JOURNAL OF
MATHEMATICAL
PHYSICS

VOLUME 7, NUMBER 5 MAY 1966

Generalized Phase-Space Distribution Functions*

LEON COHEN
Yale University, New Haven, Connecticut
(Received 20 September 1965)

A set of quasi-probability distribution functions which give the correct quantum mechanical marginal distributions of position and momentum is studied. The phase-space distribution does not have to be bilinear in the state function. The Wigner distribution is a special case. A general relationship between the phase-space distribution functions and the rule of associating classical quantities to quantum mechanical operators is derived. This allows the writing of correspondence rules at will, of which the ones presently known are particular cases. The dynamics and other properties of the generalized phase-space distribution are considered.

1. INTRODUCTION

In recent years the so-called phase-space formulation of quantum mechanics has found many applications, particularly in statistical mechanics[1] and in the study of the coherent properties of light.[2] Its basic feature is to permit one to calculate expectation values of quantum mechanical observables in the classical manner rather than through the operator formalism of quantum mechanics. That is, if

$$\langle \mathbf{G} \rangle = \int \psi^*(q) \mathbf{G}(\mathbf{q}, \mathbf{p}) \psi(q)\, dq \qquad (1.1)$$

is the expected value of the operator $\mathbf{G}$ than one attempts to write this as

$$\langle \mathbf{G} \rangle = \iint g(q, p) F(q, p)\, dq\, dp, \qquad (1.2)$$

where $g(q, p)$ is the classical function from which the quantum mechanical $\mathbf{G}$ is obtained and $F(q, p)$ is the "distribution" function.[3]

The literature has dealt almost exclusively with the Wigner[4] distribution function. It is given by

$$F_w(q, p, t) = \frac{1}{2\pi} \int \psi^*(q - \tfrac{1}{2}\tau\hbar, t) \times e^{-i\tau p} \psi(q + \tfrac{1}{2}\tau\hbar, t)\, d\tau. \qquad (1.3)$$

An integration over q or p yields the correct quantum mechanical marginal distributions. Unfortunately the Wigner distribution has some shortcomings. It can take on negative values and thus cannot be considered a true distribution. For this reason some authors have called it a quasi-probability distribution. More seriously, the Wigner distribution does not always yield the same expectation values as the correct quantum mechanical methods. For example, if one calculates the standard deviation of energy of the first excited state of the simple harmonic oscillator, using the classical Hamiltonian, a nonzero value is obtained. This is clearly in contradiction with the notion of an energy eigenstate.

* Research sponsored in part by the U. S. Air Force Office of Scientific Research.

[1] For a general review of the Wigner distribution function the article "The Wigner Function and Transport Theory" by H. More, R. Oppenheim, and J. Ross which appears in *Studies in Statistical Mechanics*, V. De Boer and G. E. Uhlenbeck, Eds. (North–Holland Publishing Company, Amsterdam, 1962), Vol. 1, may be consulted. The article contains many references.

[2] L. Mandel and E. Wolf, Rev. Mod. Phys. 37, 231 (1965).

[3] We restrict ourselves to a single dimension. Generalization to higher dimensions is direct. All integrals go from $-\infty$ to ∞.

[4] E. Wigner, Phys. Rev. 40, 749 (1932).

Another distribution given by Margenau and Hill[5] in their work on joint probability distributions and later studied in detail by Mehta[6] is

$$F(q, p, t) = \frac{1}{4\pi} \text{ Real } \psi(q, t) \times \int e^{-i\tau p}\psi^*(q - \tau\hbar)\, d\tau. \quad (1.4)$$

This distribution does give zero for the dispersion of energy for the ground and first excited states of the harmonic oscillator (although it gives nonzero values for the other eigenstates).

In the following we shall present a wide class of distribution functions of which the above two are special cases. The set of functions we propose to consider is

$$F(q, p, t; f) = \frac{1}{4\pi^2} \iiint e^{-i\theta q - i\tau p + i\theta u} f(\theta, \tau, t) \times \psi^*(u - \tfrac{1}{2}\tau\hbar, t)\psi(u + \tfrac{1}{2}\tau\hbar, t)\, d\theta\, d\tau\, du, \quad (1.5)$$

where $f(\theta, \tau, t)$ is any function[7] which satisfies

$$f(0, \tau, t) = f(\theta, 0, t) = 1. \quad (1.6)$$

It is readily seen that if $f = 1$, $\cos \frac{1}{2}\theta\tau\hbar$, we obtain (1.3) or (1.4) respectively. An integration of (1.5) over p or q yields the correct marginal distributions:

$$\int F(q, p, t; f)\, dp = |\psi(q, t)|^2, \quad (1.7)$$

$$\int F(q, p, t; f)\, dq = |\phi(p, t)|^2, \quad (1.8)$$

where $\phi(p, t)$ is the momentum state function. It should be noticed that f can be a functional of the state function itself. For example

$$f(\theta, \tau, t) = \int \psi(q - \theta\tau, t)\psi^*(q + \theta\tau, t)\, dq$$

satisfies (1.6). Thus F need not be bilinear in ψ.

2. CHARACTERISTIC FUNCTION AND CORRESPONDENCE RULES

The characteristic function $M(\theta, \tau, t)$ is defined as the Fourier transform of the distribution function.

$$M(\theta, \tau, t) = \iint F(q, p, t)e^{i\theta q + i\tau p}\, dq\, dp = \langle e^{i\theta q + i\tau p}\rangle, \quad (2.1)$$

which is the expected value of $e^{i\theta q + i\tau p}$. For a well-behaved distribution function, $M(\theta, \tau, t)$ contains as much information as the distribution function itself since (2.1) can be inverted to yield

$$F(q, p, t) = \frac{1}{4\pi^2} \iint M(\theta, \tau, t)e^{-i\theta q - i\tau p}\, d\theta\, d\tau = \frac{1}{4\pi^2} \iint \langle e^{i\theta q + i\tau p}\rangle e^{-i\theta q - i\tau p}\, d\theta\, d\tau. \quad (2.2)$$

By expanding $e^{i\theta q + i\tau p}$ of (2.1), $M(\theta, \tau, t)$ can be expressed in terms of the moments,

$$M(\theta, \tau, t) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(i\theta)^n (i\tau)^m}{n!\, m!} \langle q^n p^m\rangle_t. \quad (2.3)$$

Thus the problem of finding a quantum mechanical F reduces to finding the quantum mechanical equivalent of the classical quantity $e^{i\theta q + i\tau p}$ or $q^n p^m$. Rules which associate classical quantities to quantum mechanical operators are called correspondence rules or rules of association. There have been five such rules proposed. We list them here for completeness.

(a) Dirac's Rule of associating commutators with Poisson brackets.

(b) Von Neumann's Rules.

(c) Weyl's Rule.

$$e^{i\theta q + i\tau p} \rightarrow e^{i\theta \mathbf{q} + i\tau \mathbf{p}} \quad (2.4)$$

or equivalently as shown by McCoy[8]

$$q^n p^m \rightarrow \frac{1}{2^n} \sum_{l=0}^{n} \binom{n}{l} \mathbf{q}^{n-l}\mathbf{p}^m\mathbf{q}^l. \quad (2.5)$$

(d) Symmetrization Rule

$$e^{i\theta q + i\tau p} \rightarrow \tfrac{1}{2}\{e^{i\theta \mathbf{q}}e^{i\tau \mathbf{p}} + e^{i\tau \mathbf{p}}e^{i\theta \mathbf{q}}\} \quad (2.6)$$

or

$$q^n p^m \rightarrow \tfrac{1}{2}\{\mathbf{q}^n\mathbf{p}^m + \mathbf{p}^m\mathbf{q}^n\}. \quad (2.7)$$

(e) Rule of Born and Jordan[9]

$$q^n p^m \rightarrow \frac{1}{m+1} \sum_{l=0}^{m} \mathbf{p}^{m-l}\mathbf{q}^n\mathbf{p}^l. \quad (2.8)$$

The first four rules were considered by Schewell[10] and others. They have shown that (a) and (b) are self-contradictory. The Rule of Born and Jordan seems to have been forgotten. Moyal[11] has derived the Wigner distribution using the Weyl rule of association and Margenau and Hill have used the Symmetrization Rule to obtain (1.4). A straightforward calculation yields

$$F = \frac{2}{\hbar}\frac{1}{4\pi^2} \iiint \frac{e^{-i\tau p - i\theta q + i\theta u}}{\theta\tau} \sin \tfrac{1}{2}\theta\tau\hbar \times \psi^*(u - \tfrac{1}{2}\tau\hbar, t)\psi(u + \tfrac{1}{2}\tau\hbar, t)\, d\theta\, d\tau\, du, \quad (2.9)$$

[5] H. Margenau and R. N. Hill, Progr. Theoret. Phys. (Kyoto) **26**, 722 (1961).

[6] C. L. Mehta, J. Math. Phys. **5**, 677 (1964).

[7] f is assumed to be sufficiently well behaved so that the integrations in (1.5) can be interchanged. If f has the further property that $f^*(\theta, \tau) = f(-\theta, -\tau)$ than F will be real and the operator **G** of Sec. 2 will be Hermitian.

[8] N. H. McCoy, Proc. U. S. Natl. Acad. Sci. **18**, 674 (1932).

[9] M. Born and P. Jordan, Z. Phys. **34**, 873 (1925).

[10] J. R. Shewell, Am. J. Phys. 27, 16 (1959).

[11] J. E. Moyal, Proc. Cambridge Phil. Soc. **45**, 99 (1949).

when the rule of Born and Jordan is used. This is a special case of (1.5) with $f = \sin(\frac{1}{2}\theta\tau\hbar)/\frac{1}{2}\theta\tau\hbar$. We shall now derive a fundamental relation which exists between correspondence rules and the general phase-space distribution (1.5). This will allow us to write, at will, correspondence rules of which the above are special cases.

Let $g(q, p)$ be a classical function and $\mathbf{G}(\mathbf{q}, \mathbf{p})$ the quantum mechanical operator corresponding to it. When (1.5) is used, the expected value of g is

$$\langle g(q, p)\rangle = \iint F(q, p, t)g(q, p)\, dq\, dp \tag{2.10}$$

$$= \iiint \gamma(\theta, \tau)f(\theta, \tau)e^{i\theta u}\psi^*(u - \tfrac{1}{2}\tau\hbar) \times \psi(u + \tfrac{1}{2}\tau\hbar)\, d\theta\, d\tau\, du, \tag{2.11}$$

where we have set

$$\gamma(\theta, \tau) = \frac{1}{4\pi^2}\iint g(q, p)e^{-i\theta q - i\tau p}\, dq\, dp. \tag{2.12}$$

Now consider the expression

$$\int \psi^*(q)e^{i\theta q + i\tau p}\psi(q)\, dq, \tag{2.13}$$

which equals

$$e^{\frac{1}{2}i\theta\tau\hbar}\int \psi^*(q)e^{i\theta q}\psi(q + \tau\hbar)\, dq, \tag{2.14}$$

since

$$e^{i\theta q + i\tau p} = e^{\frac{1}{2}i\theta\tau\hbar}e^{i\theta q}e^{i\tau p} \tag{2.15}$$

and

$$e^{i\tau p}\psi(q) = \psi(q + \tau\hbar).$$

Making the substitution $u = q + \frac{1}{2}\tau\hbar$, we see that the expression (2.13) equals

$$\int \psi^*(u - \tfrac{1}{2}\tau\hbar)e^{i\theta u}\psi(u + \tfrac{1}{2}\tau\hbar)\, du. \tag{2.16}$$

Therefore (2.11) can be written in the form

$$\langle g(q, p)\rangle = \iiint \gamma(\theta, \tau)f(\theta, \tau) \times \psi^*(q)e^{i\theta q + i\tau p}\psi(q)\, d\theta\, d\tau\, dq. \tag{2.17}$$

We would like this to equal

$$\int \psi^*(q)\mathbf{G}(\mathbf{q}, \mathbf{p})\psi(q)\, dq = \langle \mathbf{G}(\mathbf{q}, \mathbf{p})\rangle. \tag{2.18}$$

Comparison then shows that the operator $\mathbf{G}$ can be taken as

$$\mathbf{G}(\mathbf{q}, \mathbf{p}) = \iint \gamma(\theta, \tau)f(\theta, \tau)e^{i\theta q + i\tau p}\, d\theta\, d\tau \tag{2.19}$$

$$= \iint e^{\frac{1}{2}i\theta\tau\hbar}\gamma(\theta, \tau)f(\theta, \tau)e^{i\theta q}e^{i\tau p}\, d\theta\, d\tau. \tag{2.20}$$

This establishes the method of obtaining the quantum mechanical $\mathbf{G}$ from the classical g if (2.18) and (2.10) are to yield the same value.

Equation (2.20) can be cast in a more operational form. Assume f to be analytic and expand it in a power series

$$f(\theta, \tau) = \sum_{rs} a_{rs}\theta^r\tau^s. \tag{2.21}$$

Equation (2.20) then becomes

$$\sum_{rsn}\frac{a_{rs}}{n!}(\tfrac{1}{2}i\hbar)^n\iint \theta^{n+r}\tau^{n+s}\gamma(\theta, \tau)e^{i\theta q}e^{i\tau p}\, d\theta\, d\tau. \tag{2.22}$$

But from 2.12

$$(\partial^{2n+r+s}/\partial q^{n+r}\partial p^{n+s})g(q, p) = i^{2n+r+s} \times \iint \theta^{n+r}\tau^{n+s}\gamma(\theta, \tau)e^{i\theta q + i\tau p}\, d\theta\, d\tau \tag{2.23}$$

or, upon inverting,

$$\theta^{n+r}\tau^{n+s}\gamma(\theta, \tau) = \frac{1}{i^{2n+r+s}}\frac{1}{4\pi^2} \times \iint \left\{\frac{\partial^{2n+r+s}}{\partial q^{n+r}\,\partial p^{n+s}}\, g(q, p)\right\}e^{-i\theta q - i\tau p}\, dq\, dp. \tag{2.24}$$

Substituting (2.24) in (2.22) we find

$$\mathbf{G}(\mathbf{q}, \mathbf{p}) = \frac{1}{4\pi^2}\sum_{rsn}\frac{a_{rs}(\tfrac{1}{2}i\hbar)^n}{n!\, i^{2n+r+s}} \times \iiiint \left\{\frac{\partial^{2n+r+s}}{\partial q'^{n+r}\,\partial p'^{n+s}}\, g(q', p')\right\} \times e^{-i\theta q' - i\tau p'}e^{i\theta q}e^{i\tau p}\, d\theta\, d\tau\, dq'dp'. \tag{2.25}$$

This expression for $\mathbf{G}$ is in normal form, that is, the $\mathbf{q}$ factors precede the $\mathbf{p}$ factors. We can therefore substitute ordinary variables q and p for the operators $\mathbf{q}$ and $\mathbf{p}$, perform the integration and resubstitute $\mathbf{q}$ and $\mathbf{p}$ for q and p, after the resulting integration has been put in a form so that the q factors precede the p factors. Carrying this out gives

$$f\left(\frac{1}{i}\frac{\partial}{\partial q}, \frac{1}{i}\frac{\partial}{\partial p}\right)e^{-\frac{1}{2}i\hbar(\partial^2/\partial q\partial p)}g(q, p). \tag{2.26}$$

Thus, to obtain a correspondence rule, chose any f which satisfies (1.6), and calculate (2.26). Then arrange the result so that the q factors precede the p factors and substitute $\mathbf{q}$ for q and $\mathbf{p}$ for p. The correspondence rules of Weyl, the rule of symmetrization, and the rule of Born and Jordan are obtained by taking f equal to 1, $\cos\frac{1}{2}\theta\tau\hbar$, $\sin(\frac{1}{2}\theta\tau\hbar)/\frac{1}{2}\theta\tau\hbar$, respectively.

3. PHASE-SPACE EIGENFUNCTIONS

If $\psi(q)$ is expanded in a complete orthonormal set $\{\varphi_n(q)\}$

$$\psi(q) = \sum_{n=0}^{\infty} a_n \varphi_n(q), \tag{3.1}$$

then the phase-space distribution function (1.5) takes the form

$$F(q, p, t; f) = \sum_{nm} a_n^*(t) a_m(t) h_{nm}(q, p; f), \tag{3.2}$$

where h_{nm} are the phase-space eigenfunctions and are given by

$$h_{nm}(q, p; f) = \frac{1}{4\pi^2} \iiint e^{-i\theta q - i\tau p + i\theta u} \times f(\theta, \tau) \varphi_n^*(u - \tfrac{1}{2}\tau\hbar) \varphi_m(u + \tfrac{1}{2}\tau\hbar)\, d\theta\, du\, d\tau. \tag{3.3}$$

In general the phase-space eigenfunctions will not form a complete orthogonal set. But if $|f| = 1$, the h_{nm}'s do form such a set in the space of q and p. It is straightforward to show that

$$\iint h_{nm}(q, p) h^*_{n'm'}(q, p)\, dq\, dp = (1/2\pi\hbar)\, \delta_{nn'} \delta_{mm'}, \tag{3.4}$$

$$\sum_{nm} h_{nm}(q, p) h^*_{nm}(q', p') = (1/2\pi\hbar)\, \delta(q - q') \delta(p - p'), \tag{3.5}$$

$$\iint h_{nm}(q, p)\, dp\, dq = \delta_{nm}, \tag{3.6}$$

$$\sum_{n} h_{nm}(q, p) = 1/2\pi\hbar. \tag{3.7}$$

4. MIXTURES

Up to now we have considered pure states given by the wavefunction ψ. In the case of a mixture the distribution function becomes

$$F(q, p, t; f) = \sum_{k=0}^{\infty} P_k F^{(k)}(q, p, t; f), \tag{4.1}$$

where P_k is the probability of having $\psi^{(k)}$, and $F^{(k)}$ is obtained from (1.5) with $\psi^{(k)}$ instead of ψ. If each $\psi^{(k)}$ is expanded in a complete orthonormal set φ_n then (4.1) becomes

$$F(q, p, t; f) = \sum_{knm} P_k a_n^{*(k)} a_m^{(k)} h_{nm}(q, p; f) \tag{4.2}$$

$$= \sum_{nm} \rho_{mn} h_{nm}(q, p; f), \tag{4.3}$$

where ρ_{nm} is the density matrix.

A necessary and sufficient condition for a given F to describe a pure state is

$$F(q, p) = \iiiint F(q', p') F(q'', p'') \times g(q' - q, q'' - q, p' - p, p'' - p)\, dq'\, dq''\, dp'\, dp'', \tag{4.4}$$

where

$$g = \frac{\hbar}{8\pi^3} \iiiint e^{i\theta(q''-q) + i\theta'(q'-q) + i\tau(p''-p) + i\tau'(p'-p)} \times e^{-\frac{1}{2}i\theta\tau'\hbar + \frac{1}{2}i\theta'\tau\hbar} f^*(\theta, \tau) f^*(\theta', \tau') \times f(\theta + \theta', \tau + \tau')\, d\theta\, d\theta'\, d\tau\, d\tau'. \tag{4.5}$$

To show that (4.4) is necessary, we use the well-known fact that

$$\rho^2 = \rho \tag{4.6}$$

is a necessary and sufficient condition for the existence of a pure state. From (4.3) and (3.4) we have

$$\rho_{mn} = 2\pi\hbar \iint F(q, p) h^*_{nm}(q, p)\, dq\, dp. \tag{4.7}$$

Substituting (4.7) in (4.6)

$$\iint F(q, p) h^*_{nm}(q, p)\, dq\, dp = 2\pi\hbar \sum_k \int F(q, p) \times F(q', p') h^*_{km}(q, p) h^*_{nk}(q', p')\, dq\, dq'\, dp\, dp'. \tag{4.8}$$

Multiplying by $h_{nm}(q'', p'')$ and summing over n and m we obtain, using (3.5),

$$F(q, p) = (2\pi h)^2 \sum_{knm} \int F(q', p') F(q'', p'') \times h^*_{km}(q'', p'') h^*_{nk}(q'p') h_{nm}(q, p)\, dq'\, dq''\, dp'\, dp''.$$

But some algebra yields

$$(2\pi h)^2 \sum_{knm} h^*_{km}(q'', p'') h^*_{nk}(q'p') h_{nm}(q, p) = g(q' - q, q'' - q, p' - p, p'' - p) \tag{4.9}$$

and therefore the necessity is established. To show that (4.4) is also sufficient, we write

$$F(q, p) = \sum_{nm} \rho_{mn} h_{nm}(q, p) \tag{4.10}$$

and substitute in (4.4)

$$\sum_{nm} \rho_{mn} h_{nm}(q, p) = \sum_{nmn'm'} \int \rho_{mn} \rho_{m'n'} h_{nm}(q'', p'') h_{n'm'}(q'p') \times g(q' - q, q'' - q, p' - p, p'' - p)\, dq'\, dq''\, dp'\, dp''. \tag{4.11}$$

But

$$\int h_{nm}(q'', p'')h_{n'm'}(q', p')g\, dq'\, dq''\, dp'\, dp'' = \delta_{nm'}h_{n'm}(q, p). \tag{4.12}$$

Thus

$$\sum_{nm} \rho_{mn}h_{nm}(q, p) = \sum_{nmn'm'} \rho_{mn}\rho_{m'n'}\delta_{nm'}h_{n'm}(q, p) = \sum_{mnn'} \rho_{mn}\rho_{nn'}h_{n'm}(q, p)$$

or

$$\rho_{mn} = \sum_{n'} \rho_{mn'}\rho_{n'n} = \rho^2_{mn},$$

which is a sufficient condition for a pure state.

5. DYNAMICS

A direct but somewhat lengthy calculation gives

$$\frac{\partial F(q, p, t; f)}{\partial t} = \frac{2}{\hbar} f^{-1}\left(i\frac{\partial}{\partial q_F}, i\frac{\partial}{\partial p_F}\right) f\left(-i\frac{\partial}{\partial q_H}, -i\frac{\partial}{\partial p_H}\right) f\left(i\frac{\partial}{\partial q_F} + i\frac{\partial}{\partial q_H}, i\frac{\partial}{\partial p_F} + i\frac{\partial}{\partial q_H}\right)$$

$$\times \sin\frac{\hbar}{2}\left[\frac{\partial}{\partial p_F}\frac{\partial}{\partial q_H} - \frac{\partial}{\partial p_H}\frac{\partial}{\partial q_F}\right]H(q, p)F(q, p, t; f) + \frac{\dot{f}\left(-i\frac{\partial}{\partial q_F}, -i\frac{\partial}{\partial p_F}\right)}{f\left(-i\frac{\partial}{\partial q_F}, -i\frac{\partial}{\partial p_F}\right)} F(q, p, t) \tag{5.1}$$

for the equation of motion of the generalized phase-space distribution. $H(q, p)$ is the classical Hamiltonian. $\partial/\partial q_H$, $\partial/\partial p_H$ operates on H only and $\partial/\partial q_F$, $\partial/\partial p_F$ on F. (5.1) reduces to the classical equation of motion if f is taken to be a function of $\hbar$ such that

$$\lim_{\hbar\to 0} f = 1$$

and

$$\lim_{\hbar\to 0} \dot{f} = 0.$$

We shall now derive the temporal transformation functions for the characteristic and distribution function in terms of the quantum mechanical Green's function. Let

$$M(\theta, \tau, t) = \iint L(\theta, \tau, t \mid \theta', \tau', t') \times M(\theta', \tau', t)\, d\theta'\, d\tau', \tag{5.2}$$

$$F(q, p, t) = \iint K(q, p, t \mid q', p', t') \times F(q', p', t')\, dq'\, dp', \tag{5.3}$$

$$\psi(q, t) = \int G(q, t \mid q', t')\psi(q', t')\, dq'. \tag{5.4}$$

From (1.5) and (2.1) we obtain for the characteristic function

$$M(\theta, \tau, t) = f(\theta, \tau)\int \psi^*(u - \tfrac{1}{2}\tau\hbar, t)e^{i\theta u} \times \psi(u + \tfrac{1}{2}\tau\hbar, t)\, du. \tag{5.5}$$

Inserting (5.4) into (5.5)

$$M(\theta, \tau, t) = f(\theta, \tau, t)\iiint e^{i\theta u}G^*(u - \tfrac{1}{2}\tau\hbar, t \mid q', t') \times G(u + \tfrac{1}{2}\tau\hbar, t \mid q, t')\psi^*(q', t')\psi(q, t)\, du\, dq\, dq'. \tag{5.6}$$

Making the substitution

$$q = u' + \tfrac{1}{2}\tau'\hbar,$$
$$q' = u' - \tfrac{1}{2}\tau'\hbar,$$

and inserting

$$1 = \frac{1}{2\pi}\iint e^{-i\theta'(u'-u'')}\, d\theta'\, du'',$$

we find that

$$M(\theta, \tau, t) = \frac{\hbar}{2\pi} f(\theta, \tau, t)\iiint\int e^{i\theta u - i\theta' u'} \times G(u + \tfrac{1}{2}\tau\hbar, t \mid u' + \tfrac{1}{2}\tau'\hbar, t') \times G^*(u - \tfrac{1}{2}\tau\hbar, t \mid u' - \tfrac{1}{2}\tau'\hbar, t') \times [M(\theta', \tau', t)/f(\theta', \tau', t')]\, d\theta'\, d\tau'\, du\, du'. \tag{5.7}$$

Comparing with (5.2) we have

$$L(\theta, \tau, t \mid \theta'\tau', t') = \frac{\hbar}{2\pi}\frac{f(\theta, \tau, t)}{f(\theta', \tau', t')}\iint e^{i\theta u - i\theta' u'} \times G(u + \tfrac{1}{2}\tau\hbar, t \mid u' + \tfrac{1}{2}\tau'\hbar, t') \times G^*(u - \tfrac{1}{2}\tau\hbar, t \mid u' - \tfrac{1}{2}\tau'\hbar, t')\, du\, du'. \tag{5.8}$$

A similar procedure applied to (5.3) gives

$$K(q, p, t \mid q', p', t') = \frac{1}{4\pi^2} \iiiint L(\theta, \tau, t \mid \theta', \tau', t') \times e^{i\theta' q' + i\tau' p' - i\theta q - i\tau p}\, d\theta\, d\tau\, d\theta'\, d\tau'. \quad (5.9)$$

6. CONCLUSION

In conclusion we would like to mention some general features of the phase-space formulation. As mentioned in the introduction, the Wigner distributoin has been widely applied. It may be of some interest to repeat some of these calculations using (1.5) to determine how sensitive the results are to the choice of f.

It is commonly held that the uncertainty principle by itself precludes the possibility of the existence of a joint distribution of position and momentum. However, this is not so. For example, the choice

$$f(\theta, \tau, t) = \frac{\iint |\psi(q)|^2\, |\phi(p)|^2\, e^{i\theta q + i\tau p}\, dq\, dp}{\int \psi^*(u - \tfrac{1}{2}\tau\hbar)\, e^{i\theta u} \psi(u + \tfrac{1}{2}\tau\hbar)\, du} \quad (6.1)$$

leads to

$$F(q, p, t) = |\psi(q)|^2\, |\phi(p)|^2, \quad (6.2)$$

which is certainly a well-defined joint distribution, and from which the uncertainty principle follows in the usual manner. The reason why a true joint distribution cannot be defined is because no choice of f yields a distribution which gives the correct quantum mechanical expectation values for all observables when calculated through phase-space integration. That is, no f exists such that, if the correspondence of quantum to classical variables derived in Sec. 2 is

$$g(q, p) \rightarrow \mathbf{G} \quad (6.3)$$

for some f, then also

$$H(g(q, p)) \rightarrow H(\mathbf{G}) \quad (6.4)$$

for the same f, where H is any function.

ACKNOWLEDGMENT

It is a pleasure for the author to thank Professor Henry Margenau for his kind help and guidance.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 7, NUMBER 5 MAY 1966

The Invariance of the Vacuum is the Invariance of the World

SIDNEY COLEMAN*
CERN, Geneva, Switzerland
(Received 30 August 1965)

The following theorem is demonstrated: If the vacuum is invariant under the group generated by the space integral of the time component of a local vector current, then the Hamiltonian is invariant also. A similar theorem holds for the group generated by the space integral of a space component of a local axial-vector current.

IN quantum field theory, we often encounter groups of transformations generated by the space integral of the fourth component of a local vector current. We know such groups generate invariances of the system if they leave the Hamiltonian invariant, for from this we can deduce that the vacuum is an invariant state, and that the transformations of field operators induced by the group are invariances of the Wightman functions. It is the purpose of this note to point out that all this follows from an apparently much more restricted assumption: we prove that, if the vacuum is invariant under such a group, then so is the Hamiltonian.[1]

Proof: If the vacuum is invariant under the group' then the generator of the group must annihilate the vacuum. That is to say,[2]

$$\int V_0(\mathbf{x}, t)\, d^3\mathbf{x}\, |0\rangle = 0. \tag{1}$$

Let $|n\rangle$ be an arbitrary state with vanishing 3-momentum. Then,

$$\langle n| \int V_0(\mathbf{x}, t)\, d^3\mathbf{x}\, |0\rangle = 0. \tag{2}$$

But because of the way we have chosen the momentum, this is equivalent to

$$\langle n|\, V_0(x)\, |0\rangle = 0, \tag{3}$$

which is the same as

$$\langle n|\, \partial^\mu V_\mu(x)\, |0\rangle = 0. \tag{4}$$

But Lorentz invariance tells us that, if this is true in any one Lorentz frame, it is true in all Lorentz frames. Since any momentum eigenstate may be obtained by applying a Lorentz transformation to a state of zero 3-momentum, Eq. (4) is valid with any momentum eigenstate on the left. Since these are a complete set,

$$\partial^\mu V_\mu(x)\, |0\rangle = 0. \tag{5}$$

But it is known[3] that any local operator which annihilates the vacuum is in fact zero. Thus,

$$\partial^\mu V_\mu = 0. \tag{6}$$

This implies that

$$\int V_0(\mathbf{x}, t)\, d^3\mathbf{x}$$

is independent of time, i.e., commutes with the Hamiltonian.

Recently, there has been some interest[4] in groups of transformations generated by the space integrals of the space components of axial vector currents. The same method can be used to show that, if the vacuum is invariant under such a group, then the Hamiltonian is also. Equation (6) is replaced by

$$\partial_\mu A_\nu - \partial_\nu A_\mu = 0. \tag{7}$$

Otherwise the proof is substantially the same.

* Alfred Sloan Research Fellow. Permanent Address: Lyman Laboratory, Harvard University, Cambridge, Massachusetts.

[1] We assume the usual axioms of field theory. See, for example, R. Streater and A. Wightman, *PCT, Spin and Statistics, and All That*, (W. A. Benjamin, Inc., New York, 1964). In addition, we assume that there are no zero-mass particles.

[2] It is easy to show that this integral exists, at least with a dense set of states (quasi-local states) on the left. Unfortunately, our subsequent manipulations involve, not quasi-local states, but momentum eigenstates; therefore the argument presented here is not a rigorous proof. I am indebted to K. Hepp for a discussion of this point. *Note added in proof:* A rigorous proof has been constructed by G.-F. Dell'Antonio (unpublished).

[3] Reference 1, Theorem 4–3.

[4] R. Feynman, M. Gell-Mann, and G. Zweig, Phys. Rev. Letters **13**, 678 (1964); B. W. Lee, *ibid.* **14**, 676 (1965); R. Dashen and M. Gell-Mann, Phys. Letters **17**, 142, 145 (1965).

Adiabatic Invariants and the "Third" Integral

GEORGE CONTOPOULOS
Institute for Space Studies, Goddard Space Flight Center, NASA, New York
AND
Thessaloniki University, Thessaloniki, Greece
(Received 23 September 1965)

In a general case of Hamiltonian systems of n degrees of freedom, depending periodically on time, n formal "third" integrals of motion are found. Their application in finding boundaries for the orbits is illustrated in a special case. Then a comparison is made between these integrals and the adiabatic invariants. Both are series expansions but the small parameter used is of different character in each case. This is shown explicitly in a simple example and the relative accuracy of the two expansions is discussed.

I. INTRODUCTION

It is well known that, under certain conditions, adiabatic invariants are constant to all orders in the small parameter,[1] i.e., they are formal integrals of motion.

On the other hand, in most time-independent Hamiltonian systems of n degrees of freedom, one can find n formal integrals of motion,[2] as power series in the coordinates and momenta. We call any such integral a "third" integral, to distinguish it from the classical energy and angular momentum integrals.

The same method can be applied to a general case of n degrees of freedom, when the potential is expressed as a series in the coordinates, and is periodic in time. Then n formal integrals are found, which are periodic in time.

There are a few similarities between these integrals and the adiabatic invariants, but there are some important differences also. These are illustrated in the last section, where both a "third" integral and an adiabatic invariant are constructed for the same dynamical system.

II. INTEGRALS OF MOTION IN PERIODIC POTENTIALS

Suppose that a potential is given as a series in the coordinates, beginning with terms of second degree, and periodic with respect to the time, with period $2\pi/\omega$, then the Hamiltonian

$$H = H_2 + H_3 + \cdots , \tag{1}$$

is also periodic in t, and it has the origin as equilibrium point. We will consider the case when the characteristic exponents of the equilibrium solution are pure imaginary and not equal. This is a most common case in applications. Then we can use a linear transformation of the variables, with coefficients periodic in t, and find a new Hamiltonian of the same form, where

$$H_2 = \sum_{i=1}^{n} \frac{\omega_i}{2} (x_i'^2 + y_i'^2) \tag{2}$$

in the new variables.[3] If we introduce further new coordinates and momenta,

$$\begin{aligned} x_i &= x_i'/(\omega_i)^{\frac{1}{2}}, \\ y_i &= (\omega_i y_i')^{\frac{1}{2}}, \end{aligned} \tag{3}$$

we find

$$H_2 = \frac{1}{2} \sum_{i=1}^{n} (\omega_i^2 x_i^2 + y_i^2), \tag{4}$$

i.e., the second-order Hamiltonian is independent of time and represents a system of n harmonic oscillators. Then H_3, H_4, $\cdots$ are homogeneous polynomials in x_i, y_i, of degree 3, 4, $\cdots$, periodic in t, with period $2\pi/\omega$; hence the coefficients of the different terms can be given in the form

$$\sum c_{\cos}^{\sin} m\omega t \Pi x_i^{a_i} y_i^{b_i}$$

with integer m, a_i, b_i, and c constant.

Let us assume further that no relation of the form

$$m_1\omega_1 + m_2\omega_2 + \cdots + m\omega = 0 \tag{5}$$

exists with integer m_1, m_2, and m equal to any of the above given values. If m takes also the value 0, we assume that (5) is not satisfied, unless $m_1 = m_2 = \cdots = 0$.

[1] See M. Kruskal, J. Math. Phys. **3**, 806 (1962) and references there.

[2] G. Contopoulos, Z. Astrophys. 49, 273 (1960); Astron. J. 68, 1 (1963).

[3] A. Liapounoff, Ann. Fac. Sci. Toulouse, 2nd Ser. 9, 281, 398 (1907).

Then we can construct, step by step, n integrals in the form of series

$$\Phi_i = \Phi_{i2} + \Phi_{i3} + \Phi_{i4} + \cdots , \tag{6}$$

where

$$\Phi_{i2} = \tfrac{1}{2}(\omega_i^2 x_i^2 + y_i^2), \tag{7}$$

and $\Phi_{i\nu}$ is a homogeneous polynomial of degree ν in x_i, y_i, periodic in t, with period $2\pi/\omega$.

In fact, any integral (6) satisfies the equation

$$\frac{\partial \Phi_i}{\partial t} + (\Phi_i, H) \equiv \frac{\partial \Phi_i}{\partial t} + \sum_{i=1}^{n} \left(\frac{\partial \Phi_i}{\partial x_i}\frac{\partial H}{\partial y_i} - \frac{\partial \Phi_i}{\partial y_i}\frac{\partial H}{\partial x_i} \right) = 0, \tag{8}$$

which can be split into the equations

$$\partial \Phi_{i2}/\partial t + (\Phi_{i2}, H_2) = 0, \tag{9}$$

$$\frac{\partial \Phi_{i,\nu+1}}{\partial t} + (\Phi_{i,\nu+1}, H_2) + (\Phi_{i,\nu}, H_3) + (\Phi_{i,\nu-1}, H_4) + \cdots = 0. \tag{10}$$

Equation (10) is a linear partial differential equation that gives $\Phi_{i,\nu+1}$ when previous terms of the series Φ_i are known. The corresponding system to this equation is

$$dt = \frac{dx_i}{y_i} = \frac{dy_i}{-\omega_i^2 x_i} = \frac{d\Phi_{i,\nu+1}}{K_{i,\nu}}, \tag{11}$$

where the function

$$K_{i,\nu} = -(\Phi_{i,\nu}, H_3) - (\Phi_{i,\nu-1}, H_4) - \cdots \tag{12}$$

is known, and is of degree $\nu + 1$.

This gives

$$\begin{aligned} x_i &= [(2\Phi_{i2})^{\frac{1}{2}}/\omega_i] \sin \omega_i(t - t_i), \\ y_i &= (2\Phi_{i2})^{\frac{1}{2}} \cos \omega_i(t - t_i). \end{aligned} \tag{13}$$

Then

$$\Phi_{i,\nu+1} = \int K_{i,\nu}\, dt, \tag{14}$$

where $K_{i,\nu}$ is written in the form

$$\sum q(2\Phi_{1*})^{\frac{1}{2}|m_1|+n_1}(2\Phi_{2*})^{\frac{1}{2}|m_2|+n_2} \cdots \begin{matrix}\sin\\ \cos\end{matrix} [m_1\omega_1(t - t_1) + m_2\omega_2(t - t_2) + \cdots + m\omega t], \tag{15}$$

where m_1, m_2, n_1, n_2 are integers, $n_1 > 0$, $n_2 > 0$, $\frac{1}{2}(|m_1| + |m_2| + \cdots) + n_1 + n_2 + \cdots = \nu + 1$, and q are constants. If no coefficient, (5) is zero, then Eq. (14) is integrated and gives $\Phi_{i,\nu+1}$ as a sum of the same form as (15). This can be expressed as a polynomial of degree $\nu + 1$ in x_i, y_i, with coefficients of the form $\begin{matrix}\sin\\ \cos\end{matrix} m\omega t$.

If H contains time-independent terms, the corresponding m is zero. Then, if $K_{i,\nu}$ contains a term with a cosine and $m_1 = m_2 = \cdots = 0$, this will give a secular term in $\Phi_{i,\nu+1}$. It can be proved, however, as in the two dimensional case,[2] that $K_{i,\nu}$ never includes cosines with $m_1 = m_2 = \cdots = 0$. Therefore n formal integrals can always be constructed, step by step.

If for certain values of m_1, m_2, $\cdots$ m, the corresponding quantity (5) is zero, the above integrals are no more valid. Then, however, the system (11) has further the integrals

$$\begin{matrix} S_M \\ C_M \end{matrix} = (2\Phi_{1*})^{\frac{1}{2}|m_1|}(2\Phi_{2*})^{\frac{1}{2}|m_2|} \cdots \begin{matrix}\sin\\ \cos\end{matrix} [m_1\omega_1(t - t_1) + m_2\omega_2(t - t_2) + \cdots + m\omega t], \tag{16}$$

which are polynomials in x_i, y_i of degree $M = |m_1| + |m_2| + \cdots$. Then one can construct two integrals of the form

$$\begin{aligned} S &= S_M + S_{M+1} + \cdots , \\ C &= C_M + C_{M+1} + \cdots , \end{aligned} \tag{17}$$

which will have also secular terms.

In a simple nonlinear case it has been proved[4] that a combination of these integrals with the above integrals Φ_i can eliminate the secular terms and give formal time-independent integrals of motion. These resonance integrals may be rather different from the above integrals (6). It seems probable that one can find such resonance integrals also in the case of time-dependent potentials of the form (1).

In many problems the Hamiltonian is given in the form

$$H = \frac{1}{2}\sum_{i=1}^{n} (\omega_i^2 x_i^2 + y_i^2) + \epsilon H_*, \tag{18}$$

where ϵ is a small parameter and H_* is of degree > 2. Then we find integrals of the form

$$\Phi_i = \bar{\Phi}_{i(0)} + \epsilon\bar{\Phi}_{i(1)} + \epsilon^2\bar{\Phi}_{i(2)} + \cdots , \tag{19}$$

where $\bar{\Phi}_{i(0)} = \Phi_{i2}$ and

$$\bar{\Phi}_{i,(\nu+1)} = -\int (\bar{\Phi}_{i,(\nu)}, H_*)\, dt. \tag{20}$$

The integrals Φ_i are useful in giving bounds for

[4] G. Contopoulos, Astron. J. 68, 763 (1963).

the orbits. For example, in a two-dimensional time-independent potential, the boundary $f(x_1, x_2) = 0$ of an orbit is found by eliminating y_1, y_2 between Eq. (19) and

$$\frac{\partial \Phi_1}{\partial y_1}\frac{\partial \Phi_2}{\partial y_2} - \frac{\partial \Phi_1}{\partial y_2}\frac{\partial \Phi_2}{\partial y_1} = 0. \tag{21}$$

In the case of a time-dependent potential, we take the set of the curves f when t takes all the values within a period. If these curves are closed, their outer boundary defines the space inside which the orbit is confined. These curves are in zero-order parallelograms. Therefore, if ϵ is sufficiently small the curves do not extend to infinity, and the orbits are confined for all times.

This is exactly true if the integrals Φ_i are convergent. In general, their convergence is unknown. However, if H is given, one can find another Hamiltonian coinciding with H up to the terms of any given degree, which has convergent integrals in a region around the origin.[5]

It seems probable that even when Φ_i are not convergent, the orbits will not go to infinity if ϵ is sufficiently small.

III. APPLICATION

As an example of the general case we consider the potential

$$V = \tfrac{1}{2}(\omega_1^2 x_1^2 + \omega_2^2 x_2^2) - \epsilon \sin \omega t \quad x_1 x_2^2. \tag{22}$$

Then

$$H = H_0 + \epsilon H_\epsilon = \tfrac{1}{2}(y_1^2 + y_2^2 + \omega_1^2 x_1^2 + \omega_2^2 x_2^2) - \epsilon \sin \omega t \quad x_1 x_2^2, \tag{23}$$

where $y_i = dx_i/dt$.

One can find now two integrals of the form (6), namely

$$\Phi_1 = \tfrac{1}{2}(\omega_1^2 x_1^2 + y_1^2) + \epsilon\left(\frac{(y_2^2 + \omega_2^2 x_2^2)}{2\omega_2^2(\omega^2 - \omega_1^2)}(\omega \cos \omega t \quad y_1 + \omega_1^2 \sin \omega t \quad x_1)\right.$$
$$- \frac{1}{2\omega_2^2[(\omega^2 - \omega_1^2 - 4\omega_2^2)^2 - 16\omega_1^2\omega_2^2]}\{\omega \cos \omega t[(\omega^2 - \omega_1^2 - 4\omega_2^2)y_1(y_2^2 - \omega_2^2 x_2^2) - 8\omega_1^2\omega_2^2 x_1 x_2 y_2]$$
$$\left. + \sin \omega t[\omega_1^2(\omega^2 - \omega_1^2 + 4\omega_2^2)x_1(y_2^2 - \omega_2^2 x_2^2) + 4\omega_2^2(\omega^2 + \omega_1^2 - 4\omega_2^2)y_1 x_2 y_2]\}\right) + \cdots, \tag{24}$$

and

$$\Phi_2 = \tfrac{1}{2}(\omega_2^2 x_2^2 + y_2^2) - \frac{2\epsilon}{\omega_2^2[(\omega^2 - \omega_1^2 - 4\omega_2^2)^2 - 16\omega_1^2\omega_2^2]}\{\omega \cos \omega t[2\omega_2^2 y_1(y_2^2 - \omega_2^2 x_2^2) - \omega_2^2(\omega^2 - \omega_1^2 - 4\omega_2^2)x_1 x_2 y_2]$$
$$+ \omega_2^2 \sin \omega t[(\omega^2 + \omega_1^2 - 4\omega_2^2)x_1(y_2^2 - \omega_2^2 x_2^2) + (\omega^2 - \omega_1^2 + 4\omega_2^2)y_1 x_2 y_2]\} + \cdots. \tag{25}$$

The sum of the intergrals Φ_1 and Φ_2 is

$$\Phi = \Phi_1 + \Phi_2 = \tfrac{1}{2}(y_1^2 + y_2^2 + \omega_1^2 x_1^2 + \omega_2^2 x_2^2) - \epsilon \sin \omega t \quad x_1 x_2^2$$
$$+ \epsilon\left(\frac{(y_2^2 + \omega_2^2 x_2^2)(\omega \cos \omega t \quad y_1 + \omega^2 \sin \omega t \quad x_1)}{2\omega_2^2(\omega^2 - \omega_1^2)} + \frac{1}{2\omega_2^2[(\omega^2 - \omega_1^2 - 4\omega_2^2)^2 - 16\omega_1^2\omega_2^2]}\right.$$
$$\times \{-\omega \cos \omega t[(\omega^2 - \omega_1^2 + 4\omega_2^2)y_1(y_2^2 - \omega_2^2 x_2^2) - 4\omega_2^2(\omega^2 + \omega_1^2 - 4\omega_2^2)x_1 x_2 y_2]$$
$$\left. - \omega^2 \sin \omega t[(\omega^2 - \omega_1^2 - 4\omega_2^2)x_1(y_2^2 - \omega_2^2 x_2^2) + 8\omega_2^2 y_1 x_2 y_2]\}\right) + \cdots. \tag{26}$$

The last integral is similar to the Hamiltonian, but there is an extra first-order term in it.

A number of orbits have been calculated in this potential, with $\omega_1^2 = 0.076$, $\omega_2^2 = 0.55$, $\epsilon = 0.206$. We use these values in order to be able to compare the orbits with those calculated previously in a model of the galactic potential near the sun,[6] where

$$H = \tfrac{1}{2}(y_1^2 + y_2^2 + \omega_1^2 x_1^2 + \omega_2^2 x_2^2) - \epsilon x_1 x_2^2 - \tfrac{2}{3}\epsilon' x_1^3. \tag{27}$$

Figures 1 and 2 give four orbits with the same initial velocity at the origin ($y_{10} = 0.0512$, $y_{20} = 0.1126$),

[5] The proof is the same as in the time-independent case, G. Contopoulos, Astrophys. J. 138, 1297 (1963).
[6] G. Contopoulos, Stockholm Obs. Ann. 20, No. 6 (1958). See also Ref. 2.

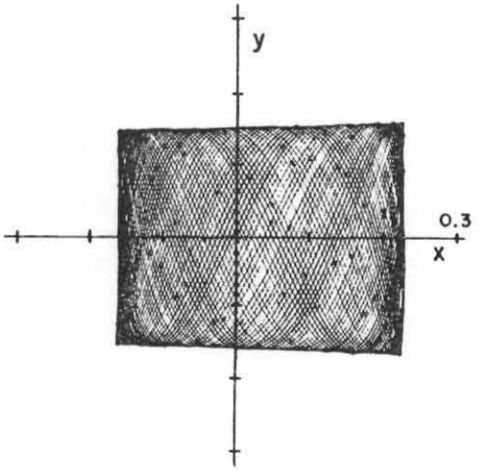

FIG. 1. Orbit in the potential $V = \frac{1}{2}(\omega_1^2 x_1^2 + \omega_2^2 x_2^2) - \epsilon x_1 x_2^2$ for $\omega_1^2 = 0.076$, $\omega_2^2 = 0.55$, $\epsilon = 0.206$, and initial conditions $x_{10} = 0$, $y_{10} = 0.0512$, $y_{20} = 0.1126$.

one for the Hamiltonian (27) with $\epsilon' = 0$ and three for the Hamiltonian (23) with $\omega = 0.1$, 1 and 10.

The calculations were made for 600 time units at least with the Runge–Kutta method in double precision and a step 0.02 and/or 0.01 time units. A comparison of the results has shown that at least seven decimal figures are always accurate.

The variations of the Hamiltonian and of the "third" integral Φ in first approximation are given in Table I. For comparison the corresponding quantities in the case of the Hamiltonian (27) are given. There we know that inclusion of higher-order terms gives a very accurate third integral.[7] The initial value of $2H$ is always the same (0.0153). It is seen that the "third" integral is always better conserved than the Hamiltonian.

The boundaries of the orbits in the time-dependent cases are oscillating, especially in Fig. 2(b). These boundaries can be found as follows: Eq. (21) gives

$$y_1 y_2 = O(\epsilon), \tag{28}$$

hence either $y_1 = O(\epsilon)$, or $y_2 = O(\epsilon)$. In the first

TABLE I. Values of the Hamiltonian and the first-order "third" integral.

	$2H_{min}$	$2H_{max}$	Φ_{init}	Φ_{min}	Φ_{max}
1 Time-indep.	0.0153	0.0153	0.00131	0.00129	0.00139
2a $\omega = 0.1$	0.0146	0.0153	0.01492	0.01478	0.01492
2b $\omega = 1$	0.0132	0.0178	0.0148	0.0147	0.0161
2c $\omega = 10$	0.0137	0.0163	0.015299	0.015296	0.015299

[7] B. Barbanis, Z. Astrophys. 56, 56 (1962).

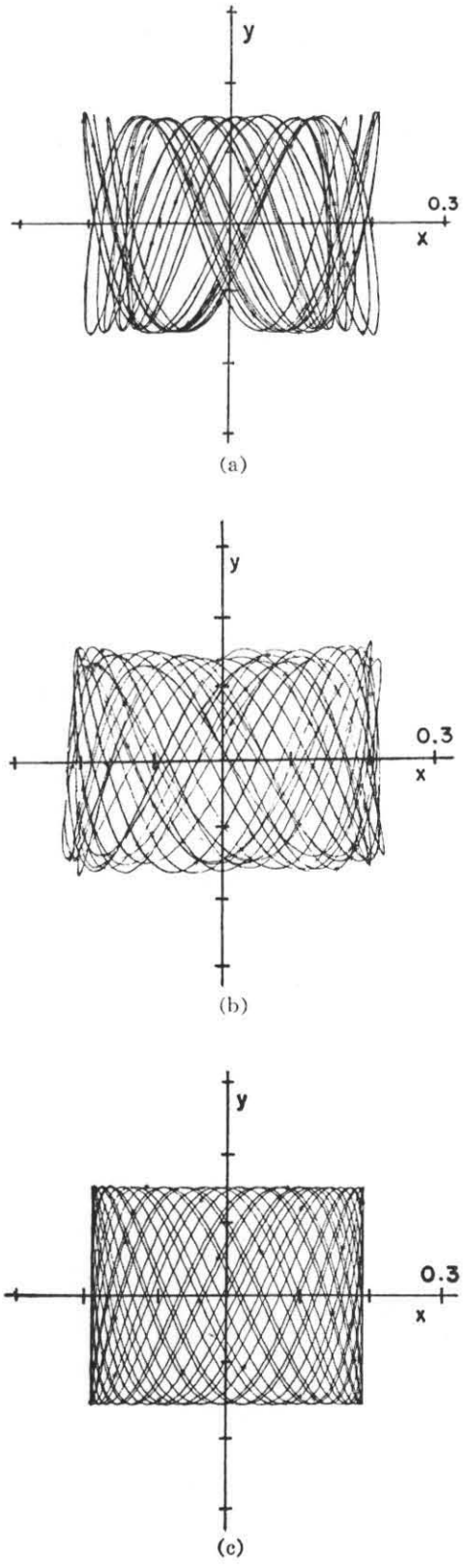

FIG. 2. Orbits in the potential $V = \frac{1}{2}(\omega_1^2 x_1^2 + \omega_2^2 x_2^2) - \epsilon \sin \omega t \; x_1 x_2^2$ for the same constants and initial conditions as in Fig. 1, and (a) $\omega = 0.1$, (b) $\omega = 1$, (c) $\omega = 10$.

case we have in first approximation

$$\Phi_1 = \tfrac{1}{2}\omega_1^2 x_1^2 + \epsilon\left\{\frac{2\Phi_{2;0}\omega_1^2 \sin \omega t \quad x_1}{2\omega_2^2(\omega^2 - \omega_1^2)}\right.$$

$$- \frac{1}{2\omega_2^2[(\omega^2 - \omega_1^2 - 4\omega_2^2)^2 - 16\omega_1^2\omega_2^2]}$$

$$\times [-8\omega\omega_1^2\omega_2^2 \cos \omega t \quad x_1x_2y_2 + \omega_1^2(\omega^2 - \omega_1^2 + 4\omega_2^2)$$

$$\left.\times \sin \omega t \quad x_1(2\Phi_{2;0} - 2\omega_2^2x_2^2)]\right\} = \Phi_{1;0}, \qquad (29)$$

where $y_2 = \pm(2\Phi_{2;0} - \omega_2^2x_2^2)^{\frac{1}{2}}$, and $\Phi_{1;0}$, $\Phi_{2;0}$ are the values of the integrals at the initial point.

In the second case

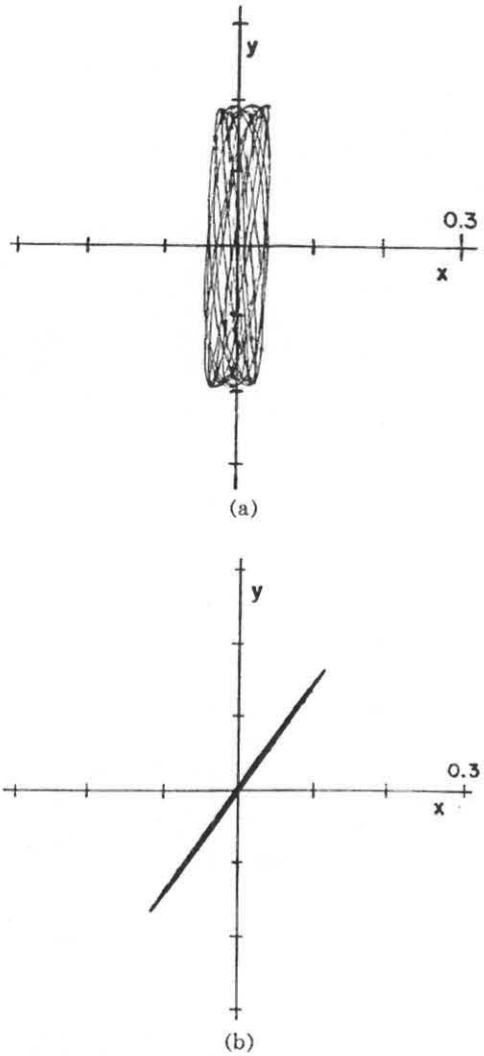

FIG. 3. Orbits in the same potential as in Fig. 2, for $\omega_1^2 = \omega_2^2 = \epsilon = 0.1$, $\omega = 1$ and initial conditions $x_{10} = x_{20} = 0$ and (a) $y_{10} = 0.013$, $y_{20} = 0.060465$; (b) $y_{10} = 0.035$, $y_{20} = 0.05099$.

$$\Phi_2 = \tfrac{1}{2}\omega_2^2x_2^2 + \frac{4\epsilon\Phi_{2;0}}{[(\omega_1^2 - \omega_1^2 - 4\omega_2^2)^2 - 16\omega_1^2\omega_2^2]}$$

$$\times [2\omega \cos \omega t \quad y_1 + (\omega^2 + \omega_1^2 - 4\omega_2^2) \sin \omega t \quad x_1]$$

$$= \Phi_{2;0}, \qquad (30)$$

where $y_1 = \pm(2\Phi_{1;0} - \omega_1^2x_1^2)^{\frac{1}{2}}$. Hence the boundaries are near the straight lines $x_1 = \pm(2\Phi_{1;0})^{\frac{1}{2}}/\omega_1$ and $y_2 = \pm(2\Phi_{2;0})^{\frac{1}{2}}/\omega_2$, varying periodically in time.

The deviations are larger when we are near the resonances $\omega = \omega_1$, $\omega = \pm 2\omega_2 \pm \omega_1$, or any higher-order resonance. In the above cases $\omega_1 = 0.27568$, $\omega_2 = 0.74162$, hence $2\omega_2 + \omega_1 \simeq 1.76$, $2\omega_2 - \omega_1 \simeq 1.21$. The last quantity is nearest to $\omega = 1$; this is the reason why the boundaries change more in case 2b. The "third" integral is better conserved when $\omega = 10$, because this value is far from resonances. Then the space filled by the orbit is very nearly a parallelogram.

Similar results were found in many other cases. Fig. 3 represents two orbits in the case $\omega_1^2 = \omega_2^2 = 0.1$, $\epsilon = 0.1$, $\omega = 1$. The orbits are rather different from those of the corresponding time-independent resonance case (27), except for the orbit 3b, which is near a periodic orbit. In general, one expects that the resonance effects of the time-independent case, which depend on the value of ω_1/ω_2, do not affect the time-dependent case, which depends on some relation of ω_1, ω_2 with ω. In the present case, it happens that $2\omega_2 + \omega_1 = 3 \times (0.3162) = 0.949$ is near $\omega = 1$, but this near resonance is of a different nature than the resonance $\omega_1 = \omega_2$. The subject is worthy of further study.

The initial conditions in the above cases are $x_{10} = y_{20} = 0$, and $y_{10} = 0.013$, $y_{20} = 0.060465$ in case 3a and $y_{10} = 0.035$, $y_{20} = 0.05099$ in case 3b. We find in case 3a: $2H_{\text{init}} = 0.003825$, $2H_{\text{min}} = 0.00336$, $2H_{\text{max}} = 0.00417$, $\Phi_{\text{init}} = 0.00319$, $\Phi_{\text{min}} = 0.00297$, $\Phi_{\text{max}} = 0.00319$, and in case 3b: $2H_{\text{init}} = 0.003825$, $2H_{\text{min}} = 0.00186$, $2H_{\text{max}} = 0.00450$, $\Phi_{\text{init}} = 0.00261$, $\Phi_{\text{min}} = 0.00240$, $\Phi_{\text{max}} = 0.00261$. It is seen that although we are near a resonance, the first-order third integral is conserved much better than the energy.

Our experience, from the time-independent cases, indicates that the conservation of the "third" integral is always improved (to a very high accuracy) as more higher-order terms are included.

IV. COMPARISON WITH THE ADIABATIC INVARIANTS

The most simple adiabatic invariants are given for one-dimensional Hamiltonians that vary *slowly*

in time. If H depends on the time through λ, where $d\lambda/dt$ is small with respect to λ/T (T is the period of the motion), then the action integral $J = \oint y_1\, dx_1$ is approximately constant.[8] Usually H is considered to vary only during a finite time interval, being constant before and after that. In a system of n degrees of freedom, n adiabatic invariants can be found when the system is separable.[8]

If H is not separable but is a slowly-varying function of the time and of some variables besides x_1, y_1, then the action $J = \oint y_1\, dx_1$ is an adiabatic invariant under certain conditions specified below.

Gardner[9] gives a general method to construct adiabatic invariants by successive approximations. We describe his method, and the conditions under which it is applicable.

We assume that H is a function of $t, x_1, y_1, x_2, y_2 \cdots$, and of a small parameter ω that satisfies the following conditions:

(1) In zero order (i.e., for $\omega = 0$) it is separable.

(2) It is a slowly varying function of the time and of the variables x_2, x_3, $\cdots$ (and eventually, but not necessarily of y_2, y_3 $\cdots$) i.e., it is a function of ωx_2, ωx_3, $\cdots$ (and, possibly, of ωy_2, ωy_3, $\cdots$), i.e.,

$$H = H_0(x_1, y_1, \omega x_2, y_2, \cdots, \omega t) + H_0^*(y_2, \omega x_2, \cdots, \omega t), \tag{31}$$

where H_0 includes all terms containing x_1, y_1, and has no zero-order terms in y_2, $\cdots$.

(3) The curves $H = \text{const}$ for $\omega = 0$ and $y_i = \text{const}$ $(i > 1)$ are closed; then they are closed also for small values of ω and fixed t, x_i, y_i $(i > 1)$.

An adiabatic invariant is constructed, step by step, by successive coordinate transformations.

If we keep t, x_i, and y_i $(i > 1)$ constant, the curves $H = \text{const}$ are transformed into circles by the following area preserving transformation

$$x_1' = r\cos\theta, \qquad y_1' = r\sin\theta, \tag{32}$$

where $r = (J/\pi)^{\frac{1}{2}}$

and

$$\theta = 2\pi \int^{x_1, y_1} ds/|\nabla H| \Big/ \oint ds/|\nabla H|;$$

J is the area inside the curve $H = \text{const}$ and ds the line element along this curve. The curves $H = \text{const}$ are circles in the new variables.

Because of the conservation of areas, $x_1'\, dy_1' + y_1\, dx_1$ is the complete differential of a generating function F_1 and

$$x_1' = \partial F_1/\partial y_1', \qquad y_1 = \partial F_1/\partial x_1. \tag{33}$$

F_1 is a function of x_1, y_1', with coefficients functions of ωt, ωx_2, y_2, $\cdots$, which are kept constant in the transformation (32). As we can add any arbitrary constant in F_1 we may write

$$F_1 = x_2 y_2' + \cdots + \psi_1(x_1, y_1', \omega x_2, y_2', \cdots \omega t), \tag{34}$$

where ψ_1 does not contain terms independent of x_1, y_1'.

Then the equations

$$\begin{aligned} x_2' &= \partial F_1/\partial y_2' = x_2 + \partial\psi_1/\partial y_2', \\ y_2 &= y_2' + \omega\partial\psi_1/\partial(\omega x_2), \\ &\cdots, \end{aligned} \tag{35}$$

together with Eqs. (33) define new canonical variables x_i', y_i', with Hamiltonian

$$H' = H + \partial F_1/\partial t = H + \omega\partial\psi_1/\partial(\omega t), \tag{36}$$

which is of the form

$$H' = H_1(x_1'^2 + y_1'^2, \omega t) + H_1^*(y_2', \omega x_2', \cdots, \omega t) + \omega\bar{H}_1(x_1', y_1', \omega x_2', y_2', \cdots, \omega t). \tag{37}$$

The term $\omega\bar{H}_1$ contains all higher-than-zero-order terms including x_1', y_1'.

The next change of variables transforms the curves $H_1 + \omega\bar{H}_1 = \text{const}$ into circles for t, x_i, y_i $(i > 1)$ constant. The deviation of these curves from circles is of order ω, therefore the difference between the two sets of variables will be order ω.

The new generating function can be written

$$F_2 = x_1' y_1'' + x_2' y_2'' + \cdots + \omega\psi_2(x_1', y_1'', \omega x_2', y_2'', \cdots \omega t), \tag{38}$$

where ψ_2 does not contain terms independent of x_1', y_1''.

Then

$$x_1'' = x_1' + \omega\partial\psi_2/\partial y_1'', \quad y_1' = y_1'' + \omega\partial\psi_2/\partial x_1'', \tag{39}$$

$$\begin{aligned} \omega x_2'' &= \omega x_2' + \omega^2\partial\psi_2/\partial y_2'', \\ y_2' &= y_2'' + \omega^2\partial\psi_2/\partial(\omega x_2'), \\ &\cdots, \end{aligned} \tag{40}$$

and the new Hamiltonian is

$$H'' = H_2(x_1''^2 + y_1''^2, \omega t) + H_2^*(y_2'', \omega x_2'', \cdots \omega t) + \omega^2\bar{H}_2(x_1'', y_1'', \omega x_2'', y_2'', \cdots \omega t). \tag{41}$$

[8] L. D. Landau and E. M. Lifschitz, *Mechanics* (Pergamon Press, New York, 1960).

[9] C. S. Gardner, Phys. Rev. 115, 791 (1959).

By repeating this procedure N times we find

$$H^{(N)} = H_N[(x_1^{(N)})^2 + (y_1^{(N)})^2, \omega t] + H_N^*(y_2^{(N)}, \omega x_2^{(N)}, \cdots, \omega t) + \omega^N \bar{H}_N(x_1^{(N)}, y_1^{(N)}, \omega x_2^{(N)}, y_2^{(N)}, \cdots, \omega t). \quad (42)$$

Then the quantity

$$J^{(N)} = [(x_1^{(N)})^2 + (x_2^{(N)})^2] \quad (43)$$

is constant to order $N - 1$, i.e., $dJ^{(N)}/dt = O(\omega^N)$. Therefore if $N \to \infty$, $J^{(N)}$ is a formal integral of motion; this is the adiabatic invariant.

If, however, both x_2 and y_2 appear in H (not through ωx_2, ωy_2) in nonlinear terms, then the above method is not applicable.

For example, in a two-dimensional system, let x_2, y_2 appear in nonlinear zero-order terms of H_0^* and let x_2 appear also in a mixed term (including x_1 and/or y_1), of degree n in ω.

Then

$$y_2 = y_2' + \partial\psi_1/\partial x_2,$$

and $\partial\psi_1/\partial x_2$ includes terms of degree n, containing x_1 and/or y_1', because, by its definition, F_1 does not contain terms independent of x_1, y_1' except x_2y_2'. Thus the expansion of H_0^* gives n-order mixed terms (containing y_2').

The next change of variables gives

$(x_2'' = x_2' + n)$-order terms

containing x_1'' and/or $y_1' + \cdots$,

and H_1^* contains again n-order mixed terms. By n such changes of variables, we can reduce all terms containing x_1 and/or y_1 up to order $n - 1$ to a function of $(x_1^{(n)})^2 + (y_1^{(n)})^2$. The $(n + 1)$ transformation of variables, however, cannot eliminate the n-order terms that include $x_1^{(n)}$, $y_1^{(n)}$; because in the new Hamiltonian the zero-order terms of H_N^* will give again mixed n-order terms including $x_1^{(n+1)}$ and/or $y_1^{(n)}$.

In Gardner's paper[9], the Hamiltonian considered is of the form (31), but the above conditions are not explicitly stated.

The adiabatic invariant $J^{(N)}$ is equal to the action J if H does not depend on the time and on the variables x_2, $y_2 \cdots$. Therefore if for $t < t_1$ and $t > t_2$ the variables x_i, y_i are such that H has zero derivatives of all orders with respect to t, x_2, y_2, $\cdots$, then the action $J = \oint y_1\, dx_1$ is well defined for $t < t_1$ and $t > t_2$; its change during time $t_2 - t_1$ is of order higher than any ω^N [it is at least of order $\exp(-\alpha/\omega)$, where α is a constant].[10,11]

In general, however, there is no time interval during which $H = \text{const}$ and no space where H is independent of x_2, $y_2 \cdots$.

In the case (27), we know that we can make the system separable by a formal variable transformation, known as the von Zeipel method. [This method gives the third integral in a somewhat different way than described above.][5] If we set $\epsilon' = 0$ we find explicitly a generating function

$$S = x_1y_1' + x_2y_2' + \frac{\epsilon}{2}\left[\frac{-y_1(y_2^2 + \omega_2^2x_2^2)}{\omega_1^2\omega_2^2} + \frac{y_1(y_2^2 - \omega_2^2x_2^2) + 4\omega_2^2x_1x_2y_2}{\omega_2^2(\omega_1^2 - 4\omega_2^2)}\right] + \cdots \quad (44)$$

that gives

$$y_1' = y_1 - 2\epsilon x_2y_2/(\omega_1^2 - 4\omega_2^2) + \cdots \quad (45)$$

etc. Then

$$J = 2\int_{x_1\min}^{x_1\max} y_1\, dx_1 = 2\int_{x_1\min}^{x_1\max} y_1^2\, dt = 2\int_{x_1\min}^{x_1\max}\left(y_1'^2 + \frac{4\epsilon y_1'x_2'y_2'}{\omega_1^2 - 4\omega_2^2} + \cdots\right) dt. \quad (46)$$

The quantity

$$2\int_{x_1\min}^{x_1\max} y_1'^2\, dt$$

is a constant, with an error of order higher than the first; in fact, we notice that it is exactly constant if $x_1\min$ and $x_2\max$ are replaced by $x_1'\min$ and $x_1'\max$, and that the value of y_1' for $x_1'\min$, $x_1'\max$ is zero. But the quantity

$$\epsilon\int_{x_1\min}^{x_1\max} y_1'x_2'y_2'\, dt$$

is not constant, in general. Therefore J has variations of the first order in ϵ. Thus the adiabatic invariant $J^{(N)}(N \to \infty)$ is more general than the action $J = \oint y_1\, dx_1$.

A comparison of the adiabatic invariants in the form $J^{(N)}$ with the "third" integral shows the following:

(a) Both are formal series expansions in terms of a small parameter (or some small parameters). When the "third" integral is given in power series in the variables and no small parameter formally appears, we may consider the energy as small

[10] F. Hertweck and A. Schlüter, Z. Naturforsch. 12a, 899 (1957).

[11] P. Vandervoort, Ann. Phys. 12, 436 (1961).

parameter. In fact, an expansion (27) may be written in dimensionless form

$$\omega_1^2\left[\frac{x_1}{(2H)^{\frac{1}{2}}}\right]^2+\left[\frac{y_1}{(2H)^{\frac{1}{2}}}\right]^2+\omega_2^2\left[\frac{x_2}{(2H)^{\frac{1}{2}}}\right]^2+\left[\frac{y_2}{(2H)^{\frac{1}{2}}}\right]^2-2\epsilon(2H)^{\frac{1}{2}}\left[\frac{x_1}{(2H)^{\frac{1}{2}}}\right]\left[\frac{y_2}{(2H)^{\frac{1}{2}}}\right]^2-\frac{2\epsilon'(2H)^{\frac{1}{2}}}{3}\left[\frac{x_1}{(2H)^{\frac{1}{2}}}\right]^3=1, \tag{47}$$

and the small parameters are essentially $\epsilon(2H)^{\frac{1}{2}}$ and $\epsilon'(2H)^{\frac{1}{2}}$; if ϵ and ϵ' are constant and the small parameter is $(2H)^{\frac{1}{2}}$.

The difference is that the small parameter in the case of the "third" integral refers to a *term*, or *terms*, while in the case of the adiabatic invariants it refers to a *variable*, or *variables*.

(b) The "third" integral is more general in the sense that it does not require H to depend on $\omega x_2 \cdots$, rather than $x_2 \cdots$. In the case of the Hamiltonian (27) we cannot find an adiabatic invariant, because both x_2 and y_2 appear in zero-order. On the other hand, if ω_2 is small with respect to ω_1, we may construct an adiabatic invariant expansion in powers of ω_2. This expansion is preferable, because the third-integral expansion is not valid when $\omega_1 \to 0$ or $\omega_2 \to 0$.

(c) The adiabatic invariants are more general in the sense that they apply also to nonperiodic time-dependent Hamiltonians.

(d) The practical construction of a "third" integral when H is a series is comparatively easy. The formulas for finding higher-order terms are given, and the necessary algebra can be performed by an electronic computer. In the case of the adiabatic invariants, the changes of variables that transform the curves $H_0 = \text{const}$ into circles cannot, in general, be given analytically in a simple form. In practice, one should expand in series of another small parameter also, which is the parameter that measures the deviations of the equipotential lines (in the $\omega_i x_i$, y_i plane) from circles, i.e., it is essentially the parameter used in the third integral.

On the other hand, if the Hamiltonian cannot be expanded in a power series, the third-integral method is not applicable, (except in special cases like the restricted three-body problem, etc.), while Gardner's method is in principle still valid.

We apply now both methods to the simple one-dimensional Hamiltonian

$$H = \tfrac{1}{2}(\omega_1^2 x_1^2 + y_1^2) - \epsilon \sin \omega t \quad x_1^2. \tag{48}$$

This case can be reduced to the well-known Mathieu equation[12]

$$d^2y/dz^2 + (a - 2q \sin 2z)y = 0, \tag{49}$$

by setting

$$y = x_1, \quad \omega t = 2z, \quad a = 4\omega_1^2/\omega^2, \quad q = 4\epsilon/\omega^2. \tag{50}$$

In this case the transformations proposed by Gardner can be explicitly carried out, and we can compare the adiabatic invariant directly with the "third" integral. The "third" integral in second-order approximation is

$$\Phi = \tfrac{1}{2}(\omega_1^2 x_1^2 + y_1^2) + \frac{2\epsilon[\omega \cos \omega t \quad x_1 y_1 - \sin \omega t(y_1^2 - \omega_1^2 x_1^2)]}{\omega^2 - 4\omega_1^2} + \frac{\epsilon^2}{2\omega_1^2(\omega^2 - 4\omega_1^2)}\{\cos 2\omega t(y_1^2 + \omega_1^2 x_1^2) - 2(y_1^2 - \omega_1^2 x_1^2) - \frac{1}{(\omega^2 - \omega_1^2)}[(\omega^2 + 2\omega_1^2) \cos 2\omega t(y_1^2 - \omega_1^2 x_1^2) + 6\omega\omega_1^2 \sin 2\omega t \quad x_1 y_1]\} + \cdots. \tag{51}$$

In applying Gardner's method we have to calculate the area J of the ellipse (48). Using formulas (32) we find

$$r = \frac{(2H)^{\frac{1}{2}}}{(\omega_1^2 - 2\epsilon \sin \omega t)^{\frac{1}{4}}}, \tag{52}$$

and

$$\theta = \cos^{-1}\left[\frac{(\omega_1^2 - 2\epsilon \sin \omega t)^{\frac{1}{4}}}{(2H)^{\frac{1}{2}}} x_1\right]. \tag{53}$$

Then

$$x_1' = (\omega_1^2 - 2\epsilon \sin \omega t)^{\frac{1}{4}} x_1, \tag{54}$$

$$y_1 = (\omega_1^2 - 2\epsilon \sin \omega t)^{\frac{1}{4}} y_1',$$

$$F_1 = (\omega_1^2 - 2\epsilon \sin \omega t)^{\frac{1}{4}} x_1 y_1', \tag{55}$$

and

$$H' = \tfrac{1}{2}(\omega_1^2 - 2\epsilon \sin \omega t)^{\frac{1}{2}}(x_1'^2 + y_1'^2) - \frac{\omega\epsilon \cos \omega t \quad x_1' y_1'}{2(\omega_1^2 - 2\epsilon \quad \sin \omega t)}. \tag{56}$$

The next change of variables is effected in a similar way. After some operations we find

$$F_2 = x_1' y_1'' - \frac{\omega\epsilon \cos \omega t(y_1''^2 - x_1'^2)}{8(\omega_1^2 - 2\epsilon \sin \omega t)^{\frac{3}{2}}}, \tag{57}$$

$$x_1'' = x_1' - \frac{\omega\epsilon \cos \omega t \quad y_1''}{4(\omega_1^2 - 2\epsilon \sin \quad \omega t)^{\frac{3}{2}}}, \tag{58}$$

$$y_1' = y_1'' + \frac{\omega\epsilon \cos \omega t \quad x_1'}{4(\omega_1^2 - 2\epsilon \sin \quad \omega t)^{\frac{3}{2}}},$$

[12] N. W. McLachlan, *Theory and Applications of Mathieu Functions* (Clarendon Press, Oxford, 1947), pp. 77, 90.

TABLE II. Comparison of the "third" integral with the adiabatic invariant.

ω_1	ω	ϵ	x_{10}	y_{10}	H_{init}	$\epsilon(2H)^{1/2}{}_{init}$	H_{max}	H_{min}	Φ_{max}	Φ_{min}	$J^{(3)}{}_{max}$	$J^{(3)}{}_{min}$
1.	0.2	0.1	0.	0.1	0.005	0.01	0.0055	0.0045	0.00499	0.00497	0.005000	0.004998
1.	0.2	0.2	0.	1.	0.5	0.2	0.59	0.38	0.50	0.48	0.500	0.499
1.	0.8	0.2	0.	1.	0.5	0.2	0.61	0.35	0.50	0.45	0.52	0.44
1.	1.2	0.2	0.	1.	0.5	0.2	0.87	0.44	0.58	0.53	0.72	0.48

and

$$H'' = \tfrac{1}{2}(\omega_1^2 - 2\epsilon \sin \omega t)^{\frac{1}{2}}(x_1''^2 + y_1''^2) + O(\omega^2). \tag{59}$$

Similarly, we find

$$x_1''' = x_1''\left\{1 + \frac{\omega^2\epsilon[11\epsilon \cos^2 \omega t - 4 \sin \omega t(\omega_1^2 - 2\epsilon \sin \omega t)]}{32(\omega_1^2 - 2\epsilon \sin \omega t)^3}\right\}, \tag{60}$$

$$y_1'' = y_1'''\left\{1 + \frac{\omega^2\epsilon[11\epsilon \cos^2 \omega t - 4 \sin \omega t(\omega_1^2 - 2\epsilon \sin \omega t)]}{8(\omega_1^2 - 2\epsilon \sin \omega t)^3}\right\}, \tag{61}$$

and

$$H''' = \tfrac{1}{2}(\omega_1^2 - 2\epsilon \sin \omega t)^{\frac{1}{2}}\left[1 - \frac{\omega^2\epsilon^2 \cos^2 \omega t}{32(\omega_1^2 - 2\epsilon \sin \omega t)^3}\right](x_1'''^2 + y_1'''^2) + O(\omega^3). \tag{62}$$

The adiabatic invariant in second-order approximation is

$$J^{(3)} = \pi(x_1'''^2 + y_1'''^2) = \frac{\pi}{(\omega_1^2 - 2\epsilon \sin \omega t)^{\frac{1}{2}}}\left(y_1^2 + (\omega_1^2 - 2\epsilon \sin \omega t)x_1^2 - \frac{\omega\epsilon \cos \omega t \quad x_1y_1}{(\omega_1^2 - 2\epsilon \sin \quad \omega t)}\right.$$
$$\left. + \frac{\omega^2\epsilon\{\epsilon \cos^2 \omega t[7x_1^2(\omega_1^2 - 2\epsilon \sin \omega t) - 5y_1^2] - 2 \sin \omega t(\omega_1^2 - 2\epsilon \sin \omega t)[x_1^2(\omega_1^2 - 2\epsilon \sin \omega t) - y_1^2]\}}{8(\omega_1^2 - 2\epsilon \sin \omega t)^3}\right). \tag{63}$$

Both expansions are equivalent if ϵ and ω are small. If we omit all terms of order higher than two in ϵ or ω we find

$$\Phi\left[1 + \frac{\epsilon^2(4\omega_1^2 + 3\omega^2)}{16\omega_1^2}\right] = \frac{J^{(3)}\omega_1}{2\pi}$$
$$= \frac{1}{2}\left\{y_1^2 + \omega_1^2x_1^2 + \frac{\epsilon \sin \omega t}{\omega_1^2}(y_1^2 - \omega_1^2x_1^2)\right.$$
$$- \frac{\epsilon\omega \cos \omega t}{\omega_1^2}x_1y_1 + \frac{\epsilon^2 \sin^2 \omega t}{2\omega_1^4}(3y_1^2 - \omega_1^2x_1^2)$$
$$- \frac{3\epsilon^2\omega \sin 2\omega t}{2\omega_1^4}x_1y_1 + \frac{\omega^2\epsilon \sin \omega t}{4\omega_1^4}(y_1^2 - \omega_1^2x_1^2)$$
$$\left. + \frac{\omega^2\epsilon^2}{8\omega_1^6}[7\omega_1^2x_1^2 - 5y_1^2 + \sin^2 \omega t(15y_1^2 - 13\omega_1^2x_1^2)]\right\}. \tag{64}$$

The above form of the "third"-integral expansion is preferable if ϵ is small (q small in Mathieu's equation) if we are not near a resonance ($\omega^2 = \omega_1^2$, $\omega^2 = 4\omega_1^2$, and in general $\omega^2 = 4\omega_1^2/n^2$, i.e., $a = n^2$). The resonance cases should be treated separately.

