}$$

is equivalent to (6.21), as seen by choosing $\lambda$ to give the strictest inequality. This proves Lemma 8.

It remains to prove Lemma 9. It is clear that the size of the cube $C$ is irrelevant. Therefore we choose $C$ to have edge length $2\pi$. We also take the origin at the center of $C$.

We first show that it is sufficient to prove the Lemma with the constant 8 replaced by 1 if $f$ and $g$ are required to be periodic with respect to opposite faces of $C$. Indeed, let

$$Q(f, g) = \left| \int_C g \, |f|^2 \, d^3r \right|^2 \Big/ \int_C |\nabla g|^2 \, d^3r \left( \int_C |\nabla f|^2 \, d^3r \right)^{\frac{1}{2}}, \tag{7.30}$$

and let $f_0$ and $g_0$ be the functions that maximize $Q$. Then $f_0$ and $g_0$ have vanishing normal derivatives on the faces of $C$, and they may be extended by reflection on the faces to functions $f_1$ and $g_1$ defined in the larger cube $C_1$, which is concentric with $C$, has faces parallel with those of $C$, and is twice the linear size. Now we put

$$\begin{cases} f_2(\mathbf{r}) = f_1(2\mathbf{r}) \\ g_2(\mathbf{r}) = g_1(2\mathbf{r}). \end{cases} \tag{7.31}$$

$f_2$ and $g_2$ are periodic in $C$, satisfy (6.26), (6.27), respectively, and

$$Q(f_2, g_2) = 1/8 Q(f_0, g_0). \tag{7.32}$$

Thus

$$Q(f, g) \leq Q(f_0, g_0) = 8 Q(f_2, g_2), \tag{7.33}$$

and it is sufficient to show

$$Q(f, g) \leq \frac{1}{\pi}, \quad f, g \text{ periodic}; \tag{7.34}$$

then by (7.33) the desired inequality (6.28) follows.

Expand $f$ and $g$ in Fourier series

$$f(\mathbf{r}) = (2\pi)^{-\frac{3}{2}} \sum_{\mathbf{n}} a_{\mathbf{n}} e^{i\mathbf{n}\cdot\mathbf{r}} \tag{7.35}$$

and

$$g(\mathbf{r}) = (2\pi)^{-\frac{3}{2}} \sum_{\mathbf{n}} b_{\mathbf{n}} e^{i\mathbf{n}\cdot\mathbf{r}}, \tag{7.36}$$

the sums running over vectors $\mathbf{n} = (n_x, n_y, n_z)$ with integer components. The normalization of $f$ is

$$\sum_{\mathbf{n}} |a_{\mathbf{n}}|^2 = 1, \tag{7.37}$$

and

$$b_0 = 0. \tag{7.38}$$

In terms of the Fourier coefficients,

$$Q(f, g) = \frac{1}{8\pi^3} \left| \sum_{\mathbf{n}} \sum_{\mathbf{m}} a_{\mathbf{n}}^* a_{\mathbf{m}} b_{\mathbf{n}-\mathbf{m}} \right| \Big/ \sum |\mathbf{n}|^2 |b_{\mathbf{n}}|^2 \left( \sum_{\mathbf{n}} |\mathbf{n}|^2 |a_{\mathbf{n}}|^2 \right)^{\frac{1}{2}}. \tag{7.39}$$

By Schwarz's inequality

$$\left| \sum_{\mathbf{n}} \sum_{\mathbf{m}} a_{\mathbf{n}}^* a_{\mathbf{m}} b_{\mathbf{n}-\mathbf{m}} \right|^2 \leq \sum_{\mathbf{k}\neq 0} |\mathbf{k}|^2 |b_{\mathbf{k}}|^2 \sum_{\mathbf{k}\neq 0} \frac{1}{|\mathbf{k}|^2} \left| \sum_{\mathbf{m}} a_{\mathbf{m}+\mathbf{k}}^* a_{\mathbf{m}} \right|^2. \tag{7.40}$$

Therefore (7.34) follows from

$$\sum_{\mathbf{k}\neq 0} \frac{1}{|\mathbf{k}|^2} \left| \sum_{\mathbf{m}} a_{\mathbf{m}+\mathbf{k}}^* a_{\mathbf{m}} \right|^2 \leq 8\pi^2 \left( \sum_{\mathbf{n}} |\mathbf{n}|^2 |a_{\mathbf{n}}|^2 \right)^{\frac{1}{2}}. \tag{7.41}$$

By Schwarz's inequality

$$\left| \sum_{\mathbf{m}} a_{\mathbf{m}+\mathbf{k}}^* a_{\mathbf{m}} \right| \leq \sum_{\mathbf{m}} |a_{\mathbf{m}}|^2 = 1 \tag{7.42}$$

for all $\mathbf{k}$. Therefore, the left-hand side of (7.41) is not more than

$$\sum_{\mathbf{k}\neq 0} \frac{1}{|\mathbf{k}|^2} \left| \sum_{\mathbf{m}} a_{\mathbf{m}+\mathbf{k}}^* a_{\mathbf{m}} \right| \leq \sum_{\mathbf{k}\neq 0} \frac{1}{|\mathbf{k}|^2} \sum_{\mathbf{m}} |a_{\mathbf{m}+\mathbf{k}}| \cdot |a_{\mathbf{m}}|. \tag{7.43}$$

It is therefore sufficient to prove that

$$\sum_{\mathbf{k}\neq 0} \sum_{\mathbf{m}} \frac{a_{\mathbf{m}+\mathbf{k}} a_{\mathbf{m}}}{|\mathbf{k}|^2} \leq 8\pi^2 \left( \sum_{\mathbf{n}} |\mathbf{n}|^2 a_{\mathbf{n}}^2 \right)^{\frac{1}{2}} \tag{7.44}$$

for any real positive $a_{\mathbf{n}}$ normalized by (7.37).

Let, for any positive $\lambda$,

$$\sigma_{\mathbf{n}} = \sigma_{\mathbf{n}}(\lambda) = \sum_{\mathbf{m}\neq 0,-\mathbf{n}} \frac{1}{|\mathbf{m}|^2 (|\mathbf{m}+\mathbf{n}|^2 + \lambda^2)}. \tag{7.45}$$

The elementary inequality

$$[|\mathbf{n}|^2 (|\mathbf{m}|^2 + \lambda^2)]^{-1} + [|\mathbf{m}|^2 (|\mathbf{n}|^2 + \lambda^2)]^{-1} \leq [|\mathbf{n}|^2 (|\mathbf{n}|^2 + \lambda^2)]^{-1} + [|\mathbf{m}|^2 (|\mathbf{m}|^2 + \lambda^2)]^{-1} \tag{7.46}$$

shows that

$$\sigma_{\mathbf{n}} \leq \sigma_0. \tag{7.47}$$

We derive an upper bound for $\sigma_0$ as follows. The function $\phi(\mathbf{r}) = |\mathbf{r}|^{-2}(|\mathbf{r}|^2 + \lambda^2)^{-1}$ satisfies $\nabla^2\phi > 0$, and therefore its value $\phi(\mathbf{n})$ at any point with integer coordinates is less than its average over a sphere centered at this point. We choose for all $\mathbf{n}$ spheres of radii $\frac{1}{2}$ so that they do not overlap. This way we obtain the estimate

$$\sigma_0 < \frac{6}{\pi} \int_{|\mathbf{r}|\geq\frac{1}{2}} \phi(r)\, d^3\mathbf{r} = \frac{24}{\lambda} \arctan 2\lambda < 12\pi^2/(1 + \pi\lambda). \tag{7.48}$$

The last inequality follows from

$$\frac{1}{\cos^2 t} > \frac{1}{(\pi/2 - t)^2} \quad \left(0 \leq t < \frac{\pi}{2}\right). \tag{7.49}$$

Integrating between 0 and $\frac{1}{2}\pi^2\lambda(1 + \pi\lambda)^{-1}$ one has

$$\tan \frac{\pi^2\lambda}{2(1 + \pi\lambda)} > 2\lambda, \tag{7.50}$$

which yields (7.48).

We now consider the sum on the left-hand side of (7.44). Let $p_{\mathbf{n}} = p_{-\mathbf{n}} > 0$. We have then by Schwarz's inequality

$$\sum_{\mathbf{k}\neq 0} \sum_{\mathbf{m}} \frac{a_{\mathbf{m}+\mathbf{k}} a_{\mathbf{m}}}{|\mathbf{k}|^2} = \sum_{\mathbf{k}\neq 0} \sum_{\mathbf{m}} \frac{a_{\mathbf{m}+\mathbf{k}} p_{\mathbf{m}+\mathbf{k}}}{|\mathbf{k}| p_{\mathbf{m}}} \times \frac{a_{\mathbf{m}} p_{\mathbf{m}}}{|\mathbf{k}|\, p_{\mathbf{m}+\mathbf{k}}} < \sum_{\mathbf{k}\neq 0} \sum_{\mathbf{m}} \frac{a_{\mathbf{m}}^2 p_{\mathbf{m}}^2}{|\mathbf{k}|^2\, p_{\mathbf{m}+\mathbf{k}}^2}. \tag{7.51}$$

Let

$$p_n = \begin{cases} |\mathbf{n}|^2 + \lambda^2 & (\mathbf{n} \neq 0), \\ 2\pi \left( \dfrac{\lambda}{\sigma_0} \right)^{\frac{1}{2}} & (\mathbf{n} = 0). \end{cases} \tag{7.52}$$

Then the right-hand side of (7.51) may be written

$$4\pi^2\lambda a_0^2 + \sum_{\mathbf{m}\neq 0} (|\mathbf{m}|^2 + \lambda^2) a_{\mathbf{m}}^2 \left( \sigma_{\mathbf{m}} + \frac{\sigma_0}{|2\pi\mathbf{m}|^2\lambda} \right) \leq 4\pi^2\lambda a_0^2 + \frac{6\pi}{\lambda} \frac{1 + 2\pi\lambda}{1 + \pi\lambda} \sum_{\mathbf{m}\neq 0} (|\mathbf{m}|^2 + \lambda^2) a_{\mathbf{m}}^2 \tag{7.53}$$

by virtue of (7.47) and (7.48). Since

$$[6\pi(1 + 2\pi\lambda)]/(1 + \pi\lambda) < 4\pi^2, \tag{7.54}$$

we get an upper bound on the right-hand side of (7.53)

$$\frac{4\pi^2}{\lambda} \sum_{\mathbf{m}} (|\mathbf{m}|^2 + \lambda^2) a_{\mathbf{m}}^2. \tag{7.55}$$

Finally, we choose $\lambda$ to minimize this upper bound. This yields the inequality (7.44) and the proof of Lemma 9 is complete.

We conclude with some informal remarks which are intended to clarify the physical meaning of Lemma 9. In our whole proof of stability, Lemma 9 is the innermost core. It is unfortunate that both the statement and the proof of Lemma 9 are couched in algebraic terms which conceal the physical motivations. Basically, the effect of Lemma 9 is to set a bound to the interaction energy between negative and positive charges, the bound depending only on the kinetic energy of the negatives and on the potential energy of the positives.

An alternative statement of the content of Lemma 9 is obtained by analyzing the inequality (7.44) which is essentially equivalent to Eq. (6.28). We can give Eq. (7.44) an intuitive meaning as follows: The function

$$R(\mathbf{x}) = \sum_{\mathbf{q}\neq 0} \frac{1}{\pi L q^2} \exp\left[\frac{2\pi i}{L}\mathbf{q}\cdot\mathbf{x}\right] \qquad (7.56)$$

is the Green's function for a cube of side $L$ with periodic boundary conditions. $R(\mathbf{x})$ is the Coulomb potential generated by a unit positive charge at each vertex of a cubic lattice, with a constant negative background density to preserve neutrality. Equation (7.44) is equivalent to the statement that a particle of mass $m$ in the periodic Coulomb potential $[-e^2R(\mathbf{r})]$ has a ground-state binding energy less than 16 Ry, or in symbols

$$e^2\int R(\mathbf{r})\,|\psi(\mathbf{r})|^2\,d^3r < (\hbar^2/2m)\int|\nabla\psi(\mathbf{r})|^2\,d^3r + 16\ \mathrm{Ry}\int|\psi(\mathbf{r})|^2\,d^3r, \qquad (7.57)$$

irrespective of the spacing $L$ of the lattice. The awkwardness of the proof of Eq. (7.44) arose from the fact that $R(\mathbf{r})$ is not spherically symmetric and so the standard argument based on Lemma 2 is not applicable. Note that $R(\mathbf{r})$ is defined to have mean value zero, so that there is no term proportional to $e^2$ in the ground-state energy.

We know that in the limit $L \to \infty$, when $R(\mathbf{r})$ tends to the Coulomb potential $(1/r)$, the ground-state energy is one Rydberg, and so Eq. (7.57) holds without the factor 16. It is also easy to verify that, as $L \to 0$, Eq. (7.57) holds with a coefficient smaller than 1 instead of the 16. It is extremely likely that Eq. (7.57) is true for all $L$ with the 16 replaced by 1. This would mean that Eq. (7.44) holds with 2 replacing 8, and likewise Eq. (6.28) would hold with 2 replacing 8 on the right side. To prove this strengthened form of Lemma 9, it would be sufficient to show that the ground-state binding energy of the potential $[-e^2R(\mathbf{r})]$ decreases monotonically with $L$. The monotonicity of the ground-state energy is physically plausible, but we have not succeeded in proving it. Our proof of Lemma 9 misses a factor of 4 through various crudities of detail, particularly in the estimate (7.48) for $\sigma_0$.

When Lemma 9 is expressed in the form (7.57), its relation to the arguments of Sec. 2 becomes clear. In Sec. 2, Theorem 11 was deduced from the fact that the ground-state binding energy of a particle in a Coulomb potential in infinite space is one Rydberg. Analogously, Lemma 9 is deduced from the fact that the ground-state binding energy in the periodic Coulomb potential is less than 16 Ry. The simple Coulomb potential is the Green's function for infinite space, while the periodic Coulomb potential is the Green's function for a finite cube. The logical structure of the proofs of the two inequalities, Theorem 11 and Lemma 9, is the same, and the logic of their use in the proofs of Theorem 12 and Theorem 5 is also the same. Only the details are more complicated for the case of the cube, and the numerical coefficients are correspondingly less precise.

It is a remarkable fact that our proof of the stability of matter, after such a tremendous detour via dissections of space and other artificial tricks, boils down in the end to an estimate of the binding energy of a single electron in a periodic Coulomb potential. We conjecture that this appearance of the periodic Coulomb potential at the kernel of the proof is not accidental. After all, the ground states of most forms of matter are crystals in which electrons are actually moving in periodic Coulomb potentials. The essence of a proof of the stability of matter should be a demonstration that an aperiodic arrangement of particles cannot give greater binding than a periodic arrangement. If this dominance of the periodic potential could be proved directly, then we would have a proof of Theorem 5 vastly simpler and more satisfactory than the one presented in this paper. Into the bargain we would also undoubtedly find a more reasonable numerical coefficient than $1.3 \times 10^{14}$.

## ACKNOWLEDGMENT

The first-named author wishes to express his grateful appreciation for the hospitality of the Institute for Advanced Study during the year 1965–66 when this work was done.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 5 MAY 1968

# Orear Behavior in Potential Scattering*

PORTER JOHNSON†
*Palmer Physical Laboratory, Princeton University, Princeton, New Jersey*

(Received 6 September 1967)

The object of this paper is to find and study a class of potentials for which the corresponding scattering amplitude $f$ decreases rapidly in energy at fixed (nonforward) angles. Specifically, we ask that $f(k, \theta) = O[\exp\{-b(\theta)k\}]$ as $k \to \infty$ for $\theta$ fixed. It is shown here that this relation is valid for certain potentials $V(r)$ which are even functions of $r$ analytic in a strip about the real $r$ axis. With further restrictions on the potentials we show that the scattering amplitude converges to its first Born approximation at high energies for fixed nonforward angles.

## 1. INTRODUCTION

It is a striking experimental result that elastic differential cross sections for processes involving nucleons decrease very rapidly with respect to energy at fixed nonforward scattering angles,[1] i.e., that they exhibit Orear behavior. This behavior appears to be a general feature of strong interaction scattering processes. It has been qualitatively described as a manifestation of the spatial extension of composite particles.[2] Since it is so general a feature of strong interactions, we have looked into nonrelativistic potential scattering to see under what conditions this behavior occurs there. We do not propose to invent a serious physical model of any sort for this behavior—we work within the context of potential scattering for guidance.[3] (We of course recall that Regge behavior, analyticity in energy, and the Mandelstam representation all have conceptual validity in potential scattering, so that the conditions for Orear behavior in potential scattering may be of some interest.) Potential scattering is an internally consistent theory, it does give unambiguous results, and finally, it is a relatively simple theory in which to work.

The basic experimental behavior pointed out by Orear is that the elastic differential cross section decreases roughly exponentially with respect to energy at fixed nonforward scattering angles. For the purposes of this paper we understand by "Orear behavior" the following asymptotic relation upon the scattering amplitude[4]:

$$f(k, \theta) \sim O[\exp(-2kb \sin \tfrac{1}{2}\theta)], \tag{1.1}$$

with $b > 0$, as the wavenumber $k \to \infty$ with the scattering angle $\theta$ fixed away from the forward direction.[5]

The object here is to find a class of (nonsingular) central potentials $V(r)$ such that (1.1) is valid for the corresponding scattering amplitude. It is obviously necessary that the first Born approximation $f_1$ obey the order relation (1.1) in order that (1.1) be valid for $f$ for sufficiently weak potentials.[6] One can easily see that $f_1$ obeys (1.1) for potentials which are even in $r$ and analytic in $r$ in a strip about the real axis. For definiteness we limit our considerations to potentials which are analytic in the region $|\mathrm{Im}\, r| < r_0$ and have the following representation for positive $r_0$[7]:

$$V(r) = \int_0^\infty d\alpha\sigma(\alpha) \exp[-\alpha(r^2 + r_0^2)]. \tag{1.2}$$

The following conditions are placed upon the weight function $\sigma(\alpha)$:

(a) $\sigma(\alpha)$ is continuous for $\alpha > 0$, and it possesses a finite derivative almost everywhere,

(b) $\lim_{\alpha \to 0^+} \sigma(\alpha) = 0$, and

(c) the integral

$$Q(\epsilon) = \int_0^\infty d\alpha\alpha^{-\frac{3}{4}} |\sigma'(\alpha)| \exp(-\alpha\epsilon) \tag{1.3}$$

converges for every $\epsilon > 0$. We then show in Sec. 2 that

* This paper is based upon a thesis by the author submitted to Princeton University in partial fulfillment of the requirements of the degree of Doctor of Philosophy. The author was a National Science Foundation Predoctoral Fellow and National Aeronautics and Space Administration Trainee while working on his thesis.

† Present address: Department of Physics, Case Western Reserve University, Cleveland, Ohio.

[1] J. Orear, Phys. Rev. Letters **12**, 112 (1964).

[2] T. T. Wu and C. N. Yang, Phys. Rev. **137**, B708 (1965).

[3] However, see M. Islam and J. Rosen, Phys. Rev. Letters **19**, 178 (1967). For considerations with singular potentials, see G. Tiktopoulos, Phys. Rev. B1550 (1965).

[4] The original Orear formula for elastic $p$–$p$ scattering, $d\sigma/d\Omega \sim \exp\{-p_\perp/p_0\}$, inevitably takes exchange scattering processes into account. One could include an exchange potential here to approximate the physical behavior more closely, but such is not our purpose.

[5] High-energy, large-angle scattering is discussed in T. T. Wu, Phys. Rev. **143**, 1110 (1966). For consideration of the related concept of Schrödinger equation models of form factors, see S. D. Drell, A. C. Finn, and M. H. Goldhaber, Phys. Rev. **157**, 1402 (1967).

[6] The fact that $f_1$ satisfies (1.1) is by no means sufficient to guarantee that $f$ satisfies (1.1) as well!

[7] Conditions (a) to (c) here are sufficient to guarantee analyticity in the strip.

if the potential has a representation of the form (1.2) subject to conditions (a) to (c), the corresponding scattering amplitude satisfies relation (1.1) for every $b < r_0$. An example of such a potential is $V(r) = (r^2 + r_0^2)^{-n}$ for $n > \frac{7}{4}$.

One may obtain more restrictive asymptotic bounds upon the scattering amplitude at higher energies if further conditions are imposed upon $\sigma$. With this in mind we place an additional condition upon $\sigma$.

(d) The following integrals converge:

(i) $\int_0^\infty d\alpha\, |\sigma(\alpha)|$,

(ii) $\int_0^\infty d\alpha\, |\sigma'(\alpha)|$.

It is shown in Sec. 2 that if $\sigma$ meets conditions (a) to (d), the scattering amplitude obeys relation (1.1) with $b = r_0$ as well. Furthermore, in Sec. 3 additional conditions are placed upon $\sigma$.

(e) $\sigma(\alpha) \geq 0$ for $\alpha \geq 0$.

(f) For positive $\alpha$ there exists some positive number $\lambda$ such that $\sigma'(\alpha) + \lambda^2\sigma(\alpha) \geq 0$.

(g) For some number $p \geq 0$ the following integrals converge for $0 \leq q \leq p$:

(i) $\int_0^\infty d\alpha\alpha^q\, |\sigma(\alpha)|$,

(ii) $\int_0^\infty d\alpha\alpha^{q+1}\, |\sigma'(\alpha)|$,

(iii) $\int_0^\infty d\alpha\alpha^{q+1}\, |\sigma''(\alpha)|$.

(h) Finally, for $p$ as chosen above there exist constants $K$ and $\alpha_0$ such that $\sigma(\alpha) \geq K/\alpha^{p+2}$ for $\alpha > \alpha_0$.

Under conditions (a) to (h) upon $\sigma$ we will establish the following limit involving the scattering amplitude $f(k, \theta)$ and its first Born approximation $f_1(k, \theta)$:

$$\lim_{\substack{k\to\infty \\ \theta \text{ fixed}}} f(k, \theta)/f_1(k, \theta) = 1. \tag{1.4}$$

It should be noted that the limit of the above function at fixed momentum transfer is known to be unity for a wide class of potentials. However, the limit here is taken at fixed angles of scattering.

An example of a potential meeting conditions (a) to (h) is $V(r) = \lambda \exp[-\alpha_0(r^2 + r_0^2)^{\frac{1}{2}}]$, for which the corresponding weight function is

$$\sigma(\alpha) = (\lambda/2\pi^{\frac{1}{2}})(\alpha_0/\alpha^{\frac{3}{2}}) \exp(-\alpha_0^2/4\alpha).$$

In Sec. 4 we speculate that relation (1.1) is valid for members of a broader class of potentials than (1.2).

## 2. OREAR BEHAVIOR

We define the scattering amplitude through the usual asymptotic limit of positive energy solutions of the time-independent Schrödinger equation. From the Schrödinger equation one can obtain the familiar Born series expansion of the scattering amplitude, which is written formally as follows:

$$f = V + VG_kV + VG_kVG_kV + \cdots, \tag{2.1}$$

where $V$ is the potential and $G_k$ is the free Green's function. We factor the potential formally through the relation $V(r) = v_1(r) \cdot v_2(r)$. Then the Born series may be written formally as

$$f = V + v_1Wv_2 + v_1W^2v_2 + \cdots, \tag{2.2}$$

where $W = v_2G_kv_1$.

Let us define the Fourier transform of the potential through the relation

$$V(\mathbf{q}) = (2\pi)^{-3}\int d\mathbf{x}V(x) \exp(-i\mathbf{q}\cdot\mathbf{x}). \tag{2.3}$$

The Fourier conjugates to $v_1$ and $v_2$ are similarly defined. Then the formal relation (2.2) can be written as

$$\begin{aligned} f(k, k') = 2\pi^2\Big[ &-V(k' - k) \\ &+ \int dp_1\, dp_2 v_2(p_1 - k)W(p_1, p_2)v_1(k' - p_2) \\ &- \int dp_1\, dp_2\, dp_3 v_2(p_1 - k)W(p_1, p_2) \\ &\times W(p_2, p_3)v_1(k' - p_3) + \cdots\Big]. \end{aligned} \tag{2.2'}$$

Then the expression for $W$ is

$$\begin{aligned} W(p_1, p_2) = (4\pi)^{-1}\int \frac{d\mathbf{x}_1\, d\mathbf{x}_2}{(2\pi)^3} v_2(x_1)v_1(x_2) \\ \times \exp\{-i(\mathbf{p}_2\cdot\mathbf{x}_1 - \mathbf{p}_1\cdot\mathbf{x}_2)\} \\ \times \exp\{ik\,|\mathbf{x}_1 - \mathbf{x}_2|\}/|\mathbf{x}_1 - \mathbf{x}_2|. \end{aligned} \tag{2.4}$$

We now require that the potential factors $v_1$ and $v_2$ have representations analogous to (1.2) with the same number $r_0$ as in the representation (1.2), i.e., there must exist $\rho_1$ and $\rho_2$ such that

$$\begin{matrix} v_1(r) \\ 2 \end{matrix} = \int_0^\infty d\alpha \begin{matrix} \rho_1(\alpha) \\ 2 \end{matrix} \exp[-\alpha(r^2 + r_0^2)]. \tag{2.5}$$

The object here is to obtain a relatively simple bound upon $W$. Let us use the representation (2.5) of the potential factors to write

$$\begin{aligned} W(p_1, p_2) = \frac{1}{32\pi^4}\iint_0^\infty d\alpha\, d\beta\rho_1(\alpha)\rho_2(\beta) \\ \times \exp[-\alpha(r^2 + r_0^2)]J(\alpha, \beta, p_1, p_2). \end{aligned}$$

The following bound upon $J$ can be obtained:

$$\begin{aligned} |J(\alpha, \beta, p_1, p_2)| \leq 4\pi^3k^{-1}(\alpha + \beta)^{-1}(\alpha\beta)^{-\frac{1}{2}} \\ \times \exp(-(p_2 - p_1)^2/4(\alpha + \beta)). \end{aligned}$$

As a result we derive a bound upon $|W|$ of the form

$$|W(p_1, p_2)| \leq \frac{K(p_2 - p_1)}{k} = (8\pi k)^{-1} \iint_0^\infty \frac{d\alpha\, d\beta}{\alpha + \beta} \frac{|\rho_1(\alpha)|\,|\rho_2(\beta)|}{[\alpha\beta]^{\frac{1}{2}}} \times \exp\left[-(\alpha+\beta)r_0^2 - (p_2 - p_1)^2/4(\alpha+\beta)\right]. \tag{2.6}$$

The bound (2.6) of $W$ may be used to bound each term of the Born series (2.2) in a relatively simple way.

Let us define functions $\bar{v}_1$ and $\bar{v}_2$ by the relations

$$\bar{v}_1(r) = \int_0^\infty d\alpha\, |\rho_1(\alpha)|\, \exp\left[-\alpha(r^2 + r_0^2)\right]$$

and let $\mathfrak{V}(r) = \bar{v}_1(r) \cdot \bar{v}_2(r)$. We define the Fourier transforms of $\bar{v}_1$ and $\bar{v}_2$ as in the relation (2.3) for $V$. Then the scattering amplitude may be bounded by the following series:

$$\begin{aligned} |f(k' - k)| &\leq 2\pi^2 \Big[ |V(k' - k)| \\ &\quad + k^{-1} \int dp_1\, dp_2 \bar{v}_2(p_1 - k) K(p_2 - p_1) \bar{v}_1(k' - p_2) \\ &\quad + k^{-2} \int dp_1\, dp_2\, dp_3 \bar{v}_2(p_1 - k) K(p_2 - p_1) \\ &\quad \times K(p_3 - p_2) \bar{v}_1(k' - p_3) + \cdots \Big]. \end{aligned}$$

The following theorem on convolution integrals may be applied:

$$\begin{aligned} &\int dp_1 \cdots dp_n a_1(q - p_1) a_2(p_1 - p_2) \\ &\qquad \times a_3(p_2 - p_3) \cdots a_n(p_{n-1} - p_n) a_{n+1}(p_n - t) \\ &= (2\pi)^{-3} \int dx a_1(x) a_2(x) a_3(x) \cdots a_n(x) a_{n+1}(x) \\ &\qquad \times \exp\left[-i(q - t) \cdot x\right], \end{aligned}$$

where $a_i(p)$ and $a_i(x)$ are Fourier conjugates as in (2.3). Thus the modulus of the scattering amplitude may be bounded as follows:

$$\begin{aligned} &|f(k, \Delta)| \\ &\leq 2\pi^2 |V(\Delta)| \\ &\quad + \left| \frac{1}{4\pi} \int dx \bar{v}_1(x) \bar{v}_2(x) \left[ \frac{K(x)}{k} + \left( \frac{K(x)}{k} \right)^2 + \cdots \right] \right. \\ &\qquad \left. \times \exp\left\{-i\Delta \cdot x\right\} \right| \\ &\leq 2\pi^2 |V(\Delta)| \\ &\quad + \frac{1}{4\pi k} \left| \int dx \mathfrak{V}(x) \frac{K(x)}{1 - [K(x)/k]} \exp\left\{-i\Delta \cdot x\right\} \right|. \end{aligned} \tag{2.7}$$

We can easily obtain the following expressions for $\mathfrak{V}(x)$ and $K(x)$ from their definitions:

$$\mathfrak{V}(x) = \iint_0^\infty d\alpha\, d\beta\, |\rho_1(\alpha)|\, |\rho_2(\beta)| \times \exp\left\{-(\alpha + \beta)(r^2 + r_0^2)\right\}, \tag{2.8}$$

$$K(x) = \pi^{\frac{1}{2}} \iint_0^\infty d\alpha\, d\beta\, |\rho_1(\alpha)|\, |\rho_2(\beta)|\, [(\alpha + \beta)/(\alpha\beta)]^{\frac{1}{2}} \times \exp\left\{-(\alpha + \beta)(r^2 + r_0^2)\right\}. \tag{2.9}$$

The formal manipulations of this section can be justified if the following conditions are met:

(a) $V(r)$ is finite for real $r$ and is square integrable as a function of $\mathbf{r}$.

(b) There exists a suitable factorization $v_1 \cdot v_2 = V$ such that $v_1(r)$, $v_2(r)$, $\bar{v}_1(r)$, $\bar{v}_2(r)$, and $\mathfrak{V}(r)$ are similarly well-defined and square-integrable functions of $\mathbf{r}$.

(c) $K(p)$ given in Eq. (2.6) is absolutely integrable as well as square integrable over $\mathbf{p}$.

(d) The conjugate function $K(x)$ given in Eq (2.9) is bounded uniformly by some number $C$ for all real $x$.

With these four conditions one can prove that the steps leading to equality (2.7) are rigorous and that the right side of (2.7) is finite for $k > C$.

One can prove that conditions (a) to (d) of this section are met if there exists a factorization $v_1(x) \cdot v_2(x) = V(x)$ such that the corresponding weight functions $\rho_1$ and $\rho_2$ satisfy the following conditions:

$$L_1(\epsilon) = \int_0^\infty d\alpha \alpha^{-\frac{3}{4}} |\rho_1(\alpha)|\, \exp\left(-\alpha\epsilon\right) \tag{2.10}$$

must be finite for every $\epsilon > 0$. We will prove (d) explicitly, to illustrate the method used in proving the others,

$$\begin{aligned} K(x) &\leq K(0) \\ &= \pi^{\frac{1}{2}} \iint_0^\infty d\alpha\, d\beta [(\alpha + \beta)/(\alpha\beta)]^{\frac{1}{2}} |\rho_1(\alpha)|\, |\rho_2(\beta)| \\ &\qquad \times \exp\left[-(\alpha + \beta) r_0^2\right]. \end{aligned}$$

We let $r_0^2 = \epsilon_1 + \epsilon_2$, where $\epsilon_1$ and $\epsilon_2$ are both positive. Thus

$$\begin{aligned} K(0) &= \pi^{\frac{1}{2}} \int_0^\infty d\alpha \alpha^{-\frac{3}{4}} |\rho_1(\alpha)|\, \exp\left[-\alpha\epsilon_1\right] \\ &\quad \times \int_0^\infty d\beta \beta^{-\frac{3}{4}} |\rho_2(\beta)|\, \exp\left[-\beta\epsilon_1\right] \mathcal{K}(\alpha, \beta, \epsilon_2). \end{aligned}$$

One can show that $\mathcal{K}(\alpha, \beta, \epsilon_2) \leq (e\epsilon_2)^{-1}$, so that $K(0) \leq (\pi^{\frac{1}{2}}/e\epsilon_2) L_1(\epsilon_1) L_2(\epsilon_1)$, which is finite.

We will now show that if conditions (a) to (c) of Sec. 1 are met, a suitable factorization subject to (2.10)

exists. We choose to factor $V(r)$ into

$$v_1(r) = (r^2 + r_0^2)V(r), \qquad v_2(r) = (r^2 + r_0^2)^{-1}. \tag{2.11}$$

One can see explicitly that $\rho_2(\alpha) = 1$ and $\rho_1$ and $\sigma$ are related by the integral equation $\sigma(\alpha) = \int_0^\alpha d\beta \rho_1(\beta)$. Under the assumption of conditions (a) and (b) of Sec. 1 the solution to this equation is

$$\rho_1(\alpha) = \sigma'(\alpha).$$

Condition (c) of Sec. 1 implies the finiteness of $L_1(\epsilon)$, whereas one may explicitly calculate $L_2$ to obtain $L_2(\epsilon) = \Gamma(\frac{1}{4})\epsilon^{-\frac{1}{4}}$. Thus this factorization meets conditions (2.10), so that the inequality is valid under conditions (a) to (c) of Sec. 1.

We wish to obtain a bound of the form (1.1) upon the scattering amplitude. For that purpose we rewrite the inequality (7) as

$$|f(k,\Delta)| \leq (4\pi)^{-1}\left|\int d\mathbf{x}\mathfrak{V}(x)\left[1 - \frac{K(x)}{k}\right]^{-1}\exp(-i\boldsymbol{\Delta}\cdot\mathbf{x})\right|. \tag{2.7'}$$

Since $\mathfrak{V}$ and $K$ depend only upon $|x|$ and are even in $|x|$, we can rewrite the inequality as

$$|f(k,\Delta)| = (2i\Delta)^{-1}\left|\int_{-\infty}^{\infty} dx\mathfrak{V}(x)\left[1 - \frac{K(x)}{k}\right]^{-1}\exp(i\Delta x)\right|. \tag{2.12}$$

We pick a number $b < r_0$, and we distort the contour of $x$ integration in (2.12) from the real axis to the line $\operatorname{Im} x = b$. (We will see that such a distortion is justified.) We let $\epsilon = r_0^2 - b^2$. Then for real $s$ we can use (2.9) to show that

$$|K(s+ib)| \leq K(ib) = \pi^{\frac{1}{2}}\int_0^\infty\int d\alpha\, d\beta\, |\rho_1(\alpha)|\,|\rho_2(\beta)|\,[(\alpha+\beta)/(\alpha\beta)]^{\frac{1}{2}} \times \exp\{-(\alpha+\beta)\epsilon\}.$$

One can use Eq. (2.10) above to show that $K(ib)$ is finite, just as it was used to prove condition (d) above. Now we require that $k > 2K(ib)$, so that

$$|[1 - k^{-1}K(s+ib)]^{-1}| < 2.$$

Then one can use (12) to bound $|f|$ as follows:

$$|f(k,\Delta)| \leq \frac{\exp(-\Delta b)}{\Delta}\int_{-\infty}^{\infty} ds(|s| + b)\,|\mathfrak{V}(s+ib)|.$$

We can use relation (2.8) for $\mathfrak{V}$ along with condition (2.10) to show that the above integral is finite. Since it is independent of $\Delta$ and $k$ it is just a number. The following bound is thus valid for sufficiently large $\Delta$ and $k$:

$$|f(k,\Delta)| \leq C\exp\{-\Delta b\},$$

from this one can easily deduce the order relation (1.1).

We now wish to show that conditions (a) to (d) of Sec. 1 allow us to prove an order relation such as (1.1) with $b = r_0$; the previous considerations allow us to conclude this only for $b < r_0$. The approach is to distort the contour of integration in (2.7') to that given by $\operatorname{Im} r = b$. One can hope to do this only if $\mathfrak{V}$ and $K$ have "finite singularities" at $r = ir_0$, such as the branch point in the function $\exp\{-\alpha_0(r^2 + r_0^2)^{\frac{1}{2}}\}$. We can show that $\mathfrak{V}(ir_0)$ and $K(ir_0)$ are finite if the following integrals of $\rho_1$ and $\rho_2$ are finite:

$$I_{\frac{1}{2}} = \int_0^\infty d\alpha\,|\rho_{\frac{1}{2}}(\alpha)|, \tag{2.13}$$

$$G_{\frac{1}{2}} = \int_0^\infty d\alpha\alpha^{-\frac{1}{2}}\,|\rho_{\frac{1}{2}}(\alpha)|. \tag{2.14}$$

It is necessary to require that $k > 2K(ir_0)$, so that $|K(u + ir_0)| < \frac{1}{2}k$ and

$$[1 - k^{-1}K(u + ir_0)]^{-1} < 2.$$

We now distort the contour in Eq. (2.12) to obtain

$$|f(k,\Delta)| \leq \Delta^{-1}\exp\{-\Delta b\} \times \int_{-\infty}^{\infty} du(|u| + r_0)\,|\mathfrak{V}(u + ir_0)|.$$

We use Eq. (2.9) to obtain the bound

$$|\mathfrak{V}(u + ir_0)| \leq \iint_0^\infty d\alpha\, d\beta\,|\rho_1(\alpha)|\,|\rho_2(\beta)|\exp\{-(\alpha+\beta)u^2\}.$$

This leads to the following bound upon $f$:

$$|f(k,\Delta)| \leq \Delta^{-1}\exp\{-\Delta r_0\}\iint_0^\infty d\alpha\, d\beta\,|\rho_1(\alpha)|\,|\rho_2(\beta)| \times \{(\alpha+\beta)^{-1} + [\pi/\alpha+\beta]^{\frac{1}{2}}r_0\}.$$

Under the assumption of conditions (2.13) and (2.14) one can show that

$$|f(k,\Delta)| \leq C\exp\{-\Delta r_0\} \tag{2.15}$$

for sufficiently large $k$ and $\Delta$. Thus (1.1) is valid for $b = r_0$.

In view of the previous requirement (2.10), the integrals (2.13) and (2.14) can fail to converge only as $\alpha \to \infty$. One can show that if (2.10) and (2.13) are finite integrals, then (2.14) must of necessity be finite. It is thus necessary to show that conditions (a) to (d) of Sec. 1 are sufficient to guarantee a factorization for which the integrals are finite. We will establish this

by exhibiting such a factorization of $V$, namely

$$v_1(r) = (r^2 + r_0^2 + \lambda^2)V(r),$$
$$v_2(r) = (r^2 + r_0^2 + \lambda^2)^{-1},$$

where $\lambda > 0$. Under the constraint $\sigma(0) = 0$, the corresponding weight functions are

$$\rho_1(\alpha) = \sigma'(\alpha) + \lambda^2\sigma(\alpha),$$
$$\rho_2(\alpha) = \exp\{-\lambda^2\alpha\}.$$

Then $I_1 \leq \int_0^\infty d\alpha\,|\sigma'(\alpha)| + \lambda^2 \int_0^\infty d\alpha\,|\sigma(\alpha)|$, which is finite by condition (d). One can calculate $I_2$ explicitly to obtain the finite result $I_2 = \lambda^{-2}$.

We have thus shown that conditions (a) to (d) are sufficient to guarantee that the scattering amplitude obeys the following relation as $k \to \infty$ at fixed angles:

$$\mathscr{f}(k, \theta) = O[\exp(-2kr_0 \sin \tfrac{1}{2}\theta)].$$

## 3. CONVERGENCE TO THE FIRST BORN APPROXIMATION

In this section we will show that the scattering amplitude asymptotically approaches the first Born approximation at high energies for a fixed angle of scattering for potentials which satisfy conditions (a) to (h) in Sec. 1. That is, for this class of potentials we will show that

$$\lim_{\substack{k\to\infty \\ \theta \text{ fixed}}} \frac{\mathscr{f}(k, \theta)}{\mathscr{f}_1(k, \theta)} = 1. \tag{1.4}$$

The proof of this will be carried out in two steps. First we will show that there is a number $K'$ such that for sufficiently large $\Delta$,

$$|f_1(\Delta)| \geq K'\Delta^{-(p+3)} \exp(-\Delta r_0), \tag{3.1}$$

where $p$ is given in condition (g). Then we will show that

$$|\mathscr{f}(k, \Delta) - \mathscr{f}_1(\Delta)| = o(k^{-(p+3)} \exp\{-\Delta r_0\}) \tag{3.2}$$

as $k$ and $\Delta$ become large. One can then conclude (1.4) from (3.1) and (3.2). We will give the explicit proof here only with $p = 0$; however, the proof can easily be extended to $p > 0$.

We begin by proving (3.1). By our conventions

$$-\mathscr{f}_1(\Delta) = \frac{\pi^{\frac{1}{2}}}{4}\int_0^\infty d\alpha\alpha^{-\frac{5}{2}}\sigma(\alpha) \exp\{-(\alpha r_0^2 + \Delta^2/4\alpha)\}.$$

Now since $\sigma(\alpha) \geq 0$ and $\sigma(\alpha) \geq K\alpha^{-2}$ for $\alpha > \alpha_0$,

$$-\mathscr{f}_1(\Delta) \geq \frac{\pi^{\frac{1}{2}}K}{4}\int_\alpha^\infty d\alpha\alpha^{-\frac{7}{2}} \exp\{-(\alpha r_0^2 + \Delta^2/4\alpha)\}$$
$$\geq \frac{\pi^{\frac{1}{2}}K}{4}\int_0^\infty d\alpha\alpha^{-\frac{7}{2}} \exp\{-(\alpha r_0 + \Delta^2/4\alpha)\}$$
$$-\int_0^{\alpha_0} d\alpha\alpha^{-\frac{7}{2}} \exp\{-(\alpha r_0^2 + \Delta^2/4\alpha)\}.$$

One can show that

$$\int_0^{\alpha_0} d\alpha\alpha^{-\frac{7}{2}} \exp\{-(\alpha r_0^2 + \Delta^2/4\alpha)\}$$
$$= O(\exp\{-\Delta^2/4\alpha_0\} = o(\Delta^{-3} \exp\{-\Delta r_0\}) \quad \text{as } \Delta \to \infty$$

and the first integral may be calculated explicitly:

$$\int_0^\infty d\alpha\alpha^{-\frac{7}{2}} \exp\{-(\alpha r_0^2 + \Delta^2/4\alpha)\}$$
$$= 2[2r_0/\Delta]^{\frac{5}{2}}K_{\frac{5}{2}}(\Delta r_0) \geq \frac{8\pi^{\frac{1}{2}}r_0^2}{\Delta^3} \exp\{-\Delta r_0\}.$$

(We have used some well-known properties of modified Bessel functions in the last step.) Thus the inequality (3.1) is proved.

For the proof of inequality (3.2) the following lemma is useful.

*Lemma:* If $p(\alpha)$ is defined for $\alpha \geq 0$ such that the integrals

$$\int_0^\infty d\alpha\,|p(\alpha)| \quad \text{and} \quad \int_0^\infty d\alpha\alpha\,|p'(\alpha)|$$

converge, then $\alpha p(\alpha)$ is bounded for $0 \leq \alpha \leq \infty$ and $\alpha p(\alpha) \to 0$ as $\alpha \to \infty$. Further, if we define

$$U(q) = \int_0^\infty d\alpha\alpha^{-\frac{3}{2}}p(\alpha) \exp\{-(\alpha r_0^2 + q^2/4\alpha)\},$$

then $U(q) = o(q^{-2} \exp\{-qr_0\})$ as $q \to \infty$.

We will now outline the proof of Eq. (3.2).

We consider potentials defined by Eq. (1.2) with $\sigma(\alpha)$ being restricted such that the following integrals converge:

(i) $\int_0^\infty d\alpha\,|\sigma(\alpha)|$,

(ii) $\int_0^\infty d\alpha\,|\sigma'(\alpha)|$,

(iii) $\int_0^\infty d\alpha\alpha\,|\sigma'(\alpha)|$,

(iv) $\int_0^\infty d\alpha\alpha\,|\sigma''(\alpha)|$.

It is convenient to choose the factorization $v_1 \cdot v_2 = V$ such that $v_1$ and $v_2$ are given as follows:

$$\rho_1(\alpha) = \sigma'(\alpha) + \lambda^2\sigma(\alpha),$$
$$\rho_2(\alpha) = \exp\{-\lambda^2\alpha\}.$$

One can then use properties (i) to (iv) to show that the following integrals involving $\rho_1$ and $\rho_2$ converge:

(v) $\int_0^\infty d\alpha\alpha\,|\rho_i'(\alpha)|$,

(vi) $\int_0^\infty d\alpha\,|\rho_i(\alpha)|$,

(vii) $\int_0^\infty d\alpha\alpha^{-\frac{1}{2}}\,|\rho_i(\alpha)|$.

We rewrite Eq. (2.8) in the form

$$\mathcal{U}(x) = \int_0^\infty d\alpha \bar{\sigma}(\alpha) \exp\{-\alpha(x^2 + r_0^2)\}, \quad (3.3)$$

where

$$\bar{\sigma}(\alpha) = \int_0^\alpha d\beta \, |\rho_1(\beta)| \, |\rho_2(\alpha - \beta)|.$$

Now (v) and (vi) are used to prove that the following integrals involving $\bar{\sigma}$ are convergent:

(viii) $\int_0^\infty d\alpha \bar{\sigma}(\alpha)$,

(ix) $\int_0^\infty d\alpha\alpha \, |\bar{\sigma}'(\alpha)|$.

Let us rewrite Eq. (2.9) in the form

$$K(x) = \int_0^\infty d\alpha \lambda(\alpha) \exp\{-\alpha(x^2 + r_0^2)\}, \quad (3.4)$$

where

$$\lambda(\alpha) = \int_0^\alpha d\beta [\pi\alpha/\beta(\alpha - \beta)]^{\frac{1}{2}} \, |\rho_1(\alpha - \beta)| \, |\rho_2(\beta)|.$$

We use (v) and (vii) to prove that the following integrals converge:

(x) $\int_0^\infty d\alpha \, |\lambda(\alpha)|$,

(xi) $\int_0^\infty d\alpha\alpha \, |\lambda'(\alpha)|$.

Let us define a weight function $\tau(\gamma, k)$ through

$$K(x)\left[1 - \frac{K(x)}{k}\right]^{-1} = \int_0^\infty d\gamma \tau(\gamma, k) \exp\{-\alpha(r^2 + r_0^2)\}. \quad (3.5)$$

It is easily shown that $\tau$ is related to $\lambda$ by the integral equation

$$\tau(\gamma, k) = \lambda(\gamma) + \frac{1}{k} \int_0^\gamma d\beta \lambda(\gamma - \beta) \tau(\beta, k).$$

A number $k_0$ is then defined as the maximum of these three finite numbers:

(a) The l.u.b. of $\alpha\lambda(\alpha)$ for $0 \leq \alpha \leq \infty$,[8]

(b) $\int_0^\infty d\alpha\alpha \, |\lambda'(\alpha)|$,

(c) $\int_0^\infty d\alpha \lambda(\alpha)$.

Then if we define a function $\tau(\gamma) \equiv \tau(\gamma, 2k_0)$, we can show that

$$0 \leq \tau(\gamma, k) \leq \tau(\gamma) \quad \text{for} \quad k > 2k_0. \quad (3.6)$$

It can be shown through the use of (x) and (xi) that the following integrals involving $\tau$ converge:

(xii) $\int_0^\infty d\gamma \, |\tau(\gamma)|$,

(xiii) $\int_0^\infty d\gamma\gamma \, |\tau'(\gamma)|$.

It is convenient to write the inequality (2.7) as follows:

$$|f(k, \Delta) - f_1(\Delta)| \leq (4\pi k)^{-1} \int d\mathbf{x} \mathcal{U}(x) \frac{K(x)}{1 - (K(x)/k)} \exp\{-i\boldsymbol{\Delta} \cdot \mathbf{x}\}. \quad (2.7)$$

We insert the representations (3.3) and (3.5) into this equation and explicitly perform the integration over $\mathbf{x}$ to obtain

$$|f(k, \Delta) - f_1(\Delta)| \leq \frac{\pi^{\frac{1}{2}}}{4k} \int_0^\infty d\beta \bar{\sigma}(\beta) \int_0^\infty d\gamma \tau(\gamma, k)(\beta + \gamma)^{-\frac{3}{2}} \times \exp\{-(\beta + \gamma)r_0^2 - \Delta^2/4(\beta + \gamma)\}. \quad (3.7)$$

Let us restrict $k$ to $k > 2k_0$ and use the inequality (3.6) to rewrite (3.7) in the form

$$|f(k, \Delta) - f_1(\Delta)| \leq \frac{\pi^{\frac{1}{2}}}{4k} \int_0^\infty d\alpha\alpha^{-\frac{3}{2}} \mu(\alpha) \exp\{-(\alpha r_0^2 + \Delta^2/4\alpha)\}, \quad (3.8)$$

where

$$\mu(\alpha) = \int_0^\alpha d\beta \bar{\sigma}(\beta) \tau(\alpha - \beta).$$

Now (viii), (ix), (xii), and (xiii) are used to prove the convergence of the following integrals involving $\mu$:

(xiv) $\int_0^\infty d\alpha \mu(\alpha)$,

(xv) $\int_0^\infty d\alpha\alpha \, |\mu'(\alpha)|$.

We can use the convergence of integrals (xiv) and (xv) along with the inequality (3.8) in the lemma to prove that

$$|f(k, \Delta) - f_1(\Delta)| = O((k\Delta^2)^{-1} \exp\{-\Delta r_0\})$$

as $\Delta \to \infty$, so that at fixed angle the relation (3.2) is valid. Thus the limit (1.4) is proved. One can also extend the proof to the case $p > 0$ with relative ease.

One can no doubt extend the validity of relation (1.4) to a wider class of potentials. We will show, however, that relation (1.4) cannot be so general as its fixed-$\Delta$ counterpart by giving an example of a potential whose second Born term is asymptotically larger than its first Born term. (The higher terms cannot cancel out this behavior for arbitrarily weak potentials.)

[8] This is guaranteed to be finite by the lemma of this section.

The example here is the Gaussian potential, $V(r) = \lambda \exp\{-ar^2\}$, for which

$$f_1(\Delta) = -(\pi^{\frac{1}{2}}\lambda/4a^{\frac{3}{2}}) \exp\{-(k^2/a)\sin^2\tfrac{1}{2}\theta\}$$

and

$$|f_2(k,\theta)| \geq \operatorname{Im} f_2(k,\theta)$$
$$= (\lambda^2\pi^3/a^2) \exp\{-k^2/a\} \sinh\{(k^2/a)\cos\tfrac{1}{2}\theta\}/k\cos\tfrac{1}{2}\theta$$

so that

$$\lim_{\substack{k\to\infty \\ 0<\theta\leq\pi}} \frac{f_1(k,\theta)}{f_2(k,\theta)} = 0.$$

There is another type of potential for which the first Born approximation oscillates, such as

$$V(r) = \int_0^\infty d\alpha \sin(m^2\alpha) \exp\{-\alpha(r^2 + r_0^2)\}$$
$$= m^2[m^4 + (r^2 + r_0^2)^2]^{-1}.$$

The corresponding first Born approximation is

$$f_1(\Delta) = -(\pi/2\Delta) \exp\{-\lambda\Delta\} \sin(\mu\Delta),$$

where $[r_0^2 - im^2]^{\frac{1}{2}} = \lambda + i\mu$ with the convention $\lambda > 0$. For such a potential the limit considered in Eq. (2.4) cannot be defined unambiguously.

## 4. SPECULATIONS

In the previous sections we have considered potentials of the form

$$V(r) = \int_0^\infty d\alpha\sigma(\alpha) \exp\{-\alpha(r^2 + r_0^2)\}. \quad (1.2)$$

We have placed conditions upon $\sigma$ sufficient to guarantee that $V(r)$ is analytic in $r = x + iy$ in the region $y^2 - x^2 < r_0^2$, and we have shown that $|f(k,\theta)| \leq \exp\{-2kr_0 \sin\tfrac{1}{2}\theta\}$ as $k \to \infty$ at fixed angles.

There are cases in which this bound is economical, as is seen in Sec. 3. On the other hand, in the case considered at the end of Sec. 3,

$$V(r) = m^2[m^4 + (r^2 + r_0^2)^2]^{-1},$$

we saw that $f_1(\Delta) = O(\exp\{-\lambda\Delta\})$ as $\Delta \to \infty$, where $\lambda = [\frac{1}{2}(r_0^2 + (r_0^4 + m^4)^{\frac{1}{2}}]^{\frac{1}{2}}$. This potential is analytic in the strip $|\operatorname{Im} r| < \lambda$. We are thus led to speculate about potentials which are analytic in a strip.

Let us consider the representation of potentials which are even in $r$ and analytic in the strip $|\operatorname{Im} r| < b$, namely

$$V(r) = \int_0^\infty d\alpha\tau(\alpha)[(r^2 + (b + i\alpha)^2)(r^2 + (b - i\alpha)^2)]^{-1}. \quad (4.1)$$

The formal expression for the first Born approximation is

$$f_1(\Delta) = -\frac{\pi}{4b} \exp\{-\Delta b\} \int_0^\infty d\alpha\tau(\alpha) \frac{\sin\Delta\alpha}{\Delta\alpha}. \quad (4.2)$$

We see that $f_1(\Delta) = O(\exp\{-\Delta b\})$ as $\Delta \to \infty$ if either of the following integrals of $\tau$ converges:

$$\int_0^\infty d\alpha\,|\tau(\alpha)|, \quad (4.3)$$

$$\int_0^\infty d\alpha\alpha^{-1}\,|\tau(\alpha)|. \quad (4.4)$$

One can show that if conditions (4.3) and (4.4) are satisfied the second Born approximation satisfies the relation

$$f_2(k,\theta) = O(\exp\{-\Delta b\})$$

as $k \to \infty$ for a fixed angle $\theta$.[9]

One is thus led to conjecture that the scattering amplitude itself exhibits Orear behavior for potentials of the form (4.1) with reasonable conditions such as (4.3) and (4.4) placed upon the weight function $\tau$. We have not yet been able to prove this conjecture.

## ACKNOWLEDGMENTS

The author wishes to thank his thesis adviser, Professor S. B. Treiman, for suggesting this topic and for providing continual guidance in its development. Also the author wishes to thank Professor G. Tiktopoulos for several stimulating discussions.

[9] The rather lengthy proof of this and certain other results are not included here. All such proofs are shown in detail in P. Johnson, thesis, Princeton University (unpublished).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 5 MAY 1968

# An Integral Equation for the Scattering Operator

M. WILNER*
*Physics Department, Lowell Technological Institute, Lowell, Massachusetts*

(Received 8 October 1967)

An integral equation, originally derived for a perturbation expansion of the $n$-point scattering function (Pugh's equation), whose (finite) solution is the renormalized perturbation result, is derived here as an exact "strong" equation for the $n$th operator derivative of the scattering *operator*, whose vacuum expectation value is the $n$-point function.

## 1. INTRODUCTION

An integral equation, Eq. (3) below, for a perturbation expansion of the $n$-point scattering function was derived by Pugh from the postulates of asymptotic quantum field theory, and the assumption that off the mass shell the scattering function should be represented by the $\varphi$ products, rather than by the time-ordered products of interpolating field operators.[1] This assumption was later shown to be a consequence of a generalization of Bogoliubov's causality condition.[2] Pugh's equation is remarkable, because with an appropriate choice of boundary conditions, its solution is identical with the result of renormalized perturbation theory, and no divergent expressions are encountered in obtaining it.[3]

After the development of the differential calculus of quantized free-field operators (operator derivatives),[4] it was shown[5] that the exact two-point operator, the operator whose vacuum expectation value is the exact two-point function, satisfies Eq. (3), with $n = 2$. The case $n > 2$ was not derived, perhaps because Eq. (3) with $n = 2$ does not imply Eq. (3) with $n > 2$, and so a derivation must start further back in the theory, perhaps because the notation was not concise enough to write compact equations with arbitrary $n$, or because it did not appear to be necessary, as the $n = 2$ operator equation was enough to yield the renormalized results.

Now that there is interest in nonperturbative solutions of Pugh's equation,[6] it is desirable to have an exact derivation. Furthermore, in his discussion of boundary conditions for Pugh's equation, Wray[3] was obliged to inquire rather deeply into the diagram structure of the $n$-point operator. He did this by means of an equation similar to the operator equation to be derived here, but valid only as an operator distribution over a left space of test functions. The equation we derive here is not so restricted. We show that Pugh's equation is satisfied, for any $n$, by the exact $n$-point operator, as a "strong" operator equation (in the sense to be explained) and therefore by the exact $n$-point scattering function.

In Sec. 2 we quote some of the results of asymptotic quantum field theory (AQFT) which are necessary for our derivation. In Sec. 3 we derive Pugh's equation, and in Sec. 4 we verify a necessary condition for the existence of solutions. We confine our discussion to neutral self-interacting scalar particles, although there are no obstacles to extending it to charged particles and up to spin one.

## 2. RESULTS OF AQFT

In terms of the (assumed) complete set of normal-ordered products of asymptotic incoming free-field operators $a(x)$ of physical mass $m$, satisfying

$$K_x a(x) \equiv (\Box - m^2)a(x) = 0, \tag{1}$$

the scattering operator has the representation

$$S = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int \omega(x_1 \cdots x_n) : a(x_1) \cdots a(x_n) : (d^4x)^n. \tag{2}$$

If $\omega(x_1 \cdots x_n)$ has a perturbation expansion

$$\omega(x_1 \cdots x_n) = \sum_{r=0}^{\infty} g^r \omega^{(r)}(x_1 \cdots x_n),$$

then $\omega^{(r)}$ satisfies Pugh's equation[1]

$$\omega^{(r)}(x_1 \cdots x_n) - \int B(x_1 \cdots x_n; y_1 \cdots y_n) \times \omega^{(r)}(y_1 \cdots y_n)(d^4y)^n = \lambda^{(r)}(x_1 \cdots x_n), \tag{3}$$

---

* This work was begun while the author was at Syracuse University, and part of it was done while he was attending the Summer Theoretical Physics Institute at the University of Colorado, whose hospitality is gratefully acknowledged.

[1] R. E. Pugh, Ann. Phys. (N.Y.) **23**, 335 (1963).

[2] T. W. Chen, F. Rohrlich, and M. Wilner, J. Math. Phys. **7**, 1365 (1966).

[3] See Ref. 1, and also J. G. Wray, Syracuse University Res. Rept. SU-67-01 and SU-67-02, where the boundary conditions are discussed in greater detail, and an error in Ref. 1 is corrected.

[4] F. Rohrlich, J. Math. Phys. **5**, 324 (1964); F. Rohrlich and M. Wilner, J. Math. Phys. **7**, 482 (1966).

[5] R. E. Pugh, J. Math. Phys. **6**, 740 (1965).

[6] An exact formal solution of an operator differential equation for the current operator, equivalent to Eq. (15) with $m = 2$, giving the current operator as an "operator integral" of the solution of the homogeneous equation, has been obtained, and will shortly be submitted for publication.

abbreviated as

$$(1 - B)\omega^{(r)}(x_1 \cdots x_n) = \lambda^{(r)}(x_1 \cdots x_n), \quad (3')$$

where

$$B = \sum_{\lambda=1}^{n} B_\lambda$$

and

$$\begin{aligned} B_\lambda(x_1 \cdots x_n; y_1 \cdots y_n) &= (-1)^n K_{x_1} \cdots K_{x_n} \Big[ \prod_{j=1}^{n}{}' \theta(x_j^0 - x_\lambda^0)\Delta(x_j - y_j) \Big] \\ &\quad \times \Delta_R(x_\lambda - y_\lambda) \\ &= (-1)^n K^n \Big( \prod_{j=1}^{n}{}' \theta_{j\lambda} \Big) \Delta^1 \cdots \Delta_R^\lambda \cdots \Delta^n, \end{aligned} \quad (4)$$

in more concise notation. The symbols used mean

$$K^n \equiv K_{x_1} \cdots K_{x_n}, \quad \theta_{j\lambda} \equiv \theta(x_j^0 - x_\lambda^0),$$
$$\Delta^i \equiv \Delta(x_i - y_i), \quad \Delta_R^\lambda \equiv \Delta_R(x_\lambda - y_\lambda),$$

and the prime on the product sign means that $j \neq \lambda$ in the product. The function $\lambda^{(r)}$ satisfies the equation

$$B\lambda^{(r)}(x_1 \cdots x_n) = 0, \quad (5)$$

which is independent of Eq. (3). $\lambda^{(r)}$ is determined by the functions $\omega^{(s)}$ with $s < r$, and is therefore known when it appears in Eq. (3). Since $B$ is idempotent, i.e.,

$$\int B_\lambda(x_1 \cdots x_n; y_1 \cdots y_n) B_\mu(y_1 \cdots y_n; z_1 \cdots z_n)(d^4y)^n = \delta_{\lambda\mu} B_\lambda(x_1 \cdots x_n; z_1 \cdots z_n), \quad (6)$$

so that $B^2 = B$, a necessary condition that a solution of Eq. (3) exist is that Eq. (5) be true, and independent of Eq. (3).

If the exact $\omega(x_1 \cdots x_n)$ is determined uniquely off as well as on the mass shell (by dynamical assumptions), and is a symmetric function of its arguments, then the $m$th operator derivative of $S$ is given by[4]

$$i^m \frac{\delta^m S}{\delta a(x_1) \cdots \delta a(x_m)} = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \times \int \omega(x_1 \cdots x_m y_1 \cdots y_n) : a(y_1) \cdots a(y_n) : (d^4y)^n, \quad (7)$$

and this is the operator whose vacuum expectation value is $\omega(x_1 \cdots x_m)$. We also have, from the definition of the operator derivative,

$$[a(x), S] = -i \int \Delta(x - y)\delta S/\delta a(y) d^4y. \quad (8)$$

If one defines a current operator

$$j(x) = iS^*\delta S/\delta a(x), \quad (9)$$

then it is possible to construct an interpolating field operator

$$A(x) \equiv a(x) + \alpha(x), \quad (10)$$

where

$$\alpha(x) = -\int \Delta_R(x - y) j(y) d^4y, \quad (11)$$

such that $A(x)$ satisfies the field equation

$$K_x A(x) = j(x), \quad (12)$$

and has the correct LSZ asymptotic limits[7]

$$\int \Delta(x - y) \frac{\overleftrightarrow{\partial}}{\partial y^0} A(y) d^3y \to \begin{cases} a(x), & y^0 \to -\infty \\ S^* a(x) S, & y^0 \to +\infty. \end{cases} \quad (13)$$

So far we have been describing an operator functional called the operator derivative. If we wish to define an operation, called "differentiating with respect to the operator $a(x)$," and denoted variously by

$$\frac{\delta}{\delta a(x_i)} \equiv \delta_{x_i} \equiv \delta_i, \quad (14)$$

such that this operation commutes with ordinary differentiation and integration, then we must distinguish between operator equations which are "strong" and those which are "weak" with respect to this operation.[5] Strong equations are those which yield valid operator equations after operator differentiation any number of times. Weak equations are those which do not. It is most convenient to take the field equations (1) and (12) as weak,[8,9] all other operator equations in this paper are strong, and therefore no special symbol will be used to indicate this fact. The strong equation we will derive is

$$(1 - B) i^m S^* \delta_1 \cdots \delta_m S = \Lambda(x_1 \cdots x_m), \quad (15)$$

where $\Lambda$ is an operator which satisfies the strong equation

$$B\Lambda(x_1 \cdots x_m) = 0. \quad (16)$$

Eq. (16) is to be proved without using Eq. (15).

### 3. DERIVATION OF EQ. (15)

We start with a representation of the scattering operator which follows from the postulates of AQFT and the strong Bogoliubov causality condition

$$\delta_y(S^* \delta_x S) = 0, \quad (17)$$

[7] H. Lehmann, K. Symanzick, and W. Zimmermann, Nuovo Cimento 1, 205 (1955); 6, 860 (1956).
[8] T. W. Chen, Ann. Phys. (N.Y.) 42, 476 (1967).
[9] J. G. Wray, Ph.D. thesis, Syracuse University, 1966 (unpublished). Chen, who first recognized the desirability of having a weak free-field equation, chose to take Eq. (10) rather than Eq. (12) as weak. We adopt Wray's choice, however.

outside the forward light cone of $(x - y)$. It is[10]

$$i^m S^* \delta_1 \cdots \delta_m S = K_1 \cdots K_m \sum_{l=0}^{m} (-1)^l T_+(A^{m-l}) T_-(a^l). \quad (18)$$

The compact notation on the right-hand side of Eq. (18) is meant to indicate that each term in the sum over $l$ is itself a sum over all partitions of the variables $x_1 \cdots x_m$, $l$ of them assigned to operators $a$, the remaining $m - l$ to operators $A$. $T_+(T_-)$ is the positively (negatively) time-ordered product. Thus

$$T_\pm(A^0) = 1, \quad T_\pm(A^1) = A, \quad T_\pm(A^m) = T_\pm(A_1 \cdots A_m),$$

$$T_+(A^{m-1}) T_-(a) = \sum_{k=1}^{m} T_+(A_1 \cdots \Lambda_k \cdots A_m) a_k$$

$$T_+(A^{m-2}) T_-(a^2) = \sum_{k_2 > k_1 \geq 1}^{m} T_+(A_1 \cdots \Lambda_{k_1 k_2} \cdots A_m) T_-(a_{k_1} a_{k_2}), \quad (19)$$

etc. The symbol $\Lambda_{k_1 k_2 \ldots}$ indicates that the factors $A_{k_1}, A_{k_2}, \cdots$ are missing from the product $A_1 \cdots A_m$. $A_i = A(x_i)$. We may set $A = a + \alpha$ in Eq. (18), and rearrange the terms, obtaining

$$T_+(A^{m-l}) = \sum_{\mu=0}^{m-l} T_+(\alpha^\mu a^{m-l-\mu}). \quad (20)$$

As in Eq. (18), each term in the sum over $\mu$ is itself a sum over all partitions of the $m - l$ variables $x_i$, $\mu$ of them assigned to operators $\alpha$, the remaining $m - l - \mu$ to operators $a$. For example, with $\mu = 1$, one of the terms in the sum over partitions is

$$T_+(\alpha a^{m-l-1}) = \sum_{k=1}^{m-l} T_+(a_{i_1} a_{i_2} \cdots \alpha_{i_k} \cdots a_{i_{m-l}}).$$

If we change the summation index in Eq. (20) from $\mu$ to $\mu' = m - \mu$, substitute in Eq. (18), and exchange the order of summation over $l$ and $\mu'$, we have

$$i^m S^* \delta_1 \cdots \delta_m S = K^m \sum_{\mu=0}^{m} \sum_{l=0}^{\mu} (-1)^l T_+(\alpha^{m-\mu} a^{\mu-l}) T_-(a^l). \quad (21)$$

We are now able to apply a remarkable ordering theorem, relating the time-ordered product to the multiple-retarded commutator,[11] which in the present case and in the present notation is

$$T_+(\alpha^{m-\mu} a^{\mu-l}) = \sum_{k=0}^{\mu-l} [a, T_+(\alpha^{m-\mu})]_R^k T_+(a^{\mu-l-k}). \quad (22)$$

Each term in the sum over $k$ is as usual a sum over all partitions of the $(\mu - l)$ operators $a(x_i)$, $k$ of them appearing in the $k$-fold retarded commutator, defined by

$$[a, T_+]_R^k = [a, [a, T_+]_R^{k-1}]_R, \quad [a, T_+]_R^1 = [a, T_+]_R, \quad [a, T_+]_R^0 = T_+, \quad (23)$$

and the remainder in the $T_+$ product. There is no sum over permutations of the operators $a(x)$ in the multiple retarded commutator, since it is already symmetric in them by virtue of the Jacobi identity. Substitution of Eq. (22) in Eq. (21) and exchanging the order of summation over $k$ and $l$ yields

$$i^m S^* \delta_1 \cdots \delta_m S = K^m \sum_{\mu=0}^{m} \sum_{k=0}^{\mu} [a, T_+(\alpha^{m-\mu})]_R^k \times \sum_{l=0}^{\mu-k} (-1)^l T_+(a^{\mu-k-l}) T_-(a^l) = K^m \sum_{\mu=0}^{m-1} [a, T_+(\alpha^{m-\mu})]_R^\mu, \quad (24)$$

since, as will be shown in the Appendix, the sum over $l$ in the second-last equation vanishes, except when $\mu - k = 0$, in which case it is the unit operator. Note that there is no term $\mu = m$, since for such a term

$$[a, T_+(\alpha^0)]_R^m = [a, 1]_R^m = 0.$$

### Extracting B

The integral kernel $B$ is contained in the term $\mu = m - 1$ of Eq. (24). We exhibit the dependence of this term on the variables $x_\lambda, x_1, \cdots x_m$ one at a time, making use of Eq. (8). Thus

$$[a, \alpha]_R^{m-1} = \sum_{\lambda=1}^{m} [a, \alpha(x_\lambda)]_R^{m-1} = \sum_{\lambda=1}^{m} [a, [a(x_1), \alpha(x_\lambda)]_R]_R^{m-2} = \sum_{\lambda=1}^{m} \theta_{1\lambda} \int \Delta(x_1 - y_1) [a, -i\delta_{y_1} \alpha(x_\lambda)]_R^{m-2} d^4 y_1. \quad (25)$$

We repeat this procedure for $x_2$, etc., and in the end use Eq. (11), and obtain

$$K^m [a, \alpha]_R^{m-1} = (-1)^m K^m (i)^{m-1} \sum_{\lambda=1}^{m} \prod_{j=1}^{m}{}' \theta_{j\lambda} \int \Delta^1 \cdots \Delta_R^\lambda \cdots \Delta^m \times \delta_{y_1} \cdots \Lambda_{y_\lambda} \cdots \delta_{y_m} j(y_\lambda) (d^4 y)^m = i^{m-1} \sum_{\lambda=1}^{m} B_\lambda \delta_1 \cdots \Lambda_\lambda \cdots \delta_m j_\lambda. \quad (26)$$

The operator derivatives commute with each other, and therefore may be written in any order, so we may abbreviate

$$\delta_1 \cdots \delta_m = \delta^m, \quad \delta_1 \cdots \Lambda_\lambda \cdots \delta_m = \delta^{m-1}.$$

It will be clear from the context which variable is to

[10] See Ref. 9. This expression is weakly equal to the $\varphi$ product of Ref. 1.

[11] F. Rohrlich and J. G. Wray, J. Math. Phys. 7, 1697 (1966).

be omitted in $\delta^{m-1}$. From the definition of $j(x)$, Eq. (9), we have

$$-i\delta^{m-1}j_\lambda = \delta^{m-1}(S^*\delta_\lambda S) = S^*\delta^m S + \sum_{k=1}^{m-1}(\delta^k S^*)(\delta^{m-1-k}\delta_\lambda S). \quad (27)$$

Each term in the summation over $k$ is itself a sum of all partitions of the operator derivatives $\delta_1, \cdots, \Lambda_\lambda, \cdots, \delta_m$, $k$ of them differentiating $S^*$, the remaining $m-1-k$ differentiating $\delta_\lambda S$. This completes the extraction of $B$, as we now have

$$K^m[a,\alpha]_R^{m-1} = Bi^m S^*\delta^m S + i^m\sum_{\lambda=1}^{m}B_\lambda\sum_{k=1}^{m-1}(\delta^k S^*)(\delta^{m-1-k}\delta_\lambda S), \quad (28)$$

so that we can rewrite Eq. (24) in the promised form

$$(1-B)i^m S^*\delta^m S = \Lambda_a + \Lambda_b, \quad (29)$$

where

$$\Lambda_a = K^m\sum_{\mu=0}^{m-2}[a, T_+(\alpha^{m-\mu})]_R^\mu, \quad (30)$$

and

$$\Lambda_b = i^m\sum_{\lambda=1}^{m}B_\lambda\sum_{k=1}^{m-1}(\delta^k S^*)(\delta^{m-1-k}\delta_\lambda S). \quad (31)$$

## 4. PROOF THAT $B(\Lambda_a + \Lambda_b) = 0$

Since $B$ is idempotent (Eq. (6)), $B\Lambda_b = \Lambda_b$. Therefore, in order to show that $B(\Lambda_a + \Lambda_b) = 0$, we show that $B\Lambda_a = -\Lambda_b$. We will have to integrate expressions of the form

$$\int\Delta^1\cdots\Delta_R^\lambda\cdots\Delta^m K^m[a, T_+(\alpha^{m-\mu})]_R^\mu(d^4y)^m.$$

Although the sum over $\lambda$ and $\mu$ of all such expressions is symmetric in the variables not integrated, $x_1\cdots x_m$, and therefore independent of the order in which the integrations are performed, any one such expression will depend on the order of integration. However, if we perform the integration in the fixed order $y_1\cdots y_m$, for each such expression, we will not run the risk of losing or duplicating terms when we sum.

We start with $\mu = 0$. Then

$$\int\Delta^1\cdots\Delta_R^\lambda\cdots\Delta^m K_y^m T_+(\alpha_y^m)(d^4y)^m = \int\Delta^1\cdots\Delta_R^\lambda\cdots\Delta^m j_1\cdots\Lambda_\lambda\cdots j_m j_\lambda(d^4y)^m, \quad (32)$$

where

$$K_y^m T_+(\alpha_y^m) \equiv K_{y_1}\cdots K_{y_m}T_+[\alpha(y_1)\cdots\alpha(y_m)],$$

and

$$j_i \equiv j(y_i).$$

*Proof:* Integrate $y_1$ by parts. The volume term vanishes, and the surface terms are

$$\int\Delta^1 K_{y_1}T_+(\alpha_y^m)d^4y_1 = (\lim_{y_1^0\to-\infty} - \lim_{y_1^0\to-\infty})\times\int\Delta^1\frac{\overleftrightarrow{\partial}}{\partial y_1^0}T_+(\alpha_y^m)d^3y_1$$
$$= \lim_{y_1^0\to-\infty}\int\Delta^1\frac{\overleftrightarrow{\partial}}{\partial y_1^0}T_+(\alpha_y^{m-1})\alpha(y_1)d^3y_1 - \lim_{y_1^0\to+\infty}\int\Delta^1\frac{\overleftrightarrow{\partial}}{\partial y_1^0}\alpha(y_1)T_+(\alpha_y^{m-1})d^3y_1. \quad (33)$$

Since $\alpha = A - a$, the asymptotic condition, Eq. (13), together with (8), (9), and (11) give

$$\int\Delta(x-y)\frac{\overleftrightarrow{\partial}}{\partial y^0}\alpha(y)d^3y \to \begin{cases}0, & y^0\to-\infty\\ -\int\Delta(x-z)j(z)d^4z, & y^0\to+\infty,\end{cases} \quad (34)$$

and so

$$\int\Delta^1 K_{y_1}T_+(\alpha_y^m)d^4y_1 = \int\Delta^1 j(y_1)T_+(\alpha(y_2)\cdots\alpha(y_m))d^4y_1. \quad (35)$$

We continue, integrating in order $y_2\cdots y_{\lambda-1}$, and obtain for the left-hand side of Eq. (32) the expression

$$\int\Delta^1\cdots\Delta_R^\lambda\cdots\Delta^m j_1\cdots j_{\lambda-1}K_{y_\lambda}\cdots K_{ym}\times T_+(\alpha(y_\lambda)\cdots\alpha(y_m))(d^4y)^m. \quad (36)$$

The integration by parts over $y_\lambda$ gives only a volume term,

$$\int\Delta_R^\lambda K_{y_\lambda}T_+(\alpha(y_\lambda)\cdots\alpha(y_m))d^4y_\lambda = -T_+(\alpha(x_\lambda)\alpha(y_{\lambda+1})\cdots\alpha(y_m)), \quad (37)$$

because both surface terms vanish, at $y_\lambda^0\to-\infty$ because of Eq. (34), and at $y_\lambda^0\to+\infty$ because of $\Delta_R^\lambda$. Once past $y_\lambda$, we may integrate the remaining $y$ variables, as in Eq. (35). At the end we substitute Eq. (11), and obtain (32).

For the general case, we show that

$$\int\Delta^1\cdots\Delta_R^\lambda\cdots\Delta^m K_y^m[a_y, T_+(\alpha_y^{m-\mu})]_R^\mu(d^4y)^m = \int\Delta^1\cdots\Delta_R^\lambda\cdots\Delta^m \times\sum_{k_\mu>\cdots>k_1\geq1}{}'j_1\cdots i\delta_{k_1},\cdots,\cdots i\delta_{k_1}\cdots j_m j_\lambda(d^4y)^m. \quad (38)$$

The operator product on the right-hand side of (38) is the same product of currents as in Eq. (32), but $j_{ki} \equiv j(y_{ki})$ is replaced by the operator derivative $i\delta_{ki} \equiv i\delta/\delta a(y_{ki})$, which acts on everything to its right, $i = 1, 2, \cdots, \mu$, $k_i \neq \lambda$, and the expression is summed over all such possible substitutions. The proof is by induction. We suppose that Eq. (38) holds for any $n \leq m$, and any $\nu < \mu$, such that $\nu < n$. (If $\nu \geq n$ there would be no such expression.) $\mu = 0$ is a special case of Eq. (38). First we extract $y_1$ from the left-hand side of (38). The method is similar to that of Eq. (33). We have

$$[a, T_+(\alpha^{m-\mu})]^\mu_R = [a(y_1), [a, T_+(\alpha^{m-\mu})]^{\mu-1}_R]_R + [a, T_+(\alpha(y_1)\alpha^{m-1-\mu})]^\mu_R. \quad (39)$$

When we integrate $y_1$ by parts the volume term vanishes, and the surface integral of the term with $a(y_1)$ vanishes when $y_1^0 \to -\infty$, because of the retarded commutator in which it appears. When $y_1^0 \to +\infty$, the retarded commutator with $a(y_1)$ becomes an ordinary commutator, so

$$\int \Delta(x-y) K_\nu [a(y), Q]_R d^4y = -\lim_{y_1^0 \to +\infty} \int \Delta(x-y) \frac{\overleftrightarrow{\partial}}{\partial y^0} [a(y), Q] d^3y = \int \Delta(x-z) i\delta_z Q d^4z, \quad (40)$$

according to Eq. (8). The surface integral of the term with $\alpha(y_1)$ vanishes when $y_1^0 \to -\infty$, because

$$T_+(\alpha(y_1)\alpha^{m-1-\mu}) \to T_+(\alpha^{m-1-\mu})\alpha(y_1),$$
$$[a, \alpha(y_1)]_R \to [a, \alpha(y_1)],$$

so that

$$[a, T_+(\alpha(y_1)\alpha^{m-1-\mu})]^\mu_R \to F(a)\alpha(y_1) + G(a)[a, \alpha(y_1)],$$

where $F$ and $G$ are operator functionals of $a$, and Eq. (34) applies. When $y_1^0 \to +\infty$,

$$[a, T_+(\alpha(y_1)\alpha^{m-1-\mu})]^\mu_R \to [a, \alpha(y_1)T_+(\alpha^{m-1-\mu})]^\mu_R \to \alpha(y_1)H(a) + [a, \alpha(y_1)]_R K(a),$$

and the retarded commutator with $\alpha(y_1)$ vanishes, so that

$$\int \Delta^1 K_{\nu_1} [a, T_+(\alpha(y_1)\alpha^{m-1-\mu})]^\mu_R d^4y_1 = -\lim_{y_1^0 \to +\infty} \int \Delta^1 \frac{\overleftrightarrow{\partial}}{\partial y_1^0} \alpha(y_1) [a, T_+(\alpha^{m-1-\mu})]^\mu_R d^3y_1 = \int \Delta^1 j(y_1) [a, T_+(\alpha^{m-1-\mu})]^\mu_R d^4y_1. \quad (41)$$

Putting together (39), (40), and (41), we have

$$\int \Delta^1 K_{\nu_1} [a, T_+(\alpha^{m-\mu})]^\mu_R d^4y_1 = \int \Delta^1 (i\delta_{\nu_1} [a, T_+(\alpha^{m-\mu})]^{\mu-1}_R + j(y_1)[a, T_+(\alpha^{m-1-\mu})]^\mu_R) d^4y_1. \quad (42)$$

Since $y_1$ has been extracted from the multiple commutator, it is not included in the sum over partitions of the remaining variables indicated by the notation. Therefore, either $\mu < m$, or Eq. (42) reduces to $0 = 0$.[12] In the former case we may set $\mu - 1 = \nu$, and $m - 1 = n$. Then $\nu < n$, which satisfies the induction hypothesis following Eq. (38), so that

$$\int \Delta^2 \cdots \Delta^\lambda_R \cdots \Delta^m K_{\nu_2} \cdots K_{\nu_m} [a, T_+(\alpha^{m-\mu})]^{\mu-1}_R (d^4y)^{m-1} = \int \Delta^2 \cdots \Delta^\lambda_R \cdots \Delta^m \times \sum'_{k_\mu > \cdots > k_2 \geq 2}{}^{m} j_2 \cdots i\delta_{k_\lambda} \cdots, \cdots i\delta_{k_\mu} \cdots j_m j_\lambda (d^4y)^{m-1}. \quad (43)$$

In the term $[a, T_+(\alpha^{m-1-\mu})]^\mu_R$, we extract $y_2$ in the same manner that we extracted $y_1$ from Eq. (39). We obtain an equation like Eq. (42), but with $x_1$, $y_1$, and $m$ replaced by $x_2$, $y_2$, and $m - 1$, respectively. Since $y_1$ and $y_2$ have now been extracted from the retarded commutator of this term, either $\mu < m - 1$ for this term, or there is no such term (i.e., $m = 1$), and the induction was completed at the step before this one. In the former case we set $\nu = \mu - 1$, and $n = m - 2$; then $\nu < n$, and $[a, T_+(\alpha^{m-1-\mu})]^{\mu-1}_R$ satisfies the induction hypothesis. This gives us an equation like (43), except that it starts with $\Delta^3$ and $j_3$ instead of $\Delta^2$ and $j_2$, and the sum over the indices $k_i$ starts with $k_\mu > \cdots > k_2 \geq 3$. These considerations reduce the left-hand side of Eq. (38) to

$$\int \Delta^1 \cdots \Delta^\lambda_R \cdots \Delta^m \times \Big\{ i\delta_1 \sum'_{k_\mu > \cdots > k_2 \geq 2}{}^{m} j_2 \cdots i\delta_{k_2} \cdots, \cdots i\delta_{k_\mu} \cdots j_m j_\lambda + j_1 i\delta_2 \sum'_{k_\mu > \cdots > k_2 \geq 3}{}^{m} j_3 \cdots i\delta_{k_2} \cdots, \cdots i\delta_{k_\mu} \cdots j_m j_\lambda + j_1 j_2 K_3 \cdots K_m [a, T_+(\alpha^{m-2-\mu})]^\mu_R \Big\} (d^4y)^m. \quad (44)$$

In the same manner we may extract the variables $y_3 \cdots y_{\lambda-1}$, satisfying the induction hypothesis at each step, and further reduce the left-hand side of Eq. (38) to

$$\int \Delta^1 \cdots \Delta^\lambda_R \cdots \Delta^m \times \Big\{ \sum_{k_1=1}^{\lambda-1} \sum'_{k_\mu > \cdots > k_2 > k_1}{}^{m} j_1 \cdots i\delta_{k_1} \cdots, \cdots i\delta_{k_\mu} \cdots j_m j_\lambda + j_1 \cdots j_{\lambda-1} K_\lambda \cdots K_m [a, T_+(\alpha^{m-\lambda+1-\mu})]^\mu_R \Big\} (d^4y)^m. \quad (45)$$

[12] The case $\mu = m$ is excluded anyway by Eq. (24).

Integration over $y_\lambda$ produces only a volume term,

$$\int \Delta_R^\lambda K_{\nu\lambda}[a, T_+(\alpha(y_\lambda)\alpha^{m-\lambda-\mu})]_R^\mu d^4y_\lambda = -[a, T_+(\alpha(x_\lambda)\alpha^{m-\lambda-\mu})]_R^\mu. \quad (46)$$

The surface terms vanish, at $y_\lambda^0 \to +\infty$ because of $\Delta_R^\lambda$, and at $y^0 \to -\infty$ because $y_\lambda$ appears only in terms $[a(y_\lambda), Q]_R$ and $\alpha(y_\lambda)$. Once past $y_\lambda$, we extract in turn the remaining $y$ variables, and at the end apply Eq. (11), thus finally reducing the left-hand side of Eq. (38) to

$$\int \Delta^1 \cdots \Delta_R^\lambda \cdots \Delta^m \sum_{k_1=1}^{m-\mu+1}{}' \sum_{k_\mu > \cdots > k_2 > k_1}^{m}{}' j_1 \cdots i\delta_{k_1} \cdots, \cdots i\delta_{k_\mu} \cdots j_m j_\lambda (d^4y)^m, \quad (47)$$

which is equal to the right-hand side of Eq. (38), because

$$\sum_{k_1=1}^{m-\mu+1}{}' \sum_{k_\mu > \cdots > k_1}^{m}{}' = \sum_{k_\mu \cdots > k_1 \geq 1}^{m}{}'.$$

The last step in the extraction, namely

$$\int \Delta^{m-\mu+1} \cdots \Delta^m K_{m-\mu+1} \cdots K_m [a, \alpha]_R^\mu (d^4y)^\mu = \int \Delta^{m-\mu+1} \cdots \Delta^m i\delta_{m-\mu+1} i\delta_{m-\mu} \cdots i\delta_m \alpha(x_\lambda)(d^4y), \quad (48)$$

offers no difficulty, since there are no volume terms, and assigning any variable to $\alpha$ except $x_\lambda$ makes the surface terms vanish as well.

Having verified Eq. (38), we supply the missing factors in front that make up $B_\lambda$, sum it over $\lambda$ and $\mu$, and so obtain

$$B\Lambda_a = \sum_{\lambda=1}^{m} B_\lambda \sum_{\mu=0}^{m-2} \sum_{k_\mu > \cdots > k_1 \geq 1}^{m}{}' j_1 \cdots i\delta_{k_1} \cdots, \cdots i\delta_{k_\mu} \cdots j_m j_\lambda. \quad (49)$$

We can use an interesting property of the $m$-point operator,[2]

$$\begin{aligned} i^m S^* \delta^m S &= (j_1 + i\delta_1)(j_2 + i\delta_2) \cdots \Lambda_\lambda \cdots (j_m + i\delta_m) j_\lambda \\ &= j_1 \cdots \Lambda_\lambda \cdots j_m j_\lambda + \sum_{k=1}^{m}{}' j_1 \cdots i\delta_k \cdots j_m j_\lambda \\ &\quad + \sum_{k_2 > k_1 \geq 1}^{m}{}' j_1 \cdots i\delta_{k_1} \cdots i\delta_{k_2} \cdots j_m j_\lambda \\ &\quad + \cdots + i^{m-1}\delta^{m-1} j_\lambda \\ &= \sum_{\mu=0}^{m-1} \sum_{k_\mu > \cdots > k_1 \geq 1}^{m} j_1 \cdots i\delta_{k_1} \cdots, \cdots i\delta_{k_\mu} \cdots j_m j_\lambda, \end{aligned} \quad (50)$$

to obtain

$$B\Lambda_a = \sum_{\lambda=1}^{m} B_\lambda (i^m S^* \delta^m S - i^{m-1}\delta^{m-1} j_\lambda). \quad (51)$$

But from Eq. (31),

$$\begin{aligned} \Lambda_b &= i^m \sum_{\lambda=1}^{m} B_\lambda (\delta^{m-1}(S^* \delta_\lambda S) - S^* \delta^m S) \\ &= \sum_{\lambda=1}^{m} B_\lambda (i^{m-1}\delta^{m-1} j_\lambda - i^m S^* \delta^m S) \\ &= -B\Lambda_a. \end{aligned}$$

This completes the proof that $B\Lambda = 0$. This was shown to be a necessary condition for the existence of a solution to Eq. (15). That it is also a sufficient condition can be seen[1] by observing that

$$i^m S^* \delta^m S = \Lambda$$

is a solution of Eq. (15).

## ACKNOWLEDGMENTS

It is a pleasure to acknowledge several stimulating discussions with Professor F. Rohrlich and Dr. J. G. Wray, and to thank Professor Rohrlich for critically reading the manuscript.

## APPENDIX

Proof that

$$\sum_{l=0}^{m} (-1)^l T_+(a^{m-l}) T_-(a^l) = \delta_{m,0}. \quad (A1)$$

For $m = 0$, we have

$$T_+(a^0) T_-(a^0) \equiv 1. \quad (A2)$$

For $m = 1$, we have

$$T_+(a) T_-(a^0) - T_+(a^0) T_-(a) \equiv a - a \equiv 0 \quad (A3)$$

Now proceed by induction. Suppose that (A1) holds when $m$ is replaced by $m - 1$. Exhibit $a_1$:

$$T_+(a^{m-l}) T_-(a^l) = T_+(a_1 a^{m-1-l}) T_-(a^l) + T_+(a^{m-l}) T_-(a_1 a^{l-1}). \quad (A4)$$

The ordering theorems[11]

$$\begin{aligned} T_+(a_1 a^n) &= T_+(a^n) a_1 + [a_1, T_+(a^n)]_R, \\ T_-(a_1 a^n) &= a_1 T_-(a^n) - [a_1, T_-(a^n)]_R \end{aligned} \quad (A5)$$

enable us to write

$$\begin{aligned} &\sum_{l=0}^{m} (-1)^l T_+(a^{m-l}) T_-(a^l) \\ &= \sum_{l=0}^{m-1} (-1)^l (T_+(a^{m-1-l}) a_1 + [a_1, T_+(a^{m-1-l})]_R) T_-(a^l) \\ &\quad + \sum_{l=1}^{m} (-1)^l T_+(a^{m-l}) (a_1 T_-(a^{l-1}) - [a_1, T_-(a^{l-1})]_R) \\ &= \sum_{l=0}^{m-1} (-1)^l \{ (T_+(a^{m-1-l}) a_1 \\ &\quad + [a_1, T_+(a^{m-1-l})]_R) T_-(a^l) - T_+(a^{m-1-l}) \\ &\quad \times (a_1 T_-(a^l) - [a_1, T_-(a^l)]_R) \} \\ &= \sum_{l=0}^{m-1} (-1)^l [a_1, T_+(a^{m-1-l}) T_-(a^l)]_R \\ &= \left[ a_1, \sum_{l=0}^{m-1} (-1)^l T_+(a^{m-1-l}) T_-(a^l) \right]_R = 0, \end{aligned} \quad (A6)$$

by hypothesis.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 5 MAY 1968

# Field-Theoretic Formulation of Quantum Statistical Mechanics*

MOORAD ALEXANIAN
*Lawrence Radiation Laboratory, University of California, Livermore, California*

(Received 7 July 1967)

The notions of strong convergence of state vectors, introduced by Haag in his formalism of axiomatic quantum field theory, are extended to the case of vectors with an infinite number of particles but finite densities. Some general properties of nonequilibrium distribution functions are derived without the use of power series expansions or any other simplifying assumption. An integral representation is obtained for the distribution functions which makes it possible to discuss their behavior for small and large energies and to obtain some information about the singularities of these functions when continued analytically.

## I. INTRODUCTION

Many of the basic problems in the theory of nonequilibrium statistical mechanics, e.g., approach to thermodynamic equilibrium, discovery of $H$-type theorems, irreversibility, etc., are reflected in the lack of understanding of the nature of a nonequilibrium system. Many of the attempts[1] have been devoted to the quantum-mechanical derivation of the Master equation for the occupation probability function which leads, in the limit of infinite time, to statistical equilibrium. Also, the early work of Bogoliubov[2] has been partially successful in obtaining a Boltzmann-like equation for the distribution function.

Beginning with any available theory, the distribution functions should, in principle, be calculated from the fundamental equations of the theory. However, the analysis of the kinetic equations of Bogoliubov, for example, is limited to power-series approximations. On the other hand, derivation of the Master equation is based on assumptions which are not entirely well understood. It appears therefore suitable to obtain statements about these distribution functions without undue use of additional assumptions.

In this work, a new formalism in axiomatic quantum statistical mechanics is presented which unifies two, so far directly unrelated, disciplines: axiomatic field theory and statistical mechanics. This unification is accomplished by the strong convergence of states in axiomatic field theory, as introduced by Haag[3] in his collision theory, when properly applied to statistical mechanics.

The main results are, therefore, the derivation of an integral representation for nonequilibrium distribution functions as the superposition of equilibrium distribution functions with different temperatures and conclusions resulting from it.

## II. ASYMPTOTIC CONDITIONS IN FIELD THEORY

Before considering the asymptotic condition relevant to quantum statistical mechanics, it seems proper to summarize rather briefly the existing asymptotic conditions in field theory. Presently, the main emphasis in axiomatic quantum field theory is the construction of scattering theories. As will become evident in what follows, the fundamental assumptions of field theory are extensible to quantum statistical mechanics and this extension, together with its implications, forms the basis for the present formalism.

The Lehmann, Symanzik, and Zimmermann[4] (LSZ) formulation of quantized fields introduced an asymptotic condition for the field operator $A(x)$ as a basic requirement of the theory. Let $\{f_\alpha(x)\}$ denote a complete and orthonormal system of positive frequency solutions of the Klein–Gordon equation

$$(\Box - m^2)f_\alpha(x) = 0, \tag{1}$$

with

$$-i\int d^3x f_\alpha(x)\overleftrightarrow{\frac{\partial}{\partial x_0}} f_\beta^*(x) = \delta_{\alpha\beta}, \tag{2}$$

where

$$f_\alpha(x)\overleftrightarrow{\frac{\partial}{\partial x_0}} f_\beta^*(x) \equiv f_\alpha(x)\frac{\partial}{\partial x_0} f_\beta^*(x) - f_\beta^*(x)\frac{\partial}{\partial x_0} f_\alpha(x). \tag{3}$$

The field operator $A(x)$ is given by

$$A(x) = \sum_\alpha \{f_\alpha(x)A^\alpha(t) + f_\alpha^*(x)A^{\alpha*}(t)\}, \tag{4}$$

with the coefficients

$$A^\alpha(t) = i\int_{x_0=t} d^3x A(x)\overleftrightarrow{\frac{\partial}{\partial x_0}} f_\alpha^*(x). \tag{5}$$

LSZ postulate the following asymptotic condition for

* Work performed under the auspices of the U.S. Atomic Energy Commission.

[1] L. van Hove, Physica **23**, 441 (1957).

[2] N. N. Bogoliubov, J. Phys. (U.S.S.R.) **10**, 256, 265, 1946; English transl. by E. Gora in *Studies in Statistical Mechanics*, J. de Boer and G. Uhlenbeck, Eds. (North-Holland Publ. Co., Amsterdam, 1962), Vol. 1.

[3] R. Haag, Phys. Rev. **112**, 669 (1958); Suppl. Nuovo Cimento **14**, 131 (1959).

[4] H. Lehmann, K. Symanzik, and W. Zimmermann, Nuovo Cimento **11**, 342 (1954); **1**, 205 (1955).

normalizable states:

$$\lim_{t\to\mp\infty}(\Phi, A^a(t)\Psi) = (\Phi, A^a_{\text{in,out}}\Psi), \tag{6}$$

where the right sides are independent of time. The fields $A_{\text{in,out}}(x)$, constructed by means of (4) with the aid of $A^a_{\text{in,out}}$, satisfy the interaction-free Klein–Gordon equation (1). The operator $A^{a*}_{\text{in}}$, when acting on the physical vacuum, creates an incoming particle with wavefunction $f_a(x)$ and the set of all such states may be used to construct a complete orthonormal system.

The smeared-out field $A^a(t)$ satisfies a weak convergence asymptotic condition[5] and establishes the particle interpretation of the quantum field theory of LSZ. It should be clear that this formulation of field theory is tailored to scattering theory and in fact the main results of the LSZ formulation are analyticity properties of the $S$-matrix elements and the many-particle structure of Green's functions.

The method of Haag[3] consists in deriving a time-like asymptotic condition for scattering states from the assumed asymptotic behavior of the vacuum expectation values at spacelike directions. Suppose $A(x)$ is an "almost-local operator field" which creates a state of a single particle from the vacuum. Then the main result of Haag is that the vector

$$\Psi_t \equiv A_{f_{\alpha_1}}(t)\cdots A_{f_{\alpha_n}}(t)\,|0\rangle, \tag{7}$$

formed by acting on the physical vacuum with the operators (5), asymptotically, $|t|\to\infty$, approaches a constant vector. This limit is independent of the choice of $A(x)$ and is given, for the case $t\to-\infty$, by $|\alpha_1\cdots\alpha_n\rangle^{(-)}$, i.e., the Heisenberg state with an initial configuration of $n$ particles which move before they collide according to the wavefunctions $f_{\alpha_1}\cdots f_{\alpha_n}$. The approach to the limit is in the sense of strong convergence.[5]

The existing asymptotic conditions in field theory are, therefore, mathematical requirements which insure every system of interacting particles to tend away from each other with increasing time so that, after a sufficiently long time, the particles are essentially free from each other's influence and, consequently, behave as interaction-free particles. It is clear that such behavior holds *only* for a system of a *finite* number of interacting particles (zero density). Therefore, the asymptotic conditions discussed so far are restricted to situations peculiar to scattering theory.

The basis of the formalism presented in this work is to extend the notions of asymptotic conditions[6] to the case of vectors with an *infinite* number of particles but finite density. These vectors are outside of the realm of interest of scattering theory but are of fundamental value in statistical mechanics.

Let the quantum-mechanical system, described by the ket vector $|\psi_t\rangle$, be specified as completely and accurately as is possible in accordance with the general laws of the theory. One makes the following basic assumptions:

(1) States of statistical equilibrium, represented by the ket vector $|\xi_t\rangle$, exist for the dynamical system.

(2) Every state $|\psi_t\rangle$ approaches, in the limit $t\to+\infty$, a unique equilibrium state $|\xi_t\rangle$.

The limit in (2) is in the sense of strong convergence,

$$\|\,|\psi_t\rangle - |\xi_t\rangle\| \to 0 \quad (t\to\infty), \tag{8}$$

i.e., the norm of the difference of the vectors in Hilbert space tends toward zero as $t\to+\infty$. Note that contrary to the usual conception of equilibrium in statistical mechanics, $|\xi_t\rangle$ represents a *single* state of the system and *not* an ensemble of states. Interestingly enough, these types of equilibrium states have also been considered by Friedrichs and are described by myriotic fields referred to as equidistribution states.[7]

An interesting feature of myriotic fields is that they do not possess vacuum states.[7] The vacuum state plays a central role in field theory as applied to scattering theory. However, the absence of vacuum states for myriotic fields is rewarded by the existence of equidistribution states[7] which, as remarked above, are analogous to the equilibrium states introduced in the present formalism.

The use of states with an infinite number of particles avoids the so-called "Poincaré cycles." However, the Hilbert space of such vectors is *not* the same as the Hilbert space usually assumed in axiomatic quantum statistical mechanics. The space assumed in these theories is an infinite-dimensional *separable* Hilbert space.[8] The more general spaces considered in the present formalism are referred to by the mathematician as *nonseparable* Hilbert spaces.[9]

[5] Let $f$ and $g$ be any two elements in the Hilbert space $H$. Denote by $(f, g)$ the scalar product of $f$ and $g$, and the norm of $f$ by $\|f\|$. The sequence $\{f_n\}$ converges strongly to $f$ if $\|f_n - f\|\to 0$. It converges weakly to $f$ if $(f_n, g)\to(f, g)$ for *any* $g$ in $H$.

[6] The more suitable asymptotic condition for this purpose is that of Haag (strong convergence approach). The LSZ asymptotic condition can be partially justified by Haag's formalism. See K. Hepp, Commun. Math. Phys. 1, 95 (1965).

[7] K. O. Friedrichs, *Mathematical Aspects of the Quantum Theory of Fields* (Interscience Publishers, Inc., New York, 1953), Part IV.

[8] See, for example, G. G. Emch, J. Math. Phys. 7, 1413 (1966). In this reference, for Hamiltonians having a purely continuous spectrum one can no longer describe physical states by means of density operators (in a separable Hilbert space). This is contrary to the present formalism where density operators are used but are bounded self-adjoint operators in a *nonseparable* (see Ref. 9) Hilbert space.

[9] A Hilbert space is *separable* if it contains a denumerable complete orthonormal set; it is *nonseparable* if complete orthonormal sets are not denumerable. F. Riesz and B. Sz.-Nagy, *Functional Analysis* (Frederick Ungar Publ. Co., New York, 1955), Chap. II.

## III. INTEGRAL REPRESENTATION

The formulation of quantum statistical mechanics introduced in the preceding section, though rather axiomatic, can be readily used to derive a useful integral representation for nonequilibrium distribution functions.

The state $|\psi_t\rangle$, as well as its asymptotic state $|\xi_t\rangle$, contains a maximal amount of information which, in principle, can be deduced from experiments. However, in practice no such measurements are ever attempted. From the asymptotic nature of the equilibrium state $|\xi_t\rangle$ one has that the set of states $\{|\xi_t\rangle\}$ forms a complete set with which to describe the dynamical system.[10] Therefore, for *any* operator $A$,

$$A = \sum_{\xi,\xi'} |\xi_t\rangle\langle\xi'_t| \langle\xi_t| A |\xi'_t\rangle. \tag{9}$$

Suppose one considers operators which limit themselves to gross measurements of the equilibrium system. These operators are functions, for example, of the density, energy, and momenta. Then there exists a group of states $|\xi_t\rangle$, denoted by $\Delta$, which give rise to the *same* values for these measurements. One usually refers to the states $|\xi_t\rangle$ as microscopic states and to the operators for gross measurements as macroscopic operators.[11] Let $G$ be one such macroscopic operator, then

$$\begin{aligned} \langle\xi_t| G |\xi'_t\rangle &= G(\Delta) \quad (\xi \in \Delta) \\ &\qquad\quad (\xi' \in \Delta) \\ &= 0 \qquad \text{otherwise.} \end{aligned} \tag{10}$$

This requirement truly defines the subspace $\Delta$ as a macroscopic state since it implies that the expectation value of *any* macroscopic operator is the same for *any* vector in $\Delta$. The subspace $\Delta$ can properly be called a macroscopic equilibrium state.

A constant is certainly a macroscopic operator. One obtains the following macroscopic orthogonality condition for the equilibrium states:

$$\begin{aligned} \langle\xi_t | \xi'_t\rangle &= \frac{1}{N_\Delta} \quad (\xi \in \Delta) \\ &\qquad\quad (\xi' \in \Delta) \\ &= 0 \qquad \text{otherwise,} \end{aligned} \tag{11}$$

where $N_\Delta$ represents the number of microstates contained in $\Delta$. It is clear that one must assume $N_\Delta$ to be finite.

For *any* macroscopic operator, (9) becomes

$$G = \sum_\Delta S_\Delta G(\Delta), \tag{12}$$

where

$$S_\Delta = \sum_{\substack{\xi' \in \Delta \\ \xi \in \Delta}} |\xi_t\rangle\langle\xi'_t|. \tag{13}$$

$G(\Delta)(\mathrm{Tr}\, S_\Delta)$ gives the expected value of the operator $G$ for states contained in $\Delta$ and $S_\Delta/\mathrm{Tr}\, S_\Delta$ represents the projection operator onto the macroscopic space $\Delta$. In fact, one has from (11) and (13) the usual equations for a projection operator,

$$S_\Delta S_{\Delta'} = S_{\Delta'} S_\Delta = 0 \quad (\Delta \neq \Delta'), \tag{14}$$

$$(S_\Delta/\mathrm{Tr}\, S_\Delta)^2 = S_\Delta/\mathrm{Tr}\, S_\Delta. \tag{15}$$

For macroscopic states, the microscopic completeness relation reduces to, after using (11) and (13),

$$\sum_\Delta \frac{S_\Delta}{\mathrm{Tr}\, S_\Delta} = 1, \tag{16}$$

which expresses macroscopic completeness.

The above considerations, although usually associated with notions in statistical mechanics; have purely field-theoretic foundations. These arise when considering the infinite number of degrees of freedom ascribed to a field.[12] In general, for states with an infinite number of particles, one must consider systems with an infinite number of degrees of freedom, thus making the infinite degrees of freedom problem in field theory of interest and relevance to statistical mechanics.

In studying the representations of the commutation (or anticommutation) relations for an infinite set of oscillators, one encounters peculiar behavior when considering the unitary transformations between different representations.[12] The Hilbert space, which is formed by a nondenumerable infinity of representations of the commutation relations, is divided into mutually exclusive *equivalence classes.* Any vector in a given class differs from another in the same class by at most a finite number of occupation numbers (so that the vectors in one equivalence class *all* have the same density). Whereas, vectors from different classes differ by a denumerable infinite number of occupation numbers and, consequently, have different densities and satisfy the orthogonality condition (11). Also, the vectors from two different equivalence classes are connected by an *improper* unitary operator, i.e., operators whose matrix elements, between vectors in the same class, vanish. The macroscopic operators introduced above are proper operators; therefore, the matrix elements of these operators satisfy Eq. (10).

[10] Strictly speaking, the completeness of the equilibrium states is an assumption on the same fundamental level as the axiom of completeness of the asymptotic states is in axiomatic quantum field theory. See D. Ruelle, Helv. Phys. Acta **35**, 147 (1962).

[11] See, for example, G. Emch, Helv. Phys. Acta **37**, 532 (1964).

[12] R. Haag, Kgl. Danske Videnskab. Selskab, Mat.-Fys. Medd. **29**, No. 12 (1955); see also G. Barton, *Introduction to Advanced Field Theory* (Interscience Publishers, Inc., New York, 1963), Chap. 13.

Since the dynamical system has constant energy

$$|\xi_t\rangle = e^{-i\tilde{H}t}|\xi\rangle, \tag{17}$$

where $\tilde{H}$ is the macroscopic Hamiltonian operator of the system and is obtained from the microscopic $H$ by

$$\tilde{H} = \sum_{\Delta} \frac{S_\Delta}{N_\Delta^2} \operatorname{Tr}(HS_\Delta), \tag{18}$$

the equilibrium state $|\xi_t\rangle$ is a *stationary state*.[13] Hence, $G(\Delta)$ and $S_\Delta$ are actually *time independent*. In what follows the subscript $t$ will be dropped in denoting the equilibrium state. Also, the stationary nature of the equilibrium states implies the completeness of the set of states $\{|\xi\rangle\}$, thus corroborating this same conclusion from the basic assumptions (1) and (2). It is worth remarking that the energy eigenvalue for an equilibrium state depends on the state through $\Delta$,

$$\tilde{H}|\bar{\xi}\rangle = \frac{\operatorname{Tr}(HS_\Delta)}{N_\Delta}|\bar{\xi}\rangle, \tag{19}$$

where

$$|\bar{\xi}\rangle = \sum_{\xi\in\Delta}|\xi\rangle \quad (\xi \in \Delta). \tag{20}$$

One can relax the maximal description of the system by introducing the density operator[8]

$$D_t = \sum_{\Delta} \frac{S_\Delta}{\operatorname{Tr} S_\Delta} P_t(\Delta), \tag{21}$$

where $D_t$ is normalized by the condition

$$\operatorname{Tr} D_t = \sum_{\Delta} P_t(\Delta) = 1. \tag{22}$$

$P_t(\Delta)$ represents the probability of finding the system in the macrostate $\Delta$ at time $t$. The density operator $D_t$ describes the state of the system macroscopically and Eq. (21) is consistent with expression (12) for an arbitrary macroscopic operator.

It is interesting to note that the density operator $D_t$ cannot satisfy the equation of motion satisfied by the usual (microscopic) density operator.[14] [Unless in the trivial case when there is no time dependence.] This can be seen directly from (13) and (19).

Consider the expected value of a microscopic operator $G$ for the system described by the density operator $D_t$,

$$\langle G\rangle_t \equiv \operatorname{Tr} GD_t = \sum_{\Delta} \tilde{G}(\Delta)P_t(\Delta) \operatorname{Tr} S_\Delta, \tag{23}$$

where the last equality follows from (10) and (11) and $\tilde{G}$ is the macroscopic operator obtained from $G$ by (18). Equation (23) expresses the desired result. It represents the expected value of any microscopic operator for any macroscopic state of the system as a superposition over the expected value of the *same* operator in *different* macroscopic equilibrium states.[15]

A specific microscopic operator is the ordinary operator for the number of particles of field theory. For the case when the momentum of the particles can take on any continuous value,

$$n(\mathbf{p}) = a^+(\mathbf{p})a(\mathbf{p}) \tag{24}$$

plays the role of the density of the number of particles in three-dimensional momentum space. The operator $a^+(\mathbf{p})[a(\mathbf{p})]$ creates [annihilates] a particle with momentum $\mathbf{p}$ and given mass $m$.[16]

As an illustration, consider a gas of infinite mass.[17] The steady-state velocity distribution function for such a gas is given by

$$f(\mathbf{v}) = Ae^{-(m/2kT)(\mathbf{v}-\mathbf{v}_0)^2}, \tag{25}$$

where all possible steady states are obtained by giving different values to the five independent constants $A$, $T$, and $\mathbf{v}_0$. The general representation (23) becomes, for this simple case,

$$f(\mathbf{v}, t) = \iiint Ae^{-(m/2kT)(\mathbf{v}-\mathbf{v}_0)^2}\sigma(\mathbf{v}_0, T, A; t)\, d\mathbf{v}_0\, dT\, dA, \tag{26}$$

which expresses the nonequilibrium distribution function for the gas as an integral over the equilibrium distribution functions. The function $\sigma(\mathbf{v}_0, T, A; t)$, to be referred to as the spectral function, is positive definite and by (22) is normalized to unity:

$$\iiint \sigma(\mathbf{v}_0, T, A; t)\, d\mathbf{v}_0\, dT\, dA = 1. \tag{27}$$

Since the probability of finding the nonequilibrium system in the equilibrium macrostate specified by $A$, $T$, and $\mathbf{v}_0$ must be finite, it follows that the worst singularity which can occur in the spectral function $\sigma(\mathbf{v}_0, T, A; t)$ is a Dirac $\delta$ function and represents the attainment of equilibrium.

The explicit dependence of $f(\mathbf{v}, t)$ on the other variables, e.g., spatial dependence, has been omitted in (26). All such dependence appears *only* in the spectral function.

---

[13] P. A. M. Dirac, *The Principles of Quantum Mechanics* (Oxford University Press, London, 1959), Chap. V.

[14] For the derivation and study of the equation of motion, see Ref. 11, Sec. 3.

[15] This result is reminiscent of the Lehmann spectral representation in field theory; see H. Lehmann, Nuovo Cimento **11**, 342 (1954).

[16] For simplicity, the states of a particle are characterized here by only a four-momenta. Of course, in general, there are other quantities to be specified, e.g., charge, spin, etc.

[17] J. H. Jeans, *The Dynamical Theory of Gases* (Dover Publ., Inc., New York, 1954), Chap. II.

It is clear that (26) is not the most general representation possible for this type of particle. In general, one will have to consider the equilibrium distribution function which includes mass rotation. This introduces an additional integration over three variables. Similarly, the case of a gas in the presence of an external field can be so treated provided one uses the appropriate equilibrium distribution function, and so on.

Recently, the work of Suchy[18] was brought to the author's attention and, with it, that of Weitzsch,[19] where the integral representation (26) is introduced as an ansatz and used for the study of strong shock waves. In the original work of Weitzsch the spectral function need not be positive definite; its sign, for different ranges of the variables, is determined by the actual solution of the equations of motion. The above derivation insures the positive definiteness of the spectral function.

Suchy was able to relate, via the integral representation (26), existing approaches for the treatment of processes with weak deviation from equilibrium with methods for the study of strong deviations. In the former class, a series expansion of (26) yields Grad's expansion[20] of the distribution function in tensorial Hermite polynomials. In the latter, one has the "two-stream" Gaussian distribution used by Liu and Lees[21] for Couette flow and the anisotropic Gaussian distribution used by Chew, Goldberger, and Low[22] for a plasma in a strong magnetic field.

The derivation of the integral representation (23) gives an understanding of the ansatz of Weitzsch. First, the ansatz is shown to be a consequence of a formalism based on fundamental physical assumptions. Second, the derivation brings forth the physical meaning of the spectral function and obtains its positive definiteness. Third, it generalizes the ansatz to other statistics.

## IV. INFINITE HOMOGENEOUS MEDIA

In the previous section, an integral representation for nonequilibrium distribution functions was derived. The representation (23) gives a physical insight into the constitution of a nonequilibrium system and may be of considerable aid in understanding nonequilibrium phenomena.

For an infinite homogeneous medium, the representation can be linked to the theory of Laplace transform from which questions of mathematical origin, but with important physical consequences, can be studied. These include, the class of functions which can be so represented, the question of the uniqueness of the spectral function, etc.

### A. Maxwell–Boltzmann Statistics

The equilibrium distribution function for a system of particles satisfying Maxwell–Boltzmann statistics is given by

$$n_{eq}^{\text{M-B}}(E, T) = 2N/\pi^{\frac{1}{2}}(1/kT)^{\frac{3}{2}}E^{\frac{1}{2}}e^{-E/kT}, \quad (28)$$

where $N$ is the total number of particles and is independent of the temperature $T$. The representation (23) becomes, when using (28) and after a trivial change of variable of integration,

$$\frac{\pi^{\frac{1}{2}}}{2N}\, k\, \frac{n^{\text{M-B}}(E, t)}{E^{\frac{1}{2}}} = \int_0^\infty e^{-E\tau}\, \rho^{\text{M-B}}(\tau, t)\, d\tau, \quad (29)$$

where

$$\rho^{\text{M-B}}(\tau, t) = \frac{\sigma^{\text{M-B}}[(k\tau)^{-1}, t]}{\tau^{\frac{1}{2}}}. \quad (30)$$

Equation (29) establishes the equivalence between the representation and expressing the distribution function as a Laplace integral.

From the general theory of the Laplace transform, one obtains the following[23]:

(1) $$\int_0^\infty \rho^{\text{M-B}}(\tau, t)\, d\tau = \infty. \quad (31)$$

(2) $n^{\text{M-B}}(E, t)/E^{\frac{1}{2}}$ is an analytic function of $E$ for Re $(E) > 0$ and it is completely monotonic for $E > 0$,

$$(-1)^k \frac{\partial^k}{\partial E^k}\left[\frac{n^{\text{M-B}}(E, t)}{E^{\frac{1}{2}}}\right] \geq 0. \quad (32)$$

(3) The (spectral) function $\rho^{\text{M-B}}(\tau, t)$ is uniquely determined by the distribution function $n^{\text{M-B}}(E, t)$ and is given by

$$\rho^{\text{M-B}}(\tau, t) = \frac{k(\pi)^{\frac{1}{2}}}{2N}\frac{1}{2\pi i}\int_{\epsilon-i\infty}^{\epsilon+i\infty} e^{E\tau}\, \frac{n^{\text{M-B}}(E, t)}{E^{\frac{1}{2}}}\, dE, \quad (33)$$

where the path of integration is the line $E = \epsilon$ and $\epsilon$ is infinitesimally small.

One can determine the small energy behavior of the distribution function from (31). Suppose $n^{\text{M-B}}(E, t)$

[18] K. Suchy, *3rd International Rarefied Gas Dynamics Symposium, Vol. I*, (Academic Press Inc., New York, 1963), p. 181.
[19] F. Weitzsch, Ann. Physik [7] **7**, 403 (1961).
[20] H. Grad, Commun. Pure Appl. Math. **2**, 325 (1949); **2**, 331 (1949).
[21] C.-Y. Liu and L. Lees, *Rarefied Gas Dynamics*, L. Talbot, Ed. (Academic Press Inc., New York, 1961), p. 391.
[22] G. F. Chew, M. L. Goldberger, and F. E. Low, Proc. Roy. Soc. (London) **A236**, 112 (1956).
[23] D. V. Widder, *The Laplace Transform* (Princeton University Press, Princeton, N.J., 1946), Chaps. II and VII.

is finite at $E = 0$, then, for $(0 \leq t < \infty)$,

$$n^{\text{M-B}}(E, t) \sim E^{\beta} f(E) \quad \begin{matrix}(E \to 0)\\ (0 \leq \beta \leq \tfrac{1}{2}),\end{matrix} \tag{34}$$

where

$$1/f(E) = o(E^{-\delta}) \quad \begin{matrix}(E \to 0)\\ (\delta > 0)\end{matrix}$$

and (35)

$$f(E) = o(E^{-\eta}) \quad \begin{matrix}(E \to 0)\\ (\eta > 0),\end{matrix}$$

with $\delta$ and $\eta$ infinitesimally small.

If one knows the distribution function as an analytic function of the energy, from (33) one determines the spectral function and, through the representation (23), the expected value of *every* macroscopic operator which refers to these particles.

Finally, suppose $\rho^{\text{M-B}}(\tau, t)$ possesses the following asymptotic expansion for high temperature and fixed[24] $t$ $(0 \leq t < \infty)$:

$$\rho^{\text{M-B}}(\tau, t) \approx \sum_{n=0}^{\infty} C_n \tau^{\lambda_n} \quad (\tfrac{1}{2} < \lambda_0 < \lambda_1 \cdots); \tag{36}$$

then for $E \to \infty$,[25]

$$\frac{(\pi)^{\frac{1}{2}} k}{2N} \frac{n^{\text{M-B}}(E, t)}{E^{\frac{1}{2}}} \approx \sum_{n=0}^{\infty} C_n \frac{\Gamma(n+1)}{E^{\lambda_n + 1}}. \tag{37}$$

Therefore, a nonequilibrium distribution function can vanish more slowly than an equilibrium distribution function as the energy approaches infinity. The integral (33) can be evaluated by enclosing the singularities to the left of $E = \epsilon$ by a suitable contour and making use of Cauchy's residue theorem. The value of the integral over the contour (at infinity) vanishes by (37) and one is left with only the contributions due to poles and cuts to the left of $E = \epsilon$.

### B. Bose-Einstein Statistics

For concreteness, a system of photons will be considered in studying the implications of the representation (23) for a system of particles satisfying B–E statistics. The analysis, of course, will not differ if other cases are considered.

The equilibrium distribution function for a system of photons is given by Planck's radiation law:

$$n_{eq}^{\text{B-E}}(E, T) = \frac{8\pi}{c^3} \frac{E^2}{h^3} \frac{1}{e^{E/kT} - 1}. \tag{38}$$

The representation (23) becomes, after a change of integration variable,

$$\frac{c^3 h^3}{8\pi} \frac{k}{E^2} n^{\text{B-E}}(E, t) = \int_0^{\infty} \frac{\rho^{\text{B-E}}(\tau, t)}{e^{E\tau} - 1} d\tau, \tag{39}$$

where

$$\rho^{\text{B-E}}(\tau, t) = \frac{\sigma^{\text{B-E}}[(k\tau)^{-1}, t]}{\tau^2}. \tag{40}$$

It is clear from (39) that for fixed $t$ $(0 \leq t < \infty)$,

$$\rho^{\text{B-E}}(\tau, t) \underset{\tau \to 0}{\longrightarrow} \tau^{\alpha} \quad (\alpha > 0). \tag{41}$$

One has, then, that the singularity of the right hand of (39) at $E = 0$ is a simple pole, just as for the equilibrium distribution. Thus, the low frequency photons are in equilibrium with the temperature

$$\int_0^{\infty} T\sigma^{\text{B-E}}(T, t)\, dT.$$

For the general Bose gas, the preservation of this singularity is of interest in studying Bose condensation.

The integral (39) represents, to the author's knowledge, a new mathematical transform which resembles the Laplace transform and is studied in some detail in the Appendix. It follows from the theorems in the Appendix that

(1) $[n^{\text{B-E}}(E, t)]/E^2$ is an analytic function of the complex variable $E$ for Re $(E) > 0$ and it is completely monotonic for $E > 0$. [Recall that $(e^x - 1)^{-1}$ is completely monotonic, see (32).]

(2) The spectral function $\sigma^{\text{B-E}}(T, t)$ is uniquely determined by the distribution $n^{\text{B-E}}(E, t)$.

In proving the uniqueness of the spectral function by Theorem 4 in the Appendix, one must assume $\alpha \geq 1$ [see (41)]. This justifies taking a limit on the energy variable inside the integral sign. For instance,

$$\frac{c^3 h^3}{8\pi} \frac{k}{E^2} n^{\text{B-E}}(E, t) \sim \int_0^{\infty} e^{-E\tau} \rho^{\text{B-E}}(\tau, t)\, d\tau \quad (E \to \infty). \tag{42}$$

Therefore, the representation reduces to that for M–B statistics, thus insuring the passage, in the limit of high energy, of a quantum-mechanical nonequilibrium system to a classical nonequilibrium system.

Analogous results as in (37) will follow from (42) if the function $\rho^{\text{B-E}}(\tau, t)$ has the behavior expressed in (36) (with $\lambda_0 > 1$). If one requires the total energy to be finite, then $\lambda_0 > 3$.

The spectral function $\sigma^{\text{B-E}}(T, t)$ is determined only by a knowledge of the high-energy behavior of the distribution and, together with (39), gives the behavior for *all* energies. This may seem somewhat surprising, but one must remember that quantum-mechanical features are already contained in the equilibrium distribution and hence in (39).

[24] The value of $\frac{1}{2}$ for the lower bound follows from the normalization condition $\int_0^{\infty} \sigma^{\text{M-B}}(T, t)\, dT = 1$ [see (27)]. If, in addition, one requires the total energy to be finite then $(\lambda_0 > \frac{3}{2})$.

[25] G. Doetsch, *Theorie und Anwendung der Laplace Transformation* (Dover Publications, Inc., New York, 1943), Chap. 12.

### C. Fermi–Dirac Statistics

The proofs of analyticity of the distribution function and uniqueness of the spectral function follow closely those for B–E statistics.

Consider the equilibrium distribution function where the number of particles is conserved,

$$n_{eq}^{\text{F-D}}(E, T) = \frac{D(E)}{e^{\mu} e^{E/kT} + 1}, \tag{43}$$

where $D(E)$ is a known phase-space factor and $\mu$ is a function of $T$ determined by normalization. Equations (23) and (43) give

$$k \frac{n^{\text{F-D}}(E, t)}{D(E)} = \int_0^{\infty} \frac{\rho^{\text{F-D}}(\tau, t)\, d\tau}{e^{\mu(\tau)} e^{\tau E} + 1}, \tag{44}$$

where

$$\rho^{\text{F-D}}(\tau, t) = \frac{\sigma^{\text{F-D}}[(k\tau)^{-1}, t]}{\tau^2}. \tag{45}$$

As before,

(1) $n^{\text{F-D}}(E, t)/D(E)$ is an analytic function of $E$ for $\text{Re}\,(E) > 0$ with the possible exception of the line $\text{Re}\,(E) = \gamma$, where it may be discontinuous.

(2) The spectral function $\sigma^{\text{F-D}}(T, t)$ is uniquely determined from the distribution $n^{\text{F-D}}(E, t)$.

To obtain (2) and, consequently, the approach to the classical limit at high energy, one must assume

$$\frac{\rho^{\text{F-D}}(\tau, t)}{1 + e^{\mu(\tau)}} \xrightarrow[\tau \to 0]{} 0. \tag{46}$$

If one considers the nonrelativistic[26] form for the energy $E$ in (43),

$$e^{-\mu(\tau)} \xrightarrow[\tau \to 0]{} \tau^{\frac{3}{2}}. \tag{47}$$

The normalization to unity of $\sigma^{\text{F-D}}(T, t)$ requires

$$\rho^{\text{F-D}}(\tau, t) \xrightarrow[\tau \to 0]{} \tau^{\alpha} \quad (\alpha > -1). \tag{48}$$

Therefore, for this case (46) is satisfied.

The existence of a line of discontinuity in (1) is associated with the large $\tau$ behavior of $e^{\mu(\tau)}$ [see Appendix]. One can resort to the previous nonrelativistic case for a determination of the exact value of $\gamma$. Since this case holds rigorously for all fermions in the limit of zero temperature[26]

$$e^{\mu(\tau)} \xrightarrow[\tau \to \infty]{} e^{-\gamma\tau}, \tag{49}$$

where

$$\gamma = \frac{h^2}{2mk} \left[\frac{3}{4\pi} \frac{N}{V}\right]^{\frac{2}{3}}. \tag{50}$$

[26] E. Schrödinger, *Statistical Thermodynamics* (Cambridge University Press, New York, 1964), Chap. VIII.

Therefore, if the spectral function $\sigma^{\text{F-D}}(T, t)$ does not vanish identically in a small neighborhood of the absolute zero, then $n^{\text{F-D}}(E, t)/D(E)$ will be discontinuous at $E = \gamma$.

## V. CONCLUSION

The basic attitude in the present work is to shift one's attention from the problem of "the approach to equilibrium" to that of obtaining general information derived from such behavior. A field-theoretic formulation of quantum statistical mechanics is presented and is used to obtain an integral representation which allows for the study of low- and high-energy behavior for nonequilibrium distribution functions, as well as analyticity and singularity structure in the complex energy variable. For instance, the pole at $E = 0$, which appears in the Bose–Einstein equilibrium distribution, persists for the nonequilibrium case.

The theory of quantum statistical mechanics presented is based on purely axiomatic field-theoretic notions. Since axiomatic field theory is developed as a relativistic quantum theory, the formalism and results obtained are easily extended to relativistic quantum statistical mechanics.

The question of the equations of motion satisfied by the probability distribution is not studied here, but is shown to correspond to quantities considered by other authors. The representation for the distribution function prescribes the form for the initial distribution. Description not in accord with such representation *cannot* lead to equilibrium. For example, for M–B statistics an initial description of the particles by means of a Gaussian distribution in *energy* is not allowable.

## ACKNOWLEDGMENTS

The author would like to thank Dr. Robert J. Gelinas for discussions and Dr. Balram Prasad of Brandeis University for bringing Ref. 18 to the author's attention.

## APPENDIX

The integral transforms (39) and (44) are special cases of the general transform

$$G(p) = \int_0^{\infty} \frac{f(t)\, dt}{e^{pt} + \lambda(t)}, \tag{A1}$$

where $\lambda(t)$ is a known (real) analytic function of $t$. The function $f(t)$ is real and together with its first derivative is sectionally continuous. For $\lambda(t) \equiv 0$, (A1) reduces to the ordinary Laplace transform.

In the domain $(0 \leq t < \infty)$ the range of $\lambda(t)$ is contained in $(-1 \leq \lambda(t) < \infty)$. Further, the function

$\lambda(t)$ is of exponential order at infinity,

$$\lambda(t) = o(e^{\gamma t}) \quad \begin{matrix}(t \to \infty)\\ (\gamma > 0).\end{matrix} \tag{A2}$$

Let $p$ be a complex variable with real and imaginary parts $\sigma$ and $\tau$, respectively. One now proves Theorem 1.

*Theorem 1:* If

$$\underset{0\le u\le\infty}{\text{u.b.}} \left| \int_0^u \frac{f(t)\,dt}{e^{p_0 t} + \lambda(t)} \right| = M < \infty \quad (\sigma_0 > \gamma), \tag{A3}$$

then (A1) converges for every $p$ for which $\sigma > \sigma_0$, and

$$\int_0^\infty \frac{f(t)\,dt}{e^{pt} + \lambda(t)} = \int_0^\infty \frac{(p - p_0)e^{(p+p_0)t} + \lambda(pe^{pt} - p_0 e^{p_0 t}) - (d\lambda/dt)(e^{pt} - e^{p_0 t})}{(e^{pt} + \lambda)^2} \beta(t)\,dt, \tag{A4}$$

where

$$\beta(t) \equiv \int_0^t \frac{f(u)\,du}{e^{p_0 u} + \lambda(u)} \quad (0 \le t \le \infty), \tag{A5}$$

the integral on the right-hand side of (A4) converging absolutely.

Definition (A5) for $\beta(t)$ gives

$$\int_0^R \frac{f(t)\,dt}{e^{pt} + \lambda(t)} = \frac{e^{p_0 R} + \lambda(R)}{e^{pR} + \lambda(R)} \beta(R) + \int_0^R \frac{(p - p_0)e^{(p_0+p)t} + \lambda(pe^{pt} - p_0 e^{p_0 t}) - (d\lambda/dt)(e^{pt} - e^{p_0 t})}{(e^{pt} + \lambda)^2} \beta(t)\,dt. \tag{A6}$$

By (A2) and (A3), the first term in (A6) goes to zero as $R \to \infty$ for $\sigma > \sigma_0$. Hypothesis (A3) implies

$$\left| \int_0^\infty \frac{(p - p_0)e^{(p_0+p)t} + \lambda(pe^{pt} - p_0 e^{p_0 t}) - (d\lambda/dt)(e^{pt} - e^{p_0 t})}{(e^{pt} + \lambda)^2} \beta(t)\,dt \right| \le M \int_0^\infty dt \left| \frac{d}{dt}\left(\frac{e^{p_0 t} - e^{pt}}{e^{pt} + \lambda(t)}\right) \right| < \infty.$$

The last inequality follows from (A2). Hence, the theorem is proved.

*Corollary 1:* If

$$\underset{0\le u\le\infty}{\text{u.b.}} \left| \int_0^u \frac{f(t)\,dt}{e^{p_0 t} + \lambda(t)} \right| = N < \infty \quad (\sigma_0 < \gamma), \tag{A7}$$

then (A7) converges for every $p$ for which $\sigma_0 < \sigma < \gamma$, and

$$\int_0^\infty \frac{f(t)\,dt}{e^{pt} + \lambda(t)} = \int_0^\infty \frac{f(t)\,dt}{e^{p_0 t} + \lambda(t)} + \int_0^\infty \frac{(p - p_0)e^{(p_0+p)t} + \lambda(pe^{pt} - p_0 e^{p_0 t}) - (d\lambda/dt)(e^{pt} - e^{p_0 t})}{(e^{pt} + \lambda)^2} \beta(t)\,dt, \tag{A8}$$

where $\beta(t)$ is defined in (A5), the integral on the right-hand side of (A8) converging absolutely.

The proof of this is exactly the same as that of Theorem 1.

*Theorem 2:* If the integral

$$G(p) = \int_0^\infty \frac{f(t)\,dt}{e^{pt} + \lambda(t)} \tag{A9}$$

converges at $p = \sigma_0 + i\tau_0$, and if $H$ and $K$ are any constants for which $H > 0$, $K > 1$, then the integral (A9) converges uniformly in the region $\Delta$ defined by the inequality

$$|p - p_0| \le K \Big/ \int_H^\infty dt \left[ e^{(\sigma_0+\sigma)t} + |\lambda| \frac{\sigma + \sigma_0}{\sigma - \sigma_0} e^{\sigma t} + 2 \left| \frac{d\lambda}{dt} \right| \frac{e^{\sigma t}}{\sigma - \sigma_0} \right] \Big/ (e^{\sigma t} - |\lambda|)^2 \quad (\sigma \ge \sigma_0). \tag{A10}$$

Let $G(p)$ converge at $p = p_0$. Define $\beta(t)$ as in (A5). One has that

$$\int_R^\infty \frac{f(t)\,dt}{e^{pt} + \lambda(t)} = \int_R^\infty \frac{(p - p_0)e^{(p_0+p)t} + \lambda(pe^{pt} - p_0 e^{p_0 t}) - (d\lambda/dt)(e^{pt} - e^{p_0 t})}{(e^{pt} + \lambda)^2} \times [\beta(t) - \beta(R)]\,dt \quad (\sigma > \sigma_0 > \gamma). \tag{A11}$$

Let $R_0 > H$ be such that

$$|\beta(t) - \beta(t')| \le \epsilon/K \quad \begin{matrix}(t > R_0)\\ (t' > R_0).\end{matrix} \tag{A12}$$

This follows from the convergence at $p_0$ [see (A3)]. If $R > R_0$, (A11) gives

$$\left| \int_R^\infty \frac{f(t)\,dt}{e^{pt} + \lambda(t)} \right| \le \frac{\epsilon}{K} \int_R^\infty dt\, |[(p - p_0)e^{(p_0+p)t} + \lambda(pe^{pt} - p_0 e^{p_0 t}) - (d\lambda/dt)(e^{pt} - e^{p_0 t})]/(e^{pt} + \lambda)^2|$$

$$\le \frac{\epsilon}{K} |p - p_0| \int_R^\infty dt \left[ e^{(\sigma_0+\sigma)t} + |\lambda| \frac{\sigma + \sigma_0}{\sigma - \sigma_0} e^{\sigma t} + 2 \left| \frac{d\lambda}{dt} \right| \frac{e^{\sigma t}}{\sigma - \sigma_0} \right] \Big/ (e^{\sigma t} - |\lambda|)^2 \quad (\sigma > \sigma_0). \tag{A13}$$

For $p = p_0$,

$$\left| \int_R^\infty \frac{f(t)\,dt}{e^{pt} + \lambda(t)} \right| = |\beta(\infty) - \beta(R)| \leq \frac{\epsilon}{K} < \epsilon. \quad \text{(A14)}$$

So that for all $p$ in $\Delta$,

$$\left| \int_R^\infty \frac{f(t)\,dt}{e^{pt} + \lambda(t)} \right| < \epsilon,$$

which proves the theorem.

An analogous theorem can be proved for $\sigma_0 < \sigma < \gamma$.

*Theorem 3:* If the integral

$$G(p) = \int_0^\infty \frac{f(t)\,dt}{e^{pt} + \lambda(t)} \quad \text{(A15)}$$

converges for $\sigma > \sigma_c < \infty$, then $G(p)$ is analytic for $\sigma > \sigma_c$.

If $p_0$ is an arbitrary point in the half-plane $\sigma > \sigma_c$, one can surround it by a circle $K$ which also lies in that half-plane. By Theorem 2, the integral (A15), and hence the series

$$G(p) = \sum_{n=0}^{\infty} \int_n^{n+1} \frac{f(t)\,dt}{e^{pt} + \lambda(t)}, \quad \text{(A16)}$$

converges uniformly in $K$. Since each term of the series is entire, one has that $G(p)$ is analytic for $\sigma > \sigma_c$.

Finally, one establishes a sufficient condition for proving the uniqueness of the spectral function $f(t)$.

*Theorem 4:* If the integral

$$G(p) = \int_0^\infty \frac{f(t)\,dt}{e^{pt} + \lambda(t)} \quad \text{(A17)}$$

converges for $p = \sigma_0 + i\tau_0$, and if $f(t)/[e^{pt} + \lambda(t)]$ is a continuous function of both variables when $t \geq 0$ and $\sigma \geq \sigma_0$, then $f(t)$ is unique.

By Theorem 1, (A17) converges for $\sigma > \sigma_0$ and uniformly in $\Delta$ [see Theorem 2]. From the uniform convergence of $G(p)$ and the assumed continuity of the integrand, one has that[27] $G(p)$ is a continuous function of $p$ for $\sigma \geq \sigma_0$, so that

$$G(p) \xrightarrow[|p| \to \infty]{} g(p) \equiv \int_0^\infty e^{-pt} f(t)\,dt. \quad \text{(A18)}$$

Since $g(p)$ is the Laplace transform of $f(t)$, therefore $f(t)$ is unique.

As an example to this theorem, let $\lambda(t)$ be independent of $t$. Now $g(p) = p^{-1}[f(t) = \theta(t)]$ gives $G(p) = [\ln(1 + \lambda)]/\lambda p$ which does not satisfy (A18) for $(\lambda \neq 0)$. However, $g(p) = e^{-pa}/p$ with $a > 0$ $[f(t) = \theta(t - a)]$ yields $G(p) = (\lambda p)^{-1} \ln [1 + e^{-pa}\lambda]$ which does satisfy (A18). In the first example the continuity hypothesis on the integrand is not satisfied since

$$\lim_{t \to 0^+} \lim_{|p| \to \infty} \frac{f(t)}{e^{pt} + \lambda} = 0,$$

$$\lim_{|p| \to \infty} \lim_{t \to 0^+} \frac{f(t)}{e^{pt} + \lambda} = \frac{f(0)}{1 + \lambda}.$$

One can reconcile this by requiring $f(0) = 0$. Note that this requirement is not satisfied by the unit step function $\theta(t)$ of the above example which violated the theorem.

[27] E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis* (Cambridge University Press, New York, 1962), Chap. IV.

# Integral Representation for Systems of Interacting Particles*

MOORAD ALEXANIAN
*University of California, Lawrence Radiation Laboratory, Livermore, California*

(Received 7 July 1967)

This is a sequel to a previous paper, designated as I, which dealt primarily with nonequilibrium systems. The ideas in I are extended and include the study of systems of interacting particles in equilibrium. An integral representation is obtained for the distribution function of interacting particles as a superposition of distribution functions for noninteracting particles. In particular, it is shown that an interacting Bose gas need not show Bose condensation and that the behavior near the point $E = 0$ *cannot* be more singular than that of a simple pole.

## I. INTRODUCTION

In a previous paper,[1] to be designated as I, an integral representation for a nonequilibrium distribution function was derived. The derivation is based on some basic postulates, which are somewhat related to the notions of strong convergence in axiomatic quantum field theory [2] and invoke no specific dynamical assumptions.

The basic ideas contained in I are sufficiently general as to entail, in some cases, a similar representation for the equilibrium distribution function for a system of interacting particles.

The study of this extension, together with some of its implications, is the subject of the present work. The notation follows I; equations labeled I refer to that article. The reader is referred to I for most bibliographic references.

## II. INTEGRAL REPRESENTATION

One of the postulates in I concerns the existence of states of statistical equilibrium. It is clear that these states describe interacting particles and, for the purpose of this paper, will be denoted by a subscript I.

In using the integral representation (I23), one needs to know the expected value of macroscopic operators for these interacting states. This is indeed a difficult problem and deserves attention on its own right. However, one can still obtain an integral representation in terms of noninteracting states for many cases of considerable interest and with obvious advantages.

Let $|\xi_I\rangle$ and $|\xi\rangle$ denote the interacting and noninteracting equilibrium states, respectively. Consider the limit which lets all the coupling constants of the interaction between the particles approach zero. Denote this limit by $\lim_{g\to 0}$. Then,

$$\lim_{g\to 0} |\xi_I\rangle = |\xi\rangle. \tag{1}$$

The general conservation laws are independent of the interaction strength; therefore, quantities like the chemical potential, for example, will appear automatically in the noninteracting states.

Consider now the case where bound states are excluded from the theory or cases where bound states can occur but that the interaction strengths are sufficiently small that there exist no bound particles among the interacting particles. Then the set of noninteracting equilibrium states $\{|\xi\rangle\}$ forms also a complete set. Equation (I12) for *every* macroscopic operator $G$ becomes, with the present notation,

$$G = \sum_{\Delta_I} S_{\Delta_I} G(\Delta_I). \tag{2}$$

From the completeness of the states $\{|\xi\rangle\}$ one has the similar equation

$$G = \sum_{\Delta} S_{\Delta} G(\Delta). \tag{3}$$

Equating (2) and (3) and after multiplying both sides by $S_{\Delta_I}/\mathrm{Tr}\, S_{\Delta_I}$ and using (I14) and (I15),

$$\frac{S_{\Delta_I}}{\mathrm{Tr}\, S_{\Delta_I}} G(\Delta_I) = \sum_{\Delta} \frac{S_{\Delta_I} S_{\Delta}}{(\mathrm{Tr}\, S_{\Delta_I})^2} G(\Delta). \tag{4}$$

Finally, by taking the trace of (4),

$$G(\Delta_I) = \sum_{\Delta} \frac{\mathrm{Tr}\,(S_{\Delta_I} S_{\Delta})}{(\mathrm{Tr}\, S_{\Delta_I})^2} G(\Delta). \tag{5}$$

The representation (5) expresses the expected value of every macroscopic operator for any interacting state of the system as a superposition over the expected value of the same operator in different noninteracting states of the system. This result is quite analogous to the representation obtained in I for nonequilibrium systems. Actually, (5) is nothing else than the change of basis from one macroscopic covering of the Hilbert space to another. Its interest lies, of course, in the meaning of the two different coverings.

The coefficients in (5) are positive definite. Using (I13) and (I11),

$$\mathrm{Tr}\,(S_{\Delta_I} S_{\Delta}) = |M|^2, \tag{6}$$

with

$$M = \sum_{\substack{\xi_I \epsilon \Delta_I \\ \xi \epsilon \Delta}} \langle \xi \,|\, \xi_I \rangle \tag{7}$$

* This work was performed under the auspices of the U.S. Atomic Energy Commission.

[1] M. Alexanian, J. Math. Phys., **9**, 725 (1968).

[2] R. Haag, Phys. Rev. **112**, 669 (1958); Suppl. Nuovo Cimento **14**, 131 (1959).

and

$$\mathrm{Tr}\, S_{\Delta_I} = N_{\Delta_I}, \tag{8}$$

where $N_{\Delta_I}$ denotes the number of microstates contained in $\Delta_I$.

The representation (I23) for nonequilibrium systems with the present notation for interacting states is

$$\langle G\rangle_t = \sum_{\Delta_I} G(\Delta_I) P_t(\Delta_I)\, \mathrm{Tr}\, S_{\Delta_I}. \tag{9}$$

With the aid of (5) one obtains

$$\langle G\rangle_t = \sum_{\Delta} G(\Delta) P_t(\Delta)\, \mathrm{Tr}\, S_{\Delta}, \tag{10}$$

where the probability of finding the nonequilibrium system in the noninteracting macroscopic equilibrium state $\Delta$ is given in terms of the similar quantity for the interacting state by

$$P_t(\Delta) = \sum_{\Delta_I} P_t(\Delta_I) \frac{\mathrm{Tr}\,(S_{\Delta_I} S_{\Delta})}{(\mathrm{Tr}\, S_{\Delta_I})(\mathrm{Tr}\, S_{\Delta})}. \tag{11}$$

From (I13) and (I16) one has that

$$\sum_{\Delta} \frac{\mathrm{Tr}\,(S_{\Delta_I} S_{\Delta})}{(\mathrm{Tr}\, S_{\Delta_I})(\mathrm{Tr}\, S_{\Delta})} = 1. \tag{12}$$

Therefore, (11) and (12) give

$$\sum_{\Delta} P_t(\Delta) = \sum_{\Delta_I} P_t(\Delta_I) = 1, \tag{13}$$

thus establishing the representation (I23) obtained in I in terms of noninteracting equilibrium states for cases where bound states are excluded from the theory.

## III. INTERACTING BOSE GAS

The representation (5) for interacting states is quite analogous to that of I for nonequilibrium states. In fact, the results of Sec. IV of I for nonequilibrium distribution functions hold true also for the equilibrium distribution function of interacting particles, provided no bound states exist in the theory.

One case which was not considered in I is the general Bose gas where the number of particles is conserved. This differs more from the Bose gas with particle nonconservation than one would think at first. This, of course, is due to the possibility of Bose condensation in an interacting Bose gas.

The equilibrium distribution function for a system of particles satisfying Bose–Einstein statistics where the number of particles is conserved is given by

$$n_{\mathrm{eq}}(E, T) = R(E)/[e^{\mu(N,T)} e^{E/kT} - 1], \tag{14}$$

where $R(E)$ is a known function of the energy $E$ of the particle and $\mu(N, T)$ is a function of the temperature $T$ and the number of particles $N$, the functional relationship being determined by normalization.[3]

The function $\mu(N, T)$ is nowhere negative and vanishes at the temperature at which Bose condensation occurs,

$$\mu(N, T_0) = 0. \tag{15}$$

For lower temperatures,

$$\mu(N, T) \equiv 0 \qquad (T < T_0). \tag{16}$$

The particle distribution function for the interacting system is obtained from (5),

$$n(E, T) = \int_0^{\infty} \frac{R(E)\sigma(T, T')\, dT'}{e^{\mu(N,T')} e^{E/kT'} - 1}. \tag{17}$$

Note that in (17) one has no integration over the number of particles since the system contains a well-defined, fixed number of particles. Hence, (17) expresses $n(E, T)$ as a superposition of systems of noninteracting bosons with different temperatures $T'$ but the *same* number of particles $N$. This leads to the following relation between *noninteracting* quantities:

$$N = n_0(T')\theta(T_0 - T') + n(T'), \tag{18}$$

where $n_0(T')$ denotes the number of particles in the condensate[4] and $N$ is independent of $T'$.

The spectral function $\sigma(T, T')$ is normalized to unity by (12),

$$\int_0^{\infty} \sigma(T, T')\, dT' = 1. \tag{19}$$

In order to study the singularity structure of $n(E, T)$ one writes (17), with the aid of (15) and (16), as follows:

$$\frac{n(E, T)}{R(E)} = \int_0^{T_0} \frac{\sigma(T, T')\, dT'}{e^{E/kT'} - 1} + \int_{T_0}^{\infty} \frac{\sigma(T, T')\, dT'}{e^{\mu(N,T')} e^{E/kT'} - 1}. \tag{20}$$

The first term in (20) has a simple pole at $E = 0$ (see Appendix) with residue

$$\alpha(T) = k \int_0^{T_0} T'\sigma(T, T')\, dT'. \tag{21}$$

As in I, the order of the pole is established by the normalization condition (19).

It is instructive to express the number of particles in the condensate for the interacting system $N_0(T)$ in terms of the spectral function. One obtains from (18) and the normalization condition (19)

$$N = \int_0^{T_0} n_0(T')\sigma(T, T')\, dT' + \int_0^{\infty} n(T')\sigma(T, T')\, dT', \tag{22}$$

so that

$$N_0(T) = \int_0^{T_0} n_0(T')\sigma(T, T')\, dT'. \tag{23}$$

The $\lambda$-transition temperature $T_\lambda$ is usually defined by[5]

$$N_0(T_\lambda) = 0. \tag{24}$$

[3] Here, as in I, one deals always with an infinite system ($V \to \infty$, $N \to \infty$ such that $N/V$ is finite). However, for the purpose of this section it is convenient to consider $N$ rather than the density $\rho$.

[4] The explicit form $n_0(T') = N[1 - (T'/T_0)^{\frac{3}{2}}]$ is obtained for nonrelativistic particles.

[5] This definition is only meaningful if $N_0(T - \epsilon) \neq 0$ for $\epsilon$ infinitesimally small and positive.

The known general properties of the spectral function (positive definiteness and normalizability) are not sufficient to establish the uniqueness of $T_\lambda$ nor rule out the possibility that (24) may hold identically for all $T_\lambda$. The physical implication of the latter supposition is that an interacting Bose gas need not exhibit Bose condensation.

Suppose the integral in (23) is continuous at $T = T_\lambda$. From the positiveness of $n_0(T')$ and $\sigma(T, T')$, (24) is equivalent to[6]

$$\sigma(T_\lambda, T') = 0 \quad (0 \leq T' < T_0). \tag{25}$$

Similarly, $\alpha(T_1) = 0$ implies

$$\sigma(T_1, T') = 0 \quad (0 < T' \leq T_0). \tag{26}$$

Hence, the vanishing of $N_0(T)$ need not imply the vanishing of $\alpha(T)$, and conversely.

An upper bound can be derived from (17) for the particle distribution function. Define the "temperature" $T^*$ by

$$T^* \equiv \int_0^\infty T' \sigma(T, T')\, dT' < \infty, \tag{27}$$

where the bound on $T^*$ is a consequence of the assumed finiteness of the total energy. From the physical meaning of the spectral function as a probability function, it is tempting to identify $T^*$ with the temperature $T$ of the system. However, the integral (27) will, in general, depend on other parameters of the theory, e.g., the strength of the interaction, which are quantities *independent* of the temperature $T$.

One gains an insight into the meaning of $T^*$ by considering the implications of its vanishing. One has from (17) and (27) that the vanishing of $T^*$ requires all the particles to occupy the level with $E = 0$. With this result in mind, $T^*$ is the analog, in the interacting case, of the temperature in the ideal Bose gas. [In fact in the limit $T^* \to 0$ they become one and the same.] From (17), (21), and (27),

$$n(E, T)/R(E) \leq kT^*/E \tag{28}$$

and

$$0 \leq \alpha(T) \leq kT^*. \tag{29}$$

The bound in (28) is reminiscent of that for the ideal case and emphasizes the analogy given above.

Finally, the representation (17) is contained in the class of integral transforms studied in the Appendix of I. Consequently, (1) $n(E, T)/R(E)$ is an analytic function of the complex variable $E$ for Re $(E) > 0$ and it is completely monotonic for $E > 0$ [see (I32)]. (2) The spectral function $\sigma(T, T')$ is uniquely determined by the distribution $n(E, T)$.

[6] Integrals containing the spectral function are understood in the Lebesgue sense and (25) is said to hold almost everywhere in the open set $(0 \leq T' < T_0)$.

## IV. DILUTE BOSE GAS

Some of the features established above for systems of interacting particles can be illustrated by an exactly soluble model.[7] The model represents a weakly interacting Bose gas, but strongly degenerate. Summarizing the results of the model for temperatures *below* the condensation temperature $T_0$, the energy spectrum of the elementary excitation is given by

$$\epsilon(p) = [E^2 + 2E\alpha]^{\frac{1}{2}}, \tag{30}$$

with $E = (2m)^{-1}p^2$ and

$$\alpha = [(4\pi a\hbar^2)/m](N/V), \tag{31}$$

where $a$ is the $S$-wave scattering amplitude. The momentum distribution of the excitations is given by the noninteracting Bose distribution

$$\bar{n}_p = \{e^{[\epsilon(p)/kT]} - 1\}^{-1} \tag{32}$$

And, finally, the momentum distribution of the actual Bose particles is given by

$$\bar{N}_p = [\bar{n}_p + A_p^2(\bar{n}_p + 1)]/(1 - A_p^2), \tag{33}$$

where

$$A_p = \alpha^{-1}[\epsilon(p) - E - \alpha]. \tag{34}$$

Ideally, one would like to express (33) by the representation (17) with a positive-definite spectral function which satisfies (19). However, in the absence of such a result, some of the implications of the representation (17) will be verified by the distribution (33).

From a knowledge of the distribution function one certainly cannot verify the strong convergence[8] required in (1). Nevertheless, strong convergence implies weak convergence. This is satisfied by (33) since

$$\lim_{\alpha \to 0} \bar{N}_p = (e^{E/kT} - 1)^{-1}. \tag{35}$$

The distribution (33) is analytic for Re $(E) > 0$. The behavior near the point $E = 0$ is given by

$$\bar{N}_p = \frac{\frac{1}{2}kT}{E} - \frac{kT}{(2\alpha E)^{\frac{1}{2}}} + p(E) \quad (T > 0), \tag{36}$$

where $p(E)$ is regular at $E = 0$. Note the simple pole at $E = 0$. The existence and order of this singularity agrees with the analytical properties of the integral representation (17) (see Appendix). Also, the next leading singularity is of the *same* form as the singularity of the second term in (20) (see Appendix). Finally, for large energies,

$$\bar{N}_p \approx \alpha^2/4E^2 \quad (E \text{ large}), \tag{37}$$

which agrees with a possible behavior already discussed in I.

[7] A. A. Abrikosov, L. P. Gorkov, and I. E. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, Inc., Englewood Cliffs, N.J., 1963), Chap. 1.

[8] The notions of weak and strong convergence are those encountered in functional analysis (see Ref. 5 in I).

The introduction of quasi-particles (elementary excitations) is a feature of the method used in obtaining (33). However, the integral representation (17) cannot discern the different effects of the interaction.

## V. CONCLUSION

The fundamental postulate in the present work and in I is that of strong convergence of interacting states. In I this assumption was made in relation to the asymptotic approach to equilibrium, whereas here it relates to the approach to noninteracting states. Strong convergence is sufficient for an approach to equilibrium in the weak sense, i.e., the approach of quantum-mechanical averages of (macroscopic) operators to the proper equilibrium values. This latter behavior is what may commonly be understood by the phrase "approach to equilibrium."

In axiomatic quantum field theory,[2] one has the strong convergence of states which, in the limit $t \to +\infty$, define asymptotic states of a *finite* number of freely moving particles. These states are constructed from the vacuum by the application of a *finite* number of time-dependent operators on the vacuum state. Further, it is clear from the nature of the operators that the states are interacting states. Consider the state constructed from the vacuum by acting on it with an *infinite* number of time-dependent operators. The work of Ref. 2 shows that asymptotically, for $t \to +\infty$, these states approach constant vectors. It is clear, if not obvious, that such states *cannot* lead (in the limit $t \to +\infty$) to an asymptotic state of an *infinite* number of freely moving particles. Herein lies the connection of axiomatic quantum field theory to the basic postulate in I and, hence, to quantum statistical mechanics. The content of the postulate in I lies in giving an answer to this question by stating that *any* interacting state of an *infinite* number of particles must approach an equilibrium state as $t \to +\infty$. This view is consistent with theorems of the nature of Poincaré's in that a system with a finite number of particles *cannot* approach equilibrium.

*Note added in proof:* The possibility of (at most) a simple pole at $E = 0$ in the distribution function supplements Bogoliubov's $1/k^2$ theorem. [H. Wagner, Z. Physik **195**, 273 (1966); P. C. Hohenberg, Phys. Rev. **158**, 383 (1967).] Bogoliubov's result is

$$n(\mathbf{p}) \geq \frac{N_0(T)}{N} \frac{m}{\beta} \frac{1}{p^2} \quad (p \to 0)$$

for temperature $\beta^{-1}$. The momentum distribution of particles $n(\mathbf{p})$ might have a singularity of the order $E^{-1}$ *at least* as $E \to 0$ (but less than $E^{-\frac{3}{2}}$). It is clear that Bogoliubov's theorem can be of use only for $0 < T < T_\lambda$.

Suppose $N_0(T) \neq 0$. From this one cannot conclude that $\alpha(T) \neq 0$. However, if one invokes the theorem then, for $0 < T < T_\lambda$, one must have $\alpha(T) \neq 0$. Consequently, the result of the present formulation requires the singularity to be *precisely* $E^{-1}$ as $E \to 0$ for $0 < T < T_\lambda$. One must then view the $1/k^2$ theorem of Bogoliubov as a condition on the *residue* of the pole and not on the degree of the singularity.

## APPENDIX

This Appendix is concerned with the singularity structure of expression (20) in the neighborhood of $E = 0$.

From the positiveness of the spectral function

$$\int_0^{T_0} \frac{\sigma(T, T')\, dT'}{e^{E/kT'} - 1} \leq \frac{\alpha(T)}{E}, \tag{A1}$$

where $\alpha(T)$ is defined by (21) and is bounded for finite temperatures [see (29)]. In (A1) the equality holds as $E$ approaches zero. Therefore, the first term in (20) has a simple pole at $E = 0$.

Similarly, for the second term in (20),

$$\int_{T_0}^{\infty} \frac{\sigma(T, T')\, dT'}{e^{\mu(N,T')} e^{E/kT'} - 1} \leq \int_{T_0}^{\infty} \frac{\sigma(T, T')\, dT'}{\mu(N, T') + E/kT'}. \tag{A2}$$

As $E \to 0$ the region of integration which is important for the value of the integral is that near the condensation temperature $T_0$ [recall (15)]. The bound in (A2) will be determined for the case of nonrelativistic energy–momentum relationship. For such cases, the behavior of the chemical potential $\mu(N, T)$ near $T_0$ is given by

$$\mu(N, T) = (\beta/T^3)[1 - (T/T_0)^{\frac{3}{2}}]^2 \quad (T \to T_0), \tag{A3}$$

where

$$\beta = \frac{2\pi^2 \hbar^6}{m^3 k^3} \left(\frac{N}{V}\right)^2. \tag{A4}$$

On substituting (A3) into (A2) one obtains

$$\int_{T_0}^{\infty} \frac{\sigma(T, T')\, dT'}{e^{\mu(N,T')} e^{E/kT'} - 1} \leq \frac{\pi \sigma(T, T_0) T_0^2}{3(k\beta E)^{\frac{1}{2}}}. \tag{A5}$$

As before, the case of the equality holds in the limit $E \to 0$. Therefore, one obtains a *weaker* approach to infinity from the second term in (20) as $E \to 0$. It is believed that such behavior is preserved in the more general case of relativistic kinematics.

The sign of the next leading singularity need not be determined by (A5). Since from (A1),

$$\int_0^{T_0} \frac{\sigma(T, T')\, dT'}{e^{E/kT'} - 1} - \frac{\alpha(T)}{E} \leq 0. \tag{A6}$$

Therefore, the square-root singularity of (A5) may appear with a *negative* sign (in the distribution function) if a simple pole exists at $E = 0$.

# Some Properties of Ladder Operators*

S. KOMY AND L. O'RAIFEARTAIGH†
*Physics Department, Syracuse University, Syracuse, New York*

(Received 6 September 1967)

The commutation relations $[H, P \pm iQ] = \pm(P \pm iQ)$ for the symmetric operators $H$, $P$, $Q$ are considered without assuming (*a priori*) any other relationship between $H$, $P$, $Q$, in particular without making any assumption concerning the commutator $[P, Q]$. It is shown that under certain mild restrictions the spectrum of $H$ is integer spaced, and that in the two particular cases $[P, Q] = i\epsilon H$, $\epsilon = 0, \pm 1$ and $H = \frac{1}{2}(P^2 + Q^2) + i\sigma[P, Q]$, corresponding to Lie groups and parastatistics, respectively, it is simple. For these two cases the explicit representations of $P$, $Q$, and $H$ are found in a simple manner. The question of the existence of a common analytic domain for $P$, $Q$, and $H$ is investigated, and some sufficient conditions for this are found.