A comparison of the values of Φ and $J^{(3)}$ has been made, in some orbits calculated numerically, by the Runge–Kutta method. In all the cases $\omega_1 = 1$, and the calculations were made for 300 time units with a step 0.02 or 0.01 time units. A check has shown that at least four significant figures in x_1, y_1 and five significant figures in H, Φ, and $J^{(3)}$ are accurate. Table II gives the data and the values of the energy H, the third-integral Φ [given by formula (51)] and the adiabatic invariant [given by formula (63)].

It is seen that for small values of ω the second-order adiabatic invariant $J^{(3)}$ is better conserved than the second-order "third"-integral Φ. However, for ω/ω_1 approaching unity, Φ is better conserved than $J^{(3)}$. This is more evident for larger ω. The conservation of the zero- and first-order "third"-integral and adiabatic invariant is always worse.

This example gives the range of values of ω for which an adiabatic invariant is useful.

If ω is near a resonance case, the above formulas are no more valid, although resonances are not apparent in formula (63). In fact, in a numerical example (in the case $\omega = \omega_1 = 1$), we have found continuous increase of the amplitude of oscillations, and $J^{(3)}$ is not even approximately conserved.

This fact indicates that the action J is not an adiabatic invariant if there is a resonance between the frequency of the perturbation and the eigenfrequency of the system. This fact is mentioned by the first authors that applied the adiabatic invariants,[13] but is rarely mentioned explicitly in modern papers.

The example discussed here shows clearly the distinction between the "third" integrals and the adiabatic invariants. They are expansions in terms of different small parameters; in the case of the third integral we have a small term, while in the case of the adiabatic invariants we have a slow dependence on the time and/or some variables.

This example shows further that the relative accuracy of the two expansions depends on the values of the parameters used. It indicates also the disadvantage of the adiabatic invariants, in that they cannot be used in resonance or near resonance cases.

[13] See, e.g., A. Sommerfeld, *Atombau and Spektrallinien* (F. Vieweg & Sohn, Braunschweig, 1951), Vol. I, 7th ed., pp. 370, 698.

ACKNOWLEDGMENTS

This work was done while the author was an NAS-NRC senior research associate under the National Aeronautics and Space Administration. He wishes to thank Dr. Robert Jastrow for the use of the facilities at the Institute for Space Studies. The author thanks also Dr. K. Prendergast and Dr. J. Kevorkian for interesting discussions.

Failure of the Mayer Irreducible Cluster Theorem with Wave Mechanics*

HUGH DE WITT AND PAUL FISHBANE†
Lawrence Radiation Laboratory, University of California, Livermore, California
(Received 26 October 1965)

The wave mechanical form of the cluster integrals in the activity expansion for the pressure of nonideal gas is written in the Wiener functional integral form. The various pieces of the nth-order cluster integral may then be easily expanded around the classical limit [Wigner–Kirkwood (WK) expansion] using the procedure of Gel'fand and Yaglom. It is shown that the reducible diagrams fail to factor at $O(\hbar^4)$, and thus the Mayer theorem that only irreducible diagrams come into the virial coefficients of the density expansion of the pressure is valid only through $O(\hbar^2)$. As an example, the WK expansion of the third virial coefficient is worked out to $O(\hbar^4)$. It is shown also how the functional integral formatism may be used to expand quantum statistical mechanical perturbation theory about the classical limit.

THE statistical mechanics of a nonideal gas developed from the grand canonical partition function gives a convenient cluster expansion for the pressure and density of the gas:

$$\beta P(\beta, \mu) = \sum_{j=1} z^j b_j(\beta), \tag{1}$$

$$\rho = \sum_{j=1} j z^j b_j, \tag{2}$$

where the activity is $z = (e^{\beta\mu})(2\pi mkT)^{\frac{3}{2}}(2\pi\hbar)^{-3} = \rho e^{\beta\delta\mu}$ with $\mu = \mu_0 + \delta\mu$ the chemical potential; $\delta\mu$ depends explicitly on the interparticle interaction. Equations (1) and (2) are parametric equations in z (or $\delta\mu$). By solving Eq. (2) for z as a function of ρ, one may eliminate the chemical potential and obtain the more convenient virial expansion for the pressure

$$\beta P(\beta, \rho) = \rho + B_2\rho^2 + B_3\rho^3 + \cdots. \tag{3}$$

The cluster integrals b_j in the activity expansion are known from Mayer's work for all j and are well defined for both classical and quantum systems.[1] The diagrammatic representation of the b_j has the important topological property that the j particles of a cluster are at least singly connected. The elimination of $\delta\mu$ in going to the virial expansion, Eq. (3), gives a remarkable topological result: the diagrams for the virial coefficients are multiply connected (irreducible); the class of singly connected diagrams has been eliminated.[2] This irreducible cluster theorem due to Mayer is known to fail for systems at densities high enough so that quantum statistics are important. However, to our knowledge the limit of validity of this theorem has not been discussed for systems of distinguishable particles interacting according to the laws of wave mechanics. There is no particular reason to expect the irreducible cluster theorem to be true in general, but for some systems it would be convenient if it is approximately true. Thus with ordinary nonideal gases at intermediate temperatures, wave mechanical effects become numerically significant in the virial coefficients.[3] Also, for high-temperature plasmas, wave mechanical effects can be important, while quantum statistics effects are negligible.

The proof of the irreducible cluster expansion for classical systems depends crucially upon the factorizability of singly connected diagrams into products of multiply connected diagrams. In this paper we show by an example that the quantum generalizations of the singly connected diagrams fail to completely factor, and that the error in the near classical limit is of order $\hbar^4$.

The discussion here will be limited to the second- and third-virial coefficients, since extension to the higher-virial coefficients will be evident. The cluster integrals in Eq. (1) may be written as

$$b_1 = 1, \qquad b_2 = \tfrac{1}{2}!\,B, \qquad b_3 = \tfrac{1}{3}!\,(D + 3C), \tag{4}$$

where B denotes the sum of all interactions of two particles. C denotes three-particle interactions in which the first interacts with the second, the second with the third, but not the first with the third. D indicates three-particle interactions in which all three interact with each other. Elimination of the chemical potential from Eqs. (1) and (2) gives

* Work performed under the auspices of the U. S. Atomic Energy Commission.

† Present address: Department of Physics, Princeton University, Princeton, N. J.

[1] T. E. Hill, *Statistical Mechanics* (McGraw-Hill Book Company, Inc., New York, 1956), Chap. 5.

[2] A complete treatment of this theorem and linear graph theory applied to cluster expansion is found in the chapter by Ford and Uhlenbeck in *Studies in Statistical Mechanics* (North Holland Publishing Company, Amsterdam, 1963), Vol. I.

[3] Hirschfelder, Curtiss, and Bird, *Molecular Theory of Gases and Liquids* (John Wiley & Sons, Inc., New York, 1954), Chap. 6.

$$\beta P = \rho - \tfrac{1}{2}B\rho^2 - \tfrac{1}{3}[D + 3(C - B^2)]\rho^3 - \cdots . \quad (5)$$

The quantum mechanical definition of B for two particles of mass m_1 and m_2 is

$$B(12) = \frac{1}{V}\frac{(2\pi\hbar)^3}{(2\pi m_1 kT)^{\frac{3}{2}}}\frac{(2\pi\hbar)^3}{(2\pi m_2 kT)^{\frac{3}{2}}} \cdot \mathrm{Tr}\,\{\exp[-\beta(K_1 + K_2 + u_{12})] - \exp[-\beta(K_1 + K_2)]\} \underset{\hbar=0}{\to} \iint \frac{d^3r_1\, d^3r_2}{V}(e^{-\beta u_{12}} - 1) = \int d^3r_{12}\, f_{12}, \quad (6)$$

where K_1 and K_2 are the particle kinetic energy operators, u_{12} is the potential interaction, and $f = e^{-\beta u} - 1$ is the usual Mayer f function. Similarly, C for three particles is

$$C(123) = \frac{1}{V}\frac{(2\pi\hbar)^3}{(2\pi m_1 kT)^{\frac{3}{2}}}\frac{(2\pi\hbar)^3}{(2\pi m_2 kT)^{\frac{3}{2}}}\frac{(2\pi\hbar)^3}{(2\pi m_3 kT)^{\frac{3}{2}}} \times \mathrm{Tr}\,\{\exp - \beta(K_1 + K_2 + K_3 + u_{12} + u_{23}) - \exp[-\beta(K_1 + K_2 + K_3 + u_{12})] - \exp[-\beta(K_1 + K_2 + K_3 + u_{23})] + \exp[-\beta(K_1 + K_2 + K_3)]\} \underset{\hbar=0}{\to} \iiint \frac{d^3r_1\, d^3r_2\, d^3r_3}{V}(e^{-\beta u_{12}} - 1)(e^{-\beta u_{23}} - 1) = \left(\int d^3r_{12}\, f_{12}\right)\left(\int d^3r_{23}\, f_{23}\right). \quad (7)$$

In the classical limit, $\hbar = 0$, the quantity C factors as shown in the last line of Eq. (7), so that $C(123) = B(12)B(23)$. C is the lowest-order example of a singly connected diagram. Since it factors for $\hbar = 0$, the quantity $C - B^2$ in the third-virial coefficient vanishes, and one finds $B_3 = -\frac{1}{3}D$, where

$$D = \frac{1}{V}\iiint d^3r_1\, d^3r_2\, d^3r_3\, f_{12}f_{23}f_{31}.$$

Because of the classical factorizability of the singly connected diagrams, it could be proved that all the virial coefficients involved only multiply connected diagrams. In the classical picture the particles are points interacting via $u(r)$, but with $\hbar \neq 0$ the particles are wave packets of extension $\lambda = \hbar/(2mkT)^{\frac{1}{2}}$ and the simple classical factorization is no longer possible.

In order to investigate the extent to which the quantum form of C factors, we would like to write B and C in a form closely resembling the classical expressions as given by the second lines of Eqs. (6) and (7). Wave mechanics as indicated by the trace operation in the definition of B and C may be conveniently expressed with the help of the Wiener functional integral. For a thorough treatment of this formulation of wave mechanics into statistical mechanical problems, the reader is referred to the very comprehensive paper of Gel'fand and Yaglom,[4] and also to the review by Brush.[5] Our notation and the mathematical manipulations follow Gel'fand and Yaglom. We use the functional integral formulation to calculate the Wigner–Kirkwood[6] expansion of B and C in powers of λ^2 about the classical limit.

As a Wiener functional integral, B may be written as

$$B(12) = \iint \frac{d^3r_1\, d^3r_2}{V}\iint d^3_{w(0,1)}s_1(v)\, d^3_{w(0,1)}s_2(v)F_{12},$$

where

$$F_{12} \equiv F\{[\mathbf{r}_1 + \lambda_1\mathbf{s}_1(v)] - [\mathbf{r}_2 + \lambda_2\mathbf{s}_2(v)]\} = \exp\left\{-\beta\int_0^1 dv\, u[\mathbf{r}_{12} + \lambda_1\mathbf{s}_1(v) - \lambda_2\mathbf{s}_2(v)] - 1\right\} \underset{\hbar=0}{\to} e^{-\beta u_{12}} - 1. \quad (8)$$

F_{12} is evidently a quantum generalization of the Mayer f function. The quantities $\lambda_1\mathbf{s}_1(v)$ and $\lambda_2\mathbf{s}_2(v)$ are the deviations of the particles from the classical paths as the particles propagate from inverse temperature "time" 0 to β ($v = 0$ to $v = 1$). For the two-particle system, one can immediately make a change of variables

$$\lambda_{12}\mathbf{n}(v) = \lambda_1\mathbf{s}_1(v) - \lambda_2\mathbf{s}_2(v), \quad (9)$$

where $\lambda_{12}^2 = \lambda_1^2 + \lambda_2^2 = \hbar^2/2\mu_{12}kT$ (μ_{12} is the reduced mass) corresponding to the relative motion. The center-of-mass motion integrates out. Thus one finds

$$B(12) = \int d^3r_{12}\int d^3_{w(0,1)}\eta(v)F[\mathbf{r}_{12} + \lambda_{12}\mathbf{n}(v)]. \quad (10)$$

One now makes a Taylor expansion of F_{12} about $\mathbf{r}_{12}$,

$$F_{12} = (e^{-U} - 1) + e^{-U}\left\{-\lambda_{12}\int_0^1 dv\, \mathbf{n}(v)\cdot\nabla U + (\lambda_{12}^2/2)\left[\int_0^1\int dv_1\, dv_2\, \mathbf{n}(v_1)\mathbf{n}(v_2) : \nabla U\nabla U - \int_0^1 dv\, \mathbf{n}(v)\mathbf{n}(v) : \nabla\nabla U\right]\cdots\right\}, \quad (11)$$

[4] I. M. Gel'fand and A. M. Yaglom, J. Math. Phys. **1**, 48 (1960).
[5] S. G. Brush, Rev. Mod. Phys. **33**, 79 (1961).
[6] L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Pergamon Press, Ltd., London, 1958), pp. 96–103.

where $U = \beta u$. The functional integrals can be performed on powers of the Cartesian components of $\mathbf{n}$ (the moments of the Wiener measure); the results for say the x component are

$$\int d^3_{w(0,1)}\eta(v)\eta_x(v_1) \cdots \eta_x(v_{2k+1}) = 0,$$

$$\int d^3_{w(0,1)}\eta(v)\eta_x(v_1) \cdots \eta_x(v_{2k}) = \sum_{i_1 \cdots i_{2k}} b(v_{i_1}, v_{i_2}) \cdots b(v_{i_{2k-1}}, v_{i_{2k}}), \quad (12)$$

where the summation is over all partitions into pairs of the indices $1, 2, \cdots, 2k$ and

$$b(v_1, v_2) = 2v_1(1 - v_2), \quad v_1 \leq v_2$$
$$= 2v_2(1 - v_1), \quad v_1 > v_2.$$

Since the odd moments, i.e., integrals of $\eta_x(v)^{2k+1}$ vanish, one is left with only even powers of λ_{12}^2. The final step is the integration over the various v variables in the Taylor expansion, Eq. (11). The result due to Yaglom is[7,8]

$$B(12) = \int d^3r_{12}\, \mathfrak{F}(\mathbf{r}_{12}),$$

$$\mathfrak{F}(\mathbf{r}_{12}) = \mathfrak{F}_{12} = f_{12} + e^{-U_{12}}[\lambda_{12}^2\chi_{12} + \lambda_{12}^4\xi_{12} + \cdots],$$

$$\chi = \tfrac{1}{12}(\nabla U)^2 - \tfrac{1}{6}\nabla^2 U,$$

$$\xi = \tfrac{1}{288}(\nabla U)^4 - \tfrac{1}{72}(\nabla U)^2\nabla^2 U - \tfrac{1}{60}\nabla U\nabla U : \nabla\nabla U + \tfrac{1}{72}(\nabla^2 U)^2 + \tfrac{1}{90}\nabla\nabla U : \nabla\nabla U + \tfrac{1}{30}\nabla U\cdot\nabla\nabla^2 U - \tfrac{1}{60}\nabla^4 U. \quad (13)$$

Some integrations by parts brings this result into the more useful form

$$B(12) = \int d^3r \left\{(e^{-U} - 1) + e^{-U}\left[-\frac{\lambda_{12}^2 U'^2}{12} + \frac{\lambda_{12}^4}{120}\left(-\frac{5}{36}U'^4 + \frac{10}{3}\frac{U'^3}{r} + \frac{2U'^2}{r^2} + U''^2\right)\cdots\right]\right\}.$$

Using the functional integral notation, the quantum generalization of *any* classical cluster integral may be written down. One needs only to replace the Mayer f_{ij} functions in the classical theory with the F_{ij} defined by Eq. (8) and integrate over the path variables, $\lambda_i\mathbf{s}_i(v)$, for each particle in the cluster. Thus C as defined by Eq. (7) becomes

$$C = \iiint \frac{d^3r_1\, d^3r_2\, d^3r_3}{V} \iiint d^3_{w(0,1)}s_1(v) \cdot d^3_{w(0,1)}s_2(v)\; d^3_{w(0,1)}s_3(v) \cdot F[\mathbf{r}_{12} + \lambda_1\mathbf{s}_1(v) - \lambda_2\mathbf{s}_2(v)]F[\mathbf{r}_{23} + \lambda_2\mathbf{s}_2(v) - \lambda_3\mathbf{s}_3(v)]. \quad (14)$$

The integrand for D differs from that of C only in that $F_{12}F_{23}$ in C becomes $F_{12}F_{23}F_{31}$ in D.

The calculation of the WK expansion of any quantum cluster in integral now requires only multiplying out the Taylor expansions of the various F_{ij} functions, grouping terms of order $\hbar^2$, $\hbar^4$, etc., using Eq. (12) for the various moments, and finally performing the "time" integrations over the v's. With cluster terms involving more than two particles, such as C, this program is straightforward though tedious, partly because the transformation to relative motion variables is no longer possible. The result of the calculation of C to $O(\hbar^4)$ is

$$C = \iiint \frac{d^3r_1\, d^3r_2\, d^3r_3}{V} \cdot\{\mathfrak{F}_{12}\mathfrak{F}_{23} + e^{-(U_{12}+U_{23})}[\lambda_2^2 G_{12,23} + \lambda_2^4 H_{12,23} + \cdots]\},$$

$$G_{12,23} = G(\mathbf{r}_{12}, \mathbf{r}_{23}) = -\tfrac{1}{6}\nabla U_{12}\cdot\nabla U_{23},$$

$$H_{12,23} = H(\mathbf{r}_{12}, \mathbf{r}_{23}) = \tfrac{1}{72}(\nabla U_{12}\cdot\nabla U_{23})^2 - \tfrac{1}{60}\nabla\nabla U_{12} : \nabla U_{23}\nabla U_{23} - \tfrac{1}{60}\nabla\nabla U_{23} : \nabla U_{12}\nabla U_{12} + \tfrac{1}{45}\nabla\nabla U_{12} : \nabla\nabla U_{23} + \tfrac{1}{30}\nabla U_{12}\cdot\nabla\nabla^2 U_{23} + \tfrac{1}{30}\nabla U_{23}\cdot\nabla\nabla^2 U_{12}. \quad (15)$$

The coefficients of λ_2^2, λ_2^4, etc. indicate the extent to which $C(123)$ does not factor into the product $B(12)B(23)$. At first glance it appears that the factorizability fails already at $O(\hbar^2)$ because of the $\lambda_2^2\nabla U_{12}\cdot\nabla U_{23}$ term. However, one immediately finds that integration over the angle between ∇U_{12} and ∇U_{23} makes the coefficient of λ_2^2 equal to zero. This same statement will be true for any cluster

[7] A. M. Yaglom, Teoriya Veroyatnostei i ee Primeneniya 1, 161 (1956).

[8] The $0(\hbar^6)$ term has been worked out by T. Kihara, Y. Midzuno, and T. Shizume, J. Phys. Soc. Japan 10, 249 (1955).

diagram which factors in the classical limit. Consequently, the failure of singly connected diagrams to factor *begins with* $O(\hbar^4)$. Integration over the angles in the $O(\hbar^4)$ terms gives the error in the factorizability of C as

$$\begin{aligned}
&C(123) - B(12)B(23)\\
&= \lambda_2^4 \iint d^3r_{12}\, d^3r_{23}\, e^{-(U_{12}+U_{23})} H_{12,23}\\
&= \frac{1}{3}\lambda_2^4 \iint d^3r_{12}\, d^3r_{23}\, e^{-(U_{12}+U_{23})}\Big\{\frac{1}{72}\, U_{12}'^2 U_{23}'^2\\
&\quad - \frac{1}{60}\, U_{12}'^2\Big(U_{23}'' + \frac{2U_{23}'}{r_{23}}\Big) - \frac{1}{60}\, U_{23}'^2\Big(U_{12}'' + \frac{2U_{12}'}{r_{12}}\Big)\\
&\quad + \frac{1}{45}\Big(U_{12}'' + \frac{2U_{12}'}{r_{12}}\Big)\Big(U_{23}'' + \frac{2U_{23}'}{r_{23}}\Big)\Big\}. \qquad (16)
\end{aligned}$$

The expansion of D can be worked out in the same way. We give here the result for the complete third-virial coefficient [from Eq. (5)] written as if the three particles are different.

$$\begin{aligned}
&B_3(123) = -\tfrac{1}{3}[D(123) + C(123) - B(12)B(23)\\
&\quad + C(231) - B(23)B(31) + C(312) - B(31)B(12)]\\
&= -\frac{1}{3}\iiint \frac{d^3r_1\, d^3r_2\, d^3r_3}{V}\{\mathfrak{F}_{12}\mathfrak{F}_{23}\mathfrak{F}_{31} + e^{-(U_{12}+U_{23}+U_{31})}\\
&\quad \times [(\lambda_2^2 G_{12,23} + \lambda_3^2 G_{23,31} + \lambda_1^2 G_{31,12})\\
&\quad + (\lambda_2^4 H_{12,23} + \lambda_3^4 H_{23,31} + \lambda_1^4 H_{31,12})\\
&\quad + \lambda_1^2\lambda_2^2(G_{12,23}G_{31,12} + I_{12,3})\\
&\quad + \lambda_1^2\lambda_3^2(G_{31,12}G_{23,31} + I_{31,2})\\
&\quad + \lambda_2^2\lambda_3^2(G_{23,31}G_{12,23} + I_{23,1})\\
&\quad + \lambda_{12}^2\lambda_3^2 G_{23,31}\chi_{12} + \lambda_{23}^2\lambda_2^2\\
&\quad \cdot G_{12,23}\chi_{13} + \lambda_{23}^2\lambda_1^2 G_{31,12}\chi_{23}]\}, \qquad (17)
\end{aligned}$$

where the additional quantity in the $O(\hbar^4)$ term is

$$I_{12,3} = -\tfrac{1}{18}\nabla U_{23}\nabla U_{31}\cdot\nabla\nabla U_{12}. \qquad (18)$$

So far in this work the functional integral formulation has been used to evaluate the trace of operators occurring in an exponential, $\exp[-\beta(K+u)]$. It is worthwhile to note that the functional integral formulation works equally well for powers of the operators, such as a perturbation expansion in powers of βu. For example, the nth term in the expansion of the second-virial coefficient from Eq. (6) is[9]

$$\begin{aligned}
B_n(12) &= \frac{1}{n!}\int d^3r \int d^3_{w(0,1)}\eta(v)\\
&\quad \cdot\Big\{-\beta\int_0^1 dv\, u[\mathbf{r} + \lambda\mathbf{n}(v)]\Big\}^n. \qquad (19)
\end{aligned}$$

A Taylor expansion of $U[\mathbf{r} + \lambda\mathbf{n}(v)]$ and evaluation of the moments of the Wiener measure gives[10,11]

$$\begin{aligned}
B_n(12) &= \frac{(-1)^n}{n!}\int d^3r\\
&\quad \cdot\Big\{U^n + \lambda^2\Big[\frac{n}{6}(\nabla^2 U)U^{n-1} + \frac{n(n-1)}{12}(\nabla U)^2 U^{n-2}\Big]\\
&\quad + \lambda^4\Big[\frac{n}{60}(\nabla^4 U)U^{n-1} + n(n-1)(\tfrac{1}{30}\nabla U:\nabla\nabla^2 U\\
&\quad + \tfrac{1}{90}\nabla\nabla U:\nabla\nabla U + \tfrac{1}{72}(\nabla^2 U)^2)U^{n-2}\\
&\quad + n(n-1)(n-2)(\tfrac{1}{60}\nabla U\nabla U:\nabla\nabla U\\
&\quad + \tfrac{1}{72}(\nabla U)^2\nabla^2 U)U^{n-3}\\
&\quad + \frac{n(n-1)(n-2)(n-3)}{288}(\nabla U)^4 U^{n-4}\Big] + \cdots\Big\}. \qquad (20)
\end{aligned}$$

Summation over n gives Eq. (13). Similar results are easily obtained for the perturbation expansions of C and D.

---

[9] Note that the n different v integrations each from 0 to 1 are the sum of $n!$ different time orderings of the interactions; all the time orderings are equal in this case. Using this fact and writing out the Wiener measure, Eq. (19) becomes

$$\begin{aligned}
B_n(12) &= (-\beta)^n \int_0^{1>v_n>\cdots>v_1} \cdots \int dv_n \cdots dv_1\\
&\quad \times (4\pi)^{\frac{3}{2}} \int \cdots \int d^3\eta_1 \cdots d^3\eta_n\\
&\quad \times u(\mathbf{r} + \lambda\mathbf{n}) \cdots u(\mathbf{r} + \lambda\eta)\\
&\quad \times \frac{e^{-n_1^2/4v_1}}{(4\pi v_1)^{\frac{3}{2}}}\frac{e^{-(n_2-n_1)^2/4(v_2-v_1)}}{[4\pi(v_2-v_1)]^{\frac{3}{2}}}\cdots\\
&\quad \times \frac{e^{-(n_n-n_{n-1})^2/4(v_n-v_{n-1})}}{[4\pi(v_n-v_{n-1})]^{\frac{3}{2}}}\frac{e^{-n_n^2/4(1-v_n)}}{[4\pi(1-v_n)]^{\frac{3}{2}}},
\end{aligned}$$

a form obtained by H. S. Green, J. Chem Phys. **20**, 1274 (1952). The quantity

$$\frac{e^{-(n_{i+1}-n_i)^2/4(v_{i+1}-v_i)}}{[4\pi(v_{i+1}-v_i)]^{\frac{3}{2}}}$$

is the free-particle propagator which takes the particle from the space "time" point $\mathbf{r} + \lambda\mathbf{n}_i$, βv_i to $\mathbf{r} + \lambda\mathbf{n}_{i+1}$, βv_{i+1}.

[10] For $n = 1$ the quantum corrections are zero, since $\int d^3r\nabla^2 u = 0$; thus B_1 is completely classical, i.e., $B_1 = -\int d^3r U$, as it should be.

[11] For nth-order perturbation theory the functional integral is very much easier to use than the method given in an earlier paper; H. E. De Witt, J. Math. Phys. **3**, 1003 (1962).

Adiabatic Switching in the Schrödinger Theory of Scattering

JOHN DAY DOLLARD*
Palmer Physical Laboratory, Princeton, N. J.
(Received 23 August 1965)

The time-dependent Schrödinger theory of scattering is studied rigorously with the potential V replaced by $e^{-\epsilon|t|}V$. Sufficient conditions are given that the Møller wave-matrices $\Omega^{\pm}$ be obtainable from this theory as $\epsilon \to 0$. The conjecture that this theory can be used to define a reasonable S-matrix when the $\Omega^{\pm}$ do not exist is shown to be false for the Coulomb potential. It is remarked that the Goldberger–Gell-Mann switching procedure also breaks down in this case.

INTRODUCTION

IN dealing with problems in the nonrelativistic theory of scattering, it is frequently convenient to switch the interaction off adiabatically by one means or another.[1] This paper studies the switching procedure in which the potential V is replaced by $e^{-\epsilon|t|}V$. As far as the author has been able to tell, no previous rigorous presentation of this procedure has been given in the literature.

The idea of the adiabatic switching procedure is the following: given the Hamiltonian

$$H = H_0 + V = \frac{-\Delta}{2m} + V \tag{1}$$

[Δ is the Laplacean, V a multiplicative operator given by the function $V(\mathbf{x})$], define a scattering matrix as follows:

(a) Replace H by the time-dependent Hamiltonian

$$H_\epsilon(t) = H_0 + e^{-\epsilon|t|}V. \tag{2}$$

(b) Solve the Schrödinger equation ($\hbar = 1$)

$$i\,\partial\psi(t)/\partial t = H_\epsilon(t)\psi(t) \tag{3}$$

by finding a unitary operator $U_\epsilon(t)$ such that

$$U(0) = 1, \qquad i\,\partial U_\epsilon(t)/\partial t = H_\epsilon(t)U_\epsilon(t). \tag{4}$$

[Then $\psi(t)$ is given by $U_\epsilon(t)\psi(0)$.]

(c) Form the operators

$$\Omega_\epsilon(t) = U_\epsilon^*(t)e^{-iH_0t} \tag{5}$$

and prove the existence of the "adiabatically switched Møller wave matrices":

$$\Omega_\epsilon^{\pm} = \lim_{t\to\pm\infty} \Omega_\epsilon(t). \tag{6}$$

(d) Prove the existence of the operators

$$\Omega_0^{\pm} = \lim_{\epsilon\to 0} \Omega_\epsilon^{\pm} \tag{7}$$

and use these to define the scattering matrix S_0:

$$S_0 = (\Omega_0^+)^*\Omega_0^-,$$

or, somewhat less ambitiously,

(d′) define the switched S-matrix S_ϵ by

$$S_\epsilon = (\Omega_\epsilon^+)^*\Omega_\epsilon^-$$

and prove that the limit

$$S_0' = \lim_{\epsilon\to 0} S_\epsilon \tag{7a}$$

exists. Then define the scattering matrix as S_0'. [Naturally, if $\Omega_\epsilon^{\pm}$ converged strongly to $\Omega_0^{\pm}$, as in (d), We would have $S_0 = S_0'$.]

This program is offered as a substitute for the usual method, which is:

(a) solve the Schrödinger equation without switching

$$i\,\partial\psi(t)/\partial t = H\psi(t) \tag{8}$$

by taking for $\psi(t)$ the function $U(t)\psi(0)$, where

$$U(t) = e^{-iHt}. \tag{9}$$

(b) Form the operator

$$\Omega(t) = e^{iHt}e^{-iH_0t}. \tag{10}$$

(c) Prove the existence of the operators

$$\Omega^{\pm} = \lim_{t\to\pm\infty} e^{iHt}e^{-iH_0t} \tag{11}$$

and use these to define the scattering matrix S

$$S = (\Omega^+)^*\Omega^-. \tag{12}$$

* Present address: Department of Mathematics, University of Rochester, Rochester, N. Y.

[1] M. Gell-Mann and M. L. Goldberger, Phys. Rev. **91**, 398 (1953). F. Coester, M. Hamermesh and K. Tanaka, Phys. Rev. **96**, 1142 (1954) and references therein; J. M. Jauch and F. Rohrlich, *Theory of Photons and Electrons* (Addison–Wesley Publishing Company, Inc., Reading, Mass., 1955), p. 134, S. Schweber, *An Introduction to Relativistic Quantum Field Theory* (Row, Peterson and Company, New York, 1961), p. 322.

It is hoped that

(1) When $\Omega^{\pm}$ exists, $\Omega_0^{\pm}$ exists and

$$\Omega_0^{\pm} = \Omega^{\pm}$$

so that the adiabatically switched theory yields the correct S-matrix as $\epsilon \to 0$.

(2) When $\Omega^{\pm}$ does not exist, $\Omega_0^{\pm}$ of (7) or at least S_0' of (7a) can still be defined, and used to produce a reasonable S-matrix for the theory.

In the following, we investigate the mathematical existence of the switched theory (that is, existence of $U_\epsilon(t)$, $\Omega_\epsilon^{\pm}$, etc.). Although giving proofs only for the case of potential scattering, we mention results for the case of n-body scattering, and give an idea of the structure of adiabatically switched theories in general. We then investigate the truth of statements (1) and (2) above.

I. ADIABATICALLY SWITCHED THEORIES

Orientation: We work in the Hilbert space $\mathfrak{L}^2$ of square-integrable functions in one three-vector variable $\mathbf{x}$. The $\mathfrak{L}^2$ norm of the function f is denoted by $||f||$. We consider a Hamiltonian H of the type given by (1) acting on this space, where V satisfies either or both of the conditions:

(1) $V(\mathbf{x})$ is square-integrable or

(2) $V(\mathbf{x})$ is locally square-integrable and bounded for large $|\mathbf{x}|$. In this case[2] V is "infinitely small" compared to $H_0 = -\Delta/2m$, by which is meant the following: denote by $\mathfrak{D}(\mathcal{O})$ the domain of the operator $\mathcal{O}$. Then, if $f \in \mathfrak{D}(H_0)$, it follows that $f \in \mathfrak{D}(V)$ and

$$||Vf|| \leq \alpha\, ||H_0 f|| + \beta\, ||f||, \tag{13}$$

where α can be chosen as small as desired, and the bound holds simultaneously for all $f \in \mathfrak{D}(H_0)$. Further, the operator H is (essentially) self-adjoint, $\mathfrak{D}(H) = \mathfrak{D}(H_0)$ is dense in $\mathfrak{L}^2$, and e^{-iHt} is unitary.

We now consider the time-dependent Hamiltonians $H_\epsilon(t)$ of Eq. (2) and try to solve the Schrödinger equation (3) by looking for operators $U_\epsilon(t)$ satisfying Eq. (4). The question of the existence of solutions to (4) for general $H_\epsilon(t)$ is not at all well understood. Even in the relatively simple case at hand the problem offers such unappetizing features as noncommutativity of $H_\epsilon(t)$ with $H_\epsilon(t')$ for $t \neq t'$. Nonetheless, the present case is covered by a theorem of Kato[3] which affirms the existence of $U_\epsilon(t)$ defined on all of $\mathfrak{L}^2$, mapping $D(H_0)$ to $D(H_0)$ and satisfying

$$(d/dt)U_\epsilon(t)\varphi = H_\epsilon(t)U_\epsilon(t)\varphi \quad \text{(strong derivative)} \tag{14}$$

for $\varphi \in D(H_0)$, i.e. a solution of (4) in the usual sense. This is a special case of Kato's Theorem 4. (Ref. 3, p. 211), with Kato's $A(t)$ equal to $-iH_\epsilon(t)$, and Kato's $U(t, 0)$ equal to our $U_\epsilon(t)$. The proof that $-iH_\epsilon(t)$ satisfies the hypotheses C_1 to C_4 of Theorem 4 is straightforward but requires much writing, and we leave it to the reader. We remark only that the simplifying properties of $H_\epsilon(t)$ which guarantee existence of $U_\epsilon(t)$ are essentially these:

(a) the very simple and smooth dependence of $H_\epsilon(t)$ on t and

(b) the fact that all the $H_\epsilon(t)$ are self-adjoint operators with the same domain $\mathfrak{D}(H_0)$, as follows from the first-mentioned paper of Kato since $e^{-\epsilon|t|}V$ is "infinitely small" compared to H_0.

It can be seen very simply, using Kato's results, that $U_\epsilon(t)$ is a unitary operator. This is indicated in the Appendix.

Having established the existence of $U_\epsilon(t)$, it is now easy to prove the existence of the operators $\Omega_\epsilon^{\pm}$ of Eq. (6).

Theorem 1: Define $\Omega_\epsilon(t)$ as in Eq. (5). Then $\Omega_\epsilon(t)$ is unitary and $\Omega_\epsilon(t)$ converges strongly as $t \to \pm\infty$ to operators $\Omega_\epsilon^{\pm}$.

Proof. $\Omega_\epsilon(t)$ is clearly unitary, since both $U_\epsilon^*(t)$ and e^{-iH_0t} are unitary. Thus, it suffices to prove convergence of $\Omega_\epsilon(t)$ on a dense set, which we choose to be $\mathfrak{D}(H_0)$. We prove convergence as $t \to +\infty$: let $\varphi \in \mathfrak{D}(H_0)$. Then, for any t_0,

$$\begin{aligned}\Omega_\epsilon(t)\varphi &= \varphi(t_0) + \int_{t_0}^{t} dt'\, \frac{d}{dt'}\,[\Omega_\epsilon(t')\varphi] \\ &= \varphi(t_0) + \int_{t_0}^{t} dt'\; U_\epsilon^*(t')e^{-\epsilon|t'|}Ve^{-iH_0t'}\varphi. \end{aligned}\tag{15}$$

We can now prove convergence of $\Omega_\epsilon(t)\varphi$ by showing that $\int_{t_0}^{\infty} dt'\; ||(d/dt')(\Omega_\epsilon(t')\varphi)||$ exists. But this is trivial, because

$$\begin{aligned}||(d/dt)[\Omega_\epsilon(t)\varphi]|| &= ||U_\epsilon^*(t)e^{-\epsilon|t|}Ve^{-iH_0t}\varphi|| \\ &= e^{-\epsilon|t|}\, ||Ve^{-iH_0t}\varphi||, \end{aligned}\tag{16}$$

and according to (13)

$$\begin{aligned}||Ve^{-iH_0t}\varphi|| &\leq \alpha\, ||H_0e^{-iH_0t}\varphi|| + \beta\, ||e^{-iH_0t}\varphi|| \\ &= \alpha\, ||H_0\varphi|| + \beta\, ||\varphi|| \end{aligned}\tag{17}$$

[in (17) we have used unitarity of e^{-iH_0t}]. Thus

$$||(d/dt)[\Omega_\epsilon(t)\varphi]|| \leq e^{-\epsilon|t|}(\alpha\, ||H_0\varphi|| + \beta\, ||\varphi||) \tag{18}$$

[2] T. Kato, Trans. Am. Math. Soc. **70**, 195 (1951).
[3] T. Kato, J. Math. Soc. Japan **5**, 208 (1953).

and the integral of the left-hand side from t_0 to ∞ exists. The proof of convergence as $t \to -\infty$ is similar.

To give a more complete insight into the structure of the adiabatically switched theory, we also prove

Theorem 2: Define $\Omega_\epsilon(t)$ as in Eq. (5). Then $\Omega_\epsilon^*(t)$ is unitary and $\Omega_\epsilon^*(t)$ converges strongly as $t \to \pm\infty$ to $\Omega_\epsilon^{\pm *}$, where $\Omega_\epsilon^\pm$ are the operators defined in Theorem 1.

Proof. $\Omega_\epsilon^*(t)$ is clearly unitary, and as in Theorem 1, we shall prove convergence on $\mathfrak{D}(H_0)$ as $t \to +\infty$. It suffices to show that $\int_{t_0}^{\infty} dt' \; ||(d/dt')\; [\Omega_\epsilon^*(t)\varphi]||$ exists for $\varphi \in \mathfrak{D}(H_0)$. But

$$||(d/dt)[\Omega_\epsilon^*(t)\varphi]|| = ||e^{iH_0 t}e^{-\epsilon|t|}VU_\epsilon(t)\varphi|| = e^{-\epsilon|t|}\,||VU_\epsilon(t)\varphi|| \qquad (19)$$

so that we will be finished if we can show that $||VU_\epsilon(t)\varphi||$ is bounded. This is not obvious but is proved in the Appendix, completing the convergence proof. That the limit of $\Omega_\epsilon^*(t)$ is $\Omega_\epsilon^{\pm *}$ is seen as follows: strong convergence of $\Omega_\epsilon(t)$ to $\Omega^\pm$ implies weak convergence of $\Omega_\epsilon^*(t)$ to $\Omega_\epsilon^{\pm *}$. Since we know that actually $\Omega_\epsilon^*(t)$ converges strongly, it must converge strongly to $\Omega_\epsilon^{\pm *}$.

It is Theorem 2 which distinguishes the adiabatically switched theories from the ordinary ones. In any theory with bound states, it is false that $\Omega^*(t)$ [see Eq. (10)] converges strongly on all of $\mathfrak{L}^2$. In fact, as is easy to see, $\Omega^*(t)$ fails to converge strongly on any bound state, but instead converges weakly to zero. This means in turn that $\Omega^{\pm *}$ annihilates the bound states, and that $\Omega^\pm$, although isometric, is not unitary. We summarize our information on the adiabatically switched theory in the following theorem.

Theorem 3: The operators $\Omega_\epsilon^\pm$ are unitary. There are no bound states in the adiabatically switched theory in the sense that $H_\epsilon(t)$ has no eigenstate for all t. The adiabatically switched theory is asymptotically complete in the sense that the asymptotic states $\Omega_\epsilon^\pm\varphi$ span all of $\mathfrak{L}^2$.

Proof. Both $\Omega_\epsilon^\pm$ and $\Omega_\epsilon^{\pm *}$ are the strong limits of sequences of unitary operators, and both are, therefore, isometric, i.e.,

$$\begin{aligned} \Omega_\epsilon^{\pm *}\;\Omega_\epsilon^\pm &= 1 \quad \text{(isometry of } \Omega_\epsilon^\pm\text{)}, \\ \Omega_\epsilon^\pm\;\Omega_\epsilon^{\pm *} &= 1 \quad \text{(isometry of } \Omega_\epsilon^{\pm *}\text{)}, \end{aligned} \qquad (20)$$

and Eqs. (20) imply that $\Omega_\epsilon^\pm$ is unitary. There are no bound states in the theory, for if ψ_B were such a state, with energy E_B, we would have

$$\Omega_\epsilon^*(t)\psi_B = e^{iH_0 t}U_\epsilon(t)\psi_B = e^{iH_0 t}e^{-iE_B t}\psi_B, \qquad (21)$$

and the right-hand side converges weakly to zero, while we know the left-hand side converges strongly to something with norm $||\psi_B||$, as $t \to \pm\infty$, as indicated.

Finally, since the range of a unitary operator acting on a Hilbert space is the entire Hilbert space, we have symbolically

$$\Omega_\epsilon^\pm \mathfrak{L}^2 = \mathfrak{L}^2 \qquad (22)$$

so that the theory is asymptotically complete.

Results in the n-body case. In the n-body case, we have only to redefine H_0 and V by

$$\begin{aligned} H_0 &= \sum_{k=1}^{n} -\frac{\Delta_k}{2m_k}, \\ V &= \sum_{i=1}^{n} V_{0i}(\mathbf{x}_i) + \sum_{1 \le i < j \le n} V_{ij}(\mathbf{x}_i - \mathbf{x}_j), \end{aligned} \qquad (23)$$

and make the same assumptions on

$$V_{ij}\ (0 \le i < j \le n)$$

as we made previously on $V(\mathbf{x})$. Naturally, we also replace $\mathfrak{L}^2$ in one three-vector variable by $\mathfrak{L}^2$ in the n three-vector variables $\mathbf{x}_1 \cdots \mathbf{x}_n$. Then defining $H_\epsilon(t)$, $U_\epsilon(t)$, etc., as before, Theorems 1, 2, and 3 also hold in the n-body case. This implies that the n-body adiabatically switched theory is a one-channel theory, i.e., there is only one possible asymptotic behavior for any wave-packet $U_\epsilon(t)\varphi$, and this behavior is evolution according to the free Schrödinger equation describing n noninteracting particles. This is clear since by Theorem 3 if $\varphi \in \mathfrak{L}^2$, there are states $\psi^\pm \in \mathfrak{L}^2$ such that $\varphi = \Omega_\epsilon^\pm\psi^\pm$. Then by Theorem 1, $U_\epsilon(t)\varphi$ approaches, as $t \to \pm\infty$, the states $e^{-iH_0 t}\psi^\pm$.

We now turn to the problem of obtaining the usual Møller wave matrices $\Omega^\pm$ from the operators $\Omega_\epsilon^\pm$ as $\epsilon \to 0$.

II. CONVERGENCE TO THE USUAL THEORY AS $\epsilon \to 0$

We first take a large class of potentials for which $\Omega^\pm$ is known to exist and show that for those potentials $\Omega^\pm$ is the limit as $\epsilon \to 0$ of $\Omega_\epsilon^\pm$. For convenience, we summarize some results in the literature.

(A) If $V(\mathbf{x})$ is square-integrable or

(B) $V(\mathbf{x})$ is locally square-integrable and for some $M > 0$

$$|V(\mathbf{x})| < \frac{C}{|\mathbf{x}|^\alpha} \quad \text{for} \quad |\mathbf{x}| \ge M, \ \text{with } \alpha > 1,$$

en the operators $\Omega^{\pm}$ of Eq. (11) exist as strong nits on all of $\mathfrak{L}^2$.[4] Convergence (as $t \to +\infty$) can be oved by showing that for φ belonging to a certain t $\mathcal{S}$ dense in $\mathfrak{L}^2$ and some t_0 [and thus for any t_0, ıce by (17) $||Ve^{-iH_0t}\varphi||$ is bounded],

$$\int_{t_0}^{\infty} \left|\left|\frac{d}{dt}\,\Omega(t)\varphi\right|\right| dt = \int_{t_0}^{\infty} ||Ve^{-iH_0t}\varphi||\, dt < \infty. \quad (24)$$

is so chosen that $\mathcal{S} \subset \mathfrak{D}(H_0)$. (For instance, $\mathcal{S}$ n be taken to be Schwartz's space $\mathcal{S}$ of infinitely ferentiable functions of rapid decrease as $|\mathbf{x}| \to \infty$). We now prove the following.

Theorem 4: If $V(\mathbf{x})$ satisfies (A) or (B), then

$$\lim_{\epsilon \to 0} \Omega_\epsilon^{\pm} \equiv \Omega_0^{\pm} = \Omega^{\pm}, \quad (25)$$

e limit holding in the sense of strong convergence; ., statement 1 of the introduction is true.

Proof: We give the proof for Ω^+. It suffices to ove convergence on a dense set, which we choose be the set $\mathcal{S}$ discussed above. For $\varphi \in \mathcal{S}$ and any > 0, we have

$$\Omega_\epsilon^+ - \Omega^+)\varphi|| \le ||(\Omega_\epsilon^+ - \Omega_\epsilon(t))\varphi||$$
$$+ ||(\Omega(t) - \Omega^+)\varphi|| + ||(\Omega_\epsilon(t) - \Omega(t))\varphi||. \quad (26)$$

e can estimate

$$\Omega_\epsilon^+ - \Omega_\epsilon(t))\varphi||$$
$$= \left|\left|\int_t^{\infty} (d/dt')[\Omega_\epsilon(t')\varphi]\, dt'\right|\right|$$
$$\le \int_t^{\infty} \left|\left|\frac{d}{dt'}\,[\Omega_\epsilon(t')\varphi]\right|\right| dt'$$
$$= \int_t^{\infty} e^{-\epsilon t'}\, ||Ve^{-iH_0t'}\varphi||\, dt'$$
$$\le \int_t^{\infty} ||Ve^{-iH_0t'}\varphi||\, dt', \quad (27)$$

ere we know the last integral converges since V isfies (A) or (B) and $\varphi \in \mathcal{S}$ so that (24) holds. ıce the right-hand side of (27) is independent of ve can, by taking t large enough, estimate the t-hand side of (27) uniformly in ϵ. Likewise, we ı make $||(\Omega(t) - \Omega^+)\varphi||$ as small as we like by ɔosing t large enough.

We now show that for fixed t the term

$$||(\Omega_\epsilon(t) - \Omega(t))\varphi||$$

vanishes as $\epsilon \to 0$. Since $\Omega_\epsilon(t) = U_\epsilon^*(t)e^{-iH_0t}$ and $\Omega(t) = e^{iHt}e^{-iH_0t}$, it is enough to show that $||[U_\epsilon^*(t) - e^{iHt}]\psi||$ vanishes as $\epsilon \to 0$ for any $\psi \in \mathfrak{D}(H_0)$. [If this has been shown, let $\varphi \in \mathcal{S}$. Then $\varphi \in \mathfrak{D}(H_0)$ and $e^{-iH_0t}\varphi = \psi \in \mathfrak{D}(H_0)$, so the result follows.] Using the fact that $U_\epsilon^*(0) = 1$ and some obvious manipulations, we have for $\psi \in \mathfrak{D}(H_0)$

$$||(U_\epsilon^*(t) - e^{iHt})\psi|| = ||(U_\epsilon^*(t)e^{-iHt} - 1)e^{iHt}\psi||$$
$$= \left|\left|\int_0^t \frac{d}{dt'}\,[U_\epsilon^*(t')e^{-iHt'}]e^{iHt}\psi\right|\right|$$
$$\le \int_0^t (1 - e^{-\epsilon t'})\, ||Ve^{-iH(t'-t)}\psi||. \quad (28)$$

Also,

$$||Ve^{-iH(t'-t)}\psi|| \le a\, ||He^{-iH(t'-t)}\psi||$$
$$+ b\, ||e^{-iH(t'-t)}\psi|| = a\, ||H\psi|| + b\, ||\psi||. \quad (29)$$

Equation (29) follows from the Appendix [Eq. (A7) at $t = 0$; remember that $H_\epsilon(0) = H$]. Thus the integral on the right-hand side of (28) is bounded by const. $\int_0^t (1 - e^{-\epsilon t'})\, dt'$ and converges to zero as $\epsilon \to 0$.

Remark: the proof that $||(U_\epsilon^*(t) - e^{iHt})\psi||$ vanishes as $\epsilon \to 0$ for $\psi \in \mathfrak{D}(H_0)$ did not depend on V satisfying (A) or (B). It is enough that V satisfies (1) or (2) at the beginning of Sec. I. We shall use this result later in the case of the Coulomb potential, which satisfies (2) but not (A) or (B).

We now complete the proof of Theorem 4 as follows: Let $\eta > 0$. Choose t so large that, for all $\epsilon > 0$, each of the first two terms on the right-hand side of (26) is less than $\frac{1}{3}\eta$. Then choose ϵ so small that the last term is less than $\frac{1}{3}\eta$. The result is that if ϵ is small enough

$$||(\Omega_\epsilon^+ - \Omega^+)\varphi|| < \eta, \quad \text{Q.E.D.} \quad (30)$$

Result in the n-body case. Theorem 4 holds in the n-body case provided that $\Omega^{\pm}$ are interpreted as the Møller wave matrices for the channel in which all particles are asymptotically free.

We now investigate the situation when $V(\mathbf{x})$ is the Coulomb potential $e_1e_2/|\mathbf{x}|$, which is of the second type mentioned at the beginning of Sec. I. Again, we summarize certain facts for the convenience of the reader.[5]

Define the "distorted free-propagation operator"

$$U_\epsilon(t) = \exp\,[-iH_{0\epsilon}(t)] \quad (31)$$

[4] J. M. Cook, J. Math. & Phys. **36**, 82 (1957).
J. M. Jauch and I. Zinnes, Nuovo Cimento **11**, 553 (1959). light extension of the work of the second paper is necessary ıandle potentials satisfying (B).

[5] J. D. Dollard, J. Math. Phys. 5, 729 (1964), and Ph.D. thesis, Princeton University (1963).

with

$$H_{0c}(t) = H_0 t + \epsilon(t)[me_1e_2/(-\Delta)^{\frac{1}{2}}] \log(-2|t|\Delta/m) = H_0 t + H'_{0c}(t), \quad (32)$$

where

$$H_0 = -\Delta/2m, \quad \epsilon(t) = \begin{cases} 1 & (t > 0), \\ -1 & (t < 0). \end{cases} \quad (33)$$

Then with

$$H_c = (-\Delta/2m) + (e_1e_2/|\mathbf{x}|), \quad (34)$$

the operator

$$\Omega_c(t) = e^{iH_c t}U_c(t) \quad (35)$$

converges strongly as $t \to \pm\infty$ to operators $\Omega_c^{\pm}$ which correctly give the Coulomb scattering matrix elements. This result implies that the operators

$$\Omega(t) = e^{iH_c t}e^{-iH_0 t}$$

cannot converge strongly, since, writing

$$\Omega(t) = e^{iH_c t}U_c(t)U_c(t)^* e^{-iH_0 t}, \quad (36)$$

it is easy to show that $\Omega(t)$ converges weakly to 0 as $t \to \pm\infty$ because $U_c(t)^* e^{-iH_0 t}$ does. This is done as follows: In the case at hand,[5] any function $\psi \in \mathfrak{L}^2$ can be written as the sum of a function ψ_B belonging to the subspace spanned by the bound states and a function φ which belongs to the range of both Ω_c^+ and Ω_c^- and can be written as $\Omega_c^{\pm}f^{\pm}$. To verify weak convergence of $\Omega(t)$ to zero, it thus suffices to verify that for any $\varphi \in \mathfrak{L}^2$ the expression $(\psi, \Omega(t)\varphi)$ converges to zero when (1) ψ is a bound state or (2) ψ has the form $\Omega_c^{\pm}f^{\pm}$. In case (1), we have $(\psi, \Omega(t)\varphi) = e^{iEt}(\psi, e^{-iH_0 t}\varphi)$ where E is the energy of ψ. Then convergence to zero is clear. (For a rigorous proof, one uses a slightly more refined version of the Riemann–Lebesgue lemma.) In case (2), using the fact that $U_c(t)^* e^{-iH_c t}\Omega_c^{\pm}$ converges strongly to $(\Omega_c^{\pm})^*\Omega_c^{\pm} = 1$ as $t \to \pm\infty$, we find that $(\psi, \Omega(t)\varphi)$ asymptotically has the form $(f^{\pm}, U_c^*(t)e^{-iH_0 t}\varphi)$ which goes to zero (again by the Riemann–Lebesgue lemma.) Thus $\Omega(t)$ does not converge strongly, and $\Omega^{\pm}$ of (11) does not exist as a strong limit.

The convergence proof for $\Omega_c(t)$ in Ref. 5 (as $t \to +\infty$) proceeds by showing that

$$\int_{t_0}^{\infty} \left\| \frac{d}{dt}\Omega_c(t)\varphi \right\| dt < \infty \quad (37)$$

for $t_0 > 1$ and a suitably chosen set of functions φ.

We will now give a counterexample to statement (2) of the Introduction by showing that the correct scattering matrices $\Omega_c^{\pm}$ *cannot* be obtained by the method of adiabatic switching outlined in the introduction. In order to do this, we first prove that the $\Omega_c^{\pm}$ *are* obtainable by a *different* switching procedure. This is done as follows: the time-derivative of $U_c(t)$ is given by

$$\frac{idU_c(t)}{dt} = \left(H_0 + \frac{me_1e_2}{(-\Delta)^{\frac{1}{2}}|t|}\right)U_c(t) \quad (t \neq 0). \quad (38)$$

We will define an adiabatically switched version of $U_c(t)$, called $U_{c\epsilon}(t)$, by requiring that $U_{c\epsilon}(t)$ should agree with $U_c(t)$ at some time t_0 and satisfy the equation

$$\frac{idU_{c\epsilon}(t)}{dt} = \left(H_0 + \frac{me_1e_2e^{-\epsilon|t|}}{(-\Delta)^{\frac{1}{2}}|t|}\right)U_{c\epsilon}(t) \quad (t \neq 0). \quad (38a)$$

Since the point $t = 0$ is clearly a "trouble point" in (38) and (38a), it is convenient to avoid talking about it by defining $U_{c\epsilon}(t)$ only for, say, $|t| \geq t_0 > 0$. We thus require

For $t \geq t_0$: $U_{c\epsilon}(t_0) = U_c(t_0)$ and $U_{c\epsilon}(t)$ satisfies (38a).

For $t \leq -t_0$: $U_{c\epsilon}(-t_0) = U_c(-t_0)$ and $U_{c\epsilon}(t)$ satisfies (38a).

From now on we deal only with the case $t \geq t_0$. The case $t \leq -t_0$ is dealt with analogously. For $t \geq t_0$ we can write

$$U_{c\epsilon}(t) = \exp(-iH_{0c\epsilon}(t)) \quad (39)$$

with

$$H_{0c\epsilon}(t) = H_0 t + \frac{me_1e_2}{(-\Delta)^{\frac{1}{2}}}\int_{t_0}^{t} dt' \frac{e^{-\epsilon t'}}{t'} + \frac{me_1e_2}{(-\Delta)^{\frac{1}{2}}}\log\frac{(-2t_0\Delta)}{m}. \quad (40)$$

It should be clear that

$$U_{c\epsilon}(t_0) = U_c(t_0) \quad (41)$$

and that $U_{c\epsilon}(t)$ as defined in (39) satisfies (38a).

We now let $U_\epsilon(t)$ be the operator (4) whose derivative is $H_\epsilon(t)U_\epsilon(t)$, where $H_\epsilon(t)$ is the adiabatically switched Coulomb Hamiltonian $-\Delta/2m + e^{-\epsilon|t|}e_1e_2/|\mathbf{x}|$. Then we have the following.

Theorem 5: Define $\Omega'_{c\epsilon}(t)$ for $|t| > t_0$ by

$$\Omega'_{c\epsilon}(t) = U_\epsilon^*(t)U_{c\epsilon}(t). \quad (42)$$

Then the strong limits

$$\lim_{t\to\pm\infty} \Omega'_{c\epsilon}(t) = \Omega'^{\pm}_{c\epsilon}$$

exist on all of $\mathfrak{L}^2$.

Proof. As usual, we prove convergence as $t \to +\infty$ on a dense set, this time the set of C-functions.[5] $f(\mathbf{x})$ is a C-function if $f(\mathbf{x})$ belongs to Schwartz's space $\mathcal{S}$ and its Fourier transform $\hat{f}(\mathbf{k})$ vanishes in a neighborhood of $\mathbf{k} = 0$. All such functions belong to the domain of H_0 and the domain of $H_{0\epsilon}(t)$. To give the proof as $t \to +\infty$ we estimate for a C-function f and $t_1 \geq t_0$, $t_1 > 1$ (this apparently irrelevant condition makes the comparison with Ref. 5 more immediate):

$$\int_{t_1}^{\infty} dt \left\| \frac{d}{dt} (\Omega'_{c\epsilon}(t) f) \right\|$$

$$= \int_{t_1}^{\infty} dt \left\| U^*_\epsilon(t) e^{-\epsilon t} e_1 e_2 \left(\frac{1}{|\mathbf{x}|} - \frac{m}{(-\Delta)^{\frac{1}{3}} t} \right) U_{c\epsilon}(t) f \right\|$$

$$= e_1 e_2 \int_{t_1}^{\infty} dt\, e^{-\epsilon t} \left\| \left(\frac{1}{|\mathbf{x}|} - \frac{m}{(-\Delta)^{\frac{1}{3}} t} \right) U_{c\epsilon}(t) f \right\|. \quad (43)$$

We could now use (13) and the fact that H_0 and $(-\Delta)^{-\frac{1}{3}}$ commute with $U_{c\epsilon}(t)$ to show that the norm in the integrand is bounded and the convergence takes place. However, it is possible to prove a stronger result; actually, the integral converges without the factor $e^{-\epsilon t}$.

$$\int_{t_1}^{\infty} dt \left\| \left(\frac{1}{|\mathbf{x}|} - \frac{m}{(-\Delta)^{\frac{1}{3}} t} \right) U_{c\epsilon}(t) f \right\| < \infty. \quad (44)$$

The proof of this fact closely resembles the proof in Ref. 5 that

$$\int_{t_1}^{\infty} dt \left\| \left(\frac{1}{|\mathbf{x}|} - \frac{m}{(-\Delta)^{\frac{1}{3}} t} \right) U_c(t) f \right\| < \infty$$

and will not be given here. One result of the proof is that the convergence in (44) can be proved with estimates independent of ϵ, so that the quantity $||(\Omega'_{c\epsilon}(t) - \Omega'^{+}_{c\epsilon}) f||$ can be made arbitrarily small uniformly in ϵ if t is taken large enough, by the same type of estimate as used in (27).

Using these results, we are now in a position to prove Theorem 6.

Theorem 6: If $\Omega^{\pm}_c$ are the operators obtained from (35) as $t \to \pm\infty$, and $\Omega'^{\pm}_{c\epsilon}$ are the operators of Theorem 5, then

$$\lim_{\epsilon \to 0} \Omega'^{\pm}_{c\epsilon} = \Omega^{\pm}_c \quad (45)$$

in the sense of strong convergence.

Proof. We prove the result for Ω^{+}_c on the set of C-functions. If f is a C-function, for any $t \geq t_0$

$$||(\Omega'^{+}_{c\epsilon} - \Omega^{+}_c) f|| \leq ||(\Omega'^{+}_{c\epsilon} - \Omega'_{c\epsilon}(t)) f||$$
$$+ ||(\Omega_c(t) - \Omega^{+}_c) f|| + ||(\Omega'_{c\epsilon}(t) - \Omega_c(t)) f||. \quad (46)$$

By the discussion above and the convergence of $\Omega_c(t)$ to Ω^{+}_c, we see that by choosing t large enough we can make the first two terms as small as we like independent of ϵ. It remains to show that for fixed t we can make $||(\Omega'_{c\epsilon}(t) - \Omega_c(t)) f||$ as small as we like. Since

$$\Omega'_{c\epsilon}(t) = U^*_\epsilon(t) U_{c\epsilon}(t),$$
$$\Omega_c(t) = e^{iH_c t} U_c(t),$$

and $U^*_\epsilon(t)$, $U_{c\epsilon}(t)$, $e^{iH_c t}$, and $U_c(t)$ are unitary, it suffices to show that for any $t \geq t_0$

$$\lim_{\epsilon \to 0} U^*_\epsilon(t) = e^{iH_c t}, \quad (47a)$$

$$\lim_{\epsilon \to 0} U_{c\epsilon}(t) = U_c(t), \quad (47b)$$

where the limits are strong limits on all of $\mathcal{L}^2$. This is true because if A_ϵ, B_ϵ are unitary, and A_ϵ, B_ϵ converge strongly to A, B, respectively, as $\epsilon \to 0$, then

$$||(A_\epsilon B_\epsilon - AB)\varphi||$$
$$\leq ||(A_\epsilon - A) B\varphi|| + ||A_\epsilon (B_\epsilon - B)\varphi||$$
$$= ||(A_\epsilon - A) B\varphi|| + ||(B_\epsilon - B)\varphi|| \to 0, \quad \epsilon \to 0. \quad (48)$$

Equation (47a) has already been proven on all functions $\psi \in \mathfrak{D}(H_0)$ [see Eq. (28)ff] and the proof extends immediately to all of $\mathcal{L}^2$. Equation (47b) is also easy to prove, for instance, by writing out $U_{c\epsilon}(t)\psi$ and $U_c(t)\psi$ in momentum space and using Lebesgue's dominated convergence theorem. This completes the proof of Theorem 6.

We now turn to the analysis of the operators $\Omega^{\pm}_\epsilon$ gotten by using the ordinary adiabatic switching procedure discussed in the introduction. The existence of these operators is guaranteed by the work of Sec. I, since the Coulomb potential satisfies (2) at the beginning of Sec. I.

Theorem 7: Let $\Omega^{\pm}_\epsilon$ be the operators obtained for the Coulomb field by the ordinary adiabatic switching procedure:

$$\Omega^{\pm}_\epsilon = \lim_{t \to \pm\infty} U^*_\epsilon(t) e^{-iH_0 t}, \quad (49)$$

where $U_\epsilon(t)$ is defined after Eq. (41). Let $\Omega'^{\pm}_{c\epsilon}$ be the operators of Theorem 5. Then

$$\Omega^{\pm}_\epsilon = \Omega'^{\pm}_{c\epsilon} \exp\left\{ \pm \frac{ime_1 e_2}{(-\Delta)^{\frac{1}{3}}} \left[\int_{t_0}^{\infty} dt' \frac{e^{-\epsilon t'}}{t'} + \log\left(\frac{-2\,|t_0|\,\Delta}{m} \right) \right] \right\}. \quad (50)$$

Further, neither $\Omega_\epsilon^\pm$ nor the switched S-matrix $S_\epsilon = (\Omega_\epsilon^+)^*\Omega_\epsilon^-$ converge strongly as $\epsilon \to 0$. Instead, all these operators converge weakly to zero as $\epsilon \to 0$.

Proof. We first prove (50) for Ω_ϵ^+. Write

$$U_\epsilon^*(t)e^{-H_0 t} = U_\epsilon^*(t)U_{c\epsilon}(t)U_{c\epsilon}^*(t)e^{-iH_0 t}. \tag{51}$$

We already know from Theorem 5 that $U_\epsilon^*(t)U_{c\epsilon}(t)$ converges strongly to $\Omega_{c\epsilon}'^+$ as $t \to +\infty$. We have only to show that

$$U_{c\epsilon}^*(t)e^{-iH_0 t} \underset{t\to+\infty}{\to} \exp\left\{\frac{ime_1e_2}{(-\Delta)^{\frac{1}{2}}} \times \left[\int_{t_0}^{\infty} dt' \frac{e^{-\epsilon t'}}{t'} + \log\left(\frac{-2t_0\Delta}{m}\right)\right]\right\} \quad \text{(strong limit)}. \tag{52}$$

But this is very simple, because a glance at (39) and (40) shows that

$$U_{c\epsilon}^*(t)e^{-iH_0 t} = \exp\left\{\frac{ime_1e_2}{(-\Delta)^{\frac{1}{2}}}\left[\int_{t_0}^{t} dt' \frac{e^{-\epsilon t'}}{t'} + \log\left(\frac{-2t_0\Delta}{m}\right)\right]\right\}. \tag{53}$$

That this operator has the limit given in (52) can be seen by passing to momentum space and using Lebesgue's dominated convergence theorem. Thus (50) is proved for Ω^+. The proof for Ω^- is similar. We now prove the statements on weak convergence to zero: it is easy to see intuitively, and not hard to verify rigorously, that the operator on the right-hand side of (52) converges weakly to zero as $\epsilon \to 0$. This is because as $\epsilon \to 0$ the integral $\int_{t_0}^{\infty} dt' e^{-\epsilon t'}/t'$ diverges and the operator "oscillates itself to death."

Denote for convenience the operator on the right-hand side of (52) by $P(\epsilon)$. Then (50) reads

$$\begin{aligned} \Omega_\epsilon^+ &= \Omega_{c\epsilon}'^+ P(\epsilon), \\ \Omega_\epsilon^- &= \Omega_{c\epsilon}'^- P(\epsilon)^*, \end{aligned} \tag{54}$$

and by the above discussion we know that $P(\epsilon)$ [and thus also $P(\epsilon)^*$] converges weakly to 0 as $\epsilon \to 0$. Now a detailed argument shows that in addition to (45) the following equations hold:

$$\lim_{\epsilon\to 0}(\Omega_{c\epsilon}'^\pm - \Omega_c^\pm)P(\epsilon) = 0 = \lim_{\epsilon\to 0}(\Omega_{c\epsilon}'^\pm - \Omega_c^\pm)P(\epsilon)^* \quad \text{(strong limit)}. \tag{55}$$

Now let $\varphi, \psi \in \mathfrak{L}^2$. Then

$$\lim_{\epsilon\to 0}(\varphi, \Omega_\epsilon^+\psi) = \lim_{\epsilon\to 0}(\varphi, \Omega_{c\epsilon}'^+P(\epsilon)\psi) = \lim_{\epsilon\to 0}(\varphi, (\Omega_{c\epsilon}'^+ - \Omega_c^+)P(\epsilon)\psi) + \lim_{\epsilon\to 0}(\varphi, \Omega_c^+P(\epsilon)\psi) = 0, \tag{56}$$

since the first term on the right-hand side goes to 0 by (55) and the second term can be rewritten as $(\Omega_c^{+*}\varphi, P(\epsilon)\psi)$ which goes to zero by the weak convergence to zero of $P(\epsilon)$. Thus Ω_ϵ^+ converges weakly to zero, and similarly for Ω_ϵ^-. We now evaluate the limit of the switched S-matrix S_ϵ: let $\varphi, \psi \in \mathfrak{L}^2$. Then

$$(\varphi, S_\epsilon\psi) = (\Omega_{c\epsilon}'^+P(\epsilon)\varphi, \Omega_{c\epsilon}'^-P(\epsilon)^*\psi). \tag{57}$$

Now, (55) shows that in the limit as $\epsilon \to 0$ the right-hand side of (57) can be replaced by $[\Omega_c^+P(\epsilon)\varphi, \Omega_c^-P(\epsilon)^*\psi]$:

$$\lim_{\epsilon\to 0}(\varphi, S_\epsilon\psi) = \lim_{\epsilon\to 0}(\Omega_c^+P(\epsilon)\varphi, \Omega_c^-P(\epsilon)^*\psi) = \lim_{\epsilon\to 0}(\varphi, P(\epsilon)^*\Omega_c^{+*}\Omega_c^-P(\epsilon)^*\psi). \tag{58}$$

However, the intertwining relations for the Coulomb–Møller wave-matrices[5]

$$H_c\Omega_c^\pm = \Omega_c^\pm H_0, \tag{59}$$

with H_0 and H_c given by (33) and (34), imply that the free Hamiltonian and, therefore, also $P(\epsilon)^*$ commute with the operator $\Omega_c^{+*}\Omega_c^-$, so we can rewrite (58) as

$$\lim_{\epsilon\to 0}(\varphi, S_\epsilon\psi) = \lim_{\epsilon\to 0}(P(\epsilon)P(\epsilon)\varphi, \Omega_c^{+*}\Omega_c^-\psi) = 0, \tag{60}$$

since it is easy to see that $[P(\epsilon)]^2$ converges weakly to zero as well as $P(\epsilon)$.

Now since by the work of Sec. I the operators $\Omega_\epsilon^\pm$ and S_ϵ are unitary, they cannot have strong limits since we have just seen that they all converge weakly to zero. This completes the proof of the theorem.

Theorem 7 shows that the program of adiabatic switching outlined in the introduction breaks down in the case of Coulomb potentials. The weak convergence to zero of $\Omega_\epsilon^\pm$ and S_ϵ makes these operators useless in defining a reasonable S-matrix for the theory.

To round out the discussion of adiabatic switching, we comment briefly on the prescription of Gell-Mann and Goldberger.[1,6] This prescription would define a switched Møller wave-matrix $\Omega_\epsilon^{+\prime\prime}$ for the Coulomb potential according to

$$\Omega_\epsilon^{+\prime\prime} = \int_0^\infty dt\, e^{-\epsilon t}e^{iH_\epsilon t}e^{-iH_0 t}. \tag{61}$$

But $\Omega_\epsilon^{+\prime\prime}$ (and likewise the corresponding $\Omega_\epsilon^{-\prime\prime}$) converges weakly to zero as $\epsilon \to 0$ because of the previously mentioned weak convergence to zero of $e^{iH_\epsilon t}e^{-iH_0 t}$ as $t \to \pm\infty$ [Eq. (35)ff]. Using slightly

[6] T. F. Jordan, J. Math. Phys. 3, 414 (1962).

more refined techniques, it is also possible to show that the switched S-matrix $(\Omega_\epsilon^{+\prime\prime})^*\Omega_\epsilon^{-\prime\prime}$ converges weakly to zero in this case. Thus the Goldberger–Gell-Mann technique also breaks down for the Coulomb potential.

It is not hard to see that the undesirable attributes of the "switched" operators for the Coulomb case are due to the long range of the Coulomb potential. This long range produces[5] an anomalous asymptotic time dependence for wave packets moving under the influence of a Coulomb potential. This anomalous time dependence is given by a change of phase $\exp[-i\,\epsilon(t)(me_1e_2/k)\log(2k^2|t|/m)]$ of the wave packet in momentum space. Instead of "erasing" the asymptotic effects of this phase, the switching procedures discussed reproduce a part of it, and in the limit as $\epsilon \to 0$, they essentially produce the phase with $|t|$ becoming infinite, thus causing weak convergence to zero. We might say that the switching procedures indicate too faithfully the distortion caused by the long range of the potential.

The above results on the Coulomb potential make it seem doubtful that adiabatic switching procedures can be used to extend nonrelativistic scattering theory in other cases in which the ordinary Møller wave matrices are not defined, i.e., cases in which $e^{iHt}e^{-iH_0t}$ does not converge. The Coulomb potential comes quite close to satisfying condition (B) at the beginning of Sec. II, which guarantees the success of the adiabatic switching method. This potential represents one of the weakest imaginable violations of condition (B), and it is disheartening to think that already for this case, which is also a case of great physical interest, the method breaks down. This makes it seem unlikely that the method will work for potentials which represent worse departures from conditions (A) or (B). In summary, the adiabatic switching method is perfectly justified for a large class of potentials for which $\Omega^\pm$ exists, but the future of attempts to use it to extend the ordinary theory looks dim.

APPENDIX: CONSTRUCTION AND PROPERTIES OF $U_\epsilon(t)$

Construction of $U_\epsilon(t)$

The construction of the operator $U_\epsilon(t)$ for $t > 0$ proceeds as follows: Consider a partition Δ of the interval $(0, t)$:

$$0 = t_0 < t_1 < \cdots < t_n = t,$$
$$t_{j-1} < \tau_j < t_j \qquad (j = 1 \cdots n). \tag{A1}$$

Let

$$X_j = \exp[(t_j - t_{j-1})(-iH_j)],$$
$$H_j = H_\epsilon(\tau_j), \qquad (j = 1 \cdots n). \tag{A2}$$

$$U(\Delta) = X_nX_{n-1} \cdots X_1. \tag{A3}$$

Kato's proof shows that as the mesh $\Delta = \max_j (t_j - t_{j-1})$ tends to zero, $U(\Delta)$ converges strongly to the desired $U_\epsilon(t)$. However, in the present case, $iH_\epsilon(t)$ satisfies Kato's hypotheses as well as $-iH_\epsilon(t)$; and Kato's proof then shows that

$$U(\Delta)^* = X_1^* \cdots X_n^* \tag{A4}$$

also converges strongly, and clearly it converges to $U_\epsilon^*(t)$. But since

$$U(\Delta)^*U(\Delta) = 1 = U(\Delta)U(\Delta)^*, \tag{A5}$$

we have

$$U_\epsilon(t)^*U_\epsilon(t) = 1 = U_\epsilon(t)U_\epsilon(t)^*, \tag{A6}$$

so that $U_\epsilon(t)$ is unitary. The construction for $t < 0$ is handled similarly.

Boundedness of $||V\,U_\epsilon(t)\,\varphi||$

We now proceed to show that the expression $||VU_\epsilon(t)\varphi||$ is bounded in t for $\varphi \in \mathfrak{D}(H_0)$, as stated in the text. First, after a little juggling of (13) and use of the triangle inequality we find that if $\varphi \in \mathfrak{D}(H_0)$, then

$$||V\varphi|| \leq a\,||H_\epsilon(t)\varphi|| + b\,||\varphi||, \tag{A7}$$

where a and b are independent of φ, ϵ, and t, and a can be chosen as small as desired.

We also notice that by the triangle inequality

$$\begin{aligned} &a\,||H_\epsilon(t)\varphi|| + b\,||\varphi|| \\ &\leq a\,||H_0\varphi|| + ae^{-\epsilon t}\,||V\varphi|| + b\,||\varphi|| \\ &\leq a(||H_0\varphi|| + ||V\varphi||) + b\,||\varphi|| \equiv M(\varphi) \quad (t > 0). \end{aligned} \tag{A8}$$

Using the notation of (A1) to (A3) we now estimate $||VX_nX_{n-1} \cdots X_1\varphi||$ for $\varphi \in \mathfrak{D}(H_0)$. In doing so, we shall repeatedly use (A7), (A8), the unitarity of X_j, the fact that $H_j = H_\epsilon(\tau_j)$ commutes with X_j, and the triangle inequality, and may omit explicit mention of some such steps. We also introduce the notation

$$d_j = e^{-\epsilon\tau_j} - e^{-\epsilon\tau_{j+1}} > 0, \tag{A9}$$

and notice that

$$H_i - H_{i+1} = d_i V. \tag{A10}$$

Lemma: If $\varphi \in \mathfrak{D}(H_0)$, and

$$n \geq 2, \text{ then } ||VX_n \cdots X_1\varphi|| \leq \prod_{i=1}^{n-1} (ad_i + 1)M(\varphi).$$

In estimating $||VX_n \cdots X_1\varphi||$ we proceed by induction on n: letting $\varphi \in \mathfrak{D}(H_0)$ we first estimate

$$\begin{aligned} ||VX_1\varphi|| &\leq a\,||H_1X_1\varphi|| \\ &\quad + b\,||\varphi|| = a\,||X_1H_1\varphi|| + b\,||\varphi|| \\ &= a\,||H_1\varphi|| + b\,||\varphi|| \leq M(\varphi) \end{aligned} \tag{A11}$$

and

$$\begin{aligned} ||VX_2X_1\varphi|| &\leq a\,||H_2X_2X_1\varphi|| \\ &\quad + b\,||X_2X_1\varphi|| = a\,||H_2X_1\varphi|| + b\,||\varphi|| \\ &\leq a\,||(H_2 - H_1)X_1\varphi|| + a\,||H_1X_1\varphi|| + b\,||\varphi|| \\ &= ad_1\,||VX_1\varphi|| + a\,||H_1\varphi|| + b\,||\varphi||. \end{aligned} \tag{A12}$$

Now using (A11) on the first term and (A8) on the second two terms, we find

$$||VX_2X_1\varphi|| \leq (ad_1 + 1)M(\varphi). \tag{A13}$$

We now assume that for some $k \geq 2$

$$||VX_kX_{k-1} \cdots X_1\varphi|| \leq \prod_{i=1}^{k-1} (ad_i + 1)M(\varphi) \tag{A14}$$

and estimate

$$\begin{aligned} ||VX_{k+1}X_k \cdots X_1\varphi|| &\leq a\,||H_{k+1}X_{k+1} \cdots X_1\varphi|| \\ &\quad + b\,||X_{k+1} \cdots X_1\varphi|| = a\,||H_{k+1}X_k \cdots X_1\varphi|| \\ &\quad + b\,||\varphi|| \leq a\,||(H_{k+1} - H_k)X_k \cdots X_1\varphi|| \\ &\quad + a\,||H_kX_k \cdots X_1\varphi|| \\ &\quad + b\,||\varphi|| = ad_k\,||VX_k \cdots X_1\varphi|| \\ &\quad + a\,||H_kX_{k-1} \cdots X_1\varphi|| + b\,||\varphi||. \end{aligned} \tag{A15}$$

Writing $H_k = H_{k-1} + (H_k - H_{k-1})$ and continuing in this manner, we get

$$\begin{aligned} ||VX_{k+1} \cdots X_1\varphi|| &\leq ad_k\,||VX_k \cdots X_1\varphi|| \\ &\quad + ad_{k-1}\,||VX_{k-1} \cdots X_1\varphi|| \\ &\quad + \cdots + ad_2\,||VX_2X_1\varphi|| \\ &\quad + ad_1\,||VX_1\varphi|| + a\,||H_1\varphi|| + b\,||\varphi||. \end{aligned} \tag{A16}$$

Using (A8) on the last two terms, (A11) on the third to last, and (A14) on the rest, we have

$$\begin{aligned} &||VX_{k+1} \cdots X_1\varphi|| \\ &\leq ad_k \prod_{i=1}^{k-1} (ad_i + 1)M(\varphi) \\ &+ ad_{k-1} \prod_{i=1}^{k-2} (ad_i + 1)M(\varphi) \\ &+ \cdots + ad_2(ad_1 + 1)M(\varphi) \\ &+ (ad_1 + 1)M(\varphi) = \prod_{i=1}^{k} (ad_i + 1)M(\varphi), \end{aligned} \tag{A17}$$

completing the proof of the lemma. To estimate the size of $\prod_{i=1}^{n-1} (a\,d_i + 1)$, let

$$\begin{aligned} L &= \log \prod_{i=1}^{n-1} (ad_i + 1) \\ &= \sum_{i=1}^{n-1} \log (ad_i + 1) \leq \sum_{i=1}^{n-1} ad_i \\ &= a \sum_{i=1}^{n-1} (e^{-\epsilon r_i} - e^{-\epsilon r_{i+1}}) = a(e^{-\epsilon r_1} - e^{-\epsilon r_n}) < 2a. \end{aligned} \tag{A18}$$

Thus

$$||VX_n \cdots X_1\varphi|| \leq e^{2a}M(\varphi). \tag{A19}$$

Notice that this result holds independent of the number n of X_i's. Now take a sequence Δ_s of partitions such that $U(\Delta_s) \to_{s\to\infty} U_\epsilon(t)$. We know that $U_\epsilon(t)\varphi \in \mathfrak{D}(H_0)$ since $U_\epsilon(t)$ maps $\mathfrak{D}(H_0)$ into itself. Thus $U_\epsilon(t)\varphi \in \mathfrak{D}(V)$. We now have

$$\begin{cases} ||VU(\Delta_s)\varphi|| \leq e^{2a}M(\varphi) & \text{(all s)} \\ U(\Delta_s)\varphi \underset{s\to\infty}{\to} U_\epsilon(t)\varphi \\ U_\epsilon(t)\varphi \in \mathfrak{D}(V) \end{cases} \tag{A20}$$

and these conditions suffice to guarantee that

$$||VU_\epsilon(t)\varphi|| \leq e^{2a}M(\varphi). \tag{A21}$$

Our proof holds for all $t > 0$, since we started from the expressions (A1), (A2), (A3), which were written for $t > 0$. However it is easy to extend the proof to $t < 0$ by similar techniques.