## 1. INTRODUCTION

We consider in this paper a set of closed symmetric operators $P$, $Q$, and $H$ in a Hilbert space $\mathcal{H}$, satisfying the commutation relations

$$[H, P \pm iQ] = \pm(P \pm iQ), \tag{1.1}$$

on a common invariant dense domain $D \subset \mathcal{H}$. $P$, $Q$, and $H$ are defined to be the closures of $P/D$, $Q/D$, and $H/D$ respectively. Equations (1.1) are the well-known relations first introduced by Cartan[1] in the study of compact simple Lie groups (for which one has in addition the relation $[P, Q] = iH$) and later introduced by Heisenberg[2] in his description of the quantum-mechanical harmonic oscillator (for which $H = \frac{1}{2}(P^2 + Q^2)$ and $[P, Q] = i$). We shall call the operators $P \pm iQ$ *ladder* operators, because if $|\lambda\rangle$ is an eigenvector of $H$ on which they are defined, and such that $P|\lambda\rangle$ and $Q|\lambda\rangle$ are in the domain $\Delta(H)$ of $H$,

$$(P \pm iQ)|\lambda\rangle = \text{const}\,|\lambda \pm 1\rangle. \tag{1.2}$$

What we wish to study here is the more general case in which the Eqs. (1.1) are not supplemented (at least *a priori*) by any other relationship between $H$, $P$, and $Q$. Further, we wish to study this case in a fairly rigorous way and obtain in a unified way some results which have been obtained previously for some special cases, in particular the case of Lie groups (with supplementary condition $[P, Q] = i\epsilon H$, $\epsilon = 0, \pm 1$) and parastatistics[3] (with supplementary condition $H = \frac{1}{2}(P^2 + Q^2) + i\sigma[P, Q]$). The latter case includes, in particular, the generalization of the quantum-mechanical commutation relations $[P, Q] = i$ suggested by Wigner[4] in connection with the Heisenberg harmonic oscillator equations.

We first show that if $\mathcal{H}$ is irreducible in the sense that any bounded operator which commutes with $P$, $Q$, and $H$ is a multiple of the identity, and if $D$ is analytic for $H$, then the spectrum of $H$ is integer spaced.

We then consider the two special cases,

$$[P, Q] = i\epsilon H, \quad \epsilon = 0, \pm 1$$

and

$$H = \tfrac{1}{2}(P^2 + Q^2) + i\sigma[P, Q],$$

and show for both cases that in addition to being integer-spaced, the spectrum of $H$ is simple.

Using this result, we find the explicit representations of $H$, $P$, and $Q$ for these two cases. In particular, for the case $[P, Q] = -iH$, which is the case of the Lie group $SU(1, 1)$, we impose the condition that $P^2 + Q^2 + H^2$ be essentially self-adjoint and rederive the five classes of unitary representations of this group rigorously obtained by Bargman,[5] while for the parastatistic case $H = \frac{1}{2}(P^2 + Q^2) + i\sigma[P, Q]$ we rederive the representations of $H$, $P$, and $Q$ which were obtained earlier in a less rigorous manner.[6,7]

Finally, we return to the general case [Eq. (1) with no subsidiary conditions] and discuss the question of the analyticity of the domain $D$ for $H$, $P$, and $Q$. We show that a sufficient condition for $D$ to be analytic for $H$, $P$, and $Q$ is that it be analytic for $H$ and that $H$ dominate $P$ and $Q$ on $D$ (i.e., that $\|Zd\| < \|Hd\| + c\|d\|$, where $Z = P$ or $Q$, $d \epsilon D$ and $c$ is a positive

* Work supported in part by the U.S. Atomic Energy Commission.

† On leave of absence from Dublin Institute for Advanced Studies, Dublin, Ireland.

[1] E. Cartan, thèse, Paris, 1894; *Oeuvres Complètes* (Gauthier-Villars, Paris, 1952), Vol. I.

[2] W. Heisenberg, Z. Physik 33, 879 (1925).

[3] For a comprehensive bibliography on parastatistics, see O. W. Greenberg and A. Messiah, Phys. Rev. 138, B1155 (1965). The free parafields are of the form $\phi(x) = \int [d^3k/(2\omega)^{\frac{1}{2}}][a_k e^{ikx} + a_k^* e^{-ikx}]$, $a_k$, $a_k^* = (P_k \pm iQ_k)$, where the $P_k$ and $Q_k$ are variables satisfying relations of the form (1.1) for each value of $k$, and the relations $[P_k, P_{k'}] = [P_k, Q_{k'}] = [Q_k, Q_{k'}] = 0$ for $k \neq k'$.

[4] E. P. Wigner, Phys. Rev. 77, 711 (1950).

[5] V. Bargmann, Ann. Math. 48, 568 (1947).

[6] T. Jordan, N. Mukunda, and S. Pepper, J. Math. Phys. 4, 1089 (1963).

[7] L. O'Raifeartaigh and C. Ryan, Proc. Roy. Irish Acad. 62A, 93 (1963).

number independent of $d$) and discuss some consequences of this result.

## 2. INTEGER-SPACED SPECTRUM

We consider the operators $H, P$, and $Q$ satisfying the commutation relations (1) and, in this section, make in addition the following two assumptions:

(i) $D$ is an analytic domain for $H$.

(ii) $\mathcal{H}$ is irreducible, in the sense that any bounded operator that commutes with $H, P$, and $Q$ is a multiple of the unit operator.

It is not assumed that $D$ is an analytic domain for $P$ and $Q$ or even that $P$ and $Q$ are self-adjoint. The question of the analyticity of $D$ with respect to $P$ and $Q$ will be discussed in Sec. 6. Using the assumptions (i) and (ii) we establish[8] the following result.

*Lemma 1:* The spectrum of $H$ is integer-spaced. Further, if $E(\lambda)$ denote the projection operators on the eigenspaces $\mathcal{H}_\lambda$ of $\mathcal{H}$, then

$$(P \pm iQ)E(\lambda)D = E(\lambda \pm 1)(P \pm iQ)D.$$

Finally, the closures of the restrictions of $P$ and $Q$ to $D(\lambda) = \bigcup_\lambda E(\lambda)D$ are equal to $P$ and $Q$, respectively.

*Proof:* From Eq. (1.1) we have

$$[H, P \pm iQ] = \pm(P \pm iQ) \quad \text{on} \quad D, \tag{2.1}$$

whence

$$(P \pm iQ)H = (H \mp 1)(P \pm iQ) \quad \text{on} \quad D, \tag{2.2}$$

and

$$(P \pm iQ)H^n = (H \mp 1)^n(P \pm iQ) \quad \text{on} \quad D. \tag{2.3}$$

Hence

$$(P \pm iQ)\sum_1^N \frac{(it)^n}{n!} H^n = \sum_1^N \frac{(it)^n}{n!}(H \mp 1)^n(P \pm iQ) \quad \text{on} \quad D. \tag{2.4}$$

Thus,

$$P\sum_1^N \frac{(it)^n}{n!} H^n = \frac{1}{2}\sum_1^N \frac{(it)^n}{n!}(H-1)^n(P+iQ) + \frac{1}{2}\sum_1^N \frac{(it)^n}{n!}(H+1)^n(P-iQ) \quad \text{on} \quad D. \tag{2.5}$$

Similar relations hold for $Q$. Since $D$ is invariant with respect to $P$ and $Q$ and analytic for $H$ (and hence for $H \pm 1$), it follows that the limits $N \to \infty$ on the right-hand side of (2.5) exist for sufficiently small $t$, $|t| < t_0$ say. Since $P$ and $Q$ are closed operators, we obtain on taking the limit

$$Pe^{itH} = \tfrac{1}{2}e^{it(H-1)}(P+iQ) + \tfrac{1}{2}e^{it(H+1)}(P-iQ) \quad \text{on} \quad D, \tag{2.6}$$

and

$$iQe^{itH} = \tfrac{1}{2}e^{it(H-1)}(P+iQ) - \tfrac{1}{2}e^{it(H+1)}(P-iQ) \quad \text{on} \quad D, \tag{2.7}$$

whence,

$$(P \pm iQ)e^{itH} = e^{it(H\mp1)}(P \pm iQ) \quad \text{on} \quad D, \tag{2.8}$$

for $|t| < t_0$.

Repetition of the argument shows that (2.8) holds for all $|t| < nt_0$, $n$ integer, and hence for all real $t$.

In particular,

$$(P \pm iQ)e^{2\pi iH} = e^{2\pi iH}(P \pm iQ) \quad \text{on} \quad D. \tag{2.9}$$

Thus $e^{2\pi iH}$ commutes with $P$ and $Q$ on $D$. Now let $f$ be any vector in the domain of $P$. Since $P$ is the closure of $P/D$ there exists a sequence $d_n \epsilon D$ such that $d_n \to f$, $Pd_n \to Pf$.

Hence, since $e^{2\pi iH}$ is bounded we have

$$e^{2\pi iH}d_n \to e^{2\pi iH}f$$

and

$$Pe^{2\pi iH}d_n = e^{2\pi iH}Pd_n \to e^{2\pi iH}Pf. \tag{2.10}$$

Thus $e^{2\pi iH}f$ is in the domain of $P$ and

$$Pe^{2\pi iH}f = e^{2\pi iH}Pf; \tag{2.11}$$

i.e., $e^{2\pi iH}$ commutes with $P$. Similarly it commutes with $Q$. Thus, since it obviously commutes with $H$, it commutes with $P$, $Q$, and $H$. From the irreducibility of $H$ as defined in (ii) above, it follows that $e^{2\pi iH}$ is a multiple of the identity, i.e.,

$$e^{2\pi iH} = e^{2\pi i\alpha}, \quad 0 \le \alpha < 1. \tag{2.12}$$

Thus the spectrum of $H$ is integer-spaced.

Now let $\lambda$ be any eigenvalue of $H$, $E(\lambda)$ the projection operator on the corresponding eigenspace, let $0 < \delta < 1$, and consider the Riemann sums

$$R_\Delta(\lambda) = \sum_\alpha \Delta(t_\alpha)\frac{e^{i(\lambda+\delta)t_\alpha} - e^{i(\lambda-\delta)t_\alpha}}{it_\alpha}e^{iHt_\alpha}, \xrightarrow[\Delta\to0]{} \int dt\,\frac{e^{i(\lambda+\delta)t} - e^{i(\lambda-\delta)t}}{it}e^{iHt} = E(\lambda). \tag{2.13}$$

From (2.6) and (2.7) we have

$$PR_\Delta(\lambda) = \tfrac{1}{2}R_\Delta(\lambda-1)(P+iQ) + \tfrac{1}{2}R_\Delta(\lambda+1)(P-iQ) \quad \text{on} \quad D, \tag{2.14}$$

$$iQR_\Delta(\lambda) = \tfrac{1}{2}R_\Delta(\lambda-1)(P+iQ) - \tfrac{1}{2}R_\Delta(\lambda+1)(P-iQ) \quad \text{on} \quad D. \tag{2.15}$$

[8] The basic idea of this section, namely showing that $\exp(2\pi iH)$ is a multiple of the identity, is due to T. Jordan *et al.* (Ref. 6).

Since $P$ and $Q$ are closed we obtain, on taking the limit $\Delta \to 0$,

$$PE(\lambda) = \tfrac{1}{2}E(\lambda - 1)(P + iQ) + \tfrac{1}{2}E(\lambda + 1)(P - iQ) \quad \text{on} \quad D, \tag{2.16}$$

$$iQE(\lambda) = \tfrac{1}{2}E(\lambda - 1)(P + iQ) - \tfrac{1}{2}E(\lambda + 1)(P - iQ) \quad \text{on} \quad D, \tag{2.17}$$

whence,

$$(P \pm iQ)E(\lambda) = E(\lambda \mp 1)(P \pm iQ) \quad \text{on} \quad D. \tag{2.18}$$

Finally, let $f$ be any vector in $\Delta(P)$. Then there exists a sequence $d_n \in D$ such that $d_n \to f$, $Pd_n \to Pf$.

Choose any fixed $E(\lambda)$ and construct the sequence of projection operators

$$K_0 = E(\lambda), \quad K_1 = E(\lambda) + E(\lambda + 1) + E(\lambda - 1),$$
$$K_2 = E(\lambda) + E(\lambda + 1) + E(\lambda - 1) + E(\lambda + 2) + E(\lambda - 2), \cdots.$$

Clearly $K_m \to 1$ as $m \to \infty$. Hence if we construct the sequence $K_m d_n$ we have

$$K_m d_n \to d_n \to f \tag{2.19}$$

and

$$PK_m d_n = K_{m+1} P d_n \to Pd_n \to Pf. \tag{2.20}$$

A similar relation holds for $Q$. Since each $K_m d_n \in \Delta[P/D(\lambda)]$ this establishes that $P$ is the closure of $P/D(\lambda)$. Similarly, $Q$ is the closure of $Q/D(\lambda)$. This completes the proof of Lemma 1.

## 3. SIMPLE SPECTRUM

In this section, we establish conditions under which the spectrum of $H$ is simple, i.e., to each eigenvalue there corresponds only one eigenvector. First we establish the following general lemma.

*Lemma 2:* The spectrum of $H$ is simple if and only if there exists a $\psi(\lambda_0) \in D(\lambda)$ such that for all $\psi_n = (P \pm iQ)^n \psi(\lambda_0)$, $n = 0, 1, 2, \cdots$,

$$(P - iQ)(P + iQ)\psi_n = c_n \psi_n, \tag{3.1}$$

$$(P + iQ)(P - iQ)\psi_n = d_n \psi_n, \tag{3.2}$$

where $c_n$ and $d_n$ are numerical coefficients.

*Proof:* The necessity of the condition is obvious. To prove the sufficiency we let $\mathcal{H}_0$ denote the closed linear span of the $\psi_n$ and let $\Lambda$ denote the projection operator on $\mathcal{H}_0$. Then for any $d \in D$, $\Lambda E_n d$ is a multiple of the vector $\psi_n$. $[E_n \equiv E[\lambda_0 + n)]$. Hence $\Lambda E_n d \in \Delta(P + iQ)$ and

$$(P \pm iQ)\Lambda E_n d = (P \pm iQ)k\psi_n = k\psi_{n+1}. \tag{3.3}$$

On the other hand,

$$\Lambda(P \pm iQ)E_n d = \Lambda E_{n\mp 1}(P \pm iQ)d = k'\psi_{n+1}. \tag{3.4}$$

Further,

$$\begin{aligned} k &= (\psi_{n+1}, (P \pm iQ)\Lambda E_n d) \\ &= ((P \mp iQ)\psi_{n+1}, \Lambda E_n d) \\ &= ((P \mp iQ)\psi_{n+1}, E_n d) \\ &= (\psi_{n+1}, (P \pm iQ)E_n d) \\ &= (\psi_{n+1}, \Lambda(P \pm iQ)E_n d) = k'. \end{aligned} \tag{3.5}$$

Therefore for all $E_n d$ we have

$$\Lambda E_n d \in \Delta(P \pm iQ)$$

and

$$(P \pm iQ)\Lambda E_n d = \Lambda(P \pm iQ)E_n d. \tag{3.6}$$

Thus $\Lambda$ commutes with $P/D(\lambda)$ and $Q/D(\lambda)$.

Now let $f$ be any vector in $\Delta(P)$. Since $P$ is the closure of $P/D(\lambda)$, there exists a sequence $f_k \in D(\lambda)$ such that $f_k \to f$ and $Pf_n \to Pf$. Hence $\Lambda f_n \to \Lambda f$ and $P\Lambda f_n = \Lambda P f_n \to \Lambda P f$. Since $P$ is closed it follows that $\Lambda f \in \Delta(P)$ and $P\Lambda f = \Lambda P f$. Thus $\Lambda$ commutes with $P$. Similarly $\Lambda$ commutes with $Q$. Since $\Lambda$ obviously commutes with $H$ it follows that $\Lambda$ commutes with $P$, $Q$, and $H$ and hence is a multiple of the identity. Since $\Lambda^2 = \Lambda$ (projection operator) it follows that $\Lambda = 1$ ($\Lambda = 0$ is excluded since $\psi(\lambda_0) \neq 0$). Thus $\mathcal{H}_0 = \mathcal{H}$, and the spectrum of $H$ is simple.

We now apply this lemma to two cases of physical interest.

*Lemma 3:* The conditions of Lemma 2 for the simplicity of the spectrum of $H$ are met in the following cases:

(a) $[P, Q] = i\epsilon H$, $\epsilon = 0, \pm 1$, $P^2 + Q^2 + H^2$ essentially self-adjoint on $D$,

(b) $H = \tfrac{1}{2}(P^2 + Q^2) + i\sigma[P, Q] \geq 0$ on $D$.

*Proof:* Case (a). According to a theorem of Nelson,[9] under conditions (a) the operators $P$, $Q$, and $H$ generate a unitary irreducible representation of a Lie group [$SU(2)$, $E_2$, and $SU(1, 1)$ for $\epsilon = 0, \pm 1$, respectively]. The operator $C = P^2 + Q^2 + \epsilon H^2$ is a Casimir operator of this group. Since $H$ is irreducible it follows that $C$ is a real $C$ number, $C$ say. We then have

$$\begin{aligned} &(P \pm iQ)(P \mp iQ)\psi_n \\ &\quad = \{P^2 + Q^2 \mp i[P, Q]\}\psi_n \\ &\quad = \{(P^2 + Q^2 + \epsilon H^2) - (\epsilon H^2 \mp \epsilon H)\}\psi_n \\ &\quad = \{C - \epsilon\lambda(\lambda \mp 1)\}\psi_n. \end{aligned} \tag{3.7}$$

Thus the conditions of Lemma 2 are satisfied and the spectrum of $H$ is simple. In addition we obtain for

[9] E. Nelson, Ann. Math. 70, 572 (1959).

the coefficients $c_n$, $d_n$ the relation

$$c_n = d_{n+1} = C - \epsilon\lambda(\lambda + 1), \quad \lambda = \lambda_0 + n. \quad (3.8)$$

Case (b). We can write $H$ in the following form:

$$H = \tfrac{1}{4}(1 - 2\sigma)(P + iQ)(P - iQ) + \tfrac{1}{4}(1 + 2\sigma)(P - iQ)(P + iQ). \quad (3.9)$$

Hence (except for the special case[10] $2\sigma + 1 = 0$, to be discussed below) we have

$$(P - iQ)(P + iQ) = [(2\sigma - 1)/(2\sigma + 1)] \times (P + iQ)(P - iQ) + [4H/(2\sigma + 1)]. \quad (3.10)$$

Since $H \geq 0$, then there exists a least eigenvalue $h_0$ of $H$. The corresponding eigenvector $\psi_0$ is such that

$$(P - iQ)\psi_0 = 0, \quad (3.11)$$

from which it follows that

$$(P + iQ)(P - iQ)\psi_0 = 0 \quad (3.12)$$

and

$$(P - iQ)(P + iQ)\psi_0 = [(4h_0)/(2\sigma + 1)]\psi_0. \quad (3.13)$$

In deriving Eq. (3.13), Eqs. (3.10) and (3.12) are used.

Suppose that Eqs. (3.1) and (3.2) are satisfied for $n = 0, 1, 2 \cdots s - 1$, then

$$(P + iQ)(P - iQ)\psi_s = (P + iQ)(P - iQ)(P + iQ)\psi_{s-1} = (P + iQ)c_{s-1}\psi_{s-1} = c_{s-1}\psi_s \quad (3.14)$$

and

$$(P - iQ)(P + iQ)\psi_s = [(2\sigma - 1)/(2\sigma + 1)]c_{s-1}\psi_s + [4(h_0 + s)/(2\sigma + 1)]\psi_s. \quad (3.15)$$

Thus Eqs. (3.1) and (3.2) are satisfied also for $n = s$. Thus, once again, the conditions of Lemma 2 are met and the spectrum of $H$ is simple.

Further, we obtain the recurrence relations

$$c_n = d_{n+1} = [(2\sigma - 1)/(2\sigma + 1)]C_{n-1} + [4(h_0 + n)/(2\sigma + 1)] \quad (3.16)$$

for the coefficients $c_n$ and $d_n$.

Finally we consider the special case $2\sigma + 1 = 0$. In this case we have

$$H = \tfrac{1}{2}(P + iQ)(P - iQ). \quad (3.17)$$

We introduce the operator

$$S = [2(H - 1)]^{-1}(P - iQ)(P + iQ), \quad (3.18)$$

which is a well-defined operator on $D$ and $D(\lambda)$. On $E(\lambda)D$ we have

$$\begin{aligned} S^2 &= [2(\lambda + 1)]^{-1}(P - iQ)(P + iQ)[2(\lambda + 1)]^{-1} \times (P - iQ)(P + iQ) \\ &= [2(\lambda + 1)^{-1}(P - iQ)[2H/2(\lambda + 1)](P + iQ) \\ &= [2(\lambda + 1)]^{-1}(P - iQ)(P + iQ) \\ &= S. \end{aligned} \quad (3.19)$$

Since $D(\lambda)$ is dense in $H$, it follows that the closure $\bar{S}$ of $S$ is a projection operator. There are then two cases to consider.

*Case 1:* $\bar{S} = I$ (where $I$ is the unit operator). Then from (3.20) we have

$$(P - iQ)(P + iQ) = 2(H + 1) \quad \text{on} \quad D. \quad (3.20)$$

Combining (3.19) and (3.22) we obtain

$$[Q, P] = i \quad \text{and} \quad H = \tfrac{1}{2}(P^2 + Q^2 + 1) \quad \text{on} \quad D. \quad (3.21)$$

But $H$ is essentially self-adjoint on $D$. Hence this is exactly the case of the harmonic oscillator in ordinary quantum mechanics.

*Case 2:* There exists at least one vector $h \in \mathcal{H}$ such that $\bar{S}h = 0$. Then since

$$[\bar{S}, H] = 0 \quad \text{on} \quad D(\lambda),$$

there exists at least one eigenspace $E(\lambda_0)\mathcal{H}$ such that $\bar{S} \neq 1$ on $E(\lambda_0)\mathcal{H}$. From this it follows that there exists at least one vector $\psi(\lambda_0) \in E(\lambda_0)D$ such that

$$S\psi(\lambda_0) = 0.$$

But then from the definition of $S$ we have

$$(P + iQ)\psi(\lambda_0) = 0. \quad (3.22)$$

Let us now construct the sequence

$$\psi_{-m} = (P - iQ)^m\psi(\lambda_0), \quad m = 0, 1, 2, \cdots. \quad (3.23)$$

Since

$$(P + iQ)(P - iQ)\psi_{-m} = 2H\psi_{-m} = 2(\lambda_0 - m)\psi_{-m} \quad (3.24)$$

and

$$\begin{aligned} (P - iQ)(P + iQ)\psi_{-m} &= (P - iQ)(P + iQ)(P - iQ)\psi_{-m+1} \\ &= (P - iQ)2(\lambda_0 - m + 1)\psi_{-m+1} \\ &= 2(\lambda_0 - m + 1)\psi_{-m}, \end{aligned} \quad (3.25)$$

[10] The special case $2\sigma + 1 = 0$ is actually the most important case since, as will be seen later, it includes the case of the quantum-mechanical harmonic oscillator and of parastatistics. Also the case $2\sigma + 1 \neq 0$ can be reduced to the case $2\sigma + 1 = 0$ in a certain sense, namely that since the spectrum of $H$ is integer-spaced and simple, it is possible to introduce new variables $P$ and $Q$ such that $[H, P' + iQ'] = \pm(P' \pm iQ')$ and $H = \frac{1}{2}(P'^2 + Q'^2) - \frac{1}{2}[P', Q']$. The question as to whether the original $(P, Q)$ or the new $(P', Q')$ are the "physical" $(P, Q)$ is left open. For details see D. Boulware and S. Deser, Nuovo Cimento **30**, 230 (1963), where the (non-canonical) transformation from $(P, Q)$ to $(P', Q')$ is given explicitly, and B. Gruber and L. O'Raifeartaigh, Proc. Roy. Irish Acad. **63A**, 69 (1964).

it is clear that this sequence satisfies the conditions of Lemma 2. Thus the spectrum of $H$ is simple.

In addition, since $H \geq 0$, the sequence (3.25) must terminate. Further, it is clear from (3.25) and (3.14) that $H = 0$ on the lowest state. Thus we obtain a finite representation for the operators $P$, $Q$, and $H$ and the spectrum of $H$ is $0, 1, 2, \cdots, n$. This is the case of parafermi statistics.

## 4. EXPLICIT REPRESENTATIONS FOR THE CASE $[P, Q] = i\epsilon H$

Since in the special cases (a) and (b) considered in the last section the spectrum of $H$ is simple, the representation of $P$, $Q$, and $H$ is completely determined by the eigenvalues of $H$ and the coefficients $c_n$ and $d_n$ of Eqs. (3.8) and (3.18). In fact the only non-vanishing matrix elements are

$$\begin{aligned} (\psi_n, H\psi_n) &= \lambda_0 + n, \\ (\psi_{n+1}, P + iQ\psi_n) &= a_n, \qquad |a_n|^2 = c_n, \\ (\psi_{n-1}, P - iQ\psi_n) &= b_n, \qquad |b_n|^2 = d_n. \end{aligned} \tag{4.1}$$

Hence in order to obtain the explicit representations in these cases it is sufficient to determine the range of $n$ and the coefficients $c_n$ and $d_n$. But since $H$, $P$, and $Q$ are symmetric the necessary and sufficient condition for the existence of a representation is

$$\lambda_0 \text{ real}, \quad a_n = b_{n+1}^*. \tag{4.2}$$

The second relation sharpens the relation $c_n = d_{n+1}$ of (3.8) and (3.18). The phases of the $a_n$ and $b_n$ are otherwise undetermined, but may be absorbed into the phases of the vectors $\psi_n$.

In the present section we determine the explicit representation in the case (a), i.e., when $[P, Q] = i\epsilon H$ and $P^2 + Q^2 + H^2$ is essentially self-adjoint on $D$. In this case, as mentioned earlier, the $P$, $Q$, and $H$ generate a unitary representation of a Lie group on $H$.

From (3.8) we have in this case

$$c_n = C - \epsilon\lambda(\lambda + 1), \quad \lambda = \lambda_0 + n. \tag{4.3}$$

We discuss the special cases $\epsilon = 0, \pm 1$ separately.

### (i) $\epsilon = 0$. Group $E_2$

In this case we obtain the representation (4.1) with $d_n = c_n = C \geq 0$. For $C = 0$ the representation is trivial with $P = Q = 0$ and $H =$ any real constant $\lambda_0$. For $C > 0$ the spectrum of $H$ is unbounded above and below. If we demand that $H$ generate a single-valued representation of $0(2) \subset E_2$ we obtain the additional restriction, $\lambda_0 =$ integer.

### (ii) $\epsilon = 1$. Group $SU(2)$

In this case we obtain (4.1) with $c_n = d_{n+1} = C - \lambda(\lambda + 1)$. The special feature of this case is that the positive operator $P^2 + Q^2 + H^2$ is a constant $(=C)$. From this it follows that $P$, $Q$, and $H$ are bounded operators. Thus the spectrum of $H$ is bounded and there must exist two values $j$ and $k$, say of $\lambda$, such that

$$\lambda(\lambda + 1) = C \quad \text{for} \quad \lambda = j, k, \quad j \geq k$$
$$j - k = \text{integer}. \tag{4.4}$$

From (4.4) we obtain immediately the well-known result that for $SU(2)$ the spectrum of $H$ is of the form $-j, -j + 1, \cdots j - 1, j$, $2j$ integer, and that the nonzero elements of $P$ and $Q$ are given by (4.1) with

$$\begin{aligned} c_n &= j(j + 1) - \lambda(\lambda + 1), \\ d_n &= j(j + 1) - \lambda(\lambda - 1). \end{aligned} \tag{4.5}$$

Note that in this case we automatically obtain the restriction $2j =$ integer. This is because $SU(2)$ is its own covering group.

### (iii) $\epsilon = -1$. Group $SU(1, 1)$

The matrix elements of $P$, $Q$, and $H$ are given by (4.1) with $c_n = d_{n+1} = C + \lambda(\lambda + 1)$. The covering group of $SU(1, 1)$ is infinitely connected. Hence to obtain single- or double-valued representations we must impose from outside the condition $2\lambda =$ integer. The possible representation can then be divided into four classes:

(a) $\lambda =$ integer, $C > 0$,

(b) $\lambda = \frac{1}{2}$-odd-integer, $C > \frac{1}{4}$.

In these two classes $C + \lambda(\lambda + 1) \geq 0$ for all $\lambda$. Hence representations are possible for all $c$, and the spectrum of $H$ is unbounded above and below.

(c) $\lambda =$ integer, $C \leq 0$,

(d) $\lambda =$ half-odd-integer, $C \leq \frac{1}{4}$.

In these two cases $C + \lambda(\lambda + 1)$ becomes negative unless for some nonnegative integer and half-odd integer, respectively, $c = j(j + 1)$. Representations are possible for these and only these values of $c$, and the spectrum of $H$ terminates either below at $j$ or above at $-(j + 1)$.

One can easily check that these four classes of representations of $SU(1, 1)$ correspond to the five classes of representations found by Bargman.[4]

## 5. EXPLICIT REPRESENTATION IN THE CASE $H = \frac{1}{2}(P^2 + Q^2) + i\sigma[P, Q]$

We now consider the parastatistics case. The matrix elements of $P$, $Q$, and $H$ are given by Eq. (4.1), with $\lambda_0 + n \geq 0$. The possibilities are restricted only by the condition $\lambda_0$ real, $c_n \geq 0$ of Eq. (4.3).

For the general case $2\sigma + 1 \neq 0$ the $C_n$ are determined by the recurrence relations (3.18). Solving these equations we obtain

$$\tfrac{1}{2}c_n = \tfrac{1}{2}d_{n+1} = (h_0 - \sigma - \tfrac{1}{2}) \times \{1 - [(2\sigma - 1)/(2\sigma + 1)]^{n+1}\} + n + 1. \quad (5.1)$$

For $\sigma \geq 0$ and $\sigma < -\frac{1}{2}$ the condition $c_n > 0$ is automatically satisfied. Representations exist and the spectrum of $H$ is unbounded above. For $-\frac{1}{2} < \sigma < 0$ the situation is more complicated. It has been analyzed in detail in Ref. 7. The main results are that no representations exist for general value of $h_0$, but for particular values of $h_0$, both finite- and infinite-dimensional representations exist.

Finally, we consider the special case $2\sigma + 1 = 0$. When this case leads to the case of the ordinary quantum-mechanical harmonic oscillator, we obviously obtain the usual matrix elements for that case, namely $h_0 = \frac{1}{2}$, $c_n = d_{n+1} = 2(n + 1)$. For the other possibility, the parafermi case, it is easy to see from the discussion of Sec. 3, that we obtain the finite matrices

$$H = \begin{bmatrix} 0 & & & & & \\ & 1 & & & & \\ & & 2 & & & \\ & & & \cdot & & \\ & & & & \cdot & \\ & & & & & n \end{bmatrix},$$

$$P = \begin{bmatrix} 0 & (1)^{\frac{1}{2}} & & & & \\ (1)^{\frac{1}{2}} & 0 & (2)^{\frac{1}{2}} & & & \\ & (2)^{\frac{1}{2}} & 0 & (3)^{\frac{1}{2}} & & \\ & & (3)^{\frac{1}{2}} & \cdot & & (n-1)^{\frac{1}{2}} \\ & & & & \cdot & \\ & & & & (n-1)^{\frac{1}{2}} & 0 \end{bmatrix},$$

$$Q = i\begin{bmatrix} 0 & (1)^{\frac{1}{2}} & & & & \\ -(1)^{\frac{1}{2}} & 0 & (2)^{\frac{1}{2}} & & & \\ & -(2)^{\frac{1}{2}} & 0 & (3)^{\frac{1}{2}} & & \\ & & -(3)^{\frac{1}{2}} & \cdot & & (n-1)^{\frac{1}{2}} \\ & & & & \cdot & \\ & & & & -(n-1)^{\frac{1}{2}} & 0 \end{bmatrix}.$$

## 6. ANALYTICITY OF THE DOMAIN $D$

So far we have assumed only that $D$ is analytic with respect to $H$. We now consider the question of its analyticity with respect to $P$ and $Q$. A sufficient condition that it be analytic with respect to $P$ and $Q$ also is given by the following lemma.

*Lemma 4:* The domain $D$ is analytic for $P$ and $Q$ if for all $d \in D$ $\|Zd\| \leq \|Hd\| + c\|d\|$, where $Z$ means $P$ or $Q$ and $c$ is a constant independent of $d$.

*Proof:* We first introduce a set of numerical coefficients $C_n^r$ by the relation

$$\sum_{r=0}^{n} C_n^r x^r = \prod_{r=0}^{n} (x + c + r - 1). \quad (6.1)$$

It is easy to see that the $C_n^r$ satisfy the recurrence relation

$$C_{n+1}^r = C_n^{r-1} + (n + c)C_n^r, \quad (6.2)$$

where it is understood that

$$C_n^{-1} = C_n^{n+1} = 0, \quad (6.3)$$

and have the following upper bound:

$$C_n^r \leq \binom{n}{r}(n + c)^{n-r}. \quad (6.4)$$

We then define $Z^r$ to be $Z_1 Z_2 \cdots Z_r$ where $Z_s = P$ or $Q$ for all $s = 1, 2, \cdots r$, and establish the relation

$$\|Z^n d\| \leq \sum_{r=0}^{n} C_n^r \|H^r d\|. \quad (6.5)$$

The proof is by induction. Suppose (6.5) is true for $n = 0, 1, \cdots N$. Then,

$$\begin{aligned} \|Z^{N+1}d\| &\leq \|(H + c)Z^N d\| \\ &\leq \|Z^N(H + c)d\| + \|[H, Z^N]d\| \\ &\leq \sum_{r=0}^{N} C_N^r \|H^r(H + c)d\| + n\|Z^N d\|, \end{aligned}$$

where in obtaining the first term we have used (6.5) and in obtaining the second we have used Eq. (1.1). Using (6.5) again, we obtain

$$\begin{aligned} \|Z^{N+1}d\| &\leq \sum_{r=0}^{N} \{C_N^r \|H^{r+1}d\| + (n + c)C_N^r \|H^r d\|\} \\ &= \sum_{r=0}^{N+1} \{C_N^{r-1} + (n + c)C_N^r\} \|H^r d\| \\ &= \sum_{r=0}^{N+1} C_{N+1}^r \|H^r d\|. \end{aligned} \quad (6.6)$$

Thus (6.5) is true for $n = N + 1$, and so for all $n$. We then have

$$\begin{aligned} \sum_{1}^{N} \frac{|t|^n}{n!} \|Z^n d\| &\leq \sum_{1}^{N} \frac{|t|^n}{n!} \sum_{r=0}^{n} C_n^r \|H^r d\| \\ &\leq \sum_{r=0}^{N} \sum_{n=r}^{N} \frac{|t|^n}{n!} C_n^r \|H^r d\| \\ &\leq \sum_{r=0}^{N} \sum_{n=r}^{N} \frac{|t|^n}{r!} \frac{(n + c)^{n-r}}{(n - r)!} \|H^r d\|. \end{aligned} \quad (6.7)$$

The last expression is obtained by using (6.4). Hence,

$$\sum_1^N \frac{|t|^n}{n!}\|Z^n d\| \le \sum_{r=0}^N \sum_{m=0}^{N-r} \frac{|t|^{m+r}}{r!}\frac{(r+m+c)^m}{m!}\|H^r d\|$$
$$\le \sum_{r=0}^N \sum_{m=0}^{N-r} \frac{|t|^r}{r!}\frac{|t|^m m^m}{m!}\left(1+\frac{r+c}{m}\right)^m \|H^r d\|$$
$$\le \sum_{r=0}^N \sum_{m=0}^{N-r} \frac{|t|^r}{r!}\frac{|t|^m m^m}{m!} e^{r+c}\|H^r d\|$$
$$\le e^c \sum_{r=0}^N \frac{e^r|t|^r}{r!}\|H^r d\| \sum_{m=0}^{N-r} \frac{|t|^m m^m}{m!}$$
$$\le e^c \sum_{r=0}^N \frac{e^r|t|^r}{r!}\|H^r d\| \sum_{m=0}^{N-r} e^m |t|^m. \qquad (6.8)$$

But $D$ is an analytic domain for $H$. Hence $d$ is an analytic vector, with radius of convergence $t_0$, say. Hence the two series on the right-hand side of (6.8) converge for $et < t_1 = \min[t_0, 1]$. Thus $d$ is an analytic vector for $Z$ with radius of convergence $\geq t_1/e$. This establishes the lemma.[11] We obtain also the following simple corollary.

*Corollary:* Let $H = \frac{1}{2}(P^2 + Q^2)$ on $D$, and $D$ be an analytic domain for $H$. Then if $H$, $P$, and $Q$ satisfy (1.1) on $D$, $D$ is an analytic domain for $H$, $P$, and $Q$.

Note that in the case $H = \frac{1}{2}(P^2 + Q^2)$ the complicated expressions (5.1) for the matrix elements of $P$ and $Q$ reduce to $c_n = d_{n+1} = 2(h_0 + n) + 1$. For $h_0 = \frac{1}{2}$ they reduce further to the ordinary quantum-mechanical values $c_n = d_{n-1} = 2(n + 1)$. The statement of the above corollary can actually be weakened by the following.

*Lemma 5:* Let $H = \frac{1}{2}(P^2 + Q^2)$ be essentially self-adjoint, and (1.1) be satisfied, on $D$. Then there exists in $\mathcal{H}$ a common dense analytic domain $A$ for $P$, $Q$, and $H$.

*Proof:* We construct the analytic domain $A$ for $H$ (which certainly exists, since $H$ is self-adjoint) and show that it is an analytic domain for $P$ and $Q$. This will follow from the proof of Lemma 4 if we can show that for all $a \in A$,

$$\|Z^{r+1}a\| \le \sum_{q=0}^{r+1} C_{r+1}^q \|H^q a\|, \qquad (6.9)$$

where the $C_r^q$ are the coefficients defined in Lemma 4. To prove this we show that it is true jointly with

$$[H, Z^r]a = \tilde{Z}_r a, \qquad (6.10)$$

where $\tilde{Z}_r$ is the $r$th-degree monomial in $P$ and $Q$ obtained by computing the same commutator on $D$.

To prove (6.9) and (6.10) we assume that they are satisfied for $r = 1 \cdots s - 1$. By definition,

$$([H, Z^s]d - \tilde{Z}_s d, a) = 0 \qquad (6.11)$$

whence, using (6.9) for $r = s - 1$ we have

$$(Hd, Z^s a) = (d, Z^s Ha - \tilde{Z}_s a). \qquad (6.12)$$

Thus $Z^s a$ is in the domain of $(H/D)^* = H$ and

$$HZ^s a = Z^s Ha - \tilde{Z}_s a. \qquad (6.13)$$

This is Eq. (6.10) for $r = s$. Further, since

$$Z^s a \in \Delta(H), \quad Z^s a \in \Delta(Z),$$

and

$$\|Z^{s+1}a\| \le \|HZ^s a\| + c\|Z^s a\|.$$

Using this equation and Eqs. (6.9) and (6.10) for $r = 1 \cdots s - 1$, we obtain exactly as in Eq. (6.6),

$$\|Z^{s+1}a\| \le \sum_{r=0}^{s+1} c_{s+1}^r \|H^r d\|.$$

This is just Eq. (6.9) for $r = s$. Thus Eqs. (6.9) and (6.10) are valid for $r = s$, and so for all $r$, as required.

This lemma generalizes a result due to Dixmier,[12] which states that if $\frac{1}{2}(P^2 + Q^2)$ is essentially self-adjoint on $D$ and $[Q, P] = i$ on $D$, then $P$ and $Q$ are essentially self-adjoint on $D$.

[11] It should perhaps be mentioned that the proof just given is not a special case of the general proof of analyticity for elements of a Lie algebra given by Nelson (Ref. 9, pp. 577, 588). In both cases $\|Z\| \le \|H\|$, but whereas in Nelson's case one uses information about $(\text{ad } Z)^n H$, here one uses information about $(\text{ad } H)^n Z$.

[12] J. Dixmier, Comp. Math. **13**, 263 (1958).

# Quantum Corrections to the Pair Distribution Function of a Classical Plasma

MARK DIESENDORF AND BARRY W. NINHAM
*Department of Applied Mathematics, University of New South Wales, Kensington, N.S.W., Australia*

(Received 8 June 1967)

Quantum-mechanical corrections to the pair distribution function of a plasma at high temperature and low density are calculated to order $e^2$ in the interaction, using standard diagram perturbation techniques. Both the effects due to quantum statistics (exchange) and the finite size of a wave packet (dynamic screening), are considered.

## 1. INTRODUCTION

In this paper calculations are made of quantum-mechanical corrections to the pair distribution function (PDF) of a plasma at high temperature and low density. Standard quantum diagram perturbation techniques[1–3] are used to derive the corrections to first order in the interaction $e^2$.

The pair distribution function $g(r)$ gives the probability of finding a particle within the distance $(r, r + dr)$ of a given particle. It is of interest because it is directly related to experimental quantities; in particular, to the cross section for the elastic scattering of an "external" particle or photon from the system, within the Born approximation.[4] As an asymptotic expansion in terms of the plasma parameter $\Lambda$, the PDF for a classical electron gas (in a uniform positive background) has the form

$$\delta g(r) \equiv g(r) - 1 = \exp\left[-\frac{\Lambda}{\kappa r}e^{-\kappa r}\right] - 1 + O(\Lambda^2) + O(\Lambda^2 \ln \Lambda), \quad (1.1)$$

where $\beta = 1/kT$, $\rho$ is the density of electrons, and $\kappa^{-1} = \lambda_D = (4\pi\rho\beta e^2)^{-\frac{1}{2}}$, the classical Debye length, is a measure of the *static* screening distance in the plasma. The plasma parameter $\Lambda$ is given by

$$\Lambda = \frac{1}{4\pi\rho\lambda_D^3} = 2\pi^{\frac{1}{2}}e^3\beta^{\frac{3}{2}}\rho^{\frac{1}{2}}. \quad (1.2)$$

For large distances $r$, Eq. (1.1) gives as leading term the well-known linearized Debye–Hückel[5] result

$$\delta g(r) = -\Lambda(\kappa r)^{-1}e^{-\kappa r}. \quad (1.3)$$

Terms beyond the first in Eq. (1.1), of order $\Lambda^2$ and $\Lambda^2 \ln \Lambda$, correspond to the Abe[6] correction to the equation of state. They have been calculated by Bowers and Salpeter, Hirt, DeWitt, and others, using diagram methods[7]; also by O'Neil and Rostoker,[8] and Lie and Ichikawa[9] using a kinetic equation approach.[10]

Apart from these classical terms, there are also quantum corrections; i.e., those involving $\hbar$. Quantum effects of two kinds persist at high temperatures: Corrections of the first kind, exchange corrections, are a consequence of quantum statistics, which give distance correlations even in the absence of interactions. The zero-order exchange correction to the PDF of a Fermi or Bose gas was found by Uhlenbeck and Gropper[11] to be, at high temperatures,

$$\delta g_0(r) = \mp e^{-\frac{1}{2}(r/\lambda\!\!\!^{-})^2}, \quad (1.4)$$

where the upper and lower signs refer to systems of fermions or bosons, respectively, $\lambda\!\!\!^{-}$ is the thermal de Broglie wavelength

$$\lambda\!\!\!^{-} = \hbar(\beta/2m)^{\frac{1}{2}} \quad (1.5)$$

and $m$ is the particle mass. In the fermion case an extra factor $\frac{1}{2}$ must multiply Eq. (1.4) to take account of the exclusion principle. The relation (1.4) is a special case of the more general result of London[12] and Placzek[13] valid for a noninteracting quantum gas at all temperatures,

$$\delta g_0(r) = \mp\left\{\frac{1}{(2\pi\hbar)^3\rho}\int d^3p e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}\frac{ze^{-\beta p^2/2m}}{1 \pm ze^{-\beta p^2/2m}}\right\}^2, \quad (1.6)$$

where the fugacity $z$ is related to the density $\rho$ by

$$\rho = \frac{1}{(2\pi\hbar)^3}\int d^3p \frac{ze^{-\beta p^2/2m}}{1 \pm ze^{-\beta p^2/2m}}. \quad (1.7)$$

[1] E. W. Montroll and J. C. Ward, Phys. Fluids **1**, 55 (1958).
[2] E. W. Montroll in *The Theory of Neutral and Ionised Gases*, C. de Witt and J. F. de Toeuf, Eds. (John Wiley & Sons, Inc., New York, 1960).
[3] J. M. Luttinger and J. C. Ward, Phys. Rev. **118**, 1417 (1960).
[4] L. Van Hove, Phys. Rev. **95**, 249 (1954), and references.
[5] P. Debye and E. Hückel, Physik. Z. **24**, 185 (1923).
[6] R. Abe, Progr. Theoret. Phys. **22**, 213 (1959).
[7] D. L. Bowers and E. E. Salpeter, Phys. Rev. **119**, 1180 (1960); C. W. Hirt, Phys. Fluids **8**, 693 (1965); H. E. DeWitt, Phys. Rev. **140**, A466 (1965).
[8] T. O'Neil and N. Rostoker, Phys. Fluids **8**, 1109 (1965).
[9] T. J. Lie and Y. H. Ichikawa, Rev. Mod. Phys. **38**, 680 (1966).
[10] There is currently some debate in the literature about one of the terms in the expansion (1.1). For large distances $r$, the diagram theories find that this term tends to infinity as $\Lambda^2 e^{-\kappa r}$, i.e., more slowly than the leading Debye term. (See, e.g., Ref. 9 for a discussion.)
[11] G. E. Uhlenbeck and L. Gropper, Phys. Rev. **41**, 79 (1932).
[12] F. London, J. Chem. Phys. **11**, 203 (1943).
[13] G. Placzek, *Proc. Second Berkeley Symp. Math. Stat. Prob.*, 581 (1950).

This formula [Eq. (1.6)] expresses the zero-order PDF in terms of the Fourier transform of the Fermi or Bose distribution. The next contribution to the PDF due to exchange is of first order in $e^2$, and is evaluated explicitly in Sec. 2. The result is

$$\delta g_E(r) \sim \frac{\beta e^2}{2r} e^{-\frac{1}{2}(r/\lambda)^2} \times [\log(r/\lambda) + \tfrac{1}{2}\log 2\gamma + \tfrac{1}{2}(\lambda/r)^2 + O(\lambda/r)^4] \quad (1.8)$$

for $r \gg \lambda$, and

$$\delta g_E(r) \sim \tfrac{1}{2}(\pi/2)^{\frac{1}{2}} \frac{\beta e^2}{\lambda} e^{-\frac{1}{2}(r/\lambda)^2} \times \left[1 + (2\pi)^{-\frac{1}{2}}\left(\frac{r}{\lambda}\right) - \tfrac{1}{6}\left(\frac{r}{\lambda}\right)^2 + \cdots\right] \quad (1.9)$$

for $r \ll \lambda$. The corresponding correction to the equation of state, which is of second order in $e^2$, is also evaluated in Sec. 2.

The second kind of quantum correction to be considered is a consequence of the quantum-mechanical treatment of classical statistics. By this we mean that the point charges which make up the classical system should strictly be treated as wave packets of spatial extent $\lambda$. These wave packets obey Maxwell–Boltzmann statistics. The net effect of making this correction is the appearance of a dynamically screened interaction[14] in place of the statically screened one. In Sec. 3 we obtain the effects of dynamical screening on the PDF up to the chain or pair approximation, which corresponds to the result of DeWitt[15] and Montroll and one of us[16] for the ring contribution to the partition function or equation of state. The result is

$$\delta g_c(r) \sim -(\beta e^2/r)\{e^{-\kappa r}[1 + \tfrac{1}{3}(\lambda/\lambda_D)^2 + O(\lambda/\lambda_D)^4] - (\lambda/r)^2 e^{-\frac{1}{2}(r/\lambda)^2}[1 + O((\lambda/r)^2, (\lambda/\lambda_D)^2)]\} \quad (1.10)$$

in the region $\lambda/\lambda_D \ll 1 \ll r/\lambda$, and

$$\delta g_c(r) \sim -(\pi/2)^{\frac{1}{2}}(\beta e^2/\lambda)[1 + O(\lambda/\lambda_D)^2 + O(r/\lambda)^2] \quad (1.11)$$

when $r/\lambda \ll \lambda/\lambda_D \ll 1$.

The calculations in this paper are performed for an electron gas immersed in a uniform background of neutralizing charge. The generalization to an electron–ion plasma is straightforward for the present contributions, and the method will be indicated briefly at the end of the appropriate sections. The system is translationally invariant.

[14] We use the terminology "dynamic screening" as a convenient mnemonic to distinguish such quantum corrections from exchange effects.

[15] H. E. DeWitt, J. Math. Phys. 3, 1216 (1962).

[16] E. W. Montroll and B. W. Ninham, "Quantum Corrections to the Debye Formula" (unpublished). The work was discussed by Professor Montroll at the February 1962 Baltimore meeting of the American Physical Society.

We use the equivalent diagram perturbation formalisms of Montroll and Ward[1,2] (MW) or Luttinger and Ward[3] (LW), since the problem has been formulated already by these authors in a convenient form for computation.

## 2. EXCHANGE CORRECTIONS

Exchange diagrams, of first order in the interaction which contribute to the PDF are shown in Figs. 1(b) and 1(c). Figure 1(a) represents the exchange term associated with the noninteracting system, which gives in the near classical limit the Uhlenbeck and Gropper formula Eq. (1.4). The diagram [Fig. 1(c)] is, in the terminology of Luttinger and Ward, anomalous, and will be shown to yield a result of higher order in $\hbar$ than Fig. 1(b). The main term to be evaluated then is the "crossed exchange" diagram of Fig. 1(b). The general quantum-mechanical expression for its contribution to the PDF is given by Ref. 2 (with an additional factor 2 for spin weighting),

$$\delta g_E(r) = \frac{1}{\rho^2(2\pi\hbar)^6\beta^2} \times \sum_{t_1 t_2=-\infty}^{\infty} \int d^3q_1 e^{i\mathbf{q}_1\cdot\mathbf{r}/\hbar} \int d^3q_2[-u(q_2)]2\Lambda_4(E), \quad (2.1)$$

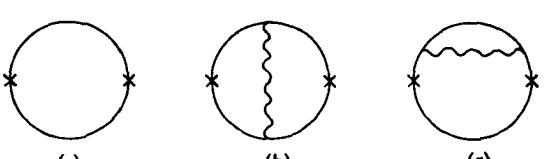
where

$$\Lambda_4(E) = \frac{1}{2\pi i}\int d^3p \int_c d\eta[(z^{-1}e^{-\beta\eta} + 1)^{-1} - ze^{\beta\eta}] \times \{[\eta + \epsilon(p)][\eta + 2\pi i t_1/\beta + \epsilon(\mathbf{p} + \mathbf{q}_1)] \times [\eta + 2\pi i(t_1 + t_2)/\beta + \epsilon(\mathbf{p} + \mathbf{q}_1 + \mathbf{q}_2)] \times [\eta + 2\pi i t_2/\beta + \epsilon(\mathbf{p} + \mathbf{q}_2)]\}^{-1} \quad (2.2)$$

and

$$u(q) = (2\pi\hbar)^{-3}\int d^3r v(r) e^{i\mathbf{q}\cdot\mathbf{r}/\hbar} = \frac{e^2}{2\hbar\pi^2 q^2} \quad (2.3)$$

is the Fourier transform of the Coulomb interaction. The contribution $ze^{\beta\eta}$ has been subtracted from the fermion function $(z^{-1}e^{-\beta\eta} + 1)^{-1}$ which usually appears in $\Lambda_n$ (Ref. 2), since a one-toron exchange diagram represents a forbidden process. In Eq. (2.2) $\epsilon(\mathbf{p}) = p^2/2m$, and the contour $c$ is to be chosen so that it separates the poles of the fermion function $(z^{-1}e^{-\beta\eta} + 1)^{-1}$ from those of the rest of the $\eta$ integrand. The usual MW procedure to evaluate the



(a) (b) (c)

FIG. 1. Exchange diagrams contributing to the pair distribution function up to first order in the interaction $e^2$. The crosses indicate unintegrated points.

propagator is to close the contour $c$ in Eq. (2.1) in an anticlockwise direction around the poles of the energy denominators which occur at

$$\eta = -\epsilon(\mathbf{p}), \quad -\epsilon(\mathbf{p}+\mathbf{q}_1) - 2\pi i t_1/\beta, \cdots. \tag{2.4}$$

For the explicit evaluation of exchange diagrams it turns out to be more convenient to close the contour of the $\eta$ integration in a clockwise direction, so that it encircles the poles of the fermion function. For electrons, we have poles due to $(z^{-1}e^{-\beta\eta}+1)^{-1}$ at

$$\eta = -\mu - (2l+1)\pi i/\beta \equiv -\zeta_l \tag{2.5}$$

with residue $1/\beta$. This leads at once to an expression for the propagator $\Lambda_4(E)$ in terms of the single-particle propagators

$$G(p, \omega_l) \equiv [\zeta_l - \epsilon(p)]^{-1} \tag{2.6}$$

of the LW formalism. Thus we have

$$\delta g_E(r) = \frac{2}{(2\pi\hbar)^6\beta^3\rho^2}\int d^3q_1 e^{i\mathbf{q}_1\cdot\mathbf{r}/\hbar}\int d^3q_2 u(q_2) \times \int d^3p S(\mathbf{p}, \mathbf{q}, \mathbf{q}_2), \tag{2.7}$$

where

$$S = \sum_{t_1,t_2,l=-\infty}^{\infty}\{[\zeta_l - \epsilon(\mathbf{p})][\zeta_{l+t_1} - \epsilon(\mathbf{p}+\mathbf{q}_1)] \times [\zeta_{l+t_1+t_2} - \epsilon(\mathbf{p}+\mathbf{q}_1+\mathbf{q}_2)] \times [\zeta_{l+t_2} - \epsilon(\mathbf{p}+\mathbf{q}_2)]\}^{-1}. \tag{2.8}$$

Note that to obtain the contribution to $\rho^{-1}(S(\mathbf{k})-1)$, where $S(\mathbf{k})$ is the static form factor[17] of the system, one simply drops the factor $(2\pi\hbar)^{-3}\int d^3q_1 \exp(i\mathbf{q}_1\cdot\mathbf{r}/\hbar)$ and puts $\mathbf{q}_1 = \hbar\mathbf{k}$.

The sums in Eq. (2.8) can be performed with the aid of the identity

$$\frac{1}{\beta}\sum_l \frac{e^{\zeta_l\alpha}}{\zeta_l - \epsilon} = \{H(\alpha)f_-(\epsilon) - H(-\alpha)f_+(\epsilon)\}e^{\alpha\epsilon}, \tag{2.9}$$

where

$$f_-(\epsilon) = 1/(z^{-1}e^{\beta\epsilon}+1), \quad f_+(\epsilon) = 1/(ze^{-\beta\epsilon}+1) = 1 - f_-(\epsilon), \tag{2.10}$$

and $H(\alpha)$ is the step function.

Changing to new summation variables

$$k_1 = l + t_1, \quad k_2 = l + t_2, \quad k_3 = l + t_1 + t_2 \tag{2.11}$$

and writing for notational convenience

$$\epsilon(\mathbf{p}) = \epsilon, \quad \epsilon(\mathbf{p}+\mathbf{q}_1) = \epsilon_1, \quad \epsilon(\mathbf{p}+\mathbf{q}_2) = \epsilon_2, \quad \epsilon(\mathbf{p}+\mathbf{q}_1+\mathbf{q}_2) = \epsilon_3, \tag{2.12}$$

we have

$$S = \sum_{l,k_1,k_2,k_3}\delta_{k_1+k_2,k_3+l}h(k_1, k_2, k_3, l) = \frac{1}{\beta}\int_0^\beta d\beta' \sum_{l,k_1,k_2,k_3} e^{2\pi i\beta'(k_1+k_2-k_3-l)/\beta}h(k_1, k_2, k_3, l), \tag{2.13}$$

where $h$ represents the summand of Eq. (2.8) and we have used an integral representation for the Kronecker-delta function. Then, using the identity (2.9) we have

$$S = \frac{1}{\beta}\int_0^\beta d\beta' \sum_{l,k_1,k_2,k_3}\frac{e^{-\beta'\zeta_l}}{(\zeta_l-\epsilon)}\frac{e^{\beta'\zeta_{k_1}}}{(\zeta_{k_1}-\epsilon_1)} \times \frac{e^{\beta'\zeta_{k_2}}}{(\zeta_{k_2}-\epsilon_2)}\frac{e^{-\beta'\zeta_{k_3}}}{(\zeta_{k_3}-\epsilon_3)} = \beta^3\int_0^\beta d\beta' f_+(\epsilon)f_-(\epsilon_1)f_-(\epsilon_2)f_+(\epsilon_3)e^{\beta'(-\epsilon+\epsilon_1+\epsilon_2-\epsilon_3)}. \tag{2.14}$$

Finally, noting that

$$f_-(\epsilon)e^{\beta\epsilon} = zf_+(\epsilon), \tag{2.15}$$

we can perform the integration over $\beta'$ in Eq. (2.14) and after substituting for $S$ in Eq. (2.7), find for the contribution to the PDF, the expression

$$\delta g_E(r) = \frac{2}{(2\pi\hbar)^6\rho^2}\int d^3q_1 e^{i\mathbf{q}_1\cdot\mathbf{r}/\hbar}\int d^3q_2 u(q_2)\int d^3p \times \frac{[f_+(\epsilon_1)f_+(\epsilon_2)f_-(\epsilon)f_-(\epsilon_3) - f_-(\epsilon_1)f_-(\epsilon_2)f_+(\epsilon)f_+(\epsilon_3)]}{(\epsilon_1+\epsilon_2-\epsilon-\epsilon_3)}. \tag{2.16}$$

A more convenient form for computation follows if we substitute for $u(q)$ from Eq. (2.3) and introduce dimensionless variables defined by the transformations

$$\mathbf{p} = \mathbf{P}(2m/\beta)^{\frac{1}{2}}; \quad \mathbf{q}_1 = \mathbf{Q}_1(2m/\beta)^{\frac{1}{2}}, \quad \mathbf{q}_2 = -(\mathbf{P}+\mathbf{Q}_1+\mathbf{Q}_2)(2m/\beta)^{\frac{1}{2}}, \tag{2.17}$$

where now

$$f_-(\mathbf{p}) = F_-(\mathbf{P}) = (z^{-1}e^{P^2}+1)^{-1}. \tag{2.18}$$

This substitution yields

$$\delta g_E(r) = \frac{2}{(2\pi\hbar)^6}(2m/\beta)^{\frac{7}{2}}\left(\frac{\beta e^2}{2\hbar\pi^2\rho^2}\right)\int d^3Q_1 e^{i\mathbf{Q}_1\cdot\mathbf{r}/\hbar}\int d^3P \times \int d^3Q_2 \frac{1}{(\mathbf{P}+\mathbf{Q}_1+\mathbf{Q}_2)^2} \times \{F_+(\mathbf{P}+\mathbf{Q}_1)F_+(\mathbf{Q}_1+\mathbf{Q}_2)F_-(\mathbf{P})F_-(Q_2) - F_-(\mathbf{Q}+\mathbf{P}_1)F_-(\mathbf{Q}_1+\mathbf{Q}_2)F_+(\mathbf{P})F_+(\mathbf{Q}_2)\} \times [(\mathbf{P}+Q_1)^2 + (\mathbf{Q}_1+\mathbf{Q}_2)^2 - P^2 - Q_2^2]^{-1}. \tag{2.19}$$

These somewhat tedious manipulations appear to be necessary to reduce the integral to a manageable form. Thus far our expression for $\delta g_E(r)$ is valid for a system of fermions at all temperatures. In the near

[17] D. Pines, *Elementary Excitations in Solids* (W. A. Benjamin, Inc., New York, 1964), p. 73.

classical limit, $F_+(P) \to 1$ and $F_-(P) \to z\exp(-P^2)$, so that the leading term is of order $z^2$. The fugacity $z$ is related to the density $\rho$ through (Montroll,[2] p. 67)

$$z = \tfrac{1}{2}\hbar^3(2\pi\beta/m)^{\frac{3}{2}}\rho\{1 - \pi^{\frac{1}{2}}e^3\beta^{\frac{3}{2}}\rho^{\frac{1}{2}} + O(\rho)\}, \quad (2.20)$$

where again we have inserted an additional factor $\frac{1}{2}$ to take care of electron spin, so that $\delta g_E(r)$ becomes

$$\delta g_E(r) = \frac{1}{4\pi^5}(\beta e^2/\lambda\!\!\!^{-})\int d^3Q_1 e^{i\mathbf{Q}_1\cdot\mathbf{r}/\lambda\!\!\!^{-}}\int d^3P \times \int d^3Q_2 \frac{1}{(\mathbf{P}+\mathbf{Q}_1+\mathbf{Q}_2)^2} \times \left\{\frac{e^{-(P^2+Q_2^2)} - e^{-(\mathbf{P}+\mathbf{Q}_1)^2}e^{-(\mathbf{Q}_1+\mathbf{Q}_2)^2}}{[(\mathbf{P}+\mathbf{Q}_1)^2 + (\mathbf{Q}_1+\mathbf{Q}_2)^2 - P^2 - Q_2^2]}\right\}. \quad (2.21)$$

Further reduction of the integral is carried out in the Appendix, and we find that $\delta g_E(r)$ has the asymptotic expansion for $r \gg \lambda\!\!\!^{-}$,

$$\delta g_E(r) \sim \tfrac{1}{2}(\beta e^2/r)e^{-\frac{1}{2}(r/\lambda\!\!\!^{-})^2} \times [\ln r/\lambda\!\!\!^{-} + \tfrac{1}{2}\ln(2\gamma) + \lambda\!\!\!^{-2}/2r^2 + O(\lambda\!\!\!^{-}/r)^4] \quad (2.22)$$

while for $r \ll \lambda\!\!\!^{-}$,

$$\delta g_E(r) \sim \tfrac{1}{2}(\pi/2)^{\frac{1}{2}}(\beta e^2/\lambda\!\!\!^{-})e^{-\frac{1}{2}(r/\lambda\!\!\!^{-})^2} \times [1 + (2\pi)^{-\frac{1}{2}}(r/\lambda\!\!\!^{-}) - \tfrac{1}{6}(r/\lambda\!\!\!^{-})^2 + O(r/\lambda\!\!\!^{-})^3]. \quad (2.23)$$

The leading term of the large $r$ expansion can become larger in magnitude than the Uhlenbeck–Gropper contribution under certain conditions.

The method of the preceding calculation can be used to compute the corresponding correction to the equation of state. If we join the two crosses of the open diagram Fig. 1(b) with an interaction line $u(q)$, we obtain the corresponding closed diagram, whose contribution to $\ln Z_G$ is

$$(\beta PV)_E \equiv \ln Z_{GE} = -\frac{2}{4}\frac{V}{(2\pi\hbar)^3\beta} \times \iiint d^3q_1 d^3q_2 d^3p u(q_1)u(q_2)S(\mathbf{p}, \mathbf{q}_1, \mathbf{q}_2) = EI_E, \quad (2.24)$$

where the propagator for the diagram is given by Eqs. (2.2) and (2.8). The factor 2 arises from spin weighting, the factor $\frac{1}{4}$ ensures that only topologically distinct diagrams are counted. The constant $E$ is given by

$$E = -\frac{1}{2}\frac{V}{(2\pi\hbar)^5}e^4z^2(2m)^2\left(\frac{2\pi m}{\beta}\right)^{\frac{1}{2}}\frac{1}{(2\pi)^{\frac{5}{2}}} = -\frac{\pi}{(2\pi)^5 r_s}\rho V(\Lambda/2\pi^{\frac{1}{2}})^{\frac{5}{3}} \quad (2.25)$$

with $r_s = me^2/\hbar^2\rho^{\frac{1}{3}}$ the usual quantum-mechanical parameter, and

$$I_E = \int \frac{d^3Q_1}{Q_1^2}\int d^3P \int \frac{d^3Q_2}{(\mathbf{P}+\mathbf{Q}_1+\mathbf{Q}_2)^2} \times \frac{[e^{-(P^2+Q_2^2)} - e^{-(\mathbf{P}+\mathbf{Q}_1)^2}e^{-(\mathbf{Q}_1+\mathbf{Q}_2)^2}]}{(\mathbf{P}+\mathbf{Q}_1)^2 + (\mathbf{Q}_1+\mathbf{Q}_2)^2 - P^2 - Q_2^2}$$
$$= \frac{(2\pi)^{\frac{7}{2}}}{4}\sum_{n=0}^{\infty}\frac{[(n-\frac{1}{2})!]^2 2^{2n}}{(2n+1)!\,(2n+1)} \quad (2.26)$$
$$= \sqrt{2}\,\pi^{\frac{11}{2}}\ln 2. \quad (2.27)$$

The evaluation of the integral $I_E$ is also carried out in the Appendix.

The contribution of the anomalous diagram Fig. 1(c) to the PDF or equation of state can be written down in the same manner and for the PDF has the form

$$\delta\overline{g_E}(r) = \frac{-2}{(2\pi\hbar)^6\beta^3\rho^2}\int d^3q_1 e^{i\mathbf{q}_1\cdot\mathbf{r}/\hbar} \times \int d^3q_2 u(q_2)\int d^3p S(\mathbf{p}, \mathbf{q}_1, \mathbf{q}_2), \quad (2.28)$$

where the propagator can be shown to be

$$S = \sum_{t_1,t_2,l}\frac{1}{(\zeta_l - \epsilon)^2}\frac{\exp\zeta_{l+t_1}0^+}{(\zeta_{l+t_1} - \epsilon_1)}\frac{\exp\zeta_{l+t_2}0^+}{(\zeta_{l+t_2} - \epsilon_2)} = -\beta^4 f_-(\epsilon_1)f_-(\epsilon_2)f_-(\epsilon)f_+(\epsilon). \quad (2.29)$$

In the near classical limit $f_-(\epsilon) \to ze^{-\beta p^2/2m}$; $f_+(\epsilon) \to 1$, and the contribution $\delta\overline{g_E}(r)$ to the PDF, while proportional to $e^2$ like $\delta g_E(r)$, is of higher order $O(z^3)$ in the fugacity, so that it can be neglected in comparison with $\delta g_E(r)$, Fig. 1(b). A similar conclusion holds for the contribution of Fig. 1(c) to the equation of state.

For a two component system in general, three PDF's exist, $g^{ee}(r)$, $g^{ii}(r)$, and $g^{ie}(r)$, where superscripts $e, i$ refer to electrons and ions, respectively. To generalize the exchange diagrams to the case of an electron–ion plasma we observe that exchange diagrams can be constructed only from identical particles. Thus each diagram of Fig. 1(b) must be composed of particle lines which are either all electron lines, contributing to $g^{ee}(r)$, or all ion lines contributing to $g^{ii}(r)$. Contributions to the ion–ion PDF will have the same form as $\delta g_E(r)$ above, differing only in the mass, and in general, in the magnitudes of density and temperature.

### 3. DYNAMIC SCREENING CORRECTIONS

In the chain (pair) approximation, the expression given by Montroll[2,18] for the PDF is

$$\delta g_c(r) = \frac{1}{(2\pi\hbar)^6\beta\rho^2}\sum_{l=-\infty}^{\infty}\int d^3q e^{i\mathbf{q}\cdot\mathbf{r}/\hbar}\frac{[-u(q)][2\Lambda_2(q,t)]^2}{1 + 2u(q)\Lambda_2(q,t)}. \quad (3.1)$$

[18] S. Fujita, A. Isihara, and E. W. Montroll, Bull. Acad. Roy. Belg. **44**, 1018 (1958).

To evaluate this expression, we use the explicit representation of $\Lambda_2$ for classical statistics[1,2]

$$\Lambda_2(q, t) = \Lambda_2(0, 0)\int_0^1 dx \exp[-\beta q^2 x(1-x)/2m] \times \exp(-2\pi i t x), \quad (3.2)$$

where

$$\Lambda_2(0, 0) = z\beta(2\pi m/\beta)^{\frac{3}{2}}. \quad (3.3)$$

The limit of high temperature and low density is equivalent to small $e^2$, since from dimensional analysis the leading correction to the perfect-gas equation of state must involve some function of the dimensionless parameter $e^2\beta\rho^{\frac{1}{3}}$. As $e^2 \to 0$, the major contribution to the integral of (3.1) comes from the region of small $q$. We therefore write

$$\Lambda_2(q, t) = \Lambda_2(0, 0) + \delta\Lambda_2(q, t), \quad (3.4)$$

where $\Lambda_2(0, 0)$ leads to the classical Debye–Hückel PDF, and

$$\delta\Lambda_2(q, t) = -\Lambda(0, 0)\int_0^1 dx \times \{1 - \exp[-\beta q^2 x(1-x)/2m]\exp(-2\pi i t x)\} \quad (3.5)$$

leads to quantum corrections due to dynamic screening. If we use Eq. (3.3) we may write the denominator of (3.1) as

$$\frac{4u(q)\Lambda_2^2(q, t)}{1 + 2u(q)\Lambda_2(q, t)} = \frac{4u(q)\Lambda_2^2(q, t)}{1 + 2u(q)\Lambda_2(0, 0)} - \frac{8[u(q)]^2\Lambda_2^2(q, t)\delta\Lambda_2(q, t)}{[1 + 2u(q)\Lambda_2(0, 0)][1 + 2u(q)\Lambda_2(q, t)]}. \quad (3.6)$$

The major corrections are contained in the first term on the right-hand side of Eq. (3.6). The second term is of higher order in both $e^2$ and $z$ and can be discarded, so that to order $z^2e^2$, after substituting Eqs. (3.2)–(3.6) into Eq. (3.1), we obtain the result

$$\delta g_c(r) = -\frac{2}{\pi}\frac{\beta e^2}{r}\int_0^\infty dq \sin(qr/\hbar)\frac{q}{(q^2 + \hbar^2\kappa^2)} \times \int_0^1 dx \exp[-\beta q^2 x(1-x)/2m] \times \int_0^1 dy \exp[-\beta q^2 y(1-y)/2m] \times \sum_{t=-\infty}^{\infty} e^{-2\pi i t(x+y)}. \quad (3.7)$$

In this expression we have performed the angular integration in (3.1) and have substituted for the fugacity $z$ the leading term from Eq. (2.20). The sum over $t$ may be carried out at once by noting that

$$\sum_{t=-\infty}^{\infty} e^{2\pi i t(x+y)} = \sum_{m=-\infty}^{\infty} \delta(x + y - m). \quad (3.8)$$

The delta function implies that $y = (m - x)$, and the restriction $0 \leq y \leq 1$ requires therefore that $0 \leq m - x \leq 1$. Since also $0 \leq x \leq 1$, $m = 1$ is the only term of the sum which contributes, so that

$$\delta g_c(r) = -\frac{2}{\pi}\frac{\beta e^2}{r}\int_0^\infty dq \sin(qr/\hbar)\frac{q}{q^2 + \hbar^2\kappa^2} \times \int_0^1 dx \exp[-2\beta q^2 x(1-x)/2m]. \quad (3.9)$$

Again the $q$ integral can be expressed in dimensionless variables by the substitution $p = q/\hbar\kappa$. Then

$$\beta q^2/2m = \alpha p^2,$$

where

$$\alpha \equiv \frac{\beta}{2m}(\hbar\kappa)^2 = (\lambda\!\!\!^{-}/\lambda_D)^2 \quad (3.10)$$

is a parameter which describes in a sense the degree of "quantumness" of the system. In the near classical limit, $\alpha$ is small compared to unity. Interchanging the order of integration, Eq. (3.9) becomes

$$\delta g_c(r) = -\frac{2}{\pi}\frac{\beta e^2}{r}\int_0^1 dx \int_0^\infty dp \sin(p\kappa r) \times \frac{p}{1 + p^2}\exp[-2\alpha p^2 x(1-x)]. \quad (3.11)$$

When $\alpha = 0$, this expression yields the classical linearized Debye–Hückel PDF

$$\delta g_c(r) \cong -\frac{\beta e^2}{r}e^{-\kappa r} \quad (3.12)$$

which diverges for small $r$. On the other hand for the quantum Boltzmann gas while the parameter $\alpha$ is always small, it is never identically zero, and $\delta g_c(r)$ does not diverge for small $r$. The asymptotic form of $\delta g_c(r)$ for large $r$ can be obtained as follows. We first perform the $p$ integration of Eq. (3.11) and find

$$\delta g_c(r) = -\frac{1}{2}\frac{\beta e^2}{r}\int_0^1 dx \exp[2\alpha x(1-x)] \times \{e^{-\kappa r}\operatorname{Erfc}(\xi_-) - e^{\kappa r}\operatorname{Erfc}(\xi_+)\}, \quad (3.13)$$

where

$$\xi_\mp = [2\alpha x(1-x)]^{\frac{1}{2}} \mp \tfrac{1}{2}\kappa r[2\alpha x(1-x)]^{-\frac{1}{2}} \quad (3.14)$$

and

$$\operatorname{Erfc}(\xi) = 2\pi^{-\frac{1}{2}}\int_\xi^\infty e^{-t^2} dt. \quad (3.15)$$

In the near classical limit $\alpha \to 0$, and since $\kappa\alpha^{-\frac{1}{2}} = 1/\lambda\!\!\!^{-}$, we have for $r/\lambda\!\!\!^{-} \gg 1$

$$\xi_\mp \to \mp\frac{r}{2\lambda\!\!\!^{-}}[2x(1-x)]^{-\frac{1}{2}}. \quad (3.16)$$

Hence using the relation

$$\text{Erfc}\,(-\xi) = 2 - \text{Erfc}\,(\xi) \tag{3.17}$$

and the asymptotic expansion

$$\text{Erfc}\,(\xi) = (\pi^{\frac{1}{2}}\xi)^{-1}e^{-\xi^2}\left(1 - \frac{1}{2\xi^2} + O(1/\xi^4)\right), \tag{3.18}$$

we find

$$\delta g_c(r) = -\frac{\beta e^2}{r}e^{-\kappa r}\int_0^1 \exp\,[2\alpha x(1-x)]\,dx + (2/\pi)^{\frac{1}{2}}(2\beta e^2\lambda\!\!\!^{-}/r^2)\int_0^1 dx[x(1-x)]^{+\frac{1}{2}} \times \exp\,[-r^2/8\lambda\!\!\!^{-2}x(1-x)]\{1 + O[(\lambda\!\!\!^{-}/r)^2, \alpha]\}, \tag{3.19}$$

where the terms neglected are of order $(\lambda\!\!\!^{-}/r)^2$, or $\alpha$, by comparison with the second term retained. The second integral can be evaluated as follows. We consider

$$I_c = e^{\frac{1}{2}(r/\lambda\!\!\!^{-})^2}\int_0^1 dx[x(1-x)]^{+\frac{1}{2}}\exp\left\{-\tfrac{1}{2}(r/\lambda\!\!\!^{-})^2\frac{1}{4x(1-x)}\right\} \tag{3.20}$$

an integral which can be cast into the convenient form

$$I_c = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}(p-1)!\,[\tfrac{1}{2}(r/\lambda\!\!\!^{-})^2]^{-p} \times \int_0^1 [x(1-x)]^{\frac{1}{2}}\left[\frac{1}{4x(1-x)} - 1\right]^{-p}dp, \qquad \text{Re}\,p > 0 \tag{3.21}$$

by using the Mellin integral representation for the exponential function. The $x$ integral may be written, with the change of variable

$$4x(1-x) = y,$$

as

$$2\int_0^{\frac{1}{2}}[x(1-x)]^{\frac{1}{2}}\left[\frac{1}{4x(1-x)} - 1\right]^{-p}dp = \frac{1}{4}\int_0^1 y^{p+\frac{1}{2}}(1-y)^{-p-\frac{1}{2}}\,dy = \frac{1}{4}\frac{\pi(p+\frac{1}{2})}{\sin\pi(p+\frac{1}{2})}, \quad -\tfrac{1}{2} < \text{Re}\,p < \tfrac{1}{2}. \tag{3.22}$$

Hence,

$$I_c = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\frac{\pi}{4}(p-1)!\,\frac{(p+\frac{1}{2})}{\sin\pi(p+\frac{1}{2})}[\tfrac{1}{2}(r/\lambda\!\!\!^{-})^2]^{-p}\,dp, \qquad 0 < \text{Re}\,p = c < \tfrac{1}{2}. \tag{3.23}$$

For large $r$, we may translate the contour to the right to get

$$I_c = \frac{(2\pi)^{\frac{1}{2}}}{4}\lambda\!\!\!^{-}/r[1 - 2(\lambda\!\!\!^{-}/r)^2 + O(\lambda\!\!\!^{-}/r)^4]. \tag{3.24}$$

The integral required is

$$I = e^{-\frac{1}{2}(r/\lambda\!\!\!^{-})^2}I_c \tag{3.25}$$

so that substituting Eqs. (3.24) and (3.25) into Eq. (3.19) and using the definition Eq. (3.10) we have

$$\delta g_c(r) = -\frac{\beta e^2}{r}\{e^{-\kappa r}[1 + \tfrac{1}{3}(\lambda\!\!\!^{-}/\lambda_D)^2 + O(\lambda\!\!\!^{-}/\lambda_D)^4] - (\lambda\!\!\!^{-}/r)^2 e^{-\frac{1}{2}(r/\lambda\!\!\!^{-})^2}[1 + O((\lambda\!\!\!^{-}/r)^2, (\lambda\!\!\!^{-}/\lambda_D)^2)]\}. \tag{3.26}$$

This result is valid in the region $\lambda\!\!\!^{-}/r \ll 1$, $\lambda\!\!\!^{-}/\lambda_D \ll 1$. For very small distances it is not difficult to show from Eq. (3.11) that

$$\delta g_c(r) = -\left(\frac{\pi}{2}\right)^{\frac{1}{2}}\frac{\beta e^2}{\lambda\!\!\!^{-}}\{1 + O(\alpha) + O(r/\lambda\!\!\!^{-})^2\}, \tag{3.27}$$

where $r/\lambda\!\!\!^{-} \ll \alpha \ll 1$.

Thus the dynamic screening correction removes the small $r$ divergence of the classical Debye–Hückel pair distribution function. The leading term, one half of whose contribution is cancelled by the exchange term Eq. (2.23), still attains very large negative values, under certain plasma conditions. This can easily be seen by writing

$$\frac{\pi^{\frac{1}{2}}}{2\sqrt{2}}\frac{\beta e^2}{\lambda\!\!\!^{-}} = \left(\frac{\pi}{2}\frac{\text{Ry}}{kT}\right)^{\frac{1}{2}} > 1; \quad T < 5\times 10^5\ ^\circ\text{K} \tag{3.28}$$

so that the small $r$ behavior of $\delta g_c(r)$ given by Eqs. (3.27) and (2.23) is not meaningful in this region.

The generalization of the dynamic screening contribution to the case of an electron–ion plasma is immediate. To do this we replace $\Lambda_2(q, t)$ in Eq. (3.1) by a sum of separate electron and ion contributions. Thus, in the numerator of (3.1),

$$[\Lambda(q,t)]^2 \to [\Lambda_e(q,t) + \Lambda_i(q,t)]^2 = \Lambda_e^2 + 2\Lambda_e\Lambda_i + \Lambda_i^2. \tag{3.29}$$

These terms lead to contributions to $\delta g^{ee}(r)$, $\delta g^{ei}(r)$, and $\delta g^{ii}(r)$, respectively. In each case the denominator of (3.1) should be replaced by $q^2 + \hbar^{22}\bar{\kappa}$, where

$$\bar{\kappa}^2 = (e^2/2\hbar\pi^2)[\Lambda_e(0,0) + \Lambda_i(0,0)] = \kappa_e^2 + \kappa_i^2. \tag{3.30}$$

When the electron and ion temperatures are equal and $\rho_e = \rho_i$, the quantum corrections to $\delta g^{ei}(r)$ and $\delta g^{ii}(r)$ are negligible, being of order $1/M$, $1/M^2$ with respect to $\delta g^{ee}(r)$. Here $M$ is the ion mass.

## ACKNOWLEDGMENT

One of the authors (B. W. N.) is indebted to the Australian Commonwealth Government for the award of a Queen Elizabeth II Fellowship.

*Note added in proof:* Since this work was completed it has been brought to the authors' attention by

Professor H. E. DeWitt that some of our results were obtained independently by Felix[19] and Trubnikov and Elesin.[20]

## APPENDIX: EXCHANGE INTEGRALS

We evaluate here the integrals of Eqs. (2.21) and (2.26). First consider the integral of Eq. (2.21) which can be written as

$$I = \int d^3Q_1 e^{iQ_1\cdot r/\lambda\!\!\!^{-}} \int d^3P \int \frac{d^3Q_2}{(\mathbf{P}+\mathbf{Q}_1+\mathbf{Q}_2)^2} \times \exp[-(\mathbf{P}+\mathbf{Q}_1)^2]\exp[-(\mathbf{Q}_1+\mathbf{Q}_2)^2]\int_0^1 dy \times \exp\{-y[P^2+Q_2^2-(\mathbf{P}+\mathbf{Q}_1)^2-(\mathbf{Q}_1+\mathbf{Q}_2)^2]\}. \tag{A1}$$

Putting $\mathbf{P}+\mathbf{Q}_1+\mathbf{Q}_2=-\mathbf{Q}_3$, $d^3Q_2=d^3Q_3$, we have

$$I = \int d^3Q_1 e^{iQ_1\cdot r/\lambda\!\!\!^{-}} \int d^3P \int \frac{d^3Q_3}{Q_3^2} \times e^{-[(\mathbf{P}+\mathbf{Q}_1)^2+(\mathbf{P}+\mathbf{Q}_3)^2]}\int_0^1 dy\, e^{-2y\mathbf{Q}_1\cdot\mathbf{Q}_3}. \tag{A2}$$

The $P$ integration is

$$\begin{aligned}
&\int d^3P \exp\{-[(\mathbf{P}+\mathbf{Q}_1)^2+(\mathbf{P}+\mathbf{Q}_3)^2]\} \\
&= 2\pi e^{-(Q_1^2+Q_3^2)}\int_0^\infty P^2\,dP e^{-2P^2}\int_{-1}^{+1} dx\, e^{-2P|\mathbf{Q}_1+\mathbf{Q}_3|x} \\
&= -2\pi\frac{e^{-(Q_1^2+Q_3^2)}}{2|\mathbf{Q}_1+\mathbf{Q}_3|}e^{\frac{1}{2}|\mathbf{Q}_1+\mathbf{Q}_3|^2} \\
&\quad\times\int_0^\infty P\,dP\{\exp[-2(P+\tfrac{1}{2}|\mathbf{Q}_1+\mathbf{Q}_3|)^2] \\
&\quad-\exp[-2(P-\tfrac{1}{2}|\mathbf{Q}_1+\mathbf{Q}_3|)^2]\} \\
&= \frac{\pi}{2}e^{-\frac{1}{2}(\mathbf{Q}_1-\mathbf{Q}_3)^2}\Big\{2\int_{\frac{1}{2}|\mathbf{Q}_1+\mathbf{Q}_3|}^\infty e^{-2x^2}dx \\
&\quad+\int_{-\frac{1}{2}|\mathbf{Q}_1+\mathbf{Q}_3|}^{+\frac{1}{2}|\mathbf{Q}_1+\mathbf{Q}_3|} e^{-2x^2}dx \\
&= \sqrt{2}\,(\pi^{\frac{3}{2}}/4)e^{-\frac{1}{2}(\mathbf{Q}_1-\mathbf{Q}_3)^2}.
\end{aligned} \tag{A3}$$

The integral $I$ then becomes

$$\begin{aligned}
I &= \frac{(2\pi)^{\frac{3}{2}}}{8}\int d^3Q_1 e^{iQ_1\cdot r/\lambda\!\!\!^{-}}\int\frac{d^3Q_3}{Q_3^2}e^{-\frac{1}{2}(\mathbf{Q}_1-\mathbf{Q}_3)^2}\int_0^1 dy\, e^{-2y\mathbf{Q}_1\cdot\mathbf{Q}_3} \\
&= \frac{(2\pi)^{\frac{3}{2}}}{4}\lambda\!\!\!^{-}/r\int_0^\infty Q_1\,dQ_1\sin(Q_1r/\lambda\!\!\!^{-}) \\
&\quad\times\int_0^\infty dQ_3 e^{-\frac{1}{2}(Q_1^2+Q_3^2)}\int_0^1 dy\int_{-1}^{+1} dz\, e^{(1-2y)Q_1Q_3z} \\
&= \frac{(2\pi)^{\frac{3}{2}}}{2}\lambda\!\!\!^{-}/r\int_0^\infty Q_1\,dQ_1\sin(Q_1r/\lambda)e^{-\frac{1}{2}Q_1^2} \\
&\quad\times\int_0^\infty dQ_3 e^{-\frac{1}{2}Q_3^2}\sum_{n=0}^\infty\frac{(Q_1Q_3)^{2n}}{(2n+1)!(2n+1)} \\
&= \frac{(2\pi)^{\frac{3}{2}}}{2}\lambda\!\!\!^{-}/r\sum_{n=0}^\infty\frac{2^{n-\frac{1}{2}}(n-\frac{1}{2})!}{(2n+1)(2n+1)!} \\
&\quad\times\int_0^\infty dQ_1 Q_1^{2n+1}e^{-\frac{1}{2}Q_1^2}\sin(Q_1r/\lambda\!\!\!^{-}) \\
&\equiv (2\pi)^{\frac{3}{2}}(\lambda\!\!\!^{-}/2r)\sigma.
\end{aligned} \tag{A4}$$

We have carried out this integral in some detail, since, unless the integrals are performed in the order indicated, $I$ is exceedingly difficult to evaluate. In order to complete the quadrature, we need to obtain expansions appropriate to both large and small $r$, and proceed as follows. Using the duplication formula

$$2^{2n}(n-\tfrac{1}{2})!\,n! = \pi^{\frac{1}{2}}(2n)!,$$

we have

$$\sigma = \tfrac{1}{4}(\pi/2)^{\frac{1}{2}}\sum_{n=0}^\infty\frac{1}{n!\,(n+\frac{1}{2})^2} \times\int_0^\infty dQ(Q^2/2)^nQe^{-\frac{1}{2}Q^2}\sin(Qr/\lambda\!\!\!^{-}). \tag{A5}$$

Then the observation that

$$\frac{1}{(n+\frac{1}{2})^2} = \int_0^1\frac{dx}{x}\int_0^x t^{n-\frac{1}{2}}\,dt \tag{A6}$$

allows us to write

$$\begin{aligned}
\sigma &= \tfrac{1}{4}(\pi/2)^{\frac{1}{2}}\int_0^1\frac{dx}{x}\int_0^x t^{-\frac{1}{2}}\,dt \\
&\quad\times\sum_{n=0}^\infty\int_0^\infty\frac{(Q^2t/2)^n}{n!}Q\,dQ\,e^{-\frac{1}{2}Q^2}\sin(Qr/\lambda\!\!\!^{-}) \\
&= \tfrac{1}{4}(\pi/2)^{\frac{1}{2}}\int_0^1\frac{dx}{x}\int_0^x t^{-\frac{1}{2}}\,dt \\
&\quad\times\int_0^\infty Q\,dQ\,e^{-\frac{1}{2}Q^2(1-t)}\sin Qr/\lambda\!\!\!^{-}.
\end{aligned} \tag{A7}$$

Hence, after completing the $Q$ integration we obtain

$$\sigma = \frac{\pi}{8}\frac{r}{\lambda\!\!\!^{-}}\int_0^1\frac{dx}{x}\int_0^x t^{-\frac{1}{2}}(1-t)^{-\frac{3}{2}}\exp\left[-\tfrac{1}{2}(r/\lambda\!\!\!^{-})^2\frac{1}{(1-t)}\right]. \tag{A8}$$

Integration by parts with respect to $x$ gives

$$\sigma = -\frac{\pi}{8}\frac{r}{\lambda\!\!\!^{-}}\int_0^1 dx\ln x\,x^{-\frac{1}{2}}(1-x)^{-\frac{3}{2}} \times\exp\left[-\tfrac{1}{2}(r/\lambda\!\!\!^{-})^2\frac{1}{(1-x)}\right] \tag{A9}$$

and a further change of variable to

$$y = \tfrac{1}{2}(r/\lambda\!\!\!^{-})^2(x/1-x)$$

[19] M. Felix in *Comptes Rendus de la VIe Conférence Internationale sur les Phénomènes d'Ionisation dans les Gaz*, P. Hubert and E. Crémien-Alcan, Eds. (Published with the support of the French Government, Paris, 1963), Vol. I, p. 185.

[20] B. A. Trubnikov and V. F. Elesin, Sov. Phys.—JETP **20**, 866 (1965).

yields the convenient form

$$\sigma = -\frac{\pi}{8}\sqrt{2}\, e^{-\frac{1}{2}(r/\lambda)^2}\int_0^\infty dy\, y^{-\frac{1}{2}}e^{-y}\left\{\ln\left[2\left(\frac{\lambda}{r}\right)^2\right] + \ln y - \ln\left[1 + 2\left(\frac{\lambda}{r}\right)^2 y\right]\right\}. \quad \text{(A10)}$$

For large $r$ we may expand the logarithm in the third term in the braces, and carry out the remaining integrals to get the complete asymptotic expansion

$$\sigma = -\pi(\sqrt{2}/8)e^{-\frac{1}{2}(r/\lambda)^2}\left\{(\pi)^{\frac{1}{2}}\ln\left[2\left(\frac{\lambda}{r}\right)^2\right] - \pi^{\frac{1}{2}}\ln(4\gamma) + \sum_{n=1}^{\infty}(-1)^n\frac{(n-\frac{1}{2})!}{n}\left[2\left(\frac{\lambda}{r}\right)^2\right]^n\right\}, \quad \text{(A11)}$$

where $\gamma$ is Euler's constant.

Collecting together Eqs. (A11), (A4), and (2.21), we have finally

$$\delta g_E(r) = \frac{1}{2}\frac{\beta e^2}{r}e^{-\frac{1}{2}(r/\lambda)^2} \times \left\{\ln(r/\lambda) + \frac{\ln 2\gamma}{2} + \tfrac{1}{2}(\lambda/r)^2 + O(\lambda/r)^4\right\}, \quad \text{(A12)}$$

which is the result quoted in Eq. (2.22).

For small $r$, we rewrite Eq. (A10) as

$$\sigma = \pi(\sqrt{2}/8)e^{-\frac{1}{2}(r/\lambda)^2}\int_0^\infty dy\, y^{-\frac{1}{2}}e^{-y}\ln\left[1 + (r/\lambda)^2\frac{1}{2y}\right]. \quad \text{(A13)}$$

The integral representation

$$\ln(1+x) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} dp\, x^{-p}\frac{\pi}{p\sin\pi p}, \quad -1 < \operatorname{Re} p < 0 \quad \text{(A14)}$$

gives the alternative form

$$\sigma = \frac{\pi\sqrt{2}}{8}e^{-\frac{1}{2}(r/\lambda)^2}\frac{1}{2\pi i} \times \int_{c-i\infty}^{c+i\infty} 2^p(\lambda/r)^{2p}\frac{\pi\, dp}{p\sin\pi p}\int_0^\infty dy\, e^{-y}y^{p-\frac{1}{2}}, \quad \text{(A15)}$$

where the $y$ integral gives $(p-\frac{1}{2})!$, which restricts the contour to the region $-\frac{1}{2} < \operatorname{Re} p < 0$. For small $r$, we close the contour to the left and obtain

$$\delta g_E(r) = \tfrac{1}{2}(\pi/2)^{\frac{1}{2}}(\beta e^2/\lambda)e^{-\frac{1}{2}(r/\lambda)^2} \times [1 + r/(2\pi)^{\frac{1}{2}}\lambda - \tfrac{1}{6}(r/\lambda)^2 + O(r/\lambda)^3], \quad \text{(A16)}$$

which is the result quoted in Eq. (2.23).

The reduction of the integral Eq. (2.26) to the form quoted,

$$I_E = \frac{(2\pi)^{\frac{7}{2}}}{4}\sum_{n=0}^{\infty}\frac{[(n-\frac{1}{2})!]^2 2^{2n}}{(2n+1)!\,(2n+1)}, \quad \text{(A17)}$$

follows in precisely the same manner. The sum can be evaluated by using the duplication formula and the identity of Eq. (A6) to write

$$\frac{\sqrt{2}\,I_E}{\pi^4} = \int_0^1\frac{dx}{x}\int_0^x dt\sum_{n=0}^{\infty}t^{n-\frac{1}{2}}\frac{(n-\frac{1}{2})!}{n!} = \pi^{\frac{1}{2}}\int_0^1\frac{dx}{x}\int_0^x dt[t(1-t)]^{-\frac{1}{2}}. \quad \text{(A18)}$$

Integration by parts gives then the result Eq. (2.27):

$$\sqrt{2}\,\pi^{-\frac{3}{2}}I_E = -\int_0^1 dx[x(1-x)]^{-\frac{1}{2}}\ln x = -\left[\frac{d}{d\alpha}\int_0^1 x^\alpha(1-x)^{-\frac{1}{2}}dx\right]_{\alpha=-\frac{1}{2}} = \pi 2\ln 2. \quad \text{(A19)}$$

# Solution of the Transport Equation with Anisotropic Scattering in Slab Geometry*

ROBERT L. BOWDEN, F. JOSEPH MCCROSSON, AND EDGAR A. RHODES†
*Department of Physics, Virginia Polytechnic Institute, Blacksburg, Virginia*

(Received 4 August 1966)

Some systematics which exist between eigenfunctions and adjoint singular integral equations arising in the solution of the transport equation in slab geometry are illustrated. The transport equation is shown to obey a singular integral equation and its relationship to the eigenfunction expansion-method solution is shown. A new method for solving for the expansion coefficients in the eigenfunction expansion method is illustrated by solving Milne's problem. The role adjoint singular integral equations play in finding appropriate weight functions for use in orthogonality relations between the eigenfunctions of the transport equation is briefly discussed.

## 1. INTRODUCTION

In recent years, exact results to various neutron transport problems in slab geometry have been obtained in which the theory of singular integral equations[1] plays a central role. In this paper we illustrate some of the systematics which exist between orthogonality of functions and adjointness[1] of singular integral equations encountered in the solution to the one-speed neutron transport equation with anisotropic scattering in slab geometry.

In particular, two different approaches have been used to obtain solutions to problems based on the above-mentioned equation: the eigenfunction expansion method due to Case[2] as presented by Mika,[3] and the transform method due to Leonard and Mullikin.[4] In the eigenfunction expansion method, the independent variables of the homogeneous integro-differential form of the transport equation are separated and the general solution is expressed as an eigenfunction expansion with arbitrary coefficients over the spectrum of the separation parameter. This spectrum consists of a discrete and a continuous part. The expansion coefficients are determined by applying boundary conditions and solving the resulting singular integral equations. In some cases the actual solving of the singular integral equations can be avoided by using orthogonality relations between the eigenfunctions. More recently, Leonard and Mullikin[4] have analyzed a problem in a finite slab by considering a subcritical assembly with a source incident on one face. This scattering problem can be stated as an inhomogeneous integral transport equation for the source function consisting of the contributions from scattered neutrons and the external sources. By generalizing the problem to include complex-valued sources, they were able to show that the source function satisfies a singular integral equation in which the space and angle variables enter as parameters.

The singular integral equations obtained by Leonard and Mullikin are adjoint[1] to the corresponding singular integral equations for the continuum coefficient in the eigenfunction expansion method, indicating a duality between the two methods. To help explain this duality, it is convenient to show that the angular flux itself satisfies a singular integral equation in which the space variable enters as a parameter (cf. Ref. 5). The relationship between the two methods then follows from the orthogonality of the eigenfunctions.

An outline of the remainder of this paper is as follows: In Sec. 2, a brief list of the results of Mika is given in order to show where similarities and differences exist between the present approach and the eigenfunction expansion method and in order to introduce results which will be used to derive the relationship between the two methods. In Sec. 3, the angular flux is shown to obey a singular integral equation. This is shown by using the homogeneous integro-differential form of the transport equation. This singular integral equation can be solved by the procedure outlined by Leonard and Mullikin.[4] In Sec. 4 the relationship between the two methods is found. This leads to another procedure for solving for the expansion coefficients in the eigenfunction expansion method. We illustrate this procedure by solving Milne's problem. In Sec. 5 the role adjoint singular integral equations play in obtaining weight functions for use in orthogonality relations between the eigenfunctions is briefly discussed.

* This work has been supported in part by the U.S. Atomic Energy commission.

† Present address: Department of Physics, Florida State University, Tallahassee, Florida.

[1] N. I. Muskelishvili, *Singular Integral Equations* (P. Noordhoff Ltd., Groningen, The Netherlands, 1953).

[2] K. M. Case, Ann. Phys. (New York) 9, 1 (1960).

[3] J. R. Mika, Nucl. Sci. Eng. 11, 415 (1961).

[4] A. Leonard and T. W. Mullikin, J. Math. Phys. 5, 399 (1964).

[5] G. J. Mitsis, Argonne National Laboratory, ANL-6768 (unpublished).

## 2. EIGENFUNCTION EXPANSION METHOD

In this section we review the eigenfunction expansion method as presented by Mika.[3] Here it is assumed that the scattering function $f_s(\Omega' \to \Omega)$, where $f_s(\Omega' \to \Omega)\, d\Omega$ represents the probability that a neutron scatters from the direction $\Omega'$ into a solid angle $d\Omega$ about the direction $\Omega$, can be expanded into a finite series of Legendre polynomials:

$$f_s(\Omega' \to \Omega) = \frac{1}{4\pi} \sum_{n=0}^{N} b_n P_n(\Omega' \cdot \Omega). \tag{2.1}$$

The $b_n$ are numbers which determine the degree of anisotropy of scattering and are restricted by the condition $f_s \geq 0$, with $b_0 = 1$. Under assumption (2.1), the source-free, one-speed neutron transport equation for the angular flux $\psi(x, \mu)$ has the form

$$\mu \frac{\partial \psi}{\partial x} + \psi = \frac{c}{2} \sum_{n=0}^{N} b_n P_n(\mu) \int_{-1}^{1} P_n(\mu') \psi(x, \mu')\, d\mu'. \tag{2.2}$$

In the above equation $x$ is the spatial coordinate measured in mean free paths, $\mu$ is the corresponding direction cosine, and $c$ is the mean number of secondary neutrons per collision.

The eigenfunctions of Eq. (2.2) are of the form

$$\phi(\nu, \mu) e^{-x/\nu}, \quad \nu \in (-1, 1) \tag{2.3a}$$

and

$$\phi(\pm z_j, \mu) e^{\mp x/z_j}, \quad j = 1, \cdots, \alpha, \tag{2.3b}$$

where the $\phi$'s are the solution to the equation

$$[\nu - \mu]\phi(\nu, \mu) = \nu k(\mu, \nu), \tag{2.4}$$

viz.,

$$\phi(\nu, \mu) = P[\nu k(\mu, \nu)]/(\nu - \mu) + \lambda(\nu)\delta(\nu - \mu), \quad \nu \in (-1, 1), \tag{2.5}$$

$$\phi(\pm z_j, \mu) = [z_j k(\mu, \pm z_j)]/(z_j \mp \mu), \quad j = 1, \cdots, \alpha, \tag{2.6}$$

$$\lambda(\nu) = 1 + P \int_{-1}^{1} \frac{\nu k(\mu, \nu)}{\mu - \nu}\, d\mu, \tag{2.7}$$

with $z_j$ $[(z_j \notin (-1, 1)]$ defined by

$$\Omega(\pm z_j) = 0, \tag{2.8}$$

where

$$\Omega(z) = 1 + \int_{-1}^{1} \frac{z k(\mu, z)}{\mu - z}\, d\mu, \tag{2.9}$$

there being $\alpha$ such pairs of roots. Mika has shown that $\alpha \leq N + 1$. In the above equations

$$k(\mu, \eta) = \frac{c}{2} \sum_{n=0}^{N} b_n P_n(\mu) h_n(\eta) \tag{2.10}$$

with

$$h_n(\eta) = \int_{-1}^{1} P_n(\mu) \phi(\eta, \mu)\, d\mu. \tag{2.11}$$

The $h_n$ can be shown to be polynomials which obey the following recursion formula:

$$(n + 1)h_{n+1}(\eta) + \eta[cb_n - (2n + 1)]h_n(\eta) + nh_{n-1}(\eta) = 0. \tag{2.12}$$

The $h_n$ are normalized so that

$$h_0(\eta) = 1 \tag{2.13}$$

from which it follows with Eqs. (2.4) and (2.11) that

$$h_1(\eta) = (1 - c)\eta. \tag{2.14}$$

Equation (2.12), together with Eqs. (2.13) and (2.14), can then be regarded as an alternate definition of the $h_n$. The symbol $P$ denotes that the Cauchy principal value is to be taken in any integration involving the term following it and $\delta(\nu - \mu)$ is the Dirac-delta function.

Mika has shown that the above solutions of Eq. (2.4) form a complete set for sufficiently well-behaved but otherwise arbitrary functions defined on the interval $(s, t)$, $-1 \leq s < t \leq 1$. One consequence of this completeness property is that the general solution of Eq. (2.2) can be written in the form

$$\psi(x, \mu) = \sum_{j=1}^{\alpha} [a_{+j}\phi(z_j, \mu)e^{-x/z_j} + a_{-j}\phi(-z_j, \mu)e^{x/z_j}] + \int_{-1}^{1} A(\nu)\phi(\nu, \mu)e^{-x/\nu}\, d\nu, \tag{2.15}$$

where the $a_{\pm j}$ and $A(\nu)$ are arbitrary expansion coefficients which can be determined by applying appropriate boundary conditions to $\psi(x, \mu)$.

Mika has also shown the following useful orthogonality properties:

$$\int_{-1}^{1} \mu\phi(\nu, \mu)\phi(\eta, \mu)\, d\mu = 0, \quad \nu \neq \eta, \tag{2.16}$$

for two arbitrary solutions of Eq. (2.4), either from the continuous or discrete parts of the spectrum, and

$$\int_{-1}^{1} \mu\phi(\nu, \mu) \int_{-1}^{1} Q(\eta)\phi(\eta, \mu)\, d\eta\, d\mu = \nu S(\nu) Q(\nu) \tag{2.17}$$

for solutions of Eq. (2.4) from the continuous part of the spectrum only. In the above, $Q(\eta)$ is a sufficiently well behaved but otherwise arbitrary function and $S(\nu)$ is defined by

$$S(\nu) \equiv \Omega^{+}(\nu)\Omega^{-}(\nu) = [\lambda(\nu)]^2 + [\pi\nu N(\nu)]^2, \tag{2.18}$$

where

$$N(\nu) \equiv k(\nu, \nu) \tag{2.19}$$

and

$$\Omega^{\pm}(\nu) = \lambda(\nu) \pm i\pi\nu N(\nu), \tag{2.20}$$

are the limits given by Plemelj's formulas[1] of $\Omega(z)$ from the upper (+) and lower (−) half-planes as $z$ tends to $\nu$, $-1 < \nu < 1$. For simplicity, we shall

assume throughout this paper that $S(\nu) \neq 0$, $-1 \leq \nu \leq 1$. The case for which this condition is not satisfied can be treated[3] by increasing the index $2\alpha$, and hence the number of discrete solutions, by the number of zeros of $S(\nu)$, $-1 \leq \nu \leq 1$. We shall also assume that

$$\Omega(\infty) \neq 0, \tag{2.21}$$

where[4]

$$\Omega(\infty) = \prod_{n=0}^{N} \left(1 - \frac{cb_n}{2n+1}\right). \tag{2.22}$$

We shall view those $c$ and $b_n$ for which the right side of Eq. (2.22) vanishes as limits of a sequence of cases where the inequality (2.21) holds.

For the discrete spectrum, the following norms are found:

$$\int_{-1}^{1} \mu[\phi(\pm z_j, \mu)]^2 \, d\mu = \pm z_j M_j, \tag{2.23}$$

where[6]

$$M_j = z_j N(z_j)[d\Omega(z)/dz]_{z=z_j}. \tag{2.24}$$

The expansion coefficients can be determined *a la* Mika by applying boundary conditions directly to Eq. (2.15) and solving the resulting singular integral equation using the general theory presented by Muskelishvili.[1] For example, in examining the critical problem, we consider an infinite slab extending from $x = -a$ to $x = +a$ and bounded by vacuum. The critical angular flux is given by Eq. (2.15) subject to the symmetry condition

$$\psi(x, \mu) = \psi(-x, \mu) \tag{2.25}$$

and the boundary condition

$$\psi(a, \mu) = 0, \quad \mu < 0. \tag{2.26}$$

Equation (2.25) expresses the fact that the angular flux must be symmetric about the $x = 0$ plane, while Eq. (2.26) states the condition that no neutrons which migrate from the slab into the vacuum return to the slab. It follows from the symmetry condition (2.25) that

$$a_{+j} = a_{-j} \tag{2.27a}$$

and

$$A(\nu) = A(-\nu). \tag{2.27b}$$

Application of the boundary condition (2.26) leads to the singular integral equation

$$\lambda(\mu)A'(\mu) + P\int_{-1}^{1} \frac{\nu k(\mu, \nu)}{\nu - \mu} A'(\nu) \, d\nu = \Psi'(\mu), \tag{2.28}$$

where

$$\Psi'(\mu) = -\sum_{j=1}^{\alpha} a_{+j}[\phi(z_j, \mu)e^{a/z_j} + \phi(z_j, -\mu)e^{-a/z_j}] - \int_0^1 A'(\nu)\phi(\nu, -\mu)e^{-2a/\nu} \, d\nu \tag{2.29}$$

and

$$A'(\nu) = A(\nu)e^{a/\nu}. \tag{2.30}$$

Equation (2.28) can be reduced by standard techniques to a Fredholm integral equation plus a set of homogeneous equations for the $a_{+j}$ for which the vanishing of the secular determinant yields the critical condition (see, for example, Ref. 7). The singular operator in Eq. (2.25) is adjoint to the singular operator obtained by Leonard and Mullikin [Ref. 4, Eqs. (3.22)] in their analysis of the critical problem. In the following section we present another method for solving transport problems which also lead to singular operators which are adjoint to those arising in the eigenfunction expansion method but which can be easily connected to the eigenfunction expansion method.

## 3. SINGULAR INTEGRAL EQUATION

In this section, we present an alternate procedure for solving Eq. (2.2) by showing that the angular flux satisfies a singular integral equation. We start by setting

$$\rho_n(x) = \int_{-0}^{1} P_n(\nu)\psi(x, \nu) \, d\nu, \tag{3.1}$$

so that we may write Eq. (2.2) as

$$\mu \frac{\partial \psi}{\partial x} + \psi = \frac{c}{2} \sum_{n=0}^{N} b_n P_n(\mu)\rho_n(x). \tag{3.2}$$

We now define

$$\Psi_0(x, \mu) = \int_{-1}^{1} \frac{k(\nu, \mu)}{\nu - \mu} [\nu\psi(x, \nu) - \mu\psi(x, \mu)] \, d\nu. \tag{3.3}$$

Noting from Eq. (3.2) that

$$\left(\mu \frac{\partial}{\partial x} + 1\right)\nu\psi(x, \nu) = \frac{c\mu}{2} \sum_{n=0}^{N} b_n P_n(\nu)\rho_n(x) + (\nu - \mu)\psi(x, \nu), \tag{3.4}$$

we find

$$\left(\mu \frac{\partial}{\partial x} + 1\right)\Psi_0(x, \mu) = \frac{c}{2} \sum_{n=0}^{N} b_n \left\{h_n(\mu) + \mu \int_{-1}^{1} \frac{k(\nu, \mu)}{\nu - \mu} [P_n(\nu) - P_n(\mu)] \, d\nu\right\} \rho_n(x). \tag{3.5}$$

[6] N. J. McCormick and I. Kuščer, J. Math. Phys. 7, 2036 (1966).

[7] F. J. McCrosson, M.S. thesis, Virginia Polytechnic Institute (1964).

Inspection of Eqs. (3.2) and (3.5) shows that if the $h_n$ satisfy the equation

$$h_n(\mu) + \mu \int_{-1}^{1} \frac{k(\nu, \mu)}{\nu - \mu} [P_n(\nu) - P_n(\mu)]\, d\nu = P_n(\mu), \quad (3.6)$$

then $\Psi_0(x, \mu)$ is a "particular solution" of Eq. (2.2). This can be shown by first noting that Eqs. (3.6) yield $h_0(\mu) = 1$ and $h_1(\mu) = (1 - c)\mu$. Using the recursion relation for Legendre polynomials, we also find that Eqs. (3.6) yield the same recursion formula as Eq. (2.12), so that the $h_n$ do indeed satisfy Eq. (3.6).

The function $\Psi_0(x, \mu)$ then is a "particular solution" of Eq. (2.2). To obtain the "general solution," we must add to $\Psi_0(x, \mu)$ a solution $\Psi_1(x, \mu)$ of the homogeneous equation

$$[\mu(\partial/\partial x) + 1]\Psi_1(x, \mu) = 0, \quad (3.7)$$

viz.,

$$\Psi_1(x, \mu) = -F(\mu)e^{-x/\mu}, \quad (3.8)$$

where $F(\mu)$ is an arbitrary function of $\mu$ whose form is to be determined by appropriate boundary conditions. Expressing the "general solution" of Eq. (2.2) as $\Psi_0(x, \mu) + \Psi_1(x, \mu)$, we write

$$\psi(x, \mu) = \int_{-1}^{1} \frac{k(\nu, \mu)}{\nu - \mu} [\nu\psi(x, \nu) - \mu\psi(x, \mu)]\, d\nu - F(\mu)e^{-x/\mu}. \quad (3.9)$$

Interpreting each part of the integral in Eq. (3.9) Cauchy principal valued integrals, we find the singular integral equation

$$\lambda(\mu)\psi(x, \mu) - P\int_{-1}^{1} \frac{\nu k(\nu, \mu)}{\nu - \mu} \psi(x, \nu)\, d\nu = -F(\mu)e^{-x/\mu} \quad (3.10)$$

in which $x$ appears only as a parameter. Equation (3.10) can be solved by a procedure very similar to that outlined by Leonard and Mullikin[4] (also see Ref. 5). For example, in the critical problem we first find the form of $F(\mu)$ by use of the symmetry and boundary conditions (2.26) and (2.25):

$$F(\mu) = \begin{cases} e^{-a/\mu} \displaystyle\int_0^1 \frac{\nu k(\nu, -\mu)}{\nu + \mu} \psi(a, \nu)\, d\nu, & \mu > 0 \\ e^{a/\mu} \displaystyle\int_0^1 \frac{\nu k(\nu, \mu)}{\nu - \mu} \psi(a, \nu)\, d\nu, & \mu < 0. \end{cases} \quad (3.11)$$

At this point, we analytically continue $\psi$ from Eq. (3.9) to the complex plane of $z$. To do this, we define a function $f(a, z)$ such that $f(a, \mu) = F(\mu)$ for $\mu > 0$ and $f(a, -\mu) = F(\mu)$ for $\mu < 0$:

$$f(a, z) = e^{-a/z} \int_0^1 \frac{\nu k(\nu, -z)}{\nu + z} \psi(a, \nu)\, d\nu. \quad (3.12)$$

We then get the functional equation

$$\Omega(z)\psi(x, z) - \int_{-1}^{1} \frac{\nu k(\nu, z)}{\nu - z} \psi(x, \nu)\, d\nu = \begin{cases} -e^{-x/z} f(a, z), & \operatorname{Re} z > 0 \\ -e^{x/z} f(a, -z), & \operatorname{Re} z < 0, \end{cases} \quad (3.13)$$

where $\Omega(z)$ is defined by Eq. (2.9). Recalling that $\Omega(z)$ has $\alpha$ pairs of zeros in the complex plane cut along $(-1, 1)$, we see that the remainder of Eq. (3.13) must vanish to the same order as $\Omega(z)$ at these points in order for $\psi(x, z)$ to be analytic on $0 < |z| < \infty$. This gives linear constraints such as

$$f(a, \pm z_j) = e^{\mp x/z_j} \int_0^1 \frac{\nu k(\nu, \mp z_j)}{\nu \pm z_j} \psi(x, \nu)\, d\nu, \quad j = 1, \cdots, \alpha, \quad (3.14)$$

for $\Omega(z_j) = 0$, $[d\Omega(z) \mid dz]_{z=z_j} \neq 0$. Setting $x = a$ in Eq. (3.13) and using the definition of $f$ in Eq. (3.12), we also find

$$f(a, z_j) = f(a, -z_j), \quad j = 1, \cdots, \alpha. \quad (3.15)$$

We note that from the general theory of singular integral equations, the solution for $\psi(x, \mu)$ can be expressed in terms of $f(a, \mu)$. But with $z = \mu$ in Eq. (3.12), $f(a, \mu)$ is defined as an integral over $\psi(a, \mu)$. Thus, the critical angular flux $\psi(x, \mu)$ at every point in the slab can be determined from the emerging critical angular flux $\psi(a, \mu)$ at the face of the slab. A singular integral equation for $\psi(a, \mu)$ can be obtained by letting $x = a$ and restricting $\mu > 0$ in Eq. (3.10):

$$\lambda(\mu)\psi(a, \mu) - P\int_0^1 \frac{\nu k(\nu, \mu)}{\nu - \mu} \psi(a, \nu)\, d\nu = -e^{-a/\mu} f(a, \mu), \quad (3.16)$$

where use has been made of the boundary condition (2.26). We note that the singular operator in this equation is adjoint to the singular operator arising in the eigenfunction expansion method. This type of singular integral equation has been analyzed by Leonard and Mullikin[4] by first treating $f(a, \mu)$ as a known function and inverting the singular operator by a slightly modified form of the theory of singular integral equations, finally obtaining a regular Fredholm equation for $f(a, \mu)$. The critical condition obtained with minor modification is that obtained by Leonard and Mullikin.[4,8]

## 4. RELATIONSHIP BETWEEN DIFFERENT APPROACHES

We are now prepared to show the relationship between the approaches in Secs. 2 and 3. We do so

[8] E. A. Rhodes, M.S. thesis, Virginia Polytechnic Institute (1965).

by noting first that if we use the explicit form of $\phi(\mu, \nu)$ from Eq. (2.5) we can write the integral

$$\int_{-1}^{1} \nu\phi(\mu, \nu)\psi(x, \nu)\, d\nu = \mu\left[\lambda(\mu)\psi(x, \mu) - P\int_{-1}^{1} \frac{\nu k(\nu, \mu)}{\nu - \mu}\, \psi(x, \nu)\, d\nu\right]. \quad (4.1a)$$

On the other hand, if we use the eigenfunction expansion (2.15) for $\psi(x, \nu)$ and then apply the orthogonality relations (2.16) and (2.17), we get

$$\begin{aligned}\int_{-1}^{1} \nu\phi(\mu, \nu)\psi(x, \nu)\, d\nu &= \int_{-1}^{1} \nu\phi(\mu, \nu)\left\{\sum_{j=1}^{\alpha}[a_{+j}\phi(z_j, \mu)e^{-x/z_j} \right. \\ &\quad \left. + a_{-j}\phi(-z_j, \mu)e^{x/z_j}] + \int_{-1}^{1} A(\eta)\phi(\eta, \nu)e^{-x/\eta}\, d\eta\right\} d\nu \\ &= \mu S(\mu)A(\mu)e^{-x/\mu}. \quad (4.1b)\end{aligned}$$

Equating the equalities of Eq. (4.1a) and (4.1b), we have

$$\lambda(\mu)\psi(x, \mu) - P\int_{-1}^{1} \frac{\nu k(\nu, u)}{\nu - u}\, \psi(x, \nu)\, d\nu = S(\mu)A(\mu)e^{-x/\mu}, \quad (4.2)$$

which when we compare with Eq. (3.10), we find

$$F(\mu) = -S(\mu)A(\mu) = [f(a, \mu) \text{ for the critical problem}]. \quad (4.3)$$

A similar procedure yields an expression for the discrete expansion coefficients:

$$\int_{-1}^{1} \nu\phi(\pm z_j, \nu)\psi(x, \nu)\, d\nu = \pm a_{\pm j} z_j M_j e^{\mp x/z_j}, \quad (4.4a)$$

or

$$\int_{-1}^{1} \frac{\nu k(\nu, \pm z_j)}{\nu \mp z_j}\, \psi(x, \nu)\, d\nu = -a_{\pm j} M_j e^{\mp x/z_j}. \quad (4.4b)$$

Letting $x = a$ in Eq. (4.3b) and recalling Eq. (3.12), we find for the critical problem

$$a_{\pm j} = -f(a, \pm z_j)/M_j. \quad (4.5)$$

Equations (4.3) and (4.5) reflect the relation between the two different methods and allow one to go from one representation to the other, i.e., substitution of Eqs. (4.3) and (4.5) into the Fredholm equation for the continuum expansion coefficient and the set of homogeneous equations for the discrete coefficients in the eigenfunction expansion method yields the critical condition indicated in the last section, provided the secular determinant of the above-mentioned equation vanished (which of course is another representation for the critical condition). The relationship between the transform method of Leonard and Mullikin[4] and the eigenfunction expansion method is obviously similar.

However, this relationship is more general. In fact, the development of Eq. (4.2) can be interpreted as a derivation of Eq. (3.10). It can also be used as an alternate procedure for obtaining the expansion coefficients in the eigenfunction expansion method. We illustrate this procedure by considering Milne's problem for the infinite half-space medium with the following boundary conditions:

$$\psi_m(0, \mu) = 0, \quad \mu > 0 \quad (4.6)$$

and

$$\lim_{x \to \infty} \psi_m(x, \mu) = B\phi(-z_k, \mu)e^{x/z_k}, \quad (4.7)$$

where $z_k$ is the largest positive eigenvalue and an arbitrary constant $B$ describes the strength of the source at infinity.

From the boundary condition at infinity (4.7), we get

$$a_{-j} = 0, \quad j \neq k. \quad (4.8)$$

We then write the solution as

$$\psi_m(x, \mu) = B\phi(-z_k, \mu)e^{x/z_k} + \sum_{j=1}^{\alpha} a_{+j}\phi(z_j, \mu)e^{-x/z_j} + \int_{-1}^{1} A_m(\nu)\phi(\nu, \mu)e^{-x/\nu}\, d\nu. \quad (4.9)$$

In a procedure similar to that used in Eq. (4.2), we obtain

$$\int_{-1}^{1} \nu\phi(\mu, \nu)\psi_m(x, \nu)\, d\nu = \mu S(\mu)A_m(\mu)e^{-x/\mu}, \quad (4.10)$$

from which we write [cf. Eq. (3.10)]

$$\lambda(\mu)\psi_m(x, \mu) - P\int_{-1}^{1} \frac{\nu k(\nu, \mu)}{\nu - \mu}\, \psi_m(x, \nu)\, d\nu = \begin{cases} -F(\mu)e^{-x/\mu}, & \mu > 0 \\ 0, & \mu < 0, \end{cases} \quad (4.11)$$

where

$$F(\mu) = -S(\mu)A_m(\mu), \quad (4.12)$$

and we have used the boundary condition (4.7) to specify that $A_m(\mu) = 0$, $\mu < 0$. Boundary condition (4.6) yields an expression $F(\mu)$:

$$F(\mu) = \int_{0}^{1} \frac{\nu k(\nu, -\mu)}{\nu + \mu}\, \psi_m(0, -\nu)\, d\nu. \quad (4.13)$$

We can analytically continue $\psi_m$ to the complex plane if we define

$$F(z) = \int_{0}^{1} \frac{\nu k(\nu, -z)}{\nu + z}\, \psi_m(0, -\nu)\, d\nu, \quad (4.14)$$

with which we write from Eq. (4.11)

$$\Omega(z)\psi_m(x, z) - \int_{-1}^{1} \frac{\nu k(\nu, z)}{\nu - z}\, \psi_m(x, \nu)\, d\nu = \begin{cases} -F(z)e^{-x/z}, & \operatorname{Re} z > 0 \\ 0, & \operatorname{Re} z < 0. \end{cases} \quad (4.15)$$

Now for $\psi_m(x, z)$ to be analytic on $0 < |z| < \infty$, $z \neq -z_k$, we must have

$$\int_{-1}^{1} \frac{\nu k(\nu, -z_j)}{\nu + z_j} \psi_m(x, \nu)\, d\nu = 0, \quad j \neq k \quad (4.16)$$

which when we let $x = 0$ and apply Eq. (4.5) we find

$$F(-z_j) = 0, \quad j \neq k. \quad (4.17)$$

Since $\psi_m(x, z)$ has a simple pole at $z = -z_k$, we find

$$F(-z_k) = 2B\Omega(\infty)X(-z_k)X(z_k)z_k^2 N(z_k) \prod_{j \neq k}^{\alpha} (z_j^2 - z_k^2), \quad (4.18)$$

where we have used the identity

$$\Omega(z) = X(z)X(-z)\Omega(\infty) \prod_{j=1}^{\alpha} (z_j^2 - z^2) \quad (4.19)$$

with

$$X(z) = \frac{1}{(1 - z)^{\alpha}} \exp \frac{1}{2\pi i} \int_0^1 \frac{\ln (\Omega^+(\mu)/\Omega^-(\mu))}{\mu - z}\, d\mu \quad (4.20)$$

as the solution of the homogeneous Hilbert problem whose boundary values along the cut (0, 1) satisfy

$$\frac{X^+(\mu)}{X^-(\mu)} = \frac{\Omega^+(\mu)}{\Omega^-(\mu)}. \quad (4.21)$$

An expression for $F(\mu)$ can be obtained by considering Eq. (4.11) with $x = 0$:

$$\lambda(\mu)\psi_m(0, -\mu) - P\int_0^1 \frac{\nu k(\nu, \mu)}{\nu - \mu} \psi_m(0, -\nu)\, d\nu = 0, \quad \mu > 0. \quad (4.22)$$

We then define the sectionally analytic function in the complex plane cut along (0, 1),

$$D(z) = \frac{1}{2\pi i} \int_0^1 \frac{\nu k(\nu, z)}{\nu - z} \psi_m(0, -\nu)\, d\nu. \quad (4.23)$$

With the use of Plemelj's formulas, we can then rewrite Eq. (4.21) as

$$\Omega^+(\mu)D^-(\mu)/\Omega^-(\mu) - D^+(\mu) = 0, \quad (4.24)$$

where $\Omega^{\pm}(\mu)$ are given by Eqs. (2.20).

Recalling Eq. (4.21), we have

$$D^+(\mu)/X^+(\mu) - D^-(\mu)/X^-(\mu) = 0. \quad (4.25)$$

We now consider the function

$$H(z) = D(z)/X(z) \quad (4.26)$$

which is analytic everywhere in the finite plane except perhaps for a cut along (0, 1). But from Plemelj's formulas and Eq. (4.25), it is obvious that $H^+(\mu) - H^-(\mu) = 0$, so that $H(z)$ is analytic everywhere in the finite plane. Since we are assuming $\Omega(\infty) \neq 0$, $k(\eta, z)$ is of degree $N$ in $z$. Therefore, from Eqs. (4.20) and (4.23) the behavior of $H(z)$ at infinity is that of a polynomial of degree $N + \alpha - 1$ at most. Hence, by Liouville's theorem, we find

$$D(z) = \frac{X(z)}{2\pi i} \sum_{l=0}^{N+\alpha-1} \beta_l z^l, \quad (4.27)$$

where the coefficients $\beta_l$ can be determined as outlined below. We note first that

$$F(z) = 2\pi i D(-z) = X(-z) \sum_{l=0}^{N+\alpha-1} \beta_l (-z)^l. \quad (4.28)$$

Therefore, from Eq. (4.17) we have

$$F(-z_j) = X(z_j) \sum_{l=0}^{N+\alpha-1} \beta_l z_j^l, \quad j \neq k, \quad (4.29)$$

so that we can write

$$F(z) = X(-z) \prod_{j=k}^{\alpha} (z_j + z) \sum_{n=0}^{N} \alpha_n z^n \quad (4.30)$$

with

$$\sum_{l=0}^{N+\alpha-1} \beta_l (-z)^l = \prod_{j \neq k}^{\alpha} (z_j + z) \sum_{n=0}^{N} \alpha_n z^n, \quad (4.31)$$

where now only the $N + 1$ $\alpha_n$ are to be determined. One of the equations necessary to do this comes from Eq. (4.18) which yields

$$\sum_{n=0}^{N} \alpha_n (-z_k)^n = 2B\Omega(\infty)X(-z_k)z_k^2 N(z_k) \prod_{j \neq k}^{\alpha} (z_j + z_k). \quad (4.32)$$

The other $N$ equations are found by eliminating the $N$ moments

$$U_j = \int_0^1 \nu^{j+1} \psi_m(0, -\nu)\, d\nu \quad (4.33)$$

from the $2N$ equations

$$\sum_{j=0}^{N-1} \omega_{ij}(y_i)(\pm 1)^{j+1} U_j = X(\pm y_i) \prod_{j \neq k}^{\alpha} (z_j \mp y_i) \sum_{n=0}^{N} \alpha_n (\mp y_i)^n, \quad (4.34)$$

where the $\pm y_i$ are the zeros of $N(z)$ and the polynomials $\omega_{ij}(\eta)$ are obtained from the relation

$$k(\nu, \pm y_i) = \sum_{j=0}^{N-1} \omega_{ij}(\pm y_i)(\pm 1)^{j+1} \nu^j (\nu \mp y_i). \quad (4.35)$$

With $F(z)$ so determined we have Eq. (4.12),

$$A(\mu) = \frac{X(-\mu)}{S(\mu)} \prod_{j \neq k}^{\alpha} (z_j + \mu) \sum_{n=0}^{N} \alpha_n \mu^n. \quad (4.36)$$

We also find

$$\int_{-1}^{1} \frac{\nu k(\nu, \mu)}{\nu - z_j} \psi_m(x, \nu)\, d\nu = -a_{+j} M_j e^{-\alpha/z_j}, \quad (4.37)$$

which with Eq. (4.14) and $x = 0$ yields

$$a_{+j} = -F(z_j)/M_j. \quad (4.38)$$

## 5. WEIGHT FUNCTIONS

It seems appropriate to conclude with a few brief remarks about the role adjoint singular integral equations play in obtaining weight functions for use in orthogonality relations between the eigenfunctions of the transport equation. The central point here is that a necessary condition is the weight function be a solution to a singular integral equation which is adjoint to the singular integral equations arising in the eigenfunction expansion method. For example, in the case of isotropic scattering, Eq. (2.4) takes the form

$$(1 - \mu/\nu)\phi(\nu, \mu) = c/2 \tag{5.1a}$$

or likewise

$$(1 - \mu/\eta)\phi(\eta, \mu) = c/2. \tag{5.1b}$$

We look for the appropriate weight function $\mu g(\mu)$ for orthogonality on $(-1, 1)$ by multiplying Eq. (5.1a) by $g(\mu)\phi(\eta, \mu)$ and Eq. (5.1b) by $g(\mu)\phi(\nu, \mu)$ and integrating on $\mu$ over $(-1, 1)$ to find

$$2(1/\nu - 1/\eta)\int_{-1}^{1} \mu\phi(\nu, \mu)\phi(\eta, \mu)g(\mu)\, d\mu = c\int_{-1}^{1} g(\mu)\phi(\eta, \mu)\, d\mu - c\int_{-1}^{1} g(\mu)\phi(\nu, \mu)\, d\mu. \tag{5.2}$$

If $\mu g(\mu)$ is the appropriate weight function, the right side of Eq. (5.2) must vanish. Using the explicit form of $\phi$, we have

$$g(\eta)\lambda(\eta) - P\int_{-1}^{1} \frac{c}{2}\frac{\eta g(u)}{\mu - \eta}\, d\mu = \text{const} \quad (\text{e.g.} = 1), \tag{5.3}$$

which is the adjoint singular integral equation having the solution $g(\mu) = 1$. In this case, the weight function is $\mu$ in agreement with the previous results of Case.[2] If, on the other hand, we wish the orthogonality to be on the interval (0, 1), we find, instead of Eq. (5.2), the equation

$$2(1/\nu - 1/\eta)\int_{0}^{1} \mu\phi(\nu, \mu)\phi(\eta, \mu)g(\mu)\, d\mu = c\int_{0}^{1} g(\mu)\phi(\eta, \mu)\, d\mu - c\int_{0}^{1} g(\mu)\phi(\nu, \mu)\, d\mu, \tag{5.4}$$

which yields analogous to Eq. (5.3) the adjoint singular integral equation

$$g(\eta)\lambda(\eta) - P\int_{0}^{1} \frac{c}{2}\frac{\eta g(\eta)}{\mu - \eta}\, d\mu = \text{const}, \tag{5.5}$$

which has the solution

$$\mu g(\mu) = (z_1 - \mu)[X^+(\mu) - X^-(\mu)]/2\pi i,$$

in agreement with the results of Kuščer, McCormick, and Summerfield.[9]

The case of anisotropic scattering is more involved. In this case the equation analogous to Eq. (5.2) for determining the weight function for orthogonality on $(-1, 1)$ has the form

$$(1/\nu - 1/\eta)\int_{-1}^{1} \mu\phi(\nu, \mu)\phi(\eta, \mu)g(\mu)\, d\mu = \int_{-1}^{1} g(\mu)k(\mu, \nu)\phi(\eta, \mu)\, d\mu - \int_{-1}^{1} g(\mu)k(\mu, \eta)\phi(\nu, \mu)\, d\mu, \tag{5.6}$$

which yields the adjoint singular integral equation

$$P_n(\eta)g(\eta)\lambda(\eta) - P\int_{-1}^{1} \frac{\eta P_n(\mu)k(\mu, \eta)}{\mu - \eta}\, g(\mu)\, d\mu = h_n(\eta), \tag{5.7}$$

which has the solution $g(\mu) = 1$. This again gives the weight function $\mu$ in agreement with the results of Mika.[3] For orthogonality on (0, 1) the above procedure leads to rather intractable results. Perhaps a better procedure is the use of the bi-orthogonality results of McCormick and Kuščer.[6] Here the weight function $\mu g(\mu)$ satisfies the dominant adjoint singular integral equation

$$g(\eta)\lambda(\eta) - P\int_{0}^{1} \frac{\eta N(\mu)}{\mu - \eta}\, g(\mu)\, d\mu = 1, \tag{5.8}$$

which has the solution

$$\mu g(\mu) = [\Omega(\infty)]^{\frac{1}{2}} \prod_{j=1}^{\alpha} (z_j - \mu)[X^+(\mu) - X^-(\mu)]/2\pi i.$$

[9] I. Kuščer, N. J. McCormick, and G. C. Summerfield, Ann. Phys. (N.Y.) **30**, 411 (1961).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 5 MAY 1968

# Two-Dimensional Measure-Preserving Mappings

JOHN M. GREENE
*Plasma Physics Laboratory, Princeton University, Princeton, New Jersey*

(Received 21 June 1967)

A particular area-preserving mapping of a plane onto itself has been studied in detail with the aid of a digital computer. A large number of fixed points, finite sets of points that transform into each other, were located and classified as elliptic or hyperbolic depending on the nature of the linearized mapping in the neighborhood. A quantity called the residue was calculated for each fixed point. This quantity can be used to predict whether other nearby fixed points are elliptic or hyperbolic. The results showed that there are considerable regions in which almost all the fixed points are hyperbolic. Further calculations were made to estimate the area enclosed by the invariant curves whose existence has been established by Moser. The boundary of this region appeared to coincide with the boundary of the region in which almost all the fixed points are hyperbolic.

## I. INTRODUCTION

Problems in many branches of physics can be reduced to the study of two-dimensional measure-preserving mappings. A most pictorial example is the behavior of magnetic field lines in a toroidal system.[1] A mapping of a cross section of the toroid onto itself is formed by following magnetic lines of force around the system from one intersection with the cross section to the next. The possible containment of a line of force within a given region for many traversals around the system can be deduced from the behavior of iterates of the corresponding two-dimensional mapping. The magnetic flux through each neighborhood is conserved by this mapping, so it is measure preserving.

A more typical example is a particle in a two-dimensional potential.[2-4] Conservation of energy restricts the motion to a three-dimensional hypersurface in the four-dimensional phase space. Any phase variable whose value recurs in the course of the motion can be used to define a two-dimensional cross section. Dynamical orbits are used to define a mapping of this surface onto itself, analogous to that described in the example above. Poincaré's invariant[5] shows that area, defined in canonical coordinates, is conserved by this mapping.

The restricted three-body problem of celestial mechanics can also be reduced to such a mapping.[6,7]

It is most interesting to know whether these two-dimensional measure-preserving mappings are ergodic or not. This means determining whether or not successive images of a given point cover densely a finite area.[8] It appears that both types of behavior occur in different regions for almost every mapping. Moser[9,10] has shown that there are isolated closed invariant curves, subject to certain restrictions detailed in Sec. III. These exact invariant curves separate regions of ergodic behavior. It is interesting that similar invariants exist for higher-dimensional mappings but they do not separate different regions of phase space.[9] The two-dimensional problem appears to be the most interesting.

Moser's paper proved the existence of invariant curves but gave no indication of the area that they might be expected to enclose. The purpose of the work described in this paper was to find some way of estimating this area.

The estimation was sought in terms of the fixed points of the mapping. These correspond to closed magnetic field lines, or to periodic orbits in a dynamical problem. Such orbits may close after one, or many, traversals through the chosen cross section. These periodic orbits are attractive for study because they are finite in length, and because there are many of them. Poincaré[2] hypothesized that they are dense in phase space, but that has not been proven either way. In any event, they are thickly scattered. The hope is that all the properties of a mapping can be understood in terms of its fixed points.

The location of the fixed points is, of course, not enough. Each fixed point must also be characterized by the linearized mapping in its immediate vicinity. It is the parameters of this linearized mapping that have been studied and in terms of which the properties of the mapping have been interpreted in this paper.

[1] L. Spitzer, Jr., Phys. Fluids **1**, 253 (1958).
[2] G. D. Birkhoff, *Collected Mathematical Papers* (American Mathematical Society, Providence, R.I., 1950), Vol. II, p. 333.
[3] M. Hénon and C. Heiles, Astron. J. **69**, 73 (1964).
[4] P. A. Sturrock, Ann. Phys. (N.Y.) **3**, 113 (1958).
[5] H. Goldstein, *Classical Mechanics* (Addison-Wesley Publ. Co., Reading, Mass., 1951), p. 247.
[6] J. Moser, Nachr. Akad. Wiss. Göttingen, II Math.-Physik. Kl., No. 6, 87 (1955).
[7] V. I. Arnol'd, Russian Math. Surveys **18**, No. 6, 85 (1963).
[8] D. R. Halmos, *Lectures on Ergodic Theory* (Mathematical Society of Japan, Tokyo, 1956), p. 25.
[9] J. Moser, *Nonlinear Problems*, R. E. Langer, Ed. (University of Wisconsin Press, Madison, Wisconsin, 1963), p. 139.
[10] J. Moser, Nachr. Akad. Wiss. Göttingen, II Math.-Physik. Kl., No. 1 (1962).

These parameters are defined and discussed in Sec. II.

In Sec. III a scheme is described for selecting certain fixed points for study.

A number of fixed points, for a particular simple choice of mapping, have been located numerically with the aid of a digital computer and the corresponding parameters have been evaluated. These computations are described in Sec. IV.

These computations indicate that there are regions in which all the fixed points are hyperbolic. The image of any given point in the neighborhood of such a point tends to move away from the fixed point on successive iterations of the mapping. Ergodic behavior appears to be restricted to these regions. This is discussed in more detail in Sec. V.

## II. DEFINITIONS

The mappings considered in this paper have the general form

$$x_1 = f(x_0, y_0), \quad y_1 = g(x_0, y_0), \tag{1}$$

or, more compactly,

$$(x_1, y_1) = T(x_0, y_0). \tag{2}$$

The functions $f$ and $g$ will be taken to be infinitely differentiable. The condition that the mapping be area preserving is

$$\det (\mathsf{J}) = 1, \tag{3}$$

where $\mathsf{J}$ is the Jacobian matrix of the partial derivatives,

$$\mathsf{J}(x, y) = \begin{pmatrix} f_x & f_y \\ g_x & g_y \end{pmatrix}. \tag{4}$$

We will consider also iterations of the mapping,

$$(x_1, y_1) = T(x_0, y_0), \quad (x_2, y_2) = T(x_1, y_1), \tag{5}$$

which can be written

$$(x_2, y_2) = T^2(x_0, y_0), \tag{6}$$

and generalized to any power of $T$.

It is frequently useful to linearize the mapping around a given initial point. This approximation can be expressed in matrix form using products of the Jacobian matrix of Eq. (4):

$$\begin{aligned} \begin{pmatrix} \hat{x}_2 - x_2 \\ \hat{y}_2 - y_2 \end{pmatrix} &= \mathsf{J}(x_1, y_1) \begin{pmatrix} \hat{x}_1 - x_1 \\ \hat{y}_1 - y_1 \end{pmatrix} \\ &= \mathsf{J}(x_1, y_1)\mathsf{J}(x_0, y_0) \begin{pmatrix} \hat{x}_0 - x_0 \\ \hat{y}_0 - y_0 \end{pmatrix}, \end{aligned} \tag{7}$$

where $(x_0, y_0)$, $(x_1, y_1)$, and $(x_2, y_2)$ are related by Eq. (5).

Special attention will be focused on fixed points. The fixed points of $T$ satisfy the equation

$$(x_f, y_f) = T(x_f, y_f). \tag{8}$$

The matrix representing the linearized mapping in the vicinity of this fixed point will be denoted by $\mathsf{M} \equiv \mathsf{J}(x_f, y_f)$. It is convenient to introduce four parameters $a$, $b$, $c$, and $d$,

$$\mathsf{M} \equiv \begin{pmatrix} a + d & c + b \\ c - b & a - d \end{pmatrix}, \tag{9}$$

in discussing such $2 \times 2$ matrices. Equation (3) then becomes

$$a^2 + b^2 - c^2 - d^2 = 1. \tag{10}$$

The quadratic form,

$$\begin{aligned} \psi \equiv (b - c)(x - x_f)^2 + 2d(x - x_f)(y - y_f) \\ + (b + c)(y - y_f)^2 \end{aligned} \tag{11}$$

is invariant under this linearized mapping. Three of the four independent parameters that will be used to specify $\mathsf{M}$ describe this quadratic. One of these is an ellipticity parameter,

$$E \equiv \frac{b^2 - c^2 - d^2}{b^2 + c^2 + d^2} = \frac{1 - a^2}{b^2 + c^2 + d^2}, \tag{12}$$

which is somewhat more convenient than the classical ellipticity. When $E = 1$ the constant-$\psi$ surfaces are circles, $E > 0$ corresponds to a set of ellipses, $E = 0$ to a set of parallel lines, and $E < 0$ to hyperbolas. Fixed points will be called elliptic or hyperbolic, depending on the sign of $E$. When the denominator of Eq. (12) vanishes, $b = c = d = 0$, $a = \pm 1$. This will be called the degenerate case.

Another parameter describes the orientation of the quadratic form. Under a rotation of coordinates, $a$ and $b$, and hence $c^2 + d^2$, are constant, but $c$ and $d$ vary as the sine and cosine of twice the angle. Thus we define

$$\theta_M \equiv \tfrac{1}{2} \tan^{-1} (c/d). \tag{13}$$

The sign of $b$ is necessary to reconstruct $\mathsf{M}$ from the full set of parameters. It also affects the orientation of the quadratic form.

The remaining important characteristic of the matrix $\mathsf{M}$ is the displacement of a point along an invariant curve. This can be deduced from a quantity that will be called the residue of the fixed point, defined by

$$\begin{aligned} R &\equiv \tfrac{1}{2} - \tfrac{1}{4} \operatorname{Tr} (\mathsf{M}), \\ &= \tfrac{1}{2}(1 - a). \end{aligned} \tag{14}$$

The interpretation of the residue depends on whether the fixed point is elliptic or hyperbolic. Indeed, the

sign of $E$ is related to the magnitude of $R$, since, from Eqs. (12) and (14),

$$E = 4R(1 - R)/(b^2 + c^2 + d^2). \tag{15}$$

For hyperbolic points the eigenvalues of the matrix M are the ratios of the distances of a point and its image when both lie on an asymptote of Eq. (11), and they depend only on $R$:

$$\lambda = 1 - 2R \pm 2[R(R - 1)]^{\frac{1}{2}}. \tag{16}$$

Note that when $R$ is small and negative $\lambda$ is nearly unity and the distance between a point and its image is small.

The behavior of elliptic points is best understood from an examination near the fixed point of iterates of the mapping $T$. The Jacobian matrix of $T^Q$ evaluated at the fixed point of $T$ is

$$\mathsf{M}_Q(x_f, y_f) = \mathsf{M}^Q = (\sin 2\pi Q\omega_0)/(\sin 2\pi\omega_0) \times \begin{pmatrix} \sin 2\pi\omega_0 \times \cot 2\pi Q\omega_0 + d & b + c \\ c - b & \sin 2\pi\omega_0 \times \cot 2\pi Q\omega_0 - d \end{pmatrix}, \tag{17}$$

where

$$\cos 2\pi\omega_0 \equiv a, \quad R = \sin^2 \pi\omega_0, \tag{18}$$

which can easily be established by induction. When $Q\omega_0 \sim 1$, $\mathsf{M}^Q$ is approximately the identity, and a given point has returned to nearly its original position. Thus $\omega_0$, which will be called the rotation number, is the average rotation of a point around the ellipse. As with hyperbolic points, small values of the residue correspond to small displacements.

Note that the residue is always real, while $\lambda$ is complex for elliptic points and $\omega_0$ is complex for hyperbolic points.

It is obvious that, aside from the singular case when $E$ vanishes and $b \neq 0$ must be specified, the matrix M is completely determined by $R$, $E$, $\theta_{\mathsf{M}}$, and sign of $b$. Namely,

$$\begin{aligned} a &= 1 - 2R, \\ b &= \pm[2R(1 - R)(1 + E)/E]^{\frac{1}{2}}, \\ c &= \sin 2\theta_{\mathsf{M}}[2R(1 - R)(1 - E)/E]^{\frac{1}{2}}, \\ d &= \cos 2\theta_{\mathsf{M}}[2R(1 - R)(1 - E)/E]^{\frac{1}{2}}. \end{aligned} \tag{19}$$

Henceforth these parameters will be used to describe M.

It follows from Eq. (15) that $0 < R < 1$ always indicates an elliptic fixed point, while if $R < 0$ or $R > 1$ the fixed point is hyperbolic. Equation (16) shows that there are two types of hyperbolic points,[11] depending on the sign of the residue. When it is positive, the eigenvalues $\lambda$ are negative and M interchanges the corresponding branches of the hyperbolas of Eq. (11). When $R$ is negative, the eigenvalues are positive and M maps each branch into itself. Thus, for small residues, the distinction between positive and negative residues corresponds to the distinction between elliptic and hyperbolic fixed points, whereas for large residues it corresponds to the difference between two types of hyperbolic points.

A quantity called the index[12] can be defined for closed curves and is useful for counting and classifying the enclosed fixed points. Consider the set of vectors connecting each point of a closed curve with its image under the mapping. The index of the curve is the number of rotations the vector makes as the initial point traverses the closed curve, taken to be positive if the rotation is in the same sense as the traversal of the curve. Similarly, the index of a discrete fixed point is defined as the index of an enclosing curve that lies in the immediate neighborhood. It can readily be established by repeated subdivision that the index of a closed curve is the sum of the indices of the enclosed fixed points. It can also be established that the index of a fixed point is equal to the sign of the residue, when the latter does not vanish.

Now consider iterations of the mapping, and particularly the fixed points of $T^2$, $T^3$, etc. If $(x_f, y_f)$ is a fixed point of $T^Q$, then $T(x_f, y_f)$ is also a fixed point. Thus the fixed points of $T^Q$ form families of $Q$ members. The quantity $Q$ is useful as an identification of a given fixed point. Again, matrices can be found that represent the linearized mapping in the vicinity of each of these fixed points. These can be expressed as products of $Q$ Jacobian matrices, Eq. (4), evaluated respectively at the successive fixed points of the corresponding family, following Eq. (7). Such matrices can be analyzed in exactly the same manner as discussed above. It follows that each fixed point of a given family has the same residue since

$$\mathsf{M}[T(x_f, y_f)] = \mathsf{J}(x_f, y_f)\mathsf{M}(x_f, y_f)\mathsf{J}^{-1}(x_f, y_f). \tag{20}$$

A family of fixed points tends to lie on a ring, or a series of rings enclosing elliptic fixed points. The indices under $T^Q$ of closed curves lying just inside and just outside such rings are almost always $+1$. Thus there must be as many families with positive as with negative residue. As we have seen above, this may

[11] G. D. Birkhoff, Ref. 2, p. 111.
[12] H. Poincaré, *Oeuvres* (Gauthier-Villars et Cie., Paris, 1928), Vol. 1, p. 85.

mean that elliptic and hyperbolic points alternate, but it does not exclude all points being hyperbolic.

## III. A HIERARCHY

This section presents a scheme for obtaining an over-all picture of the structure of any given mapping. The approximate invariant curves given by Eq. (11) serve as the starting point. These can be extended in a qualitative way into the nonlinear region where they provide a framework to discuss the mapping.

Near the fixed point the deviation from the linearized mapping of Eq. (7) can be treated as a small parameter. It is thus possible to extend Eq. (11) and obtain a formal series[11] for the approximate invariant curves $\psi = \text{const}$ in powers of $x - x_f$ and $y - y_f$. Our primary interest is in those curves that enclose elliptic fixed points. The rotation number $\omega$, defined for these closed curves as the average number of rotations around the central fixed point per iteration of the mapping in the limit as the number of iterations goes to infinity, depends on the value of $\psi$ when the mapping is approximated using higher powers of the distance from the fixed point. For surfaces very close to the central fixed point $\omega$ approaches $\omega_0$ of Eq. (18).

The shear of the mapping, defined by

$$s = d\omega/d\psi, \tag{21}$$

is very useful in classifying different mappings. For example, higher-order terms in the expansion of $\psi$ discussed above may only slightly alter the shape of the curves or may change their character completely, depending on the shear. In particular, if the rotation number $\omega_0$ of a fixed point is close to a rational number with a small denominator, and the shear is higher order than resonant nonlinear terms, the approximate invariant curves may no longer be closed.[3,8,9]

Moser[9,10] has shown that when the shear is of lower order than any resonant perturbation there are exact invariant curves enclosing the central fixed point. These invariant curves have irrational rotation numbers that satisfy the inequality[13]

$$|\omega - l/k| > \epsilon/k^{\frac{5}{2}} \tag{22}$$

for all integers $l$ and $k$, where $\epsilon$ is a constant independent of $k$.

In between these surfaces, where $\omega = P/Q$ ($P$ and $Q$ relatively prime) would be rational if it varied continuously, there must be fixed points of $T^Q$ according to a theorem of Birkhoff's.[14,15] In certain special integrable cases these fixed points may be dense on a curve and so complete a set of nested invariant curves. In such cases the residue of each fixed point vanishes. However, in almost every situation perturbations break up the surface; there are only isolated fixed points of $T^Q$ and they have nonvanishing residue.

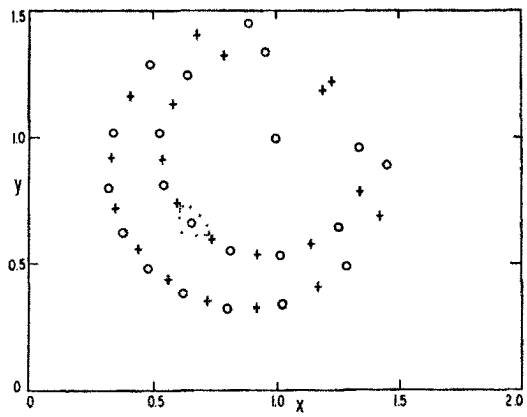



FIG. 1. Some fixed points of the mapping given by Eqs. (23)–(25). Two families for each of $P/Q = \frac{1}{6}$ and $\frac{1}{11}$ are shown enclosing the fixed point at (1, 1). Positive residue points are represented by ○ and negative residue point by +. Also shown by · are 20 points of a family, for which $Q = 180$, belonging to the third order of hierarchy and associated with the elliptic points with $P/Q = \frac{1}{6}$.

These isolated fixed points can be identified by their rotation number $P/Q$. It is possible to establish a rotation number for this finite set of points because the exact and approximate invariant curves impose considerable organization on any given mapping. Each family of fixed points will appear to lie on a closed curve that is close to an approximate invariant curve and encloses a central fixed point. A typical example is shown in Fig. 1. Thus $P - 1$ is the number of members of the family which lie between a fixed point and its image. Alternatively, $P$ is the number of rotations around the central point when following from a fixed point to its image through all $Q$ members of the family. In either case $P$ is ambiguous to an additive multiple of $Q$. A further requirement on $P$ is that two fixed points which lie close to each other should have nearly equal values of $P/Q$, which follows from the continuity of the approximate organization of the mapping.

The structure of invariant closed curves and fixed points enclosing an elliptic fixed point of $T^Q$ is similar to that enclosing a fixed point of $T$. The additional fixed points in this structure are then fixed points of $(T^Q)^N$. Their rotation number around the central fixed point of $T$ is the same as that of the fixed point of $T^Q$ with which they are associated. Thus the values of $\hat{P}$ and $\hat{Q}$ ($\omega = \hat{P}/\hat{Q}$, $\hat{Q} = QN$) are not relatively prime for these fixed points. This structure may also exist around hyperbolic points in certain cases.

[13] I. Niven, *Irrational Numbers* (Mathematical Association of America, Menasha, Wisc., 1956).
[14] G. D. Birkhoff, Ref. 2, Vol. I, p. 673.
[15] G. D. Birkhoff, Ref. 2, Vol. II, p. 252.

In order to avoid dealing with all the fixed points and invariant surfaces as a single unit, it is convenient to organize this structure recursively in a hierarchy. The zero order of this hierarchy consists of the fixed points of $T$. The first order consists of the invariant curves that enclose the fixed points of $T$ and those fixed points of $T^Q$ for which $P$ and $Q$ are relatively prime. The second order of the hierarchy includes this same structure around elliptic fixed points in the first order of the hierarchy. It is clear that there are an infinite number of orders in the hierarchy.

Some such organization into a hierarchy is necessary for making general statements. For example, in the next section a connection is found between the invariant curves of a given order of the hierarchy and the fixed points in the same order. If the fixed points of higher orders were not discarded it would be difficult to discover any meaningful relationships. However, it has not yet been possible to develop a rigorous definition of the concept. In particular, it is not clear that all fixed points can be ordered in the hierarchy.

## IV. COMPUTATIONS

### A. Description of the Chosen Mapping

The specific mapping chosen for study is

$$x_1 = (1/A)(x_0 - By_0 + Cy_0^3), \tag{23}$$

$$y_1 = Ay_0 + Bx_1 - Cx_1^3, \tag{24}$$

$$A = 1.25, \quad B = 0.1, \quad C = 0.35. \tag{25}$$

This mapping is symmetric around the line $x = y$ in the sense that if

$$T(c_1, c_2) = (c_3, c_4), \tag{26}$$

then

$$T(c_4, c_3) = (c_2, c_1). \tag{27}$$

It follows that if $(c_1, c_2)$ is a fixed point of $T^Q$, $(c_2, c_1)$ is another fixed point of $T^Q$. The mapping is also invariant under change of sign of both $x$ and $y$. (See note added in proof.) The fixed points of $T$ consist of a hyperbolic point at (0, 0) and elliptic points at (1, 1) and (−1, −1). The residue of the hyperbolic point is $R = -0.0105$. For the elliptic points $R = 0.168$ and $\omega_0^{-1} = 7.4388$.

The fixed point at (0, 0) belongs to the zero order of the hierarchy considered in the previous section. The first order of the hierarchy is composed of families of fixed points of iterates of $T$ which enclose the origin. The two elliptic fixed points of $T$ can be considered as the innermost of these families with a $P/Q$ value of $\frac{0}{1}$. These families extend continuously in $P/Q$ from this value to $\frac{1}{2}$ for fixed points of $T^2$ at $[\pm(2.35/0.35)^{\frac{1}{2}}, \mp(2.35/0.35)^{\frac{1}{2}}]$. This is a case where the central point is hyperbolic.

Most attention was concentrated on the second order of the hierarchy, the structure around the fixed points of the first order, in particular, around the elliptic fixed points at (1, 1). There are pairs of families of fixed points of iterates of $T$ for all values of $P$ and $Q$ that are relatively prime and whose ratio lies in the range $0 < P/Q < \omega_0$. Two such pairs of families are shown in Fig. 1.

By the symmetry of Eqs. (26) and (27), one member of most of these families lies on the line $x = y$. Such points are easy to locate numerically, so effort was concentrated on corresponding families.

Surrounding each positive-residue fixed point in the second order of the hierarchy are families of fixed points belonging to the third order of the hierarchy. A part of one such family is shown in Fig. 1.

### B. Calculation of Residues

Residues were calculated for approximately 150 families of fixed points, including all those for which $P \leq 11$, $P/Q \geq (9.5)^{-1}$, and $|R| > 10^{-14}$. A selected set are listed in Table I. It can be seen from Eq. (14) that small residues are calculated as the difference of two numbers of order 1. For fixed points with the smallest residues it was necessary to carry 14 to 18 decimal places in all the calculations. The two families of fixed points for a given $P/Q$ had residues that were very nearly equal and opposite, especially when they were small in magnitude. This would not be true for fixed points with $Q \leq 4$ but none such occurred in this system.

The behavior of the residues in the limit of large $Q$ can be understood by assuming that they depend exponentially on $Q$. This leads to defining a function

TABLE I. $Q/P$, location, and residue for a selected set of fixed points.

| $Q/P$ | $Q/P$ | $x = y$ | $R$ | $f = (4R)^{2/Q}$ |
|---|---|---|---|---|
| 15/2 | 7.5 | 0.940 | $-1.34 \cdot 10^{-10}$ | 0.0581[a] |
| 79/10 | 7.9 | 0.823 | $-3.2 \cdot 10^{-14}$ | 0.472 |
| 8/1 | 8.0 | 0.805 | $-1.08 \cdot 10^{-2}$ | 0.456 |
| 17/2 | 8.5 | 0.727 | $-2.25 \cdot 10^{-2}$ | 0.753 |
| 53/6 | 8.833 | 0.6948 | −0.116 | 0.971 |
| 115/13 | 8.846 | 0.6944 | 0.115 | 0.987 |
| 62/7 | 8.857 | 0.6942 | −0.254 | 1.0005 |
| 133/15 | 8.867 | 0.6940 | 0.593 | 1.013 |
| 71/8 | 8.875 | 0.6939 | −0.590 | 1.024 |
| 151/17 | 8.882 | 0.69382 | 3.31 | 1.035 |
| 80/9 | 8.889 | 0.69377 | −1.41 | 1.044 |
| 9/1 | 9.0 | 0.658 | 0.229 | 0.980 |
| 82/9 | 9.111 | 0.621 | −65.9 | 1.145 |
| 28/3 | 9.333 | 0.611 | −1.04 | 1.107 |
| 19/2 | 9.5 | 0.599 | −0.888 | 1.143 |

[a] The positive residue family with the same value of $Q/P$ has been used to calculate $f$.

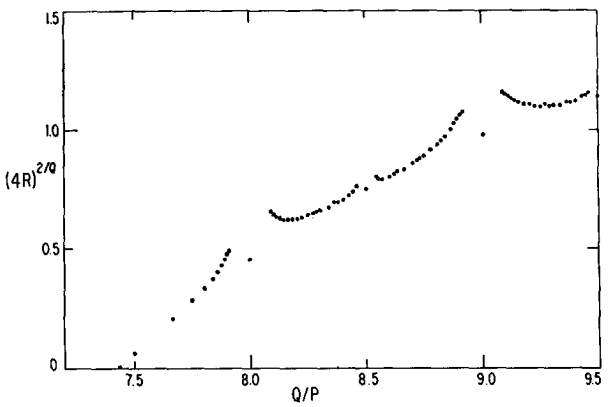



FIG. 2. A plot of $f(Q/P)$, defined in Eq. (28).

$f$ by

$$R \equiv \alpha[f(P/Q)]^{Q/2}. \tag{28}$$

A plot of $f$ vs $Q/P$ is given in Fig. 2, and selected values are listed in Table I, where $\alpha$ has been chosen to be equal to $\frac{1}{4}$ to make the curve as smooth as possible.

If $f$ were continuous it would be possible to extrapolate the calculated residues to large values of $Q$ by interpolating in $f$. In regions with $f$ greater than unity, the residues would increase indefinitely with $Q$, and with $f$ less than unity the residues would approach zero in this limit.

It is clear that Eq. (28) is a necessary first step to obtain a smooth curve suitable for interpolation, but it does not quite do the job. Apparently each fixed point creates a perturbation in its neighborhood which affects the value of $f$ for neighboring points. This is especially evident in Fig. 2 near where $Q/P$ equals 8, $8\frac{1}{2}$, and 9.