Note added in proof. Without going into details, the author wishes to remark that the results proved here can be extended without difficulty to cover switching procedures in which $e^{-\epsilon|t|}$ is replaced by another function having the same general behavior as a function of t and ϵ.

ACKNOWLEDGMENTS

I would like to thank Professor Edward Nelson and S. Gildar for interesting conversations.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 7, NUMBER 5 MAY 1966

Structure of the Crossing Matrix for Arbitrary Internal Symmetry Groups

D. B. FAIRLIE
Department of Mathematics, University of Durham, Durham, England
(Received 29 August 1965)

Some properties of crossing matrices are deduced which are independent of the particular symmetry group from which the crossing matrices are derived. In particular, a factorization of any elastic crossing matrix, analogous to the factorization of a rotation matrix in n dimensions is found, and the connection between crossing and unitary matrices elucidated. In the special case of the crossing matrix for elastic scattering of particles which transform as representations of SU_2, as well as the usual consistency requirement that each row should sum to unity, a new consistency requirement on the elements in a given column is proved. As a byproduct of this work, a possibly new quadratic identity for Racah coefficients is exhibited.

1. GENERAL PROPERTIES

IN a dynamical theory the consequences of invariance under an internal symmetry group are manifested in the crossing matrices which relate invariant amplitudes in different channels. In this paper some properties of crossing matrices are derived which are independent of any specific symmetry group in the hope that better understanding of the structure of crossing matrices may cast some light on the question of the dynamical origin of symmetries.[1,2] Most of the results obtained apply to crossing matrices for elastic scattering, where the particles exchanged have the same total quantum numbers (i.e., belong to the same representation of the internal symmetry group).

Specifically, if the s and u channels for a scattering process are described by the reactions

$$a + b \to c + d, \tag{1}$$

$$a + \bar{d} \to c + \bar{b}, \tag{2}$$

respectively, where particles a, b, c, and d transform as irreducible representations of an internal symmetry group, then the invariant S-matrix elements in the u channel $S_u(i)$ are related to those in the s channel, denoted by $S_s(j)$ by the crossing relation

$$S_u(i) = C_{ij}^{us} S_s(j), \tag{3}$$

where i labels irreducible representations in the decomposition of the Kronecker products of the representations to which particles a and b belong, and j denotes the representations in the decomposition of the product of the representations associated with particles a and d. The assumption of elastic scattering with the exchange of similar particles means that particles a and c belong to the same irreducible representation of the underlying symmetry group, while particles b and d also transform according to another irreducible representation (possibly the same). This representation is also restricted to be self-conjugate. Then the fundamental properties of the elastic crossing matrix are

$$\sum_i C_{ij} = 1, \tag{4}$$

$$\sum_i C_{ij} C_{jk} = \delta_{ik}, \tag{5}$$

with C_{ij} real.

(The superscript us denoting that the matrix connects the s and u channel amplitudes has been dropped for notational convenience.) These properties are proved[2] from the basic requirement that the matrix C^T transforms projection operators for the irreducible representations in the Kronecker product decomposition in the u channel into projection operators in the s channel.

Equation (4) is an expression of the completeness of the set of projection operators, and has the consequence that the same crossing matrix relates S-matrix elements as transition amplitudes. The second property (5) is a statement that two successive elastic crossings restore the status quo. Taking (4) and (5) as defining relations, the following lemmas are easily verified.

Lemma 1: If A and C are, respectively, $n \times n$ and $m \times m$ elastic crossing matrices, then the direct product

$$A \otimes C$$

is an $mn \times mn$ elastic crossing matrix.

[1] E. P. Wigner, Phys. Today 17, 34 (1964).
[2] A. W. Martin and W. D. McGlinn, Phys. Rev. 136, 1515 (1964). See also J. Rothleitner, Z. Physik 177, 287 964).

Lemma 2: If A and C are elastic crossing matrices of the same dimension, then so is

$$A\ C\ A.$$

Equation (5) implies that the eigenvalues of an elastic crossing matrix C are ± 1, and $+1$ is always an eigenvalue belonging to the eigenvector $\{1, 1 \cdots 1\}$ by (4). Thus

$$\text{Tr } C = n - r \qquad (r = 0, 2, \cdots, 2n-2), \tag{6}$$

where n is the dimension of the matrix C. The construction of Lemma 2 preserves the trace, because of the cyclic property of the latter, and provides a natural separation of crossing matrices into equivalence classes distinguished by their trace. The representations of the elements of period two in the permutation group on n objects by $n \times n$ matrices P_α^+ form a basic set of crossing matrices with non-negative trace. These matrices are monomial (i.e., have only one element in each row and column) and symmetric with unit elements. Clearly Tr P_α^+ = number of diagonal elements. The matrices P_α^- with negative trace which obey the defining laws (4) and (5) are seen to be obtained in terms of the matrices P_α^+ by

$$P_\alpha^- = (2/n)B - P_\alpha^+ \tag{7}$$

where $B_{ij} = 1$ for all $i, j = 1 \cdots n$. These remarks, lead to the following property of factorization.

Theorem I: Every elastic $n \times n$ crossing matrix C may be factored as follows:

$$C = A_1 A_2 \cdots A_{\frac{1}{2}n(n-1)} P_\alpha^\pm A_{\frac{1}{2}n(n-1)} \cdots A_2 A_1 \tag{8}$$

where there are $\frac{1}{2}n(n-1)$ factors both before and after the matrix $P_\alpha^\pm$, a member of the basic set referred to above chosen with trace = Tr C. A typical matrix A_r in the product is defined as follows:

$$\begin{aligned} A_{rij} &= \delta_{ij}, \quad i \neq k, \quad j \neq l; \\ A_{rii} &= A_{rki} = 0i, \qquad j \neq k, l; \\ A_{rkk} &= -A_{rii} = a_r; \\ A_{rki} &= 1 - a_r; \\ A_{rik} &= 1 + a_r. \end{aligned} \tag{9}$$

The proof of this result is straightforward, since

$$\begin{bmatrix} a & 1-a \\ 1+a & -a \end{bmatrix}$$

is the most general 2×2 elastic crossing matrix (apart from the identity),[2] it is clear that every factor in (8) is a crossing matrix. Hence by Lemma 5 the right-hand side of (8) is a crossing matrix, with trace = Tr $P_\alpha^\pm$ = Tr C by construction. Since there are $\frac{1}{2}n(n-1)$ distinct matrices A_r, corresponding to the different possible choices of k and l, it remains the show that C depends upon at most $\frac{1}{2}n(n-1)$ parameters. It is sufficient to show this for a matrix whose elements differ infinitesimally from those of a $P_\alpha^\pm$, as the elements of an arbitrary C may be regarded each as the sum of a power series in some arbitrary expansion parameter λ, the requirement that the conditions (4) and (5) be satisfied to each order in λ serves to determine the coefficients of λ^n in terms of the elements linearly dependent on λ by an iterative procedure. First, for a crossing matrix $C = P_{\alpha ij}^+ + C_{ij}$, with C_{ij} infinitesimal, the condition (5) to first order reads

$$P_{\alpha ii}^+ C_{jk} + C_{ij} P_{\alpha jk}^+ = 0 \quad (\text{all } i, k). \tag{10}$$

Since P_α^+ is monomial with unit elements, this gives

$$C_{hk} + C_{il} = 0, \tag{11}$$

where the indices h and l are specified by the choice of i and k, respectively.

Let D denote the set of indices i, j for which $P_{\alpha ij}^+ = \delta_{ij}$. If there are m indices in D then Tr $C = m$. On the account of the symmetry property of P_α^+, the same Eq. (11) arises from the h, lth component of (10) as from the i, kth component. Hence the number of independent equations is

$$\tfrac{1}{2}(n-m)^2 + m(n-m) + m^2 = \tfrac{1}{2}(n-m)(n+m) + m^2.$$

Also

$$C_{ab} = 0 \quad (\text{all } a, b \in D). \tag{12}$$

The condition that the rows of C all sum to unity is automatically satisfied by (11) for rows associated with diagonal elements of P_α^+, and again by (11) this condition yields $\frac{1}{2}(n-m)$ independent equations for the remaining rows. Hence the number of independent parameters in C is

$$\begin{aligned} n^2 - m^2 - \tfrac{1}{2}(n-m)(n+m) - \tfrac{1}{2}(n-m) &= \tfrac{1}{2}(n-m)(n+m-1) \\ &= \tfrac{1}{2}[n(n-1) - m(m-1)]. \end{aligned}$$

In the second case when C has negative trace the equations analogous to (11) to be satisfied by the infinitesimal quantities C_{ij} in the expression $C = P_{\alpha ij}^- + C_{ij} = (2/n)B_{ij} - P_{\alpha ij}^+ + C_{ij}$ are

$$\frac{2}{n} \sum_r C_{rk} - C_{hk} - C_{il} = 0 \quad (\text{all } i), \tag{13}$$

where h and l bear the same relationship to i and k as in the previous case. Write

$$C_{ij} = C'_{ij} + \sum_r C_{rj} \quad (\text{all } i).$$

Then, from (13), C'_{ij} satisfies the same Eq. (11) as in the previous calculation, together with the additional restrictions

$$\sum_i C'_{ij} = 0, \tag{15}$$

$$\sum_j C'_{ij} = 0. \tag{16}$$

Equation (11) automatically ensures that the column sums are zero for k, l in D and provides $\frac{1}{2}(n - m)$ linearly independent equations for the remaining column sums. Hence the number of independent parameters C'_{ij} is $\frac{1}{2}(n - m)(n + m - 2)$. From Eq. (13) for k, l not in D

$$\sum_r C_{rk} = \sum_r C_{rl} \tag{17}$$

and there is no requirement for $\sum_r C_{rk}$ with k in S except

$$\sum_{rk} C_{rk} = 0 \quad (\text{all } r, k).$$

Hence the number of additional parameters $\sum_r C_{rj}$ in (16) is $m - 1 + \frac{1}{2}(n - m)$. Hence the total number of parameters C_{ij} is

$$\tfrac{1}{2}(n - m)(n + m - 2) + \tfrac{1}{2}(n - m) + m - 1 = \tfrac{1}{2}n(n - 1) - \tfrac{1}{2}(m - 1)(m - 2).$$

Thus the total number of independent parameters in any C which differs infinitesimally from one of the basic crossing matrices $P^{\pm}_{\alpha}$, and hence for any C cannot exceed $\frac{1}{2}n(n - 1)$, justifying the assertion of Theorem I. When P^{+}_{α} has m diagonal elements, then all matrices A_r with nonzero off-diagonal elements only in the set $i, j \in D$, commute with P^{+}_{α}. There are $\frac{1}{2}m(m - 1)$ such matrices, and they may be supposed arranged in the product S in proximity to P^{+}_{α}. Hence the number of independent parameters in the product is $\frac{1}{2}n(n-1)-\frac{1}{2}m(m-1)$ in agreement with the above calculation. This factorization into essentially 2×2 crossing matrices is analogous to the factorization of an $n \times n$ orthogonal matrix.

It turns out that another property additional to (4) and (5) is required to characterize crossing matrices occurring physically. To see this another decomposition of a general elastic crossing matrix is required.

Lemma 3: Every elastic crossing matrix C may be expressed as

$$C = K^{-1}OK, \tag{18}$$

where O is both symmetric and orthogonal, K is symmetric, and $\sum_i K_{ij}$ is an eigenvector of O belonging to an eigenvalue $+1$.

Proof. Any matrix satisfying (5) may be written as $N\Delta N^{-1}$, where N is an arbitrary real nonsingular matrix and Δ is diagonal with elements either ± 1. Now such a matrix N may be written as $K^{-1}O_1$ where K^{-1} is symmetric and O_1 is orthogonal.[3] Hence $K^{-1}O_1\Delta O_1^T K$ satisfies (5). Now $O_1\Delta O_1^T$ is both orthogonal and symmetric by construction. The full matrix will satisfy (4) also if K is restricted so that $\sum_i K_{ij}$ is an eigenvector of O corresponding to unit eigenvalue. The lemma forms the basis of the second main result.

Theorem II: All elastic crossing matrices which occur physically may be written in the form

$$C = \begin{pmatrix} \lambda_1^{-\frac{1}{2}} & & & \\ & \lambda_2^{-\frac{1}{2}} & & \\ & & \ddots & \\ & & & \lambda_n^{-\frac{1}{2}} \end{pmatrix} O \begin{pmatrix} \lambda_1^{\frac{1}{2}} & & & \\ & \lambda_2^{\frac{1}{2}} & & \\ & & \ddots & \\ & & & \lambda_n^{\frac{1}{2}} \end{pmatrix} \tag{19}$$

with O orthogonal and symmetric, and the column vector $\{\lambda_1 \cdots \lambda_n^{\frac{1}{2}}\}$ an eigenvector of O with eigenvalue $+1$.

The proof is immediate. Squaring the total amplitude $\sum_i P^u_i S_u(i) = \sum_i P^s_i S_s(j)$, where P^u_i and P^s_j are projection operators in the U and S channels, the following equation is obtained:

$$\sum \lambda_i \, |S_u(i)|^2 = \sum \lambda_j \, |S_s(j)|^2. \tag{20}$$

Since

$$\sum P_i > < P_i = \lambda_i \delta_{ij}, \tag{21}$$

where the sum is over all intermediate states, and λ_i represents a sum over the conserved quantum numbers in the group representation labeled by i. Now in terms of the crossing relation (18) an alternative expression for (20) may be obtained, i.e.,

$$S^*_u(i)K^2_{ij}S_u(j) = S^*_s(i)K^2_{ij}S_s(j). \tag{22}$$

The equivalence of (20) and (22) requires K^2_{ij} be diagonal with positive eigenvalues. Hence K_{ij} is diagonal with diagonal elements $\lambda_1^{\frac{1}{2}}\lambda_2^{\frac{1}{2}} \cdots \lambda_n^{\frac{1}{2}}$, say. ($K$ is symmetric and may be written as $K_{ij} =$

[3] G. Birkhoff and S. Maclane, *A Survey of Modern Algebra* (The Macmillan Company, New York, 1963), p. 278.

$O_{2ik}\lambda_k^{\frac{1}{2}}O_{2kj}^T$ with O_2 orthogonal since $\lambda_i \neq 1$ in general, and K_{ij}^2 diagonal implies O_2 = identity.) In SU_2, if i is the total isotopic spin of an intermediate state, λ_i is the number of distinct values of the isotopic spin, i.e., $\lambda_i = 2i + 1$. This interpretation requires that λ_i be integral and positive. (This would appear to be a more fruitful approach to the problem of restricting the allowable 2×2 crossing matrices to those corresponding to scattering of a particle of integral or half-integral isotopic spin off another of isotopic spin $\frac{1}{2}$, than the dynamical approach of Martin and McGlinn.[2])

It does not appear from a simple parameter count that Theorem II follows simply from (4) and (5) alone, but provides a genuine additional restriction. The number of parameters required to construct an $n \times n$ symmetric orthogonal matrix of the form $\delta_{ij}n - j + 1 + S_{ij}$, with S_{ij} symmetric, is $n2(n-2)/4$ (n even) and $(n+1)(n-1)/4$ (n odd). The trace of such a matrix is zero (n even) or 1 (n odd). Hence the number of independent eigenvectors with eigenvalue $+1$ is $\frac{1}{2}n$ (n even) or $\frac{1}{2}(n+1)$ (n odd). Thus the total number of parameters in a matrix C constructed by (19) is $\frac{1}{4}n^2$ (n even) or $\frac{1}{4}(n+1)^2$ (n odd). For general n these totals are much less than the $\frac{1}{2}n(n-1)$ parameters permitted to a matrix with minimum positive trace by (4) and (5).

2. EXAMPLES

The 3×3 crossing matrix enjoys some unique properties. There are three classes; the identity, corresponding to trace $+3$, the matrix

$$\begin{bmatrix} 1-(1-a) & \dfrac{(1-a)(1+a-c)}{b-c} & \dfrac{(1-a)(b-a-1)}{b-c} \\ b & 1-\dfrac{b(1+a-c)}{b-c} & \dfrac{b(b-a-1)}{b-c} \\ c & -\dfrac{c(1+a-c)}{b-c} & 1-\dfrac{c(b-a-1)}{b-c} \end{bmatrix} \tag{23}$$

with $b \neq c$, with trace $+1$, and the matrix

$$\frac{1}{(a+b+c)}\begin{bmatrix} a-b-c & 2b & 2c \\ 2a & b-a-c & 2c \\ 2a & 2b & c-a-b \end{bmatrix} \tag{24}$$

with trace -1. All these matrices have the property that they may be represented as in (19). Also in (23), if $a + b + c = 1$, the other two columns have this property too. The anticommutator of any 2×2 crossing matrices or of any two 3×3 matrices of the form (24) is twice the identity. This may be verified most easily by observing that half the sum of any two such matrices is also a crossing matrix and thus has unit square.

Isotopic Spin Crossing Matrices

For the specific case of invariance under SU_2 there is a well-known expression for the crossing matrix for reaction (1) in terms of Racah coefficients,[4-7]

$$J_{ij}^{us} = (-1)^{a+b+c+d}(2_i + 1)W(abdc; ji), \tag{25}$$

where a, b, c, and d denote the isospin of the respective particles.

In the case of elastic crossing $a = c$, $b = d$, the proofs that the matrix (26) has unit square and admits a representation as in Theorem II are well known.[7,8] The row-sum property (4)

$$\sum_j J_{ij}^{us} = 1 \quad (\text{all } i) \tag{26}$$

may be verified by setting $a = c$, $b = d$ in the general sum rule[9]

$$\sum_i (2j+1)(-1)^{a+b-i}W(abdc; ji)W(badc; jf) = W(aifb; dc), \tag{27}$$

and choosing the free parameter f in (28) equal to zero. Rose and Yang[8] have shown that for such an elastic crossing matrix Tr $J = 0$ or 1 (n even or odd). An additional property of the matrix (25) for elastic scattering concerns the column sums.

Theorem III: If J is an $n \times n$ elastic crossing matrix for the group SU_2, the column sums satisfy

$$\sum_i (-1)^i J_{ij} = (-1)^{j+(n-1)}, \tag{28}$$

i.e., the natrix $(-1)^i J_{ij} T (-1)^{j+n-1}$ also satisfies (4) and (5). The proof depends upon another sum rule for Racah coefficients[9];

$$\sum_j (2j+1)W(ajia; bb)W(cjdb; ea)W(ajfc; be) = W(aied; bf)W(bicd; af). \tag{29}$$

Using symmetry properties and summing over i, the equation

---

[4] F. J. Dyson, Phys. Rev. **100**, 344 (1955).
[5] S. Mandelstam, J. E. Paton, R. F. Peierls, and A. Q. Sarker, Ann. Phys. (N. Y.) **18**, 198 (1962).
[6] C. N. Yang, J. Math. Phys. **4**, 52 (1963).
[7] P. A. Carruthers and J. P. Kirsch, Ann. Phys. (N. Y.) **33**, 1 (1965). This paper contains a careful discussion of phase factors.
[8] M. E. Rose and C. N. Yang, J. Math. Phys. **3**, 106 (1962).
[9] M. E. Rose, *Elementary Theory of Angular Momentum* (John Wiley & Sons, Inc., New York, 1957), p. 114.

$$\sum_j \sum_i \{(2j+1)(-1)^{i+j-2b} W(abba; ji)\} \cdot W(cjdb; ea) W(ajfc; be) = \sum_i W(aied; bf) W(bicd; af) \quad (30)$$

results.

Specializing to the case where $e = f$ and $c = d$, and using symmetry properties, Eq. (31) takes the form

$$\sum_j V_j X_{jcc} = \sum_j X_{jcc}, \quad (31)$$

with

$$\left.\begin{aligned} V_i &= \sum_j (2j+1)(-1)^{i+j-2b} W(abba; ji) \\ \text{and} \quad X_{jef} &= W(ajec; be) W(bjcc; ae). \end{aligned}\right\} \quad (32)$$

Now the set of Eqs. (32) regarded as a set linear inhomogeneous set for the unknowns V_i is overdetermined: an evident solution however is $V_i = 1$; thus it is unique and hence Theorem III is proved. This solution of (32) checks with the sum rule (28) and also with the orthogonality requirements for W coefficients. As a corollary, from (31) the general identity

$$\sum_j W(bjdc; ae) W(ajfc; be) = \sum_i W(aied; bf) W(bicd; af) \quad (33)$$

may be deduced. This result is unfamiliar to the present writer, but may be already known.

A result analogous to Theorem III is known for the crossing matrices appropriate to the scattering of particles transforming according to the adjoint representation of SU_n;[10] it would appear from this that Theorem III is capable of generalization to other symmetry groups. Along with the row-sum property (26) it provides a convenient check on elastic isospin crossing matrix calculations. The 2×2 isospin crossing matrix for scattering of an isospin $\frac{1}{2}$-particle by one of isospin j is obtained from the general 2×2 matrix by choosing the parameter a as $1/(2j+1)$;[2] the crossing matrix for the scattering of a particle of isospin 1 off one of isospin j is

$$J = \begin{bmatrix} 1-\frac{(2j-1)(j+1)}{(2j+1)j} & -\frac{1}{j} & \frac{2j+3}{2j+1} \\ -\frac{(2j-1)}{j(2j+1)} & 1-\frac{1}{j(j+1)} & \frac{2j+3}{(2j+1)(j+1)} \\ \frac{2j-1}{2j+1} & \frac{1}{j+1} & 1-\frac{(2j+3)j}{(2j+1)(j+1)} \end{bmatrix}. \quad (34)$$

The matrix (34) clearly satisfies the consistency requirement. In fact its form is determined by Eqs. (23), (19), and (28).

Inelastic crossing matrices

Foldy and Peierls[11] have shown that a general inelastic crossing matrix, must have the form

$$C^{tu}_{ij} = \mu_i^{-\frac{1}{2}} 0_{ij} \lambda_j^{\frac{1}{2}}, \quad (35)$$

where μ_i and λ_i are the dimensions of the irreducible representations of the symmetry group in the t and u channels, respectively, and 0 is an orthogonal matrix. Their proof was for the case of isotopic spin symmetry, but generalizes at once to crossing under an arbitrary symmetry group. As a step towards the further characterization of C^{tu} in terms of C^{us} the following property may be noted:

The rows of C^{tu} are left eigenvectors of C^{us}.

This property is a consequence of the fact that the result of crossing first from the s channel to the u channel and thence to the t channel is equivalent to crossing directly from the s to the t channel as C^{tu} and C^{ts} differ only by real phase factors, since three successive crossings, from the s to the u, to the t, and then back to the s channel, has the effect of conjugating all the particles. It is already known for the case of isotopic spin[8] and in that context is a consequence of Eq. (27). Some further understanding of the structure of inelastic crossing matrices may come from the extension of results (26) and (28) to inelastic processes.

[10] D. E. Neville, Phys. Rev. **132**, 844 (1963).

[11] L. L. Foldy and R. F. Peierls, Phys. Rev. **130**, 1585 (1963). See also D. Amati, L. L. Foldy, A. Stanghellini, and L. Van Hove, Nuovo Cimento **320**, 1685 (1964); and D. C. Peaslee, Phys. Rev. **136**, B1807 (1964).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 7, NUMBER 5 MAY 1966

Generalized Spherical Functions for the Noncompact Rotation Groups

J. FISCHER,* J. NIEDERLE,† AND R. RACZKA‡
International Centre for Theoretical Physics, Trieste, Italy
(Received 7 July 1965)

Special representations of arbitrary noncompact rotation groups labeled by one independent Casimir operator are considered. The explicit construction of the corresponding generalized spherical functions is given and the properties of these representations are discussed in detail.

1. INTRODUCTION

RECENTLY, the study of noncompact groups of higher symmetries has become important due to applications in the physics of elementary particles. It turns out that most of these groups either contain many noncompact rotational groups L(p, q) or appear to be subgroups of some other $L(p, q)$. For instance, the group $U(6, 6)$, which has been currently used in the last time, contains $L(6, 1), \cdots, L(6, 6)$ as its subgroups and itself is a subgroup of $L(12, 12)$. So, the investigation of the noncompact rotation groups is of great interest.[1]

However, the theory of representations of these groups has been worked out mainly from the analytic point of view, which is of little use for the theory of elementary particles. The reason is that physicists wish to know the eigenvalues of the maximal set of commuting operators in the corresponding Lie algebra and how a given representation of the considered group can be decomposed into representations of its compact or noncompact subgroups. On the other hand, in the cases of the rotational or the Lorentz group, we know for physical applications how useful it is to formulate the problem in terms of spherical functions.[2] Therefore it seems worthwhile to work out the theory of representations of the noncompact rotation groups based on the theory of generalized spherical functions.

In Sec. 2 we give a general prescription for the construction of polar coordinate systems in homogeneous spaces having an arbitrary number of dimensions and invariant under the group $L(p, q)$. We derive also the differential equation for the generalized spherical functions. In Sec. 3 we express its solution in terms of conical and spherical functions.

The special case of the groups of the Lorentz type is studied in Sec. 4. (By a group of the Lorentz type we mean a set of transformations which conserve the form $-x_1^2 - x_2^2 - \cdots - x_n^2 + x_{n+1}^2 = R^2$.) Finally, in Sec. 5 we discuss the properties and physical meaning of the continuous unitary representations of $L(p, q)$ thus obtained, and some related problems.

2. CONSTRUCTION OF POLAR COORDINATE SYSTEMS

Let us consider a homogeneous space X_{n-1} set up as the following invariant quadratic form:

$$x_1^2 + x_2^2 + \cdots + x_p^2 - x_{p+1}^2 - x_{p+2}^2 - \cdots - x_{p+q}^2 = R^2, \quad p = 1, 2, \cdots, n; \quad q = 0, 1, \cdots, n - p; \quad p + q = n, \tag{1}$$

where $x_1, \cdots, x_{p+q}$ are the Cartesian coordinates in the n-dimensional Minkowski space M_n. We denote by $L(p, q)$ the noncompact rotation group which conserves this form.

It is known in the cases of $L(3, 0)$ and $L(1, 3)$, there exist 11 and 34 different orthogonal coordinate systems in E_3 and M_4, respectively.[3] Nevertheless only a few of them are useful in physical applications. Analogously, only a suitable choice of the coordinate system in M_n allows us to construct a complete set of orthogonal functions on X_{n-1} to be the set of spherical functions, which are preferable from the physical point of view. In our case we shall construct the coordinate system in X-space by generalizing a method used by Louck and Granzow for compact homogeneous spaces.[4] Namely, we shall

* Institute of Physics of the Czechoslovak Academy of Sciences, Prague, Czechoslovakia.
† On leave of absence from Institute of Physics of the Czechoslovak Academy of Sciences, Prague, Czechoslovakia.
‡ On leave of absence from Institute of Nuclear Research, Warsaw, Poland.

[1] J. Dixmier, Bull. Soc. Math. France 89, 9 (1961). A. Kihlberg, "On the Unitary Representations of a Class of Pseudo-orthogonal Groups," Preprint, Gothenburg, 1965.

[2] A. Z. Dolginov and A. N. Moskalev, Zh. Eksperim. i Teor. Fiz. 37, 1697 (1959) [English transl.: Soviet Phys.—JETP 10, 1202 (1960); A. Z. Dolginov, *ibid.* 30, 746 (1956) [English transl.: *ibid.* 3, 589 (1956)].

[3] M. P. Olevski, Mat. Sborn. 27, 379 (1950).

[4] J. D. Louck, J. Mol. Spectr. 4, 298 (1960). K. D. Granzow, J. Math. Phys. 5, 1474 (1964).

assume that the polar coordinate system in n-dimensional space is known and then we shall derive a general rule for determining coordinates in the space of $n+1$ dimensions.

In the general case of a compact or noncompact X_n we must distinguish several possibilities. Let the metric of the $(n-1)$-dimensional homogeneous space X_{n-1} be determined by the form (1) with given p [number of pluses in (1)], q [number of minuses in (1)], $p+q=n$, and the metric of the n-dimensional homogeneous space X_n by p', q' $(p'+q'=n+1)$. Then we consider the following cases:

I. $p=n, \quad q=0; \quad p'=n+1, \quad q'=0.$

This case was solved by Granzow.[4] We shall denote it by C^+ to express the fact that X_{n-1} is a compact space and the new coordinate x_{n+1} of X_n appears in (1) with a plus sign.

II. $p=n, \quad q=0; \quad p'=n, \quad q'=1$ (the C^--type).

III. $0<p<n, q=n-p; p'=p, \quad q'=q+1$ (the N^--type because X_{n-1} is noncompact).

All the possible rotational groups are covered by these three cases.

Let us suppose now that we have constructed the polar coordinate system in the n-dimensional flat space M_n,

$$x^k = R\cdot f^k(\Theta), \qquad k=1,\cdots,n, \tag{2}$$

where[5] $\Theta \equiv (\vartheta^1, \vartheta^2, \cdots, \vartheta^{n-1})$.

We have

$$g_{kl}(M_n)\cdot f^k(\Theta)\cdot f^l(\Theta) = 1, \tag{3}$$

where $g_{kl}(M_n)$ is the metric tensor in the flat space M_n. This parametrization induces the following form of the metric tensor $g_{\kappa\lambda}(X_{n-1})$ in the homogeneous space X_{n-1}:

$$g_{\kappa\lambda}(X_{n-1}) = g_{kl}(M_n)\frac{\partial f^k(\Theta)}{\partial\vartheta^\kappa}\frac{\partial f^l(\Theta)}{\partial\vartheta^\lambda}. \tag{4}$$

Now the prescription for obtaining our polar coordinates in $n+1$ dimensions is different for different types of the above-mentioned transformations. Let us treat them separately.

[5] Coordinates in *flat* spaces are labeled by *Latin* indices, namely by $k, l, m, \cdots$, if they run from 1 to n and $r, s, t, \cdots$ if they run from 1 to $(n+1)$. Coordinates in *homogeneous* spaces are labeled by *Greek* indices, similarly, by $\kappa, \lambda, \mu, \cdots$ if they run from 1 to $n-1$ and $\rho, \sigma, \tau, \cdots$ if they run from 1 to n. If no other indication is given, we employ the Einstein summation convention over dummy indices.

I. The C^+ Case

The polar coordinate system in E_{n+1} is defined by the transformation

$$\begin{aligned} x^k &= R\sin\vartheta^n\cdot f^k(\Theta) \quad k=1,\cdots,n; \quad \vartheta^1\in\langle 0,2\pi\rangle \\ x^{n+1} &= R\cos\vartheta^n \quad \vartheta^i\in\langle 0,\pi\rangle \quad i=2,\cdots,n, \end{aligned} \tag{5}$$

i.e.,

$$x^r = R\, f^r(\Theta,\vartheta^n) \qquad r=1,\cdots,n+1,$$

where

$$f^r(\Theta,\vartheta^n) = \begin{cases} \sin\vartheta^n\cdot f^k(\Theta) & \text{if } r=k=1,\cdots,n, \\ \cos\vartheta^n & \text{if } r=n+1. \end{cases} \tag{6}$$

The metric tensor $g_{\rho\sigma}(X_n)$ of the corresponding homogeneous space X_n has the form

$$\begin{aligned} g_{\kappa\lambda}(X_n) &= \sin^2\vartheta^n\cdot g_{\kappa\lambda}(X_{n-1}), \\ g_{n\rho}(X_n) &= \delta_{n\rho} \end{aligned} \tag{7}$$

[$g_{\kappa\lambda}(X_{n-1})$ is independent of ϑ^n].

II. The C^- Case

In this case the new polar coordinates are

$$\begin{aligned} &x^k = R\cosh\vartheta^n\cdot f^k(\Theta), \qquad k=1,\cdots,n; \\ &x^{n+1} = R\sinh\vartheta^n, \\ &\vartheta^1\in\langle 0,2\pi\rangle, \qquad \vartheta^i\in\langle 0,\pi\rangle, \qquad i=1,\cdots,n-1, \\ &\vartheta^n\in(-\infty,\infty), \end{aligned} \tag{8}$$

i.e.,

$$x^r = R\, f^r(\Theta,\vartheta^n) \qquad r=1,\cdots,n+1, \tag{9}$$

where

$$f^r(\Theta,\vartheta^n) = \begin{cases} \cosh\vartheta^n\cdot f^k(\Theta) & \text{if } r=k=1,\cdots,n, \\ \sinh\vartheta^n & \text{if } r=n+1. \end{cases}$$

The corresponding metric tensor $g_{\rho\sigma}(X_n)$ is given by

$$\begin{aligned} g_{\kappa\lambda}(X_n) &= \cosh^2\vartheta^n\cdot g_{\kappa\lambda}(X_{n-1}), \\ g_{n\rho}(X_n) &= -\delta_{n\rho}, \end{aligned} \tag{10}$$

where $g_{\kappa\lambda}(X_{n-1})$ is independent of ϑ^n.

III. The N^- Case

The same transformation as in the case II holds also in the N^-case, the only difference being in the

range of variation of the coordinates. In the N^- case we have

$$\left.\begin{aligned} &\vartheta^1 \in \langle 0, 2\pi\rangle \\ &\vartheta^i \in \langle 0, \pi\rangle \quad i = 2, \cdots, p-1 \\ &\vartheta^j \in (-\infty, \infty) \quad j = p, \cdots, n \end{aligned}\right\} \text{ for } p > 1, \quad (11)$$

$$\vartheta^i \in (-\infty, \infty) \quad i = 1, 2, \cdots, n \text{ for } p = 1.$$

Using the above-mentioned formulas for metric tensors $g_{\rho\sigma}(X_n)$, we can construct an invariant operator $\Delta(X_n)$ on the homogeneous space X_n:

$$\Delta(X_n) = [g(X_n)]^{-\frac{1}{2}} \frac{\partial}{\partial\vartheta^\rho} (g(X_n))^{\frac{1}{2}} g^{\rho\sigma}(X_n) \frac{\partial}{\partial\vartheta^\sigma}, \quad (12)$$

where

$$g(X_n) = |\det (g_{\rho\sigma}(X_n))|.$$

This operator is the lowest-order Casimir operator of the group $L(p', q')$, which leaves X_n invariant. Due to the properties (7) and (10) of the metric tensor $g_{\rho\sigma}(X_n)$ the operator $\Delta(X_n)$ always decomposes into two parts. Namely, in the case C^+ we obtain

$$\Delta(X_n) = \frac{1}{\sin^{n-1}\vartheta^n} \frac{\partial}{\partial\vartheta^n} \sin^{n-1}\vartheta^n \frac{\partial}{\partial\vartheta^n} + \frac{\Delta(X_{n-1})}{\sin^2\vartheta^n}, \quad (13)$$

and in both C^- and N^-

$$\Delta(X_n) = -\frac{1}{\cosh^{n-1}\vartheta^n} \frac{\partial}{\partial\vartheta^n} \cosh^{n-1}\vartheta^n \frac{\partial}{\partial\vartheta^n} + \frac{\Delta(X_{n-1})}{\cosh^2\vartheta^n}. \quad (14)$$

The eigenvalues $(-\alpha_{n-1})$ of $\Delta(X_{n-1})$ and eigenvalues $(-\alpha_n)$ of $\Delta(X_n)$ are in all cases of the form

$$\alpha_{n-1} = l_{n-1}(l_{n-1} + n - 2); \quad \alpha_n = l_n(l_n + n - 1). \quad (15)$$

However, the range of l_k depends essentially on the circumstance whether the corresponding group is compact or noncompact. In the case of a compact group, the corresponding l_k are nonnegative integers, whereas in the case of a noncompact one they have, in addition, the form

$$l_k = -\tfrac{1}{2}(k-1) - i\lambda_k, \quad (16)$$

where λ_k is an arbitrary real number and k is the number of dimensions of X_k.[6]

In the present paper, we shall restrict ourselves to the continuous spectra of the Δ operators of the noncompact $L(p, q)$ group as well as its noncompact subgroups.

[6] Assuming the eigenvalues of the $\Delta(X_n)$ operator in the form (15) and (16) we obtain the complete spectrum. See R. Raczka, N. Limić, and J. Niederle, ICTP preprint IC/66/2, Trieste; N. Limić, J. Niederle, and R. Raczka, ICTP preprint, IC/66/18, Trieste.

3. SOLUTIONS OF DIFFERENTIAL EQUATIONS FOR GENERALIZED SPHERICAL FUNCTIONS

Since the operator $\Delta(X_n)$ decomposes into two parts as given in Eqs. (13) and (14) its eigenfunction $\Psi(\Theta, \vartheta^n)$ belonging to the eigenvalue $(-\alpha_n)$ can be taken as a product of two functions

$$\Psi(\Theta, \vartheta^n) = \psi(\Theta)\cdot\varphi(\vartheta^n), \quad (17)$$

where $\psi(\Theta)$, $\Theta \equiv (\vartheta^1, \vartheta^2, \cdots, \vartheta^{n-1})$ is the eigenfunction of $\Delta(X_{n-1})$ with eigenvalue $(-\alpha_{n-1})$. Thus the function $\varphi(\vartheta^n)$ must fulfill the equation

$$\left(\frac{1}{\sin^{n-1}\vartheta^n} \frac{d}{d\vartheta^n} \sin^{n-1}\vartheta^n \frac{d}{d\vartheta^n} - \frac{\alpha_{n-1}}{\sin^2\vartheta^n}\right)\varphi(\vartheta^n) = -\alpha_n\varphi(\vartheta^n) \quad (18)$$

in the compact case (C^+), while in the noncompact one (C^- or N^-)

$$\left(\frac{-1}{\cosh^{n-1}\vartheta^n} \frac{d}{d\vartheta^n} \cosh^{n-1}\vartheta^n \frac{d}{d\vartheta^n} - \frac{\alpha_{n-1}}{\cosh^2\vartheta^n}\right)\varphi(\vartheta^n) = -\alpha_n\varphi(\vartheta^n). \quad (19)$$

The solutions of these equations are the following:

I. C^+ case

(a) n-even

$$\varphi(\vartheta^n) = (\sin^{n-2}\vartheta^n)^{-\frac{1}{2}} P^{l_{n-1}+(n-2)/2}_{l_n+(n-2)/2}(\cos\vartheta^n); \quad (23)$$

(b) n-odd

$$\begin{aligned} n \neq 1 \quad \varphi(\vartheta^n) &= (\sin^{n-3}\vartheta^n)^{-\frac{1}{2}} \Pi^{l_{n-1}+(n-3)/2}_{l_n+(n-1)/2}(\vartheta^n) \\ n = 1 \quad \varphi(\vartheta^1) &= e^{il_1\vartheta^1}. \end{aligned} \quad (24)$$

Here, for both cases (a), (b), $l_n \geq |l_{n-1}|$ and both are nonnegative integers except for l_1, which is an integer.

II. C^- case

$$\varphi(\vartheta^n) = (\cosh^{n-1}\vartheta^n)^{-\frac{1}{2}} P^{i\lambda_n}_{l_{n-1}+(n-3)/2}(\tanh\vartheta^n), \quad (25)$$

where l_{n-1} are nonnegative integers, λ_n are real numbers.

III. N^- case

$$\varphi(\vartheta^n) = (\cosh^{n-1}\vartheta^n)^{-\frac{1}{2}} P^{i\lambda_n}_{-\frac{1}{2}+i\lambda_{n-1}}(\tanh\vartheta^n), \quad (26)$$

where λ_n and λ_{n-1} are both real numbers.

Functions $P^{ia}_{-\frac{1}{2}+ib}(x)$ are the conical functions (see, e.g., Bateman[7]) and the functions $\Pi^l_n(\vartheta)$ are the spherical orthonormalized functions $\Pi_l(n, \vartheta)$ introduced by Fock.[7]

[7] H. B. Bateman in *Higher Transcendental Functions*, edited by A. Erdélyi (McGraw-Hill Book Company, Inc., New York, 1953), Vol. I, p. 174; V. A. Fock, Z. Physik **98**, 145 (1935).

Thus the eigenfunctions $\Psi(\Theta, \vartheta^n)$ of the Casimir operator $\Delta(X_n)$ of an arbitrary group $L(p', q')$ can be expressed in terms of the conical and spherical functions in the following way

(a) p-even

$$\begin{aligned}\langle l_1, l_2, l_3, \cdots, l_{p-2}, l_{p-1}, \lambda_p, \lambda_{p+1}, \cdots, \lambda_{n-1}, \lambda_n \mid \vartheta^1, \vartheta^2, \cdots, \vartheta^n\rangle \\ = e^{il_1\vartheta^1}\cdot P^{l_1}_{l_2}(\cos\vartheta^2)\,\Pi^{l_2}_{l_3+1}(\vartheta^3)\cdots(\sin^{p-4}\vartheta^{p-2})^{-\frac{1}{2}}P^{l_{p-3}+(p-4)/2}_{l_{p-2}+(p-4)/2}(\cos\vartheta^{p-2}) \\ \cdot(\sin^{p-4}\vartheta^{p-1})^{-\frac{1}{2}}\,\Pi^{l_{p-2}+(p-4)/2}_{l_{p-1}+(p-2)/2}(\vartheta^{p-1})(\cosh^{p-1}\vartheta^p)^{-\frac{1}{2}}P^{i\lambda_p}_{l_{p-1}+(p-3)/2}(\tanh\vartheta^p)(\cosh^p\vartheta^{p+1})^{-\frac{1}{2}} \\ \cdot P^{i\lambda_{p+1}}_{-\frac{1}{2}+i\lambda_p}(\tanh\vartheta^{p+1})\cdots(\cosh^{n-2}\vartheta^{n-1})^{-\frac{1}{2}}P^{i\lambda_{n-1}}_{-\frac{1}{2}+i\lambda_{n-2}}(\tanh\vartheta^{n-1})(\cosh^{n-1}\vartheta^n)^{-\frac{1}{2}}P^{i\lambda_n}_{-\frac{1}{2}+i\lambda_{n-1}}(\tanh\vartheta^n);\end{aligned} \tag{27}$$

(b) p-odd

$$\begin{aligned}\langle l_1, l_2, l_3, \cdots, l_{p-2}, l_{p-1}, \lambda_p, \lambda_{p+1}, \cdots, \lambda_{n-1}, \lambda_n \mid \vartheta^1, \vartheta^2, \cdots, \vartheta^n\rangle \\ = e^{il_1\vartheta^1}\cdot P^{l_1}_{l_2}(\cos\vartheta^2)\cdot\Pi^{l_2}_{l_3+1}(\vartheta^3)\cdots(\sin^{p-5}\vartheta^{p-2})^{-\frac{1}{2}}\,\Pi^{l_{p-3}+(p-5)/2}_{l_{p-2}+(p-3)/2}(\vartheta^{p-2}) \\ \cdot(\sin^{p-3}\vartheta^{p-1})^{-\frac{1}{2}}P^{l_{p-3}+(p-3)/2}_{l_{p-1}+(p-3)/2}(\cos\vartheta^{p-1})(\cosh^{p-1}\vartheta^p)^{-\frac{1}{2}}P^{i\lambda_p}_{l_{p-1}+(p-3)/2}(\tanh\vartheta^p) \\ \cdot(\cosh^p\vartheta^{p+1})^{-\frac{1}{2}}P^{i\lambda_{p+1}}_{-\frac{1}{2}+i\lambda_p}(\tanh\vartheta^{p+1})\cdots(\cosh^{n-2}\vartheta^{n-1})^{-\frac{1}{2}} \\ \cdot P^{i\lambda_{n-1}}_{-\frac{1}{2}+i\lambda_{n-2}}(\tanh\vartheta^{n-1})(\cosh^{n-1}\vartheta^n)^{-\frac{1}{2}}P^{i\lambda_n}_{-\frac{1}{2}+i\lambda_{n-1}}(\tanh\vartheta^n).\end{aligned} \tag{28}$$

Since the set of Casimir operators $\Delta(X_n), \cdots, \Delta(X_1)$ forms a complete set of commuting Hermitian operators the set of their common eigenfunctions is a complete set of orthogonal normalizable functions with proper as well as improper norm. It means that every square-integrable function on X_n can be expanded in terms of (27) or (28) and of the eigenfunctions related to the discrete spectrum.

Our continuous unitary representations of the $L(p, q+1)$ group are realized in the Hilbert space of generalized Fourier tranforms[6]

$$\begin{aligned}\chi^{\lambda_p,\cdots,\lambda_n}_{l_1,\cdots,l_{p-1}} = \int_{X_n}\overline{\Psi^{\lambda_p,\cdots,\lambda_n}_{l_1,\cdots,l_{p-1}}(\theta, \vartheta^n)} \\ \cdot f(\theta, \vartheta^n)[g(X_n)]^{\frac{1}{2}}\,d\theta\,d\vartheta^n,\end{aligned} \tag{29}$$

where $\Psi^{\lambda_p,\cdots,\lambda_n}_{l_1,\cdots,l_{p-1}}(\theta, \vartheta^n)$ is a function given by (27) or (28) and $f(\theta, \vartheta^n) \in \mathsf{C}^\infty_0(X_n)$. The scalar *product* in this Hilbert space is defined by

$$\sum_{l_1,\cdots,l_{p-1}}\int g\lambda_p\cdots d\lambda_{n-1}\,\chi^{\lambda_p,\cdots,\lambda_n}_{l_1,\cdots,l_{p-1}}\,\phi^{\lambda_p,\cdots,\lambda_n}_{l_1,\cdots,l_{p-1}}$$

and the norm is finite.

4. THE GROUPS OF THE LORENTZ TYPE

The coordinate systems (5) and (8) give a complete description of any homogeneous space X invariant under the group $L(p, q)$. However, in physical applications it is desirable to have as many discrete quantum numbers as possible. For this reason if $p < q$ it is convenient to introduce another parametrization of the X. Let us consider as an example the groups of the Lorentz type $L(1, n)$, which conserve the form

$$-x_1^2 - x_2^2 - \cdots - x_n^2 + x_{n+1}^2 = R^2. \tag{30}$$

If we apply to these groups the procedure used in the N^- case we get all n quantum numbers continuous. We shall develop now a procedure which yields eigenfunctions of $\Delta(X_n)$ labeled by $n-1$ discrete quantum numbers and only one continuous.

Let us start from a homogeneous space X_{n-1} with an invariant form

$$x_1^2 + x_2^2 + \cdots + x_n^2 = R^2. \tag{31}$$

Now, assuming that the polar coordinate system of E_n has been defined by the procedure used by us in the C^+ case [Eq. (5)], we introduce the following coordinate system in M_{n+1}:

$$\begin{aligned}x^k &= R\sinh\vartheta^n\cdot f^k(\Theta) \quad k = 1, \cdots, n \quad \vartheta^1 \in \langle 0, 2\pi\rangle, \\ x^{n+1} &= R\cosh\vartheta^n \quad \vartheta^i \in \langle 0, \pi\rangle \quad i = 2, \cdots, n-1 \\ &\qquad \vartheta^n \in \langle 0, \infty\rangle,\end{aligned} \tag{32}$$

i.e.,

$$\begin{aligned}x^r &= Rf^r(\Theta, \vartheta^n) \qquad r = 1, \cdots, n+1, \\ f^r(\Theta, \vartheta^n) &= \sinh\vartheta^n\cdot f^k(\Theta) \quad \text{if} \quad r = k = 1, \cdots, n, \\ &= \cosh\vartheta^n \qquad\qquad \text{if} \quad r = n+1.\end{aligned} \tag{33}$$

The metric tensor $g_{\rho\sigma}(X_n)$ of the corresponding homogeneous space X_n has the form

$$\begin{aligned}g_{\kappa\lambda}(X_n) &= -\sinh^2\vartheta^n\cdot g_{\kappa\lambda}(x_{n-1}), \\ g_{n\rho}(X_n) &= -\delta_{n\rho},\end{aligned} \tag{34}$$

where $g_{\kappa\lambda}(X_{n-1})$ is independent of ϑ^n.

Let us construct now the invariant operator $\Delta(X_n)$ on the homogeneous space according to the formulas (12). We obtain

$$\Delta(X_n) = \frac{-1}{\sinh^{n-1}\vartheta^n}\frac{\partial}{\partial\vartheta^n}\sinh^{n-1}\vartheta^n\frac{\partial}{\partial\vartheta^n} - \frac{\Delta(X_{n-1})}{\sinh^2\vartheta^n}. \quad (35)$$

The spectrum of α_{n-1} is discrete

$$\begin{aligned} \alpha_{n-1} &= l_{n-1}(l_{n-1} + n - 2) \\ l_{n-1} &\geqq l_{n-2} \geqq \cdots \geqq |l_i| \geqq 0 \quad \text{integer}, \end{aligned} \quad (36)$$

whereas α_n is continuous

$$\begin{aligned} \alpha_n &= l_n(l_n + n - 1), \\ l_n &= -\tfrac{1}{2}(n-1) + i\lambda_n, \end{aligned} \quad (37)$$

λ_n being arbitrary real.

The separation of variables (17) leads to the following differential equation for $\varphi(\vartheta^n)$:

$$\left(-\frac{1}{\sinh^{n-1}\vartheta^n}\frac{d}{d\vartheta^n}\sinh^{n-1}\vartheta^n\frac{d}{d\vartheta^n} + \frac{\alpha_{n-1}}{\sinh^2\vartheta^n}\right)\varphi(\vartheta^n) = -\alpha_n\varphi(\vartheta^n) \quad (38)$$

and its solution is

$$\varphi(\vartheta^n) = (\sinh^{n-2}\vartheta^n)^{-\frac{1}{2}}P^{-[l_{n-1}+(n-2)/2]}_{-\frac{1}{2}+i\lambda_n}(\cosh\vartheta^n). \quad (39)$$

The function $\Psi(\Theta, \vartheta^n)$ has the form

$$\Psi(\Theta, \vartheta^n) = \psi(\Theta)\varphi(\vartheta^n), \quad (40)$$

where $\psi(\Theta)$ is the solution of the compact C^+ case in the X_{n-1} homogeneous space. For the case of the Lorentz group $L(1, 3)$ our result agrees with Smorodinskiĭ's and Vilenkin's paper.[8]

DISCUSSION

It was mentioned several times[9] that in the theory of elementary particles the most important role is played by the simplest representations of the noncompact group, namely be those for which all except one of the eigenvalues labeling Casimir operators are independent. Just these representations were constructed in this paper. It follows from the fact that the infinite-dimensional Hilbert spaces $\mathfrak{H}$ in which our continuous unitary representation $T^{\lambda_n-1}_{L(p,q)}$ of the group $L(p, q)$ acts, is fixed up by one eigenvalue α_{n-1} of the Casimir operator $\Delta(X_{n-1})$ (Sec. 2). The complete set of functions which span $\mathfrak{H}$ is determined by $(p-1)$ discrete parameters $l_1, \cdots, l_{p-1}$ and q continuous ones $\lambda_p, \cdots, \lambda_{n-1}$. Then two groups of quantum numbers are related with the invariant operators of the compact and noncompact subgroups, respectively. It is interesting that if the representation $T^{\lambda_{n-1}}_{L(p,q)}$ of $L(p, q)$ is determined by one independent Casimir operator

$$Q^1_{L(p,q)} = \Delta(X_{n-1}), \qquad Q^i_{L(p,q)} = f^i(Q^1_{L(p,q)}) \quad i = 2, \cdots, [\tfrac{1}{2}(p+q)],$$

then the representations

$$T^{\lambda_{i+j-1}}_{L(i,j)} \qquad (i = 2, \cdots, p;\ j = 1, \cdots, q-1)$$

of the groups

$$L(p, q-1), L(p, q-2), \cdots, L(p, 0), L(p-1, 0), \cdots, L(2, 0),$$

which appear in the decomposition of $T_{L(p,q)}$, are also determined by the conditions

$$Q^1_{L(i,j)} = \Delta(X_{i+j-1}) \quad \text{and} \quad Q^s_{L(i,j)} = f^s(Q^1_{L(i,j)}) \quad s = 2, \cdots, [\tfrac{1}{2}(i+j-1)].$$

Roughly speaking, the representations of the noncompact and compact subgroups $L(i, j)$ are of the same kind as the representations of the $L(p, q)$.

The discrete quantum numbers, which characterize the representations $T^{\lambda_{i+j-1}}_{L(i,j)}$ of the subgroups $L(i, j)$, are not independent. For the compact subgroups we have[10]

$$Q^1_{L(p+1,0)}(x_p) = \sum_{\mu,\nu=1}^{p+1} M_{\mu\nu}M^{\mu\nu} = Q^1_{L(p,0)}(x_{p-1}) + \sum_{a=1}^{p} M^2_{pa} \quad (41)$$

and since the last term is positive-definite we have $\lambda_p \geqq \lambda_{p-1}$.

If we take nonnegative integral values for l_n instead of $l_n = -\frac{1}{2}(n-1) + i\lambda_n$ for the eigenvalues of $\Delta(X_n)$ in the *noncompact* case, we obtain the infinite-dimensional unitary representations of the noncompact groups except for the Lorentz type groups, where we obtain the finite dimensional nonunitary representations. These representations are discussed in Ref. 6 and in the special case of the Lorentz group by Dolginov.[2] If we take for l_n a real number instead of $l_n = -\frac{1}{2}(n-1) + i\lambda_n$, then the corresponding functions are not orthogonal with respect to the measure $g^{\frac{1}{2}}\prod_{i=1}^{n}d\vartheta^i$ as was shown by Alcaras and Ferreira[10] in the case of the Lorentz group. But it

[8] N. Ya. Vilenkin and V. A. Smorodinskiĭ, Zh. Eksperim. i Teor. Fiz. **46**, 1793 (1964) [English transl.: Soviet Phys.—JETP **19**, 1209 (1964)].

[9] A. O. Barut, Phys. Rev. **139**, B1433 (1965). A. O. Barut, P. Budini, and C. Fronsdal, Preprint IC/65/34, Trieste (1965). M. Gell-mann, Lecture at Trieste Seminar (1965).

[10] J. A. C. Alcaras and P. L. Ferreira, J. Math. Phys. **6**, 578 (1965).

does not mean that the set of eigenfunctions of $\Delta(X_n)$ does not form an orthogonal basis of the Hilbert space, because we know that for the supplementary series of irreducible unitary representations we have to use the double point measure.[11] This problem will be discussed elsewhere.

We would like to emphasize that our method may be easily generalized and used to obtain other series of irreducible unitary representations of the groups considered. To achieve this we have to multiply the scalar functions $\Psi(\Theta, \vartheta^n)$ obtained here by spinor or tensor functions, similarly, as was done by Dolginov and Moskalev[2] in the case of the Lorentz group.

Recently, a number of noncompact symmetry groups have been introduced into the theory of elementary particles.[12] However, until now, mostly finite-dimensional (i.e., nonunitary) representations of these groups have been considered, whereas from the physical point of view the unitary representations are of the main interest because they conserve probability. On the other hand, if we try to use the unitary representations in the physics of elementary particles, we have to solve the following problems[12]:

(i) Determination of a set of quantum numbers characterizing a physical supermultiplet in a given unitary representation.

(ii) Decomposition of a tensor product of two representations.

(iii) Determination of representations of compact subgroups which occur in the decomposition of a product of representations of the considered noncompact group.

All these problems may be solved by our approach. In particular, the problem (ii) may be reduced to the decomposition of the product of two functions of the type (27) and (28), and the solution of the problem (iii) follows from the method we used for the construction of the complete set of commuting operators.

ACKNOWLEDGMENTS

The authors would like to thank Professor A. Salam and the International Atomic Energy Agency for the hospitality extended to them at the International Centre for Theoretical Physics.

It is a pleasure to thank Professor A. O. Barut and Professor J. Werle as well as Dr. J. Tolar for valuable discussions.

[11] I. M. Gel'fand, M. I. Graev, and N. Ya. Vilenkin, *Generalized Functions*, Vol. 5 (Academic Press Inc., New York, in press).

[12] R. Delbourgo, A. Salam, and J. Strathdee, Nuovo Cimento **36**, 689 (1965); T. Fulton and J. Wess, Phys. Letters **15**, 177 (1965). P. Budini and C. Fronsdal, *ibid.* **14**, 968 (1965). C. Fronsdal, Proc. Roy. Soc. (London) **A288**, 98 (1965). P. G. O. Freund, Phys. Rev. Letters **14**, 803 (1965).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 7, NUMBER 5 MAY 1966

Direct-Product Representations of the Canonical Commutation Relations

J. R. KLAUDER AND J. MCKENNA
Bell Telephone Laboratories, Inc., Murray Hill, New Jersey
AND
E. J. WOODS*
Department of Physics and Astronomy, University of Maryland, College Park, Maryland
(Received 5 August 1965)

We consider direct-product representations of the canonical commutation relations. An irreducible representation is defined on each of the incomplete direct-product spaces (IDPS) of von Neumann. We prove that two such representations are unitarily equivalent if and only if the corresponding IDPS are weakly equivalent, for which simple analytic tests exist. The matrix elements of these representations, coupled with a Friedrichs–Shapiro type of integral, fulfill group orthogonality relations. This classification into unitary equivalence classes also applies to direct-product representations of the canonical anticommutation relations.

1. INTRODUCTION

AMONG all the representations of the canonical commutation relations (CCR) for a scalar boson field, those defined as product representations have a certain simplicity that makes their study attractive. From a physical point of view a product representation may be associated with the ground state of a hypothetical dynamical system composed of countably many independent degrees of freedom, each having a suitable but arbitrary Hamiltonian expressed in the Schrödinger representation. Such simple dynamical systems already require the use of inequivalent representations of the CCR, the existence of which was discovered in certain examples by Friedrichs[1] and by Segal,[2] and popularized in the special models of van Hove.[3] These examples alone show the relevance of various inequivalent representations of the CCR and the importance of choosing the right representation for the right application. However, the specialized techniques used earlier to test for equivalence were unable to provide a simple, yet general, classification of product representations into unitary equivalence classes. In the present paper we extend and apply two techniques introduced elsewhere[4] (CRT V) to provide a simple and complete unitary equivalence classification of product representations.

Section 2 is devoted to a rigorous definition of a CCR representation in the Weyl form as a ray representation of a group of unitary operators. Section 3 deals with the construction of the incomplete direct product spaces (IDPS) of von Neumann[5] which provide a natural setting to define product representations. Our canonical construction assigns one irreducible representation to each IDPS, and thus is a generalization of the construction in CRT V where test function requirements limited the class of IDPS studied. In Sec. 4 we develop a kind of "group average," an integral of the Friedrichs–Shapiro type,[6] for arbitrary pairs of matrix elements of the unitary group elements from any two spaces, and established an analog of the group orthogonality relations for compact groups. In particular, we prove that the group average for matrix elements from two different representations vanishes identically if and only if the representations are unitarily inequivalent. In Sec. 5 an independent proof of the condition for unitary inequivalence is given by explicitly constructing unitary invariant c-numbers, "tags," for every product representation. By establishing conditions for the inequality of two such "tags" we are able to infer the unitary inequivalence of the corresponding representations. (The discussion in Sec. 5 does not draw on that in Sec. 4 and either section may be read after Sec. 3.)

The essence of our principal result may be stated in terms of the hypothetical dynamical systems discussed above. Suppose two such dynamical

* Supported in part by National Science Foundation Grant GP3221, and the General Research Board of the University of Maryland.

[1] K. O. Friedrichs, *Mathematical Aspects of the Quantum Theory of Fields* (Interscience Publishers, Inc. New York, 1953).

[2] See I. E. Segal, *Mathematical Problems of Relativistic Physics* (American Mathematical Society, Providence, Rhode Island, 1963) and additional references therein.

[3] L. van Hove, Physica **18**, 145 (1952).

[4] J. R. Klauder and J. McKenna, J. Math. Phys. **6**, 68 (1965), referred to hereafter as CRT V.

[5] J. von Neumann, Composito Math. **6**, 1 (1938).

[6] K. O. Friedrichs and H. N. Shapiro, "Integration of Functionals," Lecture Notes, New York Institute of Mathematical Sciences, 1957, Chap. 1.

systems are given and the normalized ground states pertaining to the nth degree of freedom are denoted by χ_n^1 and χ_n^2, respectively, both of which may be arbitrary unit vectors. Then unitary equivalence or inequivalence of the scalar field CCR for the two systems is governed by whether the sum

$$\sum_{n=1}^{\infty} \left| |(\chi_n^2, \chi_n^1)| - 1 \right|$$

converges or diverges, respectively. In the language of von Neumann's theory, our result states that canonical CCR representations defined on weakly (in)equivalent IDPS are unitarily (in)equivalent.

A certain amount of confusion exists in the literature regarding the interpretation and properties of representations pertaining to an infinite number of degrees of freedom. The convenient characterization of certain of such representations in tensor product spaces has perhaps fostered the false impression that nonseparable Hilbert spaces are required, and that in certain models "the interaction maps every state out of Hilbert space." Firstly, field representations are completely definable in separable Hilbert spaces as is illustrated by the usual direct sum formulation of the Fock representation. Secondly, any sequence of operators all of whose matrix elements vanish in the limit simply converges (weakly) to the zero operator. This is just the set of circumstances that so often prevails for the S matrix in the interaction representation when a cutoff form factor is removed. The vanishing of the matrix elements should not be regarded as "mapping the states out of Hilbert space," but is indicative of an ill-chosen representation of field operators, one for which the sequence of unitary S matrices with form factors does not converge to a unitary operator as the form factor is removed. It is for this reason that the selection of the right CCR representation for the application becomes crucial. But we must emphasize that only in product cases is it clear that an appropriate CCR representation can be found; in relativistic cases it is by no means clear that a correct representation even exists.[7]

2. REPRESENTATIONS OF THE CANONICAL COMMUTATION RELATIONS

For the sake of completeness, we recall here the rigorous mathematical definition of a representation of the CCR. The common definition of the commutation relations in the Heisenberg form proves unsuitable for rigorous mathematical study. This is due both to the distribution nature of the field operators and to the fact that they are unbounded. Both difficulties are overcome in the Weyl form of the commutation relations, the definition of which is given here.[8]

Definition 2.1: Let $\mathfrak{U}$ and $\mathfrak{W}$ be real vector spaces (the "test function" spaces) and let $f \times g \to (f, g)$, $f \in \mathfrak{U}$, $g \in \mathfrak{W}$ be a nondegenerate bilinear form from $\mathfrak{U} \times \mathfrak{W}$ into R, the field of real numbers. Let $\mathfrak{H}$ be a complex Hilbert space, and $\mathcal{G}$ the group of all unitary operators in $\mathfrak{H}$. Then a representation of the CCR is a pair of maps $\mathfrak{U} \xrightarrow{V} \mathcal{G}$ and $\mathfrak{W} \xrightarrow{W} \mathcal{G}$ which satisfy

$$W[f]W[f'] = W[f + f'], \tag{2.1}$$

$$V[g]V[g'] = V[g + g'], \tag{2.2}$$

$$V[g]W[f] = e^{-i(f,g)}W[f]V[g], \tag{2.3}$$

such that for each f, g the operators $V[tg]$ and $W[tf]$ are weakly continuous functions of the real variable t.

The spaces $\mathfrak{U}$ and $\mathfrak{W}$ are the "test function" spaces that are required by the distribution properties. The weak continuity condition ensures that the field operators can be recovered. For it follows from Stone's theorem[9] that, for each f and g, self-adjoint operators $\varphi(f)$ and $\pi(g)$ exist such that

$$W(f) = e^{i\varphi(f)},$$

$$V(g) = e^{-i\pi(g)}.$$

3. DIRECT-PRODUCT REPRESENTATIONS

In this paper we are concerned with those irreducible representations of the CCR which can be constructed as incomplete direct products of Schrödinger representations of the commutation relations for one degree of freedom. In this section we outline the construction of these representations. We assume the reader to be familiar with von Neumann's results. Since they will be needed later,

[7] However, if CCR exist for an interacting relativistic field, then Haag's theorem [see, e.g., R. F. Streater and A. S. Wightman, *PCT, Spin & Statistics, and All That*, (W. A. Benjamin Company, Inc., New York, 1964), Chap. 4] requires the representation to be unitarily inequivalent to that of a free field.

[8] More complete discussions of Definition 2.1 as well as other examples of representations of the CCR can be found in: L. Gärding and A. S. Wightman, Proc. Nat. Acad. Sci. 40, 617 (1954); A. S. Wightman and S. S. Schweber, Phys. Rev. 98, 812 (1955); I. E. Segal, Trans. Am. Math. Soc. 88, 12 (1958); J. S. Lew, thesis, Princeton University, 1960 (unpublished); H. Araki and E. J. Woods, J. Math. Phys. 4, 637 (1963). Recently D. Shelupsky has classified into unitary equivalence classes some of the representations constructed by Wightman and Schweber. D. Shelupsky, J. Math. Phys. 7, 163 (1966).

[9] F. Riesz and B. Sz.-Nagy, *Functional Analysis* (Frederick Unger Publishing Company, New York, 1955), p. 380.

we will also define irreducible, direct-product representations of the CCR for N degrees of freedom, where N is finite.

Let $\mathfrak{H}_n = L^2(R_1)$, $n = 1, 2, \cdots$, and let the Schrödinger representation be defined in $L^2(R_1)$ by the self-adjoint operators Q and P satisfying $[Q, P] = iI$. It is well known that Q and P are irreducible in $L^2(R_1)$, and the Weyl form of the representation for one degree of freedom is given by

$$V_0[q] = e^{-iqP}, \qquad W_0[p] = e^{ipQ}. \tag{3.1}$$

The Weyl form of the Schrödinger representation for N degrees of freedom ($N < \infty$) can be constructed as follows. We note that

$$L^2(R_N) = \prod_{n=1}^{N} \otimes \mathfrak{H}_n,$$

and if $\mathfrak{p} = (p_1, \cdots, p_N)$, $\mathfrak{q} = (q_1, \cdots, q_N)$ are real N-vectors, then

$$\begin{aligned} V[\mathfrak{q}] &= V_0[q_1] \otimes \cdots \otimes V_0[q_N], \\ W[\mathfrak{p}] &= W_0[p_1] \otimes \cdots \otimes W_0[p_N], \end{aligned} \tag{3.2}$$

defines the desired representation. This is an irreducible representation.

We now construct the direct-product representations for infinitely many degrees of freedom. Let $\mathrm{X} = \prod_{n=1}^{\infty} \otimes \chi_n$, where $||\chi_n|| = 1$ and $\chi_n \in \mathfrak{H}_n$, be a product vector in the complete direct-product space.[5] Let

$$\mathfrak{H}_{\mathrm{X}} = \prod_{n=1}^{\infty} \otimes^{\mathrm{X}} \mathfrak{H}_n \tag{3.3}$$

denote the incomplete direct-product space[5] (IDPS) determined by X. Define on $\mathfrak{H}_{\mathrm{X}}$ unitary operators $V_n[q_n]$ and $W_n[p_n]$, q_n, p_n real, by

$$\begin{aligned} V_n[q_n] &= 1 \otimes \cdots \otimes V_0[q_n] \otimes 1 \otimes \cdots, \\ W_n[p_n] &= 1 \otimes \cdots \otimes W_0[p_n] \otimes 1 \otimes \cdots, \end{aligned} \tag{3.4}$$

where $V_0[q_n]$ and $W_0[p_n]$ occur in the nth place in their respective products, and the other terms in both products are unit operators. Let $\{h_n(x), n = 1, 2, \cdots\}$ be a complete orthonormal set of real-valued functions in $L^2(R_1)$. Let

$$\mathcal{P} = \left\{ f(x) = \sum_{n=1}^{m} \lambda_n h_n(x),\ m < \infty \right\}, \tag{3.5}$$

where the λ_i are arbitrary real numbers, and m takes on all positive, finite, integral values. We set the test function spaces $\mathfrak{V} = \mathfrak{W} = \mathcal{P}$, and define the bilinear form

$$(f, g) = \int_{R_1} fg\, dx, \text{ where } f, g \in \mathcal{P}.$$

Then for any $f, g \in \mathcal{P}$, we define $V[g]$ and $W[f]$ by

$$\begin{aligned} V[g] &= \prod_{n=1}^{m} V_n[q_n], \qquad g = \sum_{n=1}^{m} q_n h_n(x), \\ W[f] &= \prod_{n=1}^{m'} W_n[p_n], \qquad f = \sum_{n=1}^{m'} p_n h_n(x). \end{aligned} \tag{3.6}$$

Since the $V_0[q]$ and $W_0[p]$ are irreducible on each $\mathfrak{H}_n$, the $V[g]$ and $W[f]$, for all $f, g \in \mathcal{P}$, are irreducible on the IDPS $\mathfrak{H}_{\mathrm{X}}$.

We recall von Neumann's definitions of equivalence and weak equivalence for product vectors in the complete tensor product space $\mathfrak{H}$, and state without proof some of their properties.[5]

Definition 3.1: Two product vectors $\mathrm{X} = \prod_{n=1}^{\infty} \otimes \chi_n$ and $\Lambda = \prod_{n=1}^{\infty} \otimes \lambda_n$ are said to be equivalent, $\mathrm{X} \approx \Lambda$, if and only if

$$\sum_{n=1}^{\infty} |(\chi_n, \lambda_n) - 1| < \infty. \tag{3.7}$$

Definition 3.2: Two product vectors X and Λ are said to be weakly equivalent, $\mathrm{X} \overset{w}{\approx} \Lambda$, if and only if there is a sequence of real numbers θ_n such that

$$\prod_{n=1}^{\infty} \otimes \chi_n \approx \prod_{n=1}^{\infty} \otimes e^{i\theta_n} \lambda_n.$$

Lemma 3.1: A necessary and sufficient condition for weak equivalence is

$$\sum_{n=1}^{\infty} |\,|(\chi_n, \lambda_n)| - 1| < \infty. \tag{3.8}$$

If two product vectors are not weakly equivalent then

$$\lim_{N\to\infty} \prod_{n=1}^{N} (\chi_n, \lambda_n) = 0. \tag{3.9}$$

Lemma 3.2: If $\mathrm{X} \approx \Lambda$ then $\prod_{n=1}^{N} (\chi_n, \lambda_n)$ converges and the inner product of X and Λ is defined to be

$$(\mathrm{X}, \Lambda) = \lim_{N\to\infty} \prod_{n=1}^{N} (\chi_n, \lambda_n). \tag{3.10}$$

If $\mathrm{X} \not\approx \Lambda$ then the limit of (3.10) may not exist, but the inner product of X and Λ is defined to be zero.

An IDPS is the closed linear subspace of the complete direct-product space $\mathfrak{H}$ which is spanned by all product vectors in some equivalence class. Our canonical construction has assigned to each IDPS a unique direct-product representation of the CCR, since any two equivalent vectors determine the same IDPS and hence the same repre-

sentation of the CCR. It is also true that representations defined on IDPS $\mathfrak{C}_X$ and $\mathfrak{C}_A$ in the same weak equivalence class are unitarily equivalent.[4] To prove this we use the operator

$$T = \prod_{n=1}^{\infty} \otimes e^{i\theta_n} \tag{3.11}$$

constructed by von Neumann,[5] who showed that it is a unitary map from $\mathfrak{C}_A$ onto $\mathfrak{C}_X$ (where the θ_n are given in Definition 3.2). It is clear from our construction that

$$TV_A[g]T^{-1} = V_X[g], \quad TW_A[f]T^{-1} = W_X[f]. \tag{3.12}$$

One of the main results of our paper, proved independently in Secs. 4 and 5, is that weak equivalence is also a necessary condition for unitary equivalence of the representations.

4. GROUP ORTHOGONALITY RELATIONS AND UNITARY EQUIVALENCE OF DIRECT-PRODUCT REPRESENTATIONS

In this section we construct a Friedrichs–Shapiro-type integral[6] for products of two matrix elements obtained from the irreducible direct-product representations of the CCR. In CRT V this integral was introduced and evaluated for the case where both matrix elements belonged to the same representation of the CCR. These results are extended to the case where the matrix elements belong to different representations. We obtain a direct analog of the group orthogonality relations, and use these results to complete the classification into unitary equivalence classes of the irreducible direct-product representations.

The integral is defined as follows. Let $\mathfrak{C}_{X_i}$, $i = 1, 2$ be two IDPS, and let $W_i[f]$, $V_i[g]$ be the representations of the CCR defined in $\mathfrak{C}_{X_i}$. Define

$$U_i[f, g] = V_i[g]W_i[f]. \tag{4.1}$$

If

$$f = \sum_n p_n h_n, \qquad g = \sum_n q_n h_n,$$

we define the truncated functions $f_{(N)}$, $g_{(N)}$ by

$$f_{(N)} = \sum_{n=1}^{N} p_n h_n, \qquad g_{(N)} = \sum_{n=1}^{N} q_n h_n. \tag{4.2}$$

Let $\Phi_i, \Psi_i \in \mathfrak{C}_{X_i}$ and define

$$\int (U_i[f, g]\Phi_1, \Psi_1)^*(U_2[f, g]\Phi_2, \Psi_2)\, d\mu\,(f, g) = \lim_{N\to\infty} I_N(\Phi_1, \Psi_1; \Phi_2, \Psi_2), \tag{4.3}$$

where

$$I_N(\Phi_1, \Psi_1; \Phi_2, \Psi_2) = \int (U_1[f_{(N)}, g_{(N)}]\Phi_1, \Psi_1)^* \times (U_2[f_{(N)}, g_{(N)}]\Phi_2, \Psi_2)\, d\mu\,(f_{(N)}, g_{(N)}), \tag{4.4}$$

and

$$d\mu\,(f_{(N)}, g_{(N)}) = (2\pi)^{-N} \prod_{n=1}^{N} dp_n\, dq_n. \tag{4.5}$$

Note that the $f_{(N)}$ and $g_{(N)}$ are continuous functions of the p_n, q_n, and the domain of integration is $R_N \times R_N$. It should also be noted that the set of operators $V[g_{(N)}]$ and $W[f_{(N)}]$ form a *reducible* representation of the CCR for N degrees of freedom. We now give two theorems relating to such integrals.

Theorem 4.1: Let $U[\mathbf{p}, \mathbf{q}] = V[\mathbf{q}]W[\mathbf{p}]$ be an irreducible representation of the CCR for a finite number N of degrees of freedom. Then

$$\int (U[\mathbf{p}, \mathbf{q}]\Phi_1, \Psi_1)^*(U[\mathbf{p}, \mathbf{q}]\Phi_2, \Psi_2)\, d\mu\,(\mathbf{p}, \mathbf{q}) = (\Psi_1, \Psi_2)(\Phi_2, \Phi_1), \tag{4.6}$$

where $d\mu(\mathbf{p}, \mathbf{q})$ is identical to (4.5).

Proof: A special case of this theorem when $\Phi_1 = \Phi_2$ and $||\Phi_1|| = 1$ has been rigorously proved in Ref. 10, and is embodied in Eq. (3.16) of Ref. 10. The extension of this earlier proof to cover the present case is trivial and is left to the reader.

Theorem 4.2: In a reducible representation of the CCR for a finite number of degrees of freedom we have

$$\left|\int (U[\mathbf{p}, \mathbf{q}]\Phi_1, \Psi_1)^*(U[\mathbf{p}, \mathbf{q}]\Phi_2, \Psi_2)\, d\mu\,(\mathbf{p}, \mathbf{q})\right| \leq ||\Phi_1|| \cdot ||\Phi_2|| \cdot ||\Psi_1|| \cdot ||\Psi_2||. \tag{4.7}$$

Proof: We apply von Neumann's result that any such representation is a discrete direct sum of copies of the Schrödinger representation.[11] Thus we can write

$$H = \sum_{r=1}^{L} \oplus H_r,$$

where possibly $L = \infty$, and

$$U[\mathbf{p}, \mathbf{q}] = \sum_{r=1}^{L} \oplus U_r[\mathbf{p}, \mathbf{q}],$$

[10] J. McKenna and J. R. Klauder, J. Math. Phys. 5, 878 (1964).
[11] J. von Neumann, Math. Ann. 104, 570 (1931).

where $U_r[\mathbf{p}, \mathbf{q}]$ is the Schrödinger representation for N degrees of freedom. Let

$$(U[\mathbf{p}, \mathbf{q}]\Phi, \Psi)_r = (U_r[\mathbf{p}, \mathbf{q}]\Phi_r, \Psi_r),$$

where

$$\Phi = \sum_{r=1}^{L} \oplus \Phi_r, \qquad \Psi = \sum_{r=1}^{L} \oplus \Psi_r.$$

Using Schwartz's inequality we get

$$\left| \int (U[\mathbf{p}, \mathbf{q}]\Phi_1, \Psi_1)^*(U[\mathbf{p}, \mathbf{q}]\Phi_2, \Psi_2)\, d\mu\,(\mathbf{p}, \mathbf{q}) \right|$$

$$\leq \sum_{r,s=1}^{L} \left| \int (U[\mathbf{p}, \mathbf{q}]\Phi_1, \Psi_1)_r^*(U[\mathbf{p}, \mathbf{q}]\Phi_2, \Psi_2)_s\, d\mu\,(\mathbf{p}, \mathbf{q}) \right|$$

$$\leq \sum_{r,s} \left\{ \int |(U[\mathbf{p}, \mathbf{q}]\Phi_1, \Psi_1)_r|^2\, d\mu\,(\mathbf{p}, \mathbf{q}) \right\}^{\frac{1}{2}}$$

$$\times \left\{ \int |(U[\mathbf{p}, \mathbf{q}]\Phi_2, \Psi_2)_s|^2\, d\mu\,(\mathbf{p}, \mathbf{q}) \right\}^{\frac{1}{2}}$$

$$= \left(\sum_r \left\{ \int |(U[\mathbf{p}, \mathbf{q}]\Phi_1, \Psi_1)_r|^2\, d\mu\,(\mathbf{p}, \mathbf{q}) \right\}^{\frac{1}{2}} \right)$$

$$\times \left(\sum_s \left\{ \int |(U[\mathbf{p}, \mathbf{q}]\Phi_2, \Psi_2)_s|^2\, d\mu\,(\mathbf{p}, \mathbf{q}) \right\}^{\frac{1}{2}} \right).$$

Using Theorem 4.1 the right-hand side of this inequality becomes

$$\left(\sum_r ||\Phi_1||_r||\Psi_1||_r\right)\left(\sum_s ||\Phi_2||_s||\Psi_2||_s\right)$$

$$\leq \left(\sum_r ||\Phi_1||_r^2\right)^{\frac{1}{2}}\left(\sum_r ||\Psi_1||_r^2\right)^{\frac{1}{2}}$$

$$\times \left(\sum_r ||\Phi_2||_r^2\right)^{\frac{1}{2}}\left(\sum_r ||\Psi_2||_r^2\right)^{\frac{1}{2}}$$

$$= ||\Phi_1|| \cdot ||\Psi_1|| \cdot ||\Phi_2|| \cdot ||\Psi_2||. \qquad \text{Q.E.D.}$$

It follows from Theorem 4.2 that $I_N(\Phi_1, \mathrm{X}_1; \Phi_2, \Psi_2)$ exists and

$$|I_N(\Phi_1, \Psi_1; \Phi_2, \Psi_2)| \leq ||\Phi_1|| \cdot ||\Phi_2|| \cdot ||\Psi_1|| \cdot ||\Psi_2||, \tag{4.8}$$

independent of N. We now proceed to evaluate

$$\lim_{N\to\infty} I_N(\Phi_1, \Psi_1; \Phi_2, \Psi_2)$$

for our direct product representations. [It should be noted that Eq. (4.8) is valid for *any* representations of the CCR, not just the direct-product representation of this paper. However, in general the limit does not exist.] The following discussion is a modification of the methods of CRT V.

We begin by defining some operators on the IDPS $\mathfrak{H}_\mathrm{X}$, where

$$\mathrm{X} = \prod_{n=1}^{\infty} \otimes \chi_n, \quad ||\chi_n|| = 1$$

and $\chi_n \in \mathfrak{H}_n$. Let P_{χ_n} be the projection operator on χ_n, and I_n the identity operator in $\mathfrak{H}_n$. We define a sequence of projection operators A_N, $N = 1, 2, \cdots$ by

$$A_N = \left(\prod_{n=1}^{N} \otimes I_n\right) \otimes \left(\prod_{n=N+1}^{\infty} \otimes P_{\chi_n}\right). \tag{4.9}$$

It follows from Eq. (3.10) that

$$\lim_{N\to\infty} A_n = I \qquad \text{(strong convergence)}. \tag{4.10}$$

For product vectors of the form

$$\Lambda = \prod_{n=1}^{\infty} \otimes \lambda_n, \quad \Lambda \in \mathfrak{H}_\chi,$$

we define a partial isometry B_n by

$$B_N\Lambda = \left\{\prod_{N+1}^{\infty} (\chi_n, \lambda_n)\right\} \prod_{n=1}^{N} \otimes \lambda_n; \tag{4.11}$$

B_N can be extended by linearity and continuity to all of $\mathfrak{H}_\mathrm{X}$. B_N is zero on $(1 - A_N)\mathfrak{H}_\mathrm{X}$ and is a unitary operator from $A_N\mathfrak{H}_\mathrm{X}$ onto the finite tensor product space $\mathfrak{H}_N = \prod_{n=1}^{N} \otimes \mathfrak{H}_n$. We are now prepared to evaluate our integral.

We consider first the case where X_1 and X_2 are strongly equivalent. Then $\mathfrak{H}_{\mathrm{X}_1} = \mathfrak{H}_{\mathrm{X}_2}$. Let A_N be the sequence of operators obtained from X_1. We note that the integral $I_N(\Phi_1, \Psi_1; \Phi_2, \Psi_2)$ is linear in Φ_1 and Ψ_2, and antilinear in Φ_2 and Ψ_1. We write

$$\Phi_1 = A_N\Phi_1 + (1 - A_N)\Phi_1, \tag{4.12}$$

and similarly for Φ_2, Ψ_1, Ψ_2. Then I_N consists of $2^4 = 16$ terms. We consider first the 15 terms in which $(1 - A_N)$ occurs at least once. It follows from Theorem 4.2 that these terms satisfy inequalities of the form

$$|I_N((1 - A_N)\Phi_1, A_N\Psi_1; A_N\Phi_2, A_N\Psi_2)| \leq ||(1 - A_N)\Phi_1||\, ||A_N\Psi_1|| \cdot ||A_N\Phi_2|| \cdot ||A_N\Psi_2||. \tag{4.13}$$

Since $1 - A_N \to 0$ strongly, these terms all vanish in the limit $N \to \infty$. We now consider the term $I_N(A_N\Phi_1, A_N\Psi_1; A_N\Phi_2, A_N\Psi_2)$. Since the representation $U[f_{(N)}, g_{(N)}]$ is irreducible on $A_N\mathfrak{H}_{\mathrm{X}_1}$, it follows from Theorem 4.1 that

$$I_N(A_N\Phi_1, A_N\Psi_1; A_N\Phi_2, A_N\Psi_2) = (A_N\Phi_2, A_N\Phi_1)(A_N\Psi_1, A_N\Psi_2). \tag{4.14}$$

Since $A_N \to 1$ strongly, it follows that

$$\lim_{N\to\infty} I_N(\Phi_1, \Psi_1; \Phi_2\Psi_2) = (\Phi_2, \Phi_1)(\Psi_1, \Psi_2). \tag{4.15}$$

We consider next the case where X_1 and X_2 are weakly equivalent. Let T be the unitary operator of Eq. (3.11) which maps $\mathcal{H}_{X_2}$ onto $\mathcal{H}_{X_1}$, and establishes the unitary equivalence of the representations [see Eq. (3.12)]. Then we have

$$\lim_{N\to\infty} I_N(\Phi_1, \Psi_1; \Phi_2, \Psi_2) = (T\Phi_2, \Phi_1)(\Psi_1, T\Psi_2), \quad (4.16)$$

We remark that this could also be proved directly without using the unitary equivalence.

Finally we consider the case where X_1 and X_2 are weakly inequivalent. Let A_N^i, B_N^i be the operators obtained from X_i, $i = 1, 2$. We write

$$\Phi_i = A_N^i\Phi_i + (1 - A_N^i)\Phi_i, \quad (4.17)$$

and similarly for Ψ_i. As above, the only term which does not obviously vanish in the limit $N \to \infty$ is

$$I_N(A_N^1\Phi_1, A_N^1\Psi_1; A_N^2\Phi_2, A_N^2\Psi_2).$$

We note that B_N^i maps the representation $U_i[f_{(N)}, g_{(N)}]$ on $A_N\mathcal{H}_{X_i}$ onto the Schrödinger representation on the finite tensor product space $\mathcal{H}_N$. It then follows from Theorem 4.1 that

$$I_N(A_N^1\Phi_1, A_N^1\Psi_1; A_N^2\Phi_2, A_N^2\Psi_2) = (B_N^2A_N^2\Phi_2, B_N^1A_N^1\Phi_1)(B_N^1A_N^1\Psi_1, B_N^2A_N^2\Psi_2). \quad (4.18)$$

It follows from Eq. (3.9) that the rhs of this equation vanishes in the limit $N \to \infty$ for product vectors. It then follows by a continuity argument that it vanishes for all vectors, and we have

$$\lim_{N\to\infty} I_N(\Phi_1, \Psi_1; \Phi_2, \Psi_2) = 0. \quad (4.19)$$

In Sec. 3 we noted that direct-product representations of the CCR defined on IDPS in the same weak equivalence class were unitary equivalent. If representations on weakly inequivalent spaces were unitary equivalent, then the integral in Eq. (4.19) could not vanish since it would be given by $(T\Phi_2, \Phi_1)(\Psi_1, T\Psi_2)$, where T is the unitary operator mediating the equivalence. Thus we have obtained a complete classification of the irreducible direct-product representations. We summarize our results as

Theorem 4.3: Let $U_1[f, g]$ and $U_2[f, g]$ be irreducible tensor product representations of the CCR defined on the IDPS $\mathcal{H}_1$ and $\mathcal{H}_2$. If $\mathcal{H}_1$ and $\mathcal{H}_2$ are weakly equivalent then there is a unitary operator T from $\mathcal{H}_2$ to $\mathcal{H}_1$ such that

$$TU_2[f, g]T^{-1} = U_1[f, g], \quad (4.20)$$

and

$$\lim_{N\to\infty} \int (U_1[f_{(N)}, g_{(N)}]\Phi_1, \Psi_1)^*(U_2[f_{(N)}, g_{(N)}] \times \Phi_2, \Psi_2)\, d\mu\,(f_{(N)}, g_{(N)}) = (T\Phi_2, \Phi_1)(\Psi_1, T\Psi_2), \quad (4.21)$$

where

$$d\mu\,(f_{(N)}, g_{(N)}) = (2\pi)^{-N} \prod_{n=1}^{N} dp_n\, dq_n. \quad (4.22)$$

If $\mathcal{H}_1$ and $\mathcal{H}_2$ are not weakly equivalent then the representations are unitary inequivalent and

$$\lim_{N\to\infty} \int (U_1[f_{(N)}, g_{(N)}]\Phi_1, \Psi_1)^* \times (U_2[f_{(N)}, g_{(N)}]\Phi_2, \Psi_2)\, d\mu\,(f_{(N)}, g_{(N)}) = 0 \quad (4.23)$$

for all Φ_1, Ψ_1, Φ_2, and Ψ_2.

Note added in proof. H. Araki has shown that if (for an irreducible representation) the group integral exists and does not depend on the ordering of the degrees of freedom in a fixed basis, then the representation is a direct-product representation. [H. Araki and E. J. Woods (to be published).]

5. "TAG TEST" FOR UNITARY INEQUIVALENCE OF DIRECT-PRODUCT REPRESENTATIONS

As an alternate derivation of the results on inequivalent representations derived in the last section we employ a variant of the "tag test" introduced in CRT V. Let $U_i[f, g]$, $f, g \in \mathcal{S}$, $i = 1, 2$, be two CCR representations defined on the IDPS $\mathcal{H}_{X_i}$ in the manner of Sec. 2. Consider the sequence of bounded operators

$$A_{iM} = \sum_{m=1}^{M} c_m U_i[f^{(m)}, g^{(m)}], \quad (5.1)$$

$M = 1, 2, \cdots$, defined on each space. For simplicity we do not explicitly indicate the M dependence of the sequences $\{c_m\}$, $\{f^{(m)}\}$, and $\{g^{(m)}\}$ in (5.1). If there exists a unitary map, T, of $\mathcal{H}_{X_2}$ onto $\mathcal{H}_{X_1}$ such that for all $f, g \in \mathcal{S}$

$$U_1[f, g] = TU_2[f, g]T^{-1}, \quad (5.2)$$

it then follows for all M that

$$A_{1M} = TA_{2M}T^{-1}. \quad (5.3)$$

If *any* two sequences are related by (5.3) and $\{A_{2M}\}$ converges weakly to a bounded operator B_2, then $\{A_{1M}\}$ converges weakly to $B_1 = TB_2T^{-1}$.

In particular, if $B_2 = I_2$, the identity operator on $\mathfrak{H}_{\mathrm{X}_2}$, then $B_1 = I_1$. To establish our desired inequivalence, we cook up a sequence of operators such that $B_2 = I_2$ while $B_1 = 0$, thus invalidating (5.2). Our choice of sequence is suggested by the observation that since the

$$U_0[p_n, q_n] \equiv V_0[q_n]W_0[p_n]$$

are irreducible in each $\mathfrak{H}_n$, any projection operator in $\mathfrak{H}_n$ can be obtained as a function of the $U[f, g]$. The desired "tags" are then constructed with the aid of projection operators on the components of the product reference vector.

We recall that in the proof of weak convergence, it is sufficient to show convergence on a set of vectors whose finite sums are everywhere dense, i.e., a total set. Such a set is composed of arbitrary product vectors $\Lambda_i, \Omega_i \in \mathfrak{H}_{\mathrm{X}_i}$ that differ from X_i in at most finitely many components. Thus we consider expressions of the form

$$(\Lambda_i, A_{iM}\Omega_i) = \sum_{m=1}^{M} c_m \prod_{n=1}^{\infty} (\lambda_n^i, U_0[p_n^{(m)}, q_n^{(m)}]\omega_n^i).$$

We desire to choose a family of $c_m = c_m^{RS}$, $R \geq S$, where $R, S = 1, 2, \cdots$, such that the associated A_{iM} converges weakly to the projection operator

$$A_i^{RS} = (\prod_{n<S} \otimes I) \otimes (\prod_{S\leq n\leq R} \otimes P_{\chi_n^i}) \otimes (\prod_{n>R} \otimes I).$$

Since we know[10] that

$$\int \cdots \int \prod_{S\leq n\leq R} \{(\chi_n^2, U_0[p_n, q_n]\chi_n^2)^* \times (\lambda_n^i, U_0[p_n, q_n]\omega_n^i) \frac{dp_n\, dq_n}{2\pi}\} = \prod_{S\leq n\leq R} (\lambda_n^i, \chi_n^2)(\chi_n^2, \omega_n^i),$$

we choose our sequences c_m^{RS}, $p_n^{(m)}$, $q_n^{(m)}$ such that the limit $M \to \infty$ corresponds to this integral; for example, we could use the sequence involved in a Riemann definition of the integral. At any rate, we are assured that $A_1^{RS} = TA_2^{RS}T^{-1}$ whenever (5.2) holds.

Now consider the matrix element

$$(\Lambda_i, A_i^{RS}\Omega_i) = \prod_{\substack{n<S \\ n>R}} (\lambda_n^i, \omega_n^i) \prod_{S\leq n\leq R} (\lambda_n^i, \chi_n^2)(\chi_n^2, \omega_n^i), \tag{5.4}$$

Since Ω_i and Λ_i differ from X_i on only finitely many components there exists a finite number L such that

$$\omega_n^i = \lambda_n^i = \chi_n^i, \qquad n \geq L.$$

Hence for $R \geq S \geq L$, Eq. (5.4) becomes

$$(\Lambda_i, A_i^{RS}\Omega_i) = \prod_{n<S} (\lambda_n^i, \omega_n^i) \prod_{S\leq n\leq R} |(\chi_n^i, \chi_n^2)|^2.$$

Now

$$\lim_{R\to\infty} (\Lambda_i, A_i^{RS}\Omega_i) = \prod_{n<S} (\lambda_n^i, \omega_n^i) \cdot F_i^S,$$

where

$$F_2^S = \lim_{R\to\infty} 1 = 1,$$

$$F_1^S = \lim_{R\to\infty} \prod_{S\leq n\leq R} |(\chi_n^1, \chi_n^2)|^2.$$

If X_2 is weakly inequivalent to X_1 it follows from (3.7) that $F_1^S = 0$ for all S. However if X_2 is weakly equivalent to X_1, then (3.8) implies that $\lim_{S\to\infty} F_1^S = 1$. When we combine these facts we see that

$$\lim_{S\to\infty} \lim_{R\to\infty} (\Lambda_i, A_i^{RS}\Omega_i) = b_i(\Lambda_i, \Omega_i),$$

$b_2 = 1$, and $b_1 = 1$ or 0 depending on whether $\mathrm{X}_2 \overset{w}{\approx} \mathrm{X}_1$ or not, respectively. Since this result holds for arbitrary Ω_i and Λ_i in our total set we are assured that Eq. (5.2) cannot hold for any two representations defined on weakly inequivalent spaces. Since the unitary equivalence of representations defined on weakly equivalent spaces was shown earlier, this completes the classification of unitary equivalence for our product representations.

It is clear that a trivial adaptation of the present procedure implies that analogous, irreducible direct-product representations of the canonical *anti-* commutation relations defined on weakly inequivalent IDPS (where each $\mathfrak{H}_n$ is two-dimensional) are unitarily inequivalent.

Conditions for the Existence of Closed Solutions by the Normal Ordering Method*

J. H. MARBURGER
Division of Applied Physics, Stanford University, Stanford, California
(Received 1 November 1965)

It is shown that the normal ordering method of solving boson operator equations leads to closed solutions of the Schrödinger and density operator equations only in certain cases. In particular, a necessary condition for the existence of solutions of the form $\mathfrak{N}\{\exp[G(a_i^+, a_i)]\}$ with G a finite multinomial is that the Hamiltonian be of the general form $H = \sum_{i,j}(c_{ij}a_i^+a_j + \tilde{c}_{ij}a_ia_j + \tilde{c}_{ij}^*a_i^+a_j^+) + \sum_i(c_ia_i + c_i^*a_i^+) + c$. Approximate solutions may be obtained for the density operator for arbitrary Hamiltonians.

1. INTRODUCTION

HEFFNER and Louisell[1] have recently proposed a powerful method of solving some operator equations appearing in quantum mechanics. They have shown that it is possible to transform differential equations involving certain noncommuting operators into partial differential equations involving only commuting quantities. This technique, which we shall call the normal ordering method, is discussed thoroughly in Refs. 1–3 and we do not repeat the details. The purpose of this paper is to point out an inherent limitation in the method and to list all possible cases in which it leads to closed solutions. To be precise, consider any second quantized Hamiltonian $H = H_0 + \lambda H_1$ where $H_0 = \sum_i \omega_i a_i^+ a_i$ and H_1 is a finite multinomial in the boson operators a_i^+ and a_i(i.e., $[a_i, a_j^+] = \delta_{ij}$, $[a_i, a_j] = [a_i^+, a_j^+] = 0$). We want to find that class of Hamiltonians for which the following equations are "form-solvable" by the method of normal ordering:

$$i\partial U/\partial t = HU, \tag{1}$$

$$i\partial\rho/\partial t = H\rho - \rho H. \tag{2}$$

Here U is the time-development operator for solutions of the Schrödinger equation in the Schrödinger picture and ρ is the density operator in the Schrödinger picture. By *form-solvable* we mean that the dependence of U or ρ on the operators a_i^+ and a_i may be written exactly as $\mathfrak{N}[\exp(G)]$, where G is a finite multinomial in a_i^+ and a_i with c-number coefficients, and $\mathfrak{N}$ is the normal ordering operator, defined in Ref. 1, which places all annihilation operators to the right of all creation operators without regard to commutators. It may not be possible to solve for the coefficients exactly. We shall show that the class of Hamiltonians which leads to form-solvable equations [(1) or (2)] is restricted but that there exists for *every* Hamiltonian a systematic approximation procedure capable of generating solutions of (2) valid to any given order of λ.

2. SOLUTION OF THE SCHRÖDINGER AND DENSITY OPERATOR EQUATIONS BY THE NORMAL ORDERING METHOD

Schrödinger's Equation

As in Ref. 1 we solve Eq. (1) in the interaction representation in which

$$U(t, t_0) = \exp(-iH_0t)V(t, t_0)\exp(iH_0t_0),$$

where $V(t, t_0)$ satisfies

$$i\,\partial V(t, t_0)/\partial t = H_1(a_i^+, a_i; t)V(t, t_0), \tag{3}$$

with the boundary condition $V(t, t_0) = 1$. Application of the normal ordering technique replaces (3) with a partial differential equation in the c-number variables $\bar{a}_i^+$ and $\bar{a}_i$:

$$i\,\partial\bar{V}(\bar{a}_i^+, \bar{a}_i; t)/\partial t = \bar{H}_1(\bar{a}_i^+, \bar{a}_i + \partial/\partial\bar{a}_i^+; t)\bar{V}(\bar{a}_i^+, \bar{a}_i; t), \tag{4}$$

where $\bar{V}(\bar{a}_i^+, \bar{a}_i)$ is obtained from $V(a_i^+, a_i)$ by commuting all annihilation operators to the right of all creation operators and replacing a_i^+ with $\bar{a}_i^+$ and a_i with $\bar{a}_i$. Equation (4) is to be solved by assuming the trial solution $\bar{V} = \exp[G(\bar{a}_i^+, \bar{a}_i; t)]$, where G is a multinomial in $\bar{a}_i^+$ and $\bar{a}_i$ such that both sides of the equation for G have equivalent multinomial structure. The coefficients are then determined from the system of ordinary differential equations resulting from comparison of coefficients of like monomials.

In general, it is difficult to guess the appropriate multinomial structure of G by inspection. In fact, it is often the case that no finite multinomial G

* This work was performed while the author held a NASA Traineeship at Stanford University.

[1] H. Heffner and W. H. Louisell, J. Math. Phys. **6**, 474 (1965).

[2] W. H. Louisell and L. R. Walker, Phys. Rev. **147**, B204 (1965).

[3] W. H. Louisell, *Radiation and Noise in Quantum Electronics* (McGraw-Hill Book Company, New York, 1964).

will satisfy (4). To see this, consider the differential equation for G (we choose $t_0 = 0$ for convenience):

$$i\,\partial G/\partial t = A + R_1 G, \qquad G(t = t_0 = 0) = 0. \tag{5}$$

Here A has the same structure in $\bar{a}_i^+$ and $\bar{a}_i$ as H_1 does in a_i^+ and a_i, and R_1 is a differential operator (possibly nonlinear) in $\bar{a}_i^+$ and $\bar{a}_i$. Equation (5) has the formal solution

$$G(t) = -i \exp(-itR_1) \int_0^t \exp(it'R_1) A(t')\, dt',$$

which shows that the multinomial structure of G must include that of $\sum_{n=0}^{\infty} R_1^n A (R_1^0 A \equiv A)$. Thus G has a finite multinomial structure in $\bar{a}_i^+$ and $\bar{a}_i$ if and only if there exists an integer N such that for every $n > N$ the structure of $R_1^n A$ is included in that of $\sum_{m=0}^{N} R_1^m A$. If N is the least integer for which this condition is satisfied we shall say that R_1 is *form-bounded of order N with respect to A*. Consequently, a necessary and sufficient condition that Eq. (1) be form-solvable is that R_1 in (5) be form-bounded with respect to $H_1(\bar{a}_i^+, \bar{a}_i)$.

Density operator equation

Working in the Schrödinger picture, we find for the partial differential equation resulting from (2) by the normal ordering technique

$$i\,\partial\bar{\rho}/\partial t = [\bar{H}(\bar{a}_i^+, \bar{a}_i + \partial/\partial\bar{a}_i^+; t) - \bar{H}(\bar{a}_i^+ + \partial/\partial\bar{a}_i, \bar{a}_i; t)]\bar{\rho}. \tag{6}$$

Substitution of the trial function $\bar{\rho} = \exp(G)$ leads to

$$i\,\partial G/\partial t = (T_0 + \lambda T_1) G, \tag{7}$$

where T_0 arises from H_0 and T_1 from H_1 operating on $\exp(G)$. In Eq. (7), T_0 is an example of a form-bounded operator of first order with respect to any multinomial in $\bar{a}_i^+$ and $\bar{a}_i$. In fact, since $H_0 = \sum_i \omega_i a_i^+ a_i$, then

$$T_0 = \sum_i \omega_i (\bar{a}_i^+\, \partial/\partial\bar{a}_i^+ - \bar{a}_i\, \partial/\partial\bar{a}_i).$$

It is easy to see that when this operates on any multinomial, the result is a multinomial with no new terms of different structure.

Equation (7) has the formal solution

$$G(t) = \exp[-it(T_0 + \lambda T_1)] G(0). \tag{8}$$

Since T_0 is known to be form-bounded, Eq. (8) shows that it is necessary and sufficient that T_1 be form-bounded with respect to $G(0)$ in order that (7) and hence (2) be form-solvable.

3. GENERAL FORM-BOUNDED OPERATORS

We have seen that form boundedness of the operators R_1 and T_1 in Eqs. (5) and (7) is a necessary condition for the existence of closed solutions of Eqs. (1) and (2) by the normal ordering method. In this section we state two theorems which allow us to list all Hamiltonians which lead to form-solvable equations. Both theorems are proved in the appendix.

Theorem I: Let A be the general monomial in M variables:

$$A = x_1^{n_1} x_2^{n_2} \cdots x_M^{n_M}, \tag{9}$$

and T the general linear operator of one term:

$$T = x_1^{s_1} x_2^{s_2} \cdots x_M^{s_M} (\partial/\partial x_1)^{t_1} (\partial/\partial x_2)^{t_2} \cdots (\partial/\partial x_M)^{t_M} \tag{10}$$

Then the following operators or linear combinations of them are the only linear ones form-bounded with respect to A.

(a) T as in (10) with $s_i < t_i$ for some i.
(b) T as in (10) with s_i arbitrary, but $t_j > n_j$ for some j.
(c) T as in (10) with $s_i = t_i$ for all i.

In general, it is not sufficient to consider linear operators. In fact, if the trial solution $\exp(G)$ is substituted in a linear differential equation of order l, then G satisfies an equation which includes the nonlinear term

$$T_N G = x_1^{s_1} x_2^{s_2} \cdots x_M^{s_M} (\partial G/\partial x_1)^{t_1} \times (\partial G/\partial x_2)^{t_2} \cdots (\partial G/\partial x_M)^{t_M}. \tag{11}$$

This is the only nonlinear operator we shall discuss. If T_N in (11) is form-bounded, then in all the problems considered here the other nonlinear operators in the equation for G are also form-bounded. Theorem II lists all operators (11) form-bounded with respect to the monomial (9).

Theorem II: Let A be the monomial (9) and T_N the nonlinear operator in (11) with $\sum_i t_i \geq 2$ (otherwise T_N would be linear). Then T_N is form-bounded with respect to A if and only if it falls in one of the following cases:

(a) $n_i = 0$ and $t_i \neq 0$ for some i.
(b) $n_i = 1$ and $T_N A = (\partial A/\partial x_i)^{t_i}$.
(c) $n_i = 1, n_j = 0$ for $j \neq i$ and $T_N A = x_i (\partial A/\partial x_i)^{t_i}$.
(d) $n_i = 1, n_j = 1, n_k = 0$ for $k \neq i, j$ and $T_N A = (\partial A/\partial x_i)(\partial A/\partial x_j)$.

4. SOLVABLE HAMILTONIANS

The physical requirement that the Hamiltonian be Hermitian, together with Theorems I and II, allow us to write the most general Hamiltonian leading to a form-solvable Schrödinger equation.

To do the same for the density operator equation we require additional information concerning $G(0)$.

Schrödinger's Equation

If H_1 in (3) is such that R_1 in (5) is linear, then H_1 must be linear in a_i. The most general form, which is also Hermitian, is

$$H_1 = \sum_{i,j} c_{ij} a_i^+ a_j + \sum_i (c_i a_i + c_i^* a_i^+) + c,$$

where $c_{ij} = c_{ji}^*$, and c is real. This leads to

$$R_1 = \sum_{i,j} c_{ij} \bar{a}_i^+ \, \partial/\partial \bar{a}_j^+ + \sum_i c_i \, \partial/\partial \bar{a}_i^+,$$

which according to Theorem I is form-bounded with respect to $H_1(\bar{a}_i^+, \bar{a}_i)$. (This is also easily verified by direct computation.)

If R_1 is nonlinear, it is evident that, by construction, case (a) of Theorem II never occurs. Case (b) appears for $t_i > 2$ only in combination with other terms which are not form-bounded. For example, the presence of the term $(\partial G/\partial \bar{a}_i^+)^{t_i}$ in R_1 implies the presence of the term $\bar{a}_i^2 (\partial G/\partial \bar{a}_i^+)^{t_i - 2}$, which, for $t_i > 2$, is not form-bounded with respect to $(\bar{a}_i^+)^{t_i}$, which must appear in H_1. A similar argument eliminates case (c) for $t_i > 1$. Thus only case (d) of Theorem II is allowed, giving terms in H proportional to $a_i a_j$ and $a_i^+ a_j^+$. We therefore arrive at the following conclusion: *A necessary condition for the existence of a closed solution of Eq.* (4) *of the form* $\bar{V} = \exp[G(\bar{a}_i^+, \bar{a}_i; t)]$ *is that the Hamiltonian* H *in* (1) *have the general form*

$$H = \sum_{i,j} (c_{ij} a_i^+ a_j + \bar{c}_{ij} a_i a_j + \bar{c}_{ij}^* a_i^+ a_j^+) + \sum_i (c_i a_i + c_i^* a_i^+) + c, \qquad (12)$$

where $c_{ij} = c_{ji}^*$, and c is real.

Density Operator Equation

To obtain a definite condition for the form solvability of Eq. (7), we require the general form of $G(0)$ and hence the density operator $\rho(0)$. In principle $G(0)$ is arbitrary, but in practice one always assumes that $\rho(0)$ describes an equilibrium situation in which the corresponding Hamiltonian is H_e. By "equilibrium situation" we mean that the entropy $S = -k \operatorname{Tr}\{\rho \log \rho\}$ is a maximum subject to the constraint that the average energy, and possibly other dynamic variables, is fixed. In this case it has been shown[4] that ρ is proportional to $\exp(F)$, where F is linear in the fixed variables. In particular, F is linear in H_e, which always contains the bilinear term $a_i^+ a_i$ for a boson system. By Eq. (13) below it follows that $G(0)$ always includes the term $\bar{a}_i^+ \bar{a}_i$.

The general Hamiltonian, which leads to a linear operator T_1 in Eq. (7), is the same as that obtained for the Schrödinger equation. For nonlinear operators, cases (a) and (c) of Theorem II are excluded by the discussion above. Case (b) is excluded for $t_i > 2$ because it always appears simultaneously with another operator which is never form-bounded. We therefore conclude that *a necessary condition for the existence of a closed solution of Eq.* (6) *of the form* $\bar{\rho} = \exp[G(\bar{a}_i^+, \bar{a}_i; t)]$ *is that the Hamiltonian have the general form* (12).

5. DISCUSSION

Many boson systems of physical importance have Hamiltonians which are not of the form (12). It is relevant to ask what implications our analysis has for the solvability of Eqs. (1) and (2) for these systems. In this connection, it is important to notice that the normal ordering method always yields solutions in normal form. However, there are many finite multinomials $F(a_i^+, a_i)$ such that $\exp(F)$ cannot be written as $\mathfrak{N}\{\exp[G(a_i^+, a_i)]\}$ for finite G. In fact, $\exp(F)$ can be written[3] as $\mathfrak{N}[f(a_i^+, a_i; 1)]$ where $f(\bar{a}_i^+, \bar{a}_i; t)$ is a solution of

$$\partial f/\partial t = F(\bar{a}_i^+, \bar{a}_i + \partial/\partial \bar{a}_i^+) f, \qquad (13)$$

with $f(\bar{a}_i^+, \bar{a}_i; 0) = 1$. Since this is equivalent to Eq. (4), it follows that $\exp[F(a_i^+, a_i)]$ may be written in closed form as $\mathfrak{N}\{\exp[G(a_i^+, a_i)]\}$ only if F has the general form (12). We conclude that even though a problem is not form-solvable by the normal ordering technique, a closed-form solution may still exist.

Another important consideration is the extent to which the choice of $\exp(G)$ as a trial function limits the generality of our results. It is readily verified that any other trial function leads to very complicated nonlinear equations for G. Moreover, the initial conditions become quite awkward for the density matrix equation. In view of these difficulties any nonexponential closed-form solution is to be regarded as highly exceptional.

Even when Eq. (2) is not form-solvable by the normal ordering method, it is possible to find a multinomial structure for $G(t)$ which satisfies (7) to a given finite order of λ. Expanding the formal exponential in (8), and using the form boundedness of T_0, one finds that the form $\sum_{n=0}^{N} T_1^n G(0)$ satisfies Eq. (7) to order λ^N. This approximation is subject to the usual limitations. In particular, it is not valid for times large compared with the characteristic time determined by λ^{-1}. A more fruitful approxi-

[4] E. T. Jaynes in *Statistical Physics: Brandeis Summer Institute Lectures*, 1962 (W. A. Benjamin, Inc., New York, 1963).

mation scheme might be to consider that part of the exact Hamiltonian which is of the form of Eq. (12) as solvable and treat other terms as perturbations by conventional techniques.

Finally, we remark that the general quadratic Hamiltonian (12) can be diagonalized exactly by a canonical transformation to new boson variables. While this transformation is useful for finding the exact energy spectrum, it is generally difficult to extract from it information concerning the decay of the unperturbed modes of the system. The normal ordering technique, however, is quite valuable in such cases.

The author wishes to thank Professor H. Heffner for a stimulating conversation in connection with this paper.

APPENDIX: PROOF OF THEOREMS I AND II

The following proofs ensure that the operators listed as form-bounded in Sec. 3 are the only possible ones.

Proof of Theorem I. Operating on (9) with (10) l times yields

$$T^l A = \mathfrak{N}_l x_1^{l(s_1-t_1)+n_1} x_2^{l(s_2-t_2)+n_2} \cdots x_M^{l(s_M-t_M)+n_M} \times \prod_{i=1}^{M} \theta(n_i - t_i) \times \prod_{i=1}^{M} \theta(s_i + n_i - 2t_i) \cdots \prod_{i=1}^{M} \theta[(l-1)s_i + n_i - lt_i],$$

where $\mathfrak{N}_l$ is a constant and $\theta(x)$ vanishes for $x < 0$ and is unity otherwise. If $T^l A$ is to have monomial structure included in $\sum_{m=0}^{l-1} T^m A$ then one of the following cases must be true:

Case 1. One of the θ-functions vanishes. In this case there exists an l such that $l(s_i - t_i) - s_i + n_i < 0$ for some i. Obviously, such an l always exists for $s_i < t_i$. If $s_i = t_i$ then a θ(the first) vanishes only for $t_i > n_i$. If $s_i < t_i$ the condition is $(l-1)(s_i - t_i) < t_i - n_i$ where the left side is nonnegative. Thus t_i must exceed n_i.

Case 2. No θ vanishes for any l. It is necessary to choose s_i and t_i such that $l(s_i - t_i) + n_i$ is independent of l. This is true only for $s_i = t_i$ for all i. This exhausts all possibilities for linear form-bounded operators. The admissible cases are listed in Theorem I.

Proof of Theorem II. Operating on (9) with (11) l times gives

$$T_N^l A = \mathfrak{N}_l^{(N)} \prod_{i=1}^{M} x_i^{\Delta_i \sum_{m=0}^{l-1} \psi^m + n_i \psi^l} \times \prod_{i=1}^{M} [\theta(n_i - 1)]^{t_i} \cdots \prod_{i=1}^{M} \left[\theta\left(\Delta_i \sum_{m=0}^{l-2} \psi^m + n_i \psi^{l-1} - 1 \right) \right]^{t_i},$$

where $\Delta_i \equiv s_i - t_i$, $\psi \equiv \sum_{i=1}^{M} t_i$, and $\mathfrak{N}^{(N)}$ is a constant. For form boundedness, we have the same two cases as in Theorem I:

Case 1. One of the θ functions vanishes. We require $\Delta_i R + n_i \psi^{l-1} - 1 < 0$ and $t_i \neq 0$ for some i, where $R \equiv \sum_{m=0}^{l-2} \psi^m = (\psi^{l-1} - 1)/(\psi - 1)$. First suppose that $\Delta_i < 0$ for some i. If $n_i = 0$ our condition is $\Delta_i R < 1$ which is always satisfied. If $n_i = 1$ we require $t_i > \psi + s_i - 1$ which is satisfied only for $s_i = 0$ and $\psi = t_i$ leading to case (b) of the theorem. For $n_i > 1$ we require

$$(\psi^{l-1} - 1)\,|\Delta_i| > (n_i \psi^{l-1} - 1)(\psi - 1)$$

which is never satisfied. (This is evident for $l = 1$.) For $l > 1$ our condition is

$$|\Delta_i| > [(n_i \psi^{l-1} - 1)/(\psi^{l-1} - 1)](\psi - 1),$$

where the coefficient of $\psi - 1$ is not less than 2. Thus $t_i > 2(\sum_m t_m - 1) + s_i \geq 2(t_i - 1)$, which implies $t_i < 2$ and therefore $2 > 2(\psi - 1)$ or $\psi < 2$ (which contradicts the hypothesis of the theorem).

Now suppose $\Delta_i > 0$ for some i. The requirement that a θ function vanish for some l and i becomes $\Delta_i(\psi^{l-1} - 1) < (1 - n_i \psi^{l-1})(\psi - 1)$, which is always satisfied for $n_i = 0$ and $l = 1$. For $n_i > 0$, the case $l = 1$ yields $0 < (1 - n_i)(\psi - 1)$, which is never true. For $l > 1$ we have $\Delta_i < [(1 - n_i \psi^{l-1})/(\psi^{l-1} - 1)](\psi - 1)$ which is never true, the bracketed expression being negative. Finally consider $\Delta_i = 0$. Our condition reduces to $n_i \psi^{l-1} < 1$ which is true only for $n_i = 0$.

Case 2. No θ function vanishes. Here it is necessary that $\Delta_i \sum_{m=0}^{l-1} \psi^m + n_i \psi^l$ be independent of l for all i. This requirement is satisfied if and only if the condition $-\Delta_i = n_i(\psi - 1)$ is true for all i. Summing over i gives $\sum_{i=0}^{M} n_i = \psi/(\psi - 1) - \sum_{i=0}^{M} s_i/(\psi - 1)$. Thus $\sum n_i \leq 2$. The case $\sum n_i = 0$ is excluded by the hypothesis that no θ vanish. If $\sum n_i = 1$ our condition reduces to $t_i - s_i = \psi - 1$ which leads to case (c) of the theorem. If $\sum n_i = 2$ we require $\psi = 2 - \sum s_i$ which in turn requires $\sum s_i = 0$ leading to case (d) of the theorem. This exhausts all possibilities.

Note added in proof. The formal solution immediately following Eq. (5) is correct only if R_1 is time-independent, but the condition that R_1 be form-bounded with respect to A can be shown in general.

The Initial-Value Problem for a Slab. I. Anisotropic Scattering

JANUSZ MIKA
Institute of Nuclear Research, Swierk k/ Otwocka, Poland
(Received 22 April 1965; final manuscript received 5 October 1965)

In this paper, the spectral properties of the Boltzmann operator describing the transport of monoenergetic neutrons with anisotropic scattering in a slab are considered. The Hille–Yosida theorem is applied to obtain the semi-group of operators solving the initial-value problem.

1. INTRODUCTION

IN the last few years one can observe the growing interest in the pulse neutron experiments in reactor physics. The theoretical interpretation of such experiments has been for long time based upon the diffusion approximation, in which the time-dependent neutron flux is expanded into a complete set of eigenfunctions of the diffusion equation. Lehner and Wing[1,2] first showed that for the transport equation in a slab the complete set of discrete eigenfunctions does not exist and the solution of the initial-value problem involves also the continuous spectrum. An important generalization to multigroup transport equations in a slab has been given by Pimbley.[3]

This paper is an expansion of the Lehner and Wing approach to the case of anisotropic scattering. The initial-value problem for a slab is considered. The spectrum of the Boltzmann operator is found and the Hille–Yosida theorem is applied for finding the semigroup of solution operators.

The presented analysis may be used in practical computations of the decay constants of neutron pulses in a slab. The most suitable way of doing this seems to be the normal mode expansion approach employed for isotropic scattering by Bowden and Williams.[4] The normal mode expansion method for anisotropic scattering has been developed by the author.[5]

2. THE FORMULATION OF THE INITIAL-VALUE PROBLEM

Consider an infinite slab of thickness $2a$ surrounded by a vacuum. The cross section σ is constant. At time $t = 0$ a neutron distribution $f(x, \mu)$ exists inside the slab.

The Boltzmann equation and the boundary conditions for the time-dependent neutron distribution $n(x, \mu, t)$ has the following form:

$$\frac{1}{v}\frac{\partial n}{\partial t} + \mu\frac{\partial n}{\partial x} + \sigma n(x, \mu, t) = \frac{c\sigma}{2}\int_{-1}^{+1} d\mu'\, g(\mu, \mu')n(x, \mu', t),$$

$$n(\pm a, \mu, t) = 0; \qquad \mu \lessgtr 0, \qquad t > 0,$$

$$n(x, \mu, 0) = f(x, \mu);$$

$$-a \leq x \leq a; \qquad -1 \leq \mu \leq 1. \tag{2.1}$$

The scattering function $g(\mu, \mu')$ will be considered in the following as a finite sum of Legendre polynomials:

$$g(\mu, \mu') = \sum_{k=0}^{N} b_k P_k(\mu)P_k(\mu'), \tag{2.2}$$

where b_k are the numerical coefficients with $b_0 = 1$.

In Eq. (2.1) the standard notation is used (see, e.g. Ref. 5).

It is convenient to write

$$n(x, \mu, t) = e^{\sigma v t}\psi(x, \mu, t). \tag{2.3}$$

Then choosing, for convenience, $\sigma = v = 1$ we have from Eqs. (2.1)

$$\partial\psi/\partial t = A\psi, \qquad \psi(\pm a, \mu, t) = 0; \quad \mu \lessgtr 0; \quad t > 0, \tag{2.4}$$

$$\psi(x, \mu, 0) = f(x, \mu); \quad -a \leq x \leq a; \quad -1 \leq \mu \leq 1,$$

where the operator A has the form

$$A = -\mu\frac{\partial}{\partial x}\cdot + \frac{c}{2}\sum_{k=0}^{N} b_k P_k(\mu)\int_{-1}^{+1} d\mu' P_k(\mu')\cdot\,. \tag{2.5}$$

The existence of the semigroup of the solution operators for Eq. (2.4) is related to the spectrum of the operator A. Therefore we will investigate the spectrum of A, first defining the proper space of functions to work with.

[1] J. Lehner and G. M. Wing, Commun. Pure Appl. Math. 8, 217 (1955).
[2] J. Lehner and G. M. Wing, Duke Math. J. 23, 125 (1956).
[3] G. H. Pimbley, J. Math. Mech. 8, 837 (1959).
[4] R. L. Bowden and C. D. Williams, J. Math. Phys. 5, 1527 (1964).
[5] J. Mika, Nucl. Sci. Eng. 11, 415 (1961).

3. THE SPACE OF FUNCTIONS

Let $\mathfrak{K}$ be a space of square-integrable functions in rectangle $|\mu| \leq 1$, $|x| \leq a$ with the scalar product

$$(f, g) = \int_{-a}^{+a} dx \int_{-1}^{+1} d\mu f(x, \mu)\bar{g}(x, \mu) \tag{3.1}$$

and the norm

$$||f|| = (f, f)^{\frac{1}{2}}. \tag{3.2}$$

Here $\bar{g}(x, \mu)$ denotes the complex conjugate.

The integral part of the operator A is a projection of $\mathfrak{K}$ to a subspace defined as follows:

Let L_2 be a space of square-integrable functions in $(-a, a)$ with the scalar product

$$(f, g) = \int_{-a}^{+a} dx f(x)\bar{g}(x). \tag{3.3}$$

Then define L as a product space $L = L_2 \times L_2 \times \cdots \times L_2$ (taken N times) with the scalar product

$$(f, g) = \sum_{k=0}^{N} \int_{-a}^{a} dx f_i(x)\bar{g}_i(x). \tag{3.4}$$

The norm in L will be also denoted $||f||$. It shall not lead to any confusion.

Operator A on $\mathfrak{K}$ has as the domain $D(A)$ all functions absolutely continuous in x for each fixed $\mu(|\mu| \leq 1)$ which satisfy the homogeneous boundary conditions stated in Eq. (2.4).

4. THE HALF PLANE Re $\lambda \leq 0$

We will show now that the entire half plane Re $\lambda \leq 0$ belongs to the continuous spectrum of A. In other words, there exists a set of functions $\psi_\delta \in D(A)$ such that

$$||\psi_\delta|| \geq \text{const} > 0 \tag{4.1}$$

and

$$\lim_{\delta \to 0} ||(\lambda - A)\psi_\delta|| = 0. \tag{4.2}$$

We use the same functions as Lehner and Wing[2]:

$$\psi_\delta(x, \mu) = b_\delta(\mu) \frac{x + a}{a} e^{-\lambda(x-a)/\mu} \tag{4.3}$$

where

$$b_\delta(\mu) = \begin{cases} 1/\delta & \delta^2 \leq \mu \leq \delta, \\ 0 & \text{otherwise.} \end{cases}$$

Here $0 < \delta < \frac{1}{2}$, Re $\lambda = \beta < 0$.

It is quite easy to show that both Eqs. (4.1) and (4.2) are satisfied by the functions $\psi_\delta(x, \mu)$ defined by Eq. (4.3).

Now the spectrum of A is a closed set and therefore we may state

Theorem 1: The half plane Re $\lambda \leq 0$ belongs to the spectrum of A.

5. THE SYSTEM OF INTEGRAL EQUATIONS

To investigate the right half of the spectral plane one has to use the system of integral equations instead of the integro-differential equation

$$(\lambda - A)\psi = 0 \tag{5.1}$$

with the boundary conditions stated in Eq. (2.4).

Let us solve Eq. (5.1) for $\psi(x, \mu)$ assuming the integral part to be given. We get

$$\begin{aligned} &\psi(x, \mu) \\ &= \frac{c}{2\mu} \int_{-a}^{x} e^{-\lambda(x-x')/\mu} \left[\sum_{k=0}^{N} b_k P_k(\mu)\phi_k(x') \right] dx', \quad \mu > 0; \\ &= \frac{c}{2\mu} \int_{a}^{x} e^{-\lambda(x-x')/\mu} \left[\sum_{k=0}^{N} b_k P_k(\mu)\phi_k(x') \right] dx', \quad \mu < 0. \end{aligned} \tag{5.2}$$

The moments $\phi_k(x)$ are defined as

$$\phi_k(x) = \int_{-1}^{+1} d\mu' P_k(\mu')\psi(x, \mu'). \tag{5.3}$$

To get a system of equations satisfied by the moments $\phi_k(x)$ one has to multiply both sides of Eq. (5.2) by $P_n(\mu)$ and then integrate over all μ. The change of the order of integration requires care (see Ref. [2]).

As the result one gets for Re $\lambda \geq 0$, $\lambda \neq 0$

$$\phi_n(x) = \frac{c}{2} \sum_{k=0}^{N} \int_{-1}^{+1} B_{nk}(\lambda; x - x')\phi_k(x')\, dx', \qquad n = 0, 1, \cdots N, \tag{5.4}$$

with

$$\begin{aligned} B_{nk}(\lambda; x) &= b_k[\text{sgn } (x)]^{n+k} \int_0^1 \frac{du}{u} P_n(u)P_k(u)e^{-\lambda |x|/\mu} \\ &= b_k[\text{sgn } (x)]^{n+k} \int_1^\infty \frac{dt}{t} P_n\left(\frac{1}{t}\right)P_k\left(\frac{1}{t}\right)e^{-\lambda |x| t}. \end{aligned} \tag{5.5}$$

The integrals in Eqs. (5.5) are easily recognized as linear combinations of the $E_n(x)$ functions as defined, for instance, by Case, Hoffmann, and Placzek.[6]

Equations (5.4) may be written in a vector form

$$\phi = \tfrac{1}{2}cB_\lambda\phi$$

[6] K. M. Case, F. Hoffmann, G. Placzek, *Introduction to the Theory of Neutron Diffusion* (Los Alamos Scientific Laboratory, Los Alamos, New Mexico, 1953), p. 153.

where the vector ϕ has components $\phi_k(x)$, and the matrix B_λ the elements $B_{nk}(\lambda; x - x')$.

The presented derivation shows that every solution of Eq. (5.6) yields the corresponding vector ϕ belonging to L and satisfying Eq. (5.6). By performing the derivation in the opposite direction one can check that every vector ϕ satisfying Eq. (5.6) yields the solution $\psi(x, \mu)$ of Eq. (5.1).

Thus we may state

Theorem 2: A necessary and sufficient condition that $\lambda \in P\sigma A$ on $\mathfrak{X}$ with $\operatorname{Re}\lambda \geq 0$, $\lambda \neq 0$ is that $2/c$ belongs to $P\sigma B_\lambda$ on L. The eigenelements $\psi(x, \mu)$ corresponding to λ and ϕ corresponding to $2/c$ are related to each other by Eqs. (5.2) and (5.3).

Here we use the notation $P\sigma A$, $R\sigma A$, and $C\sigma A$ for the point, residual and continuous spectrum and $\mathfrak{M}A$ for the resolvent set of the operator A.

6. THE RESOLVENT SET

Denote by Γ a set of points λ such that $\operatorname{Re}\lambda > 0$ and $\lambda \notin P\sigma A$. The points of Γ belong to $C\sigma A$, $R\sigma A$ or $\mathfrak{M}A$. Then $R_\lambda = (\lambda - A)^{-1}$ always exists. We will prove that for $\lambda \in \Gamma$, R_λ is bounded and $D(R_\lambda) = \mathfrak{X}$ or, in other words, $\Gamma \in \mathfrak{M}A$.

Consider the equation

$$(\lambda - A)u = g; \qquad g \in \mathfrak{X}. \tag{6.1}$$

It has a formal solution which can be obtained in a way similar to that used in the last section,

$$\begin{aligned} R_\lambda g &= u(x, \mu) \\ &= \frac{1}{\mu}\int_{-a}^{x} e^{-\lambda(x-x')/\mu} \\ &\quad \times \left[\frac{c}{2}\sum_{k=0}^{N} b_k P_k(\mu)\xi_k(x') + g(x', \mu)\right] dx', \quad \mu > 0, \\ &= \frac{1}{\mu}\int_{a}^{x} e^{-\lambda(x-x')/\mu} \\ &\quad \times \left[\frac{c}{2}\sum_{k=0}^{N} b_k P_k(\mu)\xi_k(x') + g(x', \mu)\right] dx', \quad \mu < 0, \end{aligned} \tag{6.2}$$

where

$$\xi_k(x) = \int_{-1}^{1} P_k(\mu')u(x, \mu')\, d\mu'. \tag{6.3}$$

Multiplying by the consecutive $P_n(\mu)$ and integrating over μ we get a system of equations

$$\xi = \tfrac{1}{2}cB_\lambda\xi + G. \tag{6.4}$$

The matrix B_λ has been defined in the last section, ξ is the vector with the components $\xi_k(x)$ and G is the vector with the components $G_k(x)$ given by

$$\begin{aligned} G_k(x) &= \int_0^1 \frac{d\mu}{\mu} P_k(\mu)\int_{-a}^{x} e^{-\lambda(x-x')/\mu} g(x', \mu)\, dx' \\ &\quad - \int_{-1}^{0} \frac{d\mu}{\mu} P_k(\mu)\int_{x}^{a} e^{-\lambda(x-x')/\mu} g(x', \mu)\, dx' \\ &= \int_0^1 H_k^{(1)}(x, \mu)\, d\mu + \int_{-1}^{0} H_k^{(2)}(x, \mu)\, d\mu \\ &= G_k^{(1)}(x) + G_k^{(2)}(x). \end{aligned} \tag{6.5}$$

We will show now that if $g \in \mathfrak{X}$ then $G \in L$. Denoting $\beta = \operatorname{Re}\lambda$ we have

$$\begin{aligned} |H_k^{(1)}(x, \mu)| &\leq \frac{|P_k(\mu)|}{\mu}\int_0^{x+a} e^{-\beta t/\mu}\, |g(x - t, \mu)|\, dt \\ &\leq \int_0^{\infty} \frac{e^{-\beta t/2\mu}}{\mu^{\frac{1}{2}}}\, |g(x-t, \mu)|\, \frac{e^{-\beta t/2\mu}}{\mu^{\frac{1}{2}}}\, dt, \end{aligned} \tag{6.6}$$

where we define $g(x, \mu) = 0$ for $|x| > a$. Applying the Schwartz inequality to Eq. (6.6) gives

$$|H_k^{(1)}(x, \mu)|^2 \leq \frac{1}{\beta\mu}\int_0^{\infty} e^{-\beta t/\mu}\, |g(x - t, \mu)|^2\, dt. \tag{6.7}$$

Now applying again the same inequality we get

$$\begin{aligned} |G_k^{(1)}(x)|^2 &\leq \int_0^1 |H_k^{(1)}(x, \mu)|^2\, d\mu \\ &\leq \frac{1}{\beta}\int_0^1 \frac{d\mu}{\mu}\int_0^{\infty} e^{-\beta t/\mu}\, |g(x-t, \mu)|^2\, dt. \end{aligned} \tag{6.8}$$

Finally,

$$\begin{aligned} &\int_{-a}^{a} |G_k^{(1)}(x)|^2\, dx \\ &\quad \leq \frac{1}{\beta}\int_0^1 \frac{d\mu}{\mu}\int_0^{\infty} e^{-\beta t/\mu}\, dt \int_{-\infty}^{\infty} |g(x - t, \mu)|^2\, dx \\ &\quad = \frac{1}{\beta^2}\int_{-a}^{a} dx \int_0^1 d\mu\, |g(x, \mu)|^2. \end{aligned} \tag{6.9}$$

Hence

$$||G_k^{(1)}|| \leq \frac{1}{\beta}\, ||g||. \tag{6.10}$$

The same estimate holds for $G_k^{(2)}(x)$. It follows then that $G \in L$.

Let us now return to Eq. (6.4) where $G \in L$ and B_λ is a compact operator. The latter follows from the fact that $||B_\lambda|| < \infty$ for $\operatorname{Re}\lambda > 0$, $\lambda \neq 0$. (See the next section.) For $\lambda \in \Gamma$, $2/c$ is not an eigenvalue of B_λ. By an alternative principle, ξ is uniquely determined for a given $g \in \mathfrak{X}$. Since $1 - c/2B_\lambda$ is bounded and one-to-one (it maps L onto itself), $(1 - c/2B_\lambda)^{-1}$ is also bounded. Therefore

$$||\xi|| \leq \text{const}\, ||G|| \leq \frac{\text{const}}{\beta}\, ||g||. \tag{6.11}$$

Now put

$$h(x, \mu) = \frac{c}{2} \sum_{k=0}^{N} b_k P_k(\mu)\xi_k(x) + g(x, \mu). \quad (6.12)$$

Then

$$||h|| \leq \tfrac{1}{2}c \max b_k \, ||\xi|| + ||g|| \leq \text{const} \, ||g||, \quad (6.13)$$

which shows that $h(x, \mu) \in \mathfrak{C}$.

Now in exactly the way used above one can show that

$$||u|| \leq \text{const} \, ||h|| \leq \text{const} \, ||g||. \quad (6.14)$$

From this we see that $u \in \mathfrak{C}$ and $u = (\lambda - A)^{-1}g$ is bounded. Therefore $R_\lambda = (\lambda - A)^{-1}$ is bounded and its domain is $\mathfrak{C}$.

Thus we have proved

Theorem 3: The open half plane $\text{Re}\,\lambda > 0$ deleting any points $\lambda \in P\sigma A$, is contained in the resolvent set $\mathfrak{R}A$.

7. THE POINT SPECTRUM

The remaining question is the nature of that part of the point spectrum which is contained in the half space $\text{Re}\,\lambda \geq 0$. We will denote it by $P\sigma_r A$.

By Theorem 2 we shall work with the integral operator B_λ defined in Sec. 5.

Theorem 4: $P\sigma_r A$ is bounded in the complex λ plane.

Proof. If $\lambda \in P\sigma_r A$, then $(1 - c/2B_\lambda)^{-1}$ does not exist.

Now

$$||B_\lambda|| \leq \sum_{k,n} ||B_{nk}|| \quad (7.1)$$

from the definition of the norm in L. Further,

$$||B_{nk}|| = \int_{-a}^{a} dx \int_{-a}^{a} dx' \, |B_{nk}(\lambda; x - x')|^2$$

$$\leq 2a \int_{-\infty}^{\infty} |B_{nk}(\lambda; x)|^2 \, dx. \quad (7.2)$$

To get the estimate for B_{nk} we write

$$\int_1^\infty \frac{dt}{t} P_n\left(\frac{1}{t}\right) P_k\left(\frac{1}{t}\right) e^{-\lambda |x| t}$$

$$= \frac{e^{-\lambda |x|}}{\lambda |x|} + \frac{1}{\lambda |x|} \int_1^\infty dt \, e^{-\lambda |x| t} \frac{d}{dt}\left[\frac{1}{t} P_n\left(\frac{1}{t}\right) P_k\left(\frac{1}{t}\right)\right]. \quad (7.3)$$

From this it follows that

$$|B_{nk}(\lambda; x)| \leq M/|\lambda x|. \quad (7.4)$$

We use this estimate for $|\lambda x| > 1$. For $|\lambda x| \leq 1$ we have

$$|B_{nk}(\lambda; x)|^2 \leq M_1(\log^2 |\lambda x| + \tfrac{1}{4}\pi^2) + M_2. \quad (7.5)$$

Equation (7.5) follows from the power expansion of E_n [see Eq. (7.10)]. M, M_1, M_2, are some positive constants.

Using both Eqs. (7.4) and (7.5) we get from Eq. (7.2)

$$||B_{nk}|| \leq \text{const}/|\lambda|. \quad (7.6)$$

Thus we see that for $\text{Re}\,\lambda \geq 0$ and $|\lambda|$ sufficiently large the Neumann series converges and $(1 - \frac{1}{2}cB_\lambda)^{-1}$ does exist. This proves the theorem.

We proceed further with investigating $P\sigma_r A$.

Lemma 1: Suppose that $P\sigma_r A$ has an accumulation point λ^*. Let $\{\lambda_p\}$ be a sequence such that $\lambda_p \to \lambda^*$. The nth components of the corresponding sequence of normalized eigen-elements $\{\phi^{(p)}\}$ form an equi-continuous set.

Proof.

$$|\phi_n^{(p)}(x + h) - \phi_n^{(p)}(x)|^2$$

$$= \frac{c^2}{4} \left| \sum_{k=0}^{N} \int_{-a}^{a} [B_{nk}(\lambda_p; x + h - x') - B_{nk}(\lambda_p; x - x')]\phi_k^{(p)}(x') \, dx' \right|^2$$

$$\leq \frac{c^2}{4} \left\{ \sum_{k=0}^{N} \int_{-a}^{a} |B_{nk}(\lambda_p; x + h - x') - B_{nk}(\lambda_p; x - x')|^2 \, dx' \right\} \left\{ \sum_{k=0}^{N} \int_{-a}^{a} |\phi_k^{(p)}(x')|^2 \, dx' \right\}$$

$$\leq \frac{c^2}{4} \sum_{k=0}^{N} \int_{-2a}^{2a} |B_{nk}(\lambda_p; u + h) - B_{nk}(\lambda_p; u)|^2 \, du. \quad (7.7)$$

As it was mentioned, each of the functions $B_{nk}(\lambda; x)$ can be written as a linear combination of functions $E_n(\lambda|x|)$. Namely

$$B_{nk}(\lambda; x) = b_k(\text{sgn}\, x)^{n+k} \sum_{\alpha=0}^{n} \sum_{\beta=0}^{k} a_{n\alpha} a_{k\beta} E_{\alpha+\beta+1}(\lambda\,|x|), \quad (7.8)$$

where $a_{n\alpha}$ are the coefficients of successive powers of Legendre polynomials:

$$P_k(\mu) = \sum_{\alpha=0}^{k} a_{k\alpha} \mu^\alpha. \quad (7.9)$$

The exponential integrals $E_n(z)$ can be expanded into power series according to the formula

$$E_n(z) = \int_1^\infty \frac{dt}{t^n} e^{-zt}$$
$$= \sum_{m=0\neq n-1}^{\infty} \frac{(-z)^m}{m!\,(n-1-m)} + (-1)^n \frac{z^{n-1}}{(n-1)!} \times (\log z - A_n + \gamma),$$
$$\gamma = 0.577216 \cdots ; \quad A_1 = 0, \quad A_n = \sum_{m=1}^{n-1} \frac{1}{m}. \tag{7.10}$$

From Eq. (7.10) it is seen that each of the functions B_{nk} can be written in the form

$$B_{nk}(\lambda; x) = g_{nk}(\lambda; x) - a_{n0}a_{k0} \log \lambda_p \tag{7.11}$$

where $g_{nk}(\lambda; x)$ are well-behaved for $\lambda = 0$ and a_{n0} are defined by Eq. (7.9). From Eq. (7.11) we see that the terms with $\log \lambda_p$ cancel in Eqs. (7.7) and we have to consider the expressions

$$|g_{nk}(\lambda_p; u + h) - g_{nk}(\lambda_p; u)|. \tag{7.12}$$

The functions $g_{nk}(\lambda; x)$ are for $\mathrm{Re}\ \lambda \geq 0$ square-integrable in any finite interval of x, e.g., $(-2a, 2a)$. This follows from the fact that the series in $g_{nk}(\lambda_p; x)$ allow for the estimates $\exp(Mx^2)$ and the logarithmic terms are obviously square-integrable.

Therefore one may use the result of Titchmarsh[7] that, for any function, square-integrable in the interval containing the interval (a, b),

$$\lim_{h\to 0} \int_a^b |f(x + h) - f(x)|^2\, dx = 0. \tag{7.13}$$

From this and Dini's theorem it follows that the last term in Eq. (7.7) converges uniformly to zero independently of x and p, $p = 1, 2 \cdots$, $|x| \leq a$. This proves the assertion.

Lemma 2: Under the same assumptions the components of eigen-elements are uniformly bounded.

Proof.

Let $\lambda^* \neq 0$.

$$|\phi_n^{(p)}(x)|^2 \leq \sum_{k=0}^{N} \int_{-a}^{a} |B_{nk}(\lambda_p; x - x')|^2\, dx' \leq \text{const}/|\lambda_p|^2. \tag{7.14}$$

Equation (7.14) follows immediately from Eq. (7.6).

For $\lambda^* = 0$ we shall proceed in a different way. Using Eq. (7.11) we may write

$$\phi_n^{(p)}(x) = \frac{c}{2} \sum_{k=0}^{N} \int_{-a}^{a} g_{nk}(\lambda_p; x - x')\phi_k^{(p)}(x')\, dx' - a_{n0} \cdot \frac{c}{2} \log \lambda_p \cdot \sum_{k=0}^{N} a_{k0} \cdot \int_{-a}^{a} \phi_k^{(p)}(x')\, dx'. \tag{7.15}$$

[7] E. C. Titchmarsh, *The Theory of Functions*, (Oxford University Press, London, 1939).

From Eq. (7.15) and the properties of g_{nk} it follows that

$$\left|1 + ac \log \lambda_p \sum_{n=0}^{N} a_{n0}^2\right| \left|\sum_{k=0}^{N} a_{k0} \int_{-a}^{a} \phi_k^{(p)}(x)\, dx\right| \leq \text{const} \tag{7.16}$$

for $|\lambda| < M$ where M is a certain positive number.

Using again Eq. (7.15) we have

$$|\phi_n^{(p)}(x)| \leq M_1 + M_2\, |\log \lambda_p| \cdot \left|\sum_{k=0}^{N} a_{k0} \int_{-a}^{a} \phi_k^{(p)}(x')\, dx'\right| \tag{7.17}$$

where M_1 and M_2 are some constants.

From Eqs. (7.16) and (7.17) we conclude finally that

$$|\phi_n^{(p)}(x)| \leq \text{const} \tag{7.18}$$

for any λ such that $\mathrm{Re}\ \lambda \geq 0$ and $|\lambda| < M$. This proves the lemma.

Lemmas 1 and 2 are the necessary and sufficient conditions for the Ascoli's theorem to be true. Thus we may state

Lemma 3: If λ^* is an accumulation point of $P\sigma_r A$, then there exists a sequence $\lambda_p \to \lambda^*$ for which the normalized eigenelements $\{\phi^{(p)}\}$, $\phi^{(p)} = c/2B_{\lambda_p}\phi^{(p)}$, possess the uniformly convergent continuous components. The normalized limit element ϕ^* has also continuous components.

Now we state

Theorem 5: The point $\lambda = 0$ is not an accumulation point of $P\sigma_r A$.

Proof.

From Eq. (5.2) we can write

$$\psi^{(p)}(x, \mu) = \frac{c}{2\mu} \int_{-a}^{x} dx'\, e^{-\lambda_p(x-x')/\mu} \times \left[\sum_{k=0}^{N} b_k P_k(\mu)\phi_k^{(p)}(x')\right], \quad \mu > 0,$$
$$= \frac{c}{2\mu} \int_{a}^{x} dx'\, e^{-\lambda_p(x-x')/\mu} \times \left[\sum_{k=0}^{N} b_k P_k(\mu)\phi_k^{(p)}(x')\right], \quad \mu < 0. \tag{7.19}$$

By the standard argument this $\psi^{(p)}(x, \mu)$ converges uniformly to the continuous limit function $\psi^*(x, \mu)$ as $p \to \infty$. Therefore if $\lambda_p \to 0$

$$\psi^*(x, \mu) = \frac{c}{2\mu} \int_{-a}^{x} dx' \left[\sum_{k=0}^{N} b_k P_k(\mu)\phi_k^*(x')\right], \quad \mu > 0,$$
$$= \frac{c}{2\mu} \int_{a}^{x} dx' \left[\sum_{k=0}^{N} b_k P_k(\mu)\phi_k^*(x')\right], \quad \mu < 0. \tag{7.20}$$

Since $||\phi^*|| = 1$ it is seen that the integrand in right-hand side of Eq. (7.20) is not identically equal to zero and it follows that $\psi^*(x,\mu)$ is not integrable with respect to μ for some given x, $|x| \leq a$.

On the other hand, for each λ_p the corresponding $\psi^{(p)}(x, \mu)$ belong to the Hilbert space and are integrable with respect to in the interval $(-1, +1)$. Then by Fatou's Lemma the limit function $\psi^*(x, \mu)$ is also summable—which contradicts the previous statement. This shows that the assumption $\lambda_p \to 0$ leads to the contradiction and proves the theorem.

From the continuity of the operator B_λ and on the basis of Lemma 3 and Theorem 5 we conclude by

Theorem 6: $P\sigma_r A$ is closed. If λ^* is an accumulation point of $P\sigma_r A$ and if $\{\lambda_p\}$ is a sequence of points such that $\lambda_p \in P\sigma_r A$ and $\lambda_p \to \lambda^*$, and if $\phi^{(p)}$ are the corresponding eigenelements, then $\phi^* = \lim_{p\to\infty}\phi^{(p)}$ is an eigen-element corresponding to the eigenvalue λ^*.

It seems that no further information may be obtained about the spectrum of the operator A, at least, by the methods developed so far in the subject. It may be noted that in the case of the multigroup approximation Pimbley[3] gets more detailed information by restricting himself to a class of symmetrizable matrices. In our case it seems that the matrix operator B_λ cannot be put into a symmetric form by any transformation.

On the other hand it is highly plausible that the discrete eigenvalues all lay on the real axis and are finite in number exactly as for the isotropic scattering. This calls, however, for further investigations.

8. SOLUTION OF THE INITIAL-VALUE PROBLEM

The final problem is to find the solution to the time-dependent transport equation with the initial condition stated in Eq. (2.4):

$$\frac{\partial\psi}{\partial t} = A\psi, \quad \psi(x, \mu, \mu) = f(x, \mu), \quad f \in D(A). \tag{8.1}$$

The domain of A includes all functions absolutely continuous in x for each fixed $|\mu| \leq$ and satisfying the homogeneous boundary conditions. They are dense in the Hilbert space since they include all continuously differentiable functions vanishing outside a compact subset of the region $|x| \leq a$, $|\mu| \leq 1$. Therefore $D(A)$ is dense in $\mathfrak{H}$.

In Sec. 6 we have proved that, for $\lambda \in \Gamma$, $R_\lambda = (\lambda - A)^{-1}$ is bounded and its domain is $\mathfrak{H}$. Then A is also closed.

To get the estimate of R_λ put

$$A = -D + c\zeta; \qquad D = \mu\, d/dx; \tag{8.2}$$
$$\zeta = \frac{1}{2}\sum_{k=0}^{N} b_k P_k(\mu) \int_{-1}^{1} d\mu'\, P_k(\mu').$$

The operator ζ is bounded since

$$||\zeta f|| \leq \frac{1}{2}\sum_{k=0}^{N} |b_k|\, ||P_k(\mu)||\, ||f|| \leq \text{const}\, ||f||. \tag{8.3}$$

Also for $u \in D(A)$

$$2\,\mathrm{Re}\,(Du, u) = \int_{-1}^{1} d\mu \int_{-a}^{a} dx\left[\mu\frac{\partial u}{\partial x}\bar{u} + \mu\frac{\partial \bar{u}}{\partial x}u\right]$$
$$= \int_0^1 \mu\, d\mu\, |u(a, \mu)|^2 - \int_{-1}^{0} \mu\, d\mu\, |u(-a, \mu)|^2 \geq 0. \tag{8.4}$$

Thus for $\beta = \mathrm{Re}\,\lambda$

$$\begin{aligned}\mathrm{Re}\,((\lambda - A)u, u) &= \beta\, ||u||^2 + \mathrm{Re}\,(Du, u) - c\,\mathrm{Re}\,(\zeta u, u)\\ &\geq \beta\, ||u||^2 - c\,\mathrm{Re}\,(\zeta u, u)\\ &\geq \beta\, ||u||^2 - c\,|(\zeta u, u)|\\ &\geq (\beta - c\,||\zeta||)\,||u||^2.\end{aligned} \tag{8.5}$$

Now set $g = (\lambda - A)u$. Then Eq. (8.5) gives

$$||g||\cdot||u|| \geq |(g, u)| \geq (\beta - c\,||\zeta||)\,||u||^2, \tag{8.6}$$

or, for $\beta > C||\zeta||$,

$$||u|| \leq ||g||/(\beta - c\,||\zeta||). \tag{8.7}$$

From Eq. (8.7) we see that

$$||R_\lambda|| \leq 1/(\beta - c\,||\zeta||). \tag{8.8}$$

With this estimate the Hille–Yosida theorem[8] assures the existence of the semigroup $T(t)$ of bounded solution operators continuous for $t \geq 0$, with $T(0) = I$. This family satisfies the differential equation

$$(d/dt)T(t)f = AT(t)f, \qquad f \in D(A), \tag{8.9}$$

and thus solves the initial-value problem for the transport equation.

We can write the solution of our problem also in the integral form

$$\psi(x, \mu, t) = \lim_{\omega\to\infty}\frac{1}{2\pi i}\int_{\alpha - i\omega}^{\alpha + i\omega} e^{\lambda t}R_\lambda f\, d\lambda,$$
$$t > 0 \quad f \in D(A), \quad \alpha > c\,||\zeta||. \tag{8.10}$$

The integral is to be understood as a strong limit of Riemann sums.

[8] E. Hille, "Functional Analysis and Semi-groups," Colloq. Am. Math. Soc. **31** (1948).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 7, NUMBER 5 MAY 1966

The Initial-Value Problem for a Slab. II. Nonuniform Slab

JANUSZ MIKA
Institute of Nuclear Research, Swierk k/Otwocka, Poland
(Received 22 June 1965; final manuscript received 5 October 1965)

This paper deals with the time-dependent neutron transport in a plane slab of material with variable nuclear properties. The one-velocity theory and the isotropic scattering are assumed. The spectral properties of the corresponding Boltzmann operator are found, and the initial-value problem is solved.

1. INTRODUCTION

IN Part I of this paper,[1] the author has considered the initial-value problem for a monoenergetic Boltzmann equation for a slab assuming the anisotropic scattering of neutrons. In this part we deal with a monoenergetic slab in which the nuclear properties of the material are dependent on position.

In practical applications one would use the results presented here mostly for the case of a system of uniform slabs. Such a situation may often happen in the experimental neutron physics.

In this paper we find some of the spectral properties of the Boltzmann operator for a considered system and solve the initial-value problem.

For simplicity, the isotropic scattering of neutrons is assumed. One could use, however, the results of I and consider the general case of a nonuniform slab with anisotropic scattering.

2. FORMULATION OF THE PROBLEM

Consider an infinite slab extending from $x = -a$ to $x = a$ surrounded by a vacuum. The time-dependent transport of monoenergetic neutrons is described by the Boltzmann equation

$$\frac{\partial N}{\partial t} + \mu \frac{\partial N}{\partial x} + \sigma(x)N(x, \mu, t) = \frac{c(x)\sigma(x)}{2} \int_{-1}^{1} d\mu' \, N(x, \mu', t). \tag{2.1}$$

The standard notation is used here. The neutron velocity is assumed equal to unity. The nuclear properties of the medium described by the total cross section $\sigma(x)$ and the mean number of secondaries per collision $c(x)$ are functions of position. The scattering is assumed to be isotropic. $N(x, \mu, t)$ describes the angular neutron distribution as a function of time, position, and the cosine of an angle between neutron velocity and the axis x.

[1] J. Mika, J. Math. Phys. 7, 833 (1966); preceding paper in this issue, hereafter referred to as I.

The boundary and initial conditions for Eq. (2.1) are the following:

$$N(\pm a, \mu, t) = 0, \qquad \mu \lessgtr 0, \qquad t > 0, \qquad N(x, \mu, 0) = f(x, \mu). \tag{2.2}$$

The solution of the initial-value problem formulated above requires the knowledge of properties of the Boltzmann operator A defined by Eq. (2.1),

$$A = -\mu \frac{\partial}{\partial x} \cdot - \sigma(x) \cdot + \frac{c(x)\sigma(x)}{2} \int_{-1}^{1} d\mu' \cdot . \tag{2.3}$$

The most important are the spectral properties of A. In other words, one has to find for which values of λ there exists a solution to the equation

$$(\lambda - A)\psi = 0, \tag{2.4}$$

where $\psi(x, \mu)$ satisfies the homogeneous boundary conditions

$$\psi(\pm a, \mu) = 0, \qquad \mu \lessgtr 0. \tag{2.5}$$

Define the Hilbert space $\mathfrak{H}$ of square-integrable functions in rectangle $|\mu| \leq 1$, $|x| \leq a$ with the scalar product

$$(f, g) = \int_{-a}^{a} dx \int_{-1}^{1} d\mu f(x, \mu)\overline{g(x, \mu)} \tag{2.6}$$

and the norm

$$||f|| = (f, f)^{\frac{1}{2}}. \tag{2.7}$$

The bar denotes a complex conjugate.

We will work also with the space L_2 of functions square-integrable in $(-a, a)$. The scalar product and the norm in L_2 are defined as follows:

$$(f, g) = \int_{-a}^{a} dx f(x)\overline{g(x)} \tag{2.8}$$

and

$$||f|| = (f, f)^{\frac{1}{2}}. \tag{2.9}$$

No confusion will arise from the fact that the same

notation is used for the scalar product and norm in $\mathfrak{H}$ and L_2 spaces.

The operator A on $\mathfrak{H}$ has as the domain $D(A)$ all functions absolutely continuous in x for each fixed μ, $|\mu| \leq 1$, satisfying the boundary conditions (2.5).

For simplicity, we will assume that both $c(x)$ and $\sigma(x)$ are nonnegative bounded functions, square-integrable in any interval contained in $(-a, a)$. These assumptions might be too restrictive from the mathematical point of view, but they are obviously satisfied by any physically relevant function.

Introduce an average value of $\sigma(x)$ in an interval (x_1, x_2) contained in $(-a, a)$.

$$\sigma(x_1, x_2) = \int_{x_1}^{x_2} \sigma(x)\, dx/(x_2 - x_1). \tag{2.10}$$

Denoting by σ_m and σ_M the minimum and maximum values of $\sigma(x)$ in $(-a, a)$, respectively, we have

$$\sigma_m \leq \sigma(x_1, x_2) \leq \sigma_M; \quad \sigma_m > 0; \quad \sigma_M > 0 \tag{2.11}$$

for $x_1, x_2 \in (-a, a)$.

Similarly,

$$c_m \leq c(x) \leq c_M; \quad c_m > 0; \quad c_M > 0 \tag{2.12}$$

for $x \in (-a, a)$.

3. THE INTEGRAL EQUATION

Equation (2.4) can be formally solved with respect to $\psi(x, \mu)$. We get

$$\begin{aligned}\psi(x, \mu) &= \frac{1}{2\mu} \int_{-a}^{x} dx' \\ &\quad \times \exp\left[-\left(\int_{x'}^{x} dx''(\lambda + \sigma(x''))\right) \Big/ \mu\right] \\ &\quad \times c(x')\sigma(x')\phi(x'), \qquad \mu > 0, \\ &= \frac{1}{2\mu} \int_{a}^{x} dx' \\ &\quad \times \exp\left[-\left(\int_{x'}^{x} dx''(\lambda + \sigma(x''))\right) \Big/ \mu\right] \\ &\quad \times c(x')\sigma(x')\phi(x'), \qquad \mu < 0,\end{aligned} \tag{3.1}$$

where

$$\phi(x) = \int_{-1}^{1} \psi(x, \mu)\, d\mu. \tag{3.2}$$

By integrating both sides of Eq. (3.1) with respect to μ, one can obtain an integral equation for $\phi(x)$ if

$$\operatorname{Re} \lambda \geq -\sigma_m; \qquad \lambda \neq -\sigma_m. \tag{3.3}$$

One may notice that if $\sigma(x', x) = \sigma_m$ on the set of zero measure then $\lambda = -\sigma_m$ is allowed.

If the above condition is satisfied we get

$$\begin{aligned}\phi(x) = \frac{1}{2} \int_{-a}^{a} dx' c(x')\sigma(x')\phi(x') \\ \times E_1[(\lambda + \sigma(x', x))\, |x - x'|],\end{aligned} \tag{3.4}$$

where the function $E_1(z)$ is defined in I.

The integral equation (3.7) will be written in an operational form

$$\phi = \tfrac{1}{2} F_\lambda \phi. \tag{3.5}$$

The definition of the operator F_λ follows from Eq. (3.4).

The presented derivation shows that every solution of Eq. (2.4) with the boundary conditions (2.5) yields the solution of Eq. (3.5). By performing the derivation in the opposite direction we see that, conversely, every solution of Eq. (3.5) yields a solution of Eq. (2.4).

In the following we will use the notation for the spectrum of a given operator T as introduced in I.

Using this notation we state the following:

Theorem 1: A necessary and sufficient condition that $\lambda \in P\sigma A$ on $\mathfrak{H}$ with $\operatorname{Re} \lambda \geq -\sigma_m$, $\lambda \neq -\sigma_m$, is that 2 belongs to $P\sigma F_\lambda$ on L_2. The eigenfunctions $\psi(x, \mu)$ and $\phi(x)$ are related to each other by Eqs. (3.1) and (3.2).

For fixed λ such that $\operatorname{Re} \lambda > -\sigma_m$, F_λ is a compact operator. To show this we can write for the norm of F_λ

$$\begin{aligned}||F_\lambda||^2 \leq \int_{-a}^{a} dx \int_{-a}^{a} dx' c^2(x')\sigma^2(x') \\ \times |E_1((\lambda + \sigma(x', x))\, |x - x'|)|^2.\end{aligned} \tag{3.6}$$

From the properties of E_1, it follows that

$$\begin{aligned}&|E_1((\lambda + \sigma(x', x))\, |u|)| \\ &\quad \leq \frac{\text{const}}{|\lambda + \sigma(x', x)|\, |u|} \leq \frac{\text{const}}{|\lambda + \sigma_m|\, |u|}.\end{aligned} \tag{3.7}$$

For $|\lambda + \sigma_m\, ||u| < 1$ we use another estimate,

$$\begin{aligned}&|E_1((\lambda + \sigma(x', x))\, |u|)|^2 \\ &\quad \leq \text{const}\, [\log^2 (|\lambda + \sigma(x', x)|\, |u|) + \tfrac{1}{4}\pi^2] \\ &\quad \leq \text{const}\, [\log^2 (|\lambda + \sigma_m|\, |u|) + \tfrac{1}{4}\pi^2].\end{aligned} \tag{3.8}$$

Let us rewrite Eq. (3.6),

$$\begin{aligned}||F_\lambda||^2 &\le c_M^2\sigma_M^2 \int_{-a}^{a} dx' \int_{-\infty}^{\infty} du \\ &\times |E_1((\lambda + \sigma(x', u + x'))\,|u|)|^2 \\ &\le c_M^2\sigma_M^2 \int_{-a}^{a} dx' \\ &\times \left[\int_{-1/|\lambda+\sigma_m|}^{1/|\lambda+\sigma_m|} + \int_{-\infty}^{-1/|\lambda+\sigma_m|} + \int_{1/|\lambda+\sigma_m|}^{\infty}\right] du \\ &\times |E_1(\cdots)|^2. \end{aligned} \tag{3.9}$$

Now from Eqs. (3.7) and (3.8) we see that each of the integrals in Eq. (3.9) exists and is majorized by an integral independent of x'. Performing the integration we have finally that

$$||F_\lambda||^2 \le \text{const}/|\lambda + \sigma_m|. \tag{3.10}$$

Thus we have proved that F_λ is a compact operator for $\text{Re}\,\lambda > -\sigma_m$.