This perturbation is illustrated in Fig. 3. The location of fixed points with $Q/P = \frac{82}{9} = 9\frac{1}{9}$ are strongly

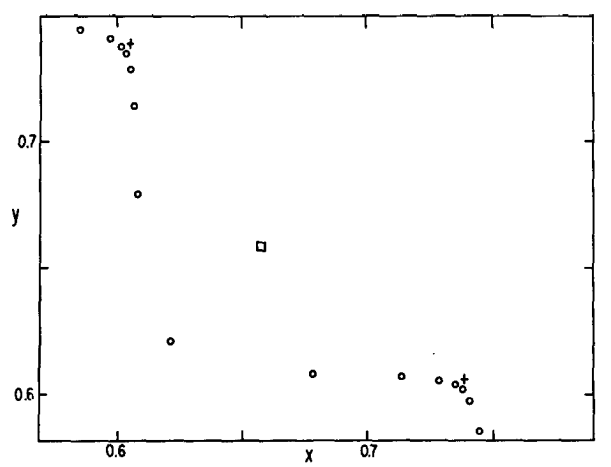



FIG. 3. A plot showing the relation between fixed points with $P/Q = \frac{9}{82}$ and nearby fixed points with $P/Q = \frac{1}{9}$. The elliptic $\frac{1}{9}$ fixed point is denoted by □, the corresponding hyperbolic fixed points by +, and part of one family of $\frac{9}{82}$ fixed points by ○.

TABLE II. Residues of fixed points near $Q/P = \frac{17}{2}$.

| $n$ | $Q/P$ | $R$ | $f = (4R)^{2/Q}$ |
|---|---|---|---|
| 3 | 59/7 | $1.83\ 10^{-5}$ | 0.724 |
| 7 | 127/15 | $4.56\ 10^{-8}$ | 0.783 |
| 11 | 195/23 | $2.97\ 10^{-9}$ | 0.829 |
| 15 | 263/31 | $1.49\ 10^{-9}$ | 0.866 |
| 19 | 331/39 | $2.21\ 10^{-9}$ | 0.894 |
| 23 | 399/47 | $5.63\ 10^{-9}$ | 0.916 |
| 27 | 467/55 | $1.65\ 10^{-8}$ | 0.932 |

affected by their proximity to fixed points with $Q/P = \frac{9}{1}$. It is reasonable that this distortion in location is accompanied by an effect on the residues, though no direct calculation is available. It is also reasonable that this distortion should depend on the ellipticity parameter $E$ of the fixed points with $Q/P = \frac{9}{1}$ since this gives the shape of the invariant curves in the vicinity.

To further illustrate the effect of distortions caused by nearby fixed points, residues were calculated for a series of fixed points with values of $Q/P$ given by the formula

$$Q/P = 8 + (2 + 1/n)^{-1}. \tag{29}$$

In Fig. 2, this series approaches $Q/P = 8.5$ from the left. These calculated residues are listed in Table II. It appears that $f$ will increase beyond unity and that for large values of $n$ the residue will become indefinitely large. This confirms that the disturbance around the fixed points $Q/P = \frac{17}{2}$ dominates the behavior of nearby fixed points. Since there is nothing special about this fixed point, it can be said with confidence that there are intervals in every range of $P/Q$ where $f$ is larger than unity.

A series of rational numbers that avoids these disturbances can be constructed with the aid of the series of numbers $F_n$,

$$F_1 = 1, \quad F_2 = 1 + \frac{1}{1}, \quad F_3 = 1 + \frac{1}{1 + 1/1}, \cdots; \tag{30}$$

the partial convergents of the infinite continued fraction with the partial quotients all unity. Table III

TABLE III. Residues of certain fixed points.

| $n$ | $Q/P$ | $R$ | $f = (4R)^{2/Q}$ |
|---|---|---|---|
| 1 | 53/6 | $1.15\ 10^{-1}$ | 0.97126 |
| 2 | 62/7 | $-2.54\ 10^{-1}$ | 1.00050 |
| 3 | 115/13 | $1.15\ 10^{-1}$ | 0.98660 |
| 4 | 177/20 | $1.22\ 10^{-1}$ | 0.99200 |
| 5 | 292/33 | $-5.54\ 10^{-2}$ | 0.98973 |
| 6 | 469/53 | $2.66\ 10^{-2}$ | 0.99048 |
| 7 | 761/86 | $5.77\ 10^{-3}$ | 0.99014 |
| 8 | 1230/139 | $-6.03\ 10^{-4}$ | 0.99025 |

is similar to Table II except that the values of $P/Q$ are chosen by the formula

$$Q/P = 9 - (5 + F_n)^{-1}. \tag{31}$$

In this series the values of $f$ are rapidly converging on a value slightly less than unity, and the residues are decreasing.

In conclusion, there appear to be significant regions where all the fixed points are hyperbolic. This is evident from the following argument. In Table II it was illustrated that values of $f$ increase with increasing proximity to other fixed points. This proximity can be measured in terms of the partial quotients of a partial fraction expansion of $P/Q$. Thus values of $f$ increase with increasing partial quotients. Then the opposite case was considered and a series with the smallest possible partial quotients was given in Table III. In this series $f$ converged on a value that fits smoothly into the series of points in Fig. 2. Thus a curve sketched through the points of Fig. 2 gives a lower bound to the values of $f$. This lower bound will be labeled $f_m(Q/P)$, a continuous function of $Q/P$. In regions where $f_m > 1$ all families of fixed points will have $f$ greater than unity, and by Eq. (28) almost all positive residue fixed points will have $R > 1$. By Eq. (15) and the discussion below Eq. (12), this is equivalent to the statement that almost all fixed points are hyperbolic in regions where $f_m > 1$.

### C. Escape Calculations

Another type of calculation that was made was to test experimentally whether a given point lay inside or outside the outer most invariant curve of the type whose existence Moser has proved. A rigorous test was devised by noting that, for the chosen mapping, any invariant curve must be symmetric around the lines $x = y$ and $x = -y$. Since invariant curves cannot intersect, any point whose image under multiple iteration crosses the line $x = -y$ must lie outside any Moser surface enclosing the fixed point at (1, 1) but not the one at (0, 0). Hence the point must lie outside all the Moser surfaces in the second order of the hierarchy. Of course, it is impossible to prove that a point lies inside a Moser surface by a finite number of iterations. A maximum of 50 000 was used.

A more annoying difficulty was the loss of significant figures, resulting from roundoff errors, when many iterations were calculated. The results did not meet the usual tests of accuracy after a few hundred to a thousand iterations were performed. However, these tests are more stringent than are required. They require that if the image of a given point $(x_0, y_0)$ under the exact mapping $T^n$ is $(x_n, y_n)$, then the calculated

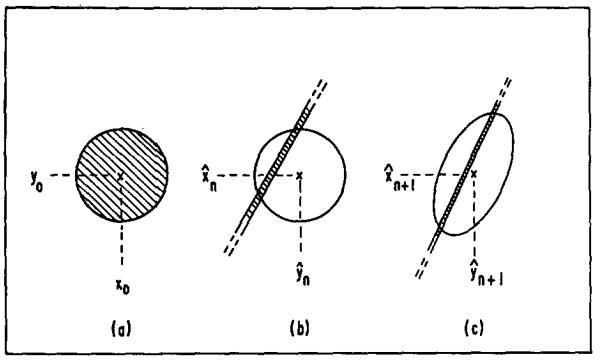



FIG. 4. A sketch of a neighborhood of $x_0$ and $y_0$ and its image after $n$ and $n + 1$ iterations of a given mapping. The central points in $b$ and $c$ represent the numerical approximation to the image of the central point in $a$. The circle in $b$ is transformed to the ellipse in $c$.

image $(\hat{x}_n, \hat{y}_n)$ should lie within a neighborhood of $(x_n, y_n)$. However, to prove that a neighborhood of $(x_0, y_0)$ lies outside a Moser surface we only need to know that the true mapping takes some point in the neighborhood of $(x_0, y_0)$ to some point in the neighborhood of $(\hat{x}_n, \hat{y}_n)$. As will now be shown, this condition is much easier to satisfy.

Figure 4 illustrates the geometry used in calculating the effect of roundoff error. The central points of Figs. 4(a), 4(b), and 4(c) are the initial point and its images as calculated numerically after $n$ and $n + 1$ iterations of $T$. A neighborhood of $(x_0, y_0)$, shown shaded in Fig. 4(a), is assumed to be transformed by the exact mapping into the shaded regions of Figs. 4(b) and 4(c). The mapping of the neighborhood of $(\hat{x}_n, \hat{y}_n)$ to a neighborhood of $(\hat{x}_{n+1}, \hat{y}_{n+1})$ will now be treated in detail in the linearized approximation, utilizing Eq. (7), to confirm the general picture.

In the linearized approximation a circle is transformed into an ellipse with a ratio of major to minor diameter of

$$D_{\max}/D_{\min} = [(c_n^2 + d_n^2)^{\frac{1}{2}} + (1 + c_n^2 + d_n^2)^{\frac{1}{2}}]^2. \tag{32}$$

Here $c_n$ and $d_n$ are defined by parameterizing $J(\hat{x}_n, \hat{y}_n)$ as in Eq. (9). In this approximation the separation of two points $\delta r_n$ is related to the separation of their images $\delta r_{n+1}$ by

$$(\delta r_{n+1})^2 = \{1 + 2c_n^2 + 2d_n^2 + [(c_n^2 + d_n^2) \times (1 + c_n^2 + d_n^2)]^{\frac{1}{2}} \cos 2(\theta - \theta_M)\}(\delta r_n)^2. \tag{33}$$

Here $\theta - \theta_M$ is the angle between the line connecting the images of the two points and the major axis of the ellipse of Fig. 4(c), and $\theta_M$ is given by Eq. (13). On the other hand, the relation connecting the distance between a point and a line $\delta r_n^*$ and the distance between their images $\delta r_{n+1}^*$ is

$$(\delta r_{n+1}^*)^2 = \{1 + 2c_n^2 + 2d_n^2 + [(c_n^2 + d_n^2) \times (1 + c_n^2 + d_n^2)]^{\frac{1}{2}} \cos 2(\theta^* - \theta_M)\}^{-1}(\delta r_n^*)^2. \tag{34}$$

On the average, assuming successive phase angles are uncorrelated, the distance between two points is increased by the factor

$$\left\langle \frac{\delta r_{n+1}}{\delta r_n} \right\rangle = \exp\left\{ \frac{1}{2\pi} \int_0^{2\pi} \tfrac{1}{2} \ln \left( \frac{\delta r_{n+1}}{\delta r_n} \right)^2 d\theta \right\} = (1 + c_n^2 + d_n^2)^{\frac{1}{2}} \quad (35)$$

while a point and a line move toward each other by the factor

$$\langle \delta r_{n+1}^* / \delta r_n^* \rangle = (1 + c_n^2 + d_n^2)^{-\frac{1}{2}}. \quad (36)$$

This describes a situation in which every region is being continually stretched into long thin fibers by successive iterations, as shown in Figs. 3(a) and 3(b).

We thus see from Eq. (35) that small errors in the distance between two points tend to grow exponentially with the number of iterations so that the stringent criterion for accuracy shows a rapid loss of significant figures. On the other hand, small errors of the distance from a point to a region tend to decrease exponentially. This would apparently lead to the conclusion that when more significant figures appear to be lost, more confidence can be placed in the results! Of course, the theory is highly simplified, but it should give some credence to the results of the "escape" calculations. These are summarized in Table IV. The first column gives the initial value for each calculation and the second column gives the number of iterations required for the image to cross the line $x + y = 0$. Those calculations that were stopped before the image crossed this line were discarded as inconclusive.

The first set of seven points in Table IV illustrates behavior outside all the structure associated with the fixed point at $x = y = 0.6580$ with $P/Q = \frac{1}{9}$. In the neighborhood of this fixed point are Moser surfaces of the third order of the hierarchy that enclose a portion of the line $x = y$. Part of this structure is illustrated in Fig. 1. Further exploration was conducted near hyperbolic fixed points to avoid these third-order Moser surfaces. The next three points were selected near the negative-residue fixed point at $x = y = 0.6938903$ with $P/Q = \frac{8}{71}$. The last point in the table was chosen near the fixed point at $x = y = 0.69417$ with $P/Q = \frac{7}{62}$. The residue of this fixed point is $R = -0.254$ (and $R = 0.253$ for the corresponding positive residue family) so that the corresponding $f$ of Table I is very slightly greater than unity. It is apparent that this point is close to the boundary of the ergodic region since its image escapes only after a very large number of iterations.

TABLE IV. Results of escape calculations.

| $x_0 = y_0$ | $N$ |
|---|---|
| 0.50 | 42 |
| 0.52 | 67 |
| 0.54 | 53 |
| 0.56 | 61 |
| 0.58 | 430 |
| 0.60 | 222 |
| 0.62 | 2 453 |
| 0.69385 | 2 155 |
| 0.69387 | 3 721 |
| 0.69389 | 3 843 |
| 0.69415 | 45 216 |

## V. CONCLUSIONS

The major features, at least, of the behavior of multiple iterations of a two-dimensional mapping can be understood in terms of the fixed points of the mapping. This paper investigated the relation between invariant curves enclosing an elliptic fixed point and the character of related families of fixed points. These families were identified as belonging to the same order of the hierarchy as the closed invariant curves. It was inferred from numerical computations that in an annulus that contains closed invariant curves an infinite number of the corresponding fixed points are elliptic. Some of these fixed points have an arbitrarily small residue. In outer regions there are no invariant curves and all but a finite number of the fixed points of the given order of the hierarchy are hyperbolic. Thus there is a close relation between containment, ergodicity, and the character of the fixed points.

This is intuitively a very satisfactory result. Consider a series of fixed points for which the residue approaches zero and $Q$ approaches infinity such as given by Eq. (31) and Table III. Such a series of fixed points approaches a closed invariant curve since the positive and negative residue points become indistinguishable in the limit as they become dense on a curve. The limit of Eq. (31) as $n$ goes to infinity is an irrational number of a type[13] that most easily satisfies Eq. (22), so that this intuition is entirely consistent with Moser's theorem.

On the other hand, hyperbolic fixed points act as scattering centers. Successive images of a given point must ultimately move away from any hyperbolic point that they approach. In a region where almost all the fixed points are hyperbolic and they are dense or almost dense, it would seem reasonable that only isolated points would have correlated images. These correlated images would be another set of fixed points.

It is useful to have some notion of the location of the region in which all the fixed points are hyperbolic without having to calculate hundreds of residues. It is fairly clear that there will be extensive ergodic

regions near fixed points for which $E$ is distinctly different from zero, so that the mapping is strongly perturbed. In the computed example this effect was strong enough to bring the boundary of the outer ergodic region inside the fixed points with $P/Q = \frac{1}{9}$. It seems reasonable that the boundary will always lie just inside a fixed point with $E$ greater than a tenth or so. The well-developed, relatively fat structures associated with these fixed points are easy to locate and thus provide a good first estimate of the stability region of the mapping. Increasing the shear, a quantity introduced in Eq. (21), tends to reduce the values of $E$. Thus the critical value of $E$ will tend to be larger in systems with small shear.

Only a few other authors have endeavored to establish some boundary to the ergodic region. Rosenbluth *et al.*[16] attempted to identify this boundary with the overlap of resonances. In the terms of this paper this translates to the statement that in regions outside a Moser surface of a given order in the hierarchy, elements of the next higher order of the hierarchy around a given point overlap with similar elements associated with neighboring points. This is not inconsistent with the results of the present paper but it seems more difficult to express quantitatively. Their criterion depends heavily on the size and shape of the secondary structure and thus depends on the parameter $E$ of this paper. Here it is deduced that $E$ plays only a secondary role in determining the boundary of the ergodic region, through its relation to $R$ by Eq. (15) and also through its effect on the location of fixed points as illustrated in Fig. 3.

Roels and Hénon[17] related the boundary of the ergodic region to an approximate radius of convergence of the asymptotic series[11] for invariant curves that was mentioned in Sec. III. Since this series is asymptotic, it has an irreducible error that increases with radius. An approximate radius of convergence of the series is a region in which this irreducible error is sharply increasing with radius. One reason that the error is irreducible is that the series is unable to handle all the details of higher orders of the hierarchy that was introduced in Sec. III. Thus it is reasonable that an approximate radius of convergence should occur at a point at which the character of higher orders of the hierarchy is changing. However, it is difficult to make any quantitative statements.

Morozov and co-workers[18–20] have noted a correspondence between ergodic regions and hyperbolic fixed points but their work has been confined to fixed points with $P = 1$. Thus they do not find an exact correspondence. In their work the mapping is generated by integrating differential equations for magnetic field lines. It is gratifying that they get results comparable to those found for the very much simpler mapping utilized in this paper.

*Note added in proof:* The information on fixed points derivable from symmetry arguments has been thoroughly discussed by de Vogelaere.[21]

## ACKNOWLEDGMENTS

The author wishes to express his gratitude to the many people who have given assistance and encouragement in the course of this work. Dr. M. D. Kruskal first suggested the use of simple algebraic models for studying two-dimensional mappings and has given much advice and stimulus. Particularly enlightening discussions with Dr. J. Moser, Dr. M. N. Rosenbluth, Dr. R. Z. Sagdeev, and Dr. J. B. Taylor have been most helpful. A. Stutz assisted with the programming.

This work was performed under the auspices of the U.S. Atomic Energy Commission, Contract AT(30-1)-1238; also use was made of computer facilities supported in part by National Science Foundation Grant NSF-GP579.

[16] M. N. Rosenbluth, R. Z. Sagdeev, J. B. Taylor, and G. M. Zaslavski, Nucl. Fusion **6**, 297 (1966).

[17] J. Roels and M. Hénon, Bull. Astron. Ser. 3 **2**, 267 (1967).

[18] I. M. Gel'fand, M. I. Graev, N. M. Zueva, M. S. Mikhailova, and A. I. Morozov, Dokl. Akad. Nauk SSSR **148**, No. 6, 1286 (1963) [Soviet Phys.—Doklady **8**, 188 (1963)].

[19] A. I. Morozov and L. S. Solov'ev, Zh. Eksp. Teor. Fiz. **45**, 955 (1963) [Soviet Phys.—JETP **18**, 660 (1964)].

[20] M. I. Graev, M. S. Mikhailova, and A. I. Morozov, Zh. Telth. Fiz. **35**, No. 7, 1189 (1965) [Soviet Phys.—Tech. Phys. **10**, 920 (1966)].

[21] R. de Vogelaere, in *Contributions to the Theory of Nonlinear Oscillations*, S. Lefschetz, Ed. (Princeton University Press, Princeton, 1958), Vol. IV, p. 53.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 5 MAY 1968

# Weyl Transform in Nonrelativistic Quantum Dynamics

BORIS LEAF
*Department of Physics, State University of New York, College at Cortland, Cortland, New York*

(Received 25 September 1967)

The Weyl transform is applied in quantum dynamics to derive and extend Moyal's statistical theory of phase-space distributions for noncommuting coordinate and momentum operators. The distinction is made between Weyl transforms in Schrödinger and Heisenberg pictures; the general case of time-dependent Hamiltonians is considered. The Wigner function for the probability distribution in a phase space of Cartesian coordinates $\mathbf{Q}$ and momenta $\mathbf{K}$ propagates according to a conditional probability $P(t, \mathbf{Q}, \mathbf{K} \mid t_0, \mathbf{Q}_0, \mathbf{K}_0)$, which is exhibited as a Feynman path integral in phase space. Properties of $P(t, \mathbf{Q}, \mathbf{K} \mid t_0, \mathbf{Q}_0, \mathbf{K}_0)$ are developed; it is expressed in terms of the quantum generalization of the classical Liouville operator. The Weyl transform of a Heisenberg operator propagates according to $P(t, \mathbf{Q}_0, \mathbf{K}_0 \mid t_0, \mathbf{Q}, \mathbf{K})$ which is also given as a Feynman path integral. An equation for the time evolution of Weyl transforms of Heisenberg operators is obtained, according to which the transform of Heisenberg coordinate and momentum operators obey a quantum form of Hamilton's equations of motion. If the initial density operator of the system commutes with the coordinate operator, then the state of the system is a mixture of pure coordinate states; the spectrum of the density operator in this case is continuous. For a Heisenberg operator $A_H(t)$ with Weyl transform $A_H(t, \mathbf{Q}, \mathbf{K})$, the function $\mathcal{A}(t, \mathbf{Q}) = \int d\mathbf{K} A_H(t, \mathbf{Q}, \mathbf{K})$ is the expectation at time $t$ of the dynamical property for a quantum system initially in a pure state of coordinate $\mathbf{Q}$; it is the quantum-mechanical generalization of the dynamical property of the system along the classical trajectory in configuration space at time $t$. The probability amplitude for the time dependence of $\mathcal{A}(t, \mathbf{Q})$ can be expressed as a Feynman path integral with a Heisenberg Lagrangian. The amplitude of the conditional probability $P(t, \mathbf{Q} \mid t_0, \mathbf{Q}_0)$ considered by Feynman is expressed as a path integral with a Schrödinger Lagrangian. The velocity in the Heisenberg Lagrangian is the negative of that in the Schrödinger Lagrangian of Feynman, but it agrees with the velocity appearing in the Hamiltonian equations. It is the Heisenberg Lagrangian that is the Lagrangian of classical dynamics. For a particle whose potential energy is a function of position, a quantum form of Newton's second law is obtained. An extension of the formalism to non-Cartesian coordinate systems is given.

## 1. INTRODUCTION

Properties of the Weyl transform have been developed in a previous paper.[1] It is the purpose of this paper to apply the transform in formulating nonrelativistic quantum dynamics, by means of probability distributions and trajectories in phase space. The probability density is given by the Wigner distribution function.[2] Some of the results on probability distributions are anticipated in a paper by Moyal[3] on quantum mechanics as a statistical theory. His treatment is simplified and extended here by use of the Weyl transform, which permits a derivation of Moyal's statistical theory from Dirac's formulation of quantum mechanics.[4] It is important to distinguish the Weyl transforms of operators in the Schrödinger and Heisenberg pictures. With this distinction, connection can be made with Feynman's path-integral formulation of quantum dynamics.[5,6] Feynman's path integrals for propagation of probability are in the Schrödinger picture; his Schrödinger Lagrangian function differs in the sign of the velocity from the Heisenberg Lagrangian. It is the Heisenberg Lagrangian which is the quantum-mechanical generalization of the classical Lagrangian. In the Heisenberg picture, the Weyl transform of the coordinate and momentum operators satisfy Hamilton's canonical equations in quantum form; when averaged over the initial phase-space distribution these equations became the Ehrenfest equations.[7]

In Sec. 2 the properties of the Weyl transform which will be needed in this paper are summarized. In Sec. 3 Weyl transforms of operators in Schrödinger and Heisenberg pictures are discussed for the case of a time-dependent Hamiltonian. The time dependence of the Wigner function is determined by

$$P(t, \mathbf{Q}, \mathbf{K} \mid t_0, \mathbf{Q}_0, \mathbf{K}_0),$$

the probability of the phase-space point $(\mathbf{Q}, \mathbf{K})$ at time $t$ conditional on $(\mathbf{Q}_0, \mathbf{K}_0)$ at $t_0$. The time dependence of the Weyl transforms of Heisenberg operators is determined by $P(t, \mathbf{Q}_0, \mathbf{K}_0 \mid t_0, \mathbf{Q}, \mathbf{K})$. For Heisenberg coordinate and momentum operators

[1] B. Leaf, J. Math. Phys. **9**, 65 (1968), additional references to the Weyl transform are given in this article. See also the paper by K. Imre, E. Özizmir, M. Rosenbaum, and P. F. Zweifel, J. Math. Phys. **8**, 1097 (1967); the "Wigner equivalent" of these authors is the same as the "Weyl transform."

[2] E. Wigner, Phys. Rev. **40**, 749 (1932).

[3] J. E. Moyal, Proc. Cambridge Phil. Soc. **45**, 99 (1949).

[4] P. A. M. Dirac, *The Principles of Quantum Mechanics* (Clarendon Press, Oxford, 1947), 3rd ed.

[5] R. P. Feynman, Rev. Mod. Phys. **20**, 367 (1948).

[6] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill Book Co., Inc., New York, 1965).

[7] See, for example, A. Messiah, *Quantum Mechanics* (North-Holland Publ. Co., Amsterdam, 1961), Vol. I, p. 216.

the Weyl transforms satisfy Hamilton's equations in quantum form. The properties of the conditional probabilities

$$P(t, \mathbf{Q}, \mathbf{K} \mid t_0, \mathbf{Q}_0, \mathbf{K}_0)$$

and

$$P(t_0, \mathbf{Q}_0, \mathbf{K}_0 \mid t_0, \mathbf{Q}, \mathbf{K})$$

are developed; they are expressed in terms of the quantum-mechanical generalization of the Liouville operator. In Sec. 4 the Markoffian properties of the conditional probabilities are considered. These properties permit formulation of the Wigner function and of the Weyl transforms of Heisenberg operators as Feynman path integrals in phase space. Feynman's path integrals are in configuration space. As discussed in Sec. 5, the transition probability considered by Feynman is $P(t, \mathbf{Q} \mid t_0, \mathbf{Q}_0)$, the probability of coordinate $\mathbf{Q}$ at time $t$ conditional on $\mathbf{Q}_0$ at $t_0$. The amplitude for this probability is given as a Feynman path integral with a Schrödinger Lagrangian function. The Lagrangian formulation of classical dynamics does not express the time evolution of probability distributions. It describes rather the time dependence of the dynamical properties of the system along the classical trajectory. For a quantum-mechanical Heisenberg operator $A_H(t)$, a function $\mathcal{A}(t, \mathbf{Q})$ can be defined; it is the integral over the momentum variables of the Weyl transform of $A_H(t)$. $\mathcal{A}(t, \mathbf{Q})$ is the expectation of the dynamical property of the quantum system initially in a pure state of coordinate $\mathbf{Q}$ at time $t$; it is the quantum-mechanical generalization of the dynamical property along the classical trajectory in configuration space at time $t$. The relevant amplitude for the time dependence of $\mathcal{A}(t, \mathbf{Q})$ can be expressed as a Feynman path integral with a Heisenberg Lagrangian function. The velocity in the Heisenberg Lagrangian is the negative of that in the Schrödinger Lagrangian of Feynman, but it agrees with the velocity appearing in the quantum form of Hamilton's equations. For a particle whose potential energy is a function of position, a quantum-mechanical form of Newton's second law is derived. Up to this point only Cartesian coordinates and momenta have been considered. In Sec. 6 path integrals in other coordinate systems are discussed

## 2. THE WEYL TRANSFORM

Properties of the Weyl transform have been developed in a previous paper.[1] In this section are summarized the more important results which will be needed in the present work.

An operator of quantum mechanics is expressed, according to the Weyl transformation, as

$$A = \int \cdots \int d\mathbf{Q}\, d\mathbf{K} \Delta(\mathbf{Q}, \mathbf{K}) A(\mathbf{Q}, \mathbf{K}), \tag{2.1}$$

where the Weyl transform $A(\mathbf{Q}, \mathbf{K})$ is given by the inverse transformation

$$A(\mathbf{Q}, \mathbf{K}) = \mathrm{Tr}\,[A\Delta(\mathbf{Q}, \mathbf{K})], \tag{2.2}$$

and

$$\Delta(\mathbf{Q}, \mathbf{K}) \equiv \int \cdots \int d\mathbf{u}\, d\mathbf{v} \times \exp\{2\pi i[\mathbf{u} \cdot (q - \mathbf{Q}) + \mathbf{v} \cdot (\mathbf{k} - \mathbf{K})]\}. \tag{2.3}$$

$\mathbf{q}$ and $\mathbf{p} = 2\pi\hbar\mathbf{k}$ are Cartesian coordinate and momentum operators for which

$$q_i p_j - p_j q_i = i\hbar\delta_{ij}. \tag{2.4}$$

$\mathbf{Q}$ and $\mathbf{K}$ are eigenvalues of $\mathbf{q}$ and $\mathbf{k}$. Alternative forms of the Hermitian operator $\Delta(\mathbf{Q}, \mathbf{K})$ are

$$\begin{aligned}\Delta(\mathbf{Q},\mathbf{K}) &= \int \cdots \int d\mathbf{v} \exp(2\pi i \mathbf{v} \cdot \mathbf{K}) \,|\mathbf{Q} + \tfrac{1}{2}\mathbf{v}\rangle\langle\mathbf{Q} - \tfrac{1}{2}\mathbf{v}| \\ &= \int \cdots \int d\mathbf{u} \exp(-2\pi i \mathbf{u} \cdot \mathbf{Q}) \,|\mathbf{K} + \tfrac{1}{2}\mathbf{u}\rangle\langle\mathbf{K} - \tfrac{1}{2}\mathbf{u}| \\ &= \exp[(i/4\pi)(\partial/\partial\mathbf{Q}) \cdot (\partial/\partial\mathbf{K})] \,|\mathbf{K}\rangle\langle\mathbf{K} \mid \mathbf{Q}\rangle\langle\mathbf{Q}|.\end{aligned} \tag{2.5}$$

Matrix elements of any operator can be obtained from (2.1) and the matrix elements of $\Delta(\mathbf{Q}, \mathbf{K})$,

$$\begin{aligned}&\langle\mathbf{Q}'|\,\Delta(\mathbf{Q}, \mathbf{K})\,|\mathbf{K}'\rangle \\ &\quad = \langle\mathbf{Q}' \mid \mathbf{K}'\rangle \exp[(-i/4\pi)(\partial/\partial\mathbf{Q}) \cdot (\partial/\partial\mathbf{K})] \\ &\qquad\qquad \times \delta(\mathbf{Q}' - \mathbf{Q})\delta(\mathbf{K}' - \mathbf{K}), \\ &\langle\mathbf{Q}''|\,\Delta(\mathbf{Q}, \mathbf{K})\,|\mathbf{Q}'\rangle \\ &\quad = \delta[\mathbf{Q} - \tfrac{1}{2}(\mathbf{Q}' + \mathbf{Q}'')] \exp[2\pi i\mathbf{K} \cdot (\mathbf{Q}'' - \mathbf{Q}')], \\ &\langle\mathbf{K}''|\,\Delta(\mathbf{Q}, \mathbf{K})\,|\mathbf{K}'\rangle \\ &\quad = \delta[\mathbf{K} - \tfrac{1}{2}(\mathbf{K}' + \mathbf{K}'')] \exp[-2\pi i\mathbf{Q} \cdot (\mathbf{K}'' - \mathbf{K}')].\end{aligned} \tag{2.6}$$

Also,

$$\mathrm{Tr}\,\Delta(\mathbf{Q}, \mathbf{K}) = 1, \quad \int \cdots \int d\mathbf{Q}\, d\mathbf{K} \Delta(\mathbf{Q}, \mathbf{K}) = \mathbf{1}, \tag{2.7}$$

$$\begin{aligned}\int d\mathbf{Q}\Delta(\mathbf{Q}, \mathbf{K}) &= \delta(\mathbf{k} - \mathbf{K}) = |\mathbf{K}\rangle\langle\mathbf{K}|, \\ \int d\mathbf{K}\Delta(\mathbf{Q}, \mathbf{K}) &= \delta(\mathbf{q} - \mathbf{Q}) = |\mathbf{Q}\rangle\langle\mathbf{Q}|.\end{aligned} \tag{2.8}$$

The Weyl transform of an operator $A(\mathbf{q})$, which is a function of $\mathbf{q}$ only, is $A(\mathbf{Q})$; the transform of $A(\mathbf{k})$, a function of $\mathbf{k}$ only, is $A(\mathbf{K})$. Such transforms are unchanged in the classical limit $\hbar \to 0$ and are therefore the same as classical dynamical functions of coordinate $\mathbf{Q}$ or momentum $\mathbf{P} = 2\pi\hbar\mathbf{K}$, respectively. The set of components of the vectors $(\mathbf{Q}, \mathbf{K})$ specify the coordinates of a point in phase space.

If $A$ and $B$ are two quantum operators, then the Weyl transform of their commutator is

$$[AB - BA](\mathbf{Q}, \mathbf{K}) = 2i\int\cdots\int d\mathbf{Q}'\, d\mathbf{K}'\delta(\mathbf{Q} - \mathbf{Q}')\delta(\mathbf{K} - \mathbf{K}') \times \sin\{(1/4\pi)[(\partial/\partial\mathbf{Q})\cdot(\partial/\partial\mathbf{K}') - (\partial/\partial\mathbf{K})\cdot(\partial/\partial\mathbf{Q}')]\} \times A(\mathbf{Q}, \mathbf{K})B(\mathbf{Q}', \mathbf{K}'). \quad (2.9)$$

$$\mathrm{Tr}\,(AB) = \int\cdots\int d\mathbf{Q}\, d\mathbf{K}A(\mathbf{Q}, \mathbf{K})B(\mathbf{Q}, \mathbf{K}), \quad (2.10)$$

$$\mathrm{Tr}\,[\Delta(\mathbf{Q}, \mathbf{K})\Delta(\mathbf{Q}', \mathbf{K}')] = \delta(\mathbf{Q} - \mathbf{Q}')\delta(\mathbf{K} - \mathbf{K}'). \quad (2.11)$$

## 3. SCHRÖDINGER AND HEISENBERG PICTURES

A time-dependent operator in the Schrödinger picture is written $A_s(t)$. The Weyl transformation (2.1) gives

$$A_s(t) = \int\cdots\int d\mathbf{Q}\, d\mathbf{K}\Delta(\mathbf{Q}, \mathbf{K})A_s(t, \mathbf{Q}, \mathbf{K}), \quad (3.1)$$

with Weyl transform

$$A_s(t, \mathbf{Q}, \mathbf{K}) = \mathrm{Tr}\,[\Delta(\mathbf{Q}, \mathbf{K})A_s(t)]. \quad (3.2)$$

Obviously, $\mathbf{Q}_s(t, \mathbf{Q}, \mathbf{K}) = \mathbf{Q}$ and $\mathbf{K}_s(t, \mathbf{Q}, \mathbf{K}) = \mathbf{K}$ are the transforms of the position and momentum operators, $\mathbf{q}_s$ and $\mathbf{k}_s$. The expectation of a dynamical property $A_s(t)$ when the system is in a state described by the probability density operator $\rho_s(t)$ is, according to (2.10),

$$\langle A(t)\rangle = \mathrm{Tr}\,[A_s(t)\rho_s(t)] = \int\cdots\int d\mathbf{Q}\, d\mathbf{K}A_s(t, \mathbf{Q}, \mathbf{K})\rho_s(t, \mathbf{Q}, \mathbf{K}). \quad (3.3)$$

The Weyl transform $\rho_s(t, \mathbf{Q}, \mathbf{K})$ is the Wigner function.[2] The density operator defined in terms of the initial states $|\psi_m, t_0\rangle$ as

$$\rho_s(t_0) = \sum_m w_m |\psi_m, t_0\rangle\langle\psi_m, t_0|, \quad w_m \geq 0, \sum_m w_m = 0, \quad (3.4)$$

becomes at time $t > t_0$

$$\rho_s(t) = U(t, t_0)\rho_s(t_0)U^\dagger(t, t_0). \quad (3.5)$$

The unitary operator $U(t, t_0)$ is the solution of the Schrödinger equation,

$$i\hbar\, dU(t, t_0)\,|\,dt = H_s(t)U(t, t_0) = U(t, t_0)H_H(t), \quad (3.6)$$

with $U(t_0, t_0) = \mathbf{1}$. $H_s(t)$ is the (time-dependent) Hamiltonian operator in the Schrödinger picture; $H_H(t)$, the Hamiltonian in Heisenberg picture. $U^\dagger(t, t_0)$, the Hermitian adjoint of $U(t, t_0)$, is the solution of the adjoint equation,

$$-i\hbar\, dU^\dagger(t, t_0)\,|\,dt = H_H(t)U^\dagger(t, t_0) = U^\dagger(t, t_0)H_s(t), \quad (3.7)$$

with $U^\dagger(t_0, t_0) = \mathbf{1}$. Accordingly, by integration and iteration of (3.6) and (3.7), for $t > t_0$,

$$U(t, t_0) = T\exp\left[-(i/\hbar)\int_{t_0}^{t} dt' H_s(t')\right] = T^r\exp\left[-(i/\hbar)\int_{t_0}^{t} dt' H_H(t')\right], \quad (3.8)$$

$$U^\dagger(t, t_0) = T\exp\left[(i/\hbar)\int_{t_0}^{t} dt' H_H(t')\right] = T^r\exp\left[(i/\hbar)\int_{t_0}^{t} dt' H_s(t')\right]. \quad (3.9)$$

$T$ indicates the time-ordered product

$$T\,H(t')H(t'') = \begin{cases} H(t')H(t''), & \text{if } t' > t'' \\ H(t'')H(t'), & \text{if } t'' > t'. \end{cases} \quad (3.10)$$

$T^r$ indicates the reversed time-ordered product for which the inequalities of (3.10) are interchanged when $T$ is replaced by $T^r$.

According to (3.5) and (3.2), the Wigner function can be written as

$$\rho_s(t, \mathbf{Q}, \mathbf{K}) = \mathrm{Tr}\,[\Delta(\mathbf{Q}, \mathbf{K})U(t, t_0)\rho_s(t_0)U^\dagger(t, t_0)] = \int\cdots\int d\mathbf{Q}_0\, d\mathbf{K}_0 P(t, \mathbf{Q}, \mathbf{K}\,|\,t_0, Q_0, \mathbf{K}_0)\rho_s(t_0, \mathbf{Q}_0, \mathbf{K}_0), \quad (3.11)$$

where the probability of $(\mathbf{Q}, \mathbf{K})$ at $t$ conditional on $(\mathbf{Q}_0, \mathbf{K}_0)$ at $t_0$ is the real quantity

$$P(t, \mathbf{Q}, \mathbf{K}\,|\,t_0, \mathbf{Q}_0, \mathbf{K}_0) = \mathrm{Tr}\,[\Delta(\mathbf{Q}, \mathbf{K})U(t, t_0)\Delta(\mathbf{Q}_0, \mathbf{K}_0)U^\dagger(t, t_0)]. \quad (3.12)$$

Like the Wigner function, $P(t, \mathbf{Q}, \mathbf{K}\,|\,t_0, \mathbf{Q}_0, \mathbf{K}_0)$ is not everywhere nonnegative.

Since $P(t, \mathbf{Q}, \mathbf{K}\,|\,t_0, \mathbf{Q}_0, \mathbf{K}_0)$ plays a central role, its properties are now considered. From the expressions for $\Delta(\mathbf{Q}, \mathbf{K})$ given in (2.5), various forms are readily obtained, as for example,

$$P(t, \mathbf{Q}, \mathbf{K}\,|\,t_0, \mathbf{Q}_0, \mathbf{K}_0) = \exp\{(i/4\pi)[(\partial/\partial\mathbf{Q})\cdot(\partial/\partial\mathbf{K}) - (\partial/\partial\mathbf{Q}_0)\cdot(\partial/\partial\mathbf{K}_0)]\} \times \exp[2\pi i(\mathbf{Q}_0\cdot\mathbf{K}_0 - \mathbf{Q}\cdot\mathbf{K})] \times \langle\mathbf{Q}|\, U(t, t_0)\,|\mathbf{Q}_0\rangle\langle\mathbf{K}_0|\, U^\dagger(t, t_0)\,|\mathbf{K}\rangle. \quad (3.13)$$

According to the Weyl transformation (2.1), from (3.12),

$$U(t, t_0)\Delta(\mathbf{Q}_0, \mathbf{K}_0)U^\dagger(t, t_0) = \int\cdots\int d\mathbf{Q}\, d\mathbf{K}\Delta(\mathbf{Q}, \mathbf{K})P(t, \mathbf{Q}, \mathbf{K}\,|\,t_0, \mathbf{Q}_0, \mathbf{K}_0), \quad (3.14)$$

$$U^\dagger(t, t_0)\Delta(\mathbf{Q}, \mathbf{K})U(t, t_0) = \int\cdots\int d\mathbf{Q}_0\, d\mathbf{K}_0\Delta(\mathbf{Q}_0, \mathbf{K}_0)P(t, \mathbf{Q}, \mathbf{K}\,|\,t_0, \mathbf{Q}_0, \mathbf{K}_0). \quad (3.15)$$

Therefore, from (2.11),

$$\int\cdots\int d\mathbf{Q}\,d\mathbf{K}P(t,\mathbf{Q},\mathbf{K}\mid t_0,\mathbf{Q}',\mathbf{K}')P(t,\mathbf{Q},\mathbf{K}\mid t_0,\mathbf{Q}'',\mathbf{K}'') = \delta(\mathbf{Q}'-\mathbf{Q}'')\delta(\mathbf{K}'-\mathbf{K}''), \quad (3.16)$$

$$\int\cdots\int d\mathbf{Q}_0\,d\mathbf{K}_0P(t,\mathbf{Q}',K'\mid t_0,\mathbf{Q}_0,\mathbf{K}_0) \times P(t,\mathbf{Q}'',\mathbf{K}''\mid t_0,\mathbf{Q}_0,\mathbf{K}_0) = \delta(\mathbf{Q}'-\mathbf{Q}'')\delta(\mathbf{K}'-\mathbf{K}''). \quad (3.17)$$

From (3.14) and (3.6),

$$\int\cdots\int d\mathbf{Q}\,d\mathbf{K}\Delta(\mathbf{Q},\mathbf{K})\partial P(t,\mathbf{Q},\mathbf{K}\mid t_0,\mathbf{Q}_0,\mathbf{K}_0)/\partial t = -(i/\hbar)[H_s(t)U(t,t_0)\Delta(\mathbf{Q}_0,\mathbf{K}_0)U^\dagger(t,t_0) - U(t,t_0)\Delta(\mathbf{Q}_0,\mathbf{K}_0)U^\dagger(t,t_0)H_s(t)],$$

so that the Weyl transformation gives

$$\begin{aligned}\partial P(t,\mathbf{Q},\mathbf{K}\mid t_0,\mathbf{Q}_0,\mathbf{K}_0)/\partial t &= -(i/\hbar)\,\mathrm{Tr}\,\{[\Delta(\mathbf{Q},\mathbf{K})H_s(t) - H_s(t)\Delta(\mathbf{Q},\mathbf{K})] \times U(t,t_0)\Delta(\mathbf{Q}_0,\mathbf{K}_0)U^\dagger(t,t_0)\}\\ &= -(i/\hbar)\int\cdots\int d\mathbf{Q}'\,d\mathbf{K}'P(t,\mathbf{Q}',\mathbf{K}'\mid t_0,\mathbf{Q}_0,\mathbf{K}_0) \times \mathrm{Tr}\,\{\Delta(\mathbf{Q},\mathbf{K})[H_s(t)\Delta(\mathbf{Q}',\mathbf{K}') - \Delta(\mathbf{Q}',\mathbf{K}')H_s(t)].\end{aligned} \quad (3.18)$$

But, from (2.9) and (2.11),

$$\mathrm{Tr}\,\{\Delta(\mathbf{Q},\mathbf{K})[H_s(t)\Delta(\mathbf{Q}',\mathbf{K}') - \Delta(\mathbf{Q}',\mathbf{K}')H_s(t)]\} = \mathfrak{L}_s(t,\mathbf{Q},\mathbf{K})\delta(\mathbf{Q}-\mathbf{Q}')\delta(\mathbf{K}-\mathbf{K}'), \quad (3.19)$$

where the quantum-mechanical Liouville operator $\mathfrak{L}_s(t,\mathbf{Q},\mathbf{K})$ is defined as

$$\mathfrak{L}_s(t,\mathbf{Q},\mathbf{K}) = 2i\int\cdots\int d\mathbf{Q}''\,d\mathbf{K}''\delta(\mathbf{Q}-\mathbf{Q}'')\delta(\mathbf{K}-\mathbf{K}'') \times \sin\{(1/4\pi)[(\partial/\partial\mathbf{Q}'')\cdot(\partial/\partial\mathbf{K}) - (\partial/\partial\mathbf{K}'')\cdot(\partial/\partial\mathbf{Q})]\} \times H_s(t,\mathbf{Q}'',\mathbf{K}''). \quad (3.20)$$

Therefore,

$$\partial P(t,\mathbf{Q},\mathbf{K}\mid t_0,\mathbf{Q}_0,\mathbf{K}_0)/\partial t = -(i/\hbar)\mathfrak{L}_s(t,\mathbf{Q},\mathbf{K})P(t,\mathbf{Q},\mathbf{K}\mid t_0,\mathbf{Q}_0,\mathbf{K}_0). \quad (3.21)$$

Since initially, according to (3.12) and (2.11),

$$P(t_0,\mathbf{Q},\mathbf{K}\mid t_0,\mathbf{Q}_0,\mathbf{K}_0) = \delta(\mathbf{Q}-\mathbf{Q}_0)\delta(\mathbf{K}-\mathbf{K}_0), \quad (3.22)$$

integration and iteration of (3.21) gives the solution, for $t > t_0$,

$$P(t,\mathbf{Q},\mathbf{K}\mid t_0,\mathbf{Q}_0,\mathbf{K}_0) = T\exp\left[-(i/\hbar)\int_{t_0}^{t}dt'\mathfrak{L}_s(t',\mathbf{Q},\mathbf{K})\right]\delta(\mathbf{Q}-\mathbf{Q}_0)(\mathbf{K}-\mathbf{K}_0). \quad (3.23)$$

According to (3.11), the Wigner function becomes

$$\rho_s(t,\mathbf{Q},\mathbf{K}) = T\exp\left[-(i/\hbar)\int_{t_0}^{t}dt'\mathfrak{L}_s(t',\mathbf{Q},\mathbf{K})\right]\rho_s(t_0,\mathbf{Q},\mathbf{K}), \quad (3.24)$$

a solution of the differential equation for the time evolution of the Wigner function,

$$\partial\rho_s(t,\mathbf{Q},\mathbf{K})/\partial t = -(i/\hbar)\mathfrak{L}_s(t,\mathbf{Q},\mathbf{K})\rho_s(t,\mathbf{Q},\mathbf{K}), \quad (3.25)$$

which is the Weyl transform of the quantum-mechanical von Neumann equation for the density operator,

$$d\rho_s/dt = -(i/\hbar)[H_s(t)\rho_s(t) - \rho_s(t)H_s(t)]. \quad (3.26)$$

Alternatively, from (3.15),

$$\int\cdots\int d\mathbf{Q}_0\,d\mathbf{K}_0\Delta(\mathbf{Q}_0,\mathbf{K}_0)\partial P(t,\mathbf{Q},\mathbf{K}\mid t_0,\mathbf{Q}_0,\mathbf{K}_0)/\partial t = (i/\hbar)[H_H(t)U^\dagger(t,t_0)\Delta(\mathbf{Q},\mathbf{K})U(t,t_0) - U^\dagger(t,t_0)\Delta(\mathbf{Q},\mathbf{K})U(t,t_0)H_H(t)],$$

so that

$$\begin{aligned}\partial P(t,\mathbf{Q},\mathbf{K}\mid t_0,\mathbf{Q}_0,\mathbf{K}_0)/\partial t &= (i/\hbar)\int\cdots\int d\mathbf{Q}_0'\,d\mathbf{K}_0'P(t,\mathbf{Q},\mathbf{K}\mid t_0,\mathbf{Q}_0',\mathbf{K}_0')\\ &\quad\times \mathrm{Tr}\,\{\Delta(\mathbf{Q}_0,\mathbf{K}_0)[H_H(t)\Delta(\mathbf{Q}_0',\mathbf{K}_0') - \Delta(\mathbf{Q}_0',\mathbf{K}_0')H_H(t)]\}.\end{aligned} \quad (3.27)$$

But (3.19) gives

$$\mathrm{Tr}\,\{\Delta(\mathbf{Q}_0,\mathbf{K}_0)[H_H(t)\Delta(\mathbf{Q}_0',\mathbf{K}_0') - \Delta(\mathbf{Q}_0',\mathbf{K}_0')H_H(t)]\} = \mathfrak{L}_H(t,\mathbf{Q}_0,\mathbf{K}_0)\delta(\mathbf{Q}_0-\mathbf{Q}_0')\delta(\mathbf{K}_0-\mathbf{K}_0'),$$

where

$$\mathfrak{L}_H(t,\mathbf{Q},\mathbf{K}) = 2i\int\cdots\int d\mathbf{Q}''\,d\mathbf{K}''\delta(\mathbf{Q}-\mathbf{Q}'')\delta(\mathbf{K}-\mathbf{K}'') \times \sin\{(1/4\pi)[(\partial/\partial\mathbf{Q}'')\cdot(\partial/\partial\mathbf{K}) - (\partial/\partial\mathbf{K}'')\cdot(\partial/\partial\mathbf{Q})]\} \times H_H(t,\mathbf{Q}'',\mathbf{K}''). \quad (3.28)$$

Therefore,

$$\partial P(t,\mathbf{Q},\mathbf{K}\mid t_0,\mathbf{Q}_0,\mathbf{K}_0)/\partial t = (i/\hbar)\mathfrak{L}_H(t,\mathbf{Q}_0,\mathbf{K}_0)P(t,\mathbf{Q},\mathbf{K}\mid t_0,\mathbf{Q}_0,\mathbf{K}_0), \quad (3.29)$$

with solution, for $t > t_0$,

$$P(t,\mathbf{Q},\mathbf{K}\mid t_0,\mathbf{Q}_0,\mathbf{K}_0) = T\exp\left[(i/\hbar)\int_{t_0}^{t}dt'\mathfrak{L}_H(t',\mathbf{Q}_0,\mathbf{K}_0)\right] \times \delta(\mathbf{Q}-\mathbf{Q}_0)\delta(\mathbf{K}-\mathbf{K}_0). \quad (3.30)$$

If the Hamiltonian is independent of time, then $H_s = H_H = H$, so that the quantum Liouville

operators of (3.20) and (3.28) are the same:

$$\mathfrak{L}_s(\mathbf{Q}, \mathbf{K}) = \mathfrak{L}_H(\mathbf{Q}, \mathbf{K}) = \mathfrak{L}(\mathbf{Q}, \mathbf{K}). \quad (3.31)$$

In this case the two expressions in (3.23) and (3.30) become, for $t > t_0$,

$$\begin{aligned} &P(t, \mathbf{Q}, \mathbf{K} \mid t_0, \mathbf{Q}_0, \mathbf{K}_0) \\ &\quad = \exp\,[-(i/\hbar)(t - t_0)\mathfrak{L}(\mathbf{Q}, \mathbf{K})]\delta(\mathbf{Q} - \mathbf{Q}_0)\delta(\mathbf{K} - \mathbf{K}_0) \\ &\quad = \exp\,[(i/\hbar)(t - t_0)\mathfrak{L}(\mathbf{Q}_0, \mathbf{K}_0)]\delta(\mathbf{Q} - \mathbf{Q}_0)\delta(\mathbf{K} - \mathbf{K}_0), \end{aligned} \quad (3.32)$$

and

$$P(t, \mathbf{Q}, \mathbf{K} \mid t_0, \mathbf{Q}_0, \mathbf{K}_0) = P(t_0, \mathbf{Q}_0, \mathbf{K}_0 \mid t, \mathbf{Q}, \mathbf{K}). \quad (3.33)$$

Equation (3.33) is equivalent to Moyal's equation (9.10)[3] for conservative systems; but it is not valid when the Hamiltonian is time dependent, because of the time-ordering operations in (3.23) and (3.30).

A Heisenberg operator $A_H(t)$ is defined as

$$A_H(t) = U^\dagger(t, t_0)A_s(t)U(t, t_0). \quad (3.34)$$

In particular, according to (3.5), the Heisenberg density operator is

$$\rho_H = \rho_s(t_0), \quad (3.35)$$

independent of time. The Weyl transform of $A_H(t)$ is

$$\begin{aligned} A_H(t, \mathbf{Q}, \mathbf{K}) &> \mathrm{Tr}\,[\Delta(\mathbf{Q}, \mathbf{K})U^\dagger(t, t_0)A_s(t)U(t, t_0)] \\ &= \int\cdots\int d\mathbf{Q}_0\, d\mathbf{K}_0 A_s(t, \mathbf{Q}_0, \mathbf{K}_0) \\ &\qquad \times \mathrm{Tr}\,[\Delta(\mathbf{Q}_0, \mathbf{K}_0)U(t, t_0)\Delta(\mathbf{Q}, \mathbf{K})U^\dagger(t, t_0)] \\ &= \int\cdots\int d\mathbf{Q}_0\, d\mathbf{K}_0 A_s(t, \mathbf{Q}_0, \mathbf{K}_0) \\ &\qquad \times P(t, \mathbf{Q}_0, \mathbf{K}_0 \mid t_0, \mathbf{Q}, \mathbf{K}). \end{aligned} \quad (3.36)$$

The Weyl transform of the Heisenberg density operator is time independent

$$\rho_H(\mathbf{Q}, \mathbf{K}) = \rho_s(t_0, \mathbf{Q}, \mathbf{K}). \quad (3.37)$$

The expectation of the dynamical property given in (3.3) becomes

$$\begin{aligned} \langle A(t)\rangle &= \int\cdots\int d\mathbf{Q}\, d\mathbf{K}\, d\mathbf{Q}_0\, d\mathbf{K}_0 A_s(t, \mathbf{Q}, \mathbf{K}) \\ &\qquad \times P(t, \mathbf{Q}, \mathbf{K} \mid t_0, \mathbf{Q}_0, \mathbf{K}_0)\rho_s(t, \mathbf{Q}_0, \mathbf{K}_0) \\ &= \int\cdots\int d\mathbf{Q}_0\, d\mathbf{K}_0 A_H(t, \mathbf{Q}_0, \mathbf{K}_0)\rho_H(\mathbf{Q}_0, \mathbf{K}_0) \\ &= \mathrm{Tr}\,[A_H(t)\rho_H]. \end{aligned} \quad (3.38)$$

From (3.36) and (3.30)

$$\begin{aligned} &A_H(t, \mathbf{Q}, \mathbf{K}) \\ &\quad = T\exp\left[(i/\hbar)\int_{t_0}^{t} dt'\mathfrak{L}_H(t', \mathbf{Q}, \mathbf{K})\right]A_s(t, \mathbf{Q}, \mathbf{K}). \end{aligned} \quad (3.39)$$

From (3.36) and (3.17)

$$\begin{aligned} &A_s(t, \mathbf{Q}, \mathbf{K}) \\ &\quad = \int\cdots\int d\mathbf{Q}_0\, d\mathbf{K}_0 A_H(t, \mathbf{Q}_0, \mathbf{K}_0)P(t, \mathbf{Q}, \mathbf{K} \mid t_0, \mathbf{Q}_0, \mathbf{K}_0), \end{aligned} \quad (3.40)$$

so that, according to (3.23), the transformation inverse to (3.39) is

$$\begin{aligned} &A_s(t, \mathbf{Q}, \mathbf{K}) \\ &\quad = T\exp\left[-(i/\hbar)\int_{t_0}^{t} dt'\mathfrak{L}_s(t', \mathbf{Q}, \mathbf{K})\right]A_H(t, \mathbf{Q}, \mathbf{K}). \end{aligned} \quad (3.41)$$

The change from $\mathfrak{L}_H$ in (3.39) to $\mathfrak{L}_s$ in (3.41) should be noted; for conservative systems (3.31) applies, so that this distinction disappears. Equation (3.39) is a solution of the differential equation

$$\begin{aligned} &\partial A_H(t, \mathbf{Q}, \mathbf{K})/\partial t \\ &\quad = T\exp\left[(i/\hbar)\int_{t_0}^{t} dt'\mathfrak{L}_H(t', \mathbf{Q}, \mathbf{K})\right]\partial A_s(t, \mathbf{Q}, \mathbf{K})/\partial t \\ &\qquad + (i/\hbar)\mathfrak{L}_H(t, \mathbf{Q}, \mathbf{K})A_H(t, \mathbf{Q}, \mathbf{K}), \end{aligned} \quad (3.42)$$

which is the Weyl transform of the quantum-mechanical Heisenberg equation for the operator $A_H(t)$,

$$\begin{aligned} dA_H(t)/dt &= \partial A_H(t)/\partial t + (i/\hbar) \\ &\qquad \times [H_H(t)A_H(t) - A_H(t)H_H(t)]. \end{aligned} \quad (3.43)$$

In addition to (3.21) and (3.29), two other expressions can be derived for $\partial P(t, \mathbf{Q}, \mathbf{K} \mid t_0, \mathbf{Q}_0, \mathbf{K}_0)/\partial t$. If $H_s(t)$ is replaced by $U(t, t_0)H_H(t)U^\dagger(t, t_0)$ in the equations leading to (3.18), it is found that

$$\begin{aligned} &\partial P(t, \mathbf{Q}, \mathbf{K} \mid t_0, \mathbf{Q}_0, \mathbf{K}_0)/\partial t \\ &\quad = -(i/\hbar)\int\cdots\int d\mathbf{Q}_0'\, d\mathbf{K}_0' P(t, \mathbf{Q}, \mathbf{K} \mid t_0, \mathbf{Q}_0', \mathbf{K}_0') \\ &\qquad \times \mathfrak{L}_H(t, \mathbf{Q}_0', \mathbf{K}_0')\delta(\mathbf{Q}_0 - \mathbf{Q}_0')\delta(\mathbf{K}_0 - \mathbf{K}_0'). \end{aligned} \quad (3.44)$$

Similarly, replacement of $H_H(t)$ by

$$U^\dagger(t, t_0)H_s(t)U(t, t_0)$$

in the equations leading to (3.27) gives

$$\begin{aligned} &\partial P(t, \mathbf{Q}, \mathbf{K} \mid t_0, \mathbf{Q}_0, \mathbf{K}_0)/\partial t \\ &\quad = (i/\hbar)\int\cdots\int d\mathbf{Q}'\, d\mathbf{K}' P(t, \mathbf{Q}', \mathbf{K}' \mid t_0, \mathbf{Q}_0, \mathbf{K}_0) \\ &\qquad \times \mathfrak{L}_s(t, \mathbf{Q}', \mathbf{K}')\delta(\mathbf{Q} - \mathbf{Q}')\delta(\mathbf{K} - \mathbf{K}'). \end{aligned} \quad (3.45)$$

From (3.44) and (3.11),

$$\partial\rho_s(t, \mathbf{Q}, \mathbf{K})/\partial t = -(i/\hbar)\int\cdots\int d\mathbf{Q}_0\, d\mathbf{K}_0 P(t, \mathbf{Q}, \mathbf{K} \mid t_0, \mathbf{Q}_0, \mathbf{K}_0) \times \mathfrak{L}_H(t, \mathbf{Q}_0, \mathbf{K}_0)\rho_H(\mathbf{Q}_0, \mathbf{K}_0), \quad (3.46)$$

an alternative to (3.25). From (3.45) and (3.36),

$$\partial A_H(t, \mathbf{Q}, \mathbf{K})/\partial t = \int\cdots\int d\mathbf{Q}_0\, d\mathbf{K}_0 P(t, \mathbf{Q}_0, \mathbf{K}_0 \mid t_0, \mathbf{Q}, \mathbf{K}) \times [\partial A_s(t, \mathbf{Q}_0, \mathbf{K}_0)/\partial t + (i/\hbar)\mathfrak{L}_s(t, \mathbf{Q}_0, \mathbf{K}_0)A_s(t, \mathbf{Q}_0, \mathbf{K}_0)], \quad (3.47)$$

an alternative to (3.42). According to (3.40), the right-hand side of (3.46) is in the form of the Weyl transform of a Schrödinger operator; according to (3.36), the right-hand side of (3.47) is in the form of the Weyl transform of a Heisenberg operator.

Equation (3.47) is particularly suitable for comparing quantum-dynamical relations with those of classical dynamics. Substitution of the time-evolution equation for the Wigner function (3.25) into (3.47) gives the quantum form of the Liouville equation,

$$\partial\rho_H(\mathbf{Q}, \mathbf{K})/\partial t = 0, \quad (3.48)$$

in agreement with (3.37). The time-rate of change of the Weyl transform of the Heisenberg Hamiltonian is determined by the explicit time dependence of $H_s(t, \mathbf{Q}, \mathbf{K})$; (3.47) gives

$$\partial H_H(t, \mathbf{Q}, \mathbf{K})/\partial t = \int\cdots\int d\mathbf{Q}_0\, d\mathbf{K}_0 P(t, \mathbf{Q}_0, \mathbf{K}_0 \mid t_0, \mathbf{Q}, \mathbf{K}) \times \partial H_s(t, \mathbf{Q}_0, \mathbf{K}_0)/\partial t. \quad (3.49)$$

Similarly, (3.47) gives the quantum form of Hamilton's canonical equations,

$$\partial\mathbf{Q}_H(t, \mathbf{Q}, \mathbf{K})/\partial t = \int\cdots\int d\mathbf{Q}_0\, d\mathbf{K}_0 P(t, \mathbf{Q}_0, \mathbf{K}_0 \mid t_0, \mathbf{Q}, \mathbf{K})(2\pi\hbar)^{-1} \times \partial H_s(t, \mathbf{Q}_0, \mathbf{K}_0)/\partial\mathbf{K}_0$$

$$\partial\mathbf{K}_H(t, \mathbf{Q}, \mathbf{K})/\partial t = -\int\cdots\int d\mathbf{Q}_0\, d\mathbf{K}_0 P(t, \mathbf{Q}_0, \mathbf{K}_0 \mid t_0, \mathbf{Q}, \mathbf{K})(2\pi\hbar)^{-1} \times \partial H_s(t, \mathbf{Q}_0, \mathbf{K}_0)/\partial\mathbf{Q}_0. \quad (3.50)$$

The Ehrenfest equations[7] for $d\langle\mathbf{Q}(t)\rangle/dt$ and $d\langle\mathbf{K}(t)\rangle/dt$ follow directly on averaging over the initial phase-space distribution, with the use of (3.38) and (3.37). It will, in fact, be shown at the end of Sec. 4, by the use of the phase-space path integral form of (3.36), that

$$\partial A_H(t, \mathbf{Q}, \mathbf{K})/\partial\mathbf{Q} = \int\cdots\int d\mathbf{Q}_0\, d\mathbf{K}_0 P(t, \mathbf{Q}_0, \mathbf{K}_0 \mid t_0, \mathbf{Q}, \mathbf{K}) \times \partial A_s(t, \mathbf{Q}_0, \mathbf{K}_0)/\partial\mathbf{Q}_0,$$

$$\partial A_H(t, \mathbf{Q}, \mathbf{K})/\partial\mathbf{K} = \int\cdots\int d\mathbf{Q}_0\, d\mathbf{K}_0 P(t, \mathbf{Q}_0, \mathbf{K}_0 \mid t_0, \mathbf{Q}, \mathbf{K}) \times \partial A_s(t, \mathbf{Q}_0, \mathbf{K}_0)/\partial\mathbf{K}_0. \quad (3.51)$$

so that (3.50) takes precisely the canonical form,

$$\partial\mathbf{Q}_H(t, \mathbf{Q}, \mathbf{K})/\partial t = (2\pi\hbar)^{-1}\partial H_H(t, \mathbf{Q}, \mathbf{K})/\partial\mathbf{K},$$
$$\partial\mathbf{K}_H(t, \mathbf{Q}, \mathbf{K})/\partial t = -(2\pi\hbar)^{-1}\partial H_H(t, \mathbf{Q}, \mathbf{K})/\partial\mathbf{Q}. \quad (3.52)$$

In the classical limit,[1] according to (3.28), the Liouville operator $\mathfrak{L}_H(t, \mathbf{Q}, \mathbf{K})$ becomes

$$\mathfrak{L}_H(t, \mathbf{Q}, \mathbf{K}) \to (i/2\pi)\{[\partial H_H(t, \mathbf{Q}, \mathbf{K})/\partial\mathbf{Q}] \cdot (\partial/\partial\mathbf{K}) - [\partial H_H(t, \mathbf{Q}, \mathbf{K})/\partial\mathbf{K}] \cdot (\partial/\partial\mathbf{Q})\}$$
$$= (\hbar/i)\{[\partial\mathbf{K}_H(t, \mathbf{Q}, \mathbf{K})/\partial t] \cdot (\partial/\partial\mathbf{K}) + [\partial\mathbf{Q}_H(t, \mathbf{Q}, \mathbf{K})/\partial t] \cdot (\partial/\partial\mathbf{Q})\}. \quad (3.53)$$

In (3.30), the operator becomes

$$T\exp\left[(i/\hbar)\int_{t_0}^{t} dt'\mathfrak{L}_H(t', \mathbf{Q}, \mathbf{K}) \to T\exp\int_{t_0}^{t} dt'\{[\partial\mathbf{K}_H(t', \mathbf{Q}, \mathbf{K})/\partial t'] \cdot (\partial/\partial\mathbf{K}) + [\partial\mathbf{Q}_H(t', \mathbf{Q}, \mathbf{K})/\partial t'] \cdot (\partial/\partial\mathbf{Q})\}\right], \quad (3.54)$$

a displacement operator along the classical trajectory. If the coordinates along the trajectory at time $t$ are defined as

$$\mathbf{Q}(t) \equiv T\exp\left\{\int_{t_0}^{t} dt'[\partial\mathbf{Q}_H(t', \mathbf{Q}, \mathbf{K})/\partial t'] \cdot (\partial/\partial\mathbf{Q})\right\}\mathbf{Q},$$
$$\mathbf{K}(t) \equiv T\exp\left\{\int_{t_0}^{t} dt'[\partial\mathbf{K}_H(t', \mathbf{Q}, \mathbf{K})/\partial t'] \cdot (\partial/\partial\mathbf{K})\right\}\mathbf{K}, \quad (3.55)$$

then, in the classical limit,

$$P(t, \mathbf{Q}_0, \mathbf{K}_0 \mid t_0, \mathbf{Q}, \mathbf{K}) \to \delta[\mathbf{Q}(t) - \mathbf{Q}_0]\delta[\mathbf{K}(t) - \mathbf{K}_0], \quad (3.56)$$

so that, from (3.36),

$$A_H(t, \mathbf{Q}, \mathbf{K}) \to A_s[t, \mathbf{Q}(t), \mathbf{K}(t)], \quad (3.57)$$

the classical dynamical variable for the system along its trajectory in phase space at time $t$. In (3.55), initially, $\mathbf{Q}(t_0) = \mathbf{Q}$ and $\mathbf{K}(t_0) = \mathbf{K}$.

## 4. MARKOFFIAN PROPERTIES AND PHASE-SPACE PATH INTEGRALS

When the Hamiltonian operator is time dependent, two forms can be distinguished for the operators $U(t, t_0)$ and $U^\dagger(t, t_0)$; these are given in (3.8) and (3.9). It is convenient to designate the forms in which $H_s$ appears as $U_s(t, t_0)$ and $U_s^\dagger(t, t_0)$, the forms in which $H_H$ appears as $U_H(t, t_0)$ and $U_H^\dagger(t, t_0)$. If $t > t_1 > t_0$, then

$$\begin{aligned} U_s(t, t_0) &= U_s(t, t_1)U_s(t_1, t_0), \\ U_s^\dagger(t, t_0) &= U_s^\dagger(t_1, t_0)U_s^\dagger(t, t_1). \end{aligned} \tag{4.1}$$

But

$$\begin{aligned} U_H^\dagger(t, t_0) &= U_H^\dagger(t, t_1)U_H^\dagger(t_1, t_0), \\ U_H(t, t_0) &= U_H(t_1, t_0)U_H(t, t_1). \end{aligned} \tag{4.2}$$

From (4.1) and (3.12),

$$\begin{aligned} P(t, \mathbf{Q}, \mathbf{K} \mid t_0, \mathbf{Q}_0, \mathbf{K}_0) &= \mathrm{Tr}\,[\Delta(\mathbf{Q}, \mathbf{K})U_s(t, t_1)U_s(t_1, t_0) \\ &\quad \times \Delta(\mathbf{Q}_0, \mathbf{K}_0)U_s^\dagger(t_1, t_0)U_s^\dagger(t, t_1)] \\ &= \mathrm{Tr}\,[U_s^\dagger(t, t_1)\Delta(\mathbf{Q}, \mathbf{K})U_s(t, t_1) \\ &\quad \times U_s(t_1, t_0)\Delta(\mathbf{Q}_0, \mathbf{K}_0)U_s^\dagger(t_1, t_0)] \\ &= \int\cdots\int d\mathbf{Q}_1\, d\mathbf{K}_1 P_s(t, \mathbf{Q}, \mathbf{K} \mid t_1, \mathbf{Q}_1, \mathbf{K}_1) \\ &\quad \times P(t_1, \mathbf{Q}_1, \mathbf{K}_1 \mid t_0, \mathbf{Q}_0, \mathbf{K}_0). \end{aligned} \tag{4.3}$$

On the other hand, from (4.2),

$$\begin{aligned} P(t, \mathbf{Q}_0, \mathbf{K}_0 \mid t_0, \mathbf{Q}, \mathbf{K}) &= \mathrm{Tr}\,[\Delta(\mathbf{Q}_0, \mathbf{K}_0)U_H(t_1, t_0)U_H(t, t_1) \\ &\quad \times \Delta(\mathbf{Q}, \mathbf{K})U_H^\dagger(t, t_1)U_H^\dagger(t_1, t_0)] \\ &= \mathrm{Tr}\,[U_H^\dagger(t, t_0)\Delta(\mathbf{Q}_0, \mathbf{K}_0)U_H(t_1, t_0) \\ &\quad \times U_H(t, t_1)\Delta(\mathbf{Q}, \mathbf{K})U_H^\dagger(t, t_1)] \\ &= \int\cdots\int d\mathbf{Q}_1\, d\mathbf{K}_1 P_H(t, \mathbf{Q}_1, \mathbf{K}_1 \mid t_1, \mathbf{Q}, \mathbf{K}) \\ &\quad \times P_H(t_1, \mathbf{Q}_0, \mathbf{K}_0 \mid t_0, \mathbf{Q}_1, \mathbf{K}_1). \end{aligned} \tag{4.4}$$

(Here, $P_s$ designates a quantity containing $U_s$; $P_H$, a quantity containing $U_H$.) For a sequence of times, $t = t_n > t_{n-1} > \cdots t_1 > t_0$, and with $\mathbf{Q}_n = \mathbf{Q}, \mathbf{K}_n = \mathbf{K}$,

$$\begin{aligned} P(t, \mathbf{Q}, \mathbf{K} \mid t_0, \mathbf{Q}_0, \mathbf{K}_0) &= \int\cdots\int d\mathbf{Q}_1\, d\mathbf{K}_1 \cdots d\mathbf{Q}_{n-1}\, d\mathbf{K}_{n-1} \\ &\quad \times \prod_{j=0}^{n-1} P_s(t_{j+1}, \mathbf{Q}_{j+1}, \mathbf{K}_{j+1} \mid t_j, \mathbf{Q}_j, \mathbf{K}_j), \end{aligned} \tag{4.5}$$

$$\begin{aligned} P(t, \mathbf{Q}_0, \mathbf{K}_0 \mid t_0, \mathbf{Q}, \mathbf{K}) &= \int\cdots\int d\mathbf{Q}_1\, d\mathbf{K}_1 \cdots d\mathbf{Q}_{n-1}\, d\mathbf{K}_{n-1} \\ &\quad \times \prod_{j=0}^{n-1} P_H(t_{j+1}, \mathbf{Q}_j, \mathbf{K}_j \mid t_j, \mathbf{Q}_{j+1}, \mathbf{K}_{j+1}). \end{aligned} \tag{4.6}$$

According to (4.5) the propagation of probability in phase space is a Markoffian stochastic process[3,8] which can be written as a continuous product of unitary transformations; from (3.11),

$$\begin{aligned} \rho_s(t, \mathbf{Q}, \mathbf{K}) &= \lim_{n\to\infty} \int\cdots\int d\mathbf{Q}_0\, d\mathbf{K}_0 \cdots d\mathbf{Q}_{n-1}\, d\mathbf{K}_{n-1} \\ &\quad \times \prod_{j=0}^{n-1} P_s(t_{j+1}, \mathbf{Q}_{j+1}, \mathbf{K}_{j+1} \mid t_j, \mathbf{Q}_j, \mathbf{K}_j)\rho_s(t_0, \mathbf{Q}_0, \mathbf{K}_0). \end{aligned} \tag{4.7}$$

According to (3.40) the transformation from Heisenberg to Schrödinger pictures proceeds by the same process; in (4.7), $\rho_s(t, \mathbf{Q}_0, \mathbf{K}_0)$ can be replaced by $\rho_H(\mathbf{Q}_0, \mathbf{K}_0)$. On the other hand, the inverse transformation from Schrödinger to Heisenberg pictures goes by way of (4.6); from (3.36),

$$\begin{aligned} A_H(t, \mathbf{Q}, \mathbf{K}) &= \lim_{n\to\infty} \int\cdots\int d\mathbf{Q}_0\, d\mathbf{K}_0 \cdots d\mathbf{Q}_{n-1}\, d\mathbf{K}_{n-1} \\ &\quad \times \prod_{j=0}^{n-1} P_H(t_{j+1}, \mathbf{Q}_j, \mathbf{K}_j \mid t_j, \mathbf{Q}_{j+1}, \mathbf{K}_{j+1})A_s(t, \mathbf{Q}_0, \mathbf{K}_0). \end{aligned} \tag{4.8}$$

Equations (4.7) and (4.8) appear as weighted averages over all paths in phase space connecting $(\mathbf{Q}_0, \mathbf{K}_0)$ at $t_0$ with $(\mathbf{Q}, \mathbf{K})$ at $t$, analogous to the Feynman path integrals in configuration space.[5] The weight attached to each path is the product of conditional probabilities for infinitesimal displacements

$$\prod_{j=0}^{n-1} P_s(t_{j+1}, \mathbf{Q}_{j+1}, \mathbf{K}_{j+1} \mid t_j, \mathbf{Q}_j, \mathbf{K}_j)$$

in (4.7), or

$$\prod_{j=0}^{n-1} P_H(t_{j+1}, \mathbf{Q}_j, \mathbf{K}_j \mid t_j, \mathbf{Q}_{j+1}, \mathbf{K}_{j+1})$$

in (4.8). In the limit as $n \to \infty$, the largest time interval $t_{j+1} - t_j$ vanishes. Accordingly, in each factor only terms to first order in $t_{j+1} - t_j$ need be retained. From (3.23) and (3.30), to first order in $t_{j+1} - t_j$

$$\begin{aligned} P_s(t_{j+1}, \mathbf{Q}_{j+1}, \mathbf{K}_{j+1} \mid t_j, \mathbf{Q}_j, \mathbf{K}_j) &= \exp\left[-(i/\hbar)\int_{t_j}^{t_{j+1}} dt' \mathfrak{L}_s(t', \mathbf{Q}_{j+1}, \mathbf{K}_{j+1})\right] \\ &\quad \times \delta(\mathbf{Q}_{j+1} - \mathbf{Q}_j)\delta(\mathbf{K}_{j+1} - \mathbf{K}_j), \\ P_H(t_{j+1}, \mathbf{Q}_j, \mathbf{K}_j \mid t_j, \mathbf{Q}_{j+1}, \mathbf{K}_{j+1}) &= \exp\left[(i/\hbar)\int_{t_j}^{t_{j+1}} dt' \mathfrak{L}_H(t', \mathbf{Q}_{j+1}, \mathbf{K}_{j+1})\right] \\ &\quad \times \delta(\mathbf{Q}_{j+1} - \mathbf{Q}_j)\delta(\mathbf{K}_{j+1} - \mathbf{K}_j). \end{aligned} \tag{4.9}$$

[8] M. C. Wang and G. E. Uhlenbeck, Rev. Mod. Phys. 17, 323 (1945).

In the classical limit, $\mathfrak{L}(t, \mathbf{Q}, \mathbf{K})/\hbar$ remains finite in (3.23) and (3.30), as seen, for example, in (3.53). The corresponding phase in the Feynman path integrals in configuration space diverges along all paths except the classical dynamical path where the action is stationary. In the classical limit, $\mathfrak{L}(t, \mathbf{Q}, \mathbf{K})/\hbar$ becomes the classical Poisson-bracket Liouville operator, which contains only first derivatives with respect to $\mathbf{Q}$ and $P = 2\pi\hbar\mathbf{K}$. Quantum effects arise from the higher-order derivatives in the quantum Liouville operators (3.20) and (3.28). According to (4.9), these effects persist, to first order in $t_{j+1} - t_j$, along the infinitesimal portion of the phase-space path between $(\mathbf{Q}_{j+1}, \mathbf{K}_{j+1})$ and $(\mathbf{Q}_j, \mathbf{K}_j)$.

The proof of (3.51) follows from the need to retain only terms to first order in the infinitesimal time intervals, $t_{j+1} - t_j$. As Feynman has pointed out, contributions from higher-order terms vanish in the limit $n \to \infty$. From (4.8) and (4.9)

$$\begin{aligned}
&\partial A_H(t, \mathbf{Q}, \mathbf{K})/\partial\mathbf{Q} \\
&= \lim_{n\to\infty} \int\cdots\int d\mathbf{Q}_0\, d\mathbf{K}_0 \cdots d\mathbf{Q}_{n-1}\, d\mathbf{K}_{n-1} (\partial/\partial\mathbf{Q}_n) \\
&\quad \times \exp\left[(i/\hbar)\int_{t_{n-1}}^{t_n} dt' \mathfrak{L}_H(t', \mathbf{Q}_n, \mathbf{K}_n)\right] \\
&\quad \times \delta(\mathbf{Q}_n - \mathbf{Q}_{n-1})\delta(\mathbf{K}_n - \mathbf{K}_{n-1}) \\
&\quad \times \exp\left[(i/\hbar)\int_{t_{n-2}}^{t_{n-1}} dt' \mathfrak{L}_H(t', \mathbf{Q}_{n-1}, \mathbf{K}_{n-1})\right] \\
&\quad \times \delta(\mathbf{Q}_{n-1} - \mathbf{Q}_{n-2})\delta(\mathbf{K}_{n-1} - \mathbf{K}_{n-2}) \cdots A_s(t, \mathbf{Q}_0, \mathbf{K}_0).
\end{aligned} \tag{4.10}$$

To first order in $t_n - t_{n-1}$,

$$\begin{aligned}
&(\partial/\partial\mathbf{Q}_n) \exp\left[(i/\hbar)\int_{t_{n-1}}^{t_n} dt' \mathfrak{L}_H(t', \mathbf{Q}_n, \mathbf{K}_n)\right] \\
&\qquad \times \delta(\mathbf{Q}_n - \mathbf{Q}_{n-1})\delta(\mathbf{K}_n - \mathbf{K}_{n-1}) \\
&= \exp\left[(i/\hbar)\int_{t_{n-1}}^{t_n} dt' \mathfrak{L}_H(t', \mathbf{Q}_n, \mathbf{K}_n)\right](\partial/\partial\mathbf{Q}_n) \\
&\qquad \times \delta(\mathbf{Q}_n - \mathbf{Q}_{n-1})\delta(\mathbf{K}_n - \mathbf{K}_{n-1}),
\end{aligned} \tag{4.11}$$

since differentiation of the exponential gives a factor $t_n - t_{n-1}$. Therefore, an integration by parts with respect to $\mathbf{Q}_{n-1}$, gives

$$\begin{aligned}
&\partial A_H(t, \mathbf{Q}, \mathbf{K})/\partial\mathbf{Q} \\
&= \lim_{n\to\infty} \int\cdots\int d\mathbf{Q}_0\, d\mathbf{K}_0 \cdots d\mathbf{Q}_{n-1}\, d\mathbf{K}_{n-1} \\
&\quad \times \exp\left[(i/\hbar)\int_{t_{n-1}}^{t_n} dt' \mathfrak{L}_H(t', \mathbf{Q}_n, \mathbf{K}_n)\right] \\
&\quad \times \delta(\mathbf{Q}_n - \mathbf{Q}_{n-1})\delta(\mathbf{K}_n - \mathbf{K}_{n-1})(\partial/\partial\mathbf{Q}_{n-1}) \\
&\quad \times \exp\left[(i/\hbar)\int_{t_{n-2}}^{t_{n-1}} dt' \mathfrak{L}_H(t', \mathbf{Q}_{n-1}, \mathbf{K}_{n-1})\right] \\
&\quad \times \delta(\mathbf{Q}_{n-1} - \mathbf{Q}_{n-2})\delta(\mathbf{K}_{n-1} - \mathbf{K}_{n-2}) \cdots A_s(t, \mathbf{Q}_0, \mathbf{K}_0).
\end{aligned} \tag{4.12}$$

Repetition of this process moves the differential down the chain until $(\partial/\partial\mathbf{Q}_0)A_s(t, \mathbf{Q}_0, \mathbf{K}_0)$ is produced at the right-hand end; (3.51) is proved.