4. THE REGION $\text{Re}\,\lambda \le -\sigma_m$

We shall now show that all the points λ such that $\text{Re}\,\lambda \le -\sigma_m$ belong to the spectrum of A. To do so it is enough to find the functions $\psi_\delta \in D(A)$ such that

$$||\psi_\delta|| \ge \text{const} > 0 \tag{4.1}$$

and

$$\lim_{\delta\to 0} ||(\lambda - A)\psi_\delta|| = 0. \tag{4.2}$$

For $\text{Re}\,\lambda \le -\sigma_m$ and $0 < \delta < \frac{1}{2}$, we choose the following functions:

$$\begin{aligned}\psi_\delta(x, \mu) &= (x + a)b_\delta(\mu) \\ &\times \exp\,[-(x - a)(\lambda + \sigma(x, a))/\mu]\end{aligned} \tag{4.3}$$

where

$$b_\delta(\mu) = \begin{cases} 1/\delta, & \delta^2 \le \mu \le \delta, \\ 0, & \text{otherwise.}\end{cases} \tag{4.4}$$

One can now easily check that $\psi_\delta(x, \mu)$, defined as above, belongs to $D(A)$, and it has both of the required properties given by Eqs. (4.1) and (4.2). In fact, $\psi_\delta(x, \mu)$ is an absolutely continuous function of x for each given μ, $|\mu| \le 1$, and

$$\psi_\delta(\pm a, \mu) = 0, \qquad \mu \lessgtr 0. \tag{4.5}$$

Further, we can show that Eqs. (4.1) and (4.2) are also satisfied.

In such way we have shown that for $\text{Re}\,\lambda \le -\sigma_M$ the points λ belong to the spectrum of A.

Now consider λ such that $-\sigma_M < \text{Re}\,\lambda < -\sigma_m$. Assume further that the minimum value of $\sigma(x)$ is attained at one of the boundaries, for instance, at $x = a$. Now select a point x_0 such that $\sigma(x) \le -\text{Re}\,\lambda$ for $x \ge x_0$ and construct for $0 < \delta < \frac{1}{2}$ the following function:

$$\bar{\psi}_\delta(x, \mu) = \begin{cases} 0, & -a \le x \le x_0, \\ (x - x_0)b_\delta(\mu)\exp[-(x - x_0)(\lambda + \sigma(x, x_0))/\mu], & \\ & x_0 \le x \le a. \end{cases} \tag{4.6}$$

This function obviously belongs to $D(A)$ since it is completely continuous and satisfies the boundary conditions (4.5).

We can now prove that $\bar{\psi}_\delta(x, \mu)$ fulfills Eqs. (4.1) and (4.2) in exactly the same way as that used for $\psi_\delta(x, \mu)$. It is sufficient only to substitute x_0 instead of $-a$ in all corresponding formulas.

The restriction concerning the minimum value of $\sigma(x)$ attained at the boundary can be relaxed in the following way. In the case when this minimum is attained inside the body we put an additional layer of black material with $c(x) \equiv 0$ and $\sigma(x) \equiv \sigma_m$. Such a layer does not change the behavior of neutrons inside the body since the boundary conditions are the same, namely, in both cases there are no incident neutrons. But the previous proof is now valid.

Finally we see that the points $-\sigma_M < \text{Re}\,\lambda < -\sigma_m$ also belong to the spectrum of A. Now σA is a closed set and we have

Theorem 2: The region $\text{Re}\,\lambda \le -\sigma_m$ is contained in the spectrum of the operator A.

5. THE RESOLVENT SET

Denote by Γ a set of points λ, such that $\text{Re}\,\lambda > -\sigma_m$ and $\lambda \notin P\sigma A$. The points of Γ belong to $R\sigma A$, $C\sigma A$, or $\mathfrak{M}A$. Then $R_\lambda = (\lambda - A)^{-1}$ always exists. We will show that R_λ is bounded and $D(R_\lambda) = \mathfrak{H}$ or, in other words, that $\Gamma \in \mathfrak{M}A$.

Consider the equation

$$(\lambda - A)u = g, \qquad g \in \mathfrak{H}. \tag{5.1}$$

This equation allows for the formal solution which may be obtained in a way similar to that of Sec. 3:

$$\begin{aligned}R_\lambda g = u(x, \mu) &= \frac{1}{\mu}\int_{-a}^{x} dx' \exp\,[-(x - x')(\lambda + \sigma(x', x))/\mu] \\ &\times [2c(x')\sigma(x')\xi(x') + g(x', \mu)], \quad \mu > 0, \\ &= \frac{1}{\mu}\int_{a}^{x} dx' \exp\,[-(x - x')(\lambda + \sigma(x', x))/\mu] \\ &\times [2c(x')\sigma(x')\xi(x') + g(x', \mu)], \quad \mu < 0,\end{aligned} \tag{5.2}$$

with

$$\xi(x) = \int_{-1}^{1} u(x, \mu)\, d\mu. \tag{5.3}$$

For $\mathrm{Re}\,\lambda > -\sigma_m$ we can integrate both sides of Eq. (5.2) with respect to μ. The result is

$$\xi = \tfrac{1}{2}F_\lambda \xi + G. \tag{5.4}$$

The operator F_λ has been defined by Eq. (3.9). The function G has the form

$$\begin{aligned} G &= G_1 + G_2 \\ &= \int_0^1 H_1(x, \mu)\, d\mu + \int_{-1}^{0} H_2(x, \mu)\, d\mu \\ &= \int_0^1 \frac{d\mu}{\mu} \int_{-a}^{x} dx' \\ &\quad \times \exp\left[-(x - x')(\lambda + \sigma(x', x))/\mu\right] g(x', \mu) \\ &\quad + \int_{-1}^{0} \frac{d\mu}{\mu} \int_{a}^{x} dx' \\ &\quad \times \exp\left[-(x - x')(\lambda + \sigma(x', x))/\mu\right] g(x', \mu). \end{aligned} \tag{5.5}$$

Now using the method similar to that of I we can show that $G \in L_2$ if $g \in \mathfrak{H}$.

In Sec. 3 we have shown that F_λ is a compact operator. For $\lambda \in \Gamma$ by virtue of Theorem 1, 2 is not an eigenvalue of F_λ. Then ξ is uniquely determined for $g \in \mathfrak{H}$ from Eq. (5.4). In other words, $(1 - \frac{1}{2}F_\lambda)^{-1}$ exists for $\lambda \in \Gamma$ and its domain is all of L_2. Since $(1 - \frac{1}{2}F_\lambda)$ is bounded and one to one, $(1 - \frac{1}{2}F_\lambda)^{-1}$ is also bounded. Therefore,

$$||\xi|| \leq \text{const}\, ||G|| \leq \text{const}/(\beta + \sigma_m)\, ||g||. \tag{5.6}$$

Now put

$$h(x, \mu) = 2c(x)\sigma(x)\xi(x, \mu) + g(x, \mu). \tag{5.7}$$

From Eq. (5.6) we have

$$||h|| \leq \text{const}\, ||g||. \tag{5.8}$$

This shows that $h \in \mathfrak{H}$.

To get the estimation of $u(x, \mu)$ as given by Eq. (5.2), we use the same procedure as that for G, substituting $h(x, \mu)$ for $g(x, \mu)$. We get

$$||u|| \leq \text{const}\, ||h||. \tag{5.9}$$

We see that $u \in \mathfrak{H}$ and is bounded. Therefore $R_\lambda = (\lambda - A)^{-1}$ is bounded and its domain is $\mathfrak{H}$.

Thus we have proved

Theorem 3: The region $\mathrm{Re}\,\lambda > -\sigma_m$, deleting points $\lambda \in P\sigma A$, is contained in the resolvent set of A.

From the estimate (3.10) one sees that for sufficiently large values of $|\lambda + \sigma_m|$ the Neumann series for F_λ converges. This shows on the basis of Theorem 3 that the spectrum of A is bounded in the region $\mathrm{Re}\,\lambda > -\sigma_m$. That result is obvious from the physical point of view since it means that there exists an upper bound for the asymptotic decay constant.

6. SOLUTION OF THE INITIAL-VALUE PROBLEM

Having found some of the spectral properties of the operator A, we turn now to the original initial value problem stated in Eqs. (2.1) and (2.2).

By the Hille–Yosida theorem (Ref. 8 of I), the initial-value problem for A generates a semigroup of solution operators if the following properties of A are satisfied:

(1) A is closed;
(2) $D(A)$ is dense in $\mathfrak{H}$;
(3) $||R_\lambda|| < (\lambda - \gamma)^{-1}$; λ real and $\gamma > 0$.

R_λ exists and is bounded. Its domain is $\mathfrak{H}$. Hence it is closed. Therefore, so is A. This proves the first property.

The second property follows immediately from the definition of the domain of A.

To prove the third property let us write

$$A = -D - \sigma + \zeta, \tag{6.1}$$

where

$$D = \mu(\partial/\partial x)\cdot; \quad \sigma = \sigma(x)\cdot; \quad \zeta = \frac{c(x)\sigma(x)}{2} \int_{-1}^{1} d\mu'\cdot \tag{6.2}$$

The operators σ and ζ are obviously bounded. Denote $k = ||\zeta|| \leq c_M \sigma_M$. For $u \in D(A)$ we have

$$\begin{aligned} &2\,\mathrm{Re}\,(Du, u) \\ &\quad = \int_{-1}^{1} d\mu \int_{-a}^{a} dx \left[\mu \frac{\partial u}{\partial x}\bar{u} + \mu \frac{\partial \bar{u}}{\partial x}\right] \\ &\quad = \int_0^1 \mu\, d\mu\, |u(a, \mu)|^2 - \int_{-1}^{0} \mu\, d\mu\, |u(-a, \mu)|^2 \geq 0. \end{aligned} \tag{6.3}$$

Thus for $\beta = \mathrm{Re}\,\lambda$

$$\begin{aligned} &\mathrm{Re}\,((\lambda - A)u, u) \\ &\quad = \beta\, ||u||^2 + (\sigma u, u) + \mathrm{Re}\,(Du, u) - \mathrm{Re}\,(\zeta u, h) \\ &\quad \geq (\beta + \sigma_m)\, ||u||^2 - |(\zeta u, u)| \\ &\quad \geq (\beta + \sigma_m - k)\, ||u||^2. \end{aligned} \tag{6.4}$$

Now put $g = (\lambda - A)u$. Then Eq. (6.4) gives

$$||u||\, ||g|| \geq |(g, u)| \geq (\beta + \sigma_m - k)\, ||u||^2, \tag{6.5}$$

or for $\beta + \sigma_m > k$,

$$||u|| \leq ||g||/(\beta + \sigma_m - k). \tag{6.6}$$

This means that for $\beta + \sigma_m > k$

$$||R_\lambda|| \leq (\beta + \sigma_m - k)^{-1}, \tag{6.7}$$

which proves the third property of A.

Now the Hille–Yosida theorem states that there exists a semigroup of bounded operators $T(t)$ continuous for $t \geq 0$ with $T(0) = 1$. This family satisfies the differential equation

$$(d/dt)T(t)f = AT(t)f, \qquad f \in D(A), \tag{6.8}$$

and thus solves the initial-value problem stated in Eqs. (2.1) and (2.2).

The solution can also be written in the form

$$N(x, \mu, t) = \lim_{\omega \to \infty} \frac{1}{2\pi i} \int_{\alpha - i\omega}^{\alpha + i\omega} e^{\lambda t}[R_\lambda f]\, d\lambda,$$

$$t > 0, \quad f \in D(A), \quad \alpha > k = ||\zeta||. \tag{6.9}$$

The integral is to be understood as a strong limit of Riemann sums.

ACKNOWLEDGMENTS

I wish to express my gratitude to Professor R. Żelazny for his constant interest in this work. Credit is also due to Dr. R. Bednarz and Dr. A. Kuszell for valuable discussions.

The Physical Regions of Many-Particle Processes*

RICHARD A. MORROW
Wilder Laboratory, Dartmouth College, Hanover, New Hampshire
(Received 22 November 1965)

Lorentz invariant expressions, in the form of determinantal conditions, are derived for the physical regions of many-particle processes. They are explicitly solved in the case of five-particle processes and the solutions are exhibited in the planes of pairs of the five independent kinematic invariants.

I. INTRODUCTION

WHEN studying a process involving a number of relativistic particles it is sometimes desirable to have a knowledge of the physical regions, that is, the physically accessible ranges of the kinematic variables used to describe the process. Byers and Yang[1] in a recent paper carried out such a study of the physical regions of N-particle processes in terms of Lorentz-invariant variables. It is proposed here to present an alternate derivation leading to simpler conditions and to apply them to the case $N = 5$. The method for doing this is based on a theorem of Omnes[2] and in essence results in a generalization of Kibble's[3] treatment of two-particle scattering.

In the next section Lorentz invariant expressions in the form of determinantal conditions are derived for the physical regions of N-particle processes. The masses of all particles are regarded as given and the kinematic invariants chosen as variables are the scalar products of the 4-momenta involved. Section III treats the case of five particles. The general expressions are solved explicitly and the solutions are exhibited as functions of two independent variables with the three remaining variables treated as parameters.

All quantities throughout are real.

II. N-PARTICLE PROCESSES

In an N-particle reaction let the 4-momentum of the ith particle of mass m_i be p_i with $p_i^2 = m_i^2$. The variables used to describe the physical regions will be a linearly independent set chosen from the set of invariant scalar products $p_i \cdot p_j \equiv x_{ij}$. The specification of which particles are incoming and which are outgoing is not made. Hence, the resulting equations will give the physical regions of all processes which involve a total of N particles (assuming a particle and its antiparticle to have the same mass). The procedure for the construction of these regions is an application of a theorem of Omnes[2] which is presented here in slightly modified form without proof.

Theorem: Given m real linearly independent 4-vectors $p_i (m \leq 3)$ with $p_i^2 = m_i^2 > 0$ there exists another real 4-vector p_{m+1} with $p_{m+1}^2 = m_{m+1}^2$ and $p_i \cdot p_{m+1} \equiv x_{im+1}$ $(i = 1, 2, \cdots, m)$ real if and only if the Gram determinant $\Delta_{m+1} \equiv \det (x_{ij})$ with $i, j = 1, 2, \cdots, m + 1$ satisfies

$$(-1)^{m+1}\Delta_{m+1} \leq 0, \tag{1}$$

where (x_{ij}) is the obvious $(m + 1) \times (m + 1)$ symmetric matrix.

It is apparent that the desired physical regions can be obtained in the (real) space of the scalar products of the 4-momenta by applying this theorem in steps beginning with $m = 1$. To be explicit the construction proceeds as follows:

(a) Given p_1 then p_2 exists if and only if the determinant

$$\Delta_2 \leq 0, \tag{2}$$

which gives the range of the kinematic invariant x_{12}.

(b) Assuming (2) holds then p_3 exists if and only if

$$\Delta_3 \geq 0. \tag{3}$$

Consequently (2) and (3) together give the ranges of the x_{ij} for $i, j, = 1, 2, 3$.

(c) Assuming (2) and (3) hold then p_4 exists if and only if

$$\Delta_4 \leq 0. \tag{4}$$

Thus (2), (3), and (4) together establish the ranges of the x_{ij} $(i, j = 1, 2, 3, 4)$ in order that the four 4-momenta exist simultaneously as real 4-vectors.

The theorem can be applied no further since the

* Partially contained in the author's Ph.D. thesis (Princeton 1963, unpublished).

[1] N. Byers and C. N. Yang, Rev. Mod. Phys. **36**, 595 (1964).

[2] R. Omnes, in *Dispersion Relations and Elementary Particles* (Les Houches 1960; John Wiley and Sons, New York, 1960), p. 341.

[3] T. W. B. Kibble, Phys. Rev. **117**, 1159 (1960).

case $m = 3$ has been reached. This is quite alright since four real linearly independent 4-vectors have been established and can be taken to span the four-dimensional space of the vectors.

(d) Thus, assuming (2), (3), and (4) hold a fifth vector, p_5 can exist as a real 4-vector if and only if it is a linear combination of the four vectors already established. This is guaranteed by requiring $\Delta_5 = 0$.

In fact, any number of additional vectors p_5, p_6, etc. can exist simultaneously as real 4-vectors if and only if all sets of five vectors are linearly dependent. Asribekov[4] has shown that if there are a total of $N - 1$ vectors then this condition of linear dependency is assured by the $\frac{1}{2}(N-5)(N-4)$ conditions

$$\det \begin{vmatrix} x_{11} & x_{12} & x_{13} & x_{14} & x_{1j} \\ x_{21} & x_{22} & x_{23} & x_{24} & x_{2j} \\ x_{31} & x_{32} & x_{33} & x_{34} & x_{3j} \\ x_{41} & x_{42} & x_{43} & x_{44} & x_{4j} \\ x_{i1} & x_{i2} & x_{i3} & x_{i4} & x_{ij} \end{vmatrix} = 0,$$

$$\text{for } i, j = 5, 6, \cdots, N-1. \qquad (5)$$

(e) There is one further condition to be imposed on an N-particle process namely conservation of 4-momentum which takes the form

$$\sum_{i=1}^{N} p_i = 0, \qquad (6a)$$

where p_i is forward (backward) timelike if the particle is incoming (outgoing).

This equation can be put in terms of the scalar products x_{ij} by taking its scalar product with all the p_i to obtain a total of N conditions:

$$\sum_{i=1}^{N} x_{ij} = 0 \quad (j = 1, 2, \cdots, N). \qquad (6b)$$

Thus the physical regions of N-particle processes in the space of the inner products of the 4-momenta is given by the intersections of the regions defined by conditions (2) through (6). The number of linearly independent scalar products can be found quite simply from the above to be $3N - 10$, a well-known result.

The general discussion is not carried any further. As a familiar example the physical regions of four-particle processes are given by the intersection of the regions defined by (2) and (3) with condition (6) holding. These are exactly the equations considered by Kibble[3] and reference is made to his analysis.

[4] V. E. Asribekov, Nucl. Phys. **34**, 461 (1962).

III. FIVE-PARTICLE PROCESSES

The physical regions of reactions involving a total of five particles are given by the intersection of the regions defined by (2), (3), and (4) with condition (6) holding. There are five linearly independent variables in this case which will be chosen to be $s = (p_1 + p_2)^2$, $t = (p_2 + p_3)^2$, $u = (p_3 + p_4)^2$, $v = (p_4 + p_5)^2$, and $w = (p_5 + p_1)^2$, and are clearly linearly related to the x_{ij} of the previous section.

Rather than attempt to treat each of these kinematic invariants on an equal footing, the physical regions are studied in the planes of pairs of them while regarding the remaining ones as parameters. It is easy to convince oneself that such a study need be carried out only in two planes, e.g., the $s - u$ and the $s - w$ planes, in order to cover all possibilities.

The procedure will be to impose conservation of 4-momentum (6a) and solve (4) leaving (2) and (3) to enter naturally. Hence using (6a) and the above choice of variables, (4) can be transformed into the symmetric determinant

$$\begin{vmatrix} 2m_1^2 & s-m_1^2-m_2^2 & v-s-t+m_2^2 & t-w-v+m_5^2 \\ \cdot & 2m_2^2 & t-m_2^2-m_3^2 & w-t-u+m_3^2 \\ \cdot & \cdot & 2m_3^2 & u-m_3^2-m_4^2 \\ \cdot & \cdot & \cdot & 2m_4^2 \end{vmatrix} \leq 0.$$

Adding appropriate rows and columns and changing the sign of the first row and column the determinant becomes (still symmetric)

$$L \equiv \begin{vmatrix} 2v & -v-t+m_1^2 & s-v-m_3^2 & v+m_4^2-m_5^2 \\ \cdot & 2t & t-m_2^2+m_3^2 & w-t-m_4^2 \\ \cdot & \cdot & 2m_3^2 & u-m_3^2-m_4^2 \\ \cdot & \cdot & \cdot & 2m_4^2 \end{vmatrix} \qquad (7)$$

with condition (4) replaced by $L \leq 0$.

L may be regarded as the determinant of a symmetric matrix whose elements a_{ij} are defined by (7). The advantage of this form of L is that the variables s, u, and w each occur only in the elements a_{13}, a_{34}, and a_{24} (and, of course, the transposed elements a_{31}, a_{43}, and a_{42}), respectively. A study of the curve $L = 0$ in the plane of any two of these variables is then relatively straightforward. In particular, the physical regions will be constructed in the $s - u$ and the $s - w$ planes by establishing conditions on the remaining parameters such that

the curve $L = 0$ will occur in these planes and such that $L \leq 0$ will define the physical regions.

The notation used in this section is the following. L is the determinant (7). $L_{ij\cdots k}$ is the principal minor of L obtained by deleting rows and columns $i, j, \cdots, k$. V_{ij} is the cofactor (signed minor) of the element a_{ij} of L; in particular $V_{ii} = L_i$. $V(ij \cdots k)_{mn}$ is the cofactor of the element a_{mn} in $L_{ij\cdots k}$.

The study of curves defined by the vanishing of a determinant like $L = 0$ is facilitated by a method developed by Tarski[5] which is based on the Jacobi ratio theorem.[6] Some of the useful formulas are derived in the Appendix and one result, Eq. (A2), is that

$$L_iL_j - L_{ij}L - V_{ij}^2 = 0. \tag{8}$$

This gives the important information that the surface $L = 0$, as a function of all five variables, exists only in regions where all L_i have the same sign.

Also of importance is the solution of $L = 0$ for the a_{ij} which is, from (A4),

$$a_{ij} = [V_{ij0} \pm (L_iL_j)^{\frac{1}{2}}]/L_{ij}, \tag{9}$$

where $V_{ij0} \equiv V_{ij}$ evaluated at $a_{ij} = 0$. This shows that the surfaces $L = 0$ and $L_i = 0$ are tangent when they meet.

For later convenience expressions for the L_{ij} are recorded below.

$$\begin{aligned}
L_{12} &= -[u - (m_3 + m_4)^2][u - (m_3 - m_4)^2],\\
L_{13} &= -[w - (t^{\frac{1}{2}} + m_4)^2][w - (t^{\frac{1}{2}} - m_4)^2],\\
L_{14} &= -[t - (m_2 + m_3)^2][t - (m_2 - m_3)^2],\\
L_{23} &= -[v - (m_4 + m_5)^2][v - (m_4 - m_5)^2],\\
L_{24} &= -[s - (v^{\frac{1}{2}} + m_3)^2][s - (v^{\frac{1}{2}} - m_3)^2],\\
L_{34} &= -[t - (v^{\frac{1}{2}} + m_1)^2][t - (v^{\frac{1}{2}} - m_1)^2].
\end{aligned} \tag{10}$$

Physical Regions in the s–u Plane

Regarded as a function of s and u, or equivalently of a_{13} and a_{34}, L is quadratic in these variables and the curve $L = 0$ is a conic section in their plane. $L_2 = 0$ is also a conic section in this plane but all other principal minor curves will either be straight lines or will not appear.

The necessary and sufficient conditions for $L \leq 0$ to define the physical regions are simply obtained. It is straightforward to put L_3 in terms of the inner products of the 4-momenta. In fact it involves p_1, p_4, and p_5 in the form of the Gram determinant Δ_3. Thus (3) requires $L_3 \geq 0$. Since this does not involve s or u, it puts a restriction on the remaining parameters. In a similar manner, L_{23} may be put in the form of the Gram determinant Δ_2 involving p_4 and p_5. Thus (2) requires $L_{23} \leq 0$. Because conservation of momentum has been built into L the conditions $L_{23} \leq 0$ and $L_3 \geq 0$ guarantee that $L \leq 0$ defines the physical regions. All that remains is to find the conditions such that $L = 0$ will actually occur in the $a_{13} - a_{34}$ plane.

First of all $L_3 \geq 0$ requires $L_{34} \leq 0$. To see this consider (8) with $i = 2$, $j = 4$, and L_3 and its cofactors replacing L and its cofactors:

$$L_3 = -\{[V(3)_{24}]^2 - L_{23}L_{34}\}/L_{234},$$

where, from (7), $L_{234} = 2v$. Thus, if $v > 0$ then L_3 can be nonnegative only if $L_{34} \leq 0$ (since $L_{23} \leq 0$). On the other hand, if $v < 0$ then from (10) $L_{34} < 0$ for all values of t. In both cases w must be chosen to make L_3 nonnegative. The permissible range of w may be found by a simple application of (A5) to L_3 with $a_{ij} = a_{24} \equiv w - t - m_4^2$.

Next, from (8), $L = 0$ occurs only in regions where all L_i have the same sign which, since L_3 is nonnegative, requires that L_4 be nonnegative. This in turn requires the condition $L_{14} \leq 0$ the proof of which is similar to that above. From (8) with $i = 1$, $j = 3$, and L_4 and its cofactors replacing L and its cofactors:

$$L_4 = -\{[V(4)_{13}]^2 - L_{14}L_{34}\}/L_{134},$$

where, from (7), $L_{134} = 2t$. Thus, if $t > 0$ then L_4 can be nonnegative only if $L_{14} \leq 0$ (since $L_{34} \leq 0$ as shown above). On the other hand, if $t < 0$ then from (10) $L_{14} < 0$. In both cases $L_4 = 0$ will occur in the s–u plane and so there will be regions where L_4 is nonnegative. The equations of these straight lines may be found by applying (9) to L_4 with $i = 1$, $j = 3$:

$$s - v - m_3^2 \equiv a_{13} = [V(4)_{130} \pm (L_{14}L_{34})^{\frac{1}{2}}]/L_{134}.$$

Furthermore with L_3 and L_4 nonnegative it can be seen from (9) with $i = 3$, $j = 4$ say, that the curve $L = 0$ will occur in the s–u plane.

Consequently, the necessary and sufficient conditions, i.e., restrictions on t, v, and w, in order that physical regions exist in the s–u plane, are L_{23}, L_{34}, $L_{14} \leq 0$, and $L_3 \geq 0$. These conditions may be imposed in the following manner: (i) choose v such that $L_{23} \leq 0$, (ii) choose t such that L_{34}, $L_{14} \leq 0$, and (iii) choose w such that $L_3 \geq 0$. The physical regions will then be given in the s–u plane by the

[5] J. Tarski, J. Math. Phys. 1, 149 (1960).
[6] A. C. Aitken, *Determinants and Matrices* (Oliver and Boyd, Edinburgh, 1956), 9th ed., p. 97.

regions $L \leq 0$. From the asymptotic form of $L = 0$ given by (A6) it can be seen that the curve will be a hyperbola if $t < 0$ and an ellipse if $t > 0$.

The form of the relevant curves in a typical example is shown in Fig. 1. The two physical regions present are shaded. Region A refers to the process $1 + 2 \rightarrow 3 + 4 + 5$ or $3 + 4 \rightarrow 1 + 2 + 5$, while region B describes the reaction $2 + 4 \rightarrow 1 + 3 + 5$ or $1 + 3 \rightarrow 2 + 4 + 5$. The convention used is that if particle n is incoming its 4-momentum is p_n while if it is outgoing its 4-momentum is $-p_n$. This can be seen from (6a) also.

Physical Regions in the s–w Plane

The form of $L = 0$ in the s–w plane, or equivalently in the $a_{13} - a_{24}$ plane, is more complicated than in the previous case: it is no longer a conic section. On the other hand, all principal minors of L appear as straight lines tangent to $L = 0$.

Now, as before, if $L_{23} \leq 0$ and $L_3 \geq 0$ then the physical regions will be given by $L \leq 0$. The implications of these conditions and the conditions necessary in order that the physical regions exist in the $a_{13} - a_{24}$ plane must be obtained. First of all it will be shown that all the lines $L_i = 0$ must occur in the plane.

Recall from the previous section that $L_{23} \leq 0$ implies that L_3 and L_4 can be nonnegative only if $L_{34} \leq 0$ and $L_{14} \leq 0$. Application of (9) to L_3 with $i = 2$, $j = 4$ then shows that $L_3 = 0$ occurs as two straight lines parallel to the a_{13} (or s) axis thus guaranteeing a region in the plane where L_3 is positive. Similarly, it is easy to show that $L_4 = 0$ occurs as two straight lines parallel to the a_{24} (or w) axis thus guaranteeing a region where L_4 is positive. Since, from (8), $L = 0$ occurs only where all L_i have the same sign, conditions for L_1 and L_2 to be nonnegative must found. Applying (8) to L_1 with $i = 2$, $j = 4$ yields

$$L_1 = -\{[V(1)_{24}]^2 - L_{12}L_{14}\}/L_{124},$$

where, from (7), $L_{124} = 2m_3^2$. Thus, L_1 can be nonnegative only if $L_{12} \leq 0$ (since $L_{14} \leq 0$). Furthermore, application of (9) to L_1 shows that $L_1 = 0$ occurs as two straight lines parallel to the a_{13} (or s) axis.

In a similar manner L_2 may be written

$$L_2 = -\{[V(2)_{13}]^2 - L_{12}L_{23}\}/L_{123},$$

where $L_{123} = 2m_4^2$. Thus L_2 can be nonnegative only if $L_{12} \leq 0$ (since $L_{23} \leq 0$). Also it is easy to show that $L_2 = 0$ occurs as two straight lines parallel to the a_{24} (or w) axis.

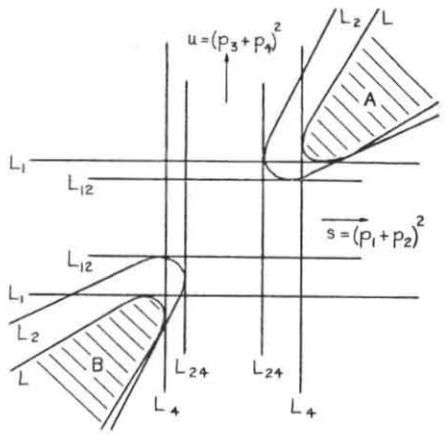

FIG. 1. Typical physical regions (shaded) for five-particle processes. A refers to $1 + 2 \rightarrow 3 + 4 + 5$ or $3 + 4 \rightarrow 1 + 2 + 5$ while B refers to $2 + 4 \rightarrow 1 + 3 + 5$ or $1 + 3 \rightarrow 2 + 4 + 5$.

Therefore, the physical regions can occur in the plane only if L_{12}, L_{23}, L_{14}, and $L_{34} \leq 0$ which also guarantee that all the curves $L_i = 0$ appear in the plane. In order that the physical regions actually do occur, there must in addition be a region or regions where all the L_i are nonnegative. This is assured if just L_1 and L_3 (or L_2 and L_4) are nonnegative. For if L_1 and L_3 are nonnegative in a region then application of (9) with $i = 1$, $j = 3$ gives a real value for a_{13} (or s) thus showing that $L = 0$ occurs there, which in turn implies that L_2 and L_4 are nonnegative in the region also. There does not appear to be any simple expression for these conditions involving only the parameters t, u, and v so they will be left as they stand.

Consequently, the necessary and sufficient conditions, i.e., restrictions on t, u, and v, in order that physical regions exist in the s–w plane, are L_{23}, L_{14}, L_{34}, $L_{12} \leq 0$ and L_1, $L_3 \geq 0$. These may be imposed as follows: (i) choose v such that $L_{23} \leq 0$, (ii) choose t such that L_{14}, $L_{34} \leq 0$, and (iii) choose u such that $L_{12} \leq 0$ and $L_1 \geq 0$ overlaps $L_3 \geq 0$ in the s–w plane. The physical regions will then be the regions $L \leq 0$.

It should be noted that in general many regions where L is negative will exist but not all will be physical regions (where all the L_i are positive). Furthermore the physical regions are bounded. For in the expansion of L_1 the coefficient of a_{24}^2 is $-L_{124} = -2m_3^2$ which shows that L_1 is positive between the lines $L_1 = 0$. Similarly the coefficient of a_{13}^2 in the expansion of L_2 is $-L_{123} = -2m_4^2$ which shows that L_2 is positive between the lines $L_2 = 0$. The region where L_1 and L_2 are positive is therefore bounded and the physical regions are contained in it.

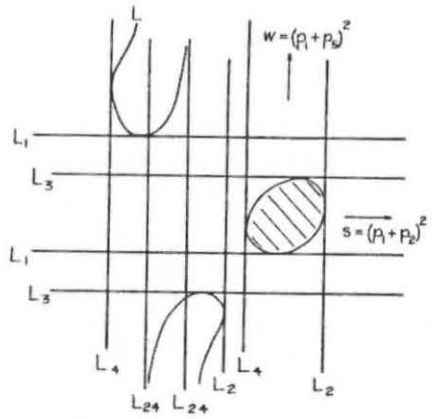

FIG. 2. A typical physical region (shaded) for the process $1 + 2 \to 3 + 4 + 5$ or $2 \to 1 + 3 + 4 + 5$.

A typical example of the configuration of curves to be found is shown in Fig. 2. There are three regions where L is negative, only one of which is the physical region (shaded). It describes the reaction $1 + 2 \to 3 + 4 + 5$ or $2 \to 1 + 3 + 4 + 5$.

IV. SUMMARY

A set of algebraic equations for the physical regions of N-particle relativistic processes has been derived in terms of kinematic invariants. The solution of the equations however was not attempted for N greater than five. In the case of $N = 5$ they were solved only by regarding them as functions of pairs of the five independent invariants treating the remaining three as parameters. The permissible ranges of these parameters were then established such that the physical regions actually can and do exist in the planes of the pairs of variables.

The author wishes to thank Professor L. F. Cook, Jr. for bringing this problem to his attention many years ago.

APPENDIX

Displayed here are some properties of determinants which proved useful in the analysis carried out in the text. Of particular importance is the Jacobi ratio theorem.[6]

Theorem: If (a_{ij}) is a square matrix with determinant

$$L \equiv \det (a_{ij}) \tag{A1}$$

then

$$[\text{adj } (a_{ij})]^{(k)} = L^{k-1} \times \text{adj}^{(k)} (a_{ij}),$$

where adj (a_{ij}) is the (adjugate) matrix obtained by replacing the elements of (a_{ij}) by their cofactors V_{ij} in L, $[\text{adj } (a_{ij})]^{(k)}$ is the kth compound[7] of adj (a_{ij}) and $\text{adj}^{(k)}$ (a_{ij}) is the kth adjugate compound[7] of (a_{ij}).

In the case $k=2$ it is easy to show that the diagonal elements of $\text{adj}^{(2)}$ (a_{ij}) are, in the notation of the text, simply L_{ij}. Noting that adj $(a_{ij}) \equiv (V_{ij})$ the diagonal elements of (A1) give

$$\begin{vmatrix} V_{ii} & V_{ij} \\ V_{ji} & V_{jj} \end{vmatrix} = LL_{ij},$$

or since $V_{ii} = L_i$ and (a_{ij}) of the text is symmetric

$$L_iL_j - L_{ij}L - V_{ij}^2 = 0. \tag{A2}$$

Another valuable equation results from an expansion of L due to Cauchy[8]

$$L = -\sum_{i,k\neq j} V(j)_{ik}a_{ij}a_{ik} + L_ja_{jj}. \tag{A3}$$

But by an expansion in terms of elements of the ith row (or column)

$$L = \sum_{i\neq j} V_{ij}a_{ij} + L_ja_{jj}.$$

Equating these two expressions for L yields

$$\begin{aligned} V_{ij} &= -\sum_{k\neq j} V(j)_{ik}a_{jk} \\ &= -\sum_{k\neq i,j} V(j)_{ik}a_{jk} - L_{ij}a_{ij} \\ &= V_{ij0} - L_{ij}a_{ij}, \end{aligned}$$

where V_{ij0} is V_{ij} evaluated at $a_{ij} = 0$.

This expression may be substituted into (A2) yielding

$$L = Aa_{ij}^2 + Ba_{ij} + C,$$

where

$$A = -L_{ij}, \quad B = 2V_{ij0}, \quad C = (L_iL_j - V_{ij0}^2)/L_{ij}.$$

Solving this quadratic equation gives

$$a_{ij} = [V_{ij0} \pm (L_iL_j)^{\frac{1}{2}}]/L_{ij} \equiv a_{ij}(\pm), \tag{A4}$$

which allows L to be written as

$$L = -L_{ij}[a_{ij} - a_{ij}(+)][a_{ij} - a_{ij}(-)]. \tag{A5}$$

In the text the asymptotic form of $L = 0$ in the $a_{13} - a_{34}$ plane is desired. This is simply obtained from (A3) by retaining only the leading terms in these variables:

$$L \sim -L_{13}a_{13}^2 - L_{34}a_{34}^2 - 2V(3)_{14}a_{13}a_{34}.$$

Assuming the coefficients to be nonvanishing, this can be equated to zero and solved, using $[V(3)_{14}]^2 - L_{13}L_{34} = -L_{134}L_3$, to obtain

$$a_{34} = a_{13}[-V(3)_{14} \pm (-L_{134}L_3)^{\frac{1}{2}}]/L_{34}, \tag{A6}$$

where $L_{134} = 2t$. Thus, with $L_3 \geq 0$, if t is negative, $L = 0$ is a hyperbola with asymptotes given by (A6), while, if t is positive, $L = 0$ is an ellipse.

[7] Reference 6, p. 90.

[8] Reference 6, p. 74.

Determination of Optical Field Correlations from Photon Counts

N. MUKUNDA*
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey
AND
T. F. JORDAN†
Department of Physics, University of Pittsburgh, Pittsburgh, Pennsylvania
(Received 1 October 1965)

It is shown that the off-diagonal matrix elements of an operator on an N-boson space are determined by diagonal matrix elements in different symmetrized product bases. The motivating application is that optical field correlations are determined by moments of the numbers of photons in different modes.

INTRODUCTION

HIGHER-order optical coherence has been defined[1] in terms of expectation values

$$\langle a_{i_1}^\dagger a_{i_2}^\dagger \cdots a_{i_N}^\dagger a_{i_{N+1}} a_{i_{N+2}} \cdots a_{i_{2N}} \rangle \tag{1}$$

of normal-ordered products of N annihilation operators a_i and N creation operators $a_i^\dagger$ for a set of modes i of the radiation field. An operational interpretation of this definition has been given[2] in terms of the moments

$$\langle b_{r_1}^\dagger(u) b_{r_1}(u) b_{r_2}^\dagger(u) b_{r_2}(u) \cdots b_{r_n}^\dagger(u) b_{r_n}(u) \rangle \tag{2}$$

of order $n \leq N$ of the probability distribution for the numbers of photons in different modes r which are related to the modes i by a unitary transformation $b_r(u) = \sum_i u_{ri} a_i$. The interpretation is equivalent to the definition if the expectation values (1) are determined when the moments (2) are known for various choices of the unitary matrix u which specifies the modes r. The moments (2) determine the expectation values

$$\langle b_{r_1}^\dagger(u) b_{r_2}^\dagger(u) \cdots b_{r_N}^\dagger(u) b_{r_1}(u) b_{r_2}(u) \cdots b_{r_N}(u) \rangle \tag{3}$$

of normal-ordered products of N annihilation operators b_r paired with the N creation operators $b_r^\dagger$ for the same modes. The theorem proved in this paper shows how the expectation values (1) are determined from the expectation values (3). A set of unitary matrices u is constructed to demonstrate this explicitly for a simple example.

* On leave of absence from Tata Institute for Fundamental Research, Bombay, India. This work was supported in part by the Summer Institute for Theoretical Physics at the University of Wisconsin, Madison, Wisconsin.
† Alfred P. Sloan Research Fellow. This work was supported by the U. S. Army Research Office, Durham, North Carolina.
[1] R. J. Glauber, Phys. Rev. 130, 2529 (1963).
[2] T. F. Jordan, Helv. Phys. Acta 37, 697 (1964).

THEOREM

Consider M operators a_i and their adjoints $a_i^\dagger$ satisfying the commutation relations

$$a_i a_j^\dagger - a_j^\dagger a_i = \delta_{ij},$$

$$a_i a_j - a_j a_i = 0,$$

for $i, j = 1, 2, \cdots, M$. For each $M \times M$ unitary matrix u let

$$b_r(u) = \sum_{i=1}^{M} u_{ri} a_i,$$

$$b_r^\dagger(u) = \sum_{i=1}^{M} u_{ri}^* a_i^\dagger$$

for $r = 1, 2, \cdots, M$. For any quantum-mechanical state (normalized, positive linear functional) of this system we prove the following.

Theorem: The $\{(N + M - 1)!/N!(M - 1)!\}^2$ expectation values

$$\langle a_{i_1}^\dagger a_{i_2}^\dagger \cdots a_{i_N}^\dagger a_{i_{N+1}} a_{i_{N+2}} \cdots a_{i_{2N}} \rangle$$

with the i's taking all values $1, 2, \cdots, M$, are determined when a suitably chosen expectation value

$$\langle b_{r_1}^\dagger(u) b_{r_2}^\dagger(u) \cdots b_{r_N}^\dagger(u) b_{r_1}(u) b_{r_2}(u) \cdots b_{r_N}(u) \rangle$$

of paired annihilation and creation operators is determined for each of $\{(N + M - 1)!/N!(M - 1)!\}^2$ suitably chosen unitary matrices u.

Let $i = 1, 2, \cdots, M$ label an orthonormal set of state vectors $|i\rangle$ for a single boson, and construct symmetrized product state vectors for N identical bosons as sums

$$|i_1, i_2, \cdots i_N\rangle = \sum_P |i_1\rangle |i_2\rangle \cdots |i_N\rangle \tag{4}$$

over all $N!$ permutations P of $i_1, i_2 \cdots, i_N$. Let

$$D(u)\,|i_1, i_2, \cdots, i_N\rangle = \sum_{i_1, i_2 \cdots i_N=1}^{M} u_{i_1 i_1} u_{i_2 i_2} \cdots u_{i_N i_N}\,|j_1, j_2, \cdots j_N\rangle$$

define a unitary operator $D(u)$ on the symmetric N-boson state vectors for each $M \times M$ unitary matrix u operating on the single-boson state vectors. Let T be an operator on the space of symmetric N-boson state vectors. By considering

$$\langle i_1, i_2, \cdots i_N|\,T\,|i_{N+1}, i_{N+2}, \cdots i_{2N}\rangle = \langle a^\dagger_{i_1} a^\dagger_{i_2} \cdots a^\dagger_{i_N} a_{i_{N+1}} a_{i_{N+2}} \cdots a_{i_{2N}}\rangle,$$

one can see that the above theorem is implied by the following.

Theorem: The $\{(N + M - 1)!/N!(M - 1)!\}^2$ matrix elements

$$\langle i_1, i_2, \cdots i_N|\,T\,|i_{N+1}, i_{N+2}, \cdots i_{2N}\rangle$$

with the i's taking all values $1, 2, \cdots, M$, are determined when a suitably chosen diagonal matrix element

$$\langle i_1, i_2, \cdots, i_N|\,D^\dagger(u) T\,D(u)\,|i_1, i_2, \cdots, i_N\rangle$$

is determined for each of

$$\{(N + M - 1)!/N!(M - 1)!\}^2$$

suitably chosen unitary matrices u.

We prove the theorem in the latter form. First we tidy up the notation. Let $\{i\}$ denote the set of indices $i_1, i_2, \cdots, i_N$. The state vectors (4) are not normalized to unity. Let $|\{i\}\rangle$ denote the same state vectors multiplied by positive numbers so that they are normalized to unity. They form an orthonormal set in the space of symmetric N-boson state vectors. Let summation over $\{i\}$ denote summation over this orthonormal set. Let

$$D_{\{j\}\{i\}}(u) = \langle\{j\}|\,D(u)\,|\{i\}\rangle$$

denote the matrix elements of the unitary operators $D(u)$. These unitary matrices form the irreducible unitary Nth-rank symmetric covariant tensor representation of the group U_M of $M \times M$ unitary matrices u.[3-5] Call this representation R.

Our goal is to determine the matrix elements

$$\langle\{j\}|\,T\,|\{k\}\rangle \tag{5}$$

from functions

$$\langle\{i\}|\,D^\dagger(u) T\,D(u)\,|\{i\}\rangle = \sum_{\{j\}\{k\}} D^*_{\{j\}\{i\}}(u) D_{\{k\}\{i\}}(u) \langle\{j\}|\,T\,|\{k\}\rangle \tag{6}$$

of the $M \times M$ unitary matrices u. The matrices $D^*_{\{j\}\{i\}}(u)$ form the irreducible unitary Nth-rank symmetric contravariant tensor representation of U_M which is the complex conjugate representation $\bar{R}$ of the representation R.[4] The product representation $\bar{R} \otimes R$ is reducible as a sum

$$D^*_{\{j\}\{i\}}(u) D_{\{k\}\{m\}}(u) = \sum_S \sum_a \sum_b C^{\bar{R}RS*}_{\{j\}\{k\}a} C^{\bar{R}RS}_{\{i\}\{m\}b} D^S_{ab}(u) \tag{7}$$

of irreducible unitary representations $D^S(u)$ of U_M labeled by S with $C^{\bar{R}RS}_{\{j\}\{k\}a}$ the Clebsch–Gordan coefficients for U_M.[4] Substituting (7) in (6) gives

$$\langle\{i\}|\,D^\dagger(u) T\,D(u)\,|\{i\}\rangle = \sum_{Sa} \sum_{\{j\}\{k\}} \langle\{j\}|\,T\,|\{k\}\rangle C^{\bar{R}RS*}_{\{j\}\{k\}a} \sum_b C^{\bar{R}RS}_{\{i\}\{i\}b} D^S_{ab}(u). \tag{8}$$

In the Appendix it is shown that $\bar{R} \otimes R$ is simply reducible; the irreducible representations labeled by S in the reduction (7) are all mutually inequivalent. Hence the unitary matrix elements $D^S_{ab}(u)$ in (7) satisfy the orthogonality relations[3,4]

$$\int du\, D^S_{ab}(u)^*\, D^{S'}_{a'b'}(u) = \delta_{SS'}\, \delta_{aa'}\, \delta_{bb'}$$

in terms of the invariant integral on the group U_M. It follows that the functions

$$\sum_b C^{\bar{R}RS}_{\{i\}\{i\}b} D^S_{ab}(u) \tag{9}$$

in (8) satisfy the orthogonality relations

$$\int du\, \{\sum_b C^{\bar{R}RS}_{\{i\}\{i\}b}\, D^S_{ab}(u)\}^* \{\sum_{b'} C^{\bar{R}RS'}_{\{i\}\{i\}b'}\, D^{S'}_{a'b'}(u)\} = \delta_{SS'}\, \delta_{aa'} \sum_b |C^{\bar{R}RS}_{\{i\}\{i\}b}|^2,$$

which show that, for fixed $\{i\}$, and for values of S such that

$$\sum_b |C^{\bar{R}RS}_{\{i\}\{i\}b}|^2 \tag{10}$$

is not zero, the functions (9) are linearly independent functions on U_M for different S and a. In the Appendix it is shown that for each S in the reduction (7) of $\bar{R} \otimes R$ there is at least one $\{i\}$ such that (10) is not zero.

Consider one S in the reduction (7), and choose $\{i\}$ such that (10) is not zero for that S. Now keep $\{i\}$ fixed and consider the set of all S for which (10)

[3] H. Weyl, *The Theory of Groups and Quantum Mechanics*, translated by H. P. Robertson (Dover Publications, Inc., New York, 1950).

[4] M. Hamermesh, *Group Theory* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1962).

[5] H. Boerner, *Representations of Groups* (North-Holland Publishing Company, Amsterdam, 1963).

is not zero. For these S the functions (9) are linearly independent for different S, a and span a linear space of functions on U_M which contains the function (8) for this fixed $\{i\}$. The dimension of this space is equal to the number of values of S, a with S in the set for which (10) is not zero. When the function (8) is determined for each of the same number of suitably chosen $M \times M$ unitary matrices u, the coefficients

$$\sum_{\{j\}\{k\}} \langle\{j\}|\, T\, |\{k\}\rangle C^{\bar{R}RS*}_{\{j\}\{k\}a} \tag{11}$$

of the linearly independent functions (9) are determined for this set of values of S and all of the associated values of a. If there is an S which was not included, then choose another $\{i\}$ such that (10) is not zero for this S and, keeping this $\{i\}$ fixed, consider the set of all S which were not included in the previous considerations and for which (10) is not zero. For these S and this fixed $\{i\}$, the functions (9) span a linear space of functions on U_M which contains the terms of the function (8) which were not determined previously. The dimension of this space is equal to the number of values of S, a corresponding to newly considered values of S. When the function (8) is determined for this number of suitably chosen unitary matrices u, the coefficients (11) of the linearly independent functions (9) are determined for these new values of S and all of the associated values of a. Repeat this procedure until the coefficients (11) are determined for all S, a. The total number of $M \times M$ unitary matrices u required is equal to the total number of values of S, a. This is the dimension of the product representation $\bar{R} \otimes R$ which is the same as the square of the dimension of the representation R which is (see Appendix)

$$\{(N + M - 1)!/N!\,(M - 1)!\}^2.$$

Multiplying (11) by $C^{\bar{R}RS}_{\{j\}\{k\}a}$ and summing over S, a, using the orthogonality property

$$\sum_{S,a} C^{\bar{R}RS*}_{\{j'\}\{k'\}a} C^{\bar{R}RS}_{\{j\}\{k\}a} = \delta_{\{j'\}\{j\}}\, \delta_{\{k'\}\{k\}}$$

of the Clebsch–Gordan coefficients,[4] yields the matrix elements (5). This completes the proof of the theorem.

EXAMPLE

There are many different suitable choices for the $M \times M$ unitary matrices u. One just needs to avoid the relatively few choices for which the linear independence of the functions $D^S_{ab}(u)$ fails.

For the case $M = 2$, $N = 2$, we exhibit explicitly nine 2×2 unitary matrices u and expectation values (3) which determine the nine expectation values (1). First we take

$$u = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix},$$

$$\begin{aligned}\langle b_1^\dagger(u) b_1^\dagger(u) b_1(u) b_1(u)\rangle &= \langle a_1^\dagger a_1^\dagger a_1 a_1\rangle \cos^4\theta + \langle a_2^\dagger a_2^\dagger a_2 a_2\rangle \sin^4\theta \\ &+ (4\langle a_1^\dagger a_2^\dagger a_1 a_2\rangle + \langle a_1^\dagger a_1^\dagger a_2 a_2\rangle + \langle a_2^\dagger a_2^\dagger a_1 a_1\rangle)\cos^2\theta\sin^2\theta \\ &+ 2(\langle a_1^\dagger a_1^\dagger a_1 a_2\rangle + \langle a_1^\dagger a_2^\dagger a_1 a_1\rangle)\cos^3\theta\sin\theta \\ &+ 2(\langle a_1^\dagger a_2^\dagger a_2 a_2\rangle + \langle a_2^\dagger a_2^\dagger a_1 a_2\rangle)\cos\theta\sin^3\theta\end{aligned}$$

for five different values of θ which allow us to determine the terms with different dependence on θ. Next we take

$$u = \begin{pmatrix} \cos\theta & i\sin\theta \\ i\sin\theta & \cos\theta \end{pmatrix},$$

$$\begin{aligned}\langle b_1^\dagger(u) b_1^\dagger(u) b_1(u) b_1(u)\rangle &= \langle a_1^\dagger a_1^\dagger a_1 a_1\rangle \cos^4\theta + \langle a_2^\dagger a_2^\dagger a_2 a_2\rangle \sin^4\theta \\ &+ (4\langle a_1^\dagger a_2^\dagger a_1 a_2\rangle - \langle a_1^\dagger a_1^\dagger a_2 a_2\rangle - \langle a_2^\dagger a_2^\dagger a_1 a_1\rangle)\cos^2\theta\sin^2\theta \\ &+ 2i(\langle a_1^\dagger a_1^\dagger a_1 a_2\rangle - \langle a_1^\dagger a_2^\dagger a_1 a_1\rangle)\cos^3\theta\sin\theta \\ &+ 2i(\langle a_1^\dagger a_2^\dagger a_2 a_2\rangle - \langle a_2^\dagger a_2^\dagger a_1 a_2\rangle)\cos\theta\sin^3\theta.\end{aligned}$$

Here three suitably chosen values of θ are sufficient to determine the terms with different dependence on θ, since the first two terms were determined previously. Finally we take

$$u = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & e^{i\pi/4} \\ 1 & -e^{i\pi/4} \end{pmatrix},$$

$$\begin{aligned}\langle b_1^\dagger(u) b_2^\dagger(u) b_1(u) b_2(u)\rangle &= \tfrac{1}{4}\langle a_1^\dagger a_1^\dagger a_1 a_1\rangle + \tfrac{1}{4}\langle a_2^\dagger a_2^\dagger a_2 a_2\rangle \\ &- \tfrac{1}{4}i(\langle a_1^\dagger a_1^\dagger a_2 a_2\rangle - \langle a_2^\dagger a_2^\dagger a_1 a_1\rangle),\end{aligned}$$

which gives us enough to determine all $\langle a^\dagger_{i_1} a^\dagger_{i_2} a_{i_3} a_{i_4}\rangle$ for $i_1, i_2, i_3, i_4 = 1, 2$. Altogether we have used nine 2×2 unitary matrices u.

ACKNOWLEDGMENTS

Our collaboration on this problem was made possible by the 1965 Summer Institute for Theoretical Physics at the University of Wisconsin, Madison, Wisconsin, where most of this work was done. We thank Professor K. W. McVoy and the Department of Physics of the University of Wisconsin for their hospitality.

APPENDIX

In order to prove the statements made in this paper regarding representations of unitary groups,

we recall some relevant facts concerning these groups. The group $U_M(SU_M)$ is the group of all $M \times M$ unitary (unimodular unitary) matrices u_{ij}. A covariant tensor of rank N transforms under either group as

$$T_{i_1 \cdots i_N} \to \sum_{i_1 \cdots i_N = 1}^{M} u_{i_1 i_1} \cdots u_{i_N i_N} T_{i_1 \cdots i_N}, \tag{A1}$$

while a contravariant tensor of rank N transforms as

$$T^{i_1 \cdots i_N} \to \sum_{i_1 \cdots i_N = 1}^{M} u^*_{i_1 i_1} \cdots u^*_{i_N i_N} T^{i_1 \cdots i_N}. \tag{A2}$$

A mixed tensor has both types of indices. A trace is taken by setting a covariant and contravariant index equal and summing from one to M.

Inequivalent irreducible unitary representations of U_M are generated from tensors as follows. If a tensor

$$T^{i_1 \cdots i_{N'}}_{i_1 \cdots i_N}$$

is covariant of rank N and contravariant of rank N', if it has a definite permutation symmetry for the covariant and contravariant indices separately, and if it is traceless, then it is an irreducible tensor. To obtain the representation in explicit unitary form, one has to choose appropriate orthonormal linear combinations of the tensor components.

For our purposes it is more convenient to consider the group SU_M for which there are two completely, antisymmetric invariant tensors $\epsilon_{i_1 \cdots i_M}$ and $\epsilon^{i_1 \cdots i_M}$ of covariant and contravariant ranks M, respectively, with

$$\epsilon_{1,2,\cdots M} = \epsilon^{1,2,\cdots M} = 1.$$

With these one can convert a contravariant index into $M - 1$ covariant indices, and vice versa. One then needs to consider only tensors of covariant type. Irreducible representations of SU_M are generated by covariant tensors whose indices possess definite permutation symmetry. The symmetry type can be specified by giving a set of M integers $f_1 \geq f_2 \geq \cdots f_M \geq 0$, which are the numbers of boxes in the rows of the Young diagram.[3-5] The rank of the tensor is $f_1 + f_2 \cdots + f_M$. The irreducible representation is denoted by $(f_1, f_2, \cdots f_M)$. Its dimension is[3-5]

$$d(f_1, f_2, \cdots f_M) = \frac{D(f_1 + M - 1, f_2 + M - 2, \cdots f_M)}{D(M - 1, M - 2, \cdots 1, 0)}, \tag{A3}$$

where

$$D(a_1, a_2, \cdots a_M) = \prod_{i<k=1}^{M} (a_i - a_k).$$

The irreducible representation of U_M given by symmetric covariant tensors of rank N, which we have called R, is the representation $(N, 0, \cdots, 0)$ of SU_M. The complex conjugate representation $\bar{R}$ is $(N, N, \cdots, N, 0)$. Their dimensions are

$$d(N, 0, \cdots, 0) = d(N, N, \cdots, N, 0) = (N + M - 1)!/N!\,(M - 1)!. \tag{A4}$$

Let $T_{i_1 \cdots i_N}$ and $V^{i_1 \cdots i_N}$ be two irreducible symmetric tensors corresponding to the representations R and $\bar{R}$ respectively. The product representation $\bar{R} \otimes R$ is given by the mixed (reducible) tensor

$$W^{i_1 \cdots i_N}_{i_1 \cdots i_N} = V^{i_1 \cdots i_N} T_{i_1 \cdots i_N}. \tag{A5}$$

By considering the mixed tensor

$$S^{i_1 \cdots i_S}_{i_1 \cdots i_S} = \sum_{k_1 \cdots k_{N-S} = 1}^{M} W^{k_1 \cdots k_{N-S} i_1 \cdots i_S}_{k_1 \cdots k_{N-S} i_1 \cdots i_S} \tag{A6}$$

obtained from (A5) by contracting $N - S$ pairs of indices, and making it traceless, we see that the representation $\bar{R} \otimes R$ contains, at least once, the irreducible representation corresponding to a mixed tensor of covariant and contravariant rank S, symmetric in both kinds of indices. Here S can take the values $0, 1, \cdots, N$. For different S these representations are inequivalent. Converting all contravariant indices in (A6) into covariant indices, we find that these are the representations $(2S, S, \cdots, S, 0)$ of SU_M. Their dimensions are

$$d(2S, S, \cdots S, 0) = \frac{(2S + M - 1)}{(M - 1)} \left\{ \frac{(S + M - 2)!}{S!\,(M - 2)!} \right\}^2. \tag{A7}$$

This can be rearranged to show that

$$d(2S, S, \cdots, S, 0) = d(S, 0, \cdots, 0)^2 - d(S - 1, 0, \cdots, 0)^2 \tag{A8}$$

for $S = 1, 2, \cdots, N$. Summing (A8) over S, we have that

$$d(N, 0, \cdots, 0)^2 = \sum_{S=0}^{N} d(2S, S, \cdots, S, 0). \tag{A9}$$

Since the left-hand side of (A9) is just the dimension of the representation $\bar{R} \otimes R$, we see that $\bar{R} \otimes R$ contains just once each of the irreducible representations corresponding to $S = 0, 1, \cdots, N$ and contains no other irreducible representations. This proves that $\bar{R} \otimes R$ is simply reducible.

Next we examine the Clebsch–Gordan coefficients

$$C^{\bar{R}RS}_{(i)(i)b}, \tag{A10}$$

which appear in (10). We denote the orthonormal basis vectors of the representations R and $\bar{R}$ by $|R, \{k\}\rangle$ and $|\bar{R}, \{i\}\rangle$ respectively. The product vectors

$$|\bar{R}, \{i\}\rangle\, |R, \{k\}\rangle \qquad \text{(A11)}$$

are an orthonormal basis for the representation $\bar{R} \otimes R$. Consider the subset of these for which $\{i\} = \{k\}$. Their expansions in terms of the orthonormal basis vectors $|S, b\rangle$ of the irreducible representations

$$|\bar{R}, \{i\}\rangle\, |R, \{i\}\rangle = \sum_{S,b} C^{\bar{R}RS}_{\{i\}\{i\}b}\, |S, b\rangle \qquad \text{(A12)}$$

define the coefficients (A10).

Consider elements of the group SU_M corresponding to $M \times M$ diagonal unitary matrices

$$u_{jk} = \delta_{jk} e^{ix_j}, \qquad \text{(A13)}$$

where x_j are real numbers such that

$$\sum_{j=1}^{M} x_j = 0 \qquad (\text{mod } 2\pi).$$

From (A1) and (A2) we see that under these transformations each component of a tensor gets multiplied by a phase factor. Of the product vectors (A11), only those for which $\{i\} = \{k\}$, namely the vectors (A12), are left invariant. We can show that, in every irreducible representation contained in $\bar{R} \otimes R$, there is at least one vector which is invariant under the phase transformations induced by the diagonal unitary matrices (A13). For, every such representation corresponds to a mixed tensor (A6) with the same number S of covariant and contravariant indices, and all components of the form

$$S^{i_1 \cdots i_S}_{i_1 \cdots i_S} \qquad \text{(no sum!)}$$

are invariant. In fact, the dimension of the subspace of invariant vectors is $(S + M - 2)!/S!(M - 2)!$.

This shows that the subspace of the representation $\bar{R} \otimes R$ spanned by the vectors (A12) has a nonempty intersection with the subspace of each irreducible representation. This means that for each irreducible representation S there is at least one $\{i\}$ and one b such that (A10) is not zero, or one $\{i\}$ such that (10) is not zero.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 7, NUMBER 5 MAY 1966

Ray Representations of Finite Nonunitary Groups

M. V. MURTHY*
Saha Institute of Nuclear Physics, Calcutta, India
(Received 15 November 1965)

An extension of the ray representation theory is formulated to embrace nonunitary groups. The coray representations are obtained by the ray representations of its unitary subgroup. Theorems of coray representations are stated. The usefulness of the formalism is discussed.

1. INTRODUCTION

INTEREST in the theory of ray representations of finite unitary groups has existed almost since their inception,[1] and persists to the present time,[2-4] and there have been numerous physical applications.[5-9] The introduction of time reversal as one of the basic symmetry elements resulted in the concept of nonunitary groups—a hybrid group which contains unitary, and antiunitary elements. Wigner's classic realization of such groups in terms of the theory of corepresentations is well

* Junior D.A.E. Fellow.

[1] J. Schur, J. Reine Angew. Math. **127**, 20 (1904); **132**, 85 (1907); **139**, 155 (1911).

[2] P. Rudra, J. Math. Phys. **6**, 1273 (1965).

[3] O. V. Kovalyev, *Irreducible Representations of Space Groups* (Kiev, 1961), in Russian.

[4] O. V. Kovalyev, and G. Ya. Lubarskii, Soviet Phys.—Tech. Phys. **28**, 1151 (1958).

[5] P. Rudra, J. Math. Phys. **6**, 1278 (1965).

[6] E. Brown, Phys. Rev. **133**, A1038 (1964).

[7] M. Lax, "Symmetry Principles in Solid State Physics" (Preprint).

[8] O. V. Kovalyev, Soviet Phys.—Solid State, **5**, 3157, 3164 (1963); Phys. Metals Metallog. (USSR) **17**, 490 (1964).

[9] J. Zak, Phys. Rev. **134**, A1602, A1607 (1964).

which appear in (10). We denote the orthonormal basis vectors of the representations R and $\bar{R}$ by $|R, \{k\}\rangle$ and $|\bar{R}, \{i\}\rangle$ respectively. The product vectors

$$|\bar{R}, \{i\}\rangle\, |R, \{k\}\rangle \qquad \text{(A11)}$$

are an orthonormal basis for the representation $\bar{R} \otimes R$. Consider the subset of these for which $\{i\} = \{k\}$. Their expansions in terms of the orthonormal basis vectors $|S, b\rangle$ of the irreducible representations

$$|\bar{R}, \{i\}\rangle\, |R, \{i\}\rangle = \sum_{S,b} C^{\bar{R}RS}_{\{i\}\{i\}b}\, |S, b\rangle \qquad \text{(A12)}$$

define the coefficients (A10).

Consider elements of the group SU_M corresponding to $M \times M$ diagonal unitary matrices

$$u_{jk} = \delta_{jk} e^{ix_j}, \qquad \text{(A13)}$$

where x_j are real numbers such that

$$\sum_{j=1}^{M} x_j = 0 \qquad (\text{mod } 2\pi).$$

From (A1) and (A2) we see that under these transformations each component of a tensor gets multiplied by a phase factor. Of the product vectors (A11), only those for which $\{i\} = \{k\}$, namely the vectors (A12), are left invariant. We can show that, in every irreducible representation contained in $\bar{R} \otimes R$, there is at least one vector which is invariant under the phase transformations induced by the diagonal unitary matrices (A13). For, every such representation corresponds to a mixed tensor (A6) with the same number S of covariant and contravariant indices, and all components of the form

$$S^{i_1 \cdots i_S}_{i_1 \cdots i_S} \qquad \text{(no sum!)}$$

are invariant. In fact, the dimension of the subspace of invariant vectors is $(S + M - 2)!/S!(M - 2)!$.

This shows that the subspace of the representation $\bar{R} \otimes R$ spanned by the vectors (A12) has a nonempty intersection with the subspace of each irreducible representation. This means that for each irreducible representation S there is at least one $\{i\}$ and one b such that (A10) is not zero, or one $\{i\}$ such that (10) is not zero.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 7, NUMBER 5 MAY 1966

Ray Representations of Finite Nonunitary Groups

M. V. MURTHY*
Saha Institute of Nuclear Physics, Calcutta, India
(Received 15 November 1965)

An extension of the ray representation theory is formulated to embrace nonunitary groups. The coray representations are obtained by the ray representations of its unitary subgroup. Theorems of coray representations are stated. The usefulness of the formalism is discussed.

1. INTRODUCTION

INTEREST in the theory of ray representations of finite unitary groups has existed almost since their inception,[1] and persists to the present time,[2-4] and there have been numerous physical applications.[5-9] The introduction of time reversal as one of the basic symmetry elements resulted in the concept of nonunitary groups—a hybrid group which contains unitary, and antiunitary elements. Wigner's classic realization of such groups in terms of the theory of corepresentations is well

* Junior D.A.E. Fellow.

[1] J. Schur, J. Reine Angew. Math. **127**, 20 (1904); **132**, 85 (1907); **139**, 155 (1911).

[2] P. Rudra, J. Math. Phys. **6**, 1273 (1965).

[3] O. V. Kovalyev, *Irreducible Representations of Space Groups* (Kiev, 1961), in Russian.

[4] O. V. Kovalyev, and G. Ya. Lubarskii, Soviet Phys.—Tech. Phys. **28**, 1151 (1958).

[5] P. Rudra, J. Math. Phys. **6**, 1278 (1965).

[6] E. Brown, Phys. Rev. **133**, A1038 (1964).

[7] M. Lax, "Symmetry Principles in Solid State Physics" (Preprint).

[8] O. V. Kovalyev, Soviet Phys.—Solid State, **5**, 3157, 3164 (1963); Phys. Metals Metallog. (USSR) **17**, 490 (1964).

[9] J. Zak, Phys. Rev. **134**, A1602, A1607 (1964).

known.[10-12] It is our purpose to present a coherent and useful formalism in which the basic ideas of ray representation theory are incorporated into nonunitary groups. Section 2 surveys the principal definitions and results. Section 3 presents the new formalism of coray representations.[13] Section 4 indicates a way to use them. Section 5 discusses the utility of the theory.

2. PRINCIPAL DEFINITIONS AND RESULTS

The state function of a physical system is represented more generally by a ray representating a direction in complex Hilbert space.[14] If Ψ is a vector corresponding to a physically realizable state, then Ψ and a constant multiple of Ψ represent one and the same state. However, when the states are normalized, a phase factor of modulus one remains undetermined, and the two vectors which differ by such a phase factor represent the same state. Now, if G is the symmetric group of the physical system, the operators $g \in G$ act on rays in Hilbert space only to map rays into rays (i.e., take from one possible state to another). Therefore we require that the matrices $D(g)$ realize the ray representation of the group G if they satisfy

$$D(g_1)D(g_2) = \omega_{g_1,g_2}D(g_3) \quad \text{for} \quad g_1g_2 = g_3, \tag{1}$$

where ω_{g_1,g_2} is a phase factor depending on g_1 and g_2 of modulus unity. If Φ and Ψ are state functions of a system invariant under the operations of G, the elements of the group g are unitary or antiunitary if they satisfy

$$(g\Psi, g\Phi) = (\Psi, \Phi) \tag{2}$$

or

$$(g\Psi, g\Phi) = (\Phi, \Psi) = (\Psi, \Phi)^*. \tag{3}$$

We denote u for unitary and a for antiunitary operators, respectively. If Φ is expanded in terms of the eigenstates of the system, the linear and the antilinear property of the unitary and antiunitary operators, respectively, becomes evident:

$$\Phi = \sum_\alpha a_\alpha \Phi_\alpha, \qquad u\Phi = \sum_\alpha a_\alpha u\phi_\alpha,$$
$$a\Phi = \sum_\alpha a_\alpha^* a\phi_\alpha. \tag{4}$$

Equation (1) provides the transformation rule for linear operators; its only new feature is the appearance of a unimodular phase factor. However, when the symmetric group under consideration contains some elements with the time reversal operator in combination with other geometrical operators, the structure of the group becomes nonunitary made up of unitary and antiunitary operators. Further, it is convienient to express the antiunitary operators as $a = V\theta$; V being unitary and θ being antiunitary time reversal. It is easily demonstrated that the product of two unitary, or two antiunitary ones is unitary, and the product of an antiunitary operator and a unitary is antiunitary with the consequence, that the nonunitary group contains equal number of unitary and antiunitary operators. The unitary operators form an invariant subgroup of index two.

Let Ψ_α^i be the αth basis function of the ith irreducible corepresentation of G. Following Wigner and adopting the summation convention, we write the corepresentation matrices as the following:

$$u\Psi_\alpha^i = D^i(u)_{\beta\alpha}\Psi_\beta^i,$$
$$a\Psi_\alpha^i = D^i(a)_{\beta\alpha}\Psi_\beta^i. \tag{5}$$

Using the above relation, one arrives at the following multiplicative structure of the corepresentations:

$$D^i(u_1)D^i(u_2) = D^i(u_1u_2),$$
$$D^i(u)D^i(a) = D^i(ua),$$
$$D^i(a)D^i(u)^* = D^i(au),$$
$$D^i(a_1)D^i(a_2)^* = D^i(a_1a_2), \tag{6}$$

and the equivalence of two corepresentations, by means of similarity transformation, is given by

$$B^i(u) = V^{-1}D^i(u)V,$$
$$B^i(a) = V^{-1}D^i(a)V^*, \tag{7}$$

where V is the transforming matrix.

3. DETERMINATION OF CORAY REPRESENTATIONS FROM THE RAY REPRESENTATIONS OF ITS UNITARY SUBGROUP

In view of the arguments put forward in Sec. 2, we assume that the matrices $D^i(g)(g \in G)$ realize the coray representation of the group G if

---

[10] E. P. Wigner, J. Math. Phys. **1**, 409, 414 (1960).

[11] E. P. Wigner, *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra* (Academic Press Inc., New York, 1959).

[12] J. O. Dimmock, J. Math. Phys. **4**, 1307 (1963).

[13] Wigner coined the word corepresentations for the vector representations of nonunitary groups. It is felt suggestive to use the terminology of coray representations (or projective corepresentations) to represent the ray representations of nonunitary groups.

[14] M. Hamermesh, *Group Theory and its Application to Physical Problems* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1962), Chap. 12, p. 458

$$\begin{aligned} D^i(u_1)D^i(u_2) &= \omega_{u_1,u_2}D^i(u_1u_2), \\ D^i(u)D^i(a) &= \omega_{u,a}D^i(ua), \\ D^i(a)D^i(u)^* &= \omega_{a,u}D^i(au), \\ D^i(a_1)D^i(a_2)^* &= \omega_{a_1,a_2}D^i(a_1a_2), \end{aligned} \tag{8}$$

where the factor system ω_{g_1,g_2} satisfies the following functional relationship according to the ray representation theory[15]:

$$\begin{aligned} \omega_{u,g_1}\omega_{ug_1,g_2} &= \omega_{u,g_1g_2}\omega_{g_1,g_2}, \\ \omega_{a,g_1}\omega^*_{ag_1,g_2} &= \omega_{a,g_1g_2}\omega^*_{g_1,g_2}. \end{aligned} \tag{9}$$

The vector representations of nonunitary groups are determined by the vector representations of its unitary subgroup. In an analogous way, the ray representations of nonunitary groups are obtained from the ray representations of its unitary subgroup obeying (8).

Avoiding repetition of arguments, we summarize the new results obtained. Closely following Wigner's analysis,[11] we denote by $\Delta^i(u)$ the irreducible ray representation of the unitary subgroup H. In fact, ω_{u_1,u_2} in Eq. (8) represents the factor system determining the irreducible ray representations of H. We denote by $\Delta^i(u)$ another ray representation of H,

$$\Delta^i(u) = \omega_{u,a_0}\omega^*_{a_0,a_0^{-1}ua_0}\Delta^{i*}(a_0^{-1}ua_0), \tag{10}$$

where a_0 is some antiunitary operator.

To determine the coray representation matrices $D^i(a)$, we begin with a set of functions forming a basis for the irreducible representation $\Delta^i(u)$

$$u\Psi^i_\alpha = \Delta^i(u)_{\beta\alpha}\Psi^i, \quad \text{for all} \quad u \in H. \tag{11}$$

The nonunitary group G is given in terms of H as

$$G = H + a_0H.$$

Accordingly, in the process of induction, the coray representations D^i of the group G are obtained from the ray representations Δ^i in one of the three following ways.

If for all $a_0 \in G$ the representations Δ^i and Δ^j are equivalent, then

$$\chi[\Delta^j(u)] = \chi[\Delta^i(u)]$$

and

$$\Delta^j(u) = \beta^{-1}\Delta^i(u)\beta,$$

which defines β, a unitary matrix satisfying

$$\beta\beta^* = \pm\omega_{a_0,a_0}\Delta^{(i)}(a_0^2). \tag{12}$$

[15] The elements g, u and a used in this relation correspond to any general element, the unitary, and the antiunitary element, respectively.

The positive and negative signs appearing in (12) correspond to two different cases.

Case (a):

If $\beta\beta^* = \omega_{a_0,a_0}\Delta^i(a_0^2)$.

Refering to (11), $a_0\Psi^i_\alpha$ reproduces the set of function Ψ^i_α. Hence the coray representation $D^i(g)$ of G corresponds to a single irreducible ray representation $\Delta^i(u)$ of H and is of the same dimentionality given by

$$D^i(u) = \Delta^i(u); \qquad D^i(a) = \omega_{aa_0^{-1},a_0}\Delta^i(aa_0^{-1})\beta. \tag{13}$$

Case (b):

If, on the other hand, $\beta\beta^* = -\omega_{a_0,a_0}\Delta^i(a_0^2)$, $a_0\Psi^i_\alpha$ gives rise to a set of functions ϕ^i_α, but, however, forms a basis for $\Delta^i(u)$ of H. $D^i(g)$ again corresponds to a single irreducible ray representation of twice the dimensionality

$$D^i(u) = \begin{pmatrix} \Delta^i(u) & 0 \\ 0 & \Delta^i(u) \end{pmatrix};$$

$$D^i(a) = \begin{pmatrix} 0 & \omega^*_{aa_0^{-1},a_0}\Delta^i(aa_0^{-1})\beta \\ -\omega^*_{aa_0^{-1},a_0}\Delta^i(aa_0^{-1})\beta & 0 \end{pmatrix}. \tag{14}$$

Case (c):

If for some antiunitary operator a_0 the representation $\Delta^i(u)$ and $\Delta^j(u)$ are nonequivalent, i.e.,

$$\chi[\Delta^i(u)] \neq \chi[\Delta^j(u)],$$

$a\Psi^i_\alpha$ gives rise to ϕ^j_α which acts as a basis for $\Delta^j(u)$ of H inequivalent to $\Delta^i(u)$. The coray representations are given by

$$D^i(u) = \begin{pmatrix} \Delta^i(u) & 0 \\ 0 & \Delta^j(u) \end{pmatrix},$$

$$D^i(a) = \begin{pmatrix} 0 & \omega_{a_0,a_0}\Delta^i(aa_0) \\ \omega^*_{a_0,a_0^{-1}a}\Delta^{j*}(a_0^{-1}a) & 0 \end{pmatrix}. \tag{15}$$

One could distinguish the three cases (a), (b), and (c) mentioned above by using the modified relation

$$m = 2/n \sum_l \chi^i(a_l^2)\omega_{a_l,a_l}, \tag{16}$$

where n is the order of the group, the factor $2/n$ represents the elements in H, and $\chi(a^2) = \chi(u)$ is the character of u.

In (16) m takes the three values 0, ± 1. The value equal to zero distinguishes case (c) from those

of (a) and (b), which are represented by $+1$ and -1, respectively, thereby serving as a useful rule.

The antiunitary operators can, however, be written as the product of unitary operators V with an antiunitary operator θ like the time reversal

$$\theta^2 = \epsilon I,$$

where I is the identity operator, and

$$\epsilon = (-1)^N,$$

N being the number of fermions in the system. Equation (16) could be transformed to represent the three cases as follows

$$\sum_i \chi[\Delta^i(u_i^2)]\omega_{u,u} = \epsilon n$$
$$= -\epsilon n$$
$$= 0. \qquad (17)$$

This result agrees well with the Frobenius–Schur criterion.[16]

The theorems leading up to the orthogonality relations for the irreducible representations of a group are here extended for coray representations.

Theorem 1: Any coray representation by nonsingular matrices is equivalent to one by unitary matrices. The proof of this theorem is identical to that for unitary groups,[11] and we only remark here that the coray representations will be unitary if the basis functions from an orthonormal set. Since any linearly independent set of functions can always be transformed into an orthonormal set, we can choose unitary matrices irrespective of whether the group in question is unitary or not.

Theorem 2: If there exists a Hermitian matrix M such that

$$MD^i(u) = D^i(u)M,$$
$$MD^i(a) = D^i(a)M^*,$$

for all the matrices of an irreducible coray representation, then M is a constant matrix, otherwise it is reducible. The first part of the proof is again identical to that given for unitary groups, and the second part of the proof differs from that given for unitary groups because of the appearence of the star, one could nevertheless proceed as in the case of corepresentations.[12]

Theorem 3: If D^i and D^j are two irreducible coray representations of the same group and of dimensionality l^i and l^j, respectively, and if there exists an $(l^i \times l^j)$ matrix X such that

$$XD^i(u) = D^j(u)X,$$
$$XD^i(a) = D^j(a)X^*,$$

then, for

$$l^i \neq l^j,$$
$$l^i = l^j$$

(X is a null matrix), the representations are equivalent. The proof is again identical with that of vector representations.[11]

Theorem 4: For two inequivalent irreducible unitary coray representations D^i and D^j

$$\sum_k D^i(u_k)_{\alpha\mu}\, D^j(u_k)_{\beta\nu} = 0,$$
$$\sum_k D^i(a_k)_{\alpha\mu}\, D^j(a_k)^*_{\beta\nu} = 0.$$

For the elements of the single unitary corepresentation, we have

$$\sum_k \{D^i(u_k)_{\alpha\mu}\, D^i(u_k)^*_{\beta\nu} + D^i(a_k)_{\alpha\nu}\, D^i(a_k)^*_{\beta\mu}\}$$
$$= n/l\ \delta_{\alpha\beta}\ \delta_{\mu\nu},$$

where n is the number of operators in G and l is the dimension of D^i. The theorem has been stated in such way as to incorporate the ray representation aspect of nonunitary groups. The proofs of orthogonality relation for irreducible ray representations[6] and corepresentations are well known[12] and this suffices as a proof for the above theorem.