## 5. COMMUTIVITY OF $\rho_H$ WITH $\mathbf{q}$, AND FEYNMAN PATH INTEGRALS

A case of special interest arises when the density operator in Heisenberg picture $\rho_H = \rho_s(t_0)$ commutes with the coordinate operator $\mathbf{q}$. Since $\rho_H$ is independent of time, the commutativity is time independent. In this case, for any two eigenkets of $\mathbf{q}$, $|\mathbf{Q}\rangle$ and $|\mathbf{Q}'\rangle$,

$$\langle\mathbf{Q}'|\, \rho_H\mathbf{q} - \mathbf{q}\rho_H\, |\mathbf{Q}\rangle = (\mathbf{Q} - \mathbf{Q}')\langle\mathbf{Q}'|\, \rho_H\, |\mathbf{Q}\rangle = 0. \tag{5.1}$$

Therefore, the matrix elements of $\rho_H$ have the form

$$\langle\mathbf{Q}'|\, \rho_H\, |\mathbf{Q}\rangle = \rho_H(\mathbf{Q})\delta(\mathbf{Q}' - \mathbf{Q}), \tag{5.2}$$

so that

$$\begin{aligned}
\rho_H &= \int\cdots\int d\mathbf{Q}\, d\mathbf{Q}'\, |\mathbf{Q}'\rangle\, \rho_H(\mathbf{Q})\delta(\mathbf{Q}' - \mathbf{Q})\langle\mathbf{Q}| \\
&= \int d\mathbf{Q}\, |\mathbf{Q}\rangle\, \rho_H(\mathbf{Q})\, \langle\mathbf{Q}| \\
&= \int d\mathbf{Q}\rho_H(\mathbf{Q})\delta(\mathbf{q} - \mathbf{Q}).
\end{aligned} \tag{5.3}$$

$\rho_H$ is a function of the operator $\mathbf{q}$ only. According to (2.8),

$$\rho_H = \int\cdots\int d\mathbf{Q}\, d\mathbf{K}\rho_H(\mathbf{Q})\Delta(\mathbf{Q}, \mathbf{K})$$

so that $\rho_H(\mathbf{Q})$ is the Weyl transform of $\rho_H$ when $\rho_H$ commutes with $\mathbf{q}$. According to (5.2), $\rho_H(\mathbf{Q})$ is an eigenvalue of $\rho_H$,

$$\rho_H\, |\mathbf{Q}\rangle = \int d\mathbf{Q}'\, |\mathbf{Q}'\rangle\langle\mathbf{Q}'|\, \rho_H\, |\mathbf{Q}\rangle = \rho_H(\mathbf{Q})\, |\mathbf{Q}\rangle, \tag{5.4}$$

so that the spectrum of $\rho_H$ is continuous. Since $\rho_H$ is a function of $\mathbf{q}$ only, the inverse Weyl transformation (2.2) gives

$$\rho_H(\mathbf{Q}) = \mathrm{Tr}\, [\rho_H\Delta(\mathbf{Q}, \mathbf{K})] = \mathrm{Tr}\, [\rho_H\delta(\mathbf{q} - \mathbf{Q})]. \tag{5.5}$$

By definition, the last equality,

$$\mathrm{Tr}\, [\rho_H\delta(\mathbf{q} - \mathbf{Q})] \equiv \int d\mathbf{Q}'\, \langle\mathbf{Q}'|\, \rho_H\, |\mathbf{Q}\rangle, \tag{5.6}$$

since in the continuous spectrum only the diagonal matrix element $\mathbf{Q}' = \mathbf{Q}$ contributes to the trace, because of the $\delta$-function singularity of the matrix elements (5.2).

For any operator $A$, the Weyl transform gives, with (2.5),

$$A = \int\cdots\int d\mathbf{Q}\, d\mathbf{K} A(\mathbf{Q},\mathbf{K}) \times \int d\mathbf{v} \exp(2\pi i \mathbf{v}\cdot\mathbf{K}) |\mathbf{Q} + \tfrac{1}{2}\mathbf{v}\rangle\langle\mathbf{Q} - \tfrac{1}{2}\mathbf{v}|$$
$$= \int\cdots\int d\mathbf{Q}'\, d\mathbf{Q}''\, d\mathbf{K} A[\tfrac{1}{2}(\mathbf{Q}' + \mathbf{Q}''), \mathbf{K}] \times \exp[2\pi i \mathbf{K}\cdot(\mathbf{Q}' - \mathbf{Q}'')] |\mathbf{Q}'\rangle\langle\mathbf{Q}''|$$
$$= \int\cdots\int d\mathbf{Q}'\, d\mathbf{Q}'' |\mathbf{Q}'\rangle\langle\mathbf{Q}''| A[\tfrac{1}{2}(\mathbf{Q}' + \mathbf{Q}''), (i/2\pi)(\partial/\partial\mathbf{Q}'')]\delta(\mathbf{Q}'' - \mathbf{Q}'). \quad (5.7)$$

If $A$ commutes with $\mathbf{q}$, its Weyl transform is $A(\mathbf{Q})$, so that (5.7) gives

$$A = \int\cdots\int d\mathbf{Q}'\, d\mathbf{Q}'' |\mathbf{Q}'\rangle\langle\mathbf{Q}''| \times A[\tfrac{1}{2}(\mathbf{Q}' + \mathbf{Q}'')]\delta(\mathbf{Q}'' - \mathbf{Q}')$$
$$= \int d\mathbf{Q} |\mathbf{Q}\rangle A(\mathbf{Q})\langle\mathbf{Q}|. \quad (5.8)$$

If $A$ commutes with $\mathbf{k}$, its Weyl transform is $A(\mathbf{K})$, so that (5.7) gives

$$A = \int\cdots\int d\mathbf{Q}'\, d\mathbf{Q}'' |\mathbf{Q}'\rangle\langle\mathbf{Q}''| \times A[(i/2\pi)(\partial/\partial\mathbf{Q}'')]\delta(\mathbf{Q}'' - \mathbf{Q}')$$
$$= \int d\mathbf{Q} |\mathbf{Q}\rangle A[(1/2\pi i)(\partial/\partial\mathbf{Q})]\langle\mathbf{Q}|. \quad (5.9)$$

Equation (5.3) is a special case of (5.8) where the operator $A$ which commutes with $\mathbf{q}$ is $\rho_H$. Just as $\rho_H(\mathbf{Q})$, the Weyl transform of $\rho_H$, is given by (5.5), so in general, if $A$ commutes with $\mathbf{q}$, its Weyl transform is

$$A(\mathbf{Q}) = \mathrm{Tr}\,[A\delta(\mathbf{q} - \mathbf{Q})]. \quad (5.10)$$

Equations (5.8) and (5.10) are the Weyl transformation and the inverse transformation for any operator which commutes with $\mathbf{q}$.

As shown in (5.3), if $\rho_H$ commutes with $\mathbf{q}$, then, in general, it is a mixture of pure coordinate states $|\mathbf{Q}\rangle\langle\mathbf{Q}|$ with weights $\rho_H(\mathbf{Q})$. The weights are nonnegative, and they are normalized to unity, since

$$\int d\mathbf{Q}\rho_H(\mathbf{Q}) = \int d\mathbf{Q}\, \mathrm{Tr}\,[\rho_H\delta(\mathbf{q} - \mathbf{Q})] = \mathrm{Tr}\,\rho_H = 1. \quad (5.11)$$

In this case, according to (3.38), the expectation at time $t$ of any observable $A(t)$ is

$$\langle A(t)\rangle = \int d\mathbf{Q}\mathcal{A}(t,\mathbf{Q})\rho_H(\mathbf{Q}), \quad (5.12)$$

where

$$\mathcal{A}(t,\mathbf{Q}) \equiv \int d\mathbf{K} A_H(t,\mathbf{Q},\mathbf{K}). \quad (5.13)$$

In a Heisenberg pure state for which $\rho_H = \delta(\mathbf{q} - \mathbf{Q}')$, $\rho_H(\mathbf{Q})$ equals $\delta(\mathbf{Q} - \mathbf{Q}')$, and $\langle A(t)\rangle = \mathcal{A}(t,\mathbf{Q}')$. $\mathcal{A}(t,\mathbf{Q})$ is, therefore, the expectation at time $t$ of the dynamical property of the system in the pure state $\rho_H = \delta(\mathbf{q} - \mathbf{Q})$. From (2.8), (3.36), and (3.12),

$$\mathcal{A}(t,\mathbf{Q}) = \mathrm{Tr}\,[A_H(t)\delta(\mathbf{q} - \mathbf{Q})]$$
$$= \int\cdots\int d\mathbf{Q}_0\, d\mathbf{K}_0 A_s(t,\mathbf{Q}_0,\mathbf{K}_0) \times \int d\mathbf{K} P(t,\mathbf{Q}_0,\mathbf{K}_0 |\, t_0,\mathbf{Q},\mathbf{K})$$
$$= \langle\mathbf{Q}|\, U^\dagger(t,t_0)A_s(t)U(t,t_0)\,|\mathbf{Q}\rangle. \quad (5.14)$$

If $A_s(t)$ commutes with $\mathbf{q}$, then according to (5.8),

$$\mathcal{A}(t,\mathbf{Q}) = \int d\mathbf{Q}_0 A_s(t,\mathbf{Q}_0)\, |\langle\mathbf{Q}_0|\, U(t,t_0)\,|\mathbf{Q}\rangle|^2. \quad (5.15)$$

If $A_s(t)$ commutes with $\mathbf{k}$, then according to (5.9),

$$\mathcal{A}(t,\mathbf{Q}) = \int d\mathbf{Q}_0 \langle\mathbf{Q}|\, U^\dagger(t,t_0)\,|\mathbf{Q}_0\rangle \times A_s[t,(1/2\pi i)(\partial/\partial\mathbf{Q}_0)]\langle\mathbf{Q}_0|\, u(t,t_0)\,|\mathbf{Q}\rangle. \quad (5.16)$$

Consider the case to which (5.15) applies, when $A_s(t)$ commutes with $\mathbf{q}$. According to (5.12), if $\rho_H$ also commutes with $\mathbf{q}$, then

$$\langle A(t)\rangle = \int\cdots\int d\mathbf{Q}_0\, d\mathbf{Q} A_s(t,\mathbf{Q}_0) \times |\langle\mathbf{Q}_0|\, U(t,t_0)\,|\mathbf{Q}\rangle|^2 \rho_s(t_0,\mathbf{Q})$$
$$= \int d\mathbf{Q} A_s(t,\mathbf{Q})\rho_s(t,\mathbf{Q}), \quad (5.17)$$

where

$$\rho_s(t,\mathbf{Q}) = \int d\mathbf{Q}_0\, |\langle\mathbf{Q}|\, U(t,t_0)\,|\mathbf{Q}_0\rangle|^2 \rho_s(t_0,\mathbf{Q}_0)$$
$$= \mathrm{Tr}\,[\rho_s(t)\delta(\mathbf{q} - \mathbf{Q})], \quad (5.18)$$

the reduced Wigner distribution function for configuration space. Comparison with (3.11) shows that if the probability of $\mathbf{Q}$ at $t$ conditional on $\mathbf{Q}_0$ at $t_0$ is defined by

$$P(t,\mathbf{Q}\,|\,t_0,\mathbf{Q}_0) \equiv \int\cdots\int d\mathbf{K}\, d\mathbf{K}_0 P(t,\mathbf{Q},\mathbf{K}\,|\,t_0,\mathbf{Q}_0,\mathbf{K}_0), \quad (5.19)$$

then $P(t,\mathbf{Q}\,|\,t_0,\mathbf{Q}_0)$ is the nonnegative quantity in (5.18),

$$P(t,\mathbf{Q}\,|\,t_0,\mathbf{Q}_0) = |\langle\mathbf{Q}|\, U(t,t_0)\,|\mathbf{Q}_0\rangle|^2. \quad (5.20)$$

This is the transition probability considered by Feynman.[5]

The probability amplitude $\langle \mathbf{Q}|\, U(t, t_0)\,|\mathbf{Q}_0\rangle$ can be exhibited as a Feynman path integral. From (4.1), with $t = t_n > t_{n-1} > \cdots t_1 > t_0$, and $\mathbf{Q}_n = \mathbf{Q}$,

$$\langle Q|\, U(t, t_0)\,|\, Q_0\rangle = \lim_{n\to\infty} \int\cdots\int d\mathbf{Q}_1 \cdots d\mathbf{Q}_{n-1} \times \prod_{j=0}^{n-1} \langle \mathbf{Q}_{j+1}|\, U_s(t_{j+1}, t_j)\,|\mathbf{Q}_j\rangle. \quad (5.21)$$

Now,

$$\langle \mathbf{Q}_{j+1}|\, U_s(t_{j+1}, t_j)\,|\mathbf{Q}_j\rangle = \langle \mathbf{Q}_j|\exp[2\pi i\mathbf{k}\cdot(\mathbf{Q}_{j+1}-\mathbf{Q}_j)]U_s(t_{j+1}, t_j)\,|\mathbf{Q}_j\rangle = \langle \mathbf{Q}_j|\exp\left\{(i/\hbar)\int_{t_j}^{t_{j+1}} dt' L_s[t', \dot{\mathbf{Q}}_j(t')]\right\}|\mathbf{Q}_j\rangle, \quad (5.22)$$

to first order in $t_{j+1} - t_j$, where the quantum-mechanical Lagrangian operator is defined as

$$L_s[t, \dot{\mathbf{Q}}_j(t)] \equiv 2\pi\hbar\mathbf{k}\cdot\dot{\mathbf{Q}}_j(t) - H_s(t), \quad (5.23)$$

with

$$\dot{\mathbf{Q}}_j(t_j) \equiv (\mathbf{Q}_{j+1} - \mathbf{Q}_j)/(t_{j+1} - t_j). \quad (5.24)$$

The matrix element (5.22) can be evaluated by the use of Weyl transforms. To first order in $t_{j+1} - t_j$,

$$U_s(t_{j+1}, t_j) = \mathbf{1} - (i/\hbar)(t_{j+1} - t_j)H_s(t_j),$$

with Weyl transform,

$$1 - (i/\hbar)(t_{j+1} - t_j)H_s(t_j\mathbf{Q}, \mathbf{K}) = \exp\left[-(i/\hbar)\int_{t_j}^{t_{j+1}} dt' H_s(t', \mathbf{Q}, \mathbf{K})\right].$$

Accordingly, to first order in $t_{j+1} - t_j$, from (2.6),

$$\langle \mathbf{Q}_{j+1}|\, U_s(t_{j+1}, t_j)\,|\mathbf{Q}_j\rangle = \int\cdots\int d\mathbf{Q}'\, d\mathbf{K}'\langle \mathbf{Q}_{j+1}|\,\Delta(\mathbf{Q}', \mathbf{K}')\,|\mathbf{Q}_j\rangle \times \exp\left[-(i/\hbar)\int_{t_j}^{t_{j+1}} dt' H_s(t, \mathbf{Q}', \mathbf{K}')\right] = \int d\mathbf{K}\exp\left\{-(i/\hbar)\int_{t_j}^{t_{j+1}} dt' H_s[t', \tfrac{1}{2}(\mathbf{Q}_j + \mathbf{Q}_{j+1}), \mathbf{K}]\right\} \times \exp[2\pi i\mathbf{K}\cdot(\mathbf{Q}_{j+1} - \mathbf{Q}_j)]. \quad (5.25)$$

The result of the integration on $\mathbf{K}$ will be a function of $\frac{1}{2}(\mathbf{Q}_j + \mathbf{Q}_{j+1})$ and $\dot{\mathbf{Q}}_j(t_j)$. It is Feynman's postulate[5] that this function has the form

$$\langle \mathbf{Q}_{j+1}|\, U_s(t_{j+1}, t_j)\,|\mathbf{Q}_j\rangle = \alpha^{-1}\exp\left\{(i/\hbar)\int_t^{t_{j+1}} dt' L_s[t', \dot{\mathbf{Q}}_j(t'), \tfrac{1}{2}(\mathbf{Q}_j + \mathbf{Q}_{j+1})]\right\}, \quad (5.26)$$

where $\alpha$ is a constant, and $L_s(t, \dot{\mathbf{Q}}_j, \mathbf{Q}_j)$ is the Lagrangian for the classical path joining $\mathbf{Q}_j$ at $t_j$ to $\mathbf{Q}_{j+1}$ at $t_{j+1}$. The postulated form is readily obtained if the Hamiltonian is quadratic in $\mathbf{K}$. For example, if

$$H_s = (2\pi\hbar\mathbf{k})^2/2m + V_s(\mathbf{q}), \quad (5.27)$$

for a particle of mass $m$ and potential energy $V_s(\mathbf{q})$, so that, from (2.5),

$$H_s(\mathbf{Q}, \mathbf{K}) = \exp[(i/4\pi)(\partial/\partial\mathbf{Q})\cdot(\partial/\partial\mathbf{K}] \times \langle \mathbf{K}\,|\,\mathbf{Q}\rangle\langle \mathbf{Q}|\, H_s\,|\mathbf{K}\rangle = (2\pi\hbar\mathbf{K})^2/2m + V_s(\mathbf{Q}),$$

then, from (5.25), to first order in $\Delta t = t_{j+1} - t_j$,

$$\langle \mathbf{Q}_{j+1}|\, U_s(t_{j+1}, t_j)\,|\mathbf{Q}_j\rangle = \exp\{-(i/\hbar)V_s[\tfrac{1}{2}(\mathbf{Q}_j + \mathbf{Q}_{j+1})]\Delta t\} \times \int d\mathbf{K}\exp\{(i/\hbar)[2\pi\hbar\mathbf{K}\cdot(\mathbf{Q}_{j+1} - \mathbf{Q}_j) - (2\pi\hbar\mathbf{K})^2\Delta t/2m]\} = (m/i2\pi\hbar\Delta t)^{\frac{3}{2}}\exp\{(i/\hbar)L_s[\dot{\mathbf{Q}}_j(t_j), \tfrac{1}{2}(\mathbf{Q}_j + \mathbf{Q}_{j+1})]\Delta t\}, \quad (5.28)$$

where the Lagrangian is

$$L_s[\dot{\mathbf{Q}}_j(t_j), \tfrac{1}{2}(\mathbf{Q}_j + \mathbf{Q}_{j+1})] = (m/2)\dot{\mathbf{Q}}_j^2(t_j) - V_s[\tfrac{1}{2}(\mathbf{Q}_j + \mathbf{Q}_{j+1})], \quad (5.29)$$

in agreement with Feynman.[9]

On the other hand, in (5.15) and (5.16), the relevant amplitude is $\langle \mathbf{Q}_0|\, U(t, t_0)\,|\mathbf{Q}\rangle$, not $\langle \mathbf{Q}|\, U(t, t_0)\,|\mathbf{Q}\rangle$. $\langle \mathbf{Q}_0|\, U(t, t_0)\,|\mathbf{Q}\rangle$ can also be exhibited as a Feynman path integral. From (4.2), with

$$t = t_n > t_{n-1} > \cdots t_1 > t_0,$$

and

$$\mathbf{Q}_n = \mathbf{Q},$$

$$\langle \mathbf{Q}_0|\, U(t, t_0)\,|\mathbf{Q}\rangle = \lim_{n\to\infty}\int\cdots\int d\mathbf{Q}_1\cdots d\mathbf{Q}_{n-1} \times \prod_{j=0}^{n-1}\langle \mathbf{Q}_j|\, U_H(t_{j+1}, t_j)\,|\mathbf{Q}_{j+1}\rangle \quad (5.30)$$

$$= \lim_{n\to\infty}\int\cdots\int d\mathbf{Q}_1\cdots d\mathbf{Q}_{n-1}\prod_{j=0}^{n-1}\langle \mathbf{Q}_j| \times \exp\left\{(i/\hbar)\int_{t_j}^{t_{j+1}} dt' L_H[t', \dot{\mathbf{Q}}_j^-(t')]\right\}|\mathbf{Q}_j\rangle, \quad (5.31)$$

where

$$L_H(t, \dot{\mathbf{Q}}_j^-(t) \equiv 2\pi\hbar\mathbf{k}\cdot\dot{\mathbf{Q}}_j^-(t) - H_H(t), \quad (5.32)$$

with

$$\dot{\mathbf{Q}}_j^-(t_j) \equiv (\mathbf{Q}_j - \mathbf{Q}_{j+1})/(t_{j+1} - t_j) = -\dot{\mathbf{Q}}_j(t_j). \quad (5.33)$$

Evaluation of the matrix element

$$\langle \mathbf{Q}_j|\, U_H(t_{j+1}, t_j)\,|\mathbf{Q}_{j+1}\rangle$$

[9] Reference 5, Eqs. (20) and (28).

by means of Weyl transforms gives, to first order in $t_{j+1} - t_j$,

$$\langle \mathbf{Q}_j | U_H(t_{j+1}, t_j) | \mathbf{Q}_{j+1} \rangle = \int d\mathbf{K} \exp\left\{-(i/\hbar)\int_{t_j}^{t_{j+1}} dt' H_H[t', \tfrac{1}{2}(\mathbf{Q}_j + \mathbf{Q}_{j+1}), \mathbf{K}]\right\} \times \exp[2\pi i \mathbf{K} \cdot (\mathbf{Q}_j - \mathbf{Q}_{j+1})]. \quad (5.34)$$

The result of integration on $\mathbf{K}$ is a function of $\frac{1}{2}(\mathbf{Q}_j + \mathbf{Q}_{j+1})$ and $\dot{\mathbf{Q}}_j^-(t_j)$. Feynman's postulate (5.26) may be modified to require that this function take the form

$$\langle \mathbf{Q}_j | U_H(t_{j+1}, t_j) | \mathbf{Q}_{j+1} \rangle = \alpha^{-1} \exp\left\{(i/\hbar)\int_{t_j}^{t_{j+1}} dt' L_H[t', \dot{\mathbf{Q}}_j^-(t'), \tfrac{1}{2}(\mathbf{Q}_j + \mathbf{Q}_{j+1})]\right\}. \quad (5.35)$$

If the Hamiltonian is time independent, then $H_H = H_s = H$, and (5.25) gives

$$\langle \mathbf{Q}_{j+1} | U_s(t_{j+1}, t_j) | \mathbf{Q}_j \rangle = \int d\mathbf{K} \exp\{-(i/\hbar)(t_{j+1} - t_j) H[\tfrac{1}{2}(\mathbf{Q}_j + \mathbf{Q}_{j+1}), \mathbf{K}] + 2\pi i \mathbf{K} \cdot (\mathbf{Q}_{j+1} - \mathbf{Q}_j)\},$$

while (5.34) gives

$$\langle \mathbf{Q}_j | U_H(t_{j+1}, t_j) | \mathbf{Q}_{j+1} \rangle = \int d\mathbf{K} \exp\{-(i/\hbar)(t_{j+1} - t_j) H[\tfrac{1}{2}(\mathbf{Q}_j + \mathbf{Q}_{j+1}), -\mathbf{K}] + 2\pi i \mathbf{K} \cdot (\mathbf{Q}_{j+1} - \mathbf{Q}_j)\}.$$

When $H(\mathbf{Q}, \mathbf{K})$ is an even function of $\mathbf{K}$ these two matrix elements are equal, and

$$L_H[\dot{\mathbf{Q}}_j^-(t), \tfrac{1}{2}(\mathbf{Q}_j + \mathbf{Q}_{j+1})] = L_s[\dot{\mathbf{Q}}_j(t), \tfrac{1}{2}(\mathbf{Q}_j + \mathbf{Q}_{j+1})].$$

This is the case in the example of (5.27) for which, with $\Delta t = t_{j+1} - t_j$,

$$\langle \mathbf{Q}_j | U_H(t_{j+1}, t_j) | \mathbf{Q}_{j+1} \rangle = (m/i2\pi\hbar\Delta t)^{\frac{3}{2}} \times \exp\{(i/\hbar) L_H[\dot{\mathbf{Q}}_j^-(t_j), \tfrac{1}{2}(\mathbf{Q}_j + \mathbf{Q}_{j+1})]\Delta t\}, \quad (5.36)$$

where

$$L_H[\dot{\mathbf{Q}}_j^-(t_j), \tfrac{1}{2}(\mathbf{Q}_j + \mathbf{Q}_{j+1})] = (m/2)(\dot{\mathbf{Q}}_j^-)^2(t_j) - V_s[\tfrac{1}{2}(\mathbf{Q}_j + \mathbf{Q}_{j+1})], \quad (5.37)$$

equal to $L_s[\mathbf{Q}_j(t_j), \tfrac{1}{2}(\mathbf{Q}_j + \mathbf{Q}_{j+1})]$ in (5.29).

Return now to the expression for $\mathcal{A}(t, \mathbf{Q})$ in (5.16), the case in which $A_s(t)$ commutes with $\mathbf{k}$. From (5.30) and (5.34),

$$A_s[t, (1/2\pi i)(\partial/\partial \mathbf{Q}_0)] \langle \mathbf{Q}_0 | U(t, t_0) | \mathbf{Q} \rangle = \lim_{n\to\infty} \int \cdots \int d\mathbf{Q}_1 \cdots d\mathbf{Q}_{n-1} \prod_{j=1}^{n-1} \langle \mathbf{Q}_j | U_H(t_{j+1}, t_j) | \mathbf{Q}_{j+1} \rangle \times A_s[t, (1/2\pi i)(\partial/\partial \mathbf{Q}_0)] \int d\mathbf{K} \times \exp\left\{-(i/\hbar)\int_{t_0}^{t_1} dt' H_H[t', \tfrac{1}{2}(\mathbf{Q}_0 + \mathbf{Q}_1), \mathbf{K}] + 2\pi i \mathbf{K} \cdot (\mathbf{Q}_0 - \mathbf{Q}_1)\right\} = \lim_{n\to\infty} \int \cdots \int d\mathbf{Q}_1 \cdots d\mathbf{Q}_{n-1} \times \prod_{j=1}^{n-1} \langle \mathbf{Q}_j | U_H(t_{j+1}, t_j) | \mathbf{Q}_{j+1} \rangle \times A_s\{t, [2\pi i(t_1 - t_0)]^{-1} \partial/\partial \dot{\mathbf{Q}}_0^-(t_0)\} \times \int d\mathbf{K} \exp\left[-(i/\hbar)\int_{t_0}^{t_1} dt' \{H_H[t', \tfrac{1}{2}(\mathbf{Q}_0 + \mathbf{Q}_1), \mathbf{K}] - 2\pi\hbar \mathbf{K} \cdot \dot{\mathbf{Q}}_0^-(t')\}\right].$$

According to the modified Feynman postulate (5.35), the last $\mathbf{K}$ integration can be expressed in terms of the Lagrangian $L_H$, so that

$$A_s[t, (1/2\pi i)(\partial/\partial \mathbf{Q}_0)] \langle \mathbf{Q}_0 | U(t, t_0) | \mathbf{Q} \rangle = \lim_{n\to\infty} \int \cdots \int d\mathbf{Q}_1 \cdots d\mathbf{Q}_{n-1} \prod_{j=1}^{n-1} \langle \mathbf{Q}_j | U_H(t_{j+1}, t_j) | \mathbf{Q}_{j+1} \rangle \times A_s\{t, (1/2\pi\hbar)\partial L_H[t_0, \dot{\mathbf{Q}}_0^-(t_0), \mathbf{Q}_0]/\partial \dot{\mathbf{Q}}_0^-(t_0)\} \times \langle \mathbf{Q}_0 | U_H(t_1, t_0) | \mathbf{Q}_1 \rangle.$$

$\mathcal{A}(t, \mathbf{Q})$ in (5.16) therefore becomes

$$\mathcal{A}(t, \mathbf{Q}) = \int d\mathbf{Q}_0 P(t, \mathbf{Q}_0(t_0, \mathbf{Q}) \times A_s\{t, (1/2\pi\hbar)\partial L_H(t_0, \dot{\mathbf{Q}}_0^-(t_0), \mathbf{Q}_0]/\partial \dot{\mathbf{Q}}_0^-(t_0)\}, \quad (5.38)$$

with

$$P(t, \mathbf{Q}_0 \mid t_0, \mathbf{Q}) = |\langle \mathbf{Q}_0 | U(t, t_0) | \mathbf{Q} \rangle|^2.$$

According to (5.15), if $A_s(t)$ commutes with $\mathbf{q}$, then

$$\mathcal{A}(t, \mathbf{Q}) = \int d\mathbf{Q}_0 P(t, \mathbf{Q}_0 \mid t_0, \mathbf{Q}) A_s(t, \mathbf{Q}_0). \quad (5.39)$$

In the example of (5.27) for which the Lagrangian $L_H$ is given in (5.37),

$$\mathcal{A}(t, \mathbf{Q}) = \int d\mathbf{Q}_0 P(t, \mathbf{Q}_0 \mid t_0, \mathbf{Q}) \times A_s[t, m\dot{\mathbf{Q}}_0^-(t_0)/2\pi\hbar], \quad (5.40)$$

according to (5.38) when $A_s(t)$ commutes with $\mathbf{k}$. For the momentum operator $\mathbf{p} = 2\pi\hbar\mathbf{k}$, $\mathcal{A}(t, \mathbf{Q})$ becomes $\mathcal{F}(t, \mathbf{Q})$, where

$$\mathcal{F}(t, \mathbf{Q}) \equiv \int d\mathbf{K} 2\pi\hbar \mathbf{K}_H(t, \mathbf{Q}, \mathbf{K}) = \int d\mathbf{Q}_0 P(t, \mathbf{Q}_0 \mid t_0, \mathbf{Q}) m\dot{\mathbf{Q}}_0^-. \quad (5.41)$$

The first Hamiltonian equation (3.50) gives, in this example,

$$\partial Q_H(t, \mathbf{Q}, \mathbf{K})/\partial t = \int\cdots\int d\mathbf{Q}_0\, d\mathbf{K}_0 \times P(t, \mathbf{Q}_0, \mathbf{K}_0 \mid t_0, \mathbf{Q}, \mathbf{K}) 2\pi\hbar\mathbf{K}_0/m.$$

Therefore, with the definition

$$\mathcal{Q}(t, \mathbf{Q}) \equiv \int d\mathbf{K} Q_H(t, \mathbf{Q}, \mathbf{K}), \tag{5.42}$$

$$\begin{aligned}\partial\mathcal{Q}(t, \mathbf{Q})/\partial t &= \int\cdots\int d\mathbf{Q}_0\, d\mathbf{K}_0 (2\pi\hbar\mathbf{K}_0/m) \times \int d\mathbf{K} P(t, \mathbf{Q}_0, \mathbf{K}_0 \mid t_0, \mathbf{Q}, \mathbf{K}) \\ &= \int d\mathbf{Q}_0 P(t, \mathbf{Q}_0 \mid t_0, \mathbf{Q}) \dot{\mathbf{Q}}_0^-,\end{aligned} \tag{5.43}$$

according to (5.14), (5.16), and (5.40). According to (5.41),

$$\mathcal{F}(t, \mathbf{Q}) = m\partial\mathcal{Q}(t, \mathbf{Q})/\partial t. \tag{5.44}$$

The second Hamiltonian equation (3.50) gives, in this example,

$$\begin{aligned}\partial\mathcal{F}(t, \mathbf{Q})/\partial t &= -\int\cdots\int d\mathbf{Q}_0\, d\mathbf{K}_0 \times \int d\mathbf{K} P(t, \mathbf{Q}_0, \mathbf{K}_0 \mid t_0, \mathbf{Q}, \mathbf{K}) \partial V_s(\mathbf{Q}_0)/\partial\mathbf{Q}_0 \\ &= -\int d\mathbf{Q}_0 P(t, \mathbf{Q}_0 \mid t_0, \mathbf{Q}) \partial V_s(\mathbf{Q}_0)/\partial\mathbf{Q}_0.\end{aligned} \tag{5.45}$$

(5.44) and (5.45) give the quantum-mechanical form of Newton's second law,

$$\begin{aligned}&m\partial^2\mathcal{Q}(t, \mathbf{Q})/\partial t \\ &= -\int d\mathbf{Q}_0 P(t, \mathbf{Q}_0 \mid t_0, \mathbf{Q}) \partial V_s(\mathbf{Q}_0)/\partial\mathbf{Q}_0 \\ &= -\int d\mathbf{K} \partial V_H(t, \mathbf{Q}, \mathbf{K})/\partial\mathbf{Q} \equiv -\partial\mathcal{U}(t, \mathbf{Q})/\partial\mathbf{Q}.\end{aligned} \tag{5.46}$$

It is clear from this example that the velocity of the system along a path in configuration space must be identified with $\dot{\mathbf{Q}}^-(t)$ in (5.33), and not with $\dot{\mathbf{Q}}(t)$ in (5.24). Furthermore, the classical Lagrangian corresponds to the quantum operator $L_H[t, \dot{\mathbf{Q}}_j^-(t)]$ in (5.32), and not to the operator $L_s[t, \dot{\mathbf{Q}}_j(t)]$ of Feynman in (5.24).

## 6. OTHER COORDINATE SYSTEMS

The discussion in the previous section has been based on the case for which the density operator $\rho_H = \rho_s(t_0)$ commutes with the coordinate operator $\mathbf{q}$. The system is prepared initially in a state which is a mixture of pure coordinate states, as given in (5.3). $\mathcal{A}(t, \mathbf{Q})$ is the expectation at time $t$ of the dynamical property of the system initially in the pure state $\rho_s(t_0) = \delta(\mathbf{q} - \mathbf{Q})$. Generalization is readily made to other initial states. Suppose $\rho_H$ commutes with a complete commuting set[10] of observables, $\mathbf{j} \equiv \{j_1, j_2 \cdots j_N\}$ with common eigenvectors $|\mathbf{J}\rangle$ specified by the set of eigenvalues $\mathbf{J}$, so that

$$\Sigma_\mathbf{J} |\mathbf{J}\rangle\langle\mathbf{J}| \equiv \Sigma_\mathbf{J} \delta(\mathbf{j} - \mathbf{J}) = \mathbf{1}. \tag{6.1}$$

Then

$$\rho_H = \Sigma_\mathbf{J} |\mathbf{J}\rangle\, \rho_H(\mathbf{J})\, \langle\mathbf{J}|, \quad \rho_H(\mathbf{J}) = \mathrm{Tr}\, [\rho_H \delta(\mathbf{j} - \mathbf{J})]. \tag{6.2}$$

The expectation at time $t$ of an operator $A(t)$ is given by

$$\langle A(t)\rangle = \Sigma_\mathbf{J} \mathcal{A}(t, \mathbf{J}) \rho_H(\mathbf{J}), \tag{6.3}$$

where

$$\begin{aligned}\mathcal{A}(t, \mathbf{J}) &= \mathrm{Tr}\, [A_H(t)\delta(\mathbf{j} - \mathbf{J})] \\ &= \langle\mathbf{J}|\, U^\dagger(t, t_0) A_s(t) U(t, t_0)\, |\mathbf{J}\rangle.\end{aligned} \tag{6.4}$$

If $A_s(t)$ also commutes with the set of $\mathbf{j}$, then

$$\begin{aligned}A_s(t) &= \Sigma_\mathbf{J} |\mathbf{J}\rangle\, A_s(t, \mathbf{J})\, \langle\mathbf{J}|, \\ A_s(t, \mathbf{J}) &= \mathrm{Tr}\, [A_s(t)\delta(\mathbf{j} - \mathbf{J})],\end{aligned} \tag{6.5}$$

so that

$$\mathcal{A}(t, \mathbf{J}) = \Sigma_{\mathbf{J}_0} A_s(t, \mathbf{J}_0) P(t, \mathbf{J}_0 \mid t_0, \mathbf{J}), \tag{6.6}$$

where

$$P(t, \mathbf{J}_0 \mid t_0, \mathbf{J}) = |\langle\mathbf{J}_0|\, U(t, t_0)\, |\mathbf{J}\rangle|^2, \tag{6.7}$$

the probability of $\mathbf{J}_0$ at $t$ conditional on $\mathbf{J}$ at $t_0$. The probability amplitude can be written as a Feynmann path integral,

$$\begin{aligned}&\langle\mathbf{J}_0|\, U(t, t_0)\, |\mathbf{J}\rangle \\ &= \lim_{n\to\infty} \Sigma_{\mathbf{J}_1} \cdots \Sigma_{\mathbf{J}_{n-1}} \prod_{i=0}^{n-1} \langle\mathbf{J}_i|\, U_H(t_{i+1}, t_i)\, |\mathbf{J}_{i+1}\rangle.\end{aligned} \tag{6.8}$$

For example, if $\rho_H$ commutes with the momentum operator $\mathbf{k}$, then

$$\begin{aligned}\mathcal{A}(t, \mathbf{K}) \equiv \mathrm{Tr}\, [A_H(t)\delta(\mathbf{k} - \mathbf{K})] &= \int d\mathbf{Q} A_H(t, \mathbf{Q}, \mathbf{K}) \\ &= \langle\mathbf{K}|\, U^\dagger(t, t_0) A_s(t) U(t, t_0)\, |\mathbf{K}\rangle,\end{aligned} \tag{6.9}$$

corresponding to (6.4) and (5.14). In momentum representation, (5.15) and (5.16) become

$$\mathcal{A}(t, \mathbf{K}) = \int d\mathbf{K}_0\, \langle\mathbf{K}|\, U^\dagger(t, t_0)\, |\mathbf{K}_0\rangle\, A_s[t, (i/2\pi)(\partial/\partial\mathbf{K}_0)] \times \langle\mathbf{K}_0|\, U(t, t_0)\, |\mathbf{K}\rangle, \tag{6.10}$$

if $A_s(t)$ commutes with $\mathbf{q}$, and

$$\mathcal{A}(t, \mathbf{K}) = \int d\mathbf{K}_0 A_s(t, \mathbf{K}_0)\, |\langle\mathbf{K}_0|\, U(t, t_0)\, |\mathbf{K}\rangle|^2, \tag{6.11}$$

[10] Reference (4), p. 57.

if $A_s(t)$ commutes with $\mathbf{k}$. The probability amplitude $\langle \mathbf{K}_0|\, U(t, t_0)\, |\mathbf{K}\rangle$ can be written as the Feynman path integral,

$$\langle \mathbf{K}_0|\, U(t, t_0)\, |\mathbf{K}\rangle = \lim_{n\to\infty} \int \cdots \int d\mathbf{K}_1 \cdots d\mathbf{K}_{n-1} \times \prod_{j=0}^{n-1} \langle \mathbf{K}_j|\, U_H(t_{j+1}, t_j)\, |\mathbf{K}_{j+1}\rangle \quad (6.12)$$

$$= \lim_{n\to\infty} \int \cdots \int d\mathbf{K}_1 \cdots d\mathbf{K}_{n-1} \prod_{j=0}^{n-1} \langle \mathbf{K}_j|\, \exp\left\{-(i/\hbar) \times \int_{t_j}^{t_{j+1}} dt' [H_H(t') + \mathbf{q}\cdot 2\pi\hbar \dot{\bar{\mathbf{K}}}_j(t')]\right\} |\mathbf{K}_j\rangle, \quad (6.13)$$

with

$$\dot{\bar{\mathbf{K}}}_j(t_j) \equiv (\mathbf{K}_j - \mathbf{K}_{j+1})/(t_{j+1} - t_j). \quad (6.14)$$

Evaluation with the use of Weyl transforms gives, to first order in $t_{j+1} - t_j$,

$$\langle \mathbf{K}_j|\, U_H(t_{j+1}, t_j)\, |\mathbf{K}_{j+1}\rangle = \int d\mathbf{Q} \exp\left[-(i/\hbar)\int_{t_j}^{t_{j+1}} dt' \left\{H_H[t', \mathbf{Q}, \tfrac{1}{2}(\mathbf{K}_j + \mathbf{K}_{j+1})] + \mathbf{Q}\cdot 2\pi\hbar\dot{\bar{\mathbf{K}}}_j(t')\right\}\right], \quad (6.15)$$

a function of $\frac{1}{2}(\mathbf{K}_j + \mathbf{K}_{j+1})$ and $\dot{\bar{\mathbf{K}}}_j(t_j)$.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 5 MAY 1968

# New Field-Theory Approach to Singular Potentials

H. H. ALY
*Department of Physics, American University of Beirut, Beirut, Lebanon*
AND
J. G. TAYLOR*
*Mathematical Institute, Oxford, University of Oxford, England*

(Received 9 June 1967)

An approach similar to the methods of renormalization of the Green's functions equations of quantum field theory is adopted to the singular potential in the Lippman–Schwinger equation. The close relation between our approach and the one used in field theory gives a method to be applied to nonrenormalizable field theories. The physical implication of this approach is discussed.

## 1. INTRODUCTION

One of the outstanding problems in high-energy physics is to give a consistent theory of nonrenormalizable interactions (e.g., weak interactions, spin-1 electrodynamics, etc.) beyond the lowest order of perturbation theory.

Various attempts to solve this problem[1] have been made without complete success. However, due to the difficulties inherent in this problem, attention has been focused on the study of the quantum theory of singular potentials, although results for such singular potentials can at most hint at the actual properties of nonrenormalizable interactions in quantum field theory. Of course, singular potentials have an intrinsic interest and do indeed raise many interesting questions.

In this paper we wish to describe a new approach to singular potentials which is intimately related to the methods of renormalization of the Green's functions equations of quantum field theory. This intimate relation allows us to hope that properties similar to those we find for singular potentials will also occur in nonrenormalizable field theories.

For several years the formulation of renormalized Green's functions equations for renormalizable interactions has been understood.[2] The crucial steps for renormalization of masses and coupling constants is achieved by a differentiation and integration procedure on the external momenta entering the propagators and vertex functions. We wish to set up equations for nonrelativistic potential scattering to which such a renormalization procedure may be applied successfully. In the field-theoretic case we deal only with scattering amplitudes (connected parts of time-order products), so that we expect it to be necessary to do the same for singular potentials.

We restrict our discussion to two particles scattering through real velocity-independent potentials. The

* Department of Physics, Queen Mary College, London, England.

[1] In the case of quantum field theory see, for example, J. G. Taylor, Suppl., Nuovo Cimento **1**, 857 (1963). In the study of potential theory see A. Bastai, L. Bertocchi, G. Furlan, S. Fubini, and M. Tonin, Nuovo Cimento **30**, 1512 (1963); W. Güttinger and E. Pfaffelhuber, "Generalized Lippman–Schwinger Scattering Equations for Singular Interactions," CERN Report 65/1211/5, Th 586; and W. Güttinger, R. Penyl, and E. Pfaffelhuber, Ann. Physik **33**, 246 (1965).

[2] J. G. Taylor, Suppl. Nuovo Cimento **1**, 857 (1963), papers I, II, and III.

if $A_s(t)$ commutes with $\mathbf{k}$. The probability amplitude $\langle \mathbf{K}_0|\, U(t, t_0)\, |\mathbf{K}\rangle$ can be written as the Feynman path integral,

$$\langle \mathbf{K}_0|\, U(t, t_0)\, |\mathbf{K}\rangle = \lim_{n\to\infty} \int \cdots \int d\mathbf{K}_1 \cdots d\mathbf{K}_{n-1} \times \prod_{j=0}^{n-1} \langle \mathbf{K}_j|\, U_H(t_{j+1}, t_j)\, |\mathbf{K}_{j+1}\rangle \quad (6.12)$$

$$= \lim_{n\to\infty} \int \cdots \int d\mathbf{K}_1 \cdots d\mathbf{K}_{n-1} \prod_{j=0}^{n-1} \langle \mathbf{K}_j|\, \exp\left\{-(i/\hbar) \times \int_{t_j}^{t_{j+1}} dt' [H_H(t') + \mathbf{q}\cdot 2\pi\hbar \dot{\bar{\mathbf{K}}}_j(t')]\right\} |\mathbf{K}_j\rangle, \quad (6.13)$$

with

$$\dot{\bar{\mathbf{K}}}_j(t_j) \equiv (\mathbf{K}_j - \mathbf{K}_{j+1})/(t_{j+1} - t_j). \quad (6.14)$$

Evaluation with the use of Weyl transforms gives, to first order in $t_{j+1} - t_j$,

$$\langle \mathbf{K}_j|\, U_H(t_{j+1}, t_j)\, |\mathbf{K}_{j+1}\rangle = \int d\mathbf{Q} \exp\left[-(i/\hbar)\int_{t_j}^{t_{j+1}} dt' \left\{H_H[t', \mathbf{Q}, \tfrac{1}{2}(\mathbf{K}_j + \mathbf{K}_{j+1})] + \mathbf{Q}\cdot 2\pi\hbar\dot{\bar{\mathbf{K}}}_j(t')\right\}\right], \quad (6.15)$$

a function of $\frac{1}{2}(\mathbf{K}_j + \mathbf{K}_{j+1})$ and $\dot{\bar{\mathbf{K}}}_j(t_j)$.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 5 MAY 1968

# New Field-Theory Approach to Singular Potentials

H. H. ALY
*Department of Physics, American University of Beirut, Beirut, Lebanon*
AND
J. G. TAYLOR*
*Mathematical Institute, Oxford, University of Oxford, England*

(Received 9 June 1967)

An approach similar to the methods of renormalization of the Green's functions equations of quantum field theory is adopted to the singular potential in the Lippman–Schwinger equation. The close relation between our approach and the one used in field theory gives a method to be applied to nonrenormalizable field theories. The physical implication of this approach is discussed.

## 1. INTRODUCTION

One of the outstanding problems in high-energy physics is to give a consistent theory of nonrenormalizable interactions (e.g., weak interactions, spin-1 electrodynamics, etc.) beyond the lowest order of perturbation theory.

Various attempts to solve this problem[1] have been made without complete success. However, due to the difficulties inherent in this problem, attention has been focused on the study of the quantum theory of singular potentials, although results for such singular potentials can at most hint at the actual properties of nonrenormalizable interactions in quantum field theory. Of course, singular potentials have an intrinsic interest and do indeed raise many interesting questions.

In this paper we wish to describe a new approach to singular potentials which is intimately related to the methods of renormalization of the Green's functions equations of quantum field theory. This intimate relation allows us to hope that properties similar to those we find for singular potentials will also occur in nonrenormalizable field theories.

For several years the formulation of renormalized Green's functions equations for renormalizable interactions has been understood.[2] The crucial steps for renormalization of masses and coupling constants is achieved by a differentiation and integration procedure on the external momenta entering the propagators and vertex functions. We wish to set up equations for nonrelativistic potential scattering to which such a renormalization procedure may be applied successfully. In the field-theoretic case we deal only with scattering amplitudes (connected parts of time-order products), so that we expect it to be necessary to do the same for singular potentials.

We restrict our discussion to two particles scattering through real velocity-independent potentials. The

* Department of Physics, Queen Mary College, London, England.

[1] In the case of quantum field theory see, for example, J. G. Taylor, Suppl., Nuovo Cimento **1**, 857 (1963). In the study of potential theory see A. Bastai, L. Bertocchi, G. Furlan, S. Fubini, and M. Tonin, Nuovo Cimento **30**, 1512 (1963); W. Güttinger and E. Pfaffelhuber, "Generalized Lippman–Schwinger Scattering Equations for Singular Interactions," CERN Report 65/1211/5, Th 586; and W. Güttinger, R. Penyl, and E. Pfaffelhuber, Ann. Physik **33**, 246 (1965).

[2] J. G. Taylor, Suppl. Nuovo Cimento **1**, 857 (1963), papers I, II, and III.

situation is described in field theory by the Bethe–Salpeter equation (BS).

The potential-theory analog of the BS equation is the Lippman–Schwinger equation (LS). Our approach consists in discussing the renormalization of the LS equation in a manner as similar as possible to that used for the BS equation.[3a]

In Sec. 2 we first discuss the renormalization of the potential analog of the $\lambda\phi^4$ theory (or other renormalizable interactions). We then show that this renormalization may be extended to more singular potentials, with the LS equation being generalized to a linear integrodifferential equation. The solutions to this equation depend on a finite number of additional parameters, which cannot be specified *a priori*. In Sec. 3 we discuss the partial-wave separation of our extended LS equation. In Sec. 4 we discuss and summarize the results and indicate the physical implication for nonrenormalizable field theories. We also raise a number of related questions which we have not been able to solve.

## 2. EXTENSION OF THE LS EQUATION

Let us first review briefly the relation between the BS equation and potential theory. We recognize three classes of field-theoretic interactions:

Class I: only finite quantities appear;

Class II: renormalizable interactions; only those divergences appear which are removable by absorption into masses and coupling constants;

Class III: nonrenormalizable interactions.

The BS equation is given by

$$M(p,p';q) = V(p,p';q) + \int d^4k V(p,k;q)M(k,p';q) \times [(k+q)^2 - m^2]^{-1}[(k-q)^2 - m^2]^{-1}, \quad (1)$$

where $q^2$ is the invariant total energy, $q \pm p$, $q \pm p'$ are the initial and final momenta of the particles, and $V(p,p';q)$ is the relativistic potential.

Class I corresponds to potentials which tend to zero for large momentum transfer and external masses. For Class II (for spinless particles) the potential is constant. In Class III, the potential may increase either (a) as a power or (b) at least exponentially in momentum transfer or external masses. We remark here that the BS equation is a well-defined integral equation under iteration for potentials of Class I. This is not the case for potentials of Class II or III.

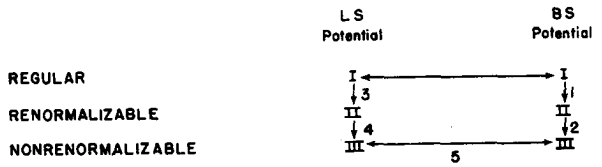



FIG. 1. Diagrammatic comparison of the three classes of LS and BS potentials.

We use the symmetric differential operator

$$\mathfrak{D} = \sum_{i,\mu=1}^{4} p_{i\mu}\partial/\partial p_{i\mu},$$

where $p_{i\mu}$ are the components of the external momenta $p_1, p_2, p_3, p_4 = q \pm p, q \pm p'$. We apply $\mathfrak{D}^{-1}\mathfrak{D}$ to both sides of Eq. (1) for a BS potential $V(p,p';q)$ arising from a renormalizable field theory (of Class II). The resulting equation is now a differentio-integral equation containing the ambiguity arising from $\mathfrak{D}^{-1}$, i.e., the value of the scattering amplitude at zero momenta for all particles. This equation is now well-behaved under iteration for BS potentials of Class II, as has been carefully discussed in the work of one of us.[3b] The value of the scattering amplitude at zero momenta cannot be determined from the differentio-integral equation, but plays the role of the renormalized charge.

It may be possible to discuss BS potentials of Class IIIa by means of $\mathfrak{D}^n$ applied to both sides of Eq. (1) and the integration by means of $\mathfrak{D}^{-n}$ (where $n-1$ is the power dependence of the "potentials" on momentum transfer and external masses). However, we may remark that it is not known if such procedure can, in fact, be done completely, nor is it known if such a discussion can be extended to the complete set of Green's functions equations in this case. We wish to see if we can carry through such a procedure for the LS equation.

The LS analog to the BS potentials again may be divided into three classes:

Class I: regular potentials behaving as does $r^{-\alpha}$ ($\alpha < 2$) when $r \to 0$;

Class II: $V(r) = r^{-\alpha}$, ($\alpha = 2$), as $r \to 0$;

Class III: (a) $V(r) = r^{-\alpha}(\alpha > 2)$ as $r \to 0$; (b) $V(r)$ has a singularity stronger than a pure power at $r = 0$.

Previous work leads us to compare Classes II and III of the LS potentials with Classes II and III of BS potentials.

In Fig. 1 we present a diagrammatic form of the situation. In the first column we describe the nature of the class. In the second column the LS equation for each class is indicated. In the third the BS equation for each class is indicated.

[3a] Reference 2, papers II and V.

[3b] Reference 2, papers I, II, and V.

We may relate corresponding classes horizontally by the standard method. We have made step (1), extending the BS equation from Class I to II by means of $\mathfrak{D}^{-1}\mathfrak{D}$. We conjectured that we can make step (2), extending it further to Class III by use of $\mathfrak{D}^{-n}\mathfrak{D}^n$ for $n > 1$. We wish to show here that it is actually possible to make both these extensions, steps (3) and (4), for the LS equation.

The LS equation is

$$f(\mathbf{p}, \mathbf{p}', \mathbf{k}^2) = \tilde{V}(|\mathbf{p} - \mathbf{p}'|) + \int \tilde{V}(|\mathbf{p} - \mathbf{q}|) \frac{f(\mathbf{q}, \mathbf{p}', \mathbf{k}^2)}{(\mathbf{k}^2 - \mathbf{q}^2)} d^3\mathbf{q}, \quad (2)$$

where

$$\tilde{V}(|\mathbf{p}|) = \int e^{i\mathbf{p}\cdot\mathbf{r}} V(|\mathbf{r}|)\, d^3\mathbf{r}. \quad (3)$$

We may also use the relativistic energy $(\mathbf{k}^2 + m^2)^{\frac{1}{2}} - m$ in place of $\mathbf{k}^2/2m$ in the LS equation (2) if, at the same time, we replace $d^3\mathbf{q}$ by the invariant measure $d^3\mathbf{q}/(\mathbf{q}^2 \pm m^2)^{\frac{1}{2}}$. This gives the same high-energy behavior as (2), so our further discussion is the same for either case. We will explicitly keep to (2).

The asymptotic behavior of $\tilde{V}(|\mathbf{p}|)$ as $|\mathbf{p}| \to \infty$ is

$$\tilde{V}(|\mathbf{p}|) \sim |\mathbf{p}|^{\alpha-3}, \quad (4)$$

where all logarithmic dependencies are neglected. We insert the behavior of Eq. (4) in Eq. (2), where we see that the equation is well-defined under iteration for $\alpha < 2$. When $\alpha = 2$, we find logarithmic divergence on iteration. This may be removed exactly, as was done for the BS equation, replacing the operator $\mathfrak{D}$ by the operator $\mathbf{d}$ defined as acting on any function of the three-vector $\mathbf{p}$ by

$$(\mathbf{d}_\mathbf{p} f)(\mathbf{p}) = [(d/d\lambda) f(\lambda\hat{p})]_{\lambda=|\mathbf{p}|},$$

where $\hat{p} = \mathbf{p}/|\mathbf{p}|$ is the unit vector along $\mathbf{p}$. Similarly, $(d^n_\mathbf{p} f)(\mathbf{p}) = [(d^n/d\lambda^n) f(\lambda\hat{p})]_{\lambda=|\mathbf{p}|}$. Then Eq. (2) becomes

$$f(\mathbf{p}, \mathbf{p}', \mathbf{k}^2) = \tilde{V}(|\mathbf{p} - \mathbf{p}'|) - \tilde{V}(|\mathbf{p}'|) + f(0, \mathbf{p}', \mathbf{k}^2) + \int_0^{|\mathbf{p}|} d\lambda \left[ \frac{d_{\mathbf{p}''} \tilde{V}(|\mathbf{p}'' - \mathbf{q}|) f(\mathbf{q}, \mathbf{p}', \mathbf{k}^2)}{(\mathbf{k}^2 - \mathbf{q}^2)} \right]_{\mathbf{p}''=\lambda\mathbf{p}} d^3\mathbf{q}. \quad (5)$$

We define

$$U(\mathbf{p}, \mathbf{p}') = \tilde{V}(|\mathbf{p} - \mathbf{p}'|) - \tilde{V}(|\mathbf{p}'|) + f(0, \mathbf{p}'). \quad (6)$$

Since $d_\mathbf{p} U(\mathbf{p}, \mathbf{p}') = d_\mathbf{p} \tilde{V}(|\mathbf{p} - \mathbf{q}|)$, we may rewrite Eq. (5) as

$$f(\mathbf{p}, \mathbf{p}', \mathbf{k}^2) = U(\mathbf{p}, \mathbf{p}') + \int_0^{|\mathbf{p}|} d\lambda \frac{d_{\mathbf{p}''} U(\mathbf{p}'', \mathbf{q}) f(\mathbf{q}, \mathbf{p}', \mathbf{k}^2)}{(\mathbf{k}^2 - \mathbf{q}^2)} d^3\mathbf{q}. \quad (7)$$

We may regard the replacement of $\tilde{V}$ [in Eq. (5)] by $U$ [in Eq. (7)] as a renormalization of the LS potential in a manner analogous to the charge renormalization for $\pi$–$\pi$ scattering discussed previously.[3]

It is evident that iteration of Eq. (5) for Class II potentials (e.g., $r^{-2}$) is convergent at every step. This is due to the differentiation acting on $\tilde{V}(|\mathbf{p}'' - \mathbf{q}|)$ under the integral sign of Eq. (7) together with integration over a finite range of the $\lambda$ variable. Thus we regard Eq. (7) as the correct equation which extends the LS equation to the $r^{-2}$-type potentials in a manner as similar as possible to the renormalized BS equation for $\lambda\phi^4$ theory.

We would like to remark that we now have an integrodifferential equation in place of the integral equation structure for the LS equation. This means that we no longer expect many properties of the LS equation to remain valid for our extended equation. We shall return to this point later.

It is now possible to see how to generalize our extended LS equation to apply to nonrelativistic potentials of Class IIIa (e.g., $r^{-\alpha}$, $\alpha > 2$). This may be achieved by the use of the integrodifferential operator $d^{-(M)}d^{+(M)}$, where $M$ (as in the relativistic case) is equal to 1 plus the singular power of the potential and for any integrable function of $p$:

$$(d^{-1}f)(\mathbf{p}) = \int_0^{|\mathbf{p}|} d\lambda f(\lambda\hat{p}).$$

We see now that the form of Eq. (2) becomes

$$f(\mathbf{p}, \mathbf{p}^1, \mathbf{k}^2) = V'(\mathbf{p}, \mathbf{p}^1, \mathbf{k}^2) + \int_0^{|\mathbf{p}|} d\lambda_1 \int_0^{\lambda_1} d\lambda_2 \cdots \int_0^{\lambda_{M-1}} d\lambda_M \times \left\{ \frac{d^M_{\mathbf{p}''} \tilde{V}(|\mathbf{p}'' - \mathbf{q}|) f(\mathbf{q}, \mathbf{p}^1, \mathbf{k}^2)}{(\mathbf{k}^2 - \mathbf{q}^2)} \right\}_{\mathbf{p}''=\lambda_M\mathbf{p}} d^3\mathbf{q}, \quad (8)$$

where

$$V'(\mathbf{p}, \mathbf{p}^1, \mathbf{k}^2) = \tilde{V}(|\mathbf{p} - \mathbf{p}'|) + \sum_{r<n} (d^r f)_{\mathbf{p}=0},$$

where only the derivatives of $f$ in $(d^r f)_{\mathbf{p}=0}$ are evaluated at $\mathbf{p} = \mathbf{0}$.

We immediately see that Eq. (8) is not finite under iteration for the corresponding Class IIIa potentials. We may remedy this as follows. We take the expansion

$$\tilde{V}(|\mathbf{p}|) = \sum_{\eta=0}^{N} a_\eta |\mathbf{p}|^\eta + W(|\mathbf{p}|), \quad (9)$$

where $W(\chi) = O(\chi^{-1})$ as $|\chi| \to \infty$. An expression of the form of Eq. (9) always exists for Class IIIa potentials when we restrict $(\alpha - 3)$ to being a nonnegative integer $N$. If, in addition, $a_n = 0$ for $n$ odd, then $d^{N+1}_\mathbf{p} \tilde{V}(|\mathbf{p} - \mathbf{q}|) = \mathbf{d}^{N+1}_\mathbf{p} W(|\mathbf{p} - \mathbf{q}|)$.[4] Thus, in

[4] This restriction is not as severe as it sounds, since, if $V(r) = e^{-\mu r} r^{-\eta}$, we have formally that $\tilde{V}(|\mathbf{p}|) = 4\pi \int_0^\infty \exp(-\mu r) r^{1-\eta} \sin(|p|r)/|r|$ is an even function of $|\mathbf{p}|$. It is natural to preserve this evenness, even under the ambiguity of the definition of $r^{-\eta}$ at $r = 0$.

this case, Eq. (8) becomes

$$f(\mathbf{p}, \mathbf{p}', \mathbf{k}^2) = V'(\mathbf{p}, \mathbf{p}', \mathbf{k}^2) + \int_0^{|\mathbf{p}|} d\lambda_1 \int_0^{\lambda_1} d\lambda_2 \cdots \int_0^{\lambda_N} d\lambda_{N+1} \times \left[ \frac{(d_{\mathbf{p}''})^{N+1} V'(|\mathbf{p}'' - \mathbf{q}|) f(\mathbf{q}, \mathbf{p}', \mathbf{k}^2)}{(\mathbf{k}^2 - \mathbf{q}^2)} \right] d^3\mathbf{q}. \quad (10)$$

Then iteration of (10) will be finite at each step, since

$$V'(\mathbf{p}, \mathbf{p}', k^2) \sim |\mathbf{p}|^N \quad \text{as} \quad |\mathbf{p}| \sim \infty.$$

If

$$f(\mathbf{q}, \mathbf{p}', \mathbf{k}^2) \sim |\mathbf{q}|^N \quad \text{as} \quad |\mathbf{q}| \sim \infty,$$

the last term on the right of (10) is finite for finite $|\mathbf{p}|$ and behaves as $|\mathbf{p}|^N$ as $|\mathbf{p}| \sim \infty$ (to within logarithmic behavior, which we neglect). Thus we see that we have a denumerable set of the extended LS equations, each one corresponding to a certain class of singular potentials, i.e., these behaving as $r^{3-N}$ as $r \to 0$, with the coefficients $a_n$ vanishing for odd $n$.

Each such extended LS equation has a number of free parameters, and we may insure unitarity by choosing these parameters as real.

Since we have not altered the factor $(\mathbf{k}^2 - \mathbf{q}^2)^{-1}$ entering in the kernel of Eq. (10) by our differentiation and integration procedure, we expect that the discussion of the existence, uniqueness and other general properties will apply exactly as for the original LS equation. We also expect that analyticity will be unchanged.

It is necessary to discuss briefly the ambiguity in the definition of the Fourier transformation of the $r^{-\alpha}$ potentials for $\alpha > 2$.

This ambiguity arises from the singularity character of $r^{-\alpha}$ at $r = 0$. We may define the singular potential $r^{-\alpha}$ by means of the pseudofunctions.[5]

The ambiguity in this definition is that of a linear sum of derivatives of $\delta(\mathbf{r})$.[6] These derivatives are of order $\alpha - 3$. This will give an added polynomial of degree $\alpha - 3$ in $\tilde{V}(|\mathbf{p}|)$. We see no natural way to reduce this ambiguity except to enforce the condition that $a_n = 0$ for odd $n$. Thus we add further arbitrary constants to the solution of our extended LS equation. So we see that there are two sources of arbitrary constants for our solution to Eq. (10) in addition to those arising in the potential. These are (1) the differentiation–integration procedure and (2) the Fourier transformation of the singular potentials. It has been shown[7] that the explicit solution of the radial Schrödinger equation with the potential $r^{-4}$ has four arbitrary constants.

We hope to discuss elsewhere the relation between the solution of the extended LS equation and the Schrödinger equation in this and other cases.

## 3. PARTIAL WAVE ANALYSIS

We now investigate the partial wave analysis of our extended LS equation. We do this in detail only for Eq. (5).

Our method evidently extends to the more general equation (10). (We shall proceed in manner very similar to that used in Ref. 3.) We use the expression

$$f(\mathbf{p}, \mathbf{p}', \mathbf{k}^2) = \sum_{l=0}^{\infty} (2l + 1) P_l(\cos \theta_{\mathbf{p}, \mathbf{p}'}) f_l(|\mathbf{p}|, |\mathbf{p}'|, \mathbf{k}^2),$$

with

$$V_l(|\mathbf{p}|, |\mathbf{p}'|) = \frac{1}{2} \int_{-1}^{+1} P_l(\cos \theta_{\mathbf{p}, \mathbf{p}'}) \tilde{V}(|\mathbf{p} - \mathbf{p}'|)\, d(\cos \theta_{\mathbf{p}, \mathbf{p}'})$$

and the recurrence relation between the Legendre polynomials to obtain

$$f_l(|\mathbf{p}|\,|\mathbf{p}'|, \mathbf{k}^2) = V_l(|\mathbf{p}|, |\mathbf{p}'|) + \int_0^{|\mathbf{p}|} d\lambda \frac{d}{d\lambda} \left[ \left( \frac{2l - 1}{2l + 1} \right) V_{l-1}(\lambda\,|\mathbf{p}|, |\mathbf{q}|) - \left( \frac{2l + 3}{2l + 1} \right) V_{l+1}(\lambda\,|\mathbf{p}|, |\mathbf{q}|) \right] \times \frac{f_l(|\mathbf{q}|, |\mathbf{p}'|, \mathbf{k}^2)}{(\mathbf{k}^2 - \mathbf{q}^2)}\, d^3q. \quad (11)$$

[We note the similarity between Eq. (11) and the corresponding equation (25) in Ref. 3, No. 5.] In Eq. (11) we see that the different partial waves are not coupled to each other, and thus the partial wave separation is sufficient to decouple separate partial waves.

However, partial wave projections of the potential which arise in the kernel of Eq. (11) are different than in the inhomogeneous term. Such is not the case for the BS equation (as was discussed in Ref. 3). We attribute this to the difference between the $\mathfrak{D}$ and $d$ operators. In particular, $\mathfrak{D}$ depends on the internal and final momenta, while $d$ does not.

## 4. SUMMARY AND CONCLUSIONS

We may summarize our results by saying that the extended LS equation, Eq. (10), is the formulation which should be used in the discussion of singular

[5] L. Schwartz, *Theorie des distributions* (Hermann & Cie., Paris, 1947).

[6] The general approach to the ambiguity in $r^{-\eta}$ ar $r = 0$ is to define a function (such as a distribution on the subspace of indefinitely differentiable functions of compact support in $r$ which are zero) together with all partial derivatives of order less than $(n - 2)$ at $r = 0$. The extension of such a distribution to all indefinitely differentiable functions of compact support involves additional terms which are derivatives of $\delta^3(r)$ with arbitrary coefficients, as mentioned in the text.

[7] H. H. Aly and H. J. W. Müller, J. Math. Phys. 7, 1 (1966).

potentials in order to come as close as possible to the renormalization process in renormalizable field theory. This completes steps (3) and (4) of Fig. 1. We use this to give some indications of the results obtained after making step (5). In other words, we can speculate about the general nature of the structure of non-renormalizable field theories. We see that these theories are expected to be of two classes with properties similar to Classes IIIa and IIIb of LS potential theory. The solution belonging to interactions of Class IIIa will, in general, depend on a finite number of additional parameters occurring in the potential. For those of Class IIIb we cannot really say anything, since we have not been able to write a suitable extended form for the LS equation which is convergent under iteration. It is possible that our restriction to an integro-differential equation which is finite under iteration is too restrictive. However, without this restriction it is difficult to say anything rigorous using the present mathematical tools.

We may surmise that, in the case of Class IIIa field theories, only a finite number of differentiations and integrations will be necessary to transform the complete set of Green's functions equations to a form convergent under iteration. We may also expect that the solutions to Class IIIb field theories will very likely depend, in general, on an infinite number of arbitrary constants (however the solutions are obtained).

We hope that weak interaction of the current–current form are of Class IIIa.

In this short paper we have left many questions unanswered—and even unasked. In particular, we would like to know the following.

(a) What is the nature of the set of bound states?
(b) What is the relation of solutions to our equation to those for the Schrödinger equation?
(c) We would like to know the properties of the linear operators which are the kernels of our extended form of the LS equation. In particular, are they completely continuous? Is the Fredholm alternative valid? And, finally, is it possible to give a general existence theory for the solutions?
(d) Is the Levinson theorem valid? It has been shown[8] that, at least for the two classes of singular potentials like $(\sinh \mu r)^{-2}$ or $r^{-4}$, the Levinson theorem is violated in the non-relativistic limit; it is possible that all singular potentials violate it.

### ACKNOWLEDGMENT

One of us (J. G. T.) would like to thank Professor A. B. Zahlan for his hospitality at the Physics Department of the American University of Beirut, where this work was done.

[8] H. H. Aly and S. Okubo, Z. Physik **200**, 20 (1967).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 5 MAY 1968

# Realization of Poincaré-Group Generators on a Light Cone

ASHER PERES*
*Department of Physics and Astronomy, University of Maryland, College Park, Maryland*

(Received 17 July 1967)

The ten generators of the proper inhomogeneous Lorentz group are explicitly constructed for spin-0 and spin-$\frac{1}{2}$ particles, in the case when wavefunctions are given on a light cone (rather than on an equal time hypersurface, as usual). The advantage of this formulation is that the generators **J** (spatial rotations) and **K** (Lorentz boosts) involve only elementary local operators. The Hamiltonian $H$ and momentum operators **P** also contain the inverse radial momentum $(\partial/\partial r)^{-1}$, but do not involve any square roots. Moreover, only two-component spinors are required for spin-$\frac{1}{2}$ particles.

## I. INTRODUCTION

The problem of covariance in quantum theory is rather different from its classical counterpart.[1] In classical theory, the dynamical variables have numerical values which are equal to the observable values of these variables, and covariance can easily be ascertained if the dynamical variables transform in a definite way (e.g., as tensor components) under the appropriate group.

On the other hand, the dynamical variables of quantum theory are linear operators acting on some Hilbert space. The state of a physical system, as seen

* Permanent address: Technion, Israel Institute of Technology, Haifa, Israel.

[1] For a full discussion of this point, the reader is referred to the lucid article by L. L. Foldy, Phys. Rev. **122**, 275 (1961). No review can do full justice to the clarity of that paper.

potentials in order to come as close as possible to the renormalization process in renormalizable field theory. This completes steps (3) and (4) of Fig. 1. We use this to give some indications of the results obtained after making step (5). In other words, we can speculate about the general nature of the structure of non-renormalizable field theories. We see that these theories are expected to be of two classes with properties similar to Classes IIIa and IIIb of LS potential theory. The solution belonging to interactions of Class IIIa will, in general, depend on a finite number of additional parameters occurring in the potential. For those of Class IIIb we cannot really say anything, since we have not been able to write a suitable extended form for the LS equation which is convergent under iteration. It is possible that our restriction to an integro-differential equation which is finite under iteration is too restrictive. However, without this restriction it is difficult to say anything rigorous using the present mathematical tools.

We may surmise that, in the case of Class IIIa field theories, only a finite number of differentiations and integrations will be necessary to transform the complete set of Green's functions equations to a form convergent under iteration. We may also expect that the solutions to Class IIIb field theories will very likely depend, in general, on an infinite number of arbitrary constants (however the solutions are obtained).

We hope that weak interaction of the current–current form are of Class IIIa.

In this short paper we have left many questions unanswered—and even unasked. In particular, we would like to know the following.

(a) What is the nature of the set of bound states?
(b) What is the relation of solutions to our equation to those for the Schrödinger equation?
(c) We would like to know the properties of the linear operators which are the kernels of our extended form of the LS equation. In particular, are they completely continuous? Is the Fredholm alternative valid? And, finally, is it possible to give a general existence theory for the solutions?
(d) Is the Levinson theorem valid? It has been shown[8] that, at least for the two classes of singular potentials like $(\sinh \mu r)^{-2}$ or $r^{-4}$, the Levinson theorem is violated in the non-relativistic limit; it is possible that all singular potentials violate it.

### ACKNOWLEDGMENT

One of us (J. G. T.) would like to thank Professor A. B. Zahlan for his hospitality at the Physics Department of the American University of Beirut, where this work was done.

[8] H. H. Aly and S. Okubo, Z. Physik **200**, 20 (1967).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 5 MAY 1968

# Realization of Poincaré-Group Generators on a Light Cone

ASHER PERES*
*Department of Physics and Astronomy, University of Maryland, College Park, Maryland*

(Received 17 July 1967)

The ten generators of the proper inhomogeneous Lorentz group are explicitly constructed for spin-0 and spin-$\frac{1}{2}$ particles, in the case when wavefunctions are given on a light cone (rather than on an equal time hypersurface, as usual). The advantage of this formulation is that the generators **J** (spatial rotations) and **K** (Lorentz boosts) involve only elementary local operators. The Hamiltonian $H$ and momentum operators **P** also contain the inverse radial momentum $(\partial/\partial r)^{-1}$, but do not involve any square roots. Moreover, only two-component spinors are required for spin-$\frac{1}{2}$ particles.

## I. INTRODUCTION

The problem of covariance in quantum theory is rather different from its classical counterpart.[1] In classical theory, the dynamical variables have numerical values which are equal to the observable values of these variables, and covariance can easily be ascertained if the dynamical variables transform in a definite way (e.g., as tensor components) under the appropriate group.

On the other hand, the dynamical variables of quantum theory are linear operators acting on some Hilbert space. The state of a physical system, as seen

* Permanent address: Technion, Israel Institute of Technology, Haifa, Israel.

[1] For a full discussion of this point, the reader is referred to the lucid article by L. L. Foldy, Phys. Rev. **122**, 275 (1961). No review can do full justice to the clarity of that paper.

by a set of observers,[2] is not specified by ascribing numerical values to the dynamical variables of that system (as in classical physics), but is now represented by a ray in Hilbert space. A different set of observers[2] will attribute to that state a different ray. The problem of covariance is to find the relationship between these rays.

If the two sets of observers are physically equivalent (e.g., related by a Lorentz transformation) this relationship is a unitary one.[3] In the important case of infinitesimal transformations, this unitary transformation can be written as

$$U = 1 + i\epsilon_j G_j, \tag{1}$$

where the $\epsilon_j$ are the infinitesimal parameters which characterize the transformation from one set of observers to another ($j = 1, \cdots, 10$ in the case of inhomogeneous Lorentz transformations), and the $G_j$ are Hermitian operators, which are called the *generators* of the transformation group. They satisfy well-defined commutation relations, depending only on that group. The problem of covariance is essentially solved when these generators have been constructed *explicitly*.[4]

However, not every realization of the Hilbert space, and hence of the Poincaré group generators, is useful. For the purpose of correspondence with classical physics, we are often interested in taking the quantum-mechanical Hilbert space for a system of particles, as *the set of square integrable functions over the classical configuration space of these particles.* Moreover, again for the purpose of correspondence with classical physics, we wish to interpret the operator $x_k$ as the $x$ coordinate of the $k$th particle. This implies that $(x_k)_{\text{op}}$ must transform exactly as the classical coordinate $x_k$, at least under transformations which map the configuration space onto itself. The latter requirement is a stringent constraint on the generators $G_j$: Those generators which map the configuration space onto itself must involve only geometrical quantities, while the dynamics of the system is represented solely by the other generators.

Now, the configuration space of a system of particles is the direct product of the individual configuration spaces of each particle,[5] and the latter are usually taken as $t$ = const hypersurfaces in space–time (with the same $t$ for all particles). In this case, the generators **P** (space translations) and **J** (space rotations) have simple forms, and are additive for several particles, while $H$ (time translations) and **K** (Lorentz boosts) convey dynamical meaning and have complicated forms (even for a single particle).[1] In particular, they involve the nonlocal operator $(p^2 + m^2)^{\frac{1}{2}}$, where $\mathbf{p} = -i\nabla$.[6]

However, it was pointed out long ago by Dirac[7] that it might be more convenient to take, as the one-particle configuration space, a hyperboloid (or possibly a light cone) invariant under homogeneous Lorentz transformations. In this case, the "simple" geometrical generators are **J** and **K** (which are additive for several particles) while $H$ and **P** are the complicated dynamical generators. Some advantages of this formulation were pointed out by Thomas.[8]

The classical generators $H$, **P**, **J**, and **K**, satisfying the Lorentz-group Poisson-bracket relations, were constructed explicitly by Dirac[7] for a single particle, in terms of the canonical variables **p** and **q**. The purpose of the present paper is to derive a quantum realization of these generators, in the special case where the configuration space is a past light cone.[9,10]

This is done in two steps. In Sec. II, we construct **J** and **K**, satisfying

$$[J_m, J_n] = i\epsilon_{mns}J_s, \tag{2}$$

$$[J_m, K_n] = i\epsilon_{mns}K_s, \tag{3}$$

$$[K_m, K_n] = -i\epsilon_{mns}J_s, \tag{4}$$

for particles of arbitrary spin $s$, whose wavefunctions are given on a past light cone.

Sections III and IV are then devoted to the construction of the Hamiltonians of spin-0 and spin-$\frac{1}{2}$ particles, respectively. Once $H$ is known, **P** is readily obtained from

$$[H, \mathbf{K}] = -i\mathbf{P}. \tag{5}$$

[2] Even in classical physics, the state of a physical system must always be referred to some set of observers (and not to a single observer). For example, the familiar formulas for Lorentz transformations relate coordinates $xyzt$ and $x'y'z't'$ in *two synchronized Lorentz frames.* (Ignoring this important point leads to "paradoxes," such as the twin paradox, which actually are not at all paradoxical when properly interpreted.)

[3] E. P. Wigner, Ann. Math. **40**, 149 (1939).

[4] As pointed out by E. Kazes, Phys. Rev. **157**, 1309 (1967), "unless the ten generators of the inhomogeneous Lorentz group can be realized, we may be dealing with an empty formalism."

[5] Or can be reduced to it by a suitable coordinate transformation (no reasonable alternative is known to the author).

[6] Locality can be restored only at the expense of adding redundant components, so that a spin-$s$ wavefunction has $2^{2s+1}$ components (with suitable constraints) rather than $2s + 1$. This artifice leads to the aesthetical feature of $\psi$ satisfying a manifestly covariant wave equation. (Such wave equations are often misinterpreted as being the essence of quantum theory.)

[7] P. A. M. Dirac, Rev. Mod. Phys. **21**, 392 (1949).