4. DETERMINATION OF FACTOR SYSTEMS

The determination of the factor systems for a particular group depends on the physical problem in question. In crystallographic space groups, these factor systems arise as a result of the existence of nonprimitive translations, associated with screw axes and glide planes. The method of obtaining the representations of a space group based on the concept of the Brillouin zone is very well known.[17] The different inequivalent points $\mathbf{k}_i$ on the Brillouin zone correspond to different subgroups $G_{\mathbf{k}_i}$ furnishing an irreducible representation of the space group. Following a development due to Lubarskii[4] we obtain expressions for the factor systems of the coray representations of magnetic groups. If $\{\mathbf{t} \mid g\}$

[16] G. Frobenius and I. Schur, S. Sitzber. Deutsch. Akad. Wiss. 49, 136 (1906).

[17] G. F. Koster, Solid State Physics, 5, 174 (1957).

is a nonprimitive translation associated with point operation g, then

$$\omega_{u,g} = e^{i[k\cdot(ut-t)]}, \qquad \omega_{a,g} = e^{i[k\cdot(at+t)]}, \tag{18}$$

where α corresponds to the unitary operation in the antiunitary operators. It is easily seen that the second relation in (18) is due to the presence of antiunitary elements such as antiscrews and antiglides in the case of a magnetic group.[18] It is interesting to see that the ω's become unity when the exponent in (18) takes the value 0 or approaches 2π. The first case corresponds to a situation when vector $\mathbf{k}$ itself is zero corresponding to the special point Γ of the Brillouin zone. In the second case there exists no nonprimitive translation corresponging to any point operation of the group $G_{\mathbf{k}}$. Lastly, when the element in $G_{\mathbf{k}}$ does not change $\mathbf{k}$ to an equivalent point, i.e., if no two points represented by $\mathbf{k}$ are equivalent, then the ω's take the value unity. This does happen when $\mathbf{k}$ lies completely within the zone.

If the spin orbit interaction is involved in the physical system, the representations of the symmetric group of the system transform as the ray representations. Consider a nonunitary group, the unitary operation that enters into the antiunitary operator a has the form $u = Ka$, where K denotes the general operator of complex conjugation. Making use of one of the Pauli matrices σ_y, one could represent $u = -K\sigma_y a$, a case in which the spin orbit interaction is taken into account. Without the use of double groups it is often more convenient to consider as before a simple group with a different factor system. This factor system could be represented in the form of the product of the factor system in (18) and that corresponding part depending on the representations of the point group ω_{g_1,g_2}, determined as follows. Each element u is compared with one of the two matrices $D^{\frac{1}{2}}(u)$ and each element $a = K\sigma_y u$ with $D^{\frac{1}{2}}(a) = K\sigma_y D^{\frac{1}{2}}(u)$, then the factor system ω_{g_1,g_2} is obtained from the equations

$$D^{\frac{1}{2}}(g_1)\, D^{\frac{1}{2}}(g_2) = \omega_{g_1,g_2}\, D^{\frac{1}{2}}(g_1 g_2). \tag{19}$$

[18] M. Atoji, Am. J. Phys., 33, 212 (1965).

5. DISCUSSION

The formalism presented here was motivated by the problem of representation of nonsymmorphic magnetic groups, in which glide planes and screws axes and their corresponding antielements exist. The theory therefore finds wide applicability for magnetic structures answering all physical questions based on symmetry of time and space. Summarizing the results of earlier investigations, (which in fact follows as special cases) it gives a general method of representation of crystallographic groups (unitary, nonunitary, symmorphic, and nonsymmorphic). In fact, the factor systems are incorporated to take care of the nonsymmorphic aspect of unitary and nonunitary groups.

The method of double group often used to achieve a similar objective[19] or when spin orbit interaction is taken into account is rather artificial and cumbersome. One could avoid in this case the construction of a covering group by considering a much simpler group with a different factor system. An application of this method to the problem of phase transitions, and energy band structures in magnetic crystals will be illustrated in the forthcoming papers.

ACKNOWLEDGMENTS

I wish to thank Professor A. K. Saha and Professor D. K. Ray for their interest in this work and Mr. P. Rudra for several informal discussions on the theory of groups. I express my gratitude to the Department of Atomic Energy, India, for financial assistance.

[19] W. Döring, Z. Naturforsh. 14, 343 (1959).

The Implications of Unitarity on Possible Higher Symmetries*

I. J. MUZINICH
Department of Physics, University of Washington, Seattle, Washington
(Received 20 August 1965)

Using the covariant spinor formalism for higher-spin particles, it is shown that the only combination of the internal symmetry group and the Lorentz group that is compatible with the unitarity condition is the direct product. Three theorems are proved which severely limit the type of symmetries that are consistent with the unitarity condition.

I. INTRODUCTION

IT has often been emphasized that it is difficult to find an example of a model theory where SU(6) or related theories and the Poincaré group are simultaneously exact symmetries.[1] The clash between the Poincaré group and SU(6) lies in the fact that one cannot construct a kinetic energy term for a Lagrangian field theory which is invariant under SU(6).[1] This fact manifests itself in the following three related ways:

(a) The commutation or anticommutation relations between field operators.
(b) The unitarity relations between physical S matrix elements.
(c) The equations of motion of off-the-mass-shell Green's functions.

We would like to point out that, under a certain class of groups, invariance of the unitarity condition limits one to consider the direct product of the Lorentz group and the internal symmetry group, and some trivial extensions as a possible symmetry group. Several examples of the conflict of SU(6) or related theories and the unitarity condition have been given in the literature;[2] however, most of the demonstrations have been confined to some definite process or model. Therefore, we feel that it is useful to turn the argument around and give a general discussion to see what possible symmetries are allowed by the invariance of the unitarity condition.

In Sec. II we prove a set of results that greatly limits the type of symmetries one can consider for the unitarity condition. We assume the elastic unitarity condition and a certain class of semisimple Lie groups; we could of course also consider (a) or (c) for the proof. The amplitudes for higher spin[3] and the covariant spinor formalism are used. Finally, it is conjectured that, as far as strong interactions are concerned, the results found here and elsewhere do not make SU(6) or related theories useless. Merely the interpretation of SU(6) as an exact symmetry of a relativistic theory is not possible.

II. KINEMATIC PRELIMINARIES

We begin by giving some kinematic and group theoretic preliminaries. We consider the relativistic scattering amplitude to be given by the covariant spinor amplitude which has the following simple transformation properties under the homogeneous Lorentz group:

$$M_{\alpha\dot{\beta}}(P', P) = D^{(s,0)}(g)_{\alpha}^{\ \alpha'} D^{(0,s)}(g)_{\dot{\beta}}^{\ \dot{\beta}'} M_{\alpha'\dot{\beta}'}(\Lambda(g)P, \Lambda(g)P'). \quad (1)$$

Here P' and P refer respectively to the set of outgoing and incoming momenta; all conventions are identical to those of Ref. 4. The quantity g is an element of the group $SL(2, C)$ (covering group of the Lorentz group); $\Lambda(g)$ and $D(g)$ are, respectively, representations of the group $SL(2, C)$ on four vectors and spinors. The number of such matrices $D(g)$ in Eq. (1) will be determined by the number of particles with spin.

Under elements of the inhomogeneous Lorentz group (Poincaré group), the spinor amplitude transforms as follows.

$$M_{\alpha'\dot{\beta}'}(P', P) = \exp\left(i[P'(g) - P(g)]\cdot a\right) D^{(s,0)}(g)_{\alpha}^{\ \alpha'} \times D^{(0,s)}(g)_{\dot{\beta}}^{\ \dot{\beta}'} M_{\alpha'\dot{\beta}'}(P'(g), P(g)),$$

* Supported by the U. S. Atomic Energy Commission under contract AT (45-1)-1388, Program B.

[1] There are many places where this is pointed out. M. A. B. Beg and A Pais, Phys. Rev. **133**, B1514 (1965); *ibid.* **138**, B692 (1965); also K. Bardakci, J. M. Cornwall, P. Freund, and B. W. Lee. Phys. Rev. Letters **13**, 698 (1964). In fact S. Coleman in a Harvard University preprint has suggested that it is impossible to make a relativistic SU(6). For further literature on SU(6) see any of the above papers.

[2] M. A. B. Beg and A. Pais, Phys. Rev. Letters **14**, 509 (1965); J. M. Cornwall, P. G. O. Freund, and K. T. Mahanthappa, *ibid.* 515; and R. Blankenbeder, M. L. Goldberger, K. Johnson, and S. B. Treiman, *ibid.* 518.

[3] For discussions of the covariant spinor amplitudes see, for example, A. O. Barut, I. J. Muzinich, and David Williams, Phys. Rev. **130**, 442 (1963); S. Weinberg, Phys. Rev. **133**, B1318 (1964); and **134**, B882 (1964).

[4] A. O. Barut, I. J. Muzinich, and David Williams, Phys. Rev. **130**, 442 (1963).

where

$$P(g) = \Lambda(g)P; \qquad P'(g) = \Lambda(g)P'. \tag{2}$$

The representations of the Lorentz group on the spinor amplitudes can be split into spin and space parts of the form

$$D(g) \otimes \mathcal{P}. \tag{3}$$

If we consider the irreducible unitary representation on physical S matrix elements or states, not spinor amplitudes, the representations of the Lorentz group are of the form

$$D[u(g, p)] \otimes \mathcal{P}. \tag{4}$$

In Eqs. (3) and (4), $\mathcal{P}$ refers to the representation of the Lorentz group on spacetime, and $u(g, P)$ is an element of the little group[5] which, for fixed timelike momentum, p is isomorphic to the three-dimensional rotation group $SO(3)$.

The spinor amplitudes provide a natural framework to discuss the transformation properties of scattering amplitudes because g and $\mathcal{P}$ are completely decoupled.

The attempts to combine internal symmetry and the Lorentz group to provide a relativistically invariant SU(6) fall into one of the following catagories:[6]

$$\mathcal{G} = G \otimes \mathcal{P}, \tag{5}$$

where $\mathcal{G}$ is a semi-simple Lie group with the following properties[6,7]

$$\mathcal{G} \supset SL(2, C) \otimes I,$$

and $\mathcal{P}$ is the Poincaré group acting on space–time. Here I is the group of internal symmetry which we will assume to be compact.

$$\mathcal{G} = G \otimes T. \tag{6}$$

In Eq. (6) $\mathcal{G}$ is semi-simple[8] and also contains $SL(2, C) \otimes I$, and T is the translation group in a vector space of dimension larger than four. The entire Poincaré group has been redefined in a larger-dimensional space.[9] The latter possibility has great difficulties associated with it; namely, continuous mass spectra and the physical interpretation of the additional translation operators. We will not discuss the latter possibility here. The first possibility which has been attempted by Pais and Beg,[1,2] and by Strathdee, Delburgo, and Salam[7] will be considered.

The action of $\mathcal{G}$ [Eq. (5)] on the spinor amplitudes is given by

$$M(P', P) = M'(P', P) = D^{\lambda}(g)M(P', P)[D^{\lambda}(g)]^{\dagger}, \tag{7a}$$

where $g \in GM'(P', P)$ is the transformed M function and $D^{\lambda}(g)$ is a representation of $\mathcal{G}$. We will assume that $\mathcal{G}$ is a good symmetry and the M function is invariant. The M function transforms under $\mathcal{P}$ in the following manner:

$$M(P', P) = M(\Lambda P', \Lambda P), \tag{7b}$$

where $\Lambda \in \mathcal{P}$.

Of course this puts some severe restrictions on the form of the M function. The M function satisfying Eqs. (7) and (8) can be decomposed into a basis set of covariants for the group $\mathcal{G}$ in the form

$$M(P', P) = \sum a_i T_i, \tag{8}$$

where

$$D^{\lambda}(g)T_i[D^{\lambda}(g)]^{\dagger} = T_i$$

and the a_i are scalar amplitudes.

We will assume that the spinor amplitude describes the scattering of some pair of super multiplets, and that, in general, several different spins are in the same multiplet. Thus, if we restrict $\mathcal{G}$ to $SL(2, C) \otimes I$, the representation $D^{\lambda}(g)$, if not completely reducible, reduces as follows:

$$D^{\lambda}(g) = \oplus_i D^{(s_i, 0)}(t) \otimes D^{\mu_i}(u), \qquad t \in SL(2, C), \quad u \in I, \tag{9}$$

where λ, s_i, and μ_i, respectively, label the irreducible representations of $\mathcal{G}$, $SL(2, C)$ (Lorentz group), and the internal symmetry group I. The labels μ_i are a set of Casimir operators for the internal symmetry group.

The elastic unitarity condition for the scattering of supermultiplets can be written:

$$M(P', P) - M^{\dagger}(P', P) = \int d\Omega\,(K)M(P', K)A(K)M^{\dagger}(K, P). \tag{10}$$

This is a many-channel unitarity condition and we assume the symmetry is exact and the thresholds

[5] E. P. Wigner, Ann. Math., **40**, 149 (1939).

[6] See L. Michel, *Proceedings of the Second Coral Gables Conference on Symmetry Principles at High Energy* (to be published).

[7] A. Salam, R. Delburgo, and T. Strathdee, Proc. Roy. Soc. (London) **A284**, 146 (1965).

[8] T. Fulton and J. Wess, Phys. Letters **14**, 57, 344 (1965); H. Bacry and J. Nuyts, CERN preprint, and W. Ruhl, CERN preprints on the group SL(6, C) and other authors that were omitted in this reference.

[9] There is yet another approach by using the equal-time commutation relations of the currents with no intention of talking about exact symmetries. R. P. Feynmann, M. Gell-Mann, and G. Zweig, Phys. Rev. Letters **13**, 678 (1964); and R. Dashen (to be published).

are degenerate. The quantity $A(K)$ is given by

$$A(K) = A(k) = \oplus_i D^{(s_i,0)}(k\cdot\sigma/M) \otimes D^{\mu_i}(l), \quad (11)$$

where

$$k\cdot\tilde{\sigma}/M = k_\mu\sigma^\mu/M = (k_4\sigma_4 - \mathbf{k}\cdot\boldsymbol{\sigma})/M,$$

and l is the identity element of internal symmetry. Also in Eq. (10)

$$d\Omega\,(K) = [d^4k_1/(2\pi)^4][d^4k_2/(2\pi)^4] \times \delta(k_1^2 - M_1^2)\delta(k_2^2 - M_2^2). \quad (12)$$

If there is more than one supermultiplet, we merely take the direct product in Eq. (11)

$$A(K) = [\oplus_i D^{(s_i,0)}(k_1\cdot\tilde{\sigma}/M_1) \otimes D^{\mu_i}(l)] \otimes [\oplus_i D^{(s_i,0)}(k_2\cdot\tilde{\sigma}/M_2) \otimes D^{\mu_i}(l)]. \quad (13)$$

The matrices $D^{(s_i,0)}(k\cdot\tilde{\sigma}/M_2)$ are again representations of the homogeneous Lorentz group $SL(2, C)$, and are simply the numerators of the Feynman propagators for higher-spin particles.[3] Also in Eq. (11) $k^2 = M^2$. In Eqs. (12) and (13), k_1 and k_2 are the momenta of the intermediate particles.

If Eq. (1) is to be invariant under the transformation Eq. (7) for M independent of g, then

$$[D^\lambda(g)]^\dagger A(k) D^\lambda(g) = A(k), \quad (14)$$

where $g \in G$ and the transformation is linear. This is certainly not the most general condition on $A(K)$ that is possible. For example, we could consider a conformal group G that acts nonlinearly on the mass hyperboloid. In this case the unitarity condition would give us a condition of the form

$$[D(g)]^\dagger A(k) D(g) = J(k(g)/k) A(k(g)), \quad (15)$$

where $J[(k(g)/k)]$ is the Jacobian of the volume element on the mass hyperboloid [Eq. (12)], and $k(g)$ is the image of k under the transformation g. We hope to report on this case in the future.

If we concern ourselves with Eq. (14), where G is a semi-simple Lie group, we can prove the following theorems to rule out such a possibility. We define an invariant of a representation λ of some group G as a matrix $\mathbf{I}^\lambda(G)$ with the property in this context.

$$[D^\lambda(g)]^\dagger \mathbf{I}^\lambda(G) D^\lambda(g) = \mathbf{I}^\lambda(G). \quad (16)$$

For compact *Lie groups*[10] of finite dimensional unitary representations, $\mathbf{I}^\lambda(G)$ could be the unit matrix with the dimensionality $N(\lambda)$ of the representation λ. Thus Eq. (16) implies that we have an invariant Hermitian form on a vector space of dimension $N(\lambda)$.

Theorem: If there exists a group G which acts irreducibly on the unitarity condition in the following manner: $[D^\lambda(g)]^\dagger A(k)\ D^\lambda(g) = A(k)g \in G$, and $D^\lambda(g)$ is irreducible for each k independent of g, then all $A(k)$ are proportional to the invariant of the representation $\mathbf{I}^\lambda(G)$.

Proof.

$$[D^\lambda(g)]^\dagger A(k_1) D^\lambda(g) = A(k_1), \quad (17a)$$

$$[D^\lambda(g)]^\dagger A(k_2) D^\lambda(g) = A(k_2) \quad (17b)$$

imply

$$D^\lambda(g) A(k_2) A^{-1}(k_1) = A(k_2) A^{-1}(k_1) D^\lambda(g), \quad (18)$$

since

$$[D^\lambda(g)]^\dagger = A(k_1)[D^\lambda(G)]^{-1} A^{-1}(k_1). \quad (19)$$

From Schur's lemma[11] the only matrix that commutes with all elements of an irreducible representation is a multiple of the unit matrix. Therefore,

$$A(k_2) = aA(k_1),$$

where a is a complex number. If there exists an invariant of this representation $\mathbf{I}^\lambda(G)$, then by the same argument

$$A(k) = a'\mathbf{I}^\lambda(G). \quad (21)$$

This is clearly an undesirable feature for a physical theory; for example, if we choose G to be compact with finite dimensional unitary representations, then $A(k)$ would be proportional to the identity of the representation. Usually $A(k_0)$ is chosen to be the invariant, where $k_0 = (\mathbf{0}, M)$, and the unitary condition will be invariant at threshold.[2]

Theorem: If there exists a group G which acts on the unitarity condition in the same manner as the previous theorem, then (a) G does not contain the Lorentz group, $SL(2, C)$, (b) G is the little group of k and contains $SL(2, C)$, or (c) $A(k)$ is the direct sum of irreducible representations of one dimension[12] of $SL(2, C)$, i.e.,

$$A(k) = \oplus_i D^{(0,0)}(k\cdot\tilde{\sigma}/M) \otimes D^{\mu_i}(l). \quad (22)$$

Proof. Let $g \in SL(2, C)$ assume $G \supset SL(2, C)$. Then

$$D^\lambda(g) = \oplus_i D^{(s_i,0)}(g) \otimes D^{\mu_i}(l) \quad (23)$$

[10] H. Weyl, *Classical Groups*, Princeton Mathematical Series (Princeton University Press, Princeton, N. J., 1946) 2nd ed.

[11] See Ref. 10 or I. Schur, Sitzber. Preuss Akad. Wiss. Physik. Math. Kl., p. 406 (1905).

[12] I would like to thank Dr. C. Itzykson for an interesting conversation on this point.

and

$$[D^{\lambda}(g)]^{\dagger}A(k)D^{\lambda}(g)$$
$$= \oplus_i\, [D^{(s_i,0)}(g)]^{\dagger}D^{(s_i,0)}(k\cdot\tilde{\sigma}/M)D^{(s_i,0)}(g)$$
$$\times D^{\mu_i}(l) = \oplus_i\, D^{(s_i,0)}(g^{+}k\cdot\tilde{\sigma}g/M) \oplus D^{\mu_i}(l). \quad (24)$$

Since $SO(3, 1)$ is homomorphic to $SL(2, C)$ of the Lorentz group,

$$g^{+}k\cdot\tilde{\sigma}g/M = k(g^{-1})\cdot\tilde{\sigma}/M = \Lambda^{-1}(g)k\cdot\tilde{\sigma}/M; \quad (25)$$

Eq. (24) becomes

$$[D^{\lambda}(g)]^{\dagger}A(k)D^{\lambda}(g) = A(\Lambda^{-1}(g)k). \quad (26)$$

Equation (26) is possible if

$$\Lambda^{-1}(g)k = k, \quad (26')$$

or $\Lambda(g)$ is a member of the little group of k, hence $g \in$ is of the little group of k which we call $SL(2, C)_k$, which is isomorphic to the rotation group in three dimensions $SO(3)$ for $k^2 = m^2 > 0$.

In Eq. (24)

$$D^{(s_i,0)}(g^{+}k\cdot\tilde{\sigma}g/M) = D^{(s_i,0)}(k\cdot\tilde{\sigma}/M) \quad \text{if} \quad s_i = 0;$$

otherwise we obtain Eq. (26). Thus Eq. (14) is possible if

$$A(k) = \oplus_i\, D^{(0,0)}(k\cdot\tilde{\sigma}/M) \otimes D^{\mu_i}(l). \quad (27)$$

The only remaining possibility is that $g \notin SL(2, C)$ and $G \supset SL(2, C)$. If we demand that the unitarity condition Eq. (14) be invariant for all g independent of k, then possibility (b) is out. And if we wish our group G to contain $SL(2, C) \otimes I \supset SL(2, C)$ we have only trivial theory of spin 0 particles and the whole problem is vacuous.

Since we know that it is impossible for G to act irreducibly on the unitarity equation for a realistic theory, then we would like to show that $D^{\lambda}(G)$ is completely reducible.[10]

Theorem: If there exists a group G which acts on the unitarity condition in the following manner:

$$[D(g)]^{\dagger}A(k)D(g) = A(k), \quad (28)$$

where $A(k)$ is defined through Eq. (11); Then

$$D(g) = \oplus_i\, D^{(s_i,0)}(l) \otimes D^{\mu_i}(u), \quad (29)$$

where $u \in I$.

Proof. Let us divide the proof into two cases:

Case (1): All representations $D^{(s_i,0)}$ in $A(k)$ are inequivalent.[13]

[13] Here by inequivalent representations of $SL(2, C)$ we mean those with different spin. This is, of course, not the only criterion for inequivalence.

From the previous theorem, $D(g)$ is not irreducible for nontrivial $A(k)$; $D(g)$ is reducible, but not completely reducible. Let us try a $D(g)$ of the form

$$D(g) = \left[\begin{array}{c|c} D_1 & B \\ \hline 0 & D' \end{array}\right]. \quad (30)$$

Let us also write

$$A(k) = D^{(s_1,0)}(k\cdot\tilde{\sigma}/M) \otimes D^{\mu_1}(l)A'(k), \quad (31)$$

where

$$A'(k) = \bigoplus_{i\neq 1} D^{(s_i,0)}(k\cdot\tilde{\sigma}/M) \otimes D^{\mu_i}(l). \quad (32)$$

Equation (28) immediately leads us to

$$\left[\begin{array}{c|c} A_1 & 0 \\ \hline 0 & A' \end{array}\right] = \left[\begin{array}{c|c} D_1^{\dagger}\, A_1\, D_1 & D_1^{\dagger}\, A_1\, B \\ \hline B^{\dagger}\, A_1\, D & B^{\dagger}A_1B + D^{\dagger\prime}A'D' \end{array}\right], \quad (33)$$

where

$$A_1 = D^{(s_1,0)}(k\cdot\tilde{\sigma}/M) \otimes D^{\mu_1}(l).$$

In order that inequivalent representations of $SL(2, C)$ will not be mixed in Eq. (33), and in order that Eq. (33) be true for all k, we have $B = 0$, and

$$D(g) = \left[\begin{array}{c|c} D_1 & 0 \\ \hline 0 & D' \end{array}\right]. \quad (34)$$

Proceeding further and repeating the argument we can completely reduce $D(g)$.

Case (2): $A(k)$ contains equivalent $D^{(s_i,0)}$, then we can write

$$A(k) = \oplus_i\, D^{(s_i,0)}(k\cdot\tilde{\sigma}/M) \otimes \oplus_j\, D^{\mu_i(j)}(l), \quad (35)$$

where j labels the irreducible representations of internal symmetry going with the same i. From the previous argument we can reduce $D(g)$ partially with respect to $A(k)$ of above.

We have the result that $D(g) = \oplus_i D_i$, where

$$D_i^{\dagger}\{A^{(s_i,0)}(k\cdot\tilde{\sigma}/M) \otimes \oplus_k\, D^{\mu_i(k)}\}D_i$$
$$= \{A^{(s_i,0)}(k\cdot\tilde{\sigma}/M) \otimes \oplus_k\, D^{\mu_i(k)}\} \quad (36)$$

for each s_i.

From previous theorems, D_i cannot be irreducible or contain the Lorentz group, unless $s_i = 0$ and we have a trivial one-dimensional representation of $SL(2, C)$. Thus, if Eq. (36) is to be valid for each s_i, then

$$D_i = D^{\prime\prime}(l) \otimes D^{\mu_i}(u), \quad (37)$$

where

$$[D^{\mu_i}(u)]^\dagger \oplus_k D^{\mu_i(k)}(l) D^{\mu_i}(u) = \oplus_k D^{\mu_i(k)}(l). \quad (37')$$

The fact that we obtain a slight modification of the result [Eq. (29)], since not all of the $D^{(s_i,0)}$ $(k\cdot\vec{\sigma}/M)$ are inequivalent, is quite natural. This means that, if our supermultiplets contain distinct internal symmetry multiplets with the same spin, we have the possibility of extending the internal symmetry by combining different multiplets with the same spin. An example is given by the 35-dimensional representation of SU(6) and the φ ω mixing. The 35-dimensional representation has the SU(3) $\otimes$ SU(2) decomposition as is well known, 35 = (8, 3) $\oplus$ (8, 1) $\oplus$ (1, 3).

We have the desired result that the symmetry group of the unitarity [Eq. (14)] is the internal symmetry group. If we allow transformations that preserve the integration on the mass hyperboloid and act linearly on the hyperboloid (keeping it invariant), we obtain Lorentz transformations and the group $SL(2, C)$. Clearly the more interesting case is Eq. (15).

We have not considered the translation part of the Poincaré group. There are trivial extensions of the Poincaré group in which the translation part is imbedded in a larger Abelian group. Such trivial extensions were considered by S. Coleman who attributes them to P. Federbush. The trivial extensions are of the form of an Abelian Lie algebra A.

$$A \supset T, \quad (38)$$

where T is spanned by P_μ and A is spanned by Q_a. The operators P_μ and Q_a satisfy

$$[P_\mu, P_\nu] = 0, \quad [P_\mu, Q_a] = 0, \quad \text{and} \quad [Q_a, Q_b] = 0. \quad (39)$$

The Lie group generated by such a Lie algebra will certainly be unitary if it is a good symmetry of the S matrix. Also, in order to have Lorentz covariance, the Q_a must transform like irreducible tensors under representations of $SL(2, C)$.

III. SUMMARY AND SPECULATIONS

It should be clear that the interpretation of SU(6)-type theories as an exact symmetry of a relativistic theory with kinetic energy terms is a more difficult assignment than previously thought. However, in some approximate sense, SU(6) might still be useful in strong interactions. The example we have in mind is the hydrogen atom and the group $O(4)$, the levels of hydrogen for fixed n have the degeneracy of $O(4)$ (n here is the principal quantum number and one of the Casimir operators for $O(4)$). The hydrogen atom has been emphasized recently by Gell-Mann[14] as a testing ground for possible noncompact symmetry groups. The point here is that the symmetry is not obvious until the dynamical problem is solved. In strong-interaction physics it will be necessary to understand more deeply the interplay between dynamics and symmetry.

[14] M. Gell-Mann, Y. Ne'eman, and Y. Dothan (to be published).

Note on the Bondi–Metzner–Sachs Group*

E. T. NEWMAN
University of Pittsburgh, Pittsburgh, Pennsylvania
AND
R. PENROSE
Birkbeck College, University of London, London, England
(Received 28 October 1965)

It is shown that, in space-times which are asymptotically flat, there are reasonable physical restrictions that allow one to impose coordinate conditions (in addition to the usual Bondi–type conditions) which restrict the allowed coordinate group to a subgroup of the Bondi–Metzner–Sachs group. This subgroup is isomorphic to the improper orthochronous inhomogeneous Lorentz group.

I. INTRODUCTION

In recent years the subject of gravitational radiation has received a great deal of attention.[1–5] One of the unusual results of this study was that, even in space-times which are asymptotically Minkowskian, one apparently cannot extract the inhomogeneous Lorentz group as an asymptotic symmetry group if finite (retarded) time intervals, only, are considered. Instead, one obtains what is known as the Bondi–Metzner–Sachs group (BMS group)—an infinite parameter group. The purpose of this note is to show that if one imposes certain apparently reasonable physical restrictions at retarded time $u = -\infty$ (or, alternatively, at $u = +\infty$), it is possible to introduce further coordinate conditions such that a well-defined (non-normal) subgroup of the BMS group—isomorphic to the improper orthochronous inhomogeneous Lorentz group—is geometrically singled out by preserving these coordinate conditions.

The physical situations we allow would appear to include a general type of scattering problem in which the sources, and perhaps some waves, can come in from infinity and can again escape to infinity after interacting. This generalizes certain situations considered by Sachs. He observed[6] that the inhomogeneous Lorentz group could be singled out at, say, $u = +\infty$ if, for example, all the matter were radiated away as zero rest-mass energy, leaving Minkowski space for sufficiently large u; alternatively, the system might ultimately simply become static. The situations we consider here appear to be much more general than this, although whether or not they are quite as general as they seem to be, depends, to some extent, on the validity of some heuristic arguments we give, which have, as their basis, some exact results in the linear theory of gravitation.

As an additional purpose of this note, we introduce an invariant differential operation on the sphere (which we denote by $\eth$) and use it to define a type of "spin s spherical harmonic." This is done in slightly greater generality than is absolutely necessary for the present work, since it is also felt that these ideas should find applications elsewhere[7] and that they could be used to simplify earlier work.[2,4,6,8]

II. THE BMS GROUP

Let u, θ, ϕ, r be standard (Bondi-type) coordinates[1,2,4,6] for asymptotically flat space–time. Thus u is a *retarded time* parameter (so that $u =$ const are null hypersurfaces opening into the future); θ, ϕ are spherical polar coordinates for the sphere at infinity on each hypersurface $u =$ const (θ, ϕ, $u =$ const giving the null geodesic generators of these hypersurfaces); r is suitably defined radial coordinate such as an affine or luminosity parameter on each of the generators of the hypersurfaces. The BMS group is defined by the following transformations[2,4,6] on the θ, ϕ, u coordinates:

$$\begin{aligned} \theta' &= \theta'(\theta, \phi), \\ \phi' &= \phi'(\theta, \phi), \\ u' &= K(\theta, \phi)\{u - \alpha(\theta, \phi)\}, \end{aligned} \tag{2.1}$$

where $(\theta, \phi) \to (\theta', \phi')$ is a conformal transformation

* Work supported, in part, by the Aerospace Research Laboratory through the European Office of the U. S. Air Force (and, in part, by Contract AF–AFOSR–454–63).

[1] H. Bondi, M. G. J. van der Burg, and A. W. K. Metzner, Proc. Roy. Soc. (London) **A269**, 21 (1962).
[2] R. K. Sachs, Proc. Roy. Soc. (London) **A270**, 103 (1962).
[3] E. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962); **4**, 998 (1963).
[4] E. Newman and T. Unti, J. Math. Phys. **3**, 891 (1962).
[5] R. Penrose, Proc. Roy. Soc. (London) **A284**, 159 (1965).
[6] R. K. Sachs, Phys. Rev. **128**, 2851 (1962).
[7] E. Newman and R. Penrose, Phys. Rev. Letters **15**, 231 (1965).
[8] A. I. Janis and E. Newman, J. Math. Phys. **6**, 902 (1965).

of the (θ, ϕ)-sphere into itself, K is the corresponding conformal factor, given by

$$d\theta'^2 + \sin^2\theta'\, d\phi'^2 = K^2(d\theta^2 + \sin^2\theta\, d\phi^2), \qquad (2.2)$$

and where α is an arbitrary (suitably smooth) real function on the sphere. (The r coordinate may also be transformed, if desired, but the transformation is somewhat arbitrary since it depends on the precise type of radial coordinate used and it is not relevant to the structure of the group.)

The particular BMS transformations for which $\theta' = \theta, \phi' = \phi$ are called, in general, *supertranslations*. Under a supertranslation, the system of null hypersurfaces u = const. is transformed into a different system of null hypersurfaces (u' = const) but no (Lorentz) rotation is involved. We may write α in terms of spherical harmonics:

$$\alpha = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} a_{m,l} Y_{l,m}(\theta, \phi),$$

where the $a_{m,l}$ are constants. The *infinite set* of parameters $a_{m,l}$ [subject to $a_{-m,l} = (-1)^m \bar{a}_{m,l}$, so that α is real] then define the supertranslation. If $a_{m,l} = 0$ for $l > 2$, so that α takes the form:

$$\alpha = \epsilon_0 + \epsilon_1 \sin\theta\cos\phi + \epsilon_2 \sin\theta\sin\phi + \epsilon_3 \cos\theta, \qquad (2.3)$$

then the supertranslations reduce to a special case, called the *translations*, with just *four* parameters $\epsilon_0, \cdots, \epsilon_3$.

This terminology is, in fact, consistent with that for the (Lorentz) translations in Minkowski space. We may call a hypersurface u = const a "good" cone in Minkowski space if it is the null cone of some point, and a "bad" cone if, on the other hand, the generators of u = const do not all meet in a point. Thus, an actual translation in Minkowski space must send "good" cones into "good" cones; under a general supertranslation, the "good" cones will be warped into "bad" cones. It is precisely the condition that α be given by (2.3), which is required to preserve the "goodness" of the Minkowski null cones.

In curved asymptotically flat space–times the difficulty is to find an appropriate analog of the Minkowskian concept of "good" and "bad" cones. It is not, in fact, necessary to do this in order to single out the translations from the remaining supertranslations, since the translations are already determined by (2.3).[9] But if we wish, in addition, to isolate the "pure" Lorentz *rotations* from Lorentz rotations which have a "supertranslation component," then some concept of a distinction between "good" and "bad" cones is necessary. For we might try to define Lorentz rotations (homogeneous Lorentz transformation) as given by (2.1) with $\alpha = 0$. The hypersurface $u = 0$ is then transformed into itself. In Minkowski space, if $u = 0$ is a "good" cone, the resulting transformation indeed represents a Lorentz rotation and it sends other "good" cones into "good" cones. If, on the other hand $u = 0$ is a "bad" cone, then we do *not* get a Lorentz rotation in general. Thus, for asymptotically flat spaces, in order to know which of the BMS transformations are to be regarded as "supertranslation-free Lorentz rotations", we must have some definition of "goodness" of u = const hypersurfaces.

In Minkowski space, the "good" cones can be characterized locally by the fact that the null rays generating them possess no shear. In asymptotically Minkowskian spaces, it will only be the *asymptotic* behavior of the (complex) shear σ of these null rays that will concern us. With null vector l^μ tangent to the null rays and complex null vector m^μ orthogonal to l^μ satisfying $m^\mu \bar{m}_\mu = -1$, l^μ and m^μ being parallelly propagated along each ray, we have[3]

$$l_{\mu;\nu} m^\mu m^\nu = \sigma(u, \theta, \phi, r) = \sigma^0(u, \theta, \phi)/r^2 + O(r^{-4}), \qquad (2.4)$$

where r is scaled so that $r, {}_\mu l^\mu \to 1$ at infinity. Thus σ^0 defines the asymptotic shear of the hypersurface given, say, by u = const. The complex quantity σ^0 is of special interest in gravitational radiation theory. It forms part of the initial data on $u = 0$ used to determine the space-time asymptotically.[2,4] Furthermore, $\partial\sigma^0/\partial u$ and $\partial^2\sigma^0/\partial u^2$ both have physical significance. We may call $\partial^2\sigma^0/\partial u^2$ the *gravitational radiation field* since it represents the r^{-1} part of the Riemann curvature field.[1,2,4] Bondi *et al*[1] and Sachs[2] call $\partial\sigma^0/\partial u$ the "news function" since it can be used as asymptotic initial data for the gravitational radiation field and $|\partial\sigma^0/\partial u|^2$ represents the flux of energy of the gravitational radiation in their analysis.

We cannot, however, attempt to define "good" cones, in general, simply by requiring $\sigma^0 = 0$. In many cases it is simply not possible to arrange $\sigma^0 = 0$ for all values of θ, ϕ, on one hypersurface, but even in cases where it is possible (e.g. in the axially symmetric, reflection symmetric cases), it is clear from the above remarks that, in the presence of gravitational radiation, if $\sigma^0 = 0$ for one value of u, we will generally have $\sigma^0 \not\equiv 0$ for a

[9] In fact, Sachs has pointed out (Ref. 6) that the translation subgroup of the BMS group is uniquely singled out by group theoretic considerations, namely as the only four-parameter normal subgroup of the BMS group.

later value of u (i.e., "goodness" would not be invariant under time translation). The idea of the present paper is that, if we make apparently reasonable physical assumptions as to how the gravitational radiation falls off at $u = -\infty$ (or alternatively at $u = +\infty$), then we can effectively minimize σ^0 at $u = -\infty$ (or alternatively at $u = +\infty$). This will restrict our coordinates to such an extent that only a subgroup of the BMS group remains—which is isomorphic to the improper orthochronous inhomogeneous Lorentz group.

III. SPIN-WEIGHTED FUNCTIONS ON A SPHERE

In order to analyze the structure of σ^0 as a function of the angular coordinates θ, ϕ, it is important first to realize that, although it is given as a scalar, it is really a tensorlike quantity. The directions of minimum and maximum shear are determined by arg σ^0. Under rotation of the spacelike vectors Re (m^μ), Im (m^μ) in their plane given by

$$(m^\mu)' = e^{i\psi} m^\mu \tag{3.1}$$

[with $(l^\mu)' = l^\mu$, ψ real], we then have

$$(\sigma^0)' = e^{2i\psi}\sigma^0. \tag{3.2}$$

We shall say that σ^0 has spin weight 2. Generally, a quantity η will be said to have *spin weight* s if it transforms as

$$\eta' = e^{si\psi}\eta \tag{3.3}$$

under (3.1). (Here, s is in general integral, but half integral values can also occur.) The quantity σ^0 has, in addition, a conformal weight of -1. That is, under (2.1) and (2.2), if we choose $\alpha = 0$ and examine the σ^0 for $u = 0$ (keeping the m^μ vectors fixed), we find[2]

$$(\sigma^0)' = K^{-1}\sigma^0. \tag{3.4}$$

Generally, a quantity η defined on the (θ, ϕ)-sphere has *conformal weight* w if under conformal transformation of the sphere with conformal factor K as in (2.2) (and with fixed m^μ vectors) we have

$$\eta' = K^w \eta. \tag{3.5}$$

[K is, in effect, the relativistic Doppler factor[2] $(c + v)^{\frac{1}{2}}(c - v)^{-\frac{1}{2}}$.] For consistency with the coordinate conditions,[1,2,4] this conformal transformation should be accompanied by

$$(l^\mu)' = Kl^\mu, \qquad r' = K^{-1}r, \tag{3.6}$$

with $(m^\mu)' = m^\mu$, whence (3.4) [cf. (2.4)].

Effectively, the concepts of spin-weight and conformal weight refer to the behavior of functions on the (θ, ϕ)-sphere at infinity only, and do not refer to the remainder of the space–time. Indeed, the concepts will apply to *any* two-dimensional abstract surface, with a Riemannian or conformal structure. Quantities with spin weights correspond to irreducible tensor quantities on the surface. The vectors Re (m^μ), Im (m^μ) may be regarded as orthogonal tangent vectors (of length $2^{-\frac{1}{2}}$) at each point of the surface. But we shall be concerned, here, only with quantities defined on a *sphere*. If spherical polar coordinates are used, a natural choice[2,4,7] for m^μ is to make Re (m^μ) and Im (m^μ) tangential, respectively, to the curves ϕ = const and θ = const. Another convenient coordinate system for the sphere is $(\zeta, \bar{\zeta})$ where the complex parameter ζ is related to (θ, ϕ) by

$$\zeta = e^{i\phi}\cot\tfrac{1}{2}\theta. \tag{3.7}$$

In this case, the natural choice for m^μ is to make Re (m^μ) and Im (m^μ) tangential, respectively, to the curves Im (ζ) = const and Re (ζ) = const.

Let η be a quantity defined on the sphere of spin weight s. Define the operator $\eth$, in a particular (θ, ϕ) coordinate system, by

$$\eth\eta = -(\sin\theta)^s\left\{\frac{\partial}{\partial\theta} + \frac{i}{\sin\theta}\frac{\partial}{\partial\phi}\right\}\{(\sin\theta)^{-s}\,\eta\}. \tag{3.8}$$

(The operator $\eth$ is effectively a covariant differentiation operator in the surface.) Under a rotation of m^μ (but, for the moment, keeping the coordinates fixed) we demand that $\eth\eta$ behave as a quantity of spin weight $s + 1$. From (3.7) and the fact that the m^μ vectors are rotated through an angle $-\phi$ in the passage from (θ, ϕ) to $(\zeta, \bar{\zeta})$, we obtain

$$\eth\eta = 2P^{1-s}\,\partial(P^s\eta)/\partial\zeta \tag{3.9}$$

as the definition of $\eth$ in the $(\zeta, \bar{\zeta})$ system, where

$$P = \tfrac{1}{2}(1 + \zeta\bar{\zeta}). \tag{3.10}$$

[The coordinate ζ and the P used here are related to the ζ and P (now called ζ' and P') of Ref. 8 by $\zeta = -\frac{1}{2}\bar{\zeta}'$, $P = \frac{1}{2}\sqrt{2}P'$.]

Now, the first important property of the operator $\eth$ is that it is invariant (with spin weight unity) under change of coordinate system which preserves the sphere metric

$$ds^2 = d\theta^2 + \sin^2\theta\, d\phi^2 = P^{-2}\, d\zeta\, d\bar{\zeta}. \tag{3.11}$$

This is most easily seen in the $(\zeta, \bar{\zeta})$ system and the result for the (θ, ϕ) system will then follow. With ψ as in (3.1), so that $\eta' = e^{si\psi}\eta$, we have

$$e^{-2i\psi} = \frac{\partial\zeta'/\partial\zeta}{\partial\bar{\zeta}'/\partial\bar{\zeta}}, \tag{3.12}$$

where $(\zeta', \bar{\zeta}')$ is the new coordinate system, with new complex tangent vector $m^{\mu'}$. Since $ds' = ds$, we have

$$\partial\zeta'/\partial\zeta = e^{-i\psi} P'/P \tag{3.13}$$

with $\partial\zeta'/\partial\bar{\zeta} = 0$. From this it follows that $\eth'\eta' \equiv 2P'^{1-s}\partial(P'^{s}\eta')/\partial\zeta' = e^{(s+1)i\psi}\,\eth\eta$, as required. [In fact, this argument applies equally to any surface, with metric written $ds^2 = P^{-2}\,d\zeta\,d\bar{\zeta}$, where P need not have the form (3.10).]

The second important property of $\eth$ concerns its behavior under a conformal transformation. Let η have conformal weight w and spin weight s (both integers or both half odd integers) where $w \geq s$. Then,

$$\eth^{w-s+1}\eta \text{ } \textit{is a quantity of conformal weight } s-1, \tag{3.14}$$

and, of course, spin weight $w + 1$. To prove this, we need only consider conformal transformations of the form

$$\zeta' = a\zeta, \tag{3.15}$$

a being a constant, with $(P')^{-2}\,d\zeta'\,d\bar{\zeta}' = K^2P^{-2}\,d\zeta\,d\bar{\zeta}$ so that

$$K = aP/P' = (1 + \zeta\bar{\zeta})/(a^{-1} + a\zeta\bar{\zeta}). \tag{3.16}$$

[The general conformal self-transformation of the sphere, excluding reflections can be built up from transformations of the type (3.15)—corresponding to a Lorentz velocity transformation in the direction of the z axis—and rotations which preserve the metric. These rotations have been already dealt with, since they correspond to metric preserving coordinate transformations.] It is required to prove that, under (3.15), $(\eth^{w-s+1}\eta)' = K^{s-1}\,\eth^{w-s+1}\eta$. [The m^{μ} vectors are unaltered under (3.15).] This can be done by a kind of inductive argument which is somewhat tedious for general values of $w - s$. For small value of $w - s$ it is not hard to verify directly that

$$(P')^{1-w}\frac{\partial}{\partial\zeta'}\left((P')^2\frac{\partial}{\partial\zeta'}\left(\cdots\frac{\partial}{\partial\zeta'}((P')^sK^w\eta)\cdots\right)\right) = K^{s-1}P^{1-w}\frac{\partial}{\partial\zeta}(\cdots(P^s\eta)\cdots),$$

where from (3.16) we use the fact that $P^2\,\partial K^{-1}/\partial\zeta$ is a constant. It is a curious fact that, given η with $w \geq s$, there is just the *one* power of $\eth$ given in (3.14) which yields a quantity correctly transforming (i.e., with a "weight") under a conformal transformation. If $w \geq -s$, then $\bar{\eth}^{w+s+1}\eta$ also correctly transforms with conformal weight $-s - 1$. Here $\bar{\eth}$ is defined [in the $(\zeta, \bar{\zeta})$ system] by

$$\bar{\eth}\eta = 2P^{1+s}\,\partial(P^{-s}\eta)/\partial\bar{\zeta}. \tag{3.17}$$

If $w > |s|$, then we can apply both operators to get a quantity with conformal weight $-w - 2$:

$$\bar{\eth}^{w+s+1}\,\eth^{w-s+1}\eta = \eth^{w-s+1}\,\bar{\eth}^{w+s+1}\eta. \tag{3.18}$$

That these two quantities are equal is not immediately obvious since the operators $\eth$ and $\bar{\eth}$ do not in general commute. In fact, we have

$$(\bar{\eth}\,\eth - \eth\,\bar{\eth})\eta = 2s\eta. \tag{3.19}$$

However, it is a consequence of the following considerations that $\eth^p\bar{\eth}^q\eta = \bar{\eth}^q\eth^p\eta$ whenever $q - p = 2s$.

The third important property of $\eth$ concerns its effect on spherical harmonics. Let $Y_{l,m}(l = 0, 1, \cdots; m = -l, \cdots, +l)$ be the usual scalar spherical harmonics. Then we can define the *spin s spherical harmonics*[7,8] (for integral spin) as follows:

$${}_sY_{l,m} = \begin{cases} \left[\dfrac{(l-s)!}{(l+s)!}\right]^{\frac{1}{2}}\eth^sY_{l,m} & (0 \leq s \leq l), \\ (-1)^s\left[\dfrac{(l+s)!}{(l-s)!}\right]^{\frac{1}{2}}\bar{\eth}^{-s}Y_{l,m} & (-l \leq s \leq 0). \end{cases} \tag{3.20}$$

The ${}_sY_{l,m}$ are not defined for $|s| > l$. In the $(\zeta, \bar{\zeta})$ system, the spin s harmonics take the form

$${}_sY_{l,m} = \frac{a_{l,m}}{[(l-s)!\,(l+s)!]^{\frac{1}{2}}}(1 + \zeta\bar{\zeta})^{-l} \times \sum_p \zeta^p(-\bar{\zeta})^{p+s-m}\binom{l-s}{p}\binom{l+s}{p+s-m}, \tag{3.21}$$

summed over integral values of p, the $a_{l,m}$ being numerical constants whose exact values are inessential for our purposes. [In fact, (3.21) applies also to "spinor harmonics" for which l, m, and s are all half odd integers.] We have ${}_s\bar{Y}_{l,m} = (-1)^{m+s}\,{}_{-s}Y_{l,-m}$ and, for *all* s with $|s| \leq l$,

$$\begin{aligned} \eth({}_sY_{l,m}) &= [(l-s)(l+s+1)]^{\frac{1}{2}}\,{}_{s+1}Y_{l,m}, \\ \bar{\eth}({}_sY_{l,m}) &= -[(l+s)(l-s+1)]^{\frac{1}{2}}\,{}_{s-1}Y_{l,m}. \end{aligned} \tag{3.22}$$

In particular, $\eth$ annihilates ${}_sY_{l,m}$ whenever $l = s$ and $\bar{\eth}$ annihilates ${}_sY_{l,m}$ if $l = -s$. We see, furthermore, that $\eth^p$ annihilates any quantity of spin weight s which is composed only of harmonics with $l < s + p$.

The ${}_sY_{l,m}$ are eigenfunctions of the operator $\bar{\eth}\eth$ for each spin weight s:

$$\bar{\eth}\eth({}_sY_{l,m}) = -(l-s)(l+s+1)\,{}_sY_{l,m}. \tag{3.23}$$

(If $s = 0$, $\bar{\eth}\eth$ is essentially the total angular momentum operator.) More generally,

$$\bar{\eth}^p \eth^p({}_sY_{l,m}) = (-1)^p \frac{(l-s)!}{(l-s-p)!} \frac{(l+s+p)!}{(l+s)!} {}_sY_{l,m}. \tag{3.24}$$

For each given s, the ${}_sY_{l,m}$ form a set of orthonormal functions of spin weight s on the sphere:

$$\int {}_sY_{l,m}\, {}_s\bar{Y}_{l',m'}\, dS = \delta_{ll'}\, \delta_{mm'}, \tag{3.25}$$

dS being the surface area element on the sphere. This can be proved (s being an integer) by induction on s using (3.22), the orthonormality of $Y_{l,m}$, and the fact that

$$\int (\eth A)B\, dS = -\int A(\eth B)\, dS \tag{3.26}$$

with (spin weight of A) + (spin weight of B) = -1. [It is immediate from (3.8) that $\int \eth C\, dS = 0$ (where C has spin weight -1) and (3.26) follows.] The orthonormal functions ${}_sY_{l,m}$ are also *complete*, for spin weight s quantities on the sphere. That is, if η has spin weight s and is suitably regular[10] on the sphere, there exist constants $\eta_{l,m}$ such that

$$\eta = \sum_{l,m} \eta_{l,m}\, {}_sY_{l,m}. \tag{3.27}$$

To see this, without loss of generality choose $s > 0$ and consider $\bar{\eth}^s\eta$. This has spin weight zero, so from the completeness of the $Y_{l,m}$, we have, for some constants $\phi_{l,m}$

$$\bar{\eth}^s\eta = \sum_{l,m} \phi_{l,m} Y_{l,m}. \tag{3.28}$$

By (3.24), we can write this

$$\bar{\eth}^s\{\eta - \sum_{\substack{l,m \\ l\geq s}} \chi_{l,m}\, \eth^s Y_{l,m}\} = \sum_{\substack{l,m \\ l<s}} \phi_{l,m} Y_{l,m}, \tag{3.29}$$

where

$$\chi_{l,m} = \frac{(l-s)!}{(l+s)!}(-1)^s \phi_{l,m} (l \geq s).$$

Moreover, the right-hand side of (3.29) must vanish since, if we multiply (3.29) by $\bar{Y}_{l',m}(l' < s)$ and integrate over the sphere, we get $\phi_{l',m}$ on the right and $\int \bar{Y}_{l'm}\, \bar{\eth}^s\{\cdots\}\, dS$ on the left. Applying (3.26) and $\eth^s Y_{l'm} = 0 (l' < s)$, we see that this vanishes. Thus, all the ϕ's (with $l' < s$) vanish, so the left-hand side of (3.29) also vanishes. Now, if μ is any suitably regular quantity on the sphere, of spin weight s' where $s' > 0$, then

$$\bar{\eth}\mu = 0 \quad \textit{implies} \quad \mu = 0. \tag{3.30}$$

For, by (3.17), $\bar{\eth}\mu = 0$ implies that $P^{-s'}\mu$ is an analytic function of ζ. Hence $\lim_{|\zeta|\to\infty} |P^{-s'}\mu| > 0$ unless $\mu = 0$. But $s' > 0$, so that $P^{-s'} \to 0$ and μ must be bounded for all ζ. Therefore $\mu = 0$. Repeated application of (3.30) to (3.29) now gives (3.27) as required, with

$$\eta_{l,m} = \{(l+s)!/(l-s)!\}^{\frac{1}{2}} \chi_{l,m}.$$

Another way of expressing this result is: given any suitably regular η on the sphere,[10] of (integral) spin weight $s > 0$, there exists ξ of spin weight zero for which

$$\eta = \eth^s\xi. \tag{3.31}$$

The relevance to this paper of the foregoing analysis lies essentially in this result.[11] (In the above, $\xi = \sum \chi_{l,m} Y_{l,m}$.) One reason this will be of interest to us here is that it enables us to define the "electric" and "magnetic" parts of η, denoted, respectively, by η_e and η_m, where

$$\eta_e = \eth^s \operatorname{Re}(\xi), \qquad \eta_m = \eth^s \operatorname{Im}(\xi). \tag{3.32}$$

[If, instead, $s < 0$, then we would use $\bar{\eth}$ in place of $\eth$ and $-s$ in place of s in (3.31), (3.32).] We would get the same η_e and η_m if a different choice of ξ were made in (3.31), since the arbitrariness in ξ lies only in harmonics $Y_{l,m}$ with $l < s$. We can characterize η_e and η_m by the fact that

$$\eta = \eta_e + \eta_m \tag{3.33}$$

with

$$\bar{\eth}^s\eta_e = \eth^s\bar{\eta}_e, \qquad \bar{\eth}^s\eta_m = -\eth^s\bar{\eta}_m. \tag{3.34}$$

The invariance properties of $\eth$ imply that the concepts of "magnetic" and "electric" parts of η are suitably invariant under rotations of the sphere or of the m^μ vectors. Furthermore if (and *only* if) η has conformal weight -1, the splitting of η into its "electric" and "magnetic" parts is also invariant under *conformal* transformation of the sphere [cf. (3.14)].

Here we shall apply these concepts to σ^0. The spin weight is 2 and the conformal weight is -1, so on each hypersurface u = const, we have a splitting $\sigma^0 = \sigma^0_e + \sigma^0_m$, which is invariant under conformal transformations of the (θ, ϕ)-sphere, i.e., under the "Lorentz" transformations (2.1) which leave the hypersurface invariant. The σ^0 and its u-derivatives describe properties of the gravitational

[10] We use the phrase "suitably regular" in the sense that η is a smooth function of ζ and $\bar{\zeta}$ in the $(\zeta, \bar{\zeta})$ system and that if we transform coordinates by, say, $\zeta' = \zeta^{-1}$, then η' is a smooth function of ζ' and $\bar{\zeta}'$. This deals with the point $\zeta = \infty$ in the $(\zeta, \bar{\zeta})$ system.

[11] This result, in the case $\eta = \sigma^0$, is considered by D. Lamb, J. Math. Phys. 7, 458 (1966).

radiation field. By analogy with the linear theory, we may suppose μ_e^0 to be associated with "electric"-type radiation (e.g., arising from changes in the mass quadrupole) and σ_m^0 to be associated with "magnetic"-type radiation (e.g., arising from changes in the angular momentum quadrupole). The analogy with electrodynamics arising here is the reason for the use of the terms "electric" and "magnetic" in (3.32). In the next section, we consider these matters more explicitly.

IV. PHYSICAL ASSUMPTIONS

We wish to impose physical restrictions on the systems under consideration, which have sufficient generality to include scattering problems in which sources can come in from and go out to spatial infinity with initial and final velocities less than c. In particular, the sources might all be confined to a bounded region of space. We shall require that the angular momentum, with respect to some origin, of each component part of the source remains bounded for all time (where it is assumed that there are a finite number of "parts" to the source). Generally, the physical restrictions that we impose are intended to be such that, $u \to -\infty$ (or alternatively as $u \to +\infty$), σ^0 behaves as

$$\sigma^0(u, \theta, \phi) \to S(\theta, \phi), \tag{4.1}$$

with S independent of u. If the analogy with the linear theory is to be trusted, we would normally expect S to have no "magnetic" part. This would add some force to the arguments but it is not actually essential. On the basis of (4.1) alone, we shall be able to extract the inhomogeneous Lorentz group.

In order to investigate what is entailed physically by the requirement that (4.1) hold, we must resort to analogies with the linear theory. The full theory is, unfortunately, insufficiently developed as yet to enable us to infer the exact relation between the motion of the sources and asymptotic quantities such as σ^0. In the linear theory, we may take $u =$ const to be "good" cones in Minkowski space. Since the shear of these cones now vanishes, the linearized σ^0, denoted σ^0_{lin}, describes a first-order deviation from "goodness" of these cones. We can then obtain σ^0_{lin} from the multipole moments $C_{l,m}(u)$ of the source distribution by[6,7]

$$\sigma^0_{\text{lin}} = \eth^2\left\{\sum_{l,m} \beta_{l,m} \frac{d^l}{du^l}\left(C_{l,m} Y_{l,m}\right) + f\right\}, \tag{4.2}$$

where $f = f(\theta, \phi)$ is real, independent of u, and corresponds to the gauge freedom in σ^0, and where the $\beta_{l,m}$ are some inessential numerics. (Both $\beta_{l,m}$ and $C_{l,m}$ have spin weight zero.)

The type of behavior for the multipole moments that we may assume as reasonable is

$$C^{\text{e}}_{l,m} \sim u^l, \tag{4.3}$$

$$C^{\text{m}}_{l,m} \sim u^{l-1}, \tag{4.4}$$

where $C^{\text{e}}_{l,m}$ and $C^{\text{m}}_{l,m}$ correspond, respectively, to the "electric" and "magnetic" parts of the multipole moments, so that $C^{\text{e}}_{l,m}Y_{l,m}$ and $C^{\text{m}}_{l,m}Y_{l,m}$ are the real and imaginary parts of $C_{l,m}Y_{l,m}$. For example, for $l = 2$, the $C^{\text{e}}_{2,m}$ and $C^{\text{m}}_{2,m}$ are, respectively, the five components of the mass quadrupole moment and spin (or angular momentum) quadrupole moment, defined so that $C^{\text{e}}_{2,m} \sim \sum MR^2$ and $C^{\text{m}}_{2,m} \sim \sum LR$, where M, L, and R are the mass, angular momentum, and distance from the origin of each component of the source. If each component is moving with uniform limiting velocity and each angular momentum is bounded, we have $R \sim u$, $L \sim 1$, whence $C^{\text{e}}_{2,m} \sim u^2$ and $C^{\text{m}}_{2,m} \sim u$. The corresponding argument for higher moments leads to (4.3) and (4.4).

By substituting (4.3) and (4.4) into (4.2), a reasonable guess for σ^0_{lin} results:

$$\sigma^0_{\text{lin}} = \eth^2\{A + iB + f\}, \tag{4.5}$$

where

$$A = \sum_{l,m} \beta_{l,m} \frac{d^l(C^{\text{e}}_{l,m} Y_{l,m})}{du^l}$$

is real and tends to a finite limit as $u \to -\infty$ (or as $u \to +\infty$), and where

$$B = \sum_{l,m} \beta_{l,m} \frac{d^l(C^{\text{m}}_{l,m} Y_{l,m})}{du^l}$$

is real and tends to zero as $u \to -\infty$ (or $u \to +\infty$). Of course, (4.5) cannot be inferred rigorously from (4.3) and (4.4) since it is not permissible simply to differentiate order of magnitudes in this way. But if the behavior (4.3) and (4.4) is sufficiently "smooth" asymptotically, then the deduction is valid. This would rule out oscillatory behavior for the sources[12] in the distant past (or future), since such behavior would certainly contradict (4.5). We assume, here, that the components of the system have no accelerations in the limit $u \to -\infty$ (or $u \to +\infty$) which would lead to a violation of (4.5).

[12] An interesting case that would be excluded by our assumptions is that of two masses revolving about one another in ever decreasing circles as they lose energy by gravitational radiation. Here the angular momentum is unbounded as $u \to -\infty$ and apparently $\sigma^0 \sim u^{3/8}$.

We can rewrite (4.5) as

$$\sigma^0_{\text{lin}} \to S_e(\theta, \phi) \tag{4.6}$$

when $u \to -\infty$ (or $u \to +\infty$), S_e being purely "electric" and independent of u. Thus, it seems not unreasonable to infer that, under similar physical assumptions in the full theory, (4.1) will hold [with $S(\theta, \phi)$ purely electric].

It may be felt that the analogy with the linear theory would be a little more trustworthy, however, if given in terms of gauge independent quantities such as the radiation field $\partial^2\sigma^0/\partial u^2$, rather than σ^0. If we assume $\partial^2\sigma^0/\partial u^2$ behaves, in the full theory, in the same way as a $\partial^2\sigma^0_{\text{lin}}/\partial u^2$ consistent with (4.6), then the only difference in the behavior of σ^0 that could arise would lie in the constants of integration. Moreover, the constant of integration involved in the passage from $\partial^2\sigma^0/\partial u^2$ to $\partial\sigma^0/\partial u$ must be zero. This follows from a requirement that the total energy radiated away, as measured according to the Bondi–Sachs formula[1,2]

$$\iiint \left|\frac{\partial\sigma^0}{\partial u}\right|^2 du\, d\theta \sin\theta\, d\phi,$$

be finite. In fact, this finiteness requirement for the gravitationally radiated energy greatly restricts the behavior of $\partial\sigma^0/\partial u$ as $u \to \pm\infty$ in any case in the full theory, in a *rigorous* way (e.g., any simple oscillatory behavior would be ruled out). However, this is not sufficient for our purposes, since the radiated energy is finite if $\sigma^0 \sim u^{\frac{1}{2}-\epsilon}(0 < \epsilon < \frac{1}{2})$ but this would violate (4.1). The remaining constant of integration arises in the passage from $\partial\sigma^0/\partial u$ to σ^0, and this would be the S of (4.1). In the linear theory the $S = S_e$ is purely "electric" and this may be reasonable in the full theory also. But, in any case, the possibility of S having a "magnetic" part can also be treated here.

We now show that, under the assumptions stated above, it is possible to introduce coordinate conditions such that $S_e(\theta, \phi) = 0$. This will be irrespective of whether the "magnetic" part S_m, of $S(= S_e + S_m)$, vanishes. Sachs[2] has shown that, under a BMS transformation (2.1), σ^0 transforms [cf. also (3.2), (3.4)] as

$$\sigma'^0(u, \theta, \phi) = K^{-1} e^{2i\psi}(\sigma^0(u, \theta, \phi) - \tfrac{1}{2}\eth^2\alpha) \tag{4.7}$$

(α having spin weight zero). It should be pointed out that $\sigma'^0(u, \theta, \phi)$ refers to the asymptotic shear of the hypersurfaces $u' = \text{const}$ of the *transformed* coordinate system, although evaluated at the (u, θ, ϕ) of the *original* coordinate system. The complete transformation to $\sigma'^0(u', \theta', \phi')$ is, of course, more complicated. Applied to (4.1), (4.7) gives

$$S'_e(\theta, \phi) = K^{-1} e^{2i\psi}(S_e(\theta, \phi) - \tfrac{1}{2}\eth^2\alpha), \tag{4.8}$$

$$S'_m(\theta, \phi) = K^{-1} e^{2i\psi} S_m(\theta, \phi), \tag{4.9}$$

since α is real. Since by (3.32) we can write $S_e(\theta, \phi) = \eth^2 G$ for some real G and since α can be chosen arbitrarily on the (θ, ϕ)-sphere, it follows that a BMS transformation exists (e.g., $\alpha = 2G$) for which $S'_e(\theta, \phi) = 0$, as required.

We thus may adopt as our coordinate condition the requirement that $S_e(\theta, \phi) = 0$. This can be done for the S_e defined *either* at $u = -\infty$ *or* at $u = +\infty$. For definiteness we could choose the condition defined at $u = -\infty$. (If $S_m = 0$, this means that in the limit $u \to -\infty$ the hypersurfaces $u = \text{const}$ become nearer and nearer to being asymptotically like the "good" cones of Minkowski space.) The BMS transformations which preserve the coordinate condition $S_e = 0$ are now those for which $\eth^2\alpha = 0$ [cf. (4.8)]. Thus α has the form (2.3) and the allowed supertranslations are simply the translations. The "Lorentz rotations" given by $\alpha = 0$ clearly do not destroy the coordinate condition $S_e = 0$. [The fact that S has conformal weight -1 is essential here, since the splitting $S = S_e + S_m$ must be invariant under conformal transformation of the (θ, ϕ)-sphere, $\eth^2$ being also suitably conformally invariant here.] The transformations (2.1) with α as in (2.3) have the same form as inhomogeneous Lorentz transformations in Minkowski space. Thus, the group of coordinate transformations which preserve all the usual coordinate conditions in asymptotic gravitational radiation theory *as well* as keeping $S_e = 0$ at $u = -\infty$ (or else at $u = +\infty$) is isomorphic to the improper orthochronous inhomogeneous Lorentz group.

V. DISCUSSION

There are several inconclusive aspects of this work (even apart from the heuristic nature of the argument concerning the physical assumptions) which should be mentioned. In particular, the ambiguity as to where the coordinate condition $S_e = 0$ is to be imposed is disturbing. If the condition is imposed at $u = -\infty$, there seems no reason at all to believe that the condition would then automatically hold also at $u = +\infty$ (although it is just conceivable that "coherence" relations between the radiation-field data at infinity, of the type considered by Friedlander,[13] might link the con-

[13] F. G. Friedlander, Proc. Roy. Soc. (London) **A279**, 386 (1964).

dition $S_e = 0$ at $u = -\infty$ with that at $u = +\infty$). In fact, the situation is worse than this since all the above arguments could be repeated using an advanced time parameter v in place of the retarded time u. We would then have four different alternative ways of choosing coordinate conditions, namely $S_e = 0$ at $u = \pm\infty$, or the corresponding $S_e = 0$ at $v = \pm\infty$, any one of which would lead to the inhomogeneous Lorentz group, but the different choices would *apparently* be quite unrelated.

It would be interesting to know whether the conditions at $u = -\infty$ are in any way related to those at $v = +\infty$. In the conformal approach to asymptotic analysis,[5] $v = -\infty$, $u = -\infty$, and $u = +\infty$ appear as distinct "points" I^-, I^0, and I^+, respectively, but $v = +\infty$ represents the *same* "points" I^0 as $u = -\infty$. Another question of interest is the relation of the present work to that of the Bergmann,[14] who concludes that if one examines asymptotically flat space–times by proceeding to infinity in *spacelike* directions, then one cannot extract the inhomogeneous Lorentz group as an asymptotic symmetry group. From the conformal point of view, this is again concerned with the "point" I^0, but, since I^0 is approached quite differently in the two analyses, there appears to be no essential conflict here.

A further unanswered question concerns the relevance of the present work to some recent suggestions[6,15,16] that possibly the BMS group can play a role in elementary-particle physics. It is not at all clear that considerations of this paper affect this possibility. The arguments here require consideration of infinite times, while the BMS group emerges if only finite (retarded) time intervals are considered. For such finite time intervals, the BMS group must still be regarded as the relevant asymptotic symmetry group for situations involving gravitational radiation in asymptotically flat space–time.

ACKNOWLEDGMENT

One of the authors, E. Newman, would like to thank Professor H. Bondi and Kings College, London, for their hospitality while this work was being completed. We would like also to thank Dr. A. Janis for having caught a serious error in the original manuscript.

[14] P. G. Bergmann, Phys. Rev. **124**, 274 (1961).

[15] E. Newman, Nature **206**, 811 (1965).

[16] A. Komar, Phys. Rev. Letters **15**, 76 (1961).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 7, NUMBER 5 MAY 1966

Spectrum of the S Operator

E. A. REMLER
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey
(Received 25 August 1965)

The structure of a scattering operator which describes events above a three-body and below a four-body threshold has been studied. It is shown that one can approximate it arbitrarily closely in the norm as the sum of two operators, each having a simple form. Using this decomposition, one obtains a general form for the spectral decomposition of this scattering operator.

INTRODUCTION

THE S-operator describing the elastic scattering of two spinless particles may be written in the form

$$S = \sum_{(Jj)} \int (dP) \exp [2i\delta(J, P^2)] \, |P(Jj)\rangle\langle P(Jj)|, \quad (1)$$

where P is the total 4-momentum of the pair, $[J(J+1)]^{\frac{1}{2}}$ is the magnitude, and j the projection of the total angular momentum vector of the pair in their center-of-mass frame. This general form is dictated entirely by kinematics and in particular by the requirements of Lorentz invariance. The content of this form, the explicit functional dependence of δ on P^2 and J, may be said to be the domain of the dynamics.

Consider next the S-operator describing all processes above a three-and below a four-particle threshold. In addition to P^2 and J, there are now other variables on which $\exp [2i\delta]$ depends. It is shown in this paper how one can approximate, *in the norm*, arbitrarily closely, this S by a sequence

$$S^{(N)} \Rightarrow S$$

$$\text{(convergence in the norm as } N \to \infty). \quad (2)$$

One can then deduce the form of the spectral decomposition of each $S^{(N)}$, obtaining the equation analogous to Eq. (1) which expresses the form of the spectral decomposition in the two-particle case.[1]

The analysis is purely on the mass shell and assumes only that S-matrix elements have the singularities implied by perturbation theory.

It is found that, for any given $S^{(N)}$ and for *fixed* values of P^2 and J, $\exp [2i\delta]$ has a continuous and possibly a discrete range. The continuous range which arises solely and directly from the disconnected graphs is *exactly* known. Essentially, it is just the elastic scattering phase shifts for each two particle pairing in the three-particle subsystem.

[1] Some information about the spectrum of S has been obtained by a number of authors from the point of view of potential scattering [see, e.g., S. Weinberg, Phys. Rev. **133**, B232 (1964)].

NOTATION

Wick's procedure[2] will be used to describe a complete set of states of three particles (respective variables will have suffixes 1, 2, or 3). These are eigenstates of total 4-momentum $P = p_1 + p_2 + p_3$; total angular momentum in the over-all center-of-mass system (Jj); total relativistic mass (squared) of a particular pair of particles, e.g., $\sigma_1 = (p_2 + p_3)^2$; angular momentum of the same pair in their center-of-mass system[3] $(J_1 j_1)$. Their normalization is

$$\langle P'(J'j')\sigma'_\alpha(J'_\alpha j'_\alpha)| \, P(Jj)\sigma_\alpha(J_\alpha j_\alpha)\rangle = \delta_{J'J}\delta_{j'j}\delta_{J_\alpha' J_\alpha}\delta_{j_\alpha' j_\alpha}\delta^4(P' - P)\delta(\sigma'_\alpha - \sigma_\alpha) \quad (3)$$

for $\alpha = 1$, 2, or 3. A complete set for the two-particle states will be, as in Eq. (1), denoted by $|P(Jj)\rangle$. In terms of these, the S-operator is given by

$$S = \sum_{(Jj)} \int (dP)[S_{22}(P^2, J) + S_{23}(P^2, J) + S_{32}(P^2, J) + S_{33}(P^2, J)], \quad (4)$$

$$S_{22}(P^2, J) = S(P^2, J) \, |P(Jj)\rangle\langle P(Jj)|, \quad (5)$$

$$S_{23}(P^2, J) = \sum_{(J_\alpha j_\alpha)} \int d\sigma_\alpha \, S(P^2, J; \sigma_\alpha(J_\alpha j_\alpha)) \times |P(Jj)\rangle\langle P(Jj)\sigma_\alpha(J_\alpha j_\alpha)|, \quad (6)$$

$$S_{32}(P^2, J) = \sum_{(J_\alpha j_\alpha)} \int d\sigma_\alpha \, S(\sigma_\alpha(J_\alpha j_\alpha); P^2 J) \times |P(Jj)\sigma_\alpha(J_\alpha j_\alpha)\rangle\langle P(Jj)|, \quad (7)$$

$$S_{33}(P^2, J) = \sum_{(J_\alpha j_\alpha)} \sum_{(J_{\alpha'} j_{\alpha'})} \int d\sigma_\alpha \, d\sigma'_\alpha \times S(\sigma'_\alpha(J'_\alpha j'_\alpha); P^2 J; \sigma_\alpha(J_\alpha j_\alpha)) \times |P(Jj)\sigma'_\alpha(J'_\alpha j'_\alpha)\rangle\langle P(Jj)\sigma_\alpha(J_\alpha j_\alpha)|. \quad (8)$$

[2] G. Wick, Ann. Phys. (N. Y.) **18**, 65 (1962).

[3] This is only roughly speaking, for exact statement see Ref. (2).

In the above equations, the upper and lower limits on the P^0 integrals, corresponding to the four- and three-body thresholds, are left implicit. Similarly, the *finite* range of σ_α is implicit in all equations. Since unitarity requires

$$S(P^2, J)S^*(P^2, J) + \sum_{(J_\alpha j_\alpha)} \int d\sigma_\alpha \times S(P^2 J; \sigma_\alpha(J_\alpha j_\alpha))S^*(P^2 J; \sigma_\alpha(J_\alpha j_\alpha)) = 1 \quad (9)$$

etc., four other equations, one can construct with these functions, for every fixed $P^2 J$, a unitary operator in a reduced Hilbert space $\tilde{\mathfrak{H}}$.[4] Thus, defining

$$\tilde{S} = \tilde{S}_{22} + \tilde{S}_{23} + \tilde{S}_{32} + \tilde{S}_{33}, \quad (10)$$

$$\tilde{S}_{22} = S(P^2, J)\ |\rangle\langle|, \quad (11)$$

$$\tilde{S}_{23} = \sum_{(J_\alpha j_\alpha)} \int d\sigma_\alpha\ S(P^2 J; \sigma_\alpha(J_\alpha j_\alpha)) \times |\rangle\langle\sigma_\alpha(J_\alpha j_\alpha)|, \quad \text{etc.}, \quad (12)$$

$$\langle\ |\ \rangle = 1, \quad (13)$$

$$\langle\sigma'_\alpha(J'_\alpha j'_\alpha)|\ \sigma_\alpha(J_\alpha j_\alpha)\rangle = \delta_{J_\alpha J'_\alpha}\delta_{j_\alpha j'_\alpha}\delta(\sigma_\alpha - \sigma'_\alpha), \quad (14)$$

one obtains

$$\tilde{S}\tilde{S}^\dagger = \tilde{1} = |\rangle\langle| + \sum_{(J_\alpha j_\alpha)} \int d\sigma_\alpha \times |\sigma_\alpha(J_\alpha j_\alpha)\rangle\langle\sigma_\alpha(J_\alpha j_\alpha)|. \quad (15)$$

This paper is essentially an analysis of the spectrum of $\tilde{S}$. Once this is known, the spectral decomposition of S can be written as

$$S = \sum_{(Jj)} \int (dP)\ |P(Jj)\rangle\ \tilde{S}\ \langle P(Jj)|, \quad (16)$$

using an obvious notation. In the preceding formula (as in the following), P, (Jj) are left implicit parameters.

Finally, since there are three different ways to couple the three particles, one has to deal with recoupling coefficients,[5] and the following abbreviation is used

$$\langle P'(J'j')\sigma'_\alpha(J'_\alpha j'_\alpha)\ |\ P(Jj)\sigma_\beta(J_\beta j_\beta)\rangle = \delta(P' - P)\delta_{J'J}\delta_{j'j}\langle\sigma'_\alpha(J'_\alpha j'_\alpha)\ |\ \sigma_\beta(J_\beta j_\beta)\rangle \quad (17)$$

for $\alpha \neq \beta$.

It is necessary and sufficient for present purposes merely to note that $\langle\sigma'_\alpha(J'_\alpha j'_\alpha)\ |\ \sigma_\beta(J_\beta j_\beta)\rangle$ is, for $\alpha \neq \beta$, a square-integrable function of σ'_α and σ_β.

[4] Equivalently, the space of vectors $(f_3(\sigma_\alpha(J_\alpha j_\alpha)), f_2)$ having finite norm $\|f\|^2 = |f_2|^2 + \sum_{(J_\alpha j_\alpha)} \int d\sigma |f_3|^2$. This norm is implied by the Dirac notation.

[5] G. Wick, Ref. 2.

ANALYSIS

It is convenient to study, instead of S, a directly related Hermetian operator defined by[6]

$$\tilde{D} = 1 - \tfrac{1}{2}(\tilde{S} + \tilde{S}^\dagger) = 1 - \tfrac{1}{2}(\tilde{S} + \tilde{S}^{-1}). \quad (18)$$

In order to proceed further, one must assume something about the structure of S; S-matrix elements will be considered to have only those singularities in the physical region implied by perturbation theory. These are just the δ-function type associated with disconnected graphs and the pole type associated with one real particle exchange.

Consequently, one can write $\tilde{D} = \tilde{D}_d + \tilde{D}_c$, where $\tilde{D}_c$ arises from connected parts and $\tilde{D}_d$ from disconnected parts.

$\tilde{D}_c$ is, however, Hilbert–Schmidt operator in $\tilde{\mathfrak{H}}$. When the pole terms in the diagrams contributing to $\tilde{D}_c$ are integrated over during the projection of $\tilde{D}_c$ onto a subspace of $\tilde{\mathfrak{H}}$ corresponding to a given (Jj), they turn into logarithmic singularities which are square integrable. Thus $\tilde{D}_c$, which contains all the true *three* particle dynamical content in S, is relatively tame. Even so, unfortunately, when $\tilde{D}_c$ is added to $\tilde{D}_d$, it can, in principle, drastically alter the spectral resolution of the latter.[7] In fact, the only aspect the spectra of $\tilde{D}$ and $\tilde{D}_d$ which must always stay the same is the limit points.[8] $\tilde{D}_c$ can, however, be approximated arbitrarily closely *in the norm* by an operator of *finite rank*. The effect of such an operator added to one of known spectral resolution is easily calculable. Specifically, it is clear that only the discrete spectrum[9] of the already known operator may change. Thus, one can say that $\tilde{D}$ is the limit in the norm of a sequence of approximations,[10] each of which has the same spectrum as $\tilde{D}_d$ except for the position and number of discrete points.

The next task is evidently to investigate the spectrum of $\tilde{D}_d$, given by

$$\tilde{D}_d = \sum_{\alpha=1}^{3} \sum_{(J_\alpha j_\alpha)} \int d\sigma_\alpha\ d_\alpha(J_\alpha, \sigma_\alpha) \times |\sigma_\alpha(J_\alpha j_\alpha)\rangle\langle\sigma_\alpha(J_\alpha j_\alpha)|. \quad (19)$$

[6] The second form in Eq. (18) is useful for inverting to find S if D is known only in an approximation.

[7] F. Riesz and B. Sz-Nagy, *Functional Analysis* (Frederick Ungar Publishing Company, New York, 1955), Sec. 134.

[8] F. Riesz and B. Sz-Nagy, Ref. 7, Sec. 133. Limit points are points of the continuous spectrum, limit points of the point spectrum and characteristic values of infinite multiplicity. The tricky point here is that a continuous spectrum might change to a denumerable everywhere dense subset of itself.

[9] Spectrum means only the eigenvalues and not the eigenvectors (which must change).

[10] The Hilbert–Schmidt operator $\tilde{D}_c$ can be approximated by taking any finite number of leading terms in its spectral decomposition.

If the $|\sigma_\alpha(J_\alpha j_\alpha)\rangle$ were orthonormal, this spectrum would be manifest. The discussion extending from this point to Eq. (39) consists in showing that $\tilde{D}_d$ may be approximated by an operator having the same continuous spectrum as $\tilde{D}_d$ would have if $|\sigma_\alpha(J_\alpha j_\alpha)\rangle$ were in fact orthonormal. Introducing the abbreviation

$$\nu = \{\alpha, (J_\alpha j_\alpha)\} \tag{20}$$

(i.e., ν tells what pairing α should be dealt with and the angular momentum numbers for that pairing), Eq. (19) can be rewritten as

$$\tilde{D}_d = \sum_\nu \int d_\nu(\sigma)\, |\sigma\nu\rangle\langle\sigma\nu|\, d\sigma = \sum_\nu \tilde{D}_{d\nu}, \tag{21}$$

where

$$d_\nu(\sigma) = 1 - \cos(2\delta_\nu(\sigma)) = 1 - \cos[2\delta_\alpha(J_\alpha, \sigma_\alpha)]. \tag{22}$$

Thus $\delta_\nu(\sigma)$ is the elastic scattering phase shift for a certain pair of particles in a certain angular momentum state at energy (squared) equal σ.

Consider next the sequence $D_d^{(N)}$; $N = 1, 2, \cdots$ formed by deleting all $\tilde{D}_{d\nu}$ in the summation in Eq. (21) having $J_\alpha > N$ (i.e., N is the highest angular momentum kept in an approximation of the two-body interactions). Thus one writes

$$\tilde{D}_d^{(N)} = \sum_{\nu=1}^{\nu_N} \tilde{D}_{d\nu}. \tag{23}$$

Then

$$\tilde{D}_d^{(N)} \Rightarrow \tilde{D}_d \qquad \text{(convergence in the norm)}. \tag{24}$$

In fact,

$$\|\tilde{D}_d - \tilde{D}_d^{(N)}\| = \|\sum_{\nu>\nu_N} \tilde{D}_{d\nu}\| \leq \sum_\alpha \|\sum_{J_\alpha>N} \sum_{i_\alpha} \tilde{D}_{d\nu}\|, \tag{25}$$

and, since different $(J_\alpha j_\alpha)$ for fixed α designate orthogonal subspaces,

$$\|\sum_{J_\alpha>N} \sum_{i_\alpha} \tilde{D}_{d\nu}\| = \text{Sup}\,\{d_\nu(\sigma);\, J_\alpha > N,\, \sigma\}. \tag{26}$$

But

$$d_\nu(\sigma) = 1 - \cos 2\delta_\nu = 2\,|\sin \delta_\nu(\sigma) e^{i\delta_\nu(\sigma)}|^2 \tag{26}$$

$$= (\text{Kinematic factor}) \times |f_\nu(\sigma)|^2, \tag{27}$$

where $f_\nu(\sigma)$ is the $(J_\alpha j_\alpha)$ expansion coefficient of the Legendre polynomial expansion of the scattering amplitude $f(\theta)$. The existence of such a function as $f(\theta)$ (square integrable over $\cos\theta$) implies that, for each α and σ,

$$\sum_{(J_\alpha i_\alpha)} |f_\nu(\sigma)|^2(2J_\alpha + 1) \equiv c_\alpha(\sigma) \leq \text{Sup}\,\{c(\sigma);\, \sigma\} \equiv c < \infty, \tag{28}$$

so that

$$d_\nu(\sigma) \leq c/(2J_\alpha + 1). \tag{29}$$

This uniform convergence [with respect to σ of the $d_\nu(\sigma)$] to zero as $J_\alpha \to \infty$ ensures, using (26), the convergence in the norm of $\tilde{D}_d^{(N)}$.

In order to clarify the nature of the spectrum of each $\tilde{D}_d^{(N)}$, note that the set $\{|\sigma\nu\rangle;\, \nu \leq \nu_N\}$ is not orthonormal, so it is convenient to introduce

$$(1 + Q_N)\, |\sigma\nu N\rangle = |\sigma\nu\rangle \tag{30}$$

such that

$$\langle\sigma'\nu' N \mid \sigma\nu N\rangle = \delta_{\nu'\nu}\delta(\sigma' - \sigma), \qquad \forall\nu \leq \nu_N. \tag{31}$$

This requires

$$\langle\sigma'\nu' N|\, (1 + Q_N)^\dagger(1 + Q_N)\, |\sigma\nu N\rangle = \langle\sigma'\nu' \mid \sigma\nu\rangle. \tag{32}$$

From the square integrability of the Wick transformation coefficients,[5] it follows that

$$\sum_{\nu\nu'}^{\nu_N} \int d\sigma\, d\sigma'\, |\langle\sigma'\nu'\, |\sigma\nu\rangle - \delta_{\nu'\nu}\delta(\sigma' - \sigma)|^2 < \infty, \tag{33}$$

so that one can write

$$\langle\sigma'\nu' \mid \sigma\nu\rangle = \sum_{n=1}^{\infty} \phi_n^{(N)}(\sigma'\nu')\phi_n^{(N)*}(\sigma\nu)(1 + \lambda_n^{(N)}), \tag{34}$$

where the $\phi_n^{(N)}$ are a complete orthonormal set of functions and

$$\sum_n \lambda_n^{(N)*} < \infty; \quad \lambda_n^{(N)} = \lambda_n^{(N)*}; \quad (1 + \lambda_n^{(N)}) \geq 0. \tag{35}$$

Thus $\langle N\sigma\nu \mid \phi_n^{(N)}\rangle = \phi_n^{(N)}(\sigma\nu)$. Equations (35) are implied by the Hilbert–Schmidt character, the Hermiticity of the kernel in Eq. (33), and the positive semidefiniteness of the norm in $\mathfrak{H}$ (respectively). A consequence of the last statement is that $1 + Q_N$ may have no unique inverse, but this is unimportant for present purposes. Thus

$$\langle\sigma'\nu' N|\, (1 + Q_N)\, |\sigma\nu N\rangle = \sum_n \phi_n^{(N)}(\sigma'\nu')\phi_n^{(N)*}(\sigma\nu)(1 + \lambda_n^{(N)})^{\frac{1}{2}} \tag{36}$$

and

$$Q_N = \sum_n [(1 + \lambda_n^{(N)})^{\frac{1}{2}} - 1]\, |\phi_n^{(N)}\rangle\langle\phi_n^{(N)}|, \tag{37}$$

showing that Q_N can always be found and is itself Hilbert–Schmidt. Writing

$$d^{(N)} = \int d\sigma \sum_{\nu \leq \nu_N} |\sigma\nu N\rangle\, d_\nu(\sigma)\, \langle \sigma\nu N|, \qquad (38)$$

one obtains

$$\tilde{D}_{\mathrm{d}}^{(N)} = d^{(N)} + [Q_N^+ d^{(N)} + d^{(N)} Q_N + Q_N^+ d^{(N)} Q_N]. \qquad (39)$$

The spectrum of $d^{(N)}$ is manifest. The term added to it is Hilbert–Schmidt because of its Q_N multipliers.[11] Thus $\tilde{D}_{\mathrm{d}}^{(N)}$ can be written as the limit in the norm of a sequence of operators of the form $d^{(N)}$ plus operator of finite rank.[10,12] It is then obvious[12] that one can find finite rank approximations such that

$$d^{(N)} + \text{finite rank operator} = \tilde{D}^{(N)} \Rightarrow \tilde{D} \text{ (in the norm)}. \qquad (40)$$

The spectrum of $\tilde{D}^{(N)}$ can consequently only differ from that of $d^{(N)}$ by the addition of a point spectrum. The continuous spectrum remains intact. From this, one can immediately deduce the general form of the spectral decomposition of $\tilde{S}^{(N)}$ [6]

$$\tilde{S}^{(N)} = \sum_{\nu}^{\nu_N} \int d\sigma \exp\,[2i\delta_\nu(\sigma)]\, |\overline{\sigma\nu N}\rangle\langle\overline{\sigma\nu N}| + \sum_{M=1}^{M_N} \gamma_M^{(N)}\, |MN\rangle\langle MN|. \qquad (41)$$

The set $\{|\overline{\sigma\nu N}\rangle,\ |MN\rangle\}$ is orthonormal, and M_N is probably finite[13] for finite N. The bar serves to remind one that $|\overline{\sigma\nu N}\rangle$ in general differ from $|\sigma\nu N\rangle$. The explicit form of this orthonormal set and the eigenvalues of the discrete spectrum γ_M depend on the details of the finite rank operator added to $d^{(N)}$ to form $\tilde{D}^{(N)}$. The $\delta_\nu(\sigma)$ are, however, *exactly* the two body elastic scattering phase shifts. $S^{(N)}$ is now given by[14]

$$S^{(N)} = \sum_{(Ji)} \int (dP)\, |P(J\hat{\jmath})\rangle \tilde{S}^{(N)} \langle P(J\hat{\jmath})|. \qquad (42)$$

DISCUSSION

The S-operator above the production threshold has been shown to be the limit in the norm of a sequence of operators of simple form. A general form for its spectral decomposition has thereby been found.

The continuous spectrum is the sum of the continuous spectra of the possible disconnected processes. It would be interesting to discover the physical significance, if any, of the discrete spectrum.

Finally, these methods and these results appear equally applicable to the S-operator beyond the four-particle threshold.

ACKNOWLEDGMENTS

The author would like to thank Professor M. Goldberger, Dr. J. Challifour, Dr. I. Halliday, and Dr. W. Hunziker for reading and discussing the manuscript.

[11] F. Riesz and B. Sz-Nagy, Ref. 7, Sec. 76.

[12] The finite rank operator in Eq. (40) is the sum of two parts, (i) a finite rank part necessary to add to $d^{(N)}$ to achieve the desired approximation to $\tilde{D}_{\mathrm{d}}{}^{(N)}$, (ii) a finite rank part approximating $\tilde{D}_{\mathrm{e}}$ to arbitrary accuracy.

[13] We know of no rigorous proof of the finiteness of M_N. However, the necessary conditions for it to be infinite appear extremely improbable.

[14] That is, $|P(Jj)\rangle|\rangle = |P(Jj)\rangle$

$|P(Jj)\rangle|\sigma_\alpha(J_\alpha j_\alpha)\rangle = |P(Jj)\sigma_\alpha(J_\alpha j_\alpha)\rangle$.

Exchange of Massive Particles in the Bethe–Salpeter Equation*

JONATHAN L. ROSNER†
Palmer Physical Laboratory, Princeton, New Jersey
(Received 23 August 1965)

The high-energy behavior of the absorptive part A(s) of the forward scattering amplitude in the $g\phi^3$ ladder-graph approximation can be obtained from a homogeneous integral equation. If A(s) grows as s^α, α is the highest characteristic value of this equation. When the mass μ of the exchanged particles is nonzero, the high-g behavior of α is radically different from that in the case of scalar photon exchange. For small g, $\alpha = -1 + g^2a - g^4b + \ldots$, where a is independent of μ. The dependence of b on μ has been calculated.

I. INTRODUCTION

SEVERAL reasons underlie the current popularity of summing Feynman graphs. Attempts to deal with renormalization problems have used such summation techniques.[1] Models which sum infinite series of graphs may have scattering cross sections behaving sensibly at high energies.[2] Finally, the connection between such field-theoretic models and their counterparts in nonrelativistic potential scattering is of great interest, for example, in the theory of Regge poles.[3]

In this paper, we have placed bounds on the leading zero-energy Regge pole $\alpha(0)$ for the sum of ladder graphs in $g\phi^3$ theory (see Fig. 1). Now $\alpha(0)$ has been obtained exactly for a scalar photon-exchange model,[4] but we are interested in drawing an analogy with the Yukawa potential and hence cannot set the mass μ of the exchanged particle equal to zero.

The leading zero-energy Regge pole $\alpha(0)$ may be thought of as the highest angular momentum of a bound state in a Yukawa potential $\lambda e^{-\mu r}/mr$. As the coupling constant is increased, the centrifugal potential $\alpha(\alpha + 1)/r^2$ must increase proportionately, so that

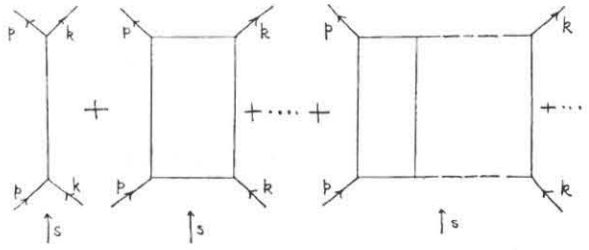

FIG. 1. Ladder graphs in $g\phi^3$ theory. Horizontal lines have mass m and vertical lines have mass μ.

$$\lim_{\lambda\to\infty} \alpha(\alpha + 1)/\lambda = 1/m\mu e. \tag{1}$$

Moreover, this property, with an altered constant on the right-hand side, is true for many potentials.[5]

In the ladder-graph model we identify λ with g^2 times a factor. If the two models were similar we would expect that as $g \to \infty$, $\alpha/g \to$ const. This indeed holds for the scalar photon-exchange model: letting m be the mass of the particles exchanging the photons, the exact expression for α is

$$\alpha = -\tfrac{3}{2} + \tfrac{1}{2}[1 + (g^2/4\pi^2m^2)]^{\frac{1}{2}} \tag{2}$$

so that

$$\lim_{g\to\infty} \alpha(g/4\pi m)^{-1} = 1. \tag{3}$$

However, Eq. (3) has no bearing on an analogy with the Yukawa potential. What we find instead of (3), when $\mu \neq 0$, is that

$$\lim_{g\to\infty} \alpha(g/\mu)^{-\frac{1}{2}} = \text{const}, \tag{4}$$

which is strikingly different from (1) or (3).

We have also examined the effects of a nonzero μ on the weak-coupling limit of α. Defining $\lambda \equiv g^2/16\pi^2m^2$, we have found that

$$\alpha = -1 + \lambda - \lambda^2 f(\mu), \tag{5}$$

* Part of a thesis submitted to Princeton University in partial fulfillment of the degree of Doctor of Philosophy. Work supported by the National Science Foundation.

† Present address: Department of Physics, University of Washington, Seattle, Washington.

[1] T. D. Lee and C. N. Yang, Phys. Rev. **128**, 885 (1962); T. D. Lee, *ibid.* 899; G. Feinberg and A. Pais, *ibid.* **132**, 2724 (1963); **133**, B477 (1964); K. Johnson, M. Baker, and R. Willey, *ibid.* **136**, B1111 (1964).

[2] L. Bertocchi, S. Fubini, and M. Tonin, Nuovo Cimento **25**, 626 (1962); C. Ceolin, F. Duimio, R. Stroffolini, and S. Fubini, *ibid.* **26**, 247 (1962); D. Amati, A. Stanghellini, and S. Fubini, *ibid.* 896; A. Bastai, L. Bertocchi, G. Furlan, and M. Tonin, *ibid.* **30**, 1532 (1963).

[3] B. W. Lee and R. F. Sawyer, Phys. Rev. **127**, 2266 (1962); M. Baker and I. J. Muzinich, *ibid.* **132**, 2291 (1963); R. F. Sawyer, *ibid.* **131**, 1384 (1963); B. W. Lee and A. R. Swift, *ibid.* 1857.

[4] N. Nakanishi, Phys. Rev. **135**, B1430 (1964); *ibid.*, **136**, B1830 (1964).

[5] G. Tiktopoulos and S. B. Treiman, Phys. Rev. **135**, B711 (1964).

where $f(\mu)$ is a definite integral which we have calculated analytically for $\mu = m$, and numerically in general. Note that the $O(\lambda)$ term is independent of the mass of the exchanged particle.[3]

The zero-energy Regge pole at $l = \alpha(0)$ gives rise to a behavior of the forward scattering amplitude as $s^{\alpha(0)}$ in the crossed channel when s, the square of the center-of-mass energy in that channel, becomes large. Assuming that a dispersion relation with the appropriate number of subtractions holds, the high-energy behavior of the scattering amplitude and that of its imaginary or absorptive part are the same.[2] The absorptive part $A(s)$ satisfies an especially simple linear inhomogeneous integral equation,[2] whose kernel and inhomogeneous term are both positive. This enables one to place upper and lower bounds on the solution using certain trial-function methods.[5]

By placing such bounds on $A(s)$, we show that α is a characteristic value of a homogeneous linear integral equation. This equation cannot be solved exactly, but standard methods can be used to get upper and lower bounds on α, leading to the results stated above.

In Sec. II we show that the high-s behavior of $A(s)$ can be learned from a characteristic value problem. Methods for bounding this characteristic value are described in Sec. III, and applied in Sec. IV. Section V is devoted to a brief discussion of our results.

II. THE HOMOGENEOUS EQUATION

The absorptive part due to the Born term (the first diagram on the left in Fig. 1) is just $A^{(2)} = \pi g^2 \delta(s - \mu^2)$. (The superscript refers to the order in perturbation theory.) The absorptive part A of the forward scattering amplitude arising from the sum of the remaining terms in Fig. 1 is a function of s, p^2, and k^2. In what follows we have taken $k^2 = 0$ for kinematical convenience. Then A satisfies the following integral equation:[5]

$$A(s, p^2) = A^{(4)}(s, p^2) + \frac{g^2}{(2\pi)^3} \int d^4q\, \delta(q^2 - \mu^2)\theta(q^0) \frac{A[s', (p-q)^2]}{[m^2 - (p-q)^2]^2}. \quad (6)$$

The inhomogeneous term $A^{(4)}$ corresponds to the second diagram on the left in Fig. I, and is given by

$$A^{(4)}(s, p^2) = \frac{g^4}{16\pi} \frac{\{s(s-4\mu^2)\}^{\frac{1}{2}}\theta(s-4\mu^2)}{[(s-p^2)(sm^2 - \mu^2p^2) + s(m^2-\mu^2)^2]}. \quad (7)$$

It is convenient to express A as a function of the invariants $u \equiv (s - p^2) \equiv 2p\cdot k$ and $x \equiv -p^2$. After multiplying the integral in (6) by

$$1 = \int dx'\, \delta[x' + (p-q)^2] \times \int du'\, \delta[u' - 2(p-q)\cdot k], \quad (8)$$

and performing the q integration, one obtains

$$A(u, x) = A^{(4)}(u, x) + \frac{g^2}{16\pi^2 u} \int du'\, \theta(x_+ - x_-) \int_{x_-}^{x_+} dx' \frac{A(u', x')}{(m^2 + x')^2}, \quad (9)$$

where

$$x_+ = u' - 4\mu^2 \quad (10)$$

and

$$x_- = u'/u\{x + \mu^2[1 - (u'/u)]^{-1}\}. \quad (11)$$

For kinematical convenience we have taken external lines massless. The physical process being considered, the scattering of two massless particles (via the sum of all but the leftmost of the graphs of Fig. 1) corresponds to $x = 0$. The solution to (9) must hence be obtained for all $u \geq 4\mu^2$ and all $x \geq 0$. Of special interest is the high-u behavior of the solution $A(u, x)$ for fixed x. This is of course identical to the high-s behavior of A.

Upper and lower bounds will now be placed on the solution of (9), using a method described by Tiktopoulos and Treiman.[5] Suppose

$$\psi = \phi + K\psi, \quad (12)$$

where $\phi \geq 0$ and $K(x, x') \geq 0$ for all x and x'. Let $\bar{\psi}$ be a trial function, and let $\bar{K}$ be a kernel such that $\bar{K} \geq K$ for all x and x'. Suppose

$$\bar{\psi} - \bar{K}\bar{\psi} \geq \phi = \psi - K\psi \quad \text{for all } x. \quad (13)$$

Then $(\bar{\psi} - \psi) - K(\bar{\psi} - \psi) + (K - \bar{K})\bar{\psi} \geq 0$, so certainly $(\bar{\psi} - \psi) - K(\bar{\psi} - \psi) \geq 0$. If now $1 + K + K^2 + \cdots +$ exists, and is a positive operator,

$$\bar{\psi} \geq \psi. \quad (14)$$

The inequalities can of course be reversed. Equation (9) satisfies all the criteria named above. Its kernel is a Volterra kernel, and $1 + K + K^2 + \cdots +$ is in fact a finite series at any given s.

A. The Upper Bound

We can replace the kernel and inhomogeneous term of (9) by larger ones, and are guaranteed that the solution $\bar{A}(u, x)$ of the resulting equation is an upper bound to $A(u, x)$. In the rest of this paper

take $m = 1$ and $\lambda = g^2/16\pi^2$. The following replacements are convenient: (1) Replace x_+ by ∞. (2) Replace the inhomogeneous term $A^{(4)}(u, x)$ by a larger one

$$\bar{A}^{(4)}(u, x) = (g^4/16\pi u)[u/(x + 4\mu^2)]^{\alpha+1}, \quad (15)$$

where $\alpha + 1 \geq 0$. The majorized integral equation is then

$$\bar{A}(u, x) = \frac{g^4}{16\pi u}\left[\frac{u}{x + 4\mu^2}\right]^{\alpha+1} + \frac{\lambda}{u}\int_0^u du' \int_{x_-}^{\infty} dx' \frac{\bar{A}(u', x')}{(1 + x')^2}, \quad (16)$$

which has a solution of the form

$$\bar{A}(u, x) = u^\alpha \phi_\alpha(x) \quad (17)$$

as long as $\phi_\alpha(x)$ satisfies

$$\phi_\alpha(x) = \frac{g^4}{16\pi}\frac{1}{(x + 4\mu^2)^{\alpha+1}} + \lambda\int_0^1 \xi^\alpha\, d\xi \int_{\xi[x+\mu^2/(1-\xi)]}^{\infty} \frac{\phi_\alpha(x')}{(1 + x')^2}\, dx'. \quad (18)$$

Define $\chi_\alpha(x) \equiv \phi_\alpha(x)/(1 + x)$,

$$g_\alpha(x) \equiv [1/(1 + x)](g^4/16\pi)[1/(x + 4\mu^2)^{\alpha+1}],$$

and perform the ξ integration in (18). Then

$$\chi_\alpha(x) = g_\alpha(x) + \frac{\lambda}{\alpha + 1}\int_0^\infty dx'\, L(x, x')\frac{\chi_\alpha(x')}{(1 + x)(1 + x')}, \quad (19)$$

where

$$L(x, x') = \left[\frac{x + x' + \mu^2 - \{(x + x' + \mu^2)^2 - 4xx'\}^{\frac{1}{2}}}{2x}\right]^{\alpha+1}. \quad (20)$$

Since the term $g_\alpha(x)$ is square integrable on $(0, \infty)$, and since the kernel of (19) is square integrable, the Born series solution of (19) converges whenever λ is less than a certain value λ_0.[6] The quantity λ_0 is the lowest value of λ for which (when α is fixed) there is a solution to the homogeneous equation:

$$\chi_\alpha^h(x) = \frac{\lambda}{\alpha + 1}\int_0^\infty L(x, x')\frac{\chi_\alpha^h(x')}{(1 + x)(1 + x')}\, dx'. \quad (21)$$

Since

$$g_\alpha(x) > 0,$$

$$\chi_\alpha(x) = [(1 + \lambda K + \lambda^2 K^2 + \cdots +)g_\alpha](x) > 0,$$

as is necessary in order that $\bar{A}(u, x) = u^\alpha(1+x)\chi_\alpha(x)$ provides an upper bound to $A(u, x)$.

Since the kernel of (19) is monotonically decreasing in α, the Born series converges for all $\alpha \geq \alpha_c$ if it converges for $\alpha = \alpha_c$. For a fixed value of λ, there is a highest value of α for which (21) has a solution. Call this value α_0. The Born series of (19) converges for all $\alpha > \alpha_0$.

The upper bound to $A(u, x)$ is now constructed by letting $\alpha = \alpha_0 + \epsilon$, where $\epsilon > 0$. Then (19) has a positive solution $\chi_\alpha(x)$, and

$$A(u, x) \leq (1 + x)u^\alpha \chi_\alpha(x). \quad (22)$$

The bound (22) is useful for all x, $0 \leq x < \infty$, only if $\chi_\alpha(x) < \infty$, $0 \leq x < \infty$. The solution to (19) is not *a priori* finite at the origin, but only square integrable on $(0, \infty)$.[7] However, using the Schwarz inequality, we find

$$\chi_\alpha(x) < g_\alpha(x) + \frac{\lambda}{\alpha + 1}\left[\int_0^\infty \frac{\{L(x, x')\}^2\, dx'}{(1 + x')^2}\right]^{\frac{1}{2}} \times \left[\int_0^\infty \{\chi_\alpha(x')\}^2\, dx'\right]^{\frac{1}{2}};$$

and since $L(x, x') < 1$,

$$\chi_\alpha(x) < g_\alpha(x) + \text{const}\, x\, ||\chi_\alpha|| < \infty,$$

where $||\chi_\alpha|| \equiv [\int_0^\infty \{\chi_\alpha(x')\}^2\, dx']^{\frac{1}{2}}$. In particular, $\chi_\alpha(0)$ exists.

Since ϵ may be made as small as desired, we have shown:

$$\text{If } \lim_{u\to\infty} A(u, x)/u^\alpha = \text{const}, \quad \alpha \leq \alpha_0. \quad (23)$$

B. The Lower Bound

If a solution $A(u, x)$ to (9) is guessed, and it turns out that $A - \lambda KA \leq A^{(4)}$, then $A \leq A$, as was mentioned above. Assume a trial function

$$A(u, x) = c\theta[u - s_0 - (1 + a)x]u^\alpha \phi_\alpha(x). \quad (24)$$

The restrictions on the parameters are:

$$c > 0, \qquad \alpha > -1, \qquad a > 0, \qquad s_0 > 4\mu^2(1 + a). \quad (25)$$

Let s_1 be an arbitrary parameter, subject to the condition

$$\frac{s_1}{1 + a} \geq \left\{2\left[\frac{s_0}{1 + a}\right]^{\frac{1}{2}} + \mu\right\}^2. \quad (26)$$

(The reason for this condition is given in Appendix

[6] S. Weinberg, Phys. Rev. **133**, B232 (1964).

[7] F. Riesz and B. Sz-Nagy, *Functional Analysis* (Frederick Ungar Publishing Company, New York, 1955), p. 150.

A.) The definitions $\sigma_0 \equiv s_0/(1 + a)$ and $\sigma_1 \equiv s_1/(1 + a)$ are useful.

We now choose the parameters so that

$$A - \lambda KA \leq A^{(4)} \quad \text{for all } u, x, \tag{27}$$

directing all efforts toward obtaining the highest possible value of α consistent with (27).

First, for $s(=u - x) \leq s_1 + ax$, demand that

$$A(u, x) \leq A^{(4)}(u, x). \tag{28}$$

Since $K \geq 0$, this will imply (27). Now (28) can be satisfied as long as

$$\phi_\alpha(x) \leq M/(1 + x)^{\alpha+1} \tag{29}$$

for some M, so let us adopt this restriction on the ϕ_α to be used in (24). To satisfy (28) is then merely a matter of choosing small enough c. For $s \geq s_1 + ax$, demand that

$$A(u, x) - \frac{\lambda}{u}\int du'\ \theta(x_+ - x_-) \times \int_{x_-}^{x_+} dx'\ \frac{A(u', x')}{(1 + x')^2} < 0. \tag{30}$$

This condition will certainly imply $A - KA \leq A^{(4)}$. Since $s_0 \geq 4\mu^2(1 + a)$ and $u \geq s_1 + 4\mu^2(1 + a)$, if we define

$$x_2 \equiv \frac{\sigma_1 u'/u}{1 + (\sigma_0/x)} - \sigma_0 \tag{31}$$

we have that $x_+ \geq x_2$, so a condition that implies (30) is

$$A(u, x) - \frac{\lambda}{u}\int du'\ \theta(x_2 - x_-) \times \int_{x_-}^{x_2} dx'\ \frac{A(u, x')}{(1 + x')^2} \leq 0. \tag{32}$$

This will be satisfied by the trial function

$$A(u, x) = c\theta(s - s_0 - ax)u^\alpha\phi_\alpha/(x)$$

for

$$s \geq s_1 + ax \ [\text{or } u \geq s_1 + (1 + a)x]$$

if

$$\phi_\alpha(x) - \lambda\int \xi^\alpha\, d\xi\ \theta(\chi_2 - \chi_-) \times \int_{x_-}^{x_2} dx'\ \frac{\phi_\alpha(x')}{(1 + x')^2} \leq 0, \tag{33}$$

where

$$\begin{aligned} x_2 &= \sigma_1\xi/[1 + (\sigma_0/x)] - \sigma_0 \\ \text{and} \qquad x_- &= \xi[x + \mu^2/(1 - \xi)]. \end{aligned} \tag{34}$$

The above manipulations eliminate the variable u, reducing the problem to a one-dimensional one. Now find the maximum value of α for which equality can hold in (33). [In Appendix A we show that the conditions (26) and (31) guarantee that such a value always exist.] Let ϕ_α be the corresponding eigenfunction. Since the kernel of (33) is everywhere non-negative, this eigenfunction can be chosen everywhere non-negative (see Appendix A). We may choose $\phi_\alpha(0) = 1$. If α is then increased by an infinitesimal amount, the inequality (33) will be satisfied, since the kernel of (33) is monotonically decreasing in α. It is shown in Appendix A that the eigenfunction ϕ_α satisfies $\phi_\alpha \leq M/(1 + x)^{\alpha+1}$. Hence, the trial function satisfies $A - KA < A^{(4)}$ for all values $u \geq 4\mu^2$, $x \geq 0$, and thus provides a lower bound to $A(u, x)$. We shall now show how to obtain the "best possible" lower bound.

Recall the definition $\chi_\alpha(x) \equiv \phi_\alpha(x)/(1 + x)$. Equality in (33) is equivalent to the equation

$$\chi_\alpha(x) = \frac{\lambda}{\alpha + 1}\int_0^\infty L_1(x, x')\,\frac{\chi_\alpha(x')}{(1 + x)(1 + x')}, \tag{35}$$

where

$$L_1(x, x') = \xi_b^{\alpha+1} - \xi_a^{\alpha+1}, \tag{36}$$

$$\xi_b = \{x + x' + \mu^2 - [(x + x' + \mu^2)^2 - 4xx']^{\frac{1}{2}}\}(2x)^{-1}, \tag{37}$$

and

$$\xi_a = x'[1 + (\sigma_0/x')][1 + (\sigma_0/x)]\sigma_1^{-1}, \tag{38}$$

as can be seen by reversing the order of integration in (33) and performing the ξ integral. Let α_1 denote the highest value of α for which (35) has a solution. It is a continuous function of σ_1. As $\sigma_1 \to \infty$, the kernel of (35) remains bounded and the traces of all its iterates exist. In fact,

$$\lim_{\sigma_1\to\infty} L_1(x, x') = L(x, x'), \tag{39}$$

where $L(x, x')$ is just the kernel of (21), defined in (20). Hence as $\sigma_1 \to \infty$, α_1 approaches α_0 arbitrarily closely from below. This means that, by choosing σ_1 large enough [and satisfying (25), (26), and a restriction on c], we can exhibit a lower bound to $A(u, x)$ of the form

$$A(u, x) = cu^{\alpha_0-\epsilon}\theta[u - s_0 - (a + 1)x]\phi_\alpha(x), \tag{40}$$

where ϵ is arbitrary small, so that if

$$\lim_{u\to\infty} A(u, x)/u^\alpha = \text{const}, \quad \alpha \geq \alpha_0. \tag{41}$$

Combining (41) with (23), we have the desired result: *let α_0 be the highest value of α for which* (21) *has a solution. Then if*

$$\lim_{u\to\infty} A(u, x)/u^{\alpha} = \text{const}, \quad \alpha = \alpha_0.$$

The kernel of (21) is monotonically decreasing in α, so $\alpha_0(\lambda)$ is a monotonically increasing function. Its domain is $0 \leq \lambda < \infty$ and its range is $-1 \leq \alpha < \infty$. Hence it has a unique inverse $\lambda_0(\alpha)$ which is the lowest value of λ for which (21) has a solution (for fixed α between -1 and ∞). We shall now obtain upper and lower bounds on $\lambda_0(\alpha)$, and then invert for the bounds on $\alpha_0(\lambda)$.

III. BOUNDS ON EIGENVALUES

A *polar* kernel $K(x, x')$ is one which can be expressed as

$$K(x, x') = f(x)S(x, x')[f(x')]^{-1}, \tag{42}$$

where $S(x, x') = S(x', x)$. The kernel of (21) is polar, for example. Consider the equation

$$\phi(x) = \lambda \int K(x, x')\phi(x')\, dx', \tag{43}$$

where K is polar. Its eigenvalues are the same as those of

$$\psi(x) = \lambda \int S(x, x')\psi(x')\, dx', \tag{44}$$

and its solutions are $\phi(x) = f(x)\psi(x)$. Now, a symmetric kernel $S(x, x')$ has at least one characteristic value λ_0,[8] given by

$$\lambda_0 = \inf_{\psi \epsilon L^2} \frac{(\psi, \psi)}{(\psi, S\psi)}, \tag{45}$$

where a function belongs to L^2 if

$$||\psi|| \equiv \left[\int_0^{\infty} dx\, \{\psi(x)\}^2\right]^{\frac{1}{2}} < \infty.$$

The inner product (ψ, ϕ) is defined by

$$(\psi, \phi) \equiv \int_0^{\infty} dx\, \psi(x)\phi(x).$$

For any trial function $\psi_\tau \epsilon L^2$, by (45),

$$\lambda_0 \leq (\psi_\tau, \psi_\tau)/(\psi_\tau, S\psi_\tau), \tag{46}$$

providing a simple upper bound to λ_0.

In principle, one can compute the lowest eigenvalue of a polar kernel K as closely as desired.[9] It can be shown that

$$H_n \equiv \text{Tr}\,(K^n) = \sum_{i=1}^{\infty} \frac{1}{(\lambda_i)^n}, \qquad n \geq 2, \tag{47}$$

where λ_i are the characteristic values of K. Since the eigenvalues of a polar kernel are real (if it is related to a symmetric kernel via a real function), $H_{2n} \geq 0$ for $n \geq 1$. Suppose the eigenvalue λ_0 occurs with multiplicity r_0. Rewrite (47) as

$$H_{2n} = \frac{r_0}{\lambda_0^{2n}}(1 + \epsilon_n), \text{ where } \epsilon_n = \frac{1}{r_0}\sum_{j=r_0}^{\infty}\left(\frac{\lambda_0}{\lambda_j}\right)^{2n}.$$

Since $\lambda_j > \lambda_0$ for $j \geq r_0$, we see that $\epsilon_n \to 0$ as $n \to \infty$. Then

$$\lim_{n\to\infty} (H_{2n})^{-1/n} = \lambda_0^2 \tag{48}$$

and

$$|\lambda_0| \geq (r_0/H_{2n})^{\frac{1}{2n}} \quad \text{for all} \quad n. \tag{49}$$

The above methods will now be applied to obtain bounds on the lowest characteristic value $\lambda_0(\alpha)$ of Eq. (21).

IV. RESULTS

A. Strong-Coupling Limit

Equation (21) may be rewritten as an equation with symmetric kernel using the definitions

$$y \equiv \mu^2/x,$$
$$\psi(y) \equiv y^{-(\alpha+3)/2}\chi(\mu^2/y),$$
$$z \equiv \alpha + 1;$$

we then obtain

$$\psi(y) = \frac{\lambda}{z\mu^2}\int_0^{\infty} \frac{\psi(y')\, dy'}{[1 + (y/\mu^2)][1 + (y'/\mu^2)]} \times [a - \{a^2 - 1\}^{\frac{1}{2}}]^z, \tag{50}$$

where

$$a \equiv \cosh v + t, \tag{51}$$

$$v \equiv \tfrac{1}{2}\log(y'/y), \tag{52}$$

and

$$t \equiv \tfrac{1}{2}(yy')^{\frac{1}{2}}. \tag{53}$$

Minorization of the kernel increases the lowest eigenvalue, as can be seen from (45). The kernel of (50) may be minorized using the inequality

$$a - \{a^2 - 1\}^{\frac{1}{2}} \geq \exp(-\{v^2 + 2t\}^{\frac{1}{2}}), \tag{54}$$

which is equivalent to

$$\tfrac{1}{2}[\log(a + \{a^2 - 1\}^{\frac{1}{2}})]^2 \leq (v^2/2) + t \tag{55}$$

[8] R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience Publishers, Inc., 1953), Vol. I, p. 122.
[9] S. G. Mikhlin, *Integral Equations* (Pergamon Press, Inc., New York, 1957), p. 90.

(equality holds in (55) for $t = 0$, and the inequality may be verified for $t \geq 0$ by differentiating both sides with respect to t). An equation with lowest eigenvalue larger than that of (50) is hence

$$\psi(y) = \frac{\lambda}{z\mu^2} \int_0^\infty \frac{\psi(y')\, dy'}{[1 + (y/\mu^2)][1 + (y'/\mu^2)]} \times \exp\left(-z\{v^2 + 2t\}^{\frac{1}{2}}\right). \quad (56)$$

To obtain an upper bound on λ_0, one can now use (46), with S the kernel of (56). Choose a trial function

$$\psi_\tau(y) = \theta[(\gamma^2/z^2) - y], \quad (57)$$

where γ is an arbitrary parameter. Then

$$(\psi_\tau, \psi_\tau) = \gamma^2/z^2,$$

and

$$(\psi_\tau, S\psi_\tau) \geq 2\{(z\mu^2[1 + (\gamma^2/\mu^2 z^2)^2]\}^{-1} \times \int_0^{\gamma^2/z^2} dy \int_0^y dy' \exp\left(-z\{v^2 + \gamma^2/z^2\}^{\frac{1}{2}}\right)$$

$$= \frac{2\gamma^4}{z^6\mu^2[1 + (\gamma^2/\mu^2 z^2)^2]} \int_0^\infty e^{-2\beta/z}\, d\beta\, e^{-\{\beta^2+\gamma^2\}^{\frac{1}{2}}}.$$

Now

$$\int_0^\infty d\beta\, e^{-\{\beta^2+\gamma^2\}^{\frac{1}{2}}} = \gamma K_1(\gamma),^{10}$$

so

$$\lim_{z\to\infty} \frac{z^4(\psi_\tau, S\psi_\tau)}{(\psi_\tau, \psi_\tau)} \geq \frac{2}{\mu^2} \max_\gamma \{\gamma^3 K_1(\gamma)\}. \quad (58)$$

Since $\max_\gamma\{\gamma^3 K_1(\gamma)\} \simeq 1.16$ at $\gamma \simeq 2.45$, we have finally

$$\lim_{z\to\infty} z/(\lambda/\mu^2)^{\frac{1}{2}} \geq (2 \times 1.16)^{\frac{1}{2}} = 1.23. \quad (59)$$

To obtain a lower bound on λ_0, return to (21). Define $H_n \equiv \mathrm{Tr}\,(K^n)$ where K is the kernel of (21). From (49) (since K and λ_0 are positive) we have

$$\lambda_0 \geq (H_n)^{-1/n} \quad \text{for all } n \geq 1. \quad (60)$$

Define

$$\eta(x, x') = \frac{x + x' + \mu^2 - \{(x + x' + \mu^2)^2 - 4xx'\}^{\frac{1}{2}}}{2(xx')^{\frac{1}{2}}}. \quad (61)$$

Then H_n can be written

$$H_n = z^{-n} \int_0^\infty \cdots \int_0^\infty \prod_{i=1}^n \left\{\frac{dx_i}{(1 + x_i)^2}\, [\eta(x_i, x_{i+1})]^z\right\} \quad (62)$$

(identify x_{n+1} with x_1). The following change of variables is convenient. Multiply (62) by

$$1 = \int_{-\infty}^\infty dW\, \delta(W)$$

and define ρ and ω_i $(i = 1, \cdots, n)$ by

$$x_i^{-1} = \rho \exp(\omega_i), \quad (63)$$

$$W = \sum_{i=1}^n \omega_i. \quad (64)$$

Then

$$H_n = nz^{-n} \int_0^\infty \rho^{n-1}\, d\rho \int_{-\infty}^\infty \cdots \int_{-\infty}^\infty \delta\left(\sum_{i=1}^n \omega_i\right) \times \prod_{i=1}^n \left\{\frac{d\omega_i}{(1 + \sigma e^{\omega_i})^2}\, [a_i - (a_i^2 - 1)^{\frac{1}{2}}]^z\right\}, \quad (65)$$

where

$$a_i = \cosh \tfrac{1}{2}(\omega_i - \omega_{i+1}) + \tfrac{1}{2}\rho\mu^2 \exp[\tfrac{1}{2}(\omega_i + \omega_{i+1})]. \quad (66)$$

Now divide the region of integration over the ω_i into two portions. Call E the region where $-\epsilon \leq \omega_i \leq \epsilon$ for all i and some fixed $\epsilon > 0$. Call R the complement of this region. For large z, we shall show that the contribution of region R to H_n is no greater than $z^{-n} \exp(-2\epsilon z/n)$, while that of region E is no greater than a constant times z^{1-4n}.

For even n, if $\sum_{i=1}^n \omega_i = 0$ and at least one $\omega_i \geq \epsilon$, then one or more of the quantities $|\omega_i - \omega_{i+1}|$ is at least $4\epsilon/n$. This implies that in R one or more a_i is at least $\cosh(2\epsilon/n)$, so one or more $a_i - (a_i^2 - 1)^{\frac{1}{2}}$ is at most $\exp(-2\epsilon/n)$. Denoting the contribution of region R to H_n by $(H_n)_R$, we see from (62) that

$$(H_n)_R \leq z^{-n} \exp(-2\epsilon z/n). \quad (67)$$

In region E, every a_i is greater than or equal to the corresponding a_i defined by

$$a_i \equiv \cosh \tfrac{1}{2}(\omega_i - \omega_{i+1}) + \tfrac{1}{2}\rho\mu^2 e^{-\epsilon}. \quad (68)$$

Hence

$$(H_n)_E \leq nz^{-n} \int_0^\infty \rho^{n-1}\, d\rho \int_{-\epsilon}^\epsilon \cdots \int_{-\epsilon}^\epsilon \delta\left(\sum_{i=1}^n \omega_i\right) \times \prod_{i=1}^n \left\{\frac{d\omega_i}{(1 + \sigma e^{\omega_i})^2}\, [a_i - (a_i^2 - 1)^{\frac{1}{2}}]^z\right\}. \quad (69)$$

This can be further majorized by neglecting the factors in the denominator and replacing the limits on ω_i by $-\infty$ and ∞. Since the ith factor in the product is then a function only of ρ and the difference $\omega_i - \omega_{i+1}$, the majorized integral can be

[10] A. Erdelyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi, *Higher Transcendental Functions* (McGraw-Hill Book Company, Inc., New York, 1953), Vol. 2, p. 82.

simplified considerably. Perform the integral over ω_n to eliminate the δ-function. Then $\omega_n = -\sum_{i=1}^{n-1} \omega_i$. Define

$$v_i = \frac{\omega_i - \omega_{i+1}}{2}, \qquad i + 1, \cdots, n - 2;$$

$$v_{n-1} = \frac{\omega_{n-1} - \omega_n}{2} = \frac{1}{2}\left(\omega_{n-1} + \sum_{i=1}^{n-1} \omega_i\right),$$

so that $dv_1 \cdots dv_{n-1} = (n/2^{n-1})\, d\omega_1 \cdots d\omega_{n-1}$. Note that $a_n - (a_n^2 - 1)^{\frac{1}{2}} \leq 1$. Then

$$(H_n)_E \leq \frac{1}{2}\left[\frac{2}{z}\right]^n \int_0^\infty \rho^{n-1}\, d\rho\, [C(\rho)]^{n-1},$$

where

$$C(\rho) \equiv \int_{-\infty}^{\infty} dv\, \{a - (a^2 - 1)^{\frac{1}{2}}\}^s$$

and

$$a \equiv \cosh v + \frac{\rho\mu^2}{2} e^{-\epsilon}.$$

We shall now scale out z (in order to pass to the $z \to \infty$ limit) and μ^2. Define β and γ by

$$\beta \equiv vz,$$
$$\gamma^2 \equiv \rho\mu^2 e^{-\epsilon} z^2.$$

Then

$$(H_n)_E \leq \left[\frac{2e^\epsilon}{\mu^2 z^4}\right]^n z \int_0^\infty \gamma^{2n-1}\, d\gamma\, [D(\gamma)]^{n-1}, \qquad (70)$$

where

$$D(\gamma) \equiv \int_{-\infty}^{\infty} d\beta\, \{a - (a^2 - 1)^{\frac{1}{2}}\}^z \qquad (71)$$

and

$$a = \cosh(\beta/z) + (\gamma^2/2z^2). \qquad (72)$$

We have chosen β and γ so that $\lim_{z\to\infty} D(\gamma)$ is finite and nonzero for all $\gamma \geq 0$. The quantity $\{a - (a^2 - 1)^{\frac{1}{2}}\}^z$ may be written $\exp\{-z \log [a + (a^2 - 1)^{\frac{1}{2}}]\}$. Expand $\log(a + \{a^2 - 1\}^{\frac{1}{2}})$ in a power series in $1/z$; the result is

$$\log(a + \{a^2 - 1\}^{\frac{1}{2}}) = (1/z)(\beta^2 + \gamma^2)^{\frac{1}{2}} + O(z^{-3}).$$

Hence

$$\lim_{z\to\infty} D(\gamma) = \int_{-\infty}^{\infty} d\beta\, e^{-(\beta^2+\gamma^2)^{\frac{1}{2}}} = 2\gamma K_1(\gamma),^{10}$$

and

$$\lim_{z\to\infty} [(H_n)_E z^{4n-1}] \leq \frac{1}{2}\left[\frac{4e^\epsilon}{\mu^2}\right]^n \int_0^\infty \gamma^{3n-2}\, d\gamma\, [K_1(\gamma)]^{n-1}. \qquad (73)$$

Since (49) holds for all μ and for all z, we may take the large-z limit while letting n tend to infinity as well. If we choose n = largest even integer smaller than $z^{\frac{1}{2}-\Delta}$, where $\Delta > 0$, we see from (67) that $z^4[(H_n)_R]^{1/n}$ tends to 0 as $z \to \infty$. However, $z^4[(H_n)_E]^{1/n}$ tends to (at most)

$$\frac{4e^\epsilon}{\mu^2} \max_\gamma \{\gamma^3 K_1(\gamma)\}$$

as $z \to \infty$ for this choice of n, and since $\lim_{n\to\infty} (p^n + q^n)^{1/n} = \max(p, q)$, we have (letting ϵ be as small as we like)

$$\lim_{\substack{z\to\infty \\ n\to\infty}} (z^4 \{H_n\}^{1/n}) \leq \frac{4}{\mu^2} \max_\gamma \{\gamma^3 K_1(\gamma)\}$$

or

$$\lim_{z\to\infty} z \Big/ \left(\frac{\lambda}{\mu^2}\right)^{\frac{1}{2}} \leq (4 \times 1.16)^{\frac{1}{2}} = 1.47. \qquad (74)$$

From the bounds (74) and (59) we see that $\alpha = z - 1$ behaves as $(g/\mu)^{\frac{1}{2}}$ for large g.

B. Weak-Coupling Limit

To obtain an upper bound on λ_0, apply (46) to (50) using the trial function

$$\psi_r(y) = [1 + (y/\mu^2)]^{-1}.$$

Then (46) implies

$$\lambda_0 \leq z[I(\mu, \tfrac{1}{2}z)]^{-1}, \qquad (75)$$

where

$$I(\mu, z) \equiv \int_0^\infty \frac{dy/\mu^2}{(1 + y/\mu^2)^2} \times \int_0^\infty \frac{dy'/\mu^2}{(1 + y'/\mu^2)^2}\, [a - (a^2 - 1)^{\frac{1}{2}}]^z, \qquad (76)$$

and a is given by (51), (52), and (53).

To obtain a lower bound to λ_0, apply (49) to (50) for $n = 1$, obtaining

$$\lambda_0 \geq z[I(\mu, z)]^{-\frac{1}{2}}. \qquad (77)$$

The bounds (75) and (77) are especially close to one another for small z, as can be seen by expanding $I(\mu, z)$ in a Taylor's series in z about $z = 0$. Note that $I(\mu, 0) = 1$. Writing

$$I(\mu, z) = 1 + zI_z(\mu, 0) + \tfrac{1}{2}z^2 I_{zz}(\mu, 0) + \cdots +,$$

we have

$$[I(\mu, z)]^{-\frac{1}{2}} = 1 - \tfrac{1}{2}zI_z(\mu, 0) + O(z^2)$$

and

$$[I(\mu, \tfrac{1}{2}z)]^{-1} = 1 - \tfrac{1}{2}zI_z(\mu, 0) + O(z^2),$$

TABLE I. The function $f(\mu)$ for selected values of μ^2.

μ^2	$f(\mu)$
0	1.00
0.01	1.02
0.05	1.07
0.10	1.12
0.50	1.37
1.00	1.56
5.00	2.31
10.00	2.77
50.00	4.07

which means that

$$\alpha = -1 + \lambda - \lambda^2 f(\mu) + O(\lambda^3), \tag{78}$$

where

$$f(\mu) = -\tfrac{1}{2} I_s(\mu, 0) = \int_0^\infty \frac{dy/\mu^2}{(1 + y/\mu^2)^2} \times \int_0^\infty \frac{dy'/\mu^2}{(1 + y'/\mu^2)^2} \log\left(a + \{a^2 - 1\}^{\frac{1}{2}}\right), \tag{79}$$

and a is given by (51), (52), and (53).

To order λ (78) is just Lee and Sawyer's weak-coupling limit.[3] The coefficient of λ is independent of the mass μ of the exchanged particle; the dependence on μ first enters in the λ^2 coefficient.

For $\mu = 1$, (76) may be done in closed form (see Appendix B). Its value is

$$I(1, z) = 1 + \tfrac{1}{9} z[2z\psi'(\tfrac{1}{3}\{z + 3\}) - (z + 4)\psi'(\tfrac{1}{3}\{z + 1\}) - (z - 4)\psi'(\tfrac{1}{3}\{z + 2\})], \tag{80}$$

where

$$\psi'(z) = (d^2/dz^2) \log \Gamma(z). \tag{81}$$

Then

$$I_s(1, 0) = (4/9)[\psi'(\tfrac{2}{3}) - \psi'(\tfrac{1}{3})] = -3.125,$$[11]

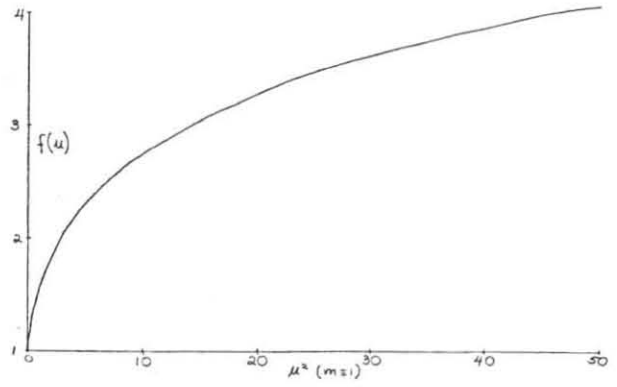

FIG. 2. Function $f(\mu)$.

[11] E. Jahnke and F. Emde, *Tables of Functions with Formulae and Curves* (Dover, New York, 1945), 4th ed., p. 17.

so

$$f(1) = 1.56. \tag{82}$$

The behavior of $I(\mu, z)$ as $\mu \to 0$ is interesting since it is not analytic in μ^2 at $\mu = 0$, as may be verified by computing $I_s(\mu, 0)$ and letting $\mu \to 0$. The integral then diverges logarithmically. The limit of $I(\mu, z)$, as $\mu \to 0$ exists, is

$$I(0, z) = 1 - 2z + 2z^2\psi'(1 + z) \tag{83}$$

(see Appendix C), so (78) gives

$$\alpha = -1 + \lambda - \lambda^2 + O(\lambda^3) \tag{84}$$

in agreement with the result obtained from Eq. (2). One can show that $f(\mu) = 1 - O(\mu^2 \log \mu^2)$ as $\mu \to 0$.

The values of $f(\mu)$ for selected values of μ have been computed numerically and are given in Table I. A graph of these values is given in Fig. 2.

V. SUMMARY AND CONCLUSIONS

We have been concerned with the high-energy behavior of the absorptive part of the forward scattering amplitude in the ϕ^3 ladder-graph approximation. Some consequences of letting the exchanged particles be massive were noted.

The absorptive part satisfies an inhomogeneous linear integral equation (9), whose solutions can be bounded from above and below. By placing such bounds we have shown that the high-s behavior is determined by a homogeneous equation (21). If $A(s)$ grows as s^α, α is the highest characteristic value of this equation. Equation (21) was first studied by Ceolin *et al.*,[2] but we felt that a rigorous proof of its relevance would be useful.

Note added in proof: An equation equivalent to (21) may also be obtained by applying a type of Laplace transform to (9) to reduce it to a Fredholm equation in the single variable x.[12]

Upper and lower bounds have been placed on α in both strong- and weak-coupling cases. The results are:

(1) Strong coupling:

$$1.23 \leq \lim_{g \to \infty} \frac{\alpha}{(g/4\pi\mu)^{\frac{1}{2}}} \leq 1.47, \tag{85}$$

(2) Weak coupling:

$$\alpha = -1 + \lambda - \lambda^2 f(\mu), \tag{78}$$

[12] S. Nussinov and J. Rosner, J. Math. Phys. (to be published).

where

$$f(\mu) = 1 - O(\mu^2 \log \mu^2) \quad \text{as} \quad \mu \to 0,$$

$$f(1) = 1.56,$$

and the general $f(\mu)$ is given by Fig. 2 and Eq. (79).

The dependence of α on the ratio g/μ for strong coupling is entirely reasonable. In potential theory $1/\mu$ corresponds to the range of the potential, so we find that, in the strong-coupling limit, α is a function of the (dimensionless) product of strength and range. Since g has the dimensions of a mass, we see that, in the strong-coupling limit, α is independent of the mass of the particles forming the bound state (i.e., the horizontal lines in Fig. 1, of mass m).

The analogy between a Yukawa interaction and the ladder-graph approximation in $g\phi^3$ theory cannot be applied to Regge-pole behavior in the strong-coupling limit, as a comparison of (1) and (4) shows (identifying λ with g^2 times a factor).

The sensitivity to the mass of the exchanged particle is peculiar to the ϕ^3 theory. It stems from the Fredholm nature of the ladder-graph Bethe–Salpeter equation.[3] In a similar ϕ^4 model the Bethe–Salpeter equation is singular.[3] The ϕ^4 interaction is analogous to a singular potential behaving as $1/r^2$ at the origin, whose range one would expect to be less crucial in determining the highest possible angular momentum of a bound state. This is the case: the high-energy behavior in a ϕ^4 model analogous to the ladder-graph model is independent of whether or not any of the internal particles are massive.[3]

ACKNOWLEDGMENTS

Part of this work was done at the 1964 and 1965 Summer Institutes for Theoretical Physics at the University of Wisconsin. I am grateful to Professor R. G. Sachs and Professor K. W. McVoy for extending the hospitality of the department and arranging for computer time at the UWCC. At Princeton, my thanks to F. Gilman, J. Noble, O. Lanford, and Dr. G. F. Tiktopoulos for helpful advice. Finally, my special thanks to Professor Sam B. Treiman, whose constant interest and critical suggestions aroused my curiosity and channeled my energies.

APPENDIX A. HOMOGENEOUS EQUATION FOR THE LOWER BOUND

We show in this Appendix, (1) that (35) has a highest characteristic value $\alpha = \alpha_1$, and (2) that the corresponding eigensolution can be chosen nonnegative and obeys (29).

Defining $\psi_\alpha(x) = x^{\frac{1}{2}(\alpha+1)}\chi_\alpha(x)$, (35) becomes an equation for ψ with symmetric kernel. For any fixed $\alpha > -1$ there is, hence, a smallest value of λ (call it λ_1) for which the equation has a solution, as long as its kernel does not vanish everywhere. (The condition (26) insures this.)

As α ranges from -1 to ∞, λ_1 will range from 0 to ∞. The kernel of (35) is monotonically decreasing in α. Hence for a fixed λ, $0 < \lambda < \infty$, there will be a largest value of α (call it α_1) for which (35) has a solution. This solution is the function for which the expression on the right-hand side of (45) actually attains its minimum. If the solution were oscillatory we could construct a smaller such expression by substituting $|\psi|$ for ψ. Hence it is nonoscillatory and can be taken nonnegative.

We shall now prove the bound (29) for the solution to (35). First of all, $|\phi_\alpha(x)| < \infty$ for all x. This may be demonstrated using the fact that $||\psi_\alpha|| < \infty$, where $\psi_\alpha(x) = x^{\frac{1}{2}(\alpha+1)}\phi_\alpha(x)/(1+x)$, and applying the Schwarz inequality to (35). Now we can prove (29) inductively. Assume

$$\phi_\alpha(x) \leq M'/(1+x)^\beta, \qquad 0 \leq \beta \leq \alpha. \tag{A1}$$

Using the inequality derived from (35):

$$\phi_\alpha(x) \leq \frac{\lambda}{\alpha+1}\int_0^\infty \frac{dx'}{(1+x')^2} \times \left[\frac{2x'}{x+x'+\mu^2}\right]^{\alpha+1}\phi_\alpha(x), \tag{A2}$$

it is easy to show

$$\phi_\alpha(x) \leq M''/(1+x)^{\beta+1-\epsilon},$$

where $\epsilon > 0$. A final application of (A2) then gives the desired result (29).

APPENDIX B. $I(1,z)$

In this Appendix we perform the integral (76) for $\mu = 1$. The following change of variables is convenient: define u and ω by

$$\cosh\omega = \tfrac{1}{2}[(y/y')^{\frac{1}{2}} + (y'/y)^{\frac{1}{2}}], \tag{B1}$$

$$\cosh u = \tfrac{1}{2}[(y/y')^{\frac{1}{2}} + (y'/y)^{\frac{1}{2}}] + \tfrac{1}{2}(yy')^{\frac{1}{2}}, \tag{B2}$$

so that (76) becomes

$$I(1,z) = 16\int_0^\infty du \sinh u e^{-2zu} J(u), \tag{B3}$$

where

$$J(u) = \int_0^u d\omega \frac{\cosh u - \cosh \omega}{(1 - 4 \cosh u \cosh \omega + 4 \cosh^2 u)^2}. \tag{B4}$$

It can be shown that

$$J(u) = \tfrac{1}{3} \sinh^3 u (d^2/du^2)(u/\{1 - e^{-6u}\}) \tag{B5}$$

so that, integrating (B3) twice by parts, noting that the boundary terms vanish, and substituting $v = 2u$, we obtain

$$I(1, z) = \frac{1}{3} \int_0^\infty \frac{v\, dv}{1 - e^{-3v}} \times [(z - 2)^2 e^{-(z-2)v} - 4(z - 1)^2 e^{-(z-1)v} + 6z^2 e^{-zv} - 4(z + 1)^2 e^{-(z+1)v} + (z + 2)^2 e^{-(z+2)v}]. \tag{B6}$$

Now

$$\int_0^\infty \frac{t^n e^{-zt}\, dt}{1 - e^{-t}} = (-1)^{n+1} \psi^{(n)}(z), \tag{B7}$$

where

$$\psi^{(n)}(z) = \frac{d^{n+1}}{dz^{n+1}} \log \Gamma(z) = (-1)^n n! \sum_{m=0}^{\infty} \frac{1}{(z + m)^{n+1}} \tag{B8}$$

(see Ref. 13, Chap. 1), so, finally,

$$I(1, z) = (1/27)[(z - 2)^2 \psi'(\tfrac{1}{3}\{z - 2\}) - 4(z - 1)^2 \psi'(\tfrac{1}{3}\{z - 1\}) + 6z^2 \psi'(\tfrac{1}{3}\{z\}) - 4(z + 1)^2 \psi'(\tfrac{1}{3}\{z + 1\}) + (z + 2)^2 \psi'(\tfrac{1}{3}\{z + 2\})]. \tag{B9}$$

[13] A. Erdelyi *et al.*, Op. cit., Vol. 1.

Using the identity

$$\psi'(1 + z) = \psi'(z) - z^{-2},$$

this can be rewritten as

$$I(1, z) = 1 + \tfrac{1}{9} z [2z\psi'(\tfrac{1}{3}\{z + 3\}) - (z + 4)\psi'(\tfrac{1}{3}\{z + 1\}) - (z - 4)\psi'(\tfrac{1}{3}\{z + 2\})] \tag{B10}$$

[see Eq. (80)], which shows that $I(1, z)$ is free of singularities for $z > -1$, and hence can be expanded in a Taylor's series about $z = 0$.

APPENDIX C. $I(0,z)$

Substitute $v = \mu^{-2}(yy')^{\frac{1}{2}}$,

$$\omega = \tfrac{1}{2} \log (y'/y) \text{ in (76).}$$

Then, letting $\mu \to 0$,

$$I(0, z) = 4 \int_0^\infty d\omega \int_0 v\, dv \times (1 + 2v \cosh \omega + v^2)^{-2} e^{-2z\omega}. \tag{C1}$$

Now

$$\int_0^\infty v\, dv\, (1 + 2v \cosh \omega + v^2)^{-2} = \frac{1}{2} \frac{d^2}{d\omega^2} \left[\frac{\omega}{1 - e^{-2\omega}}\right],$$

so integrating (C1) by parts,

$$I(0, z) = -1 - 2z + 8z^2 \int_0^\infty \frac{\omega e^{-2\omega z}\, d\omega}{1 - e^{-2\omega}} = -1 - 2z + 2z^2 \psi'(z),$$

or, using (B10),

$$I(0, z) = 1 - 2z + 2z^2 \psi'(1 + z). \tag{C2}$$

Properties of Quantum Statistical Expectation Values*

D. A. UHLENBROCK
The Institute for Advanced Study, Princeton, New Jersey
(Received 26 May 1965)

The semigroup of statistical operators, and the unitary group of time translation operators generated by the same Hamiltonian of a nonrelativistic fermion field, are naturally imbedded in a holomorphic half-plane semigroup. The statistical expectation values of products of time-dependent operators are then boundary values of holomorphic functions which allow a Bochner integral representation with a Cauchy kernel. The general time–temperature-dependent Green's functions permit a concise spectral representation. It is suggested that a thermodynamic perturbation theory should treat the Heisenberg and Bloch equations simultaneously in terms of a perturbation theory for holomorphic half-plane semigroups generated by semibounded self-adjoint Hamiltonians.

1. INTRODUCTION

THE traditional approach to quantum statistical mechanics is based on the unique canonical quantization of classical Hamiltonians for systems with finitely many degrees of freedom together with the ensemble averaging in terms of traces involving a statistical operator.[1] The core of the problem lies in establishing the existence of a thermodynamic limit (such as $N/V = \text{const}$, $V \to \infty$, $N =$ number of degrees of freedom, $V =$ volume) and its evaluation for the quantities of interest.[2]

On the other hand, inspired by relativistic quantum field theory, attempts have been made to construct what might be called a nonrelativistic statistical quantum field theory, which from the very start works with systems of infinitely many degrees of freedom. The quantization is achieved in terms of a representation of the canonical commutation relations (CCR) for bosons or anticommutation relations (CAR) for fermions. It has been shown in several detailed studies of the commutation relations that the uniqueness up to unitary equivalence is lost for systems with infinitely many degrees of freedom.[3] Thus the problem arises of constructing the appropriate representation for a given system. Cyclic representations of the CCR for the infinite free Bose gas were treated by Araki and Woods,[4] while Araki and Wyss[5] dealt with the CAR. We shall be concerned with the CAR. According to the last-named authors their cyclic representations are equivalent to cyclic *-representations of the abstract CAR-algebra $\mathfrak{A}_C(\mathfrak{R})$ in the algebra $B(\mathfrak{K})$ of all bounded linear operators on a separable complex (H)-space $\mathfrak{K}$. Such representations can be characterized up to unitary equivalence by positive linear functionals on $\mathfrak{A}_C(\mathfrak{R})$ to the complex numbers C. The reconstruction proceeds via the Gel'fand construction which due to the special nature of the CAR still leads to bounded operators even though $\mathfrak{A}_C(\mathfrak{R})$ is not necessarily a complete normed algebra. If $\mathcal{O}$ is some open connected subset of l-dimensional ($l = 3$ for cases of physical interest) Euclidean space, then $\mathfrak{R}$ could, for example, be realized by the Schwartz test function-spaces[6] $\mathfrak{R} = S(\mathcal{O}) \oplus S(\mathcal{O})$ (two components of spin). If $\mathcal{O} = R^l$ a representation α: $\mathfrak{A}_C(\mathfrak{R}) \to B(\mathfrak{K})$ gives rise to a representation of the CAR as operator-valued tempered distributions in the space arguments if the appropriate continuity is ensured. Certain invariance groups (translations, rotations, etc.) are then also easily incorporated in the representation by defining the action of these groups on $\mathfrak{R}$ and hence on $\mathfrak{A}_C(\mathfrak{R})$ in the obvious fashion and requiring the representation functional to be invariant with respect to these transformations. Furthermore, a charge concept can be introduced via an involution on $\mathfrak{R}$ (i.e., complex conjugation of test functions) and leads to the "representations with charge conservation," which are

* Supported by the National Science Foundation.

[1] The uniqueness up to unitary equivalence of the operator representation of the canonical commutation relations (CCR) in the Weyl form is a well-known result due to J. von Neumann, Math. Ann. 104, 570 (1931), who also championed the statistical operator [or density matrix; see for example his book, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, New Jersey, 1955)].

[2] For recent results see L. van Hove, Physica 15, 951 (1949); C. N. Yang and T. D. Lee, Phys. Rev. 87, 404 (1952); also D. Ruelle, Theoretical Physics Institute lecture notes, summer 1963, University of Colorado, Boulder, Colorado; and J. Ginibre, J. Math. Phys. 6, 238, 252 (1965).

[3] L. Gårding and A. S. Wightman, Proc. Natl. Acad. Sci. U. S. 40, 6, 622 (1954); R. Haag, Kgl. Danske Videnskab. Selskab, Mat.-Fys. Medd. 29, No. 12 (1955); I. E. Segal, Trans. Am. Math. Soc. 88, 12 (1958); H. Araki, J. Math. Phys. 1, 492 (1960), where also other references are given.

[4] H. Araki and E. J. Woods, J. Math. Phys. 4, 637 (1963).

[5] H. Araki and W. Wyss, Helv. Phys. Acta 37, 136 (1964). See also D. Shale and W. Forrest Stinespring, Ann. Math. 80, 365 (1964).

[6] L. Schwartz, *Théorie des Distributions* (Hermann & Cie., Paris, 1957/59).

equivalent to the cyclic *-representations of the algebra of "observables" $\mathfrak{A}_0(\mathfrak{R})$. The latter is *-isomorphic to the zero charge part of $\mathfrak{A}_C(\mathfrak{R})$ and its representations are again characterized by positive linear functionals. We assume given a representation of the CAR for $\mathfrak{O}$ not necessarily unbounded or equal to R^l. The dynamics shall be specified by a self-adjoint lower semibounded linear operator H with domain $D(H) \subset \mathfrak{K}$. We are not concerned with the problem of whether and how this "Hamiltonian" H can be expressed as a "function" of the CAR operators.[7] H will characterize the equilibrium distribution in terms of the statistical operator $W(\beta) = \exp(-\beta H)$, β = inverse temperature, as well as the unitary time translation operators $U(t) = \exp(itH)$.

In Sec. 2 we deal with the relation between $W(\beta)$ and $U(t)$ by means of the theory of operator semigroups holomorphic in a half-plane. In Sec. 3 we pursue the restrictions on H due to the requirement that $W(\beta)$ be an operator in the trace class.[8] Also statistical expectation values (or correlation functions) of time-dependent operators in $\mathfrak{A}_C(\mathfrak{R})$ [we do not distinguish notationally between $\mathfrak{A}_C(\mathfrak{R})$ and its given representation] turn out to be boundary values of functions holomorphic in a very simple tubular domain. A formula identical with Bochner's formula for square integrable boundary values is given which relates the values of a certain class of holomorphic functions to their values on the distinguished boundary of the tube by means of an integral with a simple Cauchy kernel. In Sec. 4 we describe a general and concise spectral representation of the statistical Green's functions at finite temperatures and valid in the physical region of the time–temperature variables.

We observe that the simple relation between the time and temperature dependence of the statistical expectation values (expressed in terms of half-plane semigroups) has been obscured by the somewhat uncritical use of the formal analogy between the Heisenberg equation of motion and the Bloch equation (the analogy is of the same limited character as that between the "classical" Schrödinger equation and the heat equation). The simple fact is that the difference of a factor i $[=(-1)^{\frac{1}{2}}]$ induces all of the difference between a (strongly) continuous unitary group and a self-adjoint semigroup restricted to the positive real axis. The above formal analogy has been used in the thermodynamic perturbation theory of the Bloch equation which emulates the methods of the Dyson–Feynman–Schwinger perturbation theory for the Heisenberg equation of motion. In particular, that perturbation theory is based on the introduction of temperature-dependent quantities,

$$a(\beta) = e^{\beta H} a(0) e^{-\beta H},$$

where $a(0)$ is some (fermion)-operator and β a real number (the inverse temperature). While such a transformation is quite appropriate when e^{iHt} is used instead of $e^{\beta H}$ it is dangerous here, since for physical positive β the operator $e^{\beta H}$ is unbounded together with H. As a consequence, the above transformation in no way represents an equivalence transformation, as it maps bounded operators into unbounded ones in general.

It is our contention that it is undesirable to work with the Bloch and Heisenberg equations (for the same H) independently. Instead of rotating the time variables into temperature variables, treating the resulting pure Bloch problem with the aid of a perturbation theory based on the above transformation, and returning to real values of the time variables by analytic continuation in the temperature variables, one should treat both equations simultaneously. This is very natural and simple since both the Heisenberg and the Bloch equations are but two aspects of the same mathematical object: the holomorphic half-plane semigroup generated by H. While the Bloch equation describes the behavior of this semigroup along the real "spine" in terms of the self-adjoint semigroup $e^{-\beta H}$, $\beta > 0$, the Heisenberg equation describes its behavior along the boundary in terms of the unitary "boundary" group e^{iHt}, t real. The perturbation theory is thus naturally based on the unique canonical representation of this half-plane semigroup as a Laplace–Stieltjes integral over the projection valued measure belonging to H. *Most of our notational conventions are collected in the Appendix.*

2. STATISTICAL SEMIGROUP

Given any positive self-adjoint $W \in B(\mathfrak{K})$, then for $\epsilon \in R_+$, the resolvent $R(\lambda; W)$ is holomorphic in

[7] F. Coester and R. Haag, Phys. Rev. 117, 1137 (1960); H. Araki in Ref. 3. It is worth noting in this connection that in the treatment of the infinite free Bose gas in Ref. 4 the unitary time translation operators $U(t) = \exp(iHt)$ were found *not* to belong to the von Neumann algebra generated by all CCR operators in the Weyl form.

[8] Under these circumstances H must have a pure point spectrum of finite multiplicity, and if H is to be a Hamiltonian the corresponding system must generally have a finite space extension. There still remains the need to establish a thermodynamic limit for sequences of such systems, whose volume tends to infinity and whose limiting system is no longer describable by a density matrix of the trace class. Such a program was recently formulated and studied by Ruelle (D. Ruelle, Lectures at Institut d'Etudes Scientifiques, Cargese, Corsica, summer 1965).

the complement of the sector V: $0 \leq |\lambda| \leq ||W||$, $|\arg \lambda| \leq \epsilon$ such that $||R(\lambda; W)|| \leq (\sin \epsilon \cdot |\lambda|)^{-1}$ in $0 < |\lambda| < ||W||$, $|\arg \lambda| > \epsilon$. If ∂V is the positively oriented boundary of V, the formula

$$U(z) = \frac{1}{2\pi i} \int_{\partial V} R(\lambda; W)\lambda^{-iz}\, d\lambda, \qquad z \in C_+ \tag{2.1}$$

is seen to define a semigroup holomorphic in C_+ such that $U(in) = W^n$ for $n \in I_+$.[9] This illustrates that any candidate W for a statistical operator can always be imbedded in a holomorphic half-plane semigroup. It will be seen that the maximal domain of holomorphic existence cannot be made larger than a half-plane if, in particular, the infinitesimal generator is self-adjoint and bounded from below but not from above. Namely, if H is any such Hamiltonian, the (unnormalized) statistical operator for the (grand) canonical ensemble can be represented by means of the operational calculus for self-adjoint operators as an exponential "function" of H: $W(s) = e^{-sH}$ for $s \in R_+$. This exponential has the semigroup property $W(s + s') = W(s) \cdot W(s')$, $s, s' \in R_+$ and tends strongly to 1 as s tends to zero in R_+.

The following theorem states that this form of $W(s)$ is the most general form if only the set $\{W(s) \mid s \in R_+\}$ is required to form a self-adjoint semigroup satisfying some weak conditions.

Theorem: Let $S(R_+) \equiv \{W(s) \mid s \in R_+, W(s) = W(s)^*\} \subset B(\mathfrak{H})$ form a self-adjoint semigroup such that

(a) for some $\alpha \in R_+$

$$\sup_{\alpha < s \leq \max[\alpha+1, 2\alpha]} ||W(s)|| < \infty,$$

(b) for some $s_0 \in R_+$, $W(s_0)^{-1}$ is an unbounded linear operator, then

(c) $h \equiv -\lim_{s\to\infty}(s^{-1} \log ||W(s)||) > -\infty$, $||W(s)|| = \exp(-sh)$, $s \in R_+$,

(d) $s\text{-}\lim_{s\to 0_+} W(s) = 1$,

(e) $S(R_+)$ has a holomorphic extension $S(C_+) \equiv \{U(z) \mid z \in C_+\}$ with maximal domain of holomorphic existence C_+,

(f) $S(R_+)$ has a self-adjoint infinitesimal generator $-H$ with domain $D(H)$ and associated resolution of the identity $E(\lambda)$:

$$D(H) = \left\{\varphi \;\middle|\; \int_h^\infty \lambda^2\, d\, ||E(\lambda)\varphi||^2\right\} \subset \mathfrak{H},$$
$$H\varphi = \int_h^\infty \lambda\, dE(\lambda)\varphi \quad \text{for} \quad \varphi \in D(H). \tag{2.2}$$

Also $U(z)$, $z = t + is\epsilon C_+$ is uniquely represented as

$$U(z) = \int_h^\infty e^{i\lambda z}\, dE(\lambda) \tag{2.3}$$

in the sense of strong operator convergence.

Remarks:

(i) Condition (a) can be replaced by (a') $S(R_+)$ is weakly measurable;

(ii) if in (b) $W(s_0)^{-1}$ is assumed to exist as a bounded linear operator, then $S(R_+)$ will have a holomorphic extension to an analytical group on all of C and H will be bounded. Of course, the usual case of physical interest is the one where H is unbounded.

(iii) The condition (d) means that $S(R_+)$ is of class (C_0) in the sense of (HP). Also the transformation $W(s) \to \exp(s \cdot h)W(s)$ does not change the class of the semigroup and is equivalent to the transformation of the generator $H \to H - h1$. Thus the transformed $H - h1$ is nonnegative and $h1 - H$ is "dissipative". In the following, we shall often assume that this transformation has been effected, whereupon $\{W(s)\}$ becomes a "contraction" semigroup.

(iv) The "converse" of the theorem is easily seen to hold in the sense that any self-adjoint linear (unbounded) operator H with lower bound h will give rise to a semigroup defined by Eq. (2.3) with the properties appearing in the theorem.

The proof of this theorem results from an adaptation of (HP 22.3) to our purposes and notation. The infinitesimal generator $-H$ is defined through

$$D(H) = \{\varphi \mid \lim_{s\to 0+} [s^{-1}(1 - W(s))]\varphi \text{ exists}\} \subset \mathfrak{H}$$
$$H\varphi = \lim_{s\to 0+} [s^{-1}(1 - W(s))]\varphi \quad \text{for} \quad \varphi \in D(H). \tag{2.4}$$

Condition (b) insures that $\mathfrak{H}_0 \equiv \bigcup_{s\in R_+}(W(s)\mathfrak{H})$ is dense in $\mathfrak{H}$ and hence implies (d). By (HP 16.7) the spectral classification of the spectral value $\lambda = 0$ is identical for all $s \in R_+$. Therefore $0 \notin \sigma_p(W(s))$ for $s \in R_+$ and $W(s)^{-1}$ exists. Due to (b), $S(R_+)$ cannot be imbedded in a strongly continuous group on all of R; therefore $0 \in \sigma_c(W(s))$ and $D(W(s)^{-1})$ is dense since $0 \notin \sigma_p(W(s)) \cup \rho(W(s))$. Then by (HP 11.4) $W(s)$ cannot tend to 1 uniformly as $s \to 0_+$ and H is unbounded. It follows that the maximal

[9] The general reference for semigroups is E. Hille and F. S. Phillips, *Functional Analysis and Semigroups*, second edition, Am. Math. Soc. Coll. Publ. 31, 1957. It will, in the following, be referred to as (HP a–b), where a–b denotes the particular subsection in question. The above embedding statement is easily proved with the aid of the resolvent identity and Cauchy's formula and is a special case of (HP 17–7).

domain of analytical existence of $U(z)$ is C_+ and not C.

It appears therefore that the semigroup property together with the mild restrictions (a), (b) already leads to a type of statistical operator characteristic of a canonical ensemble, where the variable s is identified as an inverse temperature.

We briefly quote some of the relevant results from (HP 10.3, 6; 17.5), where it is shown that $D(H) \supset \mathfrak{H}_0$ is the countable union of nowhere dense sets in $\mathfrak{H}$, while $\bigcap_{n\in I_+} D(H^n)$ is still dense in $\mathfrak{H}$. Also $U(z)$ is uniformly continuous and uniformly differentiable on C_+, uniformly on compact subsets, such that

$$-(d/dz)[U(z)\varphi] = HU(t)\varphi = U(z)H\varphi \text{ for } \varphi \in D(H);$$

and for

$$\varphi \in \bigcap_{n\in I_+} D(H^n)$$

the series

$$\sum_{n\in I_+} \frac{z^n}{n!}(-H)^k\varphi$$

is asymptotic to $U(z)\varphi$ for $\epsilon < \arg z < \pi - \epsilon$, $\epsilon > 0$ in the sense that

$$\lim_{|z|\to 0} z^{-n}\left[U(z)\varphi - \sum_{k\in I_{n-1}} \frac{z^k}{k!}(-H)^k\varphi\right] = \frac{1}{n!}(-H)^n\varphi$$

$$\text{for} \quad n \in I_+.$$

There is a number of other formulas which also reflect one or another of the properties of the exponential function.

Since $||U(z)\varphi||^2 \leq e^{-2hs}\int_h^\infty d\,||E(\lambda)\varphi||^2$ for $z = t + is$, it follows that $||U(z)||$ is bounded for $|t| < 1$ and $0 < s \leq 1$. Then, by (HP 17.9), the set $T(R) \equiv \{U(t)\ |t \in R\}$ defined by $U(t) \equiv s\text{-lim}_{s\to 0_+} U(z)$ forms the strongly continuous unitary "boundary" group of the semigroup $S(C_+)$ with the infinitesimal generator iH. $T(R)$ commutes elementwise with $S(C_+)$ and $U(z) = W(s)\cdot U(t)$. Of course, these statements can be derived immediately from the representation (2.3) of $U(z)$ and agree with Stone's theorem concerning the integral representation of strongly continuous unitary groups.