[8] L. H. Thomas, Phys. Rev. **85**, 868 (1952).

[9] The generators for a future light cone are simply obtained by reversing the signs of **K** and **P**.

[10] Possibly, the use of a past light cone might obviate to what I. Bloch [Phys. Rev. **156**, 1377 (1967)] calls "some relativistic oddities in the quantum theory of observation." In his paper, Bloch points out that some contradictions "could be avoided if the transition from pure state to mixture somehow took place along a light cone . . ." See also W. C. Davidon and H. Ekstein, J. Math. Phys. **5**, 1588 (1964).

However, it is not at all trivial to construct $H$, because we must satisfy all the other commutation relations, namely

$$[H, \mathbf{J}] = 0, \tag{6}$$

$$[P_m, K_n] = -i\delta_{mn}H, \tag{7}$$

$$[H, \mathbf{P}] = 0, \tag{8}$$

$$[J_m, P_n] = i\epsilon_{mns}P_s, \tag{9}$$

$$[P_m, P_n] = 0. \tag{10}$$

While (6) is satisfied by any rotational scalar, (7) implies that $H$ satisfies

$$[K_m, [K_n, H]] = -\delta_{mn}H, \tag{11}$$

and (8) leads to the nonlinear condition

$$[H, [H, K_m]] = 0, \tag{12}$$

the fulfillment of which is the main difficulty of our problem.

Equations (9) and (10) are consequences of the preceding ones and of the Jacobi identity.

Finally, Sec. V is devoted to a brief discussion of our results, and of possible generalizations. Some auxiliary formulas are listed in an Appendix.

Throughout this paper, we use natural units: $\hbar = c = 1$.

## II. THE HOMOGENEOUS LORENTZ GROUP

In this section, we construct $\mathbf{J}$ and $\mathbf{K}$. We consider first the case of spinless particles, and use polar coordinates, so that the state of a particle is described by a wavefunction $\psi(r, \theta, \varphi)$.

Rotations, generated by $\mathbf{J}$, are transformations of $\theta$ and $\varphi$ leaving the quadratic form $d\theta^2 + \sin^2\theta\, d\varphi^2$ invariant. They are generated, as usual, by[11]

$$L_x = i\left(\sin\varphi\frac{\partial}{\partial\theta} + \cot\theta\cos\varphi\frac{\partial}{\partial\varphi}\right), \tag{13a}$$

$$L_y = i\left(-\cos\varphi\frac{\partial}{\partial\theta} + \cot\theta\sin\varphi\frac{\partial}{\partial\varphi}\right), \tag{13b}$$

$$L_z = -i\frac{\partial}{\partial\varphi}, \tag{13c}$$

with $\mathbf{J} \equiv \mathbf{L}$ for spinless particles.

More generally, $\psi$ will have $(2s + 1)$ components, and then

$$\mathbf{J} = \mathbf{L} + \mathbf{S}, \tag{14}$$

where the $\mathbf{S}$ are (numerical) spin matrices.[12]

We now turn to Lorentz transformations proper. On the past light cone $r = -t$, we have

$$ds^2 \equiv dt^2 - dr^2 - r^2(d\theta^2 + \sin^2\theta\, d\varphi^2) = -r^2(d\theta^2 + \sin^2\theta\, d\varphi^2), \tag{15}$$

so that homogeneous Lorentz transformations, which keep $ds^2$ invariant, induce conformal transformations of the unit sphere.[13,14]

Taking the special case of an infinitesimal boost in the $z$ direction, we have (omitting spin)

$$x' = x, \tag{16a}$$

$$y' = y, \tag{16b}$$

$$z' = z - vt = z + vr. \tag{16c}$$

We thus see that $\delta z = vr$ is generated by $rp_z$, i.e., $K_z = rp_z$. More generally, it is easily seen that $\mathbf{K} = r\mathbf{p}$ satisfies Eqs. (3) and (4) and hence is the solution of our problem for $s = 0$. (The Hermiticity problem will be discussed at the end of this section.)

To generalize this result for nonzero $s$, it is convenient to separate, in $\mathbf{K} = r\mathbf{p}$, radial and angular variables by means of the identity

$$\mathbf{r} \times \mathbf{L} = \mathbf{r}(\mathbf{r} \cdot \mathbf{p}) - r^2\mathbf{p}. \tag{17}$$

We define

$$\mathbf{n} = r^{-1}\mathbf{r} = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta), \tag{18}$$

and further define $p$ as being the Hermitian part of $\mathbf{n} \cdot \mathbf{p}$ [it will be given explicitly in Eq. (23), after we derive the form of the Lorentz-invariant scalar product].

Classically, we have

$$\mathbf{K} = r\mathbf{p} = \mathbf{n}rp - \mathbf{n} \times \mathbf{L}, \tag{19}$$

from which we guess

$$\mathbf{K} = \tfrac{1}{2}(rp + pr)\mathbf{n} + \tfrac{1}{2}(\mathbf{J} \times \mathbf{n} - \mathbf{n} \times \mathbf{J}). \tag{20}$$

It is now a matter of routine to verify that (20) indeed satisfies Eqs. (3) and (4) for any $s$, and therefore is the solution of our problem. Some auxiliary formulas have been listed in the Appendix.

To complete this section, we still have to define a Lorentz-invariant scalar product, i.e., to find the weight function $\rho(r, \theta, \varphi)$ in

$$\langle\psi_1 \mid \psi_2\rangle = \int \psi_1^\dagger(r, \theta, \varphi)\psi_2(r, \theta, \varphi)\rho(r, \theta, \varphi)\, dr\, d\theta\, d\varphi, \tag{21}$$

such that $\langle\psi_1 \mid \psi_2\rangle$ will be invariant under proper homogeneous Lorentz transformations. To this effect, we note that both $r^2\sin\theta\, dt\, dr\, d\theta\, d\varphi$ and

[11] L. I. Schiff, *Quantum Mechanics* (McGraw-Hill Book Co., Inc., New York, 1955), p. 75.

[12] The sum (14) is reducible, but its irreducible parts are not local, and we shall not consider them.

[13] R. K. Sachs, Phys. Rev. **128**, 2851 (1962).

[14] A. Komar, Am. J. Phys. **33**, 1024 (1965).

$\delta(t^2 - r^2)$ are Lorentz scalars, and therefore their product is a scalar. Integration over $t$ then shows that

$$\rho = r \sin\theta, \tag{22}$$

is the desired weight function, up to a multiplicative constant (rather than $r^2 \sin\theta$, as usual).

The result (22) can also be derived by noting that mere rotational symmetry implies that $\rho = f(r) \sin\theta$. Then, direct use of (16) shows that $r \sin\theta$, $dr\,d\theta$ and $d\varphi$ are separately invariant under boosts in the $z$ direction. Q.E.D.

It follows from (21) and (22) that all the components of $\mathbf{L}$ and hence of $\mathbf{J}$ are Hermitian. Likewise

$$p = -i\left(\frac{\partial}{\partial r} + \frac{1}{2r}\right), \tag{23}$$

is also Hermitian, and therefore all the components of $\mathbf{K}$, Eq. (20), are Hermitian.

It is also convenient to define the Hermitian operator

$$q = \tfrac{1}{2}(rp + pr) = -i(\partial/\partial r)r, \tag{24}$$

in terms of which we can readily write down the Casimir invariants

$$\mathbf{J} \cdot \mathbf{K} = \mathbf{J} \cdot \mathbf{n}q, \tag{25}$$

and

$$\mathbf{J}^2 - \mathbf{K}^2 = (\mathbf{J} \cdot \mathbf{n})^2 - q^2 - 1. \tag{26}$$

We thus see that $q$ and $\mathbf{J} \cdot \mathbf{n}$ are invariant under proper homogeneous Lorentz transformations.

## III. THE SPIN-0 HAMILTONIAN

The next, and much more difficult problem, is to find an $H$ satisfying (6), (11), and (12). In the classical case, with Poisson brackets instead of commutators, it was shown by Dirac[7] that

$$H = \frac{1}{2}\left(p + \frac{m^2}{p} + \frac{J^2}{rpr}\right). \tag{27}$$

This expression is Hermitian, but, unfortunately, satisfies neither (11) nor (12) in quantum theory (the right-hand sides are of the order of $\hbar^3$).

The correct expression for $H$ can however be easily found, by seeking the "missing terms" necessary to satisfy (11) and (12). It is

$$H = \frac{1}{2}\left(p + \frac{m^2}{p + (\frac{1}{4}rpr)} + \frac{J^2 + \frac{1}{4}}{rpr}\right), \tag{28}$$

from which we derive

$$\mathbf{P} = -H\mathbf{n} + \mathbf{K}(r^{-1} - i/2rpr), \tag{29}$$

$$= -\tfrac{1}{2}(H\mathbf{n} + \mathbf{n}H) + \tfrac{1}{2}(\mathbf{K}r^{-1} + r^{-1}\mathbf{K}) + (\mathbf{n}/4rpr). \tag{30}$$

[The correction terms in (28) and the last term in (30) would have $\hbar^2$ factors, if we had not set $\hbar = 1$.] Some auxiliary formulas are given in the Appendix.

## IV. THE SPIN-½ HAMILTONIAN

The preceding results are valid only for spin-0, because, when substituted in (12), the right-hand side of (12) is proportional to $\mathbf{n} \cdot \mathbf{J}$, which vanishes only for spin-0.

A possible approach to get a Hamiltonian for spin-½ particles is to transform the Dirac equation to coordinates $r$, $\theta$, $\varphi$, and $u = r + t$, and to set $H = i\partial/\partial u$.[15]

We first go over to the equal time polar coordinates $tr\theta\varphi$. The Dirac equation then reads[16]

$$i\partial\psi/\partial t = -i\alpha_r(\partial\psi/\partial r) + i\alpha_r r^{-1}\boldsymbol{\sigma}' \cdot \mathbf{L}\psi + \beta m\psi, \tag{31}$$

where $\alpha_r = \boldsymbol{\alpha} \cdot \mathbf{n}$ and the $\boldsymbol{\sigma}'$ are $4 \times 4$ block-diagonal matrices, the blocks of which are Pauli's $\boldsymbol{\sigma}$ (in the representation where $\beta$ is diagonal, the blocks being $I$ and $-I$). We now introduce the null coordinate $u = r + t$ and Eq. (31) becomes

$$i(1 + \alpha_r)(\partial\psi/\partial u) = (\alpha_r p_r + i\alpha_r r^{-1}\beta k + \beta m)\psi, \tag{32}$$

where[16]

$$p_r = -i[(\partial/\partial r) + r^{-1}], \tag{33}$$

and

$$k = \beta(\boldsymbol{\sigma}' \cdot \mathbf{L} + 1). \tag{34}$$

Note that $k$ commutes with $r$, $p_r$, $\alpha_r$, and $\beta$, and that $k^2 = J^2 + \frac{1}{4}$.

We now define two projection operators

$$P_\pm = \tfrac{1}{2}(1 \pm \alpha_r), \tag{35}$$

satisfying $P_\pm^2 = P_\pm$ and $P_+P_- = 0$, and write

$$\psi = \psi_+ + \psi_-, \tag{36}$$

where

$$\psi_\pm = P_\pm\psi. \tag{37}$$

Noting that $\alpha_r = P_+ - P_-$ and that

$$P_\pm\beta = \beta P_\mp, \tag{38}$$

Eq. (32) becomes

$$2i(\partial\psi_+/\partial r) = p_r(\psi_+ - \psi_-) - ir^{-1}\beta k(\psi_+ - \psi_-) + \beta m(\psi_+ + \psi_-). \tag{39}$$

Multiplying on the left by $P_-$, we obtain

$$\psi_- = (p_r)^{-1}\beta(m - ir^{-1}k)\psi_+. \tag{40}$$

[15] This method does not work with the Klein–Gordon equation, presumably because of different normalizations of $\psi$. Actually, even the Dirac $\psi$ does not satisfy the normalization (21), and the success of this method for spin-½ particles may be just a lucky accident.

[16] Reference 11, p. 334.

On the other hand, multiplying (39) by $P_+$ yields

$$2i(\partial\psi_+/\partial u) = p_r\psi_+ + ir^{-1}\beta k\psi_- + \beta m\psi_-, \quad (41)$$

$$= p_r\psi_+ + (m + ir^{-1}k)p_r^{-1}(m - ir^{-1}k)\psi_+. \quad (42)$$

From the above result, which is exact, we guess[17]

$$H = \tfrac{1}{2}[p + (kr^{-1} - im)p^{-1}(kr^{-1} + im)], \quad (43)$$

$$= \frac{1}{2}\left(p + \frac{m^2}{p} + \frac{J^2 + \frac{1}{4}}{rpr} - \frac{mk}{rppr}\right). \quad (44)$$

It is now a matter of tedious, but straightforward, calculations to check that (11) and (12) are satisfied. Details are given in the Appendix.

Moreover, from Eq. (34), we note that $k$ is block diagonal, the blocks being $\pm(\boldsymbol{\sigma}\cdot\mathbf{L} + 1)$. Hence $H$ is reducible, and it is obviously possible to retain only one of the blocks, e.g.,

$$k = \boldsymbol{\sigma}\cdot\mathbf{L} + 1 = \boldsymbol{\sigma}\cdot\mathbf{J} - \tfrac{1}{2}, \quad (45)$$

so that two-component spinors are indeed adequate for spin-$\frac{1}{2}$ particles (in conformity with Sec. II). This result is the main advantage of the present formulation of quantum mechanics.

The two possible forms of the Hamiltonian, corresponding to opposite signs of $k$, may be interpreted as pertaining to particles and antiparticles, respectively.

Finally, we write the explicit form of the momentum operators:

$$\mathbf{P} = \tfrac{1}{2}[-\mathbf{n}p + r^{-1}\mathbf{K} + \mathbf{K}r^{-1} - (kr^{-1} - im)\mathbf{n}p^{-1} \times (kr^{-1} + im)], \quad (46)$$

$$= -\tfrac{1}{2}(\mathbf{n}H + H\mathbf{n}) + \tfrac{1}{2}(r^{-1}\mathbf{K} + \mathbf{K}r^{-1}) + (\mathbf{n}/4rpr) - m\boldsymbol{\sigma}\times\mathbf{n}(rp + pr)/4rppr. \quad (47)$$

## V. OUTLOOK

There is a striking similarity between Eqs. (28–30) which give $H$ and $\mathbf{P}$ for spin-0, and Eqs. (43–47), which refer to spin-$\frac{1}{2}$. However, there are also important differences, and there is no obvious generalization to arbitrary spin. More powerful methods will therefore be needed in order to obtain $H$ and $\mathbf{P}$ for higher spin.

Another important problem, which we have not discussed, is that of interactions.[1,7,8] It was pointed out by Dirac[7] that the "point form" of quantum mechanics (of which the present work is a special case) might possibly be more suitable than the more familiar "instant form" to discuss relativistic interactions between particles. Further calculations will clearly be needed to investigate this point.

[17] Note the difference between Eqs. (23) and (33), and see Ref. 15 above.


## ACKNOWLEDGMENT

It is a pleasure for the author to express his gratitude to Professor J. Weber for the warm hospitality shown to him at the University of Maryland.


## APPENDIX

The following is an assortment of formulas which were useful in our calculations. We recall that $\mathbf{J}$ is given by Eq. (14) and $\mathbf{K}$ by Eq. (20).

$$\tfrac{1}{2}(\mathbf{J}\times\mathbf{n} - \mathbf{n}\times\mathbf{J}) = \mathbf{J}\times\mathbf{n} - i\mathbf{n} = -\mathbf{n}\times\mathbf{J} + i\mathbf{n}, \quad (A1)$$

$$= \tfrac{1}{2}i[J^2, \mathbf{n}]. \quad (A2)$$

$$[n_a, (J\times n)_b] = i(\delta_{ab} - n_an_b), \quad (A3)$$

$$[(J\times n)_a, (J\times n)_b] = -i\epsilon_{abc}J_c, \quad (A4)$$

$$[n_a, \mathbf{J}\cdot\mathbf{n}] = 0, \quad (A5)$$

$$[K_a, n_b] = i(n_an_b - \delta_{ab}), \quad (A6)$$

$$[\mathbf{K}, r^{-1}] = i\mathbf{n}/r, \quad (A7)$$

$$[\mathbf{K}, p] = i\mathbf{n}p, \quad (A8)$$

$$[\mathbf{K}, (rpr)^{-1}] = i\mathbf{n}/rpr, \quad (A9)$$

$$[\mathbf{K}, p^{-1}] = -i\mathbf{n}/p, \quad (A10)$$

$$[K, (rppr)^{-1}] = 0, \quad (A11)$$

$$[K_a, \sigma_b] = i(n_b\sigma_a - \delta_{ab}\mathbf{n}\cdot\boldsymbol{\sigma}), \quad (A12)$$

$$[K_a, \mathbf{n}\cdot\boldsymbol{\sigma}] = 0, \quad (A13)$$

$$[p^{-1}, r^{-1}] = -i/rppr = -i/prrp, \quad (A14)$$

$$[r^{-1}, (rpr)^{-1}] = i/(rpr)^2, \quad (A15)$$

$$k = \boldsymbol{\sigma}\cdot\mathbf{L} + 1 = \boldsymbol{\sigma}\cdot\mathbf{J} - \tfrac{1}{2}, \quad (A16)$$

$$[K_a, k] = i[\tfrac{1}{2}\epsilon_{abc}\sigma_bn_c(rp + pr + i) - n_ak], \quad (A17)$$

$$[n_a, k] = i\epsilon_{abc}\sigma_bn_c, \quad (A18)$$

$$[K_a, [K_b, k]] = -\delta_{ab}k, \quad (A19)$$

$$[K_a, k/r] = i\epsilon_{abc}\sigma_bn_cp, \quad (A20)$$

$$[k^2, [k, \mathbf{n}]] = k\mathbf{n} + \mathbf{n}k. \quad (A21)$$

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 5 MAY 1968

# Properties of Higher-Order Commutator Products and the Baker-Hausdorff Formula

ERIK ERIKSEN
*Institute of Physics, University of Oslo, Blindern, Norway*

(Received 2 October 1967)

The element $z = \log e^x e^y$, which is known to be an element of the Lie-algebra generated by $x$ and $y$, is expressed as a commutator series in $x$ and $y$ with coefficients given in terms of certain fixed polynomials. The result is given explicitly to sixth order. Useful recurrence relations are obtained. The method is based on certain properties of higher-order commutator products, particularly their idempotent character.

## I. INTRODUCTION

The problem of compounding exponentials of noncommuting quantities is of interest in various fields of mathematics[1] and theoretical physics,[2] for instance: in group theory, perturbation theory, transformation theory, and statistical mechanics. The problem has its origin in group theory where it led to the Baker–Hausdorff theorem:

If $e^x e^y = e^z$, then $z$ is a Lie element generated by $x$ and $y$; that is, $z$ is a sum of repeated commutators of $x$ and $y$ (including $x$ and $y$ itself).[3]

The theorem was proved by Baker[4] and Hausdorff[5] by a recursive construction of $z$. A simpler proof has been given by Magnus[1] by means of a characterization of Lie elements found by Friedrichs.[6]

The present investigation was stimulated by the need of an explicit commutator expansion for practical use. The author found only scattered results available in the literature, except the systematic approach due to Goldberg.[7] His expansion, however, is not an explicit commutator expansion, but an expansion in terms of the monomials

$$1, x, y, x^2, xy, yx, y^2, x^3, x^2y, xyx, \text{ etc.},$$

which are linearly independent elements (basic ring elements) in the free associative ring generated by $x$ and $y$. The coefficient for any specified monomial is expressed by Goldberg as an integral over a product of certain fixed polynomials.

One of the main results of the present investigation is a commutator expansion, with coefficients which are closely related to those of Goldberg. There is a new feature in the commutator expansion compared to the expansion in terms of monomials; namely, while the manifold of monomials by definition are linearly independent, the corresponding commutator products are not. Consider, for instance, the monomials $xyxy$ and $xy^2x$, which are linearly independent elements in the free associative ring. The corresponding commutator products, however, exhibit linear dependence. In fact, we have the identity

$$[[[x, y], x], y] = [[[x, y], y], x].$$

This means that one and the same result can be expressed in numerous ways depending on which Lie elements are chosen as a basis.

A problem of special interest occurs when one of the generators, say $y$, is considered as "small," for instance, when $y$ is attached to a small parameter. Then we may be interested in a commutator expansion

$$z = z_{(0)} + z_{(1)} + \cdots + z_{(n)} + \cdots,$$

where the terms are of increasing degree with respect to $y$. For this situation we have obtained a recurrence relation connecting $z_{(n)}$ with a certain part of $z_{(n-1)}$.

The starting point leading to the results above is a study of the structure of higher-order commutator products. In Sec. II we prove various properties of these products, particularly that the so called "curly bracket operator" is essentially idempotent when applied to a homogeneous function of the generators. This idempotency has been proved earlier (and independently of each other) by Dynkin,[8] Specht,[9] and Wever,[10] using methods which differ somewhat from the one employed here.

[1] W. Magnus, Commun. Pure Appl. Math. 7, 649 (1954).
[2] K. Kumar, J. Math. Phys. 6, 1923, 1928 (1965).
[3] To be precise, $x$ and $y$ should be considered as free generators of an associative ring $R$ over the field of real numbers. Then the element $z$ in $R$ defined by $e^z = e^x e^y$ is a Lie element.
[4] H. F. Baker, Proc. London Math. Soc., Second Series 3, 24 (1904).
[5] F. Hausdorff, Ber. Verhandl. Sächs Akad. Wiss., Leipzig, Math. Phys. Kl. 58, 19 (1906).
[6] K. O. Friedrichs, Commun. Pure Appl. Math. 6, 1 (1953).
[7] K. Goldberg, Duke J. Math. 23, 13 (1956).
[8] E. B. Dynkin, Dokl. Akad. Nauk. SSSR 57, 323 (1947).
[9] W. Specht, Math. Z. 51, 367 (1949).
[10] F. Wever, Math. Ann. 120, 563 (1947–49).

## II. COMMUTATOR-IDENTITIES AND LIE ELEMENTS

### A. Definition of Curly Bracket Operator and Left-Ordered Commutator Product

Let $f$ and $g$ be elements in the ring generated by the operators $x_1, x_2, \cdots, x_N$, and let $c_i$ be an element in the field of coefficients. Following Magnus[1] we define a curly bracket operator $\{\ \}$ with the properties,

$$\{c_i\} = 0, \tag{1a}$$

$$\{x_i\} = x_i, \tag{1b}$$

$$\{x_i \cdots x_j x_k\} = [\{x_i \cdots x_j\}, x_k], \tag{1c}$$

$$\{c_1 f + c_2 g\} = c_1\{f\} + c_2\{g\}. \tag{1d}$$

The following consequences are obvious:

$$\{x_i x_j\} = [x_i, x_j] = -\{x_j x_i\} \quad \text{(ANTISYMMETRY)}, \tag{2}$$

$$\{x_i x_j x_k\} + \{x_j x_k x_i\} + \{x_k x_i x_j\} = 0 \quad \text{(JACOBI IDENTITY)}, \tag{3}$$

$$\{x_i x_j \cdots x_k x_l\} = [[\cdots [x_i, x_j], \cdots, x_k], x_l]. \tag{4}$$

A repeated commutator product of the last type, with all brackets [ standing to the left, will be called a left-ordered commutator product (German: links normiert). The transformation from a left-ordered to a right-ordered product is trivial:

$$[[\cdots [x_1, x_2], \cdots, x_{n-1}], x_n] = (-1)^{n-1}[x_n, [x_{n-1}, \cdots, [x_2, x_1] \cdots]].$$

The above definition of the curly bracket operator is unique, provided we know which elements are considered as basic arguments in the commutator operation (the elements $x_1, x_2, \cdots, x_N$ above). In the following we need to introduce various sets of arguments with respect to which the curly bracket operator is defined. In order to avoid clumsy notation we take the risk of introducing the following convention: A curly bracket without index means that the arguments of the commutator operation are a certain specified set of generators ($x_1, x_2, \cdots, x_N$ in the rest of the present section). If the curly bracket is equipped with one or more indices, say $u$ and $v$, then $u$ and $v$ appear explicitly as arguments in the function and are considered as basic and independent arguments in the commutator operation.

An example will make the point clear. Let $u$ and $v$ be the elements $u = x_1 x_2$, $v = x_3 x_1$. Then we have

$$\{uv\} = [[[x_1, x_2], x_3], x_1],$$

$$\{uv\}_u = [[x_1 x_2, x_3], x_1],$$

$$\{uv\}_{u,v} = [x_1 x_2, x_3 x_1].$$

### B. Properties of the Curly Bracket Operator

Let $f$, $g$, and $h$ be elements in the free associative ring generated by $x_1, x_2, \cdots, x_N$, and let $F$, $G$, and $G_n$ be the elements

$$F = \{f\}, \quad G = \{g\}, \quad G_n = \{g_n\},$$

where $g_n$ is the $n$th-degree part of $g$. Then, if the zeroth-degree part of $f$ is zero, the following identities are true:

$$\{f\ g\} = \{F\ g\}_F, \tag{5}$$

$$\{f\ G\} = [F, G], \tag{6}$$

$$\{f\ G\ h\} = \{f\ G\ h\}_G, \tag{7}$$

$$\{G_n\ h\} = n\{G_n\ h\}_{G_n}. \tag{8}$$

For proof see the Appendix.

For $h = 1$, the last equation reads

$$\{G_n\} = nG_n, \tag{9}$$

$$\{\{g_n\}\} = n\{g_n\}. \tag{10}$$

Thus the curly bracket operation on a homogenous function is an essentially idempotent operation.

From the idempotency property we deduce the following statement:

Let $r_n$ be a homogenous function of $n$th degree. Then $\{r_n\} = 0$ if and only if there exists a function $g_n$ such that

$$r_n = g_n - (1/n)\{g_n\}. \tag{11}$$

*Proof:* If (11) is true, we have $\{r_n\} = \{g_n\} - (1/n)\{\{g_n\}\} = 0$. If $\{r_n\} = 0$, we put $g_n = r_n$, and (11) is true.

Some examples of vanishing curly brackets for the case of two generators are given below. In order to be in agreement with the notation in the Baker–Hausdorff formula, the generators are denoted by $x$ and $y$:

$$\{x^n\} = 0, \quad n = 2, 3, 4, \cdots, \tag{12a}$$

$$\{xy + yx\} = 0, \tag{12b}$$

$$\{xy(xy - yx)\} = 0, \tag{12c}$$

$$\{xyx(x^2y - 2xyx + yx^2)\} = 0, \tag{12d}$$

$$\{xy(3xyxy - 3yxyx + y^2x^2 - x^2y^2)\} = 0. \tag{12e}$$

### C. Application to Lie Elements

The Lie elements (or Lie functions) with respect to a set of generators $x_1, x_2, \cdots, x_N$ are defined recursively as follows. The commutator product $[u, v]$ of the Lie elements $u$ and $v$ is a Lie element. Any linear combination (finite or infinite) of Lie elements is a Lie element. The generators are Lie elements (of degree one).

The following theorem is of importance in connection with the Baker–Hausdorff formula.

*Theorem:* An element $L$ is a Lie element if and only if there exists an element $l$ such that

$$L = \{l\}. \tag{13}$$

*Proof:* If $L$ is a Lie element of first or second degree, the existence of the element $l$ is obvious, since $x_i = \{x_i\}$ and $[x_i, x_j] = \{x_i x_j\}$. According to Eq. (6), the commutator product of two curly bracket expressions is a curly bracket expression. Hence, from the recursive definition of Lie elements, it follows that any Lie element may be written as a curly bracket expression. The reverse statement, that Eq. (13) implies that $L$ is a Lie element, is an obvious consequence of the definitions of the curly bracket operator and of the Lie element.

We can now formulate the following characterization of a homogenous Lie element:

A homogenous element $L_n$ of $n$th degree is a Lie element if and only if

$$\{L_n\} = nL_n. \tag{14}$$

*Proof:* If $L_n$ is a Lie element, then by Eq. (13) there exists an element $l_n$ such that $L_n = \{l_n\}$. Hence, $\{L_n\} = \{\{l_n\}\} = n\{l_n\} = nL_n$. Conversely, if Eq. (14) is true, $L_n$ can be expressed as $L_n = (1/n)\{L_n\}$, which by Eq. (13) implies that $L_n$ is a Lie element.

Consider a Lie function $L$ which is written as a sum of terms of increasing degree,

$$L = \sum_{n=1}^{\infty} L_n. \tag{15}$$

Then, by means of Eq. (14),

$$\{L\} = \sum_{n=1}^{\infty} nL_n. \tag{16}$$

## III. THE BAKER-HAUSDORFF FORMULA

### A. Some Symmetry Properties

Henceforth, the generators are $x$ and $y$. They are arguments in the curly bracket operation when nothing else is denoted.

As is well known,[1,4,5] the function $z$ defined by

$$e^z = e^x e^y, \quad z = f(x, y) \tag{17}$$

is a Lie function of $x$ and $y$. By means of the inverted equation $\exp(-z) = \exp(-y)\exp(-x)$, we obtain the symmetry property

$$z = f(x, y) = -f(-y, -x). \tag{18}$$

If we make the decomposition

$$z = S + A, \quad S \text{ ODD DEGREE}, \quad A \text{ EVEN DEGREE}, \tag{19}$$

the symmetry property (18) tells us that the odd part is symmetric (with respect to permutation of the arguments), while the even part is antisymmetric:

$$z = f(x, y) = S + A, \quad f(y, x) = S - A. \tag{20}$$

Next consider the equation

$$e^{f(y,x)} = e^y e^x = e^{-x} e^z e^x = \exp(e^{-x} z e^x)$$

from which we extract

$$f(y, x) = e^{-x} z e^x. \tag{21}$$

By means of a well-known commutator expansion, we get in our notation

$$f(y, x) = \{ze^x\}_z, \tag{22}$$

That is,

$$S - A = \{Se^x\}_S + \{Ae^x\}_A,$$

$$\{A(e^x + 1)\}_A = \{S(1 - e^x)\}_S. \tag{23}$$

This equation can be solved with respect to $A$ by multiplication from the right by $(e^x + 1)^{-1}$, followed by curly bracket operation

$$A = -\{S \tanh(x/2)\}_S. \tag{24a}$$

By permutation of $x$ and $y$,

$$A = \{S \tanh(y/2)\}_S. \tag{24b}$$

These are useful relations, by which $A$ can easily be obtained to one degree higher than $S$. They exhibit a peculiar property, viz., that $A$ can always be expressed in such a way that (say) $y$ is to the right in the curly bracket.

Combining the last two equations, we get the interesting relation

$$\{S(\tanh(x/2) + \tanh(y/2))\}_S = 0. \tag{25}$$

The author has not been able to utilize this relation to obtain a recursive solution for $S$.

### B. Expansion of $z = \log(e^x e^y)$ in a Commutator Series

The idea of the present method is to utilize the properties of Lie functions with respect to the curly bracket operation. If the Lie function $z$ is known as an expansion in terms of basic ring elements (monomials),

$$z = \sum_{n=1}^{\infty} z_n \quad (z_n, \text{ TERMS OF DEGREE } n), \tag{26}$$

then, by Eq. (14), $z$ can be expressed as an explicit commutator series,

$$z = \sum_{n=1}^{\infty} (1/n)\{z_n\}. \tag{27}$$

An expansion of the first type was, in fact, found by Goldberg,[7] who attacked the problem from a combinatorial point of view. The present derivation of

Goldberg's result is based on an integral representation, and is presented here because of its simplicity and to exhibit some intermediate results.

If we imagine that $z$ is written as a sum of monomials, it is obvious that we can make a unique decomposition,

$$z = x\xi(x, y) + y\eta(x, y), \tag{28}$$

into terms $x\xi(x, y)$ beginning with a power of $x$, and terms $y\eta(x, y)$ beginning with a power of $y$. From Eq. (18), we get the symmetry transform $\eta(x, y) = \xi(-y, -x)$ by which $z$ can be expressed by one of the parts, say

$$z = x\xi(x, y) + y\xi(-y, -x). \tag{29}$$

The result of curly bracket operation is

$$\{z\} = \sum_{n=1}^{\infty} nz_n = \{x\xi(x, y)\} + \{y\xi(-y, -x)\}. \tag{30}$$

Hence, if $\xi(x, y)$ is known as a power series in $x$ and $y$, $z_n$ may easily be expressed as a sum of repeated commutator products.

Our starting point is the identity

$$z = \int_0^1 g(z, p)\, dp. \tag{31}$$

The function

$$g(z, p) = [(e^z - 1)^{-1} + 1 - p]^{-1} \tag{32}$$

is of fundamental importance in the present paper, as well as in Goldberg's. Making use of the notation

$$a = e^x - 1, \quad b = e^y - 1, \quad \beta = 1 - e^{-y},$$

and introducing $e^z = e^x e^y$ in the integral (31), we obtain

$$\begin{aligned} z &= (e^x e^y - 1)\int_0^1 [1 + (e^x e^y - 1)(1 - p)]^{-1}\, dp \\ &= (e^x - e^{-y})\int_0^1 [e^{-y} + (e^x - e^{-y})(1 - p)]^{-1}\, dp \\ &= (a + \beta)\int_0^1 [1 - a(p - 1) - \beta p]^{-1}\, dp. \end{aligned} \tag{33}$$

From this result we select terms beginning with a power of $x$:

$$\begin{aligned} x\xi(x, y) &= a\int_0^1 [1 - a(p - 1) - \beta p]^{-1}\, dp \\ &= a\int_0^1 dp \sum_{n=0}^{\infty} [a(p - 1) + \beta p]^n \\ &= \int_0^1 dp \sum (p - 1)^{\tau-1} p^{\sigma} a^{s_1}\beta^{s_2}a^{s_3}\cdots(a \vee \beta)^{s_m}, \end{aligned} \tag{34}$$

$$\tau = s_1 + s_3 + \cdots, \quad \sigma = s_2 + s_4 + \cdots. \tag{35}$$

The symbol $a \vee \beta$ means $a$ if $m$ is odd, and $\beta$ if $m$ is even. The sum is taken over all positive integer values of $m$ and $s_j$, such that the possible terms in the sum are

$$a^{s_1},\ a^{s_1}\beta^{s_2},\ a^{s_1}\beta^{s_2}a^{s_3},\ \cdots$$

with the exponents running independently over all positive integers.

From Eq. (35) we could proceed by first performing the integration. This would result in an expansion in powers of $a$ and $\beta$:

$$x\xi(x, y) = \sum (-1)^{r-1} B(\tau, \sigma + 1) a^{s_1}\beta^{s_2}a^{s_3}\cdots(a \vee \beta)^{s_m}. \tag{36}$$

The function $B$ is the Beta function:

$$\begin{aligned} B(u, v) = B(v, u) &= \int_0^1 p^{u-1}(1 - p)^{v-1}\, dp \\ &= \Gamma(u)\Gamma(v)/\Gamma(u + v). \end{aligned}$$

This and other expansions can be made by starting from Eq. (34) or Eq. (35). However, we shall proceed along the following lines. In Eq. (35) we sum over $s_1, s_2, \cdots, s_m$, keeping $m$ constant. Then, by the symmetry property

$$g(-y, 1 - p) = -g(y, p), \tag{37}$$

the result can be presented as

$$\begin{aligned} x\xi(x, y) = \sum_{m=1}^{\infty} \int_0^1 dp\, p^{m'}(p - 1)^{m''} g(x, p)g(y, p) \\ \times g(x, p)\cdots[g(x, p) \vee g(y, p)], \end{aligned} \tag{38}$$

where $m' = [m/2]$ and $m'' = [(m - 1)/2]$. The number of $g$ functions is $m$.

For the sake of completeness we also give the formula

$$\begin{aligned} y\eta(x, y) = \sum_{m=1}^{\infty} \int_0^1 dp\, p^{m''}(p - 1)^{m'} g(y, p)g(x, p) \\ \times g(y, p)\cdots[g(y, p) \vee g(x, p)], \end{aligned} \tag{39}$$

which is obtained from Eq. (38) by a symmetry transform.

Following Goldberg,[7] to whom we refer for further details, we introduce the polynomials $G_n(p)$, $n = 1, 2, \cdots$, defined by the generating function

$$g(x, p) = \sum_{n=1}^{\infty} G_n(p)x^n. \tag{40}$$

Then, from (38) and (39) we obtain Goldberg's result:

$$x\xi(x, y) = \sum c_x(s_1, s_2, \cdots, s_m)x^{s_1}y^{s_2}x^{s_3}\cdots(x \vee y)^{s_m}, \tag{41}$$

$$y\eta(x, y) = \sum c_y(s_1, s_2, \cdots, s_m)y^{s_1}x^{s_2}y^{s_3}\cdots(y \vee x)^{s_m}, \tag{42}$$

$$c_x(s_1, s_2, \cdots, s_m) = (-1)^{n+1}c_y(s_1, s_2, \cdots, s_m) = \int_0^1 dp\, p^{m'}(p-1)^{m''}G_{s_1}(p)G_{s_2}(p)\cdots G_{s_m}(p),$$
$$n = s_1 + s_2 + \cdots + s_m. \tag{43}$$

The summation is over all positive integer values of $m$ and $s_j$.

Now, the sum of (41) and (42) has the extremely well-hidden property of being a Lie function, the property which can be made explicit by the curly bracket operator:

$$\{z\} = \sum_{n=1}^{\infty} nz_n = \{x\xi(x, y)\} + \{y\eta(x, y)\}.$$

From the definition of the curly bracket operator it is obvious that terms like $\{x^{s_1}y^{s_2}\cdots\}$, $s_1 > 1$, are zero. Thus, from Eq. (41)

$$\{x\xi(x, y)\} = x + \sum_{m>1} c_x(1, s_2, \cdots, s_m) \times \{xy^{s_2}x^{s_3}\cdots(x \vee y)^{s_m}\}.$$

By means of the symmetry property (29), our result may be expressed as

$$z = x + y + \{x\varphi(x, y)\} + \{y\varphi(-y, -x)\}, \tag{44}$$
$$\varphi(x, y) = \sum b(r_1, r_2, \cdots, r_m)y^{r_1}x^{r_2}\cdots(y \vee x)^{r_m}, \tag{45}$$

where the coefficients are given by

$$b(r_1, r_2, \cdots, r_m) = \left(1 + \sum_{i=1}^{m} r_i\right)^{-1} c_x(1, r_1, r_2, \cdots, r_m) \tag{46}$$

and the sum is taken over all positive integer values of $m$ and $r_i$.

It should be emphasized that an expansion in terms of commutators may be expressed in an infinite number of ways. This is due to the linear dependence of the commutators, see Eqs. (12, a–e). For instance, we know from Eqs. (24) that the antisymmetric part $A$ of $z$ can always be expressed as the curly bracket of a function whose last factor is $x$ (or $y$). The expansion which is given above has a supernumerary number of higher-order commutators. It seems rather difficult to obtain an explicit expansion such that $z_n$ is expressed as a curly bracket operating on the fewest possible number of basic ring elements. To sixth order we have obtained the results:

$$z_1 = x + y, \qquad z_2 = \tfrac{1}{2}\{xy\},$$
$$z_3 = (\tfrac{1}{12})\{xy(y - x)\}, \quad z_4 = -(\tfrac{1}{24})\{xyxy\},$$
$$z_5 = (\tfrac{1}{720})\{xy(x^3 - 2x^2y + 6xyx - 6yxy + 2y^2x - y^3)\},$$
$$z_6 = (\tfrac{1}{1440})\{xy(x^3 - 2x^2y + 6xyx - xy^2 + 2y^2x)y\}. \tag{47}$$

### C. A Recursive Solution

Various recursive solutions may be written down. The following one is relatively simple, and well adapted for practical use.

We start with the identity

$$(e^y + e^{-x})z = (e^y - e^{-x})z \coth(z/2), \tag{48}$$

which is easily shown to be true by multiplication with $e^x$ from the left. As usual we consider $x$ and $y$ as arguments and $z$ as a function of them. The curly bracket operation gives the result

$$\{(2 + y - x)z\} = \{(x + y)z \coth(z/2)\},$$

which is true due to the fact that terms starting with $x^n$ or $y^n$, $n > 1$, are annihilated by the curly bracket operator. With the notation $u = x + y$, $v = x - y$, the result reads

$$\{z\} = \tfrac{1}{2}\{vz\} + \{u(z/2)\coth(z/2)\}. \tag{49}$$

That is, we have a recursive solution,

$$z_1 = u, \quad nz_n = \tfrac{1}{2}\{vz_{n-1}\} + \{u[(z/2)\coth(z/2)]_{n-1}\}. \tag{50}$$

The index $n - 1$ denotes the $(n - 1)$th-degree part of the term. We can take advantage of Eq. (7), by which any of the Lie functions $z_i$ may be considered as an independent argument in the curly bracket operation. For instance,

$$\{vz_{n-1}\} = \{vz_{n-1}\}_{z_{n-1}} = [v, z_{n-1}].$$

In the usual commutator notation the first recurrence relations read:

$$z_1 = u,$$
$$2z_2 = \tfrac{1}{2}[v, z_1],$$
$$3z_3 = \tfrac{1}{2}[v, z_2],$$
$$4z_4 = \tfrac{1}{2}[v, z_3] + (\tfrac{1}{12})[[u, z_2], z_1],$$
$$5z_5 = \tfrac{1}{2}[v, z_4] + (\tfrac{1}{12})[[u, z_2], z_2] + (\tfrac{1}{12})[[u, z_3], z_1]. \tag{51}$$

### D. Expansion in Powers of $y$

We shall derive a recurrence relation which gives $z$ as a commutator series

$$z = z_{(0)} + z_{(1)} + \cdots + z_{(n)} + \cdots \tag{52}$$

of increasing degree $(n)$ with respect to $y$. The function $z_{(n)}$ is obviously a Lie function for all $n$, and $z_{(0)}$ is the element $x$. By the symmetry property, $z = f(x, y) = -f(-y, -x)$, the expansion above can easily be translated into one of increasing powers of $x$.

If we let the parameter $t$ be attached to the generator $y$, $z$ becomes a function of $t$,

$$e^z = e^x e^{vt}, \tag{53}$$

and Eq. (52) becomes an expansion in powers of $t$. We shall obtain an expression for the derivative $\dot{z}$ with respect to $t$.

In general one has the expansion

$$e^{w+\Delta w} = e^w + e^w\{\Delta w(e^w - 1)/w\}_{\Delta w, w} + \text{TERMS OF HIGHER DEGREE},$$

which gives the formula

$$de^w/dt = e^w\{\dot{w}(e^w - 1)/w\}_{\dot{w},w}. \tag{54}$$

We take the derivative of Eq. (53) with respect to $t$, and multiply from the left by $e^{-z}$. The result is the equation

$$\{\dot{z}(e^z - 1)/z\}_{\dot{z},z} = y,$$

which may be solved with respect to $\dot{z}$:

$$\dot{z} = \{yz(e^z - 1)^{-1}\}_z = \{yz(e^z - 1)^{-1}\}.$$

We put $t = 1$ and get

$$\sum_{n=1}^{\infty} nz_{(n)} = \{yz(e^z - 1)^{-1}\}. \tag{55}$$

The expression $z(e^z - 1)^{-1}$ is related in a simple way to the function $\xi(x, y)$ [and to the function $\eta(x, y)$] defined by Eq. (29). If we write $z$ in the form

$$\begin{aligned} z &= (e^x e^y - 1)(e^z - 1)^{-1}z \\ &= (e^x - 1)e^y(e^z - 1)^{-1}z + (e^y - 1)(e^z - 1)^{-1}z, \end{aligned}$$

we find the relation

$$(e^z - 1)^{-1}z = e^{-y}x(e^x - 1)^{-1}\xi(x, y). \tag{56}$$

That is

$$\sum_{n=1}^{\infty} nz_{(n)} = \{yx(e^x - 1)^{-1}\xi(x, y)\}, \tag{57}$$

which gives a recurrence relation of formally simple structure,

$$nz_{(n)} = \{yx(e^x - 1)^{-1}\xi(x, y)_{(n-1)}\}. \tag{58}$$

The main difficulty in application is, of course, to select the term $\xi(x, y)_{(n-1)}$ from $z_{(n-1)}$.

The first step goes as follows:

$$z_{(0)} = x, \quad \text{i.e.,} \quad \xi_{(0)} = 1,$$

$$z_{(1)} = \{yx(e^x - 1)^{-1}\} = y + \sum_{n=1}^{\infty}(B_n/n!)\{yx^n\} \tag{59}$$

($B_n$ BERNOULLI NUMBER).

For the next step, consider a term like $\{yx^n\}$, $n > 0$. The term starting with $y$ is $yx^n$, and the term starting with a power of $x$ is

$$\{yx^n\} - yx^n = -\sum_{k=1}^{n} x\{yx^{n-k}\}x^{k-1}. \tag{60}$$

Hence,

$$x\xi(x, y)_{(1)} = \{yx(e^x - 1)^{-1}\} - yx(e^x - 1)^{-1}, \tag{61}$$

$$\xi(x, y)_{(1)} = -\sum_{n=1}^{\infty}\sum_{k=1}^{n}(B_n/n!)\{yx^{n-k}\}x^{k-1}, \tag{62}$$

$$z_{(2)} = \tfrac{1}{2}\{yx(e^x - 1)^{-1}\xi(x, y)_{(1)}\}. \tag{63}$$

We shall not proceed any further with the iteration process. It should be pointed out that Eq. (57) can be solved by one of the previous results, for instance, by Eq. (34),

$$x(e^x - 1)^{-1}\xi(x, y) = \int_0^1 dp[1 - a(p - 1) - \beta p]^{-1} \qquad (a = e^x - 1, \quad \beta = 1 - e^{-y}).$$

This gives the equation

$$\sum_{n=1}^{\infty} nz_{(n)} = \int_0^1 dp\{y[1 - a(p - 1) - \beta p]^{-1}\} \tag{64}$$

from which $z_n$ may be obtained by the expansion technique which we have previously used. The result is essentially of the same type as previous results in the present paper.

## APPENDIX: PROOF OF THE IDENTITIES (5)–(8)

Because of the linearity property (1d) in the definition of the curly bracket operator, it is sufficient to prove the identities for the case that $f$, $g$, and $h$ are monomials. Further, we note that the identities are trivially true when $g$ is of degree zero, i.e., $G = \{g\} = 0$. Thus we let $f$, $g$, and $h$ be the monomials:

$$\begin{aligned} f &= x_a x_b \cdots x_c \quad (\text{DEGREE} > 0), \\ g &= x_i x_j \cdots x_k \quad (\text{DEGREE} > 0), \\ h &= x_p x_q \cdots x_r \quad (\text{DEGREE} > 0), \end{aligned}$$

or

$$h = h_0 \quad (\text{DEGREE ZERO}).$$

**1.** *Proof of Eq.* (5): $\{fg\} = \{Fg\}_F$: From Eq. (4), we have (as a matter of notation)

$$\begin{aligned} \{fg\} &= \{x_a x_b \cdots x_c x_i x_j \cdots x_k\} \\ &= [\cdots [F, x_i], \cdots, x_k] \\ &= \{Fx_i \cdots x_k\}_F = \{Fg\}_F. \end{aligned}$$

**2.** *Proof of Eq.* (6): $\{fG\} = [F, G]$: The equation is true when $G$ is of first degree, that is when $G = \{x_i\} = x_i$. In that case we have from Eq. (5)

$$\{fx_i\} = [F, x_i] = \{Fx_i\}_F.$$

The general proof goes by induction, assuming the equation to be true when $G$ is of $n$th degree. We put $G_{n+1} = [G_n, x_m]$ and get

$$\begin{aligned} \{fG_{n+1}\} &= \{fG_n x_m\} - \{fx_m G_n\} \\ &= \{fG_n\}x_m - x_m\{fG_n\} - [\{fx_m\}, G_n] \\ &= [F, G_n]x_m - x_m[F, G_n] - [[F, x_m], G_n] \\ &= [F, [G_n, x_m]] = [F, G_{n+1}]. \end{aligned}$$

Q.E.D.

3. *Proof of Eq. (7)*: $\{fGh\} = \{fGh\}_G$: From Eqs. (6) and (1c),

$$\{fG\} = [F, G] = [\{f\}, G] = \{fG\}_G.$$

Thus, Eq. (7) is true when $h$ is of degree zero. Now, let $h$ be a monomial of positive degree. Then, from the last equation, we get by repeated commutator multiplication with the factors in $h$,

$$[\cdots [[\{fG\}, x_p], x_q], \cdots, x_r] = [\cdots [[\{fG\}_G, x_p], x_q], \cdots, x_r].$$

That is

$$\{fGh\} = \{fGh\}_G. \qquad \text{Q.E.D.}$$

4. *Proof of Eq. (8)*: $\{G_n h\} = n\{G_n h\}_{Gn}$: We first prove the equation $\{G_n\} = nG_n$, which is obviously true for $n = 1$. The proof is carried out by induction. We assume the equation to be true for $G_n$, and express $G_{n+1}$ as the commutator $G_{n+1} = [G_n, x_m]$. The result is

$$\begin{aligned}\{G_{n+1}\} &= \{G_n x_m\} - \{x_m G_n\} \\ &= \{G_n\}x_m - x_m\{G_n\} - \{x_m G_n\}_{G_n} \\ &= nG_n x_m - x_m nG_n - x_m G_n + G_n x_m \\ &= (n+1)[G_n, x_m] = (n+1)G_{n+1}.\end{aligned}$$

Thus, Eq. (8) is true when $h$ is of degree zero. The generalization to the case that $h$ is a monomial of positive degree is obvious in light of the technique used in the proof of Eq. (7) above.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 5 MAY 1968

# Two- and Three-Particle Coulomb Systems*

GERALD L. NUTT†
*Department of Physics, University of Pittsburgh, Pittsburgh, Pennsylvania*

(Received 2 August 1967)

The Coulomb $T$ matrix is examined in a two-body and a three-body Hilbert space. The purpose is to demonstrate how this function may be used to construct the Faddeev three-body kernel. In the process the apparent paradox of the indefiniteness of the Coulomb $T$ matrix on the energy shell is resolved. The question of off-shell unitarity is discussed.

## I. INTRODUCTION

Recent work[1] on the three-particle problem has used the off-energy-shell scattering amplitudes of the two-body subsystems in the construction of the resolvent for the three-particle system. It is reasonable, therefore, to expect that the Coulomb Green's function can supply the necessary two-body information for the solution of atomic three-body problems.

It is easy to obtain the Coulomb scattering amplitude from the Coulomb Green's function in all the necessary generality. Unfortunately, the Coulomb scattering amplitude is not a very simple function in a two-body Hilbert space, with the result that application in a three-body Hilbert space involves some mathematical difficulty. The primary problem associated with the off-energy-shell Coulomb $T$ matrix is that, in the momentum representation, it has a regular singularity of imaginary exponent lying squarely on the energy shell, thus making it impossible to define for "physical" scattering.[2]

We can illustrate these comments by solving

$$G_0 T = GV$$

for the Coulomb $T$ matrix, where $G_0$ is the free-particle Green's function, $T$ the desired $T$ matrix, $G$ is the Coulomb Green's function, and $V$ the Coulomb potential.[3] For an attractive potential the result is

$$\begin{aligned}&\langle \mathbf{k}_2|\, T(k^2)\,|\mathbf{k}_1\rangle \\ &\quad = -\frac{e^2}{2\pi^2}\frac{1}{|\mathbf{k}_2 - \mathbf{k}_1|^2} \\ &\qquad \times \left[1 - \frac{4i\nu}{e^{2\pi\nu} - 1}\int_{C_0} dt\, t^{-i\nu}\frac{1}{\epsilon(1-t)^2 - 4t}\right], \quad (1)\end{aligned}$$

where

$$\nu = \frac{me^2}{k^2}, \quad k^2 = 2mE, \quad \epsilon = \frac{(k^2 - k_1^2)(k^2 - k_2^2)}{k^2|\mathbf{k}_1 - \mathbf{k}_2|^2}.$$

$E$ is the energy appearing in the Green's function.

* This research was sponsored by the Office of Naval Research, Contract No. Nonr-624(16).

† Present address: Royal Military College, Kingston, Ontario.

[1] For example, L. D. Faddeev, Zh. Ehsp. Teor. Phys. **39**, 1459 (1960) [Sov. Phys.—JETP **12**, 1014 (1961)].

[2] W. F. Ford, Phys. Rev. **133**, B1616 (1964); J. Math. Phys. **7**, 626 (1966).

[3] L. Hostler, J. Math. Phys. **5**, 591 (1964); J. Schwinger, J. Math. Phys. **5**, 1606 (1964).

3. *Proof of Eq. (7)*: $\{fGh\} = \{fGh\}_G$: From Eqs. (6) and (1c),

$$\{fG\} = [F, G] = [\{f\}, G] = \{fG\}_G.$$

Thus, Eq. (7) is true when $h$ is of degree zero. Now, let $h$ be a monomial of positive degree. Then, from the last equation, we get by repeated commutator multiplication with the factors in $h$,

$$[\cdots [[\{fG\}, x_p], x_q], \cdots, x_r] = [\cdots [[\{fG\}_G, x_p], x_q], \cdots, x_r].$$

That is

$$\{fGh\} = \{fGh\}_G. \qquad \text{Q.E.D.}$$

4. *Proof of Eq. (8)*: $\{G_n h\} = n\{G_n h\}_{Gn}$: We first prove the equation $\{G_n\} = nG_n$, which is obviously true for $n = 1$. The proof is carried out by induction. We assume the equation to be true for $G_n$, and express $G_{n+1}$ as the commutator $G_{n+1} = [G_n, x_m]$. The result is

$$\begin{aligned}\{G_{n+1}\} &= \{G_n x_m\} - \{x_m G_n\} \\ &= \{G_n\}x_m - x_m\{G_n\} - \{x_m G_n\}_{G_n} \\ &= nG_n x_m - x_m nG_n - x_m G_n + G_n x_m \\ &= (n+1)[G_n, x_m] = (n+1)G_{n+1}.\end{aligned}$$

Thus, Eq. (8) is true when $h$ is of degree zero. The generalization to the case that $h$ is a monomial of positive degree is obvious in light of the technique used in the proof of Eq. (7) above.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 5 MAY 1968

# Two- and Three-Particle Coulomb Systems*

GERALD L. NUTT†
*Department of Physics, University of Pittsburgh, Pittsburgh, Pennsylvania*

(Received 2 August 1967)

The Coulomb $T$ matrix is examined in a two-body and a three-body Hilbert space. The purpose is to demonstrate how this function may be used to construct the Faddeev three-body kernel. In the process the apparent paradox of the indefiniteness of the Coulomb $T$ matrix on the energy shell is resolved. The question of off-shell unitarity is discussed.

## I. INTRODUCTION

Recent work[1] on the three-particle problem has used the off-energy-shell scattering amplitudes of the two-body subsystems in the construction of the resolvent for the three-particle system. It is reasonable, therefore, to expect that the Coulomb Green's function can supply the necessary two-body information for the solution of atomic three-body problems.

It is easy to obtain the Coulomb scattering amplitude from the Coulomb Green's function in all the necessary generality. Unfortunately, the Coulomb scattering amplitude is not a very simple function in a two-body Hilbert space, with the result that application in a three-body Hilbert space involves some mathematical difficulty. The primary problem associated with the off-energy-shell Coulomb $T$ matrix is that, in the momentum representation, it has a regular singularity of imaginary exponent lying squarely on the energy shell, thus making it impossible to define for "physical" scattering.[2]

We can illustrate these comments by solving

$$G_0 T = GV$$

for the Coulomb $T$ matrix, where $G_0$ is the free-particle Green's function, $T$ the desired $T$ matrix, $G$ is the Coulomb Green's function, and $V$ the Coulomb potential.[3] For an attractive potential the result is

$$\begin{aligned}&\langle \mathbf{k}_2|\, T(k^2)\,|\mathbf{k}_1\rangle \\ &\quad = -\frac{e^2}{2\pi^2}\frac{1}{|\mathbf{k}_2 - \mathbf{k}_1|^2} \\ &\qquad \times \left[1 - \frac{4i\nu}{e^{2\pi\nu} - 1}\int_{C_0} dt\, t^{-i\nu}\frac{1}{\epsilon(1-t)^2 - 4t}\right], \quad (1)\end{aligned}$$

where

$$\nu = \frac{me^2}{k^2}, \quad k^2 = 2mE, \quad \epsilon = \frac{(k^2 - k_1^2)(k^2 - k_2^2)}{k^2|\mathbf{k}_1 - \mathbf{k}_2|^2}.$$

$E$ is the energy appearing in the Green's function.

* This research was sponsored by the Office of Naval Research, Contract No. Nonr-624(16).

† Present address: Royal Military College, Kingston, Ontario.

[1] For example, L. D. Faddeev, Zh. Ehsp. Teor. Phys. **39**, 1459 (1960) [Sov. Phys.—JETP **12**, 1014 (1961)].

[2] W. F. Ford, Phys. Rev. **133**, B1616 (1964); J. Math. Phys. **7**, 626 (1966).

[3] L. Hostler, J. Math. Phys. **5**, 591 (1964); J. Schwinger, J. Math. Phys. **5**, 1606 (1964).

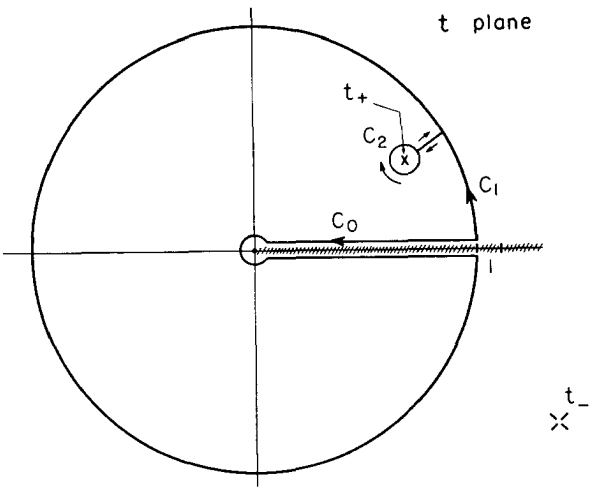



FIG. 1. Analytic configuration of the integrand in Eq. (1).

The integration contour $C_0$ starts at $t = 1$, slightly above the real axis. It moves to the origin and around the origin once in the positive sense. It finally moves to $t = 1$, slightly below the real axis. $C_0$ can be deformed as shown in Fig. 1, into the unit circle excluding the one pole of the integrand lying within the unit circle.

The integrand in Eq. (1) can be factored into two simple poles, one lying inside the unit circle and the other outside:

$$\langle \mathbf{k}_2 | \, T(k^2) \, | \mathbf{k}_1 \rangle = -\frac{e^2}{2\pi^2} \frac{1}{|\mathbf{k}_2 - \mathbf{k}_1|^2} \times \left[ 1 - \frac{4i\nu}{e^{2\pi\nu} - 1} \frac{1}{\epsilon} \int_{C_0} dt \frac{t^{-i\nu}}{(t - t_+)(t - t_-)} \right], \quad (2)$$

where $t_+ = t_-^{-1} = [(1 - \epsilon)^{\frac{1}{2}} - 1]/[(1 + \epsilon)^{\frac{1}{2}} + 1]$.

If we take $t_+$ to lie on the principal sheet with $\arg t_+ = \theta$, $t_-$ lies on the lower sheet with $\arg t_- = -\theta$. If we denote by $C_1$ the contour taken around the unit circle in the positive sense, and by $C_2$ the contour about $t_+$ in the negative sense, we have by Cauchy's theorem

$$\left( \int_{C_0} dt - \int_{C_1} dt - \int_{C_2} dt \right) \frac{t^{-i\nu}}{(t - t_+)(t - t_-)} = 0. \quad (3)$$

The energy-shell condition is given by

$$k_1^2 = k^2 = k_2^2, \quad (4)$$

for which $\epsilon$ and therefore $t_+$ vanishes. The energy shell can be approached in two ways. If $t_+$ lies on the principal sheet in the $t$ plane, the value of the Coulomb $T$ matrix approaches the value at the energy shell discontinuously as $t_+$ squeezes out of the contour $C_2$ and moves to the branch point. If, on the other hand, $t_+$ does not lie on the principal sheet, there is no contribution at all from $C_2$ and the value at the energy shell is approached continuously. On shell, then, we take only contributions from $C_1$.

For small $\epsilon$ the integrand in Eq. (1) can be expanded in a series, uniformly convergent on the unit circle. Thus the term in the square bracket in Eq. (1) becomes

$$1 + \frac{4i\nu}{e^{2\pi\nu} - 1} \int_{C_1} dt \frac{t^{-i\nu-1}}{4} \left[ 1 - \frac{\epsilon(1 - t)^2}{4t} \right]^{-1} = 1 + \frac{i\nu}{e^{2\pi\nu} - 1} \int_{e^{0i}}^{e^{2\pi i}} dt \, t^{-i\nu-1} \times [1 + \tfrac{1}{4}\epsilon(t - 2 + t^{-1}) + O(\epsilon^2)] = \tfrac{1}{2}\epsilon(1 + \nu^2)^{-1} + O(\epsilon^2). \quad (5)$$

We have the result that the value of the Coulomb $T$ matrix in the momentum representation is zero on the energy shell. This result is quite useful in the three-particle Hilbert space but should not be interpreted as a physical result in the two-particle Hilbert space. This apparent paradox is resolved by observing that momentum eigenstates are not acceptable asymptotic scattering states for the Coulomb potential. A brief look at time-dependent scattering theory will make this clear.

In the remainder of this paper we shall show how the physical scattering amplitude is obtained from formal scattering theory using the momentum representative of $T$ given by Eq. (1). This will be followed by a derivation of the off-shell unitarity condition for $T$. Finally, we shall examine the series for the amplitude for elastic scattering of a particle by a bound two-particle system developed in terms of the Faddeev three-body kernel. Taking a representative term of this series we use the results of the off-shell unitarity condition to express this term in the proper asymptotic representation.

## II. COULOMB SCATTERING

According to formal scattering theory, the physical scattering amplitude is obtained by letting the energy variable in the $T$ matrix approach the real axis from above. In our case, where the $T$ matrix is given by Eq. (1), this would correspond to giving $k^2$ a vanishingly small imaginary part. From this point of view we should be able to obtain the physical scattering amplitude from Eq. (1) by expanding $\langle \mathbf{k}_1 | \, T(k^2) \, | \mathbf{k}_2 \rangle$ about the energy shell and taking the limit as $k_1^2$ and $k_2^2$ approach $k^2$ continuously.

Unfortunately, this approach is not possible in the case of the Coulomb $T$ matrix because as we have seen, it does not approach its value on the energy

shell continuously. For $\epsilon$, small but not zero, we find that the path of integration over the parameter $t$ along $C_0$ can be deformed into an integral along the contours $C_1$ and $C_2$. The contribution from $C_2$ is just the residue at $t = t_+$. For $\epsilon$ sufficiently small we have

$$t_+ \approx \tfrac{1}{4}\epsilon, \tag{6}$$

and

$$\langle \mathbf{k}_2 | \, T(k^2) \, | \mathbf{k}_1 \rangle = -\frac{e^2}{2\pi^2} \frac{1}{|\mathbf{k}_2 - \mathbf{k}_1|^2} \frac{\epsilon}{2} \frac{1}{1+\nu^2} - \frac{e^2}{\pi} \nu \frac{1}{|\mathbf{k}_1 - \mathbf{k}_2|^2} \times \frac{1}{e^{2\pi\nu} - 1} \left[ \frac{(k^2 - k_1^2)(k^2 - k_2^2)}{4k^2 |\mathbf{k}_1 - \mathbf{k}_2|^2} \right]^{-i\nu} + O(\epsilon^2). \tag{7}$$

Clearly, the second term in this equation approaches no limit continuously as $k_1^2$ and $k_2^2$ are made to approach $k^2$ for $\nu$ real.[2]

This difficulty can be easily straightened out by taking a look at time-dependent scattering theory. We will not develop all the results quoted below but we will refer the reader to a standard treatment of time-dependent scattering theory whenever the result is not obvious.

The amplitude for scattering of a particle from an asymptotic $\delta$-function-normalized initial state to an asymptotic $\delta$-function-normalized final state is given by

$$S_{fi} = \lim_{t \to \infty} \langle \Phi_f | \, U(t, 0) U(0, -t) \, | \Phi_i \rangle, \tag{8}$$

where the asymptotic state vectors are time-independent, and $U(t, 0)$ and $U(0, -t)$ are unitary operators which in the limit approach the (not necessarily unitary) Møller operators. If we express the asymptotic state as a wave packet composed of momentum components clustering about the value $\mathbf{k}_{1,2}$,

$$|\Phi_{i,f}\rangle = \int d\mathbf{k} \, \langle \mathbf{k} \mid \Phi_{i,f} \rangle \, |\mathbf{k}\rangle, \tag{9}$$

we can usually replace the asymptotic states $|\Phi_i\rangle$ and $|\Phi_f\rangle$ by momentum eigenstates. The physical scattering amplitude is then accurately given by

$$S_{\mathbf{k}_2, \mathbf{k}_1} = \lim_{t \to \infty} \langle \mathbf{k}_2 | \, U(t, 0) U(0, -t) \, | \mathbf{k}_1 \rangle. \tag{10}$$

In order that Eq. (10) may be substituted for Eq. (8) we require that $\langle \mathbf{k} \mid \Phi_{i,f} \rangle$ be (1) sharply peaked about the values $\mathbf{k}_1$ and $\mathbf{k}_2$, respectively, and (2) $\langle \mathbf{k} \mid \Phi_{i,f} \rangle$ be well-behaved.[4] The requirement that $\langle \mathbf{k} \mid \Phi_{i,f} \rangle$ be well-behaved arises from the requirement that if (10) is to replace (8) it is necessary that terms like

$$\lim_{t \to \infty} \int d\mathbf{k} \exp [\pm i(E_b - E_k)t] \langle \psi_b \mid \mathbf{k} \rangle \langle \mathbf{k} \mid \Phi_{i,f} \rangle, \quad E_b - E_k < 0 \tag{11}$$

vanish by the Riemann–Lebesgue lemma. The states $|\psi_b\rangle$ are bound eigenstates of the system. There are similar terms arising from the continuum states which must also vanish but it is sufficient for the purpose of this argument to consider only the terms like (11).

If we are considering a Coulomb-scattering problem we surely may expect to find asymptotic states which are strongly peaked about some momentum eigenstate. On the other hand, the momentum representative of the asymptotic Coulomb state is not at all well-behaved.

A glance at the asymptotic Coulomb modified plane-wave states will make this clear. These states are obtained from a result of Guth and Mullin[5] which gives the momentum representation of the Coulomb scattering state:

$$\begin{aligned} \langle \mathbf{k}' \mid \psi_k \rangle_c^+ &= (2\pi)^{-3} \int d\mathbf{r} \, e^{-i\mathbf{k}'\cdot\mathbf{r}} e^{i\mathbf{k}\cdot\mathbf{r}} \, F_1(\mp i\nu, 1, i(kr - \mathbf{k}\cdot\mathbf{r})) \\ &= \lim_{\eta \to 0} \delta(\mathbf{k}' - \mathbf{k})(1 + i\nu)\{[k'^2 + (\eta - ik)^2]/\eta^2\}^{i\nu} \\ &\quad - (2\pi^2)^{-1}\{[k'^2 + (\eta - ik)^2]^{1-i\nu}\}^{-1} \\ &\quad \times \{2\nu k/[\eta^2 + (\mathbf{k}' - \mathbf{k})^2]^{1+i\nu}\}. \end{aligned} \tag{12}$$

The last term is easily interpreted as a diverging spherical wave times the physical Coulomb scattering amplitude. The first term is the desired Coulomb modified plane wave for which the asymptotic region is defined by $k' \approx k$. Approaching no definite limit asymptotically and with an infinite phase, the asymptotic Coulomb state does not qualify as a test function which would allow (11) to vanish by the Riemann–Lebesgue lemma.

In order for these asymptotic Coulomb states to be suitable for use in Eq. (8), they must have $\delta$-function normalization. The normalization constants are easily determined to within a phase factor by

$$NN^* \int d\mathbf{k}' \delta(\mathbf{k}' - \mathbf{k}_1)\delta(\mathbf{k}' - \mathbf{k}_2)(1 + i\nu)(1 - i\nu) \times \{[k'^2 + (\eta - ik_1)^2]/\eta^2\}^{i\nu}\{[k'^2 + (\eta - ik_2)]/\eta^2\}^{-i\nu} = \delta(\mathbf{k}_1 - \mathbf{k}_2), \tag{13}$$

for which a possible solution is

$$N = (1 + i\nu)^{-1}. \tag{14}$$

Thus, the asymptotic Coulomb-scattering states appearing in Eq. (8) are

$$\langle \mathbf{k}_1' | \Phi_{\mathbf{k}_1} \rangle^+ = \delta(\mathbf{k}_1' - \mathbf{k}_1)\{[k_1'^2 + (\eta - ik_1)^2]/\eta^2\}^{i\nu},$$
$$\langle \mathbf{k}_2' | \Phi_{\mathbf{k}_2} \rangle^- = \delta(\mathbf{k}_2 - \mathbf{k}_2')\{[k_2'^2 + (\eta - ik_2)^2]/\eta^2\}^{-i\nu}. \tag{15}$$

[4] See S. S. Schweber, *An Introduction to Relativistic Quantum Field Theory* (Row, Peterson, Evanston, Illinois, 1961), especially Sec. 11c.

[5] E. Guth and C. J. Mullin, Phys. Rev. **83**, 667 (1951).

The $\pm$ superscripts are to signify that these states correspond respectively to a diverging or converging spherical part in the total scattering eigenstate.

We now take the momentum representative of the Coulomb-scattering amplitude in the region of the energy shell and allow it to approach the unitary axis from above. From Eq. (7),

$$\begin{aligned}\langle \mathbf{k}_2' |\, U(\infty,0)U(0,-\infty)\,|\mathbf{k}_1'\rangle &= -2\pi i\delta\left(\frac{k_1'^2}{2m}-\frac{k_2'^2}{2m}\right)\left(-\frac{e^2}{2\pi}\right)\frac{1}{|\mathbf{k}_1'-\mathbf{k}_2'|^2}\\ &\times\left[\frac{(k_2'^2+(\eta-ik)^2)(k_1'^2+(\eta-ik)^2)}{4k^2|\mathbf{k}_1'-\mathbf{k}_2'|^2}\right]^{-i\nu}\\ &\times e^{-\pi\nu}|\Gamma(1+i\nu)|^2+O(\epsilon). \qquad (16)\end{aligned}$$

Thus, from Eqs. (8), (15), and (16),

$$\begin{aligned}S_{\mathbf{k}_2,\mathbf{k}_1} &= \int d\mathbf{k}_1'\,d\mathbf{k}_2'\,\langle\Phi_{\mathbf{k}_2}\,|\,\mathbf{k}_2'\rangle\\ &\times\langle\mathbf{k}_2'|\,U(\infty,0)U(0,-\infty)\,|\mathbf{k}_1'\rangle\langle\mathbf{k}_1'\,|\,\Phi_{\mathbf{k}_1}\rangle\\ &= -2\pi i\delta\left(\frac{k_1^2}{2m}-\frac{k_2^2}{2m}\right)\left\{-\frac{e^2}{2\pi^2}\frac{1}{|\mathbf{k}_1-\mathbf{k}_2|^{2+2i\nu}}\right\}\\ &\times e^{2i\nu\ln 2k/\eta^2}|\Gamma(1+i\nu)|^2\,e^{-\pi\nu}. \qquad (17)\end{aligned}$$

The term in the curly brackets can be compared with the Coulomb-scattering amplitude by multiplying by $-4\pi^2 m$ and we get exact agreement. The infinite logarithmic phase factor can be thought of as belonging to the normalization of the asymptotic Coulomb states; hence it should be absorbed in $N$. The remaining factors are recognized as arising from the normalization of the complete Coulomb-scattering states. If Eq. (12) had the conventional normalization they would not appear.

By displacing the energy from the real axis with a positive imaginary part, and taking the Coulomb $T$ matrix in the proper asymptotic representation, we have shown that the Coulomb $T$ matrix yields the correct Coulomb-scattering amplitude.

## III. UNITARITY

The Coulomb $T$ matrix given by Eq. (1) can be derived from the following familiar equations

$$\begin{aligned}T(E) &= V+VG_0(E)T(E)\\ &= V+VG(E)V. \qquad (18)\end{aligned}$$

$G(E)$, the Coulomb Green's function, satisfies

$$\begin{aligned}G(E) &= G_0(E)+G_0(E)VG(E)\\ &= G_0(E)+G(E)VG_0(E). \qquad (19)\end{aligned}$$

Using Eqs. (18) and (19) we can derive an off-shell unitarity condition for $T(E)$:

$$\begin{aligned}&T(E+i\eta)-T(E-i\eta)\\ &\quad= V(G(E+i\eta)-G(E-i\eta))V\\ &\quad= 2i\eta V(1+G(E+i\eta))G_0(E+i\eta)\\ &\qquad\times G_0(E-i\eta)(1+VG(E-i\eta))V\\ &\quad= T(E+i\eta)(G_0(E+i\eta)-G_0(E-i\eta))T(E-i\eta), \qquad (20)\end{aligned}$$

where

$$\begin{aligned}&\langle\mathbf{k}'|\,G_0(E+i\eta)-G_0(E-i\eta)\,|\mathbf{k}''\rangle\\ &\qquad= -2\pi i\delta[E-(k'^2/2m)]\delta(\mathbf{k}'-\mathbf{k}''). \qquad (21)\end{aligned}$$

The limit implied by $E\pm i\eta$ means that $\arg k = 0$ or $\pi$, respectively. Combining Eqs. (20) and (21), and taking matrix elements we have the result:

$$\begin{aligned}&\langle\mathbf{k}_2|\,T(k^2)\,|\mathbf{k}_1\rangle-\langle\mathbf{k}_2|\,T((-k)^2)\,|\mathbf{k}_1\rangle\\ &\quad= -2\pi i\int d\mathbf{k}'\,\langle\mathbf{k}_2|\,T(k^2)\,|\mathbf{k}'\rangle\,\delta[(k^2/2m)-(k'^2/2m)]\\ &\qquad\times\langle\mathbf{k}'|\,T((-k)^2)\,|\mathbf{k}_1\rangle. \qquad (22)\end{aligned}$$

If we substitute Eq. (1) for $\langle\mathbf{k}_2|\,T(k^2)\,|\mathbf{k}_1\rangle$, we see that the $\delta$ function on the right-hand side of Eq. (22) requires that

$$\frac{(k^2-k_1^2)(k^2-k'^2)}{k^2|\mathbf{k}_1-\mathbf{k}'|^2}=\frac{(k^2-k'^2)(k^2-k_2^2)}{k^2|\mathbf{k}'-\mathbf{k}_2|^2}=0. \qquad (23)$$

Under these conditions we find that

$$\begin{aligned}\langle\mathbf{k}_2|\,T(k_1^2)\,|\mathbf{k}_1\rangle &= -\frac{e^2}{2\pi}\frac{1}{|\mathbf{k}_1-\mathbf{k}_2|^2}\\ &\times\left[1+\frac{i\nu}{e^{2\pi\nu}-1}\int_{e^{0i}}^{e^{2\pi i}}dt\,t^{-i\nu-1}\right]=0, \qquad (24)\end{aligned}$$

so that the right-hand side of Eq. (22) vanishes identically. Thus, from the form of the off-shell scattering amplitude in Eq. (2),

$$\begin{aligned}&\langle\mathbf{k}_2|\,T(k^2)\,|\mathbf{k}_1\rangle-\langle\mathbf{k}_2|\,T((-k)^2)\,|\mathbf{k}_1\rangle\\ &=\frac{e^2}{2\pi^2}\frac{1}{|\mathbf{k}_1-\mathbf{k}_2|^2}\frac{4i\nu}{e^{2\pi\nu}-1}\frac{1}{\epsilon}\int_{C_0}dt\,t^{-i\nu}\frac{1}{(t-t_+)(t-t_-)}\\ &+\frac{4i\nu}{e^{-2\pi\nu}-1}\frac{1}{\epsilon}\int_{C_0}dt\,t^{i\nu}\cdot\frac{1}{(t-t_+)(t-t_-)}=0. \qquad (25)\end{aligned}$$

$C_0$ can now be deformed into the two contours $C_1$ and $C_2$. The contribution to the $t$ integral from $C_1$ can be evaluated by expanding the integrand in a series uniformly convergent within the unit circle as long as

$|t_+| < 1$:

$$-\frac{e^2}{2\pi^2}\frac{4i\nu}{|\mathbf{k}_1-\mathbf{k}_2|^2}\frac{t_+}{t_+^2+1}\frac{1}{\epsilon}\int_{e^{0i}}^{e^{2\pi i}}dt \times\left[\frac{t^{-i\nu-1}}{e^{2\pi\nu}-1}+\frac{t^{i\nu-1}}{e^{-2\pi\nu}-1}\right]\left[1+\frac{t_+}{t_+^2+1}\left(t+\frac{1}{t}\right)+\cdots+\left(\frac{t_+}{t_+^2+1}\right)^n\left(t+\frac{1}{t}\right)^n+\cdots\right].$$

In any term of the series, say, the one with $(t+t^{-1})^n$, there is a symmetry between the positive and negative powers of $t$. That is, for every power $t^m$ there occurs a power $t^{-m}$ with the same coefficient. For every term proportional to

$$\frac{4i\nu}{e^{2\pi\nu}-1}\int_{e^{i0}}^{e^{2\pi i}}dt\,t^{-i\nu-1+m}=\frac{4i\nu_1}{-i\nu+m},$$

there occurs one proportional to

$$\frac{4i\nu}{e^{-2\pi\nu}-1}\int_{e^{i0}}^{e^{2\pi i}}dt\,t^{i\nu-1-m}=\frac{4i\nu}{i\nu-m}$$

for $m$, a positive or a negative integer. Thus the contribution from the contour $C_1$ cancels term by term. We are left with the result

$$\frac{e^2}{2\pi^2}\frac{1}{|\mathbf{k}_1-\mathbf{k}_2|^2}\left[\frac{4i\nu}{e^{2\pi\nu}-1}\frac{1}{\epsilon}\int_{C_2}dt\frac{t^{-i\nu}}{(t-t_+)(t-t_-)}+\frac{4i\nu}{e^{-2\pi\nu}-1}\frac{1}{\epsilon}\int_{C_2}dt\frac{t^{i\nu}}{(t-t_+)(t-t_-)}\right]=0. \quad (26)$$

Equation (25) tells us that there is no discontinuity in the momentum representative of the off-energy-shell Coulomb $T$ matrix along the unitary axis. In particular, Eq. (26) shows that there is no discontinuity arising from that contribution coming from the contour $C_2$.