The relationship between the statistical semigroup $S(R_+)$ and the unitary group of time translations $T(R)$ is thus briefly this: They are generated respectively by $-H = i(iH)$ and iH and are both related to the same half-plane semigroup $S(C_+)$, the former forming the self-adjoint "spine" semigroup while the latter gives the unitary boundary group. We note that the approach to the boundary is continuous in the strong but not in the uniform topology of $B(\mathfrak{H})$ and that every element of $S(C_+)$ has an inverse which is densely defined but unbounded. $S(C_+)$ is maximal in this sense.

3. STATISTICAL EXPECTATION VALUES

The statistical expectation value (SEV) of $a \in B(\mathfrak{H})$ is obtained by

$$\langle a\rangle = \text{Tr}\,(Wa)/\text{Tr}\,W,$$

where W is any (unnormalized) statistical operator. If for some $z_0 \in C_+$, $W = U(z_0)$ is in the trace class $\mathfrak{C}_1$ (Appendix) then it follows with (2.3) that $U(z) \in \mathfrak{C}_1$ for all $z \in C_+$. We assume $U(z) \in \mathfrak{C}_1$ from now on. Then $U(z)$ is compact and its spectrum is discrete with finite multiplicity away from zero:

$$U(z) = \sum_{j\in I_+} e^{i\lambda_j z}E_j,$$

$\lambda_j \geq h = 0$, $\lambda_j > \lambda_{j'}$ for $j > j'$; $\dim E_j = TrE_j \equiv \nu_j < \infty$. It follows from the monotonicity and continuity of the exponential function together with the fact that the eigenvalues of $\mathfrak{U}(z)$ accumulate only at zero that the eigenvalues λ_j of H accumulate only at $+\infty$.

$T(R)$ induces the obvious group of *-automorphisms on $B(\mathfrak{H})$ by

$$a(t) \equiv U(t)aU(t)^{-1}, \quad a \in B(\mathfrak{H}), \quad U(t) \in T(R). \tag{3.1}$$

This can be regarded as the boundary behavior of the transformation

$$a(z) = U(z)aU(z)^*, \qquad z \in \bar{C}_+. \tag{3.2}$$

We note that for $z = i\beta$ (3.2) reduces to

$$a(i\beta) = W(\beta)aW(\beta) = e^{-\beta H}ae^{-\beta H} \neq e^{\beta H}ae^{-\beta H},$$

and that the alternative to (3.2), $a(z) = U(z)aU(z)^{-1}$, which also agrees with (3.1) on the boundary of $\bar{C}_+$, is unacceptable, since $U(z)^{-1}$ is unbounded for $z \in C_+$. *A priori* it is not clear that the transformation (3.1) restricted to $a \in \mathfrak{A}_C(\mathfrak{R})$ will be a *-automorphism for $\mathfrak{A}_C(\mathfrak{R})$, since $a(t)$ might not even be in $\mathfrak{A}_C(\mathfrak{R})$. However, if H is a "function" of elements in $\mathfrak{A}_C(\mathfrak{R})$, [its spectral projectors are contained in $\mathfrak{A}_C(\mathfrak{R})$] this will be the case. At any rate the transformation (3.2) does not have this property for $z \in C_+$. The SEV of interest are of the form ($n \in I_+$ fixed)

$$\text{Tr}\left[W(\beta)\prod_{i\in I_n}^{\rightarrow} a_{j_i}(l_i)\right], \quad \beta \in R_+, \quad l_i \in R,$$

$$a_{j_i} \in \mathfrak{A}_C(\mathfrak{R}) \quad \text{for} \quad i \in I_n. \tag{3.4}$$

Due to the cyclic invariance of the trace, Eq. (3.1), (3.2), and the semigroup property of $U(z_i)$, these expressions are recognized as special boundary values of functions defined in the tubular domain

$$\mathfrak{T}_n \equiv \{z \in C^n \mid z = t + is,\ t \in R^n, s \in R_+^n\} = (C_+)^n:$$

$$\mathrm{Tr}\,[U(\bar{l}_1 - \bar{l}_n + i\beta)a_{j_1} \cdots U(\bar{l}_n - \bar{l}_{n-1})a_{j_n}] \equiv V(\beta, j, t) \equiv V(j, z), \tag{3.5}$$

where z lies in the "physical region" $\mathcal{P}_n(\beta)$ of $\mathfrak{T}_n$ defined by

$$\begin{aligned} z_1 &= t_1 + is_1 = (\bar{l}_1 - \bar{l}_n) + i\beta, \\ z_2 &= t_2 + is_2 = \bar{l}_2 - \bar{l}_1, \cdots, \\ z_n &= t_n + is_n = \bar{l}_n - \bar{l}_{n-1}, \end{aligned} \tag{3.6}$$

i.e., $\mathcal{P}_n(\beta)$ is the $(n-1)$-dimensional submanifold of C^n on the boundary of $\mathfrak{T}_n$ characterized by

$$\begin{aligned} &\text{(i)} \quad \sum_{i \in I_n} \mathrm{Re}\, z_i = 0 \\ &\text{(ii)} \quad \mathrm{Im}\, z_1 = \beta, \quad \mathrm{Im}\, z_i = 0 \quad \text{for} \quad i \in I_n'. \end{aligned} \tag{3.7}$$

The tube $\mathfrak{T}_n$ is a domain of holomorphy for $V(j, z)$ as a function of z, since its base R_+^n is convex and open. Furthermore, in order to show holomorphy of $V(j, z)$ in $\mathfrak{T}_n$, it suffices (due to Hartogs' theorem) to show holomorphy in each variable z_i, $i \in I_n$ separately.[10] Again, due to the cyclic invariance of the trace, it is enough to show this for expressions like $\mathrm{Tr}\,[U(z_1)A]$ for arbitrary $A \in B(\mathfrak{H})$, $z_1 \in \mathfrak{T}_1 = C_+$.

$$\mathrm{Tr}\,[U(z_1)A] = \sum_{i \in I_+} e^{i\lambda_i \cdot z_1} A_i,$$

where $A_i = \mathrm{Tr}\,(E_i A)$ is absolutely convergent for fixed $z_1 \in C_+$. For any compact subset D of C_+ the above trace is absolutely and uniformly over D, majorized by

$$\sum_{i \in I_+} \exp\,[-\lambda_i\, \mathrm{Im}\, z(D)]\ |A_i| < \infty,$$

where $z(D)$ is any $z \in D$ satisfying $\mathrm{Im}\, z(D) = \inf_{z \in D} (\mathrm{Im}\, z) > 0$. This proves that $V(j, z)$ is holomorphic in $\mathfrak{T}_n$. Also it is clear that

$$U(z)A = s\text{-}\lim_{n \to \infty} \sum_{i \in I_n} e^{i\lambda_i z} E_i A,$$

and $\mathrm{Tr}\,[U(z)A] < \infty$ implies the convergence of the expression for $U(z)A$ in the trace norm and hence in the operator norm. Then, upon interchange of summation, product, and trace, $V(j, z)$ can be written as the Fourier–Laplace transform of a distribution:

$$V(j, z) = \mathrm{Tr}\left[\sum_{k \in (I_+)^n} e^{iz \cdot \lambda_k} \vec{\prod_{i \in I_n}} (E_{k_i} a_{j_i})\right] = \int_{\lambda \in R^n} d\lambda\, e^{iz \cdot \lambda} V(j, \lambda), \tag{3.8}$$

$$V(j, \lambda) = \sum_{k \in (I_+)^n} \delta_n(\lambda - \lambda_k) \cdot \mathrm{Tr} \vec{\prod_{i \in I_n}} (E_{k_i} a_{j_i}) \in \mathfrak{D}'_{(\lambda)}$$

such that

$$\mathrm{supp}\ V(j, \lambda) = \bigcup_{k \in (I_+)^n} \lambda_k \subset (\bar{R}_+)^n$$

and

$$\delta_n(\lambda - \lambda_k) = \bigotimes_{i \in I_n} \delta(\lambda_i - \lambda_{k_i}).$$

We note the estimates

$$\left|\mathrm{Tr} \vec{\prod_{i \in I_n}} (E_{k_i} a_{j_i})\right| \leq |E_{k_1}|_1 \cdot \left\| a_{j_1} \vec{\prod_{i \in I_n'}} (E_{k_i} a_{j_i}) \right\| \leq \nu_{k_1} \cdot \prod_{i \in I_n} \|a_{j_i}\| < \infty,$$

and therefore also

$$\left|\mathrm{Tr} \vec{\prod_{i \in I_n}} (E_{k_i} a_{j_i})\right| \leq \alpha \cdot \min_{i \in I_n} \nu_{k_i}, \quad \text{with} \quad \alpha = \prod_{i \in I_n} \|a_{j_i}\|.$$

We shall assume that the spacing between the point eigenvalues λ_k of H and their multiplicities ν_k are restricted such that $V(i, \lambda) \in \mathcal{S}'_{(\lambda)}$. There are a number of criteria for determining whether an element of $\mathfrak{D}'$ is tempered,[6] but we will not pursue this here.

Since the base R_+^n of $\mathfrak{T}_n$ is connected and convex as well as selfconjugate and supp $V(j, \lambda) \subset (\bar{R}_+)^n$ Vladmirinov's theorem (see Appendix) gives the result that $V(j, z) \in H_1(0, R_+^n)$, and that

$$V(j, t) = \lim_{s \to 0,\, s \in R_+^n} V(j, t + is)$$

exists as the unique boundary value in $\mathcal{S}'$ as s goes to zero (nontangentially) in R_+^n. Here $H_1(0; R_+^n)$ is the class of functions $f(z)$ holomorphic in $\mathfrak{T}_n$, such that for every compact subcone $\mathcal{C}$ of R_+^n there exist $M(\mathcal{C}) < \infty$ and $\mathcal{C}$-independent $\alpha > 0$, $\beta > 0$ for which

$$|f(z)| \leq M(\mathcal{C})(1 + |z|)^{\beta}(1 + |s|^{-\alpha}) \quad \text{for} \quad z \in \mathfrak{T}^{\mathcal{C}}.$$

Also $f \in H_1(0; R_+^n)$ implies that $D^{\alpha} f$, af^p, $a \in C$, and $p \geq 1$ are in $H_1(0; R_+^n)$.

Let $\chi_{\bar{R}_+^n}(\lambda)$ be the characteristic function of $\bar{R}_+^n$ in R^n. $\chi_{\bar{R}_+^n}(\lambda) \in \mathcal{S}'$ and, since

$$\mathrm{supp}\,[V(j, \lambda)e^{-\lambda s}] \subset \bar{R}_+^n,$$

the two distributions can be multiplied to give a distribution in $\mathcal{S}'$ such that

$$\begin{aligned} V(j, z) &= \mathfrak{F}[(V(j, \lambda)e^{-\lambda s}) \cdot (\chi_{\bar{R}_+^n}(\lambda)); t] \\ &= \mathfrak{F}[(V(j, \lambda))(e^{-\lambda s}\chi_{\bar{R}_+^n}(\lambda)); t] \\ &= \{\mathfrak{F}[V(j, \lambda)] * \mathfrak{F}[e^{-\lambda s}\chi_{\bar{R}_+^n}(\lambda)]\}(t). \end{aligned}$$

[10] S. Bochner and W. T. Martin, *Several Complex Variables*, (Princeton University Press, Princeton, New Jersey, 1948).

Since

$$\mathfrak{F}[e^{-\lambda s}\chi_{\bar{R}_+^n}(\lambda); t] = i^n(\prod_{i\in I_n} z_i)^{-1},$$

by easy direct evaluation the folding formula gives the result

$$V(j, z) = \left(\frac{i}{2\pi}\right)^n \int_{R^n} \frac{V(j, l + i0)\, dl}{\prod_{i\in I_n}(z_i - l_i)}, \qquad (3.9)$$

when the integral is extended over the appropriately oriented distinguished boundary R^n of the tube $\mathfrak{I}_n$. This integral representation agrees with Bochner's formula[11] derived for $f(z)$ such that

$$\int_{R^n} dt\, |f(t + is)|^2 \leq \text{constant (of } s).$$

4. SPECTRAL REPRESENTATIONS

Since $V(j, z) \in H_1(0; R_+^n)$, clearly it is in $\mathcal{O}_M(t)$ for fixed $s \in R_+^n$, and it can be multiplied with the distribution (see appendix) $\eta_\sigma(z) \in \mathcal{S}'_{(t)}$ to give

$$V_\sigma(j, z) \equiv \eta_\sigma(z) \cdot V(j, z) \in \mathcal{S}'_{(t)}. \qquad (4.1)$$

Then

$$V_\sigma(j, \lambda, s) \equiv \mathfrak{F}^{-1}[V_\sigma(j, z); \lambda]$$
$$= (2\pi i)^{1-n} \int d\bar{\lambda}\, [\prod_{i\in I_{n'}} (\bar{\lambda}_i - \lambda_i - i\epsilon_i \cdot \sigma_i)^{-1} V(j, \bar{\lambda}) e^{-s\bar{\lambda}}|_{\substack{\bar{\lambda}_i = \lambda_i \\ \epsilon_i \to 0+}} \qquad (4.2)$$

is seen to be the particular boundary value characterized by the sequence σ of a function $V(j, \lambda, s)$, which, due to the support properties of $V(j, \lambda)$, is holomorphic for $\lambda' = [\lambda_2 \cdots \lambda_n]$ in the Cartesian product of $(n - 1)$ complex planes cut along the positive real axes. It is clear that the weight function $V(j, \lambda)e^{-s}$ can be reconstructed from the discontinuities across the cuts by the formula

$$V(j, \lambda)e^{-s\lambda} = \lim_{\epsilon_i \to 0+} \sum_{\substack{\sigma_i \\ i\in I_{n'}}} (\prod_{i\in I_{n'}} \sigma_i) \times V(j, \lambda + i\sigma\epsilon, s)|_{\epsilon_1 = 0}. \qquad (4.3)$$

In this way all the inverse Fourier transforms of the 2^{n-1} mixed—advanced—retarded functions $V_\sigma(j, z)$ are boundary values of one and the same holomorphic function of λ'

$$V(j, \lambda, s), \quad \lambda' \in (C_{\text{cut}})^{n-1}.$$

Similarly, one can arrive at a spectral representation for the (unnormalized) time-ordered SEV or Green's functions in the physical region $\mathcal{P}(\beta)$.[12] Let the a_{j_i}, $i \in I_n$ now be elements of $\mathfrak{A}_C(\mathfrak{R})$ which anticommute. $\epsilon_\pi = \pm 1$ is the one-dimensional antisymmetric representation of γ_n. It is seen from (3.5) and (3.7) that, in $\mathcal{P}_n(\beta)$, $V(j, z)$ depends only on $t' = [t_2, \cdots, t_n] = [\bar{t}_2 - \bar{t}_1, \cdots, \bar{t}_n - \bar{t}_{n-1}]$. The Green's functions are then defined in $\mathcal{P}_n(\beta)$ through (T is the time-ordering operator)

$$G(\beta, j, t') = \mathrm{Tr}\,[W(\beta)T(a_{j_1}(\bar{t}_1) \cdots a_{j_n}(\bar{t}_n))]$$
$$= \int dt_1\, \delta(\sum_{i\in I_n} t_i) \sum_{\pi\in\gamma_n} \epsilon_\pi \pi[\eta(t) \cdot V(\beta, j, t)]$$
$$= \mathfrak{F}'[G(\beta, j, \lambda'); t']. \qquad (4.4)$$

If one uses the cyclic invariance of the trace $(\pi_c^{k-1}V)(j, \lambda) = V(j, \lambda)$, and introduces the quantities

$$g_\epsilon(\beta, \lambda, \bar{\lambda}) \equiv \sum_{k\in I_n} ((-1)^k 2\pi i)^{1-n} e^{-\beta\bar{\lambda}_k} \bigotimes_{\substack{j\in I_n \\ j\neq k}} (\lambda_j - \lambda_k - \bar{\lambda}_j + \bar{\lambda}_k + i\epsilon_j)^{-1} \qquad (4.5)$$

$$V'(j, \bar{\lambda}) \equiv \sum_{\pi'\in\gamma_{n'}} \epsilon_{\pi'} \cdot (\pi' V)(j, \bar{\lambda}), \qquad (4.6)$$

then the Fourier transform in (4.4) can be inverted by

$$G(\beta, j, \lambda') = 2\pi \sum_{\pi\in\gamma_n} \epsilon_\pi \mathfrak{F}^{-1}[\delta(\Sigma t_i)\pi(\eta(t)\, V(\beta, j, t)); \lambda]|_{\lambda_1 = 0}$$
$$= \sum_{\pi\in\gamma_n} \epsilon_\pi \int_R da \int_{R^n} d\bar{\lambda}\, \{(-1)^{n-1}\, \delta(\lambda_{\alpha_1} - \bar{\lambda}_{\alpha_1} - a) e^{-\beta\bar{\lambda}_{\alpha_1}} \bigotimes_{\substack{i\in I_n \\ i\neq\alpha_1}} [(2\pi i)(\lambda_i - \bar{\lambda}_i - a + i\epsilon_i)]^{-1}\} (\pi V)(j, \bar{\lambda})_{\lambda_1 = 0}$$
$$= \sum_{\pi'\in\gamma_{n'}} \epsilon_{\pi'} \int_R da \int_{R^n} d\bar{\lambda}\, \{\sum_{k\in I_n} (-1)^{(n-1)k}\, \delta(\lambda_k - \bar{\lambda}_k - a) e^{-\beta\bar{\lambda}_k} \bigotimes_{\substack{i\in I_n \\ i\neq k}} [(2\pi i)(\lambda_i - \bar{\lambda}_i - a + i\epsilon_i)]^{-1}\} (\pi' V)(j, \bar{\lambda})|_{\lambda_1 = 0}$$

[11] S. Bochner, Ann. Math. 45, 686 (1944).

[12] For $n = 2$ this was first derived by L. D. Landau, J. Exptl. Theoret. Phys. (USSR) 34, 262 (1958) [Sov. Phys. J. Exptl. Theoret. Phys. 7, 182 (1958)]. For $n = 3$ this was written down in a formula with six terms by V. L. Bonch-Bruevich, Doklady Akad. Nauk (SSSR) 126, 539 (1959) [English transl.: Sov. Phys. Dokl. 4, 596 (1959)]. D. H. Kobe, Ann. Phys. 19, 448 (1962) treated the problem for zero temperature.

or finally

$$G(\beta, j, \lambda') = \lim_{\substack{\lambda_1=0 \\ \epsilon\to 0+}} \int_{R^n} d\bar{\lambda}\, g_\epsilon(\beta, \lambda, \bar{\lambda}) V'(j, \bar{\lambda}). \qquad (4.7)$$

Formula (4.7) gives a very concise and simple expression for $G(\beta, j, \lambda')$ in terms of all $V(j, \lambda)$. The kernel of the integral representation $g_\epsilon(\beta, \lambda, \bar{\lambda})$ is only symmetrically dependent on the components of λ and $\bar{\lambda}$ due to the summation over the n contributions from the cyclic invariance of the trace. $V'(j, \bar{\lambda})$ is the antisymmetric sum over those permutations of the original spectral function $V(j, \lambda)$ where only the last $(n-1)$ arguments are permuted.

CONCLUSION

Correlation functions (SEV) and, in particular Green's functions figure prominently in contemporary treatments of the many-body problem. For zero-temperature systems the latter occur naturally in the perturbation theory à la Dyson–Feynman.[13] A variant of this technique based on introducing the temperature (imaginary time) dependence in the "interaction picture" of the Bloch equation through

$$a(\beta) = e^{\beta H} a e^{-\beta H}, \quad \beta \in R_+, \quad a \in \mathfrak{A}_c(\mathfrak{R}) \qquad (4.8)$$

has been proposed by Matsubura.[14] The underlying idea of this scheme is to solve the resulting equations for the purely temperature-dependent quantities and to obtain information about these quantities for physical values of the arguments via analytic continuation in the temperature variables justified and guided by the spectral representations.

As pointed out in the introduction Eq. (4.8) is incompatible with our point of view of characterizing the relationship between the Bloch and Heisenberg equations for the same H in terms of the holomorphic half-plane semigroup generated by H. It is clear that Eq. (4.8) is harmless for *bounded* self-adjoint H (in which case the semigroup can be extended to an entire analytical group) but that $e^{\beta H}$ and therefore $a(\beta)$ is unbounded if H is unbounded. As a consequence the semigroup property is lost in general, since for example $e^{-\beta H} \cdot e^{\beta' H} \neq e^{-(\beta-\beta')H}$ if $\beta > \beta' > 0$ because the left-hand side is an unbounded operator while the right-hand side is not. Furthermore if one wishes to work meaningfully with SEV's depending on the temperature through Eq. (4.8) one is forced to adopt a modified definition of the trace. On the one hand, one has to extend the trace concept of the previous sections so as to have it also defined for a certain class of unbounded operators. This is achieved by means of the definition in terms of the sum of diagonal matrix-elements with respect to a suitable complete orthonormal basis in $\mathfrak{H}$. Then on the other hand, since for a given unbounded (densely defined) linear operator on $\mathfrak{H}$ there always exist normalized vectors not in the domain of the operator (there exist even complete orthonormal basis systems none of whose elements is in the domain of the operator) it is clear that the trace definition has to be restricted to such basis systems for which all elements are in the domain of the operator. This means in particular that $\mathrm{Tr}\,(UAU^{-1}) = \mathrm{Tr}\,A$ does not hold in general (in a fixed basis system) for all unitary U but only for those U which do not violate the above condition. Similar limitations apply to the cyclic invariance of the trace of products of operators when evaluated in a fixed basis system. It follows that the SEV's defined with the more restricted trace [of the trace class in $B(\mathfrak{H})$] cannot be rendered meaningful outside of the tube $\mathfrak{T}_n$. In particular the values of the time–temperature variables reached by Eq. (4.8) cannot be connected by analytic continuation with the physical values on the boundary of $\mathfrak{T}_n$.

We note that the "thermodynamic perturbation theory" discussed above is not unique in the sense that there are methods within the framework of the perturbation theory of semigroups which are properly applied here. However, the results for general semigroups are not specific enough to be of much use in this connection.[15] The very special nature of the semigroup involved here, expressed partly by relation (2.3)

$$U(z) = \int_h^\infty e^{iz\lambda}\, dE(\lambda) \equiv e^{izH}, \quad z \in C_+,$$

implies, of course, that the perturbation theory of $U(z)$ is intimately related to that of the infinitesimal

[13] D. Pines, *The Many Body Problem* (W. A. Benjamin, Inc., New York, 1961); and the lucid exposition in P. Nozières, *Le problème à N corps* (Dunod Cie., Paris, 1963).

[14] The H in (4.8) is the unperturbed Hamiltonian. The basic reference and some more detailed and recent accounts of the method are T. Matsubara, Progr. Theoret. Phys. (Kyoto) **14**, 351 (1955); L. P. Kadanoff and G. Baym, *Quantum Statistical Mechanics*, (W. A. Benjamin, Inc., New York, 1962), see 1–2 and Appendix, who start from the work by P. C. Martin and J. Schwinger, Phys. Rev. **115**, 1342 (1959); A. A. Abrikosov, L. P. Gor'kov, and I. E. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics*, translated by R. Silverman (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1963) Ch. III; C. De Dominicis and P. C. Martin, J. Math. Phys. **5**, 14, 31 (1964).

[15] Thus the perturbation theory of Phillips gives an explicit expression for the perturbed semigroup in terms of an infinite sum of integrals, but the perturbed operator is only allowed to vary in a rather restricted class of operators. This is discussed in (HP) chapter XIII and E. Nelson, "Operator Differential Equations," lecture notes, Princeton University, 1964–65.

generator. The latter problem is well studied, though subtle in mathematical details, and can (under suitable conditions on the operators) be solved by a perturbation algorithm.

ACKNOWLEDGMENTS

The author gratefully acknowledges the comments of Dr. M. Guenin, Dr. H. Reeh, Professor J. Łopuszanski and a conversation with Professor C. N. Yang. He also enjoyed a conversation with Professor Oppenheimer and wishes to thank him for his generous hospitality at The Institute for Advanced Study.

APPENDIX

I is the ring of integers, I_+ the strictly positive subset, $\bar{I}_+ = \{0\} \cup I_+$, $I_n = \{1, 2, \cdots, n\}$ for $n \in I_+$ and $I_n' = I_n - \{1\}$. Similarly R is the field of reals, R_+ the strictly positive subset, $\bar{R}_+$ its closure. C is the field of complex numbers, C_+ the (strict) upper half-plane with closure $\bar{C}_+$.

$\mathfrak{H}$ is a separable complex Hilbert space with elements φ, χ, and inner product (φ, χ). $B(\mathfrak{H})$ is the algebra of all bounded linear operators from $\mathfrak{H}$ to $\mathfrak{H}$, normed with the operator norm $||a||$, unit element 1. For a linear operator A, $R(\lambda; A) = (\lambda 1 - A)^{-1}$ is the resolvent, $\rho(A)$ the resolvent set, $\sigma(A)$ the spectrum.

The set of all operators $W \in B(\mathfrak{H})$ for which the (basis independent) sum $\mathrm{Tr}\, W = \sum_{i \in I_+} (\varphi_i, W\varphi_i)$, with $\{\varphi_i \mid j \in I_+\}$ a complete orthonormal base in $\mathfrak{H}$, is finite and then absolutely convergent forms the two-sided *-ideal $\mathfrak{C}_1$ of trace class operators in $B(\mathfrak{H})$.[16] Every $W \in \mathfrak{C}_1$ is compact and $\mathfrak{C}_1$ is a Banach *-algebra when normed with the symmetric trace norm $|W|_1 = \mathrm{Tr}\, [(W^*W)^{\frac{1}{2}}]$, which majorizes the trace and the operator norm and satisfies $|aW|_1 \leq ||a||.\ |W|_1$ for $a \in B(\mathfrak{H})$, $W \in \mathfrak{C}_1$.

For any fixed $n \in I_+$, n-tuples of objects will be denoted as $[t_1 \cdots t_n] = [t]_n = [t] = t$ if there is no risk of confusion with the components. Given a $[t]_n$ the associated $(n-1)$-tuple $[t_2 \cdots t_n]$ will be $[t]_n'$ or t'. Let γ_n be the symmetric group of degree n. Then $\pi \in \gamma_n$ can be characterized by

$$\pi = \begin{bmatrix} 1 & \cdots & n \\ \alpha_1 & \cdots & \alpha_n \end{bmatrix} = \begin{bmatrix} \bar{\alpha}_1 & \cdots & \bar{\alpha}_n \\ 1 & \cdots & n \end{bmatrix},$$

$$\alpha_j \in I_n, \quad j \in I_n, \quad \bigcup_{j \in I_n} \alpha_j = I_n,$$

the action of π on the n-tuples is defined by

$$(\pi^{-1}[t])_i \equiv t_{\alpha_i}, \quad j \in I_n,$$

while for functions of n-tuples of arguments $f([t], [s])$ (for example)

$$(\pi f)([t], [s]) \equiv f(\pi^{-1}[t], \pi^{-1}[s]).$$

The n-cycles of γ_n can be generated as powers π_c^{k-1}, $k \in I_n$, where $\pi_c = \left(\begin{smallmatrix} 1 & 2 & \cdots & n-1 & n \\ 2 & 3 & \cdots & n & 1 \end{smallmatrix}\right)$ and every $\pi \in \gamma_n$ has a unique decomposition as $\pi = \pi' \cdot \pi_c^{k-1}$, $k \in I_n$ such that π' is of the form

$$\pi' = \begin{bmatrix} 1 & 2 & \cdots & n \\ 1 & \beta_2 & \cdots & \beta_n \end{bmatrix}.$$

The set of all such π' is obviously isomorphic to γ_{n-1} and will be denoted by γ_n'.

Products of n-tuples of (noncommuting) elements in $B(\mathfrak{H})$ are written as $\prod_{j \in I_n}^{\rightarrow} a_j$, $a_j \in B(\mathfrak{H})$, where the arrow indicates the ordering of the factors a_j as j ascends in I_n. Functions of several complex variables: In notation we follow the book by Vladimirov.[17] A set $\mathcal{C} \subset R^n$ is a *cone* (with vertex at 0) if $s \in \mathcal{C} \to \lambda s \in \mathcal{C}$ for $\lambda \in R_+$. The *projection* $pr(\mathcal{C})$ of $\mathcal{C}$ is the set $\{s \mid s \in \mathcal{C} \text{ and } |s| = 1\}$. $\mathcal{C}'$ is a *compact subcone* of $\mathcal{C}$ if $pr(\mathcal{C}') \subset pr(\mathcal{C})$. $\mathcal{C}^*$ is the *conjugate* of $\mathcal{C}$ defined as $\mathcal{C}^* = \{s \mid s \cdot s' \geq 0 \text{ for all } s' \in \mathcal{C}\}$. Any $\mathcal{C}^*$ is closed and convex. The function $\mu_{\mathcal{C}}(s) = \sup_{s' \in pr(\mathcal{C})}(-s \cdot s')$ is the *indicatrix* of $\mathcal{C}$. If $h(\mathcal{C})$ is the convex hull of $\mathcal{C}$, the convexity index $\rho_{\mathcal{C}}$ is the number

$$\sup_{s \in R^n - \mathcal{C}^*} [\mu_{h(\mathcal{C})}(s)/\mu_{\mathcal{C}}(s)],$$

which equals 1 for convex cones. A subset $\mathfrak{T}^{\mathcal{C}}$ of C^n is a tube with base $\mathcal{C}$, if $\mathcal{C}$ is a connected open cone in R^n and $\mathfrak{T}^{\mathcal{C}} = R^n + i\mathcal{C}$. Given a tube $\mathfrak{T}^{\mathcal{C}}$, $p \geq 1$, $a \in \bar{R}_+$ one defines a class $H_p(a; \mathcal{C})$ of functions $f(z)$ by

$$f(z) \in H_p(a; \mathcal{C})$$

if

(i) $f(z)$ holomorphic in $\mathfrak{T}^{\mathcal{C}}$,

(ii) for every compact subcone $\mathcal{C}'$ of $\mathcal{C}$ there exist a finite constant $M(\mathcal{C}')$, $(\alpha \in R_+, \beta \in R_+)$ independent of $\mathcal{C}'$, such that for $z \in \mathfrak{T}^{\mathcal{C}}$

$$|f(z)| \leq M(\mathcal{C}')(1 + |z|)^{\beta}(1 + |s|^{-\alpha})e^{a|y|^p}.$$

One also defines

$$H_p(a_+; \mathcal{C}) = \bigcap_{a' > a} H_p(a'; \mathcal{C}).$$

[16] N. Dunford and J. T. Schwartz, *Linear Operators* (Interscience Publishers, Inc., New York, 1963), Sec. XI.9; C. E. Rickart, *General Theory of Banach Algebras*, (D. van Nostrand Company, Inc., Princeton, New Jersey, 1960), Sec. A1.4.

[17] V. S. Vladimirov, *Metody teorii funktsiĭ mnogikh kompleksnykh peremennykh*, (Naŭka, Moscow 1964), Secs. 25, 26, in particular p. 275, 280. The Fourier–Laplace transforms of distributions were treated in L. Schwartz, Medd. Lunds Univ. Mat. Semin. (Suppl.), 196 (1952).

In particular the class $H_1(0; \mathcal{C})$ has the property that if $f \in H_1(0; \mathcal{C})$ then any derivative $D^a f$ as well as power af^p, $p \geq 1$, $a \in C$ of f is also in $H_1(0; \mathcal{C})$. A function $f(z)$ holomorphic in the tube $\mathfrak{T}^{\mathcal{C}}$ is called the Fourier–Laplace transform of the spectral distribution $g(\lambda) \in \mathcal{D}'_{(\lambda)}$ if

(i) $g(\lambda)e^{-\lambda s} \in \mathcal{S}'_{(\lambda)}$ for $s \in \mathcal{C}$, $z = t + is$

(ii) $f(z) = \mathfrak{F}[g(\lambda)e^{-\lambda s}; t] = \int d\lambda\, g(\lambda)e^{iz\lambda}$, $z \in \mathfrak{T}^{\mathcal{C}}$.

Vladimirov proves the

Theorem: Let $f(z) \in H_1(a_+; \mathcal{C})$, $a \geq 0$, $\mathcal{C}$ a connected cone, then its spectral distribution $g(\lambda)$ is in $\mathcal{S}'_\lambda$ and $g(\lambda) = 0$ if $\mu_{\mathcal{C}}(\lambda) > a$. Conversely if $g(\lambda) \in \mathcal{S}'$, $g(\lambda) = 0$ for $\mu_{\mathcal{C}}(\lambda) >$ a for some $a \geq 0$ and cone $\mathcal{C}$, then all $D^a f(z)$ of its Fourier–Laplace transform $f(z)$ belong to $H_1(\rho_C a; h(\mathcal{C}))$.

Furthermore, under the assumptions of the converse part of the theorem it follows that $f(z)$ has unique boundary values $f(t)$

$$f(t) = \lim_{s \to 0,\, s \in \mathcal{C}'} f(t + is)$$

in $\mathcal{S}'$, which are independent of the way in which s tends to zero in $\mathcal{C}'$ compact in $\mathcal{C}$.

Let $\sigma_i = \pm 1$, $i \in I'_n$ be a collection of signs and denote by σ the $(n-1)$-tuple $[\sigma_2 \cdots \sigma_n]$. If $\theta(t)$ is the Heaviside distribution, put

$$\eta_{\sigma_i}(z_i) = \sigma_i \theta(\sigma_i \operatorname{Re} z_i), \quad i \in I'_n$$

and

$$\eta_\sigma(z) \equiv 1 \bigotimes_{i \in I_n'} \eta_{\sigma_i}(z_i) \in \mathcal{S}'_{(t)}$$

as well as

$$\eta(t) \equiv 1 \bigotimes_{i \in I_n'} \theta(-t_i).$$

One easily evaluates that

$$\mathfrak{F}^{-1}[\eta_\sigma(z); \lambda] = \delta(\lambda_1) \bigotimes_{i \in I_n'} [(2\pi i)(\lambda_i - i\epsilon_i\sigma_i)]^{-1}|_{\epsilon_i \to 0+}.$$

Rigid Motions in Einstein Spaces*

HUGO D. WAHLQUIST AND FRANK B. ESTABROOK
Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California
(Received 26 July 1965)

The dyadic formulation of general relativity is used systematically to discuss rigid congruences in Einstein space–time. For space–time of uniform curvature, the quotient space metrics of rotating and accelerating rigid bodies are obtained. For Einstein space–time of nonuniform curvature, all irrotational, nonisometric, rigid motions are explicitly displayed. They have one degree of freedom, and occur only in degenerate static metrics of Class B. Rotating rigid congruences in Einstein space–time of nonuniform curvature are shown to have no degrees of freedom. Their evolution is in fact found to be governed by a complete set of 14 first-order total differential equations, linear in the time derivatives of the dyadic variables. Such rotating motions are shown further to be constrained by a set of algebraic conditions, and the implication of this for the validity of the Herglotz–Noether theorem in Einstein space–time is discussed.

I. INTRODUCTION

THE investigation of rigid motion in relativity was initiated by the early papers of M. Born,[1] G. Herglotz,[2] and F. Noether.[3] Little further work on the subject appeared until much later when its possible importance for the general theory of relativity emerged, as in, for example, the remarks of Synge.[4] The Born constraint of vanishing rate-of-strain may be carried over directly into general relativity, cf. Rosen[5] and Salzmann and Taub.[6] So defined, rigid motions in curved space–time are of considerable interest; mathematically, for their possible use in the invariant geometric analysis of the structure and global properties of Riemannian manifolds, analogous to that based on the more restricted isometries; and physically, for attaining a better understanding of the imprecise concept of rigidity so often implicitly involved in the interpretation of relativistic experiments, e.g., the Pound–Rebka red-shift experiment.

In 1959 C. B. Rayner[7] showed that the proper energy density and the *magnitude* of the angular velocity of a rigid body must be constant in time. Recently F. A. E. Pirani and G. Williams[8] have included this latter in a set of six relations between the time behavior of the angular velocity vector of the body and certain components of the space–time curvature, relations which follow from time constancy of the orthogonal metric of the rigid body. R. H. Boyer[9] has carefully discussed the entire problem and has extended this work to obtain sufficient, but not necessary, conditions for the validity of the Herglotz–Noether theorem in curved manifolds.

In the present paper we obtain some new results for test rigid motions in Einstein space, using the dyadic formalism for general relativity presented in what we will refer to as Paper I.[10] The basic dyadic equations for this problem are given in Sec. II.

In Sec. III, partly as an illustration of dyadic techniques, we prove the Herglotz–Noether theorem in space–time of uniform curvature (de Sitter space) —a proof already far from trivial in covariant language. We derive in this space the general quotient space (three-dimensional or inner) metrics of rotating and accelerating test rigid bodies.

In Sec. IV we show that nonrotating nonisometric rigid motions can only occur in a restricted class of Type D Einstein spaces, and that such motions are allowed but one degree of freedom. The metrics and the rigid frames are exhibited. These are the only rigid motions having any degree of freedom whatever in Einstein spaces which are neither flat nor uniformly curved.

In Sec. V we explicitly attack the Herglotz–Noether theorem in Einstein space. The six time-derivative equations of Pirani and Williams[8] are shown to be included in a full set of fourteen, which

* This work was sponsored by the National Aeronautics and Space Administration under Contract No. NAS7-100.

[1] M. Born, Ann. Physik **30**, 1 (1909).
[2] G. Herglotz, Ann. Physik **31**, 393 (1910).
[3] F. Noether, Ann. Physik. **31**, 919 (1910).
[4] J. L. Synge, *Relativity: The General Theory* (North-Holland Publishing Company, Amsterdam, 1960).
[5] N. Rosen, Phys. Rev. **71**, 54 (1947).
[6] G. Salzman and A. H. Taub, Phys. Rev. **95**, 1659 (1954).
[7] C. B. Rayner, Compt. Rend. **248**, 929 (1959a).
[8] F. A. E. Pirani and G. Williams, Séminaire JANET, 5ième année, 1961/2, No. 8–9.
[9] R. H. Boyer, Proc. Roy. Soc. (London) **A283**, 343 (1965).
[10] F. Estabrook and H. Wahlquist, J. Math. Phys. **5**, 1629 (1964). Referred to as Paper I. For expositions of several special techniques of calculation in the dyadic notation, as well as an extended discussion of the rigidity problem, see H. D. Wahlquist and F. B. Estabrook, Report No. 32-868, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California (1966).

may be in principle solved for the local time-rate-of-change of the fourteen dyadic components involved, subject to a set of subsidiary algebraic conditions. A rigid frame can rotate only in the unlikely event that the subsidiary conditions are all fulfilled throughout the motion; the equations then determine the complete evolution of the motion and the metric from any initial 3-space. It remains to show that the subsidiary conditions can be satisfied only by isometries, if the Herglotz–Noether theorem is to be stated for general Einstein spaces in the usual simple fashion.

II. THE DYADIC EQUATIONS FOR RIGID REFERENCE FRAMES IN EINSTEIN SPACE

The dyadic formalism results from the introduction of an orthonormal tetrad of basis vectors, one timelike and three spacelike, at each event in space–time. The timelike tetrad vector $\mathbf{u}^0$ is everywhere aligned with a given timelike congruence and so coincides with the unit tangent to the congruence. If the world lines of the congruence represent the history of a material medium, $\mathbf{u}^0$ will be the 4-velocity field of the matter. The spacelike triad $\mathbf{u}^\alpha$ [or $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$] then provides a locally co-moving frame of reference and spans the local 3-space orthogonal to the given congruence at each event. A formal partitioning of the space–time manifold is thus achieved, and, by projecting all tensor fields of interest into these orthogonal 3-spaces, we arrive at the 3-dimensional vector and dyadic formalism. A complete discussion of this approach and the dyadic notation is contained in Paper I.

It has the advantage that the field quantities are always explicitly displayed as arrays of physically interpretable components; further, as will appear in the present paper, it allows the choice of intrinsic reference vectors [as well as the use of intrinsic (holonomic) coordinates] in analyzing the algebraic aspects of relativistic problems before integration of their partial differential equations.

A general timelike congruence may be described at each point by its absolute acceleration vector $\mathbf{a}$, angular velocity vector $\boldsymbol{\Omega}$, and symmetric rate-of-strain dyadic S. Following Born, the rigid congruence is defined by $\mathsf{S} = 0$, so that both the shear and expansion vanish. In space–time this is equivalent to constraining the world lines of the timelike congruence to have constant orthogonal separations. As a geometrical consequence of this constraint, the orthogonal 3-space metric on a rigid body in co-moving coordinates is constant throughout time, $(h_{\alpha\beta})^{\cdot} = 0$. By adopting orthonormal basis triads which co-rotate with the rigid body [$\boldsymbol{\omega} = \boldsymbol{\Omega}$, cf. Eqs. (C42), (D25), and (D26) of Paper I] and are thus fixed in it, we ensure that the anholonomic affinity of the orthogonal 3-space is time-independent: $\dot{\mathsf{N}} = 0$. [In Paper I the dyadic affinity was denoted by $\mathsf{N}^\star$, the bare symbol N representing only its symmetric part (up to a trace). This notation has proved slightly clumsy, and in the present paper N replaces the $\mathsf{N}^\star$ of Paper I.] In this situation it is permissible to replace the many local orthogonal 3-spaces with the concept of a single metric "quotient space"—the 3-dimensional manifold of co-moving coordinates, χ^α. Geometrically, we may picture the quotient space as a reduction of space–time obtained when all events lying on each world line of the timelike congruence are identified. The orthogonal metric $h_{\alpha\beta}$ and the affinity N are then applied to this quotient space.

The curvature equation for this 3-dimensional metric space becomes

$$\nabla \times \mathsf{N} = -\tfrac{1}{2}\mathsf{N}^T \overset{\times}{\times} \mathsf{N} + \mathsf{E}. \tag{1}$$

The symmetric dyadic E is the conservative dyadic of the quotient space, satisfying the Bianchi identity

$$\nabla \cdot \mathsf{E} = 0. \tag{2}$$

Its diagonal elements are the Riemannian curvatures of the quotient space, based on the orthonormal triad vectors, while all six components are sufficient to express the complete 3-space curvature tensor. Since $\dot{\mathsf{N}} = 0$, we can apply the commutation relation, Eq. (4), to Eq. (1) to show that

$$\dot{\mathsf{E}} = 0, \tag{3}$$

independently of the coordinate system employed. In co-moving coordinates this is, of course, seen as a trivial consequence of $(h_{\alpha\beta})^{\cdot} = 0$.

The commutation rules for covariant and time differentiation in the quotient space are somewhat unfamiliar in general as a result of the use of anholonomic reference systems. And although the geometry of the quotient space is constant in time, we may have time-dependent objects in the space and for these the order of space and time differentiation is of consequence. Letting ψ, $\mathbf{V}$, and M symbolize, respectively, an arbitrary scalar, vector, and dyadic, we have

$$\begin{aligned} (\nabla\psi)^{\cdot} - \nabla(\dot{\psi}) &= \mathbf{a}\dot{\psi}, \\ (\nabla\mathbf{V})^{\cdot} - \nabla(\dot{\mathbf{V}}) &= \mathbf{a}\dot{\mathbf{V}}, \\ (\nabla \times \mathsf{M})^{\cdot} - \nabla \times (\dot{\mathsf{M}}) &= \mathbf{a} \times \dot{\mathsf{M}}. \end{aligned} \tag{4}$$

Note that for any time-independent quantity these operations do commute; this is a special property of rigid congruences.

Commutation of the 3-dimensional covariant differentiation operator ∇ with itself is also somewhat unusual, except when the angular velocity $\boldsymbol{\Omega}$ of the rigid body vanishes. (In this latter case the 4-velocity of the body is derivable from a scalar potential and its world lines are the orthogonal trajectories of a family of hypersurfaces in space–time. Consequently, the quotient space becomes isometric to all the members of a family of immersed Riemannian 3-spaces, which could be parametrized by t.) Samples of the space–space commutation relations are

$$\begin{aligned} \nabla \times \nabla \psi &= 2\psi\boldsymbol{\Omega}, \\ \nabla \cdot (\nabla \times \mathbf{V}) &= 2\boldsymbol{\Omega} \cdot \dot{\mathbf{V}}, \\ \nabla \times (\nabla \mathbf{V}) &= -\mathsf{E} \times \mathbf{V} + 2\boldsymbol{\Omega}\dot{\mathbf{V}}. \end{aligned} \tag{5}$$

We impose the condition for Einstein space–time in the form

$$G_{\mu\nu} = R_{\mu\nu} - \tfrac{1}{2}Rg_{\mu\nu} = \Lambda g_{\mu\nu}, \tag{6}$$

where $\Lambda = -\frac{1}{4}R$ is the cosmological constant. In orthonormal components this equation leads to the following expressions for the contracted curvature quantities of Eq. (D11) of Paper I:

$$\mathsf{T} = \tfrac{1}{2}\Lambda \mathsf{I}, \tag{7}$$

$$\mathsf{t} = 0, \tag{8}$$

$$\rho = \tfrac{1}{2}\Lambda. \tag{9}$$

Ten of the twenty curvature components are specified by these equations. The remaining ten, comprised in the Weyl tensor, are represented in two symmetric, traceless, dyadics A and B which remain arbitrary. The curvature of the quotient space is related to the space–time curvature by a generalized Gauss equation,

$$\mathsf{E} = -\mathsf{A} + \tfrac{1}{3}\Lambda \mathsf{I} - 3\boldsymbol{\Omega}\boldsymbol{\Omega}. \tag{10}$$

This equation, together with Eq. (3), has important consequences for allowed rigid motions, which will be discussed fully in later sections.

As a vector field in the quotient space, the acceleration of the rigid body satisfies the dyadic equation

$$\nabla \mathbf{a} = -\mathbf{a}\mathbf{a} + \boldsymbol{\Omega}\boldsymbol{\Omega} - \dot{\boldsymbol{\Omega}} \times \mathsf{I} + \mathsf{A} - (\Omega^2 + \tfrac{1}{3}\Lambda)\mathsf{I}. \tag{11}$$

The trace of this equation reads

$$\nabla \cdot \mathbf{a} = -a^2 - 2\Omega^2 - \Lambda, \tag{12}$$

and its antisymmetric part is

$$\nabla \times \mathbf{a} = 2\dot{\boldsymbol{\Omega}}. \tag{13}$$

This last equation is one integrability condition for introduction of a scalar potential ϕ and a vector potential $\mathbf{A}$:

$$\phi \mathbf{a} = -\nabla\phi + \dot{\mathbf{A}}. \tag{14}$$

The corresponding set of equations for the angular velocity or vorticity field becomes

$$\nabla\boldsymbol{\Omega} = -2\mathbf{a}\boldsymbol{\Omega} + (\mathbf{a} \cdot \boldsymbol{\Omega})\mathsf{I} + \mathsf{B}, \tag{15}$$

$$\nabla \cdot \boldsymbol{\Omega} = \mathbf{a} \cdot \boldsymbol{\Omega}, \tag{16}$$

$$\nabla \times \boldsymbol{\Omega} = 2\boldsymbol{\Omega} \times \mathbf{a}, \tag{17}$$

with the vector potential equation

$$2\phi\boldsymbol{\Omega} = \nabla \times \mathbf{A}. \tag{18}$$

The scalar and 3-vector potentials, ϕ and $\mathbf{A}$, introduced in Eqs. (14) and (18) can be related to the expression of the space–time interval in a co-moving system of coordinates (x^α, t),

$$ds^2 = -(1/\phi^2)\, dt^2 + (2/\phi^2) A_\alpha\, dx^\alpha\, dt + [h_{\alpha\beta} - (1/\phi^2) A_\alpha A_\beta]\, dx^\alpha\, dx^\beta,$$

where $A_\alpha = \mathbf{A} \cdot \mathbf{e}_\alpha$ and $h_{\alpha\beta}$ is the metric of the 3-dimensional quotient space defined in terms of natural basis vectors $\mathbf{e}_\alpha$ by $h_{\alpha\beta} = \mathbf{e}_\alpha \cdot \mathbf{e}_\beta$ with $\mathbf{e}_\alpha \cdot \mathbf{e}^\beta = \delta_\alpha^\beta$ and $\mathbf{e}^\alpha = \nabla x^\alpha$. The existence of sets of these potentials for the inertial fields $\mathbf{a}$ and $\boldsymbol{\Omega}$ is in fact guaranteed by the general availability of co-moving coordinates; cf., e.g., Lichnerowicz.[11] For any such system a set results, defined and interpreted as absolute derivatives of the t coordinate, respectively, along and orthogonal to the t congruences:

$$\phi \equiv \frac{dt}{d\tau}, \qquad \mathbf{A} \equiv \nabla t,$$

where τ is proper time along a line of the t congruence.

The gravitational field equations, or Bianchi identities, have the Maxwellian-like dyadic forms

$$\nabla \times \mathsf{A} - \dot{\mathsf{B}} = -\mathbf{a} \times \mathsf{A} + \mathsf{A} \times \mathbf{a} - \boldsymbol{\Omega} \times \mathsf{B} - 2\mathsf{B} \times \boldsymbol{\Omega}, \tag{19}$$

$$\nabla \times \mathsf{B} + \dot{\mathsf{A}} = -\mathbf{a} \times \mathsf{B} + \mathsf{B} \times \mathbf{a} + \boldsymbol{\Omega} \times \mathsf{A} + 2\mathsf{A} \times \boldsymbol{\Omega}. \tag{20}$$

The antisymmetric parts of these two equations are the vector equations

$$\nabla \cdot \mathsf{A} = -3\boldsymbol{\Omega} \cdot \mathsf{B}, \tag{21}$$

$$\nabla \cdot \mathsf{B} = 3\boldsymbol{\Omega} \cdot \mathsf{A}. \tag{22}$$

Finally, since we have specialized to Einstein space,

[11] A. Lichnerowicz, *Théories Relativistes de la Gravitation et de l'Électroagnétmisme* (Masson et Cie., Paris, 1955).

the contracted Bianchi identities are trivial. They merely insist on the constancy of the cosmological constant;

$$\dot{\Lambda} = \nabla\Lambda = 0. \tag{23}$$

III. RIGID MOTION IN SPACE-TIMES OF CONSTANT CURVATURE

We first quickly prove the Herglotz–Noether theorem proper, which states that rigid motions, when there is rotation, are necessarily isometries of the manifold. The only original contribution here is a derivation of the quotient space metrics intrinsic to a rotating and accelerating rigid body in the space–time of constant curvature.

Setting the conformal curvature dyadics A and B to zero, we obtain the equations appropriate to a space–time of constant curvature, $K = \frac{1}{3}\Lambda$. Combining Eqs. (3) and (10) we find that

$$\dot{\boldsymbol{\Omega}} = 0. \tag{24}$$

Since our equations are written in body-fixed axes, this result shows immediately that the angular velocity is a fixed vector in the body; however, since $\dot{\boldsymbol{\Omega}} + \boldsymbol{\omega}\times\boldsymbol{\Omega}$ also vanishes ($\boldsymbol{\omega} = \boldsymbol{\Omega}$), the angular velocity vector is fixed in Fermi propagated axes as well. We note also that Eq. (13) now reads simply, $\nabla\times\mathbf{a} = 0$.

Using the fact that space and time derivatives of $\boldsymbol{\Omega}$ will commute, it follows from time differentiating Eq. (16) that $(\mathbf{a}\cdot\boldsymbol{\Omega})^{\cdot} = 0$. Then treating Eq. (15) similarly one gets the dyadic equation

$$\dot{\mathbf{a}}\boldsymbol{\Omega} = 0. \tag{25}$$

Thus, if the angular velocity does not vanish, we must have $\dot{\mathbf{a}} = 0$ so that the acceleration vector is body-fixed also. The two equations now satisfied by $\mathbf{a}$ when $\boldsymbol{\Omega} \neq 0$, viz., $\nabla\times\mathbf{a} = \dot{\mathbf{a}} = 0$, together with the constraint of rigidity $\mathsf{S} = 0$, are the necessary and sufficient conditions for an isometric or Killing vector congruence. Thus the Herglotz–Noether theorem applies to all space–times of constant curvature. The full ten parameter group of isometries exists for these spaces and is well known; Herglotz[2] in particular has given a detailed treatment of the timelike isometries in flat space–time.

The curvature of the spatial geometry in a rotating rigid body depends on the angular velocity, as shown by the Gauss equation (10) for the 3-space curvature dyadic E,

$$\mathsf{E} = -3\boldsymbol{\Omega}\boldsymbol{\Omega} + \tfrac{1}{3}\Lambda\mathsf{I}. \tag{26}$$

The scalar curvature of the space is given by $-2\,\mathrm{Tr}\,\mathsf{E}$ and so has the value $-2(\Lambda - 3\Omega^2)$. The metric of the quotient 3-space on a rotating rigid disk has been a subject of occasional interest in the literature of relativity; several references to early discussions are given by Berenda.[12] In constant curvature space–time, the most general rotating and accelerating rigid body is a typical example of a situation providing its own intrinsic comoving coordinates, and the quotient metrics are easily obtained with this approach.

For the rotating body, three intrinsic scalars may exist; a^2, Ω^2, and $\mathbf{a}\cdot\boldsymbol{\Omega}$, all of which must be time independent as we have shown, so that co-moving coordinates may be constructed from them. The gradients of these scalars are obtained directly from Eqs. (11) and (15) and can be combined to express the following vectors:

$$\mathbf{a} = -\tfrac{1}{4}\nabla \ln [a^2 - \Omega^2 + \tfrac{1}{3}\Lambda]^2$$
$$= -\tfrac{1}{4}\nabla \ln [(\mathbf{a}\cdot\boldsymbol{\Omega})^2 + \tfrac{1}{3}\Lambda\Omega^2], \tag{27}$$

$$\boldsymbol{\Omega} = [\Omega^2/2(\mathbf{a}\cdot\boldsymbol{\Omega})]\nabla \ln [\Omega^2/(a^2 - \Omega^2 + \tfrac{1}{3}\Lambda)^2], \tag{28}$$

$$\mathbf{a}_\perp \equiv \mathbf{a} - [(\mathbf{a}\cdot\boldsymbol{\Omega})/\Omega^2]\boldsymbol{\Omega}$$
$$= \tfrac{1}{4}\nabla \ln [(a^2 - \Omega^2 + \tfrac{1}{3}\Lambda)^2/(\Omega^2)^2]. \tag{29}$$

Comparison of the gradients further results in the integral

$$(\mathbf{a}\cdot\boldsymbol{\Omega})^2 + \tfrac{1}{3}\Lambda\Omega^2 = k^2(a^2 - \Omega^2 + \tfrac{1}{3}\Lambda)^2, \tag{30}$$

revealing that the three scalars are not independent so that only two intrinsic coordinates can actually be constructed. Since space–times of negative constant curvature have very peculiar physical properties, we shall here consider $\Lambda \geq 0$ only. By the form of (30) the integration constant is then required to be positive, so is written k^2, and we take $k \geq 0$.

It is clear that two especially convenient intrinsic coordinates are obtained by taking one of them to be some function of the argument of the logarithm in Eq. (28) and the other a function of the argument in Eq. (29). For then the gradient vectors of such coordinates will be orthogonal, leading to a diagonal metric. We select the functions by keeping in mind the flat space–time example of the rigid disk rotating about a fixed axis, and requiring the general metric to reduce in that limit to a familiar form in cylindrical co-moving coordinates.

This is accomplished for the argument in Eq. (28) by setting

$$\Omega^2/(a^2 - \Omega^2 + \tfrac{1}{3}\Lambda)^2 = (k^2/\alpha^2)\sin^2 [\alpha(1/\omega_0 k - z)], \tag{31}$$

where $\alpha \equiv (\frac{1}{3}\Lambda)^{\frac{1}{2}}$, ω_0 is a new constant, and z is

[12] C. W. Berenda, Phys. Rev. 62, 280 (1942).

introduced as the coordinate whose surfaces have $\boldsymbol{\Omega}$ as normal vector. The other coordinate r is introduced by putting

$$(a^2 - \Omega^2 + \tfrac{1}{3}\Lambda)/\Omega^2 = -(1/2k^2)[-1 + \beta\cos(2k\omega_0 r)], \tag{32}$$

where $\beta \equiv (1 + 4k^2)^{\frac{1}{2}}$, and $\mathbf{a}_\perp$ is thus orthogonal to the coordinate surfaces r = constant. Now solving Eqs. (30), (31), and (32) simultaneously we express the intrinsic scalars for the general case in terms of these intrinsic coordinates;

$$\begin{aligned} a^2 &= a_\perp^2 + (\mathbf{a}\cdot\boldsymbol{\Omega})^2/\Omega^2, \\ a_\perp^2 &= \alpha^2\beta^2 R^{-2}\sin^2(2k\omega_0 r)\csc^2 Z, \\ \Omega^2 &= 4\alpha^2 k^2 R^{-2}\csc^2 Z, \\ (\mathbf{a}\cdot\Omega)^2 &= 4\alpha^4 k^2 R^{-2}\csc^2 Z\cot^2 Z, \end{aligned} \tag{33}$$

where

$$\begin{aligned} R &\equiv -1 + \beta\cos(2k\omega_0 r), \\ Z &\equiv \alpha(1/\omega_0 k - z). \end{aligned}$$

Differentiating (31) and (32), substituting in Eqs. (28) and (29), and using the expressions in Eq. (33), we find

$$\boldsymbol{\Omega} = |\boldsymbol{\Omega}|\,\nabla z, \tag{34}$$

showing ∇z as a unit vector, and

$$\mathbf{a}_\perp = -|\mathbf{a}_\perp|\,(k\omega_0/\alpha)\sin Z\,\nabla r. \tag{35}$$

A third intrinsic coordinate does not exist; however, applying the dyadic Eqs. (11) and (15) for $\nabla\mathbf{a}$ and $\nabla\boldsymbol{\Omega}$, we easily prove that

$$(\mathbf{a}\times\boldsymbol{\Omega})\cdot[\nabla\times(\mathbf{a}\times\boldsymbol{\Omega})] = 0,$$

so that a scalar potential also exists for this vector which is orthogonal to both $\boldsymbol{\Omega}$ and $\mathbf{a}_\perp$. Letting the potential be denoted by θ, we can show that a consistent choice is

$$\mathbf{a}\times\boldsymbol{\Omega} = |\mathbf{a}_\perp|\,|\boldsymbol{\Omega}|\,(\beta/\omega_0)[\sin(2k\omega_0 r)/(2\,|R|)^{\frac{1}{2}}] \times (k\omega_0/\alpha)\sin Z\,\nabla\theta. \tag{36}$$

The gradients of the coordinates constitute the natural basis vector triad $\mathbf{e}^\alpha$, so putting

$$\mathbf{e}^1 = \nabla r, \qquad \mathbf{e}^2 = \nabla\theta, \qquad \mathbf{e}^3 = \nabla z,$$

we use Eqs. (34), (35), and (36) to determine the dual set as

$$\begin{aligned} \mathbf{e}_1 &= -(k\omega_0/\alpha)\sin Z(\mathbf{a}_\perp/|\mathbf{a}_\perp|), \\ \mathbf{e}_2 &= (k\omega_0/\alpha)\sin Z(\beta/\omega_0)[\sin(2k\omega_0 r)/(2\,|R|)^{\frac{1}{2}}] \times (\mathbf{a}\times\boldsymbol{\Omega}/|\mathbf{a}\times\boldsymbol{\Omega}|), \\ \mathbf{e}_3 &= \boldsymbol{\Omega}/|\boldsymbol{\Omega}|, \end{aligned}$$

satisfying the inversion $\mathbf{e}^\alpha\cdot\mathbf{e}_\beta = \delta^\alpha_\beta$. Using $h_{\alpha\beta} = \mathbf{e}_\alpha\cdot\mathbf{e}_\beta$ the general quotient space metric for a rotating, accelerating, rigid body in constant curvature space–time becomes

$$dl^2 = (k\omega_0/\alpha)^2\sin^2 Z \times \{dr^2 + (\beta/\omega_0)^2[\sin^2(2k\omega_0 r)/2\,|R|]\,d\theta^2\} + dz^2. \tag{37}$$

Limits are required on the ranges of both z and r coordinates; for z

$$0 < Z = \alpha(1/\omega_0 k - z) < \pi, \tag{38}$$

and for r

$$0 \le 2k\omega_0 r \le \pi, \tag{39}$$

but this latter range is subdivided by the vanishing of the quantity R in the denominator at

$$\cos(2k\omega_0 r) = \beta^{-1}.$$

Since the intrinsic scalars become infinite at this point, it is best to give two different metrics for each range.

The first becomes

$$\text{I.} \qquad 0 \le r \le (1/2k\omega_0)\cos^{-1}(1/\beta),$$

$$dl^2 = \left(\frac{k\omega_0}{\alpha}\right)^2\sin^2 Z \times\left\{dr^2 + \left(\frac{\beta}{\omega_0}\right)^2\frac{\sin^2(2k\omega_0 r)}{2[\beta\cos(2k\omega_0 r) - 1]}\,d\theta^2\right\} + dz^2. \tag{40}$$

Letting $\Lambda \to 0$ we find for flat space–time

$$dl^2 = (1 - \omega_0 kz)^2 \times\left\{dr^2 + \left(\frac{\beta}{\omega_0}\right)^2\frac{\sin^2(2k\omega_0 r)}{2[\beta\cos(2k\omega_0 r) - 1]}\,d\theta^2\right\} + dz^2 \qquad (\Lambda = 0). \tag{41}$$

And if we further let k vanish so that $\mathbf{a}\cdot\boldsymbol{\Omega} = 0$, we get

$$dl^2 = dr^2 + [r^2/(1 - \omega_0^2 r^2)]\,d\theta^2 + dz^2 \qquad (\Lambda = k = 0), \tag{42}$$

which is the quotient metric on a fixed axis rotating rigid disk in the form obtained by Berenda.[12] Compare also the discussion by Møller.[13]

To write the metric for the second allowed range, it is desirable to introduce a new coordinate $\bar{r}$ according to

$$2k\omega_0\bar{r} = \pi - 2k\omega_0 r,$$

[13] C. Møller, *The Theory of Relativity* (Oxford University Press, London, 1952), Chap. VIII.

so that

II. $$0 \leq \bar{r} \leq (1/2k\omega_0)\cos^{-1}(-1/\beta),$$

$$dl^2 = \left(\frac{k\omega_0}{\alpha}\right)^2 \sin^2 Z \times \left\{d\bar{r}^2 + \left(\frac{\beta}{\omega_0}\right)^2 \frac{\sin^2(2k\omega_0\bar{r})}{2[\beta\cos(2k\omega_0\bar{r})+1]}\, d\theta^2\right\} + dz^2. \quad (43)$$

The limiting case $\Lambda = k = 0$ is of little interest here, since it can be shown from Eq. (33) that it corresponds to vanishing angular velocity in flat space–time, and consequently a flat quotient space results.

IV. IRROTATIONAL RIGID FRAMES

The class of irrotational rigid motions proves to be the most interesting for arbitrary exterior gravitational fields. Although we have not quite been able to carry through a complete proof of the Herglotz–Noether theorem, it will be shown in Sec. V that the motion of the *rotating* rigid frame in any exterior field is totally determined by initial conditions on a spacelike hypersurface. Only among the irrotational frames can we find any motions allowing arbitrary functions of time, or in other words, degrees of freedom. We shall concentrate here on finding all such nonisometric normal rigid motions which may exist in an arbitrary Einstein space. The variety of exterior metrics admitting such motions will also be obtained; they turn out to be quite special, falling into a well-known class of Einstein spaces.

The equations for these motions are, of course, obtained by putting $\boldsymbol{\Omega} = 0$ throughout Sec. II. We note first of all that the potential equation (18) for $\boldsymbol{\Omega}$ becomes $\nabla \times \mathsf{A} = 0$, and accordingly we may select a gauge such that $\mathbf{A} = 0$. This simply corresponds to the possibility of using a time coordinate whose surfaces coincide with the normal hypersurfaces of the congruence. Next we find that Eq. (15) reduces to $\mathsf{B} = 0$, and the gravitational field is thus described by A alone. Here we insist that $\mathsf{A} \neq 0$, since the manifold would otherwise reduce to a space–time of constant curvature. This structure of the field dyadics can only be obtained in Type I and D Einstein spaces in the Petrov–Pirani–Sachs classification scheme, as pointed out previously by Pirani and Williams.[8]

Since we are not presently interested in isometric motions, we assume henceforth that $\dot{\mathbf{a}} \neq 0$, and attempt to construct solutions obeying this constraint. Now the algebraic properties of the dyadic A are examined. Using the fact that $\dot{\mathsf{A}} = 0$ [Eq. (20)], we time-differentiate the Bianchi identity Eq. (19) and find

$$(\nabla \times \mathsf{A})^{\cdot} = -\dot{\mathbf{a}} \times \mathsf{A} + \mathsf{A} \times \dot{\mathbf{a}} = 0, \quad (44)$$

so that the symmetric part of the dyadic $\dot{\mathbf{a}} \times \mathsf{A}$ vanishes. From this equation and $\mathrm{Tr}\ \mathsf{A} = 0$, it can be shown that $\dot{\mathbf{a}}$ is an eigenvector of A and that we may solve for A as

$$\mathsf{A} = \alpha^3[3(\dot{\mathbf{a}}\dot{\mathbf{a}}/\dot{a}^2) - \mathsf{I}], \quad (45)$$

where $2\alpha^3$ is the eigenvalue. It will be convenient to define a unit vector, $\mathbf{u} = \dot{\mathbf{a}}/|\dot{\mathbf{a}}|$, and so to write Eq. (45) as

$$\mathsf{A} = \alpha^3[3\mathbf{u}\mathbf{u} - \mathsf{I}]. \quad (46)$$

It is easily verified that $\mathsf{A} : \mathsf{A} = 6\alpha^6$, and the time derivative of this shows $\dot{\alpha} = 0$. Now time differentiating Eq. (46) and using the fact that $\mathbf{u}$ is a unit vector, we find $\dot{\mathbf{u}} = 0$ so that $\mathbf{u}$ is a body-fixed unit vector.

The vector $\mathbf{u}$ will be adopted as one of the intrinsic triad basis vectors; we now find two more to complete the triad. Let $\mathbf{a}_\perp$ be the component of the acceleration perpendicular to $\dot{\mathbf{a}}$, so that

$$\mathbf{a}_\perp \equiv \mathbf{a} - (\mathbf{a}\cdot\mathbf{u})\mathbf{u}.$$

We may verify that this is a time-independent, body-fixed vector and so can use it to define a second intrinsic triad vector $\mathbf{v}$ by writing

$$\mathbf{a}_\perp = |\mathbf{a}_\perp|\,\mathbf{v} = (\mathbf{a}\cdot\mathbf{v})\mathbf{v}, \quad (47)$$

where

$$v^2 = 1, \quad \mathbf{u}\cdot\mathbf{v} = 0, \quad \dot{\mathbf{v}} = 0, \quad (\mathbf{a}\cdot\mathbf{v})^{\cdot} = 0. \quad (48)$$

The third triad vector $\mathbf{w}$ is defined simply by $\mathbf{w} = \mathbf{u} \times \mathbf{v}$, and, of course, $\mathbf{a}\cdot\mathbf{w} = \dot{\mathbf{a}}\cdot\mathbf{w} = 0$. This construction will fail if $\mathbf{a}$ and $\dot{\mathbf{a}}$ are collinear, but for now we assume this is not the case. We shall return to consider the collinear case later; it is fairly easily obtained as a limit of the more general solution.

Having adopted a basis triad, we proceed to investigate its differential properties. From its definition we may write

$$\nabla\mathbf{u} = \nabla(\dot{\mathbf{a}}/|\dot{\mathbf{a}}|) = (1/|\dot{\mathbf{a}}|)[\nabla(\dot{\mathbf{a}}) - (\nabla\,|\dot{\mathbf{a}}|)\mathbf{u}]. \quad (49)$$

The right-hand side of this equation is evaluated by combining the commutation relation Eq. (4) and the time derivative of Eq. (11). We have

$$\nabla\mathbf{u} = -\mathbf{u}[\mathbf{a} - (\mathbf{a}\cdot\mathbf{u})\mathbf{u}] = -\mathbf{u}\mathbf{a}_\perp = -(\mathbf{a}\cdot\mathbf{v})\mathbf{u}\mathbf{v}, \quad (50)$$

and the corollaries

$$\nabla \times \mathbf{u} = (\nabla \mathbf{u}) \dot{\times} \mathsf{I} = -(\mathbf{a}\cdot\mathbf{v})\mathbf{w}, \qquad \nabla \cdot \mathbf{u} = (\nabla \mathbf{u}) : \mathsf{I} = 0. \tag{51}$$

In like manner we evaluate $\nabla \mathbf{v} = \nabla(\mathbf{a}_\perp/|\mathbf{a}_\perp|)$. The result is

$$\nabla \mathbf{v} = (\mathbf{a}\cdot\mathbf{v})\mathbf{u}\mathbf{u} - [(\alpha^3 + \tfrac{1}{3}\Lambda)/(\mathbf{a}\cdot\mathbf{v})]\mathbf{w}\mathbf{w}, \tag{52}$$

and

$$\nabla \times \mathbf{v} = 0, \qquad \nabla \cdot \mathbf{v} = [1/(\mathbf{a}\cdot\mathbf{v})][(\mathbf{a}\cdot\mathbf{v})^2 - \alpha^3 - \tfrac{1}{3}\Lambda]. \tag{53}$$

Finally, for $\nabla \mathbf{w}$, we have

$$\nabla \mathbf{w} = [(\alpha^3 + \tfrac{1}{3}\Lambda)/(\mathbf{a}\cdot\mathbf{v})]\mathbf{w}\mathbf{v}, \tag{54}$$

with

$$\nabla \times \mathbf{w} = -[(\alpha^3 + \tfrac{1}{3}\Lambda)/(\mathbf{a}\cdot\mathbf{v})]\mathbf{u}, \qquad \nabla \cdot \mathbf{w} = 0. \tag{55}$$

The geometrical content of these equations may be summarized by the statements that $\mathbf{v}$ is a 2-space normal, geodesic vector of the quotient space, while $\mathbf{u}$ and $\mathbf{w}$ are 2-space normal, Killing vectors of the quotient space.

We are now prepared to calculate the gradients of all the independent intrinsic scalars which can be formed; viz., ϕ, $\mathbf{a}\cdot\mathbf{u}$, $\mathbf{a}\cdot\mathbf{v}$, and α. All other scalars are algebraic functions or derivatives of these. The potential equation (14) gives $\nabla\phi$ directly,

$$\nabla \phi = -\phi[(\mathbf{a}\cdot\mathbf{u})\mathbf{u} + (\mathbf{a}\cdot\mathbf{v})\mathbf{v}]. \tag{56}$$

The next two are evaluated as

$$\nabla(\mathbf{a}\cdot\mathbf{u}) = \nabla\mathbf{a}\cdot\mathbf{u} + \nabla\mathbf{u}\cdot\mathbf{a} = -(a^2 - 2\alpha^3 + \tfrac{1}{3}\Lambda)\mathbf{u} - (\mathbf{a}\cdot\mathbf{u})(\mathbf{a}\cdot\mathbf{v})\mathbf{v}, \tag{57}$$

and similarly

$$\nabla(\mathbf{a}\cdot\mathbf{v}) = -[(\mathbf{a}\cdot\mathbf{v})^2 + \alpha^3 + \tfrac{1}{3}\Lambda]\mathbf{v}. \tag{58}$$

The gradient of α is most easily derived by substituting the explicit form for A given in Eq. (46) into the field equation (21) to obtain

$$\nabla \alpha = -\alpha(a\cdot v)\mathbf{v}. \tag{59}$$

Comparison of Eqs. (58) and (59) now reveals that an integral exists, and using these equations it is found that

$$\nabla\{[(\mathbf{a}\cdot\mathbf{v})^2 - 2\alpha^3 + \tfrac{1}{3}\Lambda]/\alpha^2\} = 0. \tag{60}$$

Integrating and solving for $(\mathbf{a}\cdot\mathbf{v})^2$, this may be written as

$$(\mathbf{a}\cdot\mathbf{v})^2 = \epsilon k^2\alpha^2 + 2\alpha^3 - \tfrac{1}{3}\Lambda, \tag{61}$$

where we use an indicator $\epsilon = \pm 1$ to account for sign and may restrict the integration constant k so that $k \geq 0$. We have, of course, $\nabla k = 0$, but since all other quantities in Eq. (61) have already been shown to be time-independent, it follows that $\dot{k} = 0$ as well.

It is not quite so obvious that the set of Eqs. (56)–(59) leads also to another integral. Using them together with the integral (61) we may verify, however, that

$$\nabla\{[(\mathbf{a}\cdot\mathbf{u})^2 + \epsilon k^2\alpha^2]/\phi^2\} = 0,$$

and so obtain

$$(\mathbf{a}\cdot\mathbf{u})^2 + \epsilon k^2\alpha^2 = \epsilon'(k')^2\phi^2,$$

where ϵ' is another sign indicator, $\epsilon' = \pm 1$, and $k' \geq 0$ with $\nabla k' = 0$. Here k' might be a function of time. Since it enters only as a factor of the potential ϕ, however, it may be absorbed in the potential itself. The integral now becomes simply

$$(\mathbf{a}\cdot\mathbf{u})^2 = \epsilon'\phi^2 - \epsilon k^2\alpha^2. \tag{62}$$

We note that this equation forbids the combination: $\epsilon' = -1$, $\epsilon = 1$; all other combinations are in general allowed. Adding Eqs. (61) and (62) we obtain the square of the acceleration vector

$$a^2 = \epsilon'\phi^2 + 2\alpha^3 - \tfrac{1}{3}\Lambda. \tag{63}$$

Physically reasonable space–times are usually associated with $\Lambda \geq 0$, and, for these cases, we see from Eqs. (61) and (63) that if α is negative, the only allowed combination of signs is $\epsilon' = \epsilon = 1$.

At this point all the essential intrinsic relations have been derived; it remains only to adopt a system of co-moving coordinates, and obtain the corresponding metrics of the quotient space and of space–time. As mentioned before, all three of the adopted triad vectors fortunately turn out to be 2-space normal. We may consequently expect to be able to introduce an orthogonal coordinate system in the quotient space.

Two of the intrinsic scalars, $\mathbf{a}\cdot\mathbf{v}$ and α, offer themselves as candidates for intrinsic co-moving coordinates, since we have verified that they are time independent. Actually, since these are related by Eq. (61), only one independent intrinsic coordinate is available. Any function of α and $\mathbf{a}\cdot\mathbf{v}$ might be selected, but a convenient choice proves to be $-(k\alpha)^{-1}$. According to Eq. (59) the triad vector $\mathbf{v}$ is normal to the 2-spaces α = constant, so we adopt the symbol y for this coordinate and solve for the scalars in terms of it:

$$\alpha = -1/ky, \tag{64}$$

$$(\mathbf{a}\cdot\mathbf{v})^2 = (\epsilon/y^2) - [2/(ky)^3] - \tfrac{1}{3}\Lambda. \tag{65}$$

Differentiating Eq. (64) we have

$$\nabla\alpha = (1/ky^2)\nabla y,$$

so that Eq. (59) may be transformed to

$$\mathbf{e}^2 \equiv \nabla y = y(\mathbf{a}\cdot\mathbf{v})\mathbf{v}, \tag{66}$$

where we have introduced the natural basis vector $\mathbf{e}^2$.

Intrinsic co-moving potentials for the other triad vectors are not available; this is connected with the fact that $\mathbf{u}$ and $\mathbf{w}$ were found to be Killing vectors. We know, however, that the space congruences generated by $\mathbf{u}$ and $\mathbf{w}$ are each body-fixed and 2-space-normal, and of course orthogonal to each other. Accordingly we may introduce orthogonal, body-fixed (co-moving) coordinates, say x and z respectively, by any arbitrary time-independent labeling of the 2-surfaces normal to these congruences. Introducing the natural basis vectors $\mathbf{e}^1$ and $\mathbf{e}^3$ associated with a particular labeling we can write

$$\mathbf{e}^1 \equiv \nabla x = (1/y)\mathbf{u}, \tag{67}$$

$$\mathbf{e}^3 \equiv \nabla z = [1/y(\mathbf{a}\cdot\mathbf{v})]\mathbf{w}. \tag{68}$$

The dual natural triad is clearly

$$\mathbf{e}_1 = y\mathbf{u}, \quad \mathbf{e}_2 = [1/y(\mathbf{a}\cdot\mathbf{v})]\mathbf{v}, \quad \mathbf{e}_3 = y(\mathbf{a}\cdot\mathbf{v})\mathbf{w}, \tag{69}$$

and the nonzero covariant metric coefficients of the quotient space are consequently given by

$$\begin{aligned} h_{11} &= \mathbf{e}_1\cdot\mathbf{e}_1 = y^2, \\ h_{22} &= \mathbf{e}_2\cdot\mathbf{e}_2 = [1/y^2(\mathbf{a}\cdot\mathbf{v})^2] \\ &= [\epsilon - (2/k^3y) - \tfrac{1}{3}\Lambda y^2]^{-1}, \\ h_{33} &= \mathbf{e}_3\cdot\mathbf{e}_3 = y^2(\mathbf{a}\cdot\mathbf{v})^2 = [\epsilon - (2/k^3y) - \tfrac{1}{3}\Lambda y^2]. \end{aligned} \tag{70}$$

Since the constant k occurs here only as k^{-3}, we shall henceforth write m, where $m = k^{-3}$ is a positive constant. Note that the metric coefficients are independent of x and z; this expresses the symmetry properties we expected from the existence of two Killing vectors, $\mathbf{u}$ and $\mathbf{w}$.

To write the space–time metric in co-moving coordinates as in Sec. II, we need only obtain an explicit expression for the potential ϕ, since we are employing the gauge $\mathbf{A} = 0$. To find ϕ we write the obvious equality

$$y\phi[\mathbf{a} - (\mathbf{a}\cdot\mathbf{u})\mathbf{u} - (\mathbf{a}\cdot\mathbf{v})\mathbf{v}] = 0, \tag{71}$$

and insert $\mathbf{a}$ from Eq. (14), $(\mathbf{a}\cdot\mathbf{u})$ from (62), $\mathbf{u}$ from (67), and $(\mathbf{a}\cdot\mathbf{v})\mathbf{v}$ from (66). The result of all these substitutions is

$$\nabla(y\phi) \pm (y\phi)[\epsilon'(y\phi)^2 - \epsilon]^{\frac{1}{2}}\nabla x = 0. \tag{72}$$

Consider first the case $\epsilon' = \epsilon = 1$; from Eq. (62) we find $(y\phi)^2 \geq 1$, so that an appropriate substitution is

$$y\phi = \operatorname{cosec}\theta.$$

With this substitution Eq. (72) becomes

$$\nabla\theta = \pm\nabla x,$$

and so

$$\theta = \pm[x + f(t)], \tag{73}$$

where $\nabla f = 0$, but f is allowed to be any function of the t coordinate, since, again, we have chosen the gauge $\mathbf{A} \equiv \nabla t = 0$. So the space–time interval in these co-moving coordinates has the form

$$\begin{aligned} \text{I.} \qquad & \epsilon = \epsilon' = 1, \qquad (y\phi