It is difficult to see how Eq. (26) can be satisfied identically for all points along the unitary axis unless the contribution from $C_2$ vanishes identically. This occurs if in some way we can continuously deform the contour $C_0$ into the contour $C_1$ without passing over a pole lying inside the unit circle and on the principal sheet. If $k^2$ approaches the unitary axis in such a way that the pole inside the unit circle moves off the principal sheet, then $C_0$ can be deformed into $C_1$. We shall see an example of this in the next section.

## IV. THREE-PARTICLE SYSTEMS

For a three-particle system, the amplitude for the elastic scattering of particle 2 by a bound subsystem composed of particles 1 and 3 can be iterated in terms of the Faddeev kernel as follows[6]:

$$T=T_1+T_3+T_1G_0T_3+T_3G_0T_1+\cdots. \quad (27)$$

$G_0$ is the free-particle Green's function operator for three particles, and $T_1$ is the two particle $T$ matrix for scattering of particles 2 and 3, etc. If these particles interact via the Coulomb potential, the momentum representative of Eq. (27) can be written down, to any order of iteration, in terms of Eq. (1). Consider, for example, the first term:

$$\langle\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3|\,T_1\,|\mathbf{k}_1'\mathbf{k}_2'\mathbf{k}_3'\rangle=\langle\mathbf{k}_{23}|\,T_1(k_1^2)\,|\mathbf{k}_{23}'\rangle\,\delta(\mathbf{p}_1-\mathbf{p}_1'). \quad (28)$$

The notation is explained as follows: the *vector* $\mathbf{k}_i$ is the momentum of particle $i$; $\mathbf{k}_{23}$ is the relative momentum of particles 2 and 3; $\mathbf{p}_1$ is the momentum of particle 1 relative to the center of mass of particles 2 and 3; $k_1^2$ is not $\mathbf{k}_1\cdot\mathbf{k}_1$ but rather $2\mu_{23}E-(\mu_{23}/\mu_1)p_1^2$,

$$\mu_{23}=m_2m_3/(m_2+m_3),\qquad \mu_1=m_1(m_2+m_3)/(m_1+m_2+m_3), \quad (29)$$

and $E$ is the energy of the three-particle system.

When dealing with three particles we have two additional manifolds which describe the system just as well as $(\mathbf{k}_{23},\mathbf{p}_1)$. They are obtained by permutation of the particles. We shall signify them by the obvious notation $(\mathbf{k}_{12},\mathbf{p}_3)$ and $(\mathbf{k}_{31},\mathbf{p}_2)$.

Unfortunately, the momentum representative of Eq. (28) is not very useful. Again we must take inner products of Eq. (18) with the correct asymptotic scattering states in order to obtain a physically interesting three-body scattering amplitude. If we assume that the (1, 3) system is bound in the Bohr energy level with 1 for the principal quantum number in both the initial and final states, the asymptotic wavefunctions are

$$\langle\mathbf{k}_{31}\mathbf{p}_2'\,|\,\mathbf{p}_2\psi_{100}\rangle=(8\lambda^5)^{\frac{1}{2}}\delta(\mathbf{p}_2-\mathbf{p}_2')/\pi(k_{31}^2+\lambda^2)^2, \quad (30)$$

where $\lambda^2=\mu_{13}e^2$ is the reciprocal Bohr radius.

For the first term in the three-particle elastic-scattering amplitude we are lead to the integral

$$\frac{8\lambda^5}{\pi^2}\int\frac{d\mathbf{k}_{31}}{(k_{31}^2+\lambda^2)^2}\int\frac{d\mathbf{k}_{31}'}{(k_{31}'^2+\lambda^2)^2}\times\delta(\mathbf{p}_1-\mathbf{p}_1')\langle\mathbf{k}_{23}|\,T_1(k_1^2)\,|\mathbf{k}_{23}'\rangle, \quad (31)$$

the trivial integrals over $\mathbf{p}_2$ and $\mathbf{p}_2'$ having been carried out. The integral over $\mathbf{k}_{31}'$ also becomes trivial if we use the identity

$$\mathbf{p}_1-\mathbf{p}_1'=\mathbf{k}_{31}-\mathbf{k}_{31}'-[m_1/(m_3+m_1)](\mathbf{p}_2-\mathbf{p}_2'). \quad (32)$$

[6] H. Ekstein, Phys. Rev. **101**, 880 (1956); C. Lovelace, Phys. Rev. **135**, B1225 (1964).

The next step is to express the Coulomb $T$ matrix given by Eq. (1) in terms of the three-body variables, and in particular, the variables in the $(\mathbf{k}_{31}, \mathbf{p}_2)$ manifold. We have the identity

$$\begin{aligned}\mathbf{k}_{23} - \mathbf{k}'_{23} &= -\mathbf{p}_2 + \mathbf{p}'_2 - [m_2/(m_2 + m_3)](\mathbf{p}_1 - \mathbf{p}'_1) \\ &= -\mathbf{p}_2 + \mathbf{p}'_2, \end{aligned} \tag{33}$$

the last equality resulting from the $\delta$ function in (31). Thus,

$$\langle \mathbf{k}_{23}| T_1(k_1^2) |\mathbf{k}'_{23}\rangle = -\frac{e^2}{2\pi^2}\frac{1}{|\mathbf{p}_2 - \mathbf{p}'_2|^2} \times \left[1 - \frac{4i\nu_1}{e^{2\pi\nu_1} - 1}\int_{C_0} dt \frac{t^{-i\nu_1}}{\epsilon_1(1-t)^2 - 4t}\right], \tag{34}$$

where

$$\nu_1^2 = \frac{\mu_{23}^2 e^4}{k_1^2},$$

$$k_1^2 = 2\mu_{23}E - \frac{\mu_{23}}{\mu_1}\left(\mathbf{k}_{31} - \frac{m_1}{m_3 + m_1}\mathbf{p}_2\right)^2, \tag{35}$$

$$E = -(\lambda^2/2\mu_{31}) + (p_2^2/2\mu_2), \tag{36}$$

and where we have used the identity

$$\mathbf{p}_1 = \mathbf{k}_{31} - [m_1/(m_3 + m_1)]\mathbf{p}_2$$

in the equation for $k_1^2$.

$\epsilon_1$ is given by

$$\begin{aligned}\epsilon_1 &= \frac{(k_1^2 - k_{23}^2)(k_1^2 - k_{23}'^2)}{k_1^2 |\mathbf{p}_2 - \mathbf{p}'_2|} \\ &= \left(\frac{\mu_{23}}{\mu_{31}}\right)^2 \frac{(k_{31}^2 + \lambda^2)(k_{31}'^2 + \lambda^2)}{k_1^2 |\mathbf{p}_2 - \mathbf{p}'_2|^2}. \end{aligned} \tag{37}$$

In the second equality we have used the identity

$$\begin{aligned}E - (2\mu_1)^{-1}p_1^2 - (2\mu_{23})^{-1}k_{23}^2 &= E - (2\mu_2)^{-1}p_2^2 - (2\mu_{31})^{-1}k_{31}^2 \\ &= -(2\mu_{31})^{-1}(\lambda^2 + k_{31}^2). \end{aligned} \tag{38}$$

The integral (31) can now be put in the form

$$-\frac{e^2}{2\pi^2}\frac{8\lambda^5}{\pi^2}\int \frac{d\mathbf{k}_{31}}{(k_{31}^2 + \lambda^2)^2(k_{31}'^2 + \lambda^2)^2}\frac{1}{|\mathbf{p}_2 - \mathbf{p}'_2|^2} \times \left[1 - \frac{4i\nu_1}{e^{2\pi\nu_1} - 1}\int_{C_0} dt \frac{t^{-i\nu_1}}{\epsilon_1(1-t)^2 - 4t}\right], \tag{39}$$

where $\mathbf{k}'_{31} \to \mathbf{k}_{31} + [m_1/(m_3 + m_1)](\mathbf{p}_2 - \mathbf{p}'_2)$.

As shown in the Appendix, the integral over $k_{31}$ can be made into a contour integral by extending the range of integration from $0 < k_{31} < \infty$ to $-\infty < k_{31} < \infty$. By this device we convert the integral along the positive real axis into an integral around the various singularities of the integrand lying in the upper half-plane.

The singularities contributing to the integral are of the following possible types: (1) poles occurring in the asymptotic wavefunctions at $k_{31} = i\lambda$ and $k'_{31} = i\lambda$; (2) if the potential in $T_1$ is attractive, poles arise from each of the possible bound states of particles 2 and 3.

In case a branch point should occur at $k_1 = 0$, we would have to include an integral of the discontinuity along the cut. We saw that the unitarity condition gave us zero for this discontinuity. It is instructive, however, to see just how the integral representation of Eq. (2) eliminates the contribution to the $T$ matrix arising from the contour $C_2$ when $k_1$ lies on the unitary axis.

The pole in the $t$ integrand, from which we get the contour $C_2$, we have called $t_+$:

$$t_+ = t_-^{-1} = [(1 + \epsilon_1)^{\frac{1}{2}} - 1]/[(1 + \epsilon_1)^{\frac{1}{2}} + 1]. \tag{40}$$

Thus,

$$\begin{aligned}|t_+| &= \left\{\frac{[(1 + \epsilon_1)^{\frac{1}{2}} - 1][(1 + \epsilon_1)^{\frac{1}{2}*} - 1]}{[(1 + \epsilon_1)^{\frac{1}{2}} + 1][(1 + \epsilon_1)^{\frac{1}{2}*} + 1]}\right\}^{\frac{1}{2}} \\ &= \left[\frac{|(1 + \epsilon_1)^{\frac{1}{2}}|^2 + 1 - 2\,\mathrm{Re}\,(1 + \epsilon_1)^{\frac{1}{2}}}{|(1 + \epsilon_1)^{\frac{1}{2}}|^2 + 1 + 2\,\mathrm{Re}\,(1 + \epsilon_1)^{\frac{1}{2}}}\right]^{\frac{1}{2}} \end{aligned} \tag{41}$$

so that $|t_+| < 1$ for $\mathrm{Re}\,(1 + \epsilon_1)^{\frac{1}{2}} > 0$ and

$$\begin{aligned}&\arg t_+ \\ &= \arg[(1 + \epsilon_1)^{\frac{1}{2}} - 1] - \arg[(1 + \epsilon_1)^{\frac{1}{2}} + 1] \\ &= \sin^{-1}\frac{\mathrm{Im}\,(1 + \epsilon_1)^{\frac{1}{2}}}{(\{\mathrm{Re}\,[(1 + \epsilon_1)^{\frac{1}{2}} - 1]\}^2 + [\mathrm{Im}\,(1 + \epsilon_1)^{\frac{1}{2}}]^2)^{\frac{1}{2}}} \\ &\quad - \sin^{-1}\frac{\mathrm{Im}\,(1 + \epsilon_1)^{\frac{1}{2}}}{(\{\mathrm{Re}\,[(1 + \epsilon_1)^{\frac{1}{2}} + 1]\}^2 + [\mathrm{Im}\,(1 + \epsilon_1)^{\frac{1}{2}}]^2)^{\frac{1}{2}}}. \end{aligned} \tag{42}$$

Clearly,

$\arg t_+ > 0$,

$$\text{if } \begin{cases}\mathrm{Im}\,(1 + \epsilon_1)^{\frac{1}{2}} > 0 \quad \text{and} \quad \mathrm{Re}\,(1 + \epsilon_1)^{\frac{1}{2}} > 0 & \text{(43a)} \\ \mathrm{Im}\,(1 + \epsilon_1)^{\frac{1}{2}} < 0 \quad \text{and} \quad \mathrm{Re}\,(1 + \epsilon_1)^{\frac{1}{2}} < 0, & \text{(43b)}\end{cases}$$

$\arg t_+ < 0$,

$$\text{if } \begin{cases}\mathrm{Im}\,(1 + \epsilon_1)^{\frac{1}{2}} > 0 \quad \text{and} \quad \mathrm{Re}\,(1 + \epsilon_1)^{\frac{1}{2}} < 0 & \text{(43c)} \\ \mathrm{Im}\,(1 + \epsilon_1)^{\frac{1}{2}} < 0 \quad \text{and} \quad \mathrm{Re}\,(1 + \epsilon_1)^{\frac{1}{2}} > 0. & \text{(43d)}\end{cases}$$

We shall assume that when $E$ has an infinitesimal positive imaginary part, $t_+$ lies on the principal sheet of the $t$ Riemann surface. As $k_{31}$ moves to $+\infty$, we find that $k_1$ and therefore $\epsilon_1$ become infinite. $\arg k_1$ then equals $\frac{1}{2}\pi - \delta$, where $\delta$ is an infinitesimal positive angle. Under the condition that $k_{31}$ is infinite, Eq. (37) shows us that $(1 + \epsilon_1)^{\frac{1}{2}}$ is proportional to $k_1$. If we call the position $k_{31} = +\infty$ point 1, we see that

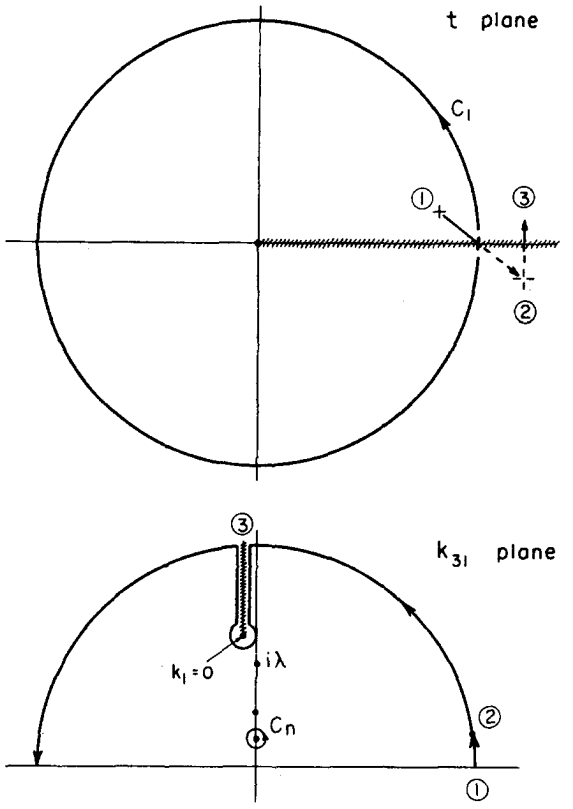



FIG. 2. Trajectory of $t_+$ as $k_{31}$ is continued into the upper half-plane. The trajectory of $t_-$ is obtained from $t_- = t_+^{-1}$.

conditions (43a) are met and point 1 maps into the point shown in Fig. 2.

As $k_{31}$ moves from point 1 along a semicircle in the upper half-plane, $\arg k_1$ quickly increases past $\frac{1}{2}\pi$, with the result that condition (43a) goes over into condition (43c). At this point $t_+$ and $t_-$ change places. We now have $t_+$ lying outside the unit circle and $t_-$ inside the unit circle. We denote this point by 2. As we move along the semicircle in the upper $k_{31}$ plane we come to a point at which $\arg k_1 = \pi + \delta$, which we call point 3. At point 3 we obtain conditions (43b) where $t_-$ moves once more onto the lower sheet but remains within the unit circle. As we move down the cut, around the branch point and up the other side of the cut, $t_-$ remains inside the unit circle without once crossing the contour $C_1$. We have taken the discontinuity across the unitary cut to be the difference between the $T$ matrix evaluated at a point on the *physical* sheet with $\arg k_1 = \delta$, and the $T$ matrix evaluated at a point just below on the *unphysical* sheet with $\arg k_1 = \pi + \delta$. This demonstrates how we are able to eliminate the contribution from $C_2$ when $k_1^2$ lies on the unitary axis in accordance with Eq. (26).

Turning to the pole singularities in the upper $k_{31}$ plane, we consider first those coming from $k_{31} = i\lambda$ and $k'_{31} = i\lambda$. From Eq. (37) we see that we are now interested in the region around $\epsilon_1 = 0$. Using Eq. (5) in (39) we get

$$-\frac{e^2}{4\pi^2}\frac{8\lambda^5}{\pi^2}\left(\frac{\mu_{23}}{\mu_{31}}\right)^2\frac{1}{|\mathbf{p}_2 - \mathbf{p}_2'|^4} \times \int \frac{d\mathbf{k}_{31}}{(k_{31}^2 + \lambda^2)(k_{31}'^2 + \lambda^2)}\frac{1}{k_1^2 + \lambda^2}. \quad (44)$$

Equations (32) and (35) allow us to integrate this expression by the Feynman method. One must be careful to take only the residues at $k_{31}^2 + \lambda^2 = 0$ and $k_{31}'^2 + \lambda^2 = 0$, but avoid the poles at $k_1^2 + \lambda^2 = 0$. The latter belong to the set of poles coming from the Coulomb bound states.

The poles at the bound states occur whenever $iv$ is a positive integer, say $n$. In this case the branch point in the $t$ integrand becomes a pole of order $n$. We deform the contour $C_0$ continuously into a small circle about the origin. Thus, we have the simple result for the $t$ integration:

$$\int_{C_0} dt\, t^{-n}\frac{1}{(t - t_+)(t - t_-)} = \frac{2\pi i}{(n-1)!}\left(\frac{\partial}{\partial t}\right)^{n-1}\frac{1}{(t - t_+)(t - t_-)}\bigg|_{t=0} = 2\pi i\frac{t_+^n - t_-^n}{t_+ - t_-}. \quad (45)$$

If, in addition, we use the representation

$$\frac{1}{e^{2\pi v_1} - 1} = \frac{e^{-\pi v_1}}{2\pi i}\left(-\frac{1}{iv_1} + \sum_{n=1}^{\infty}(-1)^n\frac{2iv_1}{v_1^2 + n^2}\right), \quad (46)$$

the contribution to (39) coming from the bound-state poles is

$$\frac{e^2}{4\pi^2}\frac{8\lambda^5}{\pi^2}\sum_{n=1}^{\infty}\int d\Omega_{31}\int_{C_n} dk_{31}\frac{k_{31}^2}{(k_{31}^2 + \lambda^2)^3(k_{31}'^2 + \lambda^2)^3} \times k_1^2\frac{8(iv_1)^2(-1)^n e^{-\pi v_1}}{v_1^2 + n^2}\frac{t_+^n - t_-^n}{t_+ - t_-} = \frac{16\lambda^5 e^2}{\pi^4}\sum_{n=1}^{\infty}\int d\Omega_{31}\int_{C_n} dk_{31}\frac{k_{31}^2}{(k_{31}^2 + \lambda^2)^3(k_{31}'^2 + \lambda^2)^3} \times \frac{\lambda^4}{n^4}\frac{1}{k_1^2 + (\lambda^2/n^2)}\frac{t_+^n - t_-^n}{t_+ - t_-}. \quad (47)$$

$C_n$ is taken to be a small circle in the $k_{31}$ plane enclosing the pole at $k_1^2 + \lambda^2/n^2 = 0$.

By taking the sum of (44) and (47) we are able to express (31) in terms of more or less elementary integrals. It does not appear to be practical to try to evaluate these integrals in general. However, they have

been done for particular systems where simplifying approximations can be made. The same methods can be used to simplify the higher-order terms in the iteration series (27); the extension is fairly obvious.

Coulomb effects in three-body systems have been studied by Schulman[7] where two of the particles interact via a Coulomb potential, but the remaining pair of interactions are of short range. This approach does not seem to allow itself to be generalized to systems of three charged particles.

It is these systems of three charged particles that we hope to be able to handle through the results of this paper. In particular we have studied systems such as electron–hydrogen scattering and positron–hydrogen scattering. In either case we can assign $m_1$ and $m_2$ the values of the electron mass, and $m_3$ the proton mass. In expressions such as Eqs. (32), (33), and (35) we can neglect terms of order $m_1/m_3$ and $m_2/m_3$ to a high degree of accuracy. The resulting integrals can be easily evaluated and can be expected to realistically represent the scattering amplitude within the limitations imposed by truncating the series (27).

Because the evaluation of the integrals (44) and (47) is most easily done when they are restricted to particular physical systems, we will reserve a discussion of these results for a more appropriate journal.

## V. CONCLUSION

In spite of the mathematical complexity of the momentum representation of the off-shell Coulomb $T$ matrix, we have shown how it is useful in calculating physical scattering amplitudes. For two-body scattering we obtain the physical amplitude by using the correct representation of the scattering states. The well-known on-shell singularities of the momentum representation do not arise. Again in the case of three-body scattering amplitudes, the proper asymptotic states and the unitarity condition allows a simplification of the off-shell two-body $T$ matrix.

## ACKNOWLEDGMENT

The author gratefully acknowledges the helpful comments and discussions contributed by Dr. J. D. Garcia during the course of this work.

[7] L. Schulman, Phys. Rev. **156**, 1129 (1967).

## APPENDIX

Consider an integral of the following form:

$$\int d\mathbf{k} f(|\mathbf{k}-\mathbf{a}|^2) = \int_0^{2\pi} d\phi \int_{-1}^{1} d(\cos\theta) \int_0^{\infty} k^2\, dk\, f(k^2 + a^2 - 2ka\cos\theta). \tag{A1}$$

We change the variables of integration $k$, $\cos\theta$, to new variables $k'$ and $\cos\theta'$:

$$k = -k', \quad \cos\theta = -\cos\theta' \tag{A2}$$

Thus,

$$\begin{aligned}\int d\mathbf{k} f(|\mathbf{k}-\mathbf{a}|^2) &= \int_0^{2\pi} d\phi \int_1^{-1} d(\cos\theta') \\ &\quad \times \int_0^{-\infty} k'^2\, dk'\, f(k'^2 + a^2 - 2k'a\cos\theta') \\ &= \int_0^{2\pi} d\phi \int_{-1}^{1} d(\cos\theta') \\ &\quad \times \int_{-\infty}^{0} k'^2\, dk'\, f(k'^2 + a^2 - 2k'a\cos\theta'). \end{aligned} \tag{A3}$$

Averaging Eqs. (A1) and (A2), and dropping the primes on the dummy variables $k'$ and $\cos\theta'$ we get

$$\int d\mathbf{k} f(|\mathbf{k}-\mathbf{a}|^2) = \frac{1}{2}\int_0^{2\pi} d\phi \int_{-1}^{1} d(\cos\theta) \int_{-\infty}^{\infty} k^2\, dk\, f(|\mathbf{k}-\mathbf{a}|^2). \tag{A4}$$

This result can be generalized to integrands involving more than one vector. For example, the range of integration of integrands of the form

$$g(|\mathbf{k}-\mathbf{a}|^2; |\mathbf{k}-\mathbf{b}|^2)$$

can be expanded from $0 < k < \infty$ to $-\infty < k < \infty$ in a way similar to the method just given. In addition to the changes of variable (A2) we must change the azimuthal angle as follows: $\phi = \phi' + \pi$, where both $\phi$ and $\phi'$ have the range 0 to $2\pi$

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 5 MAY 1968

# Nonexistence of Finite-Energy Stationary Quantum States in Nonlinear Field Theories*

GERALD ROSEN
*Drexel Institute of Technology, Philadelphia, Pennsylvania*

(Received 29 June 1967)

A simple proof is given which precludes the existence of any finite-energy stationary quantum-state solution to the Schrödinger equation for a physically interesting Lorentz-covariant self-interacting scalar theory.

Supplemented with suitably smooth prescribed initial data, the nonlinear $c$-number field equation

$$\ddot{\theta} - \nabla^2\theta + m^2\theta + \tfrac{1}{2}gp\,|\theta|^{p-2}\theta = 0 \qquad (1)$$

for a scalar function $\theta = \theta(\mathbf{x}, t)$ (either complex valued or real valued and $C^2$ for $\mathbf{x}$ and all $t > 0$) has a *global solution* if $2 < p < 6$ with $m^2$ and $g$ nonnegative constant parameters.[1] However, it has been shown that the classical field equation (1) has no global solutions which are localized in space and periodic in time with $g \geq 0$ and $p \geq 2$.[2] Does the nonexistence of localized periodic solutions to Eq. (1) evince a corresponding nonexistence in the quantum field theory? That is, are stationary quantum states of finite energy generally precluded in self-interacting scalar field theories based on a Lagrangian density

$$\mathcal{L} = |\dot{\theta}|^2 - |\nabla\theta|^2 - m^2\,|\theta|^2 - g\,|\theta|^p \qquad (2)$$

of the form associated with Eq. (1)? Our purpose is to report a proof which shows that no finite-energy stationary quantum states exist for model theories based on a Lagrangian density of the form (2) with $g \geq 0$ and $p \geq 2$; hence, the vacuum and all other stationary states have an energy which is patently infinite.

To prove the nonexistence of any finite-energy stationary quantum state for a theory based on (2) with $g \geq 0$ and $p \geq 2$, let the $\theta$ (boson) field be diagonalized for all values of $\mathbf{x}$ at a fixed instant of time and consider the *energy functionality*[3]

$$E = E\{\Psi\} \equiv \int \Psi^* H \Psi \mathfrak{D}(\theta) \Big/ \int |\Psi|^2\, \mathfrak{D}(\theta) \qquad (3)$$

associated with a state functional $\Psi = \Psi[\theta]$, where

$$H = \int (\pi^*\pi + \nabla\theta^* \cdot \nabla\theta + m^2\theta^*\theta + g(\theta^*\theta)^{p/2})d^3\mathbf{x},$$

$$\left[\pi^* = -i\frac{N}{2}\hbar\frac{\delta}{\delta\theta^*}, \quad \pi = -i\frac{N}{2}\hbar\frac{\delta}{\delta\theta}, \right. \qquad (4)$$

$$\left. N = \begin{bmatrix} 1 \text{ for } \theta \text{ real} \\ 2 \text{ for } \theta \text{ complex} \end{bmatrix}\right]$$

is the Hamiltonian operator derived from (2) and $\mathfrak{D}(\theta)$ is a (real nonnegative displacement-invariant) measure for the functional integrations over all fields $\theta = \theta(\mathbf{x})$. Both $\Psi$ and $\mathfrak{D}(\theta)$ are defined to within normalization factors independent of $\theta$; it is assumed that the numerator and denominator in (3) exist as finite quantities, $E$ being finite for the $\Psi$'s under consideration here. The energy functionality (3) is stationary with respect to variations in $\Psi$ about a physical state functional, by virtue of the Schrödinger equation. In particular, for a variation in $\Psi$ induced by a transformation of the field $\theta \to \hat{\theta}$,

$$\Psi = \Psi[\theta] \to \hat{\Psi} = \hat{\Psi}[\theta], \qquad \hat{\Psi}[\hat{\theta}] \equiv \Psi[\theta], \qquad (5)$$

we have

$$\begin{aligned} E\{\hat{\Psi}\} &= \int \hat{\Psi}[\theta]^* H \hat{\Psi}[\theta] \mathfrak{D}(\theta) \Big/ \int |\hat{\Psi}[\theta]|^2\, \mathfrak{D}(\theta) \\ &= \int \hat{\Psi}[\hat{\theta}]^* \hat{H} \hat{\Psi}[\hat{\theta}] \mathfrak{D}(\hat{\theta}) \Big/ \int |\hat{\Psi}[\hat{\theta}]|^2\, \mathfrak{D}(\hat{\theta}) \\ &= \int \Psi^* \hat{H} \Psi \mathfrak{D}(\theta) \Big/ \int |\Psi|^2\, \mathfrak{D}(\theta), \end{aligned} \qquad (6)$$

provided that $\mathfrak{D}(\theta)$ and $\mathfrak{D}(\hat{\theta})$ only differ in normalization and where $\hat{H}$ is the Hamiltonian operator (4) expressed in terms of $\hat{\theta}$ and the associated $\hat{\pi}$. It follows that

$$\begin{aligned} E\{\hat{\Psi}\} - E\{\Psi\} &= \int \Psi^*(\hat{H} - H)\Psi \mathfrak{D}(\theta) \Big/ \int |\Psi|^2\, \mathfrak{D}(\theta) \\ &\equiv \langle \hat{H} - H \rangle \end{aligned} \qquad (7)$$

vanishes to first order with respect to transformations of the field $\theta \to \hat{\theta}$, provided that $\mathfrak{D}(\theta)$ and $\mathfrak{D}(\hat{\theta})$ only differ in normalization. Two such transformations of the field are of special importance.[4]

1. Dilatation induced: $\hat{\theta}(\mathbf{x}) \equiv \theta(\lambda\mathbf{x}), \hat{\pi}(\mathbf{x}) = \lambda^3\pi(\lambda\mathbf{x})$, $\lambda$ real and positive.

We find

$$\hat{H} = \int (\lambda^3\pi^*\pi + \lambda^{-1}\nabla\theta^* \cdot \nabla\theta + \lambda^{-3}m^2\theta^*\theta + \lambda^{-3}g(\theta^*\theta)^{p/2})d^3\mathbf{x} \qquad (8)$$

* Work supported by a National Science Foundation grant.

[1] K. Jörgens, Z. Math. 77, 295 (1961).
[2] G. Rosen, J. Math. Phys. 7, 2066 (1966).
[3] G. Rosen, Phys. Rev. Letters 16, 704 (1966).
[4] Such transformations have also been considered by H. Schiff, Proc. Roy. Soc. (London) A269, 277 (1962).

by changing the dummy integration variable $(\lambda \mathbf{x}) \to \mathbf{x}$. Hence, since (7) vanishes to first order about $\lambda = 1$,

$$(\partial\langle \hat{H}\rangle/\partial\lambda)_{\lambda=1} = \int \langle 3\pi^*\pi - \nabla\theta^* \cdot \nabla\theta - 3m^2\theta^*\theta - 3g(\theta^*\theta)^{p/2}\rangle d^3\mathbf{x} = 0. \quad (9)$$

2. Scale induced: $\hat{\theta}(\mathbf{x}) \equiv \xi\theta(\mathbf{x})$, $\hat{\pi}(\mathbf{x}) = \xi^{-1}\pi(\mathbf{x})$, $\xi$ real and positive.

We find

$$\hat{H} = \int (\xi^{-2}\pi^*\pi + \xi^2\nabla\theta^* \cdot \nabla\theta + \xi^2 m^2\theta^*\theta + \xi^p g(\theta^*\theta)^{p/2}) d^3\mathbf{x}, \quad (10)$$

and hence

$$(\partial\langle \hat{H}\rangle/\partial\xi)_{\xi=1} = \int \langle -2\pi^*\pi + 2\nabla\theta^* \cdot \nabla\theta + 2m^2\theta^*\theta + pg(\theta^*\theta)^{p/2}\rangle d^3\mathbf{x} = 0. \quad (11)$$

Now by adding $\frac{2}{3}$ of Eq. (9) to Eq. (11), we obtain

$$\int \langle \tfrac{4}{3}\nabla\theta^* \cdot \nabla\theta + (p-2)g(\theta^*\theta)^{p/2}\rangle d^3\mathbf{x} = 0, \quad (12)$$

a relation which implies the necessary condition $p < 2$ for existence of a finite-energy stationary state with $g > 0$. Therefore, all stationary states have an infinite energy in a quantum field theory with $g \geq 0$ and $p \geq 2$. Conditions for a local essentially nonlinear scalar field theory [based on a Lorentz-invariant Lagrangian density more general in form than (2)] to admit stationary quantum states have been reported elsewhere.[5]

[5] G. Rosen, Phys. Rev. **160**, 1278 (1967); **165**, 1934 (1968); **167**, 1395 (1968).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 5 MAY 1968

# Improved Method for Quantum-Mechanical Three-Body Problem. III. Use of Sturmian Functions

LEONARD EYGES AND JOHN R. JASPERSE*
*Air Force Cambridge Research Laboratories, Office of Aerospace Research, Bedford, Massachusetts*

(Received 11 August 1967)

We extend previous work on the ground state of the symmetric three-body problem by expanding the two-body orbitals $\varphi(k, \kappa)$ or $\varphi(\mathbf{k}, \boldsymbol{\kappa})$ (one and three dimensions, respectively) in a set of Sturmian functions. This retains the advantages of the previous expansions, and gains several new ones as well. Among these are a simplification of the equations and a transparent way of estimating convergence. The one-dimensional problem is reduced to an infinite set of coupled integral equations in one variable and the three-dimensional one to a doubly infinite set. As an application and a test of convergence we have solved the one-dimensional equations numerically in successive truncations. We find that keeping only the first term of the set yields results typically accurate to a fraction of a percent.

## I. INTRODUCTION

In previous papers by one of us,[1-4] we have treated the quantum-mechanical problem of three identical particles bound by identical interparticle potentials. In the first of these we pointed out the advantages of writing the wavefunction in a special way, as a sum of three parts or, as we called them, "two-body orbitals," one part for each interparticle distance. This idea proved fruitful for the bound-state problem and it was also applied by Mitra,[5] Fadeev,[6] and others to three-body scattering problems. Now these two-body orbitals derived their name from the fact that they satisfied an equation that resembled a two-body equation, and this observation made it natural to try to expand them in a complete set of two-body functions. This was done in Ref. 4. As we emphasized there, the most advantageous two-body set to use was not that generated by the interparticle potential that entered the three-body equation, but rather that for a potential of the same shape, but with enhanced strength. Even so there remained one problem. This (or any other ordinary two-body set) has both a discrete spectrum of eigenvalues and a continuous one, and with it one is faced in principle with the nasty problem of treating the continuous eigenfunctions. In this joint paper then, we get around this final difficulty by expanding the two-body orbitals in a set of two-body Sturmian functions. This is a set which

* Part of this research was in partial fulfillment of the requirement for the Ph.D. degree in physics at Northeastern University.

[1] L. Eyges, Ann. phys. **2**, 101 (1957).
[2] L. Eyges, Phys. Rev. **115**, 1643 (1959).
[3] L. Eyges, Phys. Rev. **121**, 1744 (1961).
[4] L. Eyges, J. Math. Phys. **6**, 1320 (1965).
[5] A. N. Mitra, Nucl. Phys. **32**, 529 (1962).
[6] L. D. Fadeev, Zh. Eksp. Teor. Fiz. **39**, 1459 (1960) [Sov. Phys.—JETP **12**, 1014 (1961)]; Dokl. Akad. Nauk SSSR **138**, 561 (1961); **145**, 301 (1962) [Sov. Phys.—Dokl. **6**, 384 (1961); **7**, 600 (1963)].

by changing the dummy integration variable $(\lambda \mathbf{x}) \to \mathbf{x}$. Hence, since (7) vanishes to first order about $\lambda = 1$,

$$(\partial\langle \hat{H}\rangle/\partial\lambda)_{\lambda=1} = \int \langle 3\pi^*\pi - \nabla\theta^* \cdot \nabla\theta - 3m^2\theta^*\theta - 3g(\theta^*\theta)^{p/2}\rangle d^3\mathbf{x} = 0. \quad (9)$$

2. Scale induced: $\hat{\theta}(\mathbf{x}) \equiv \xi\theta(\mathbf{x})$, $\hat{\pi}(\mathbf{x}) = \xi^{-1}\pi(\mathbf{x})$, $\xi$ real and positive.

We find

$$\hat{H} = \int (\xi^{-2}\pi^*\pi + \xi^2\nabla\theta^* \cdot \nabla\theta + \xi^2 m^2\theta^*\theta + \xi^p g(\theta^*\theta)^{p/2}) d^3\mathbf{x}, \quad (10)$$

and hence

$$(\partial\langle \hat{H}\rangle/\partial\xi)_{\xi=1} = \int \langle -2\pi^*\pi + 2\nabla\theta^* \cdot \nabla\theta + 2m^2\theta^*\theta + pg(\theta^*\theta)^{p/2}\rangle d^3\mathbf{x} = 0. \quad (11)$$

Now by adding $\frac{2}{3}$ of Eq. (9) to Eq. (11), we obtain

$$\int \langle \tfrac{4}{3}\nabla\theta^* \cdot \nabla\theta + (p-2)g(\theta^*\theta)^{p/2}\rangle d^3\mathbf{x} = 0, \quad (12)$$

a relation which implies the necessary condition $p < 2$ for existence of a finite-energy stationary state with $g > 0$. Therefore, all stationary states have an infinite energy in a quantum field theory with $g \geq 0$ and $p \geq 2$. Conditions for a local essentially nonlinear scalar field theory [based on a Lorentz-invariant Lagrangian density more general in form than (2)] to admit stationary quantum states have been reported elsewhere.[5]

[5] G. Rosen, Phys. Rev. **160**, 1278 (1967); **165**, 1934 (1968); **167**, 1395 (1968).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 5 MAY 1968

# Improved Method for Quantum-Mechanical Three-Body Problem. III. Use of Sturmian Functions

LEONARD EYGES AND JOHN R. JASPERSE*
*Air Force Cambridge Research Laboratories, Office of Aerospace Research, Bedford, Massachusetts*

(Received 11 August 1967)

We extend previous work on the ground state of the symmetric three-body problem by expanding the two-body orbitals $\varphi(k, \kappa)$ or $\varphi(\mathbf{k}, \boldsymbol{\kappa})$ (one and three dimensions, respectively) in a set of Sturmian functions. This retains the advantages of the previous expansions, and gains several new ones as well. Among these are a simplification of the equations and a transparent way of estimating convergence. The one-dimensional problem is reduced to an infinite set of coupled integral equations in one variable and the three-dimensional one to a doubly infinite set. As an application and a test of convergence we have solved the one-dimensional equations numerically in successive truncations. We find that keeping only the first term of the set yields results typically accurate to a fraction of a percent.

## I. INTRODUCTION

In previous papers by one of us,[1-4] we have treated the quantum-mechanical problem of three identical particles bound by identical interparticle potentials. In the first of these we pointed out the advantages of writing the wavefunction in a special way, as a sum of three parts or, as we called them, "two-body orbitals," one part for each interparticle distance. This idea proved fruitful for the bound-state problem and it was also applied by Mitra,[5] Fadeev,[6] and others to three-body scattering problems. Now these two-body orbitals derived their name from the fact that they satisfied an equation that resembled a two-body equation, and this observation made it natural to try to expand them in a complete set of two-body functions. This was done in Ref. 4. As we emphasized there, the most advantageous two-body set to use was not that generated by the interparticle potential that entered the three-body equation, but rather that for a potential of the same shape, but with enhanced strength. Even so there remained one problem. This (or any other ordinary two-body set) has both a discrete spectrum of eigenvalues and a continuous one, and with it one is faced in principle with the nasty problem of treating the continuous eigenfunctions. In this joint paper then, we get around this final difficulty by expanding the two-body orbitals in a set of two-body Sturmian functions. This is a set which

* Part of this research was in partial fulfillment of the requirement for the Ph.D. degree in physics at Northeastern University.

[1] L. Eyges, Ann. phys. **2**, 101 (1957).
[2] L. Eyges, Phys. Rev. **115**, 1643 (1959).
[3] L. Eyges, Phys. Rev. **121**, 1744 (1961).
[4] L. Eyges, J. Math. Phys. **6**, 1320 (1965).
[5] A. N. Mitra, Nucl. Phys. **32**, 529 (1962).
[6] L. D. Fadeev, Zh. Eksp. Teor. Fiz. **39**, 1459 (1960) [Sov. Phys.—JETP **12**, 1014 (1961)]; Dokl. Akad. Nauk SSSR **138**, 561 (1961); **145**, 301 (1962) [Sov. Phys.—Dokl. **6**, 384 (1961); **7**, 600 (1963)].

is complete, but which has no continuum. It has been exploited by Rotenberg[7] for three-body scattering problems, and is discussed below. In addition to doing away with the continuum, the Sturmian set has several other advantages that will appear later.

The method works for particles in any number of dimensions. It is then convenient to start with three bodies in one dimension, since the equations contain the essence of the method without the somewhat superfluous complications that three dimensions brings.

## II. GENERAL THEORY

The basic equation for the one-dimensional three-body problem is Eq. (3) of Ref. 4 for $\varphi(k, \kappa)$, the Fourier transform of the two-body orbital. This can be written

$$\varphi(k,\kappa) = -[2\pi(k^2 + \tfrac{3}{4}\kappa^2 + K^2)]^{-1}\iint v(x)\Big\{\varphi(k',\kappa) \times e^{ix(k'-k)} + \varphi\Big(\frac{k'}{2}, k' - 2\kappa\Big) \times [e^{ix(k'-k-\frac{3}{2}\kappa)} + e^{ix(-k'-k+\frac{3}{2}\kappa)}]\Big\}\,dk'\,dx. \quad (1)$$

It is convenient now to make a minor change in notation, and show explicitly both that the interparticle potential $v(x)$ is negative and that it has a length $a$ associated with it; instead of Eq. (4) of Ref. 4, we write

$$v(x) = -v_0 u(x/a).$$

Then Eq. (1) becomes, with $x/a = z$,

$$\varphi(k,\kappa) = \frac{w}{2\pi(k^2 + \frac{3}{4}\kappa^2 + \beta^2)}\iint u(z)\Big\{\varphi(k',\kappa) \times e^{iz(k'-k)} + \varphi\Big(\frac{k'}{2}, k' - 2\kappa\Big) \times [e^{iz(k'-k-\frac{3}{2}\kappa)} + e^{iz(-k'-k+\frac{3}{2}\kappa)}]\Big\}\,dk'\,dz. \quad (2)$$

Now $k$ and $\kappa$ are dimensionless, as are the energy parameter $\beta^2$ and potential-strength parameter $w$:

$$\beta^2 = m\,|E|\,a^2/\hbar^2, \quad w = mV_0a^2/\hbar^2.$$

Our problem now is to extract $\beta^2$ from Eq. (2) as a function of $w$ for the ground state. Consider then $\varphi(k, \kappa)$ as a function of $k$. Much as we have done before,[4] we want to expand this dependence in a complete set of functions; as we have remarked we shall take this to be a Sturmian set corresponding to a two-body potential. This set is described in detail in the Appendix: suffice it here to say in general how it arises. If we consider a two-body problem involving a potential[8] $\tilde{w}u(z)$ and corresponding energy parameter $\tilde{\beta}$, then one ordinarily considers the spectrum of eigenvalues $\tilde{\beta}_1, \tilde{\beta}_2, \tilde{\beta}_3 \cdots$ for a fixed $\tilde{w}$. We can turn this around however, and imagine $\tilde{\beta}$ fixed, and then ask for the set of eigenvalues $\tilde{w}_l$ and corresponding eigenfunctions. This set is the Sturmian set.

For the ground state of the three-body problem for which $\varphi$ is an even function of $k$, we need the even functions of this set in momentum space; we call them $T_l(k)$.[9] From the Appendix we take two basic properties of these functions. First, they satisfy the Schrödinger equation in momentum space for $\tilde{\beta}$ fixed and $\tilde{w}_l$ considered as an eigenvalue:

$$T_l(k) = \frac{\tilde{w}_l}{2\pi(k^2 + \tilde{\beta}^2)}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u(z)T_l(k')e^{iz(k'-k)}\,dz\,dk'; \quad (3)$$

second, they satisfy the orthogonality relation

$$\int_{-\infty}^{\infty}(k^2 + \tilde{\beta}^2)T_l^*(k)T_n(k)\,dk = \tilde{w}_l\delta_{nl}. \quad (4)$$

With these functions then we expand $\varphi(k, \kappa)$

$$\varphi(k,\kappa) = \sum_{l=1}^{\infty} T_l(k)f_l(\kappa), \quad (5)$$

and put this into (2) to get

$$\Big(k^2 + \frac{3\kappa^2}{4} + \beta^2\Big)\sum_{l=1}^{\infty} T_l(k)f_l(\kappa) = \frac{w}{2\pi}\iint u(z)\Big\{e^{iz(k'-k)}\sum_{l=1}^{\infty} T_l(k')f_l(\kappa) + [e^{iz(k'-k-3\kappa/2)} + e^{iz(-k'-k+3\kappa/2)}] \times \sum_{l=1}^{\infty} T_l\Big(\frac{k'}{2}\Big)f_l(k' - 2\kappa)\Big\}\,dk'\,dz. \quad (6)$$

We multiply through by $T_n^*(k)$ and integrate with respect to $k$ to obtain

$$\int\Big[T_n^*(k)(k^2 + \tfrac{3}{4}\kappa^2 + \beta^2)\sum_{l=1}^{\infty} T_l(k)f_l(\kappa)\Big]\,dk = \frac{w}{2\pi}\iiint u(z)T_n^*(k)e^{iz(k'-k)}\Big\{\sum_{l=1}^{\infty} T_l(k')f_l(\kappa)\Big\}\,dk'\,dz\,dk + \frac{w}{2\pi}\iiint u(z)T_n^*(k)[e^{iz(k'-k-3\kappa/2)} + e^{iz(-k'-k+3\kappa/2)}] \times \Big\{\sum_{l=1}^{\infty} T_l\Big(\frac{\kappa'}{2}\Big)f_l(k' - 2\kappa)\Big\}\,dk'\,dz\,dk. \quad (7)$$

We can then simplify this equation considerably as follows. For the left-hand side we define constants

[7] M. Rotenberg, Ann. Phys. 19, 262 (1962).

[8] As before, we shall generally use the tilde (~) to distinguish a quantity that pertains to the two-body problem from its three-body counterpart.

[9] We label the first member of this set with $l = 1$, not $l = 0$.

$c_{nl}$ and $d_{nl}$ by

$$c_{nl} = \int T_n^*(k) T_l(k)\, dk, \tag{8}$$

$$d_{nl} = \int T_n^*(k) T_l(k) k^2\, dk, \tag{9}$$

whereupon it becomes

$$\sum_{l=1}^{\infty} f_l(\kappa)((\tfrac{3}{4}\kappa^2 + \beta^2)c_{nl} + d_{nl}),$$

which is the left-hand side of Eq. (7). The first triple integral on the right-hand side reduces immediately on using Eq. (A8) of the Appendix, and is

$$w f_n(\kappa),$$

the first term on the right-hand side of Eq. (7). For the remaining term on the right-hand side, we use a slight variant of (3), viz.,

$$\iiint u(z) T_n^*(k) e^{-iz(k-\xi)}\, dk\, dz = \frac{2\pi(\xi^2 + \tilde{\beta}^2) T_n^*(\xi)}{\tilde{w}_n},$$

and the second triple integral on the right-hand side reduces to

$$\frac{w}{\tilde{w}_n} \int \left[\left(k' - \frac{3\kappa}{2}\right)^2 + \beta^2\right] \times \left[T_n^*\left(k' - \frac{3\kappa}{2}\right) + T_n^*\left(-k' + \frac{3\kappa}{2}\right)\right] \times \left(\sum_{l=1}^{\infty} T_l\left(\frac{k'}{2}\right) f_l(k' - 2\kappa)\right) dk',$$

which is the second term on the right-hand side of Eq. (7). We put all this together now, recognize that $T_n(k) = T_n(-k)$ and change the variable of integration in the last expression from $k'$ to $y = k' - 2\kappa$. We get

$$\sum_{l=1}^{\infty} f_l(\kappa)\left[\left(\frac{3\kappa^2}{4} + \beta^2\right)c_{nl} + d_{nl}\right] - w f_n(\kappa) = \frac{2w}{\tilde{w}_n} \int \left[\left(y + \frac{\kappa}{2}\right)^2 + \beta^2\right] T_n^*\left(y + \frac{\kappa}{2}\right) \times \left(\sum_{l=1}^{\infty} T_l\left(\frac{y}{2} + \kappa\right) f_l(y)\right) dy. \tag{10}$$

There is one further simplification we can make. From Eq. (A9) we have

$$\tilde{\beta}^2 c_{nl} + d_{nl} = \tilde{w}_n \delta_{nl}. \tag{11}$$

Now the parameter $\tilde{\beta}$ that enters the two-body equation is still at our disposal. We have already indicated[3,4] on physical grounds the desirability of choosing it to equal $\beta$, the three-body binding energy. Now we see a mathematical reason as well, for if we do this Eq. (11) becomes

$$\beta^2 c_{nl} + d_{nl} = \tilde{w}_n \delta_{nl}, \tag{12}$$

and this can be used to simplify the left-hand side of Eq. (10) so that it becomes

$$f_n(\kappa)(\tilde{w}_n - w) + \tfrac{3}{4}\kappa^2 \sum_{l=1}^{\infty} f_l(\kappa) c_{nl} = \frac{2w}{\tilde{w}_n} \int_{-\infty}^{\infty} dy [(y + \tfrac{1}{2}\kappa)^2 + \beta^2] T_n^*(y + \tfrac{1}{2}\kappa) \times \left[\sum_{l=1}^{\infty} f_l(y) T_l(\tfrac{1}{2}y + \kappa)\right]. \tag{13}$$

This then is the final set of integral equations in the functions $f_n(\kappa)$.

With this at hand we can discuss the three-dimensional problem; its general outline is very similar to the one-dimensional one, except that vector variables replace scalar ones. As in one dimension we write a typical interparticle potential $V(r)$ as

$$V(r) = -V_0 u(r/a),$$

and use the same definition for the energy parameter $\beta^2$ and potential strength parameter $w$. Then the basic equation for the two-body orbital $\varphi(\mathbf{k}, \boldsymbol{\kappa})$, Eq. (2) of Ref. 4, becomes

$$\phi(\mathbf{k}, \boldsymbol{\kappa}) = \frac{w}{(2\pi)^3(k^2 + \frac{3}{4}\kappa^2 + \beta^2)} \iint u(r) \Big\{\phi(\mathbf{k}', \boldsymbol{\kappa}) \times e^{i\mathbf{r}\cdot(\mathbf{k}'-\mathbf{k})} + \phi(\tfrac{1}{2}\mathbf{k}', \mathbf{k}' - 2\boldsymbol{\kappa}) \times [e^{i\mathbf{r}\cdot(\mathbf{k}'-\mathbf{k}-\frac{3}{2}\boldsymbol{\kappa})} + e^{i\mathbf{r}\cdot(-\mathbf{k}'-\mathbf{k}+\frac{3}{2}\boldsymbol{\kappa})}]\Big\}\, d\mathbf{k}'\, d\mathbf{r}. \tag{14}$$

We shall again expand $\varphi$ as a function of $\mathbf{k}$, in a complete set of Sturmian functions. These are defined as the eigenfunctions of the relative two-body problem for a potential $-\tilde{V}_0 u(r/a)$ with the energy parameter $\tilde{\beta}^2 = 2m|E|/\hbar^2$ fixed and the potential strength parameter $\tilde{w} = 2m\tilde{V}_0/\hbar^2$ as eigenvalue. They are discussed in detail in the Appendix. As in one dimension, and for the same reasons, we shall choose the energy parameter $\tilde{\beta}$ to be equal to $\beta$.

We shall be concerned mainly with the three-dimensional Sturmian functions in momentum space; we call them $T_{\mathbf{s}}(\mathbf{k})$. The vector index $\mathbf{s}$ stands for the multiple quantum numbers that are appropriate to the three dimensions; we shall specify them more closely later. But writing them in the present form for the moment we retain a very useful similarity with the one-dimension case. Thus we expand $\varphi(\mathbf{k}, \boldsymbol{\kappa})$ much as in Eq. (5),

$$\varphi(\mathbf{k}, \boldsymbol{\kappa}) = \sum_{\mathbf{s}} T_{\mathbf{s}}(\mathbf{k}) F_{\mathbf{s}}(\boldsymbol{\kappa}), \tag{15}$$

and inserting this in (14) can use essentially the same algebra and subsidiary equations as before. We therefore omit the details, and simply record the basic set of equations, the analogs of Eqs. (13), for the functions $F_s(\boldsymbol{\varkappa})$. These are

$$F_s(\boldsymbol{\varkappa})(\tilde{w}_s - w) + \frac{3\kappa^2}{4}\sum_t F_t(\boldsymbol{\varkappa})C_{st} = \frac{w}{\tilde{w}_s}\int[(\mathbf{y}+\boldsymbol{\varkappa}/2)^2+\beta^2] \times \{T_s^*(\mathbf{y}+\boldsymbol{\varkappa}/2) + T_s^*(-\mathbf{y}-\boldsymbol{\varkappa}/2)\} \times \sum_p F_p(\mathbf{y})T_p(\mathbf{y}/2+\boldsymbol{\varkappa})\,d\mathbf{y}, \quad (16)$$

where

$$C_{st} = \int T_s^*(\mathbf{k})T_t(\mathbf{k})\,d\mathbf{k}. \quad (17)$$

Now the form (15) for $\varphi(\mathbf{k}, \boldsymbol{\varkappa})$ is really too general if we are interested in the ground state, which presumably corresponds to zero angular momentum. For we have shown previously that the most general form for this case is

$$\varphi(\mathbf{k}, \boldsymbol{\varkappa}) = \sum_{l=0}^{\infty}\varphi_l(k, \kappa)P_l(\cos\gamma), \quad (18)$$

where $\gamma$ is the angle between $\mathbf{k}$ and $\boldsymbol{\varkappa}$. This limits the form of the functions $F_s(\boldsymbol{\varkappa})$, as we show now. As we discuss in the Appendix the vector index $\mathbf{s}$ that labels the functions $T_s(\mathbf{k})$ can be taken to stand for the three conventional quantum numbers $nlm$, with $T_s(\mathbf{k})$ of the form

$$T_s(\mathbf{k}) \to T_{nl}(k)Y_{lm}(\Omega_k).$$

Thus the expansion (15) can be written

$$\varphi(\mathbf{k}, \boldsymbol{\varkappa}) = \sum_{nlm} T_{nl}(k)Y_{lm}(\Omega_k)F_{nlm}(\boldsymbol{\varkappa}).$$

If we compare this to (18), in which we imagine $P_l(\cos\gamma)$ expanded by the addition theorem, we conclude that $F_{nlm}(\boldsymbol{\varkappa})$ is of the form

$$F_{nlm}(\boldsymbol{\varkappa}) \to F_{nl}(\kappa)Y_{lm}^*(\Omega_\kappa). \quad (19)$$

If then we put (19) into (16) we get

$$F_{nl}(\kappa)Y_{lm}^*(\Omega_\kappa)(\tilde{w}_{nl} - w) + \frac{3\kappa^2}{4}\sum_{n'l'm'} F_{n'l'}(\kappa)Y_{l'm'}^*(\Omega_\kappa)C_{nlmn'l'm'} = \frac{w}{\tilde{w}_{nl}}\int[(\mathbf{y}+\boldsymbol{\varkappa}/2)^2+\beta^2](1+(-)^l) \times T_{nl}^*(|\mathbf{y}+\boldsymbol{\varkappa}/2|)Y_{lm}^*(\Omega_{y+\kappa/2}) \times \left\{\sum_{n'l'm'} F_{n'l'}(y)Y_{l'm'}^*(\Omega_y)T_{n'l'}(|\mathbf{y}/2+\boldsymbol{\varkappa}|) \times Y_{l'm'}(\Omega_{y/2+\kappa})\right\}y^2\,dy\,d\Omega_y. \quad (20)$$

This equation can now be simplified in several ways. First, for the summation over $m'$ under the integral sign we can write

$$\sum_{m'} Y_{l'm'}^*(\Omega_y)Y_{l'm'}(\Omega_{y/2+\kappa}) = [(2l'+1)/4\pi]P_{l'}(\cos\mu),$$

where $\mu$ is the angle between $\mathbf{y}$ and $\mathbf{y}/2 + \boldsymbol{\varkappa}$. Moreover, since Eq. (20) holds for arbitrary $m$ we can conveniently set $m = 0$, whereupon

$$Y_{l0}(\Omega_\kappa) \to [(2l+1)/4\pi]^{\frac{1}{2}}P_l(\cos\theta_\kappa)$$

where $\theta_\kappa$ is the angle between $\boldsymbol{\varkappa}$ and the $z$ axis. If we then multiply by $\sin\theta_\kappa$ and integrate over $\theta_\kappa$ we get

$$F_{nl}(\kappa)(\tilde{w}_{nl} - w) + \frac{3\kappa^2}{4}\sum_{n'} F_{n'l}(\kappa)C_{nl0,n'l0} = \frac{w}{\tilde{w}_{nl}}\frac{(2l+1)}{8\pi}\int[(\mathbf{y}+\boldsymbol{\varkappa}/2)^2+\beta^2](1+(-)^l) \times T_{nl}(|\mathbf{y}+\boldsymbol{\varkappa}/2|)P_l(\cos\theta_{y+\kappa/2})P_l(\cos\theta_\kappa) \times \left\{\sum_{n'l'}(2l'+1)F_{n'l'}(y)T_{n'l'}(|\mathbf{y}/2+\boldsymbol{\varkappa}|)P_{l'}(\cos\mu)\right\} \times y^2\sin\theta_\kappa\,dy\,d\Omega_y\,d\theta_\kappa. \quad (21)$$

Then this is the final set of equations.

## III. APPLICATIONS

As an example we apply Eqs. (13) to a square-well interparticle potential[10] which is defined by

$$u(z) = \begin{cases} 1, & |z| < 1 \\ 0, & |z| > 1. \end{cases}$$

The two-body functions $T_n(k)$, which are discussed in the Appendix, are then

$$T_n(k) = \frac{2}{(2\pi)^{\frac{1}{2}}[1+(\sin 2\alpha_n/2\alpha_n)]^{\frac{1}{2}}}\left\{\frac{(\tilde{\beta}^2+\alpha_n^2)}{(\tilde{\beta}^2+k^2)(\alpha_n+k)} \times \left[\frac{\alpha_n\sin(\alpha_n-k)}{\alpha_n-k} + \cos\alpha_n\sin k\right]\right\}. \quad (22)$$

The relation between the energy parameter $\tilde{\beta}$ and well strength $\tilde{w}_n$ is given by

$$\alpha_n\tan\alpha_n = \tilde{\beta} \quad (23)$$

with

$$\alpha_n = (\tilde{w}_n - \tilde{\beta}^2)^{\frac{1}{2}}. \quad (24)$$

With these results, we are now ready to discuss the numerical solutions of Eqs. (13). Since this is an infinite set of equations, we must of course truncate it. We shall begin by truncating it as severely as possible,

[10] In Ref. 7 Rotenberg has remarked that the Sturmian set is not complete in regions of space where the potential is zero; however, this seems to cause no difficulty in the present problem.

i.e., by assuming only $f_1$ is different from zero; later we discuss the error that this involves.

With this assumption Eqs. (13) reduce to the single one

$$f_1(\kappa) = \frac{2w}{\tilde{w}_1\left(\tilde{w}_1 - w + \frac{3\kappa^2}{4}c_{11}\right)} \int \left[\left(y + \frac{\kappa}{2}\right)^2 + \beta^2\right] \times T_1^*\left(y + \frac{\kappa}{2}\right) T_1(y/2 + \kappa) f_1(y)\, dy. \quad (25)$$

According to the prescription above in which $\tilde{\beta}$ is taken to be equal to $\beta$, the parameter $\tilde{w}_1$ is determined according to (24) from

$$(\tilde{w}_1 - \beta^2)^{\frac{1}{2}} \tan (\tilde{w}_1 - \beta^2)^{\frac{1}{2}} = \beta.$$

If then we consider that this equation is to be satisfied by fixing $\beta$ and then solving for $\tilde{w}$, we can consider $\beta$ as known. The functions $T_1(y + \kappa/2)$ and $T_1(y/2 + \kappa)$ that enter (13), and that are defined by Eq. (22), are then well-determined. If we put them into (25) explicitly we get the basic equation

$$f_1(\kappa) = \frac{4w(\beta^2 + \alpha_1^2)^2}{\pi\tilde{w}_1(\tilde{w}_1 - w + \frac{3}{4}\kappa^2 c_{11})[1 + (\sin 2\alpha_1/2\alpha_1)]} \times \int_{-\infty}^{\infty} \frac{f_1(y)}{\left\{\beta^2 + \left[\frac{y}{2} + \kappa\right]^2\right\}\left[\alpha_1 + y + \frac{\kappa}{2}\right]\left[\alpha_1 + \frac{y}{2} + \kappa\right]} \times \left\{\frac{\alpha_1 \sin (\alpha_1 - y/2 - \kappa)}{(\alpha_1 - y/2 - \kappa)} + \cos \alpha_1 \sin (y/2 + \kappa)\right\} \times \left\{\frac{\alpha_1 \sin (\alpha_1 - y - \kappa/2)}{(\alpha_1 - y - \kappa/2)} + \cos \alpha_1 \sin (y + (\kappa/2))\right\} dy. \quad (26)$$

To solve this equation we have approximated the integral by a sum, and thereby converted it to a set of linear equations whose vanishing determinant gives the relation between $\beta$ and $w$ that we seek. To check the approximation made in truncating the set of Eqs. (13) we have also similarly solved the analogous pair and triad of equations that are obtained by truncating less severely, i.e., when $f_1$ and $f_2$ are assumed nonzero, and likewise for $f_1$, $f_2$, and $f_3$. The results are given in Table I.

We see that there is a very small difference in the numerical values for the three different cases we consider. The approximation of keeping only $f_1$ is thus a very good one. We see this from another point of view if we look at the functions $f_1$, $f_2$, and $f_3$ which are plotted in Fig. 1. The small differences among the numerical values in Table I are reflected in the fact that $f_2$ and $f_3$ are indeed small compared with $f_1$.

Although these results are quite satisfactory,

TABLE I. The binding-energy parameter $\beta^2 = m|E|a^2/\hbar^2$ vs the potential strength parameter $w = mV_0a^2/\hbar^2$, for one-dimensional, three-body problem with square-well interparticle potentials of strength $V_0$ and width $2a$. The last three columns give the results obtained by truncating Eqs. (13) successively, i.e., by first keeping only $f_1$, then $f_1$ and $f_2$, and finally $f_1$ and $f_2$ and $f_3$.

| | $\beta^2$ | | |
|---|---|---|---|
| $w$ | With $f_1$ | With $f_1$ and $f_2$ | With $f_1, f_2$, and $f_3$ |
| 0.5 | 0.57920 | 0.57920 | 0.57895 |
| 1.0 | 1.6210 | 1.6210 | 1.6231 |
| 1.5 | 2.8204 | 2.8217 | 2.8214 |
| 2.0 | 4.0958 | 4.0998 | 4.0997 |

considerable computing capacity is needed to solve the equations when $f_1$, $f_2$, and $f_3$ are included. Since $f_2$ and $f_3$ are indeed small, it is then natural to see if they can be obtained to sufficient accuracy from $f_1$ itself, in an iterative way. This is, if possible, clearly desirable, since it reduces the problem from that of solving several coupled integral equations to that of solving a single one with some kind of subsequent integration.

To show that this is indeed possible, we do it. In (13) then we set $n = 2$, restrict the sum on the left-hand side to $l = 1$ and $l = 2$, and keep only the presumed dominant term corresponding to $l = 1$ under the integral sign. Then our approximate expression for $f_2$ is

$$f_2(\kappa) \approx -\frac{3\kappa^2 c_{21} f_1(\kappa)}{4\left[\tilde{w}_2 - w + \frac{3\kappa^2}{4}c_{22}\right]} + \frac{2w}{\tilde{w}_2\left[\tilde{w}_2 - w + \frac{3\kappa^2}{4}c_{22}\right]} \times \int_{-\infty}^{\infty} \left[\left(y + \frac{\kappa}{2}\right)^2 + \beta^2\right] T_2^*\left(y + \frac{\kappa}{2}\right) \times T_1\left(\frac{y}{2} + \kappa\right) f_1(y)\, dy. \quad (27)$$

We have calculated $f_2(\kappa)$ numerically from this equation for the case $w = 2$; the results are presented in Fig. 2 and compared with the "exact" result for $f_2$ obtained as described above. The agreement is quite good.

It is instructive to look at the qualitative reasons that $f_2$ is everywhere small, since very similar reasons will apply for the three-dimensional case. Part of the reason is the fact that $\tilde{w}_2$ is fairly large, usually an order of magnitude larger than $w_1$. For example, for $w = 1$, $\tilde{w}_2 = 13.3$; for $w = 2$, $\tilde{w}_2 = 14.4$. With

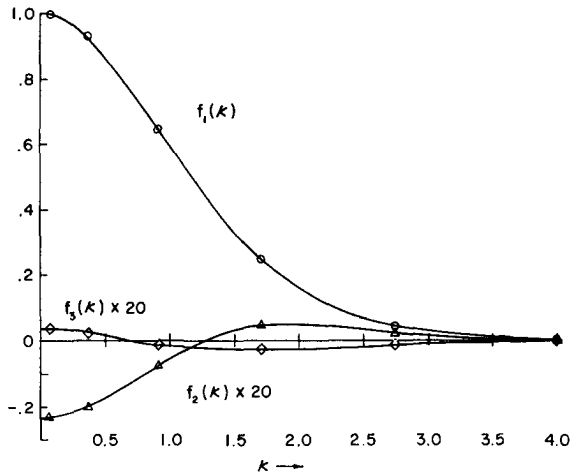



FIG. 1. The functions $f_1(\kappa)$, $f_2(\kappa)$, $f_3(\kappa)$ as calculated numerically from the truncated set of Eqs. (13); $w = 1$.

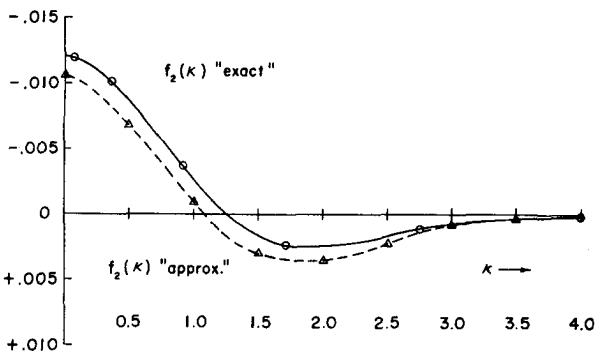



FIG. 2. The function $f_2$ calculated approximately from Eq. (19) and the zeroth-order expression for $f_1$, compared with the same as calculated as for Fig. 1.

this in mind we see that the first term on the right-hand side of Eq. (19) is small on two accounts; first, there is the factor $\tilde{w}_2$ in the denominator; second, since $f_1(\kappa)$ peaks at the origin (and is in fact normalized to unity there), $\kappa^2$ is large when $f_1$ is small and vice versa, and their product turns out always to be small. The term on the right-hand side that contains the integral is small because the integral is of order unity, but the factor that multiplies it is of order $w/\tilde{w}_2^2$. The higher functions $f_3$, etc., can be estimated in much the same way, and they are still smaller for similar but stronger reasons.

The solution of Eqs. (21) for the three-dimensional case is begun along the same lines; it is of course more complicated in that a double truncation, in the two indices $n$ and $l$ are involved. One would expect, *a priori*, almost as good convergence as for the one-dimensional case since we have already found[3,4] that the truncation in $l$ causes small error, and the truncation as a function of $n$ is very similar to that in one dimension. We have made a preliminary investigation of this in that we have solved numerically the truncated Eq. (21) keeping only $F_{10}$. The answers we get do not, however, agree with those due to Kalos[11]; we have still to resolve this. It may be that Kalos' results are in error, although this seems unlikely, or that the truncation we have been forced to make is much worse than in one dimension,[12] or even that there may be some difficulty due to the fact that the Sturmian functions for a cutoff potential do not form a complete set for expansion outside the cutoff region. But these matters demand more discussion than is appropriate for the present paper, and we shall leave them in abeyance for the moment, contenting ourselves with the one-dimensional results and with the fact that a generalization of this method has proved itself elsewhere[13] for the He atom, i.e., for Coulomb interparticle potentials.

## IV. DISCUSSION

The present method has now been expounded in several papers; it seems worthwhile at this point to summarize it, and highlight its basic features, so that the essentials stand out. Moreover, Fadeev[6] has discussed the scattering problem with a method that has some resemblance to this one, and it is worth pointing out both the similarities and the differences.

The first feature is the splitting of the wavefunction into three parts, as in Eq. (12) of Ref. 2. The motivation for this is described at some length in the earlier papers, and we shall not repeat it. We note, however, that a similar technique was later applied by Fadeev, who in the scattering problem split the $t$ matrix analogously into three parts. This splitting has by now been proven advantageous in several ways; it turns out that these "partial wavefunctions" generally satisfy simpler equations than the total wavefunctions. For example, the scattering problem for separable potentials can be solved with it,[5] as can the problem of the three repulsive $\delta$ functions in a box,[12] and with it, as one of us has shown,[13] even the problem of the He atom becomes tractable.

The next point of the method, which is conceptually independent of the previous one, is the expansion of the partial wavefunctions or "two-body orbitals" in a complete set of two-body functions. These are most naturally taken to be the eigenfunctions that correspond to the interparticle potential acting between any pair. Analogously, in the scattering problem Fadeev

[11] M. H. Kalos, Phys. Rev. **128**, 1791 (1962).
[12] L. Eyges, J. Math. Phys. **7**, 938 (1966).
[13] J. Jasperse, thesis, Northeastern University (1966); J. Jasperse and M. Friedman, Phys. Rev. **159**, 69 (1967).

expresses the three-body $t$ matrix in terms of all possible two-body $t$ matrices. The advantage of this procedure in the present case is that it reduces the equations for the orbitals to fairly simple and transparent form, since one can use equations like (A8) to simplify what would otherwise be complicated integrals. Thus, the expansion in a two-body set removes any explicit appearance of the potentials from the three-body equations.

The third feature of the method is essentially the observation that the two-body expansion basis we have just discussed may be the most natural, but it is not necessarily the most advantageous. As we have shown, we can expand in eigenfunctions of not only the true interparticle potential, but also of one with the same *shape* and enhanced *strength*. This gives an extra parameter, which can be exploited to considerable advantage. As we have seen if we adopt the prescriptions given in the text for this parameter, we simplify the form of the equations and insure the correctness of the asymptotic form of the two-body orbitals.

The fourth step in the method, and this is the innovation of the present paper, is of course, to choose not an ordinary set of two-body functions, but the Sturmian set. One then retains all the previous advantages, and also does away with continuum problems Moreover, the problem of truncating the equations becomes relatively transparent, as we have seen, in that estimates of the higher-order functions are given in terms of the Sturmian eigenvalues, for example, as in Eq. (27).

## ACKNOWLEDGMENTS

The authors would like to thank Paul Tsipouras of the Data Processing Branch of this Laboratory for his competent help in the computer programming.

## APPENDIX: STURMIAN FUNCTIONS

We first treat one dimension for which the Schrödinger equation[14] for the potential

$$V(x) = -V_0 u(x/a) \tag{A1}$$

and binding energy $E = -|E|$ is

$$-\frac{\hbar^2}{m}\frac{d^2\psi}{dx^2} - V_0 u(x/a)\psi = -|E|\,\psi. \tag{A2}$$

It will be general enough for us to assume that the potential $u(x/a)$ is symmetric about the origin. Then the eigenfunctions of $\psi$ split up into two groups: the symmetric ones which we call $S_l$, and the antisymmetric ones, which we do not need to consider. Now we set

$$z = x/a, \quad \tilde{\beta}^2 = m|E|\,a^2/\hbar^2, \tilde{w} = mV_0 a^2/\hbar^2; \tag{A3}$$

the tildes ($\sim$) over the parameters $\beta$ and $w$ emphasize that they refer to a two-body problem. We also put a subscript on $\tilde{w}$ to emphasize that it is now one of a set of eigenvalues, and rewrite Eq. (A2) for the specifically symmetric function $S_l$,

$$\frac{d^2S_l}{dz^2} - (\tilde{w}_l u(z) + \tilde{\beta}^2)S_l = 0. \tag{A4}$$

The Fourier transform of $S_l$ is $T_l(k)$,

$$T_l(k) = \frac{1}{(2\pi)^{\frac{1}{2}}}\int_{-\infty}^{\infty} S_l(z)e^{-ikz}\,dz, \tag{A5}$$

and it satisfies

$$T_l(k) = \frac{\tilde{w}_l}{2\pi(k^2+\tilde{\beta}^2)}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u(z)T_l(k')e^{iz(k'-k)}dz\,dk'. \tag{A6}$$

The orthogonality relation for the $S_l$, obtained from (A4) by multiplying it by $S_m^*$, integrating, etc., is not quite the standard one but reads

$$\int_{-\infty}^{\infty} S_m^*(z)S_l(z)u(z)\,dz = \delta_{ml}. \tag{A7}$$

This equation becomes, in terms of $T_m^*$ and $T_l$,

$$\frac{1}{2\pi}\int\int\int T_m^*(k)T_l(k')e^{iz(k'-k)}u(z)\,dk\,dk'\,dz = \delta_{ml}. \tag{A8}$$

This last equation, used in conjunction with (A6), gives the further useful relation

$$\int_{-\infty}^{\infty}(k^2+\tilde{\beta}^2)T_m^*(k)T_l(k)\,dk = \delta_{ml}\tilde{w}_m. \tag{A9}$$

For the one-dimensional application in the text we need the symmetric Sturmian set for the potential

$$u(z) = \begin{cases} 1, & |z| < 1 \\ 0, & |z| > 1. \end{cases}$$

This is elementary to calculate; the functions are

$$S_n(z) = \begin{cases} N_n \cos\alpha_n z, & |z| < 1, \\ N_n \cos\alpha_n \exp[\tilde{\beta}(1-|z|)], & |z| > 1, \end{cases}$$

where

$$\alpha_n = (\tilde{w}_n - \tilde{\beta}^2)^{\frac{1}{2}}.$$

The normalization constant $N_n$ is

$$N_n = \left(1 + \frac{\sin\alpha_n\cos\alpha_n}{\alpha_n}\right)^{-\frac{1}{2}},$$

[14] Note the factor $m$ instead of the usual $2m$ in the denominator of this equation. This comes in because $x$ is really a relative variable and $m$ is really the reduced mass.

and the relation between the energy parameter $\tilde{\beta}$ and the strength parameter $\tilde{w}_n$ obtained by matching solutions at $|z| = 1$ is

$$\alpha_n \tan \alpha_n = \tilde{\beta}.$$

The Fourier transform $T_n(k)$ turns out to be

$$T_n(k) = \frac{2N_n}{(2\pi)^{\frac{1}{2}}} \left\{ \frac{\tilde{\beta}^2 + \alpha_n^2}{(\tilde{\beta}^2 + k^2)(\alpha_n + k)} \left[ \frac{\alpha_n \sin(\alpha_n - k)}{\alpha_n - k} + \cos \alpha_n \sin k \right] \right\}. \quad \text{(A10)}$$

Now we turn to the three-dimensional case, which is closely analogous to the one-dimensional one, except that scalar variables become vector ones. Thus, consider the Schrödinger equation for the relative motion of two particles with interparticle distance $r'$, and interparticle potential $V(r')$. With

$$V(r') = -V_0 u(r'/a),$$

we have with

$$r = r'/a$$

and the same definitions as in Eq. (A3) for $\tilde{\beta}$ and $\tilde{w}$,

$$\nabla^2 \psi - (\tilde{w} u(r) + \tilde{\beta}^2)\psi = 0. \quad \text{(A11)}$$

The set of solutions of this equation with $\tilde{\beta}$ fixed and $\tilde{w}$ considered as eigenvalues are the Sturmian functions, which we call $S(\mathbf{r})$. They are labeled by the multiple quantum numbers of three-dimensions, which we can represent by a vector subscript $\mathbf{l}$. Thus as a function of $\mathbf{r}$ they satisfy the relabeled Eq. (A11),

$$\nabla^2 S_{\mathbf{l}} - (\tilde{w}_{\mathbf{l}} u(r) + \tilde{\beta}^2) S_{\mathbf{l}} = 0. \quad \text{(A12)}$$

We call the Fourier transform $T_{\mathbf{l}}(\mathbf{k})$,

$$T_{\mathbf{l}}(\mathbf{k}) = \frac{1}{(2\pi)^{\frac{3}{2}}} \int S_{\mathbf{l}}(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}}\, d\mathbf{r}. \quad \text{(A13)}$$

With this vector notation the properties of the $S_{\mathbf{l}}(\mathbf{r})$ and $T_{\mathbf{l}}(\mathbf{k})$ are very close to those of their one-dimensional counterparts. Thus the normalization condition analogous to (A7) reads

$$\int S_{\mathbf{n}}^*(\mathbf{r}) S_{\mathbf{l}}(\mathbf{r}) u(r)\, d\mathbf{r} = \delta_{\mathbf{nl}}. \quad \text{(A14)}$$

Similarly, the analogs of (A6) and (A9) are

$$T_{\mathbf{l}}(\mathbf{k}) = \frac{\tilde{w}_{\mathbf{l}}}{(2\pi)^3 (k^2 + \tilde{\beta}^2)} \iint u(r) T_{\mathbf{l}}(\mathbf{k}') e^{i\mathbf{r}\cdot(\mathbf{k}'-\mathbf{k})}\, d\mathbf{r}\, d\mathbf{k}',$$

and

$$\int (k^2 + \beta^2) T_{\mathbf{m}}^*(\mathbf{k}) T_{\mathbf{l}}(\mathbf{k})\, d\mathbf{k} = \delta_{\mathbf{ml}} \tilde{w}_{\mathbf{m}}.$$

For practical applications we must of course specify the functions $T_{\mathbf{l}}(\mathbf{k})$ more concretely. If then we consider the vector subscript $\mathbf{l}$ to stand for the three conventional quantum numbers $nlm$ of three dimensions we can think of the functions $T_{\mathbf{l}}(\mathbf{k})$ as products of a function of $k$ and a spherical harmonic

$$T_{\mathbf{l}}(\mathbf{k}) \to T_{nl}(k) Y_{lm}(\Omega_k).$$

Similarly,

$$S_{\mathbf{l}}(\mathbf{r}) \to S_{nl}(r) Y_{lm}(\Omega_r) \quad \text{and} \quad \tilde{w}_{\mathbf{l}} \to w_{nl}.$$

Now we consider specifically the Sturmian set for the square well

$$u(r) = \begin{cases} 0, & 0 < r < 1 \\ 1, & r > 1. \end{cases}$$

We shall be particularly interested in the case $l = 0$ for which the functions $S_{n0}$ are readily found to be, with $\alpha_n = (\tilde{w}_{n0} - \tilde{\beta}^2)^{\frac{1}{2}}$, $\alpha_n \cot \alpha_n = -\tilde{\beta}$,

$$S_{n0} = \begin{cases} N_{n0} \sin \alpha_n r/r, & r < 1 \\ N_{n0} \sin \alpha_n \exp \tilde{\beta}(1 - r)/r, & r > 1. \end{cases} \quad \text{(A15)}$$

The normalization constant $N_{n0}$ is determined from the relation (A14) which reads

$$\int_0^1 |S_{n0}(r)|^2 r^2\, dr = 1, \quad \text{(A16)}$$

and from which we find

$$N_{n0} = [2/(1 - (\sin 2\alpha_n / 2\alpha_n))]^{\frac{1}{2}} \quad \text{(A17)}$$

From the result (A15) we also find the expression for $T_{n0}(k)$,

$$T_{n0}(k) = \frac{2(\beta^2 + \alpha_n^2)}{\sqrt{\pi}\left(1 - \frac{\sin 2\alpha_n}{2\alpha_n}\right)^{\frac{1}{2}} k(\alpha_n + k)(\beta^2 + k^2) \times \left\{ \frac{\alpha \sin(\alpha_n - k)}{(\alpha_n - k)} - \sin \alpha_n \cos k \right\}}.$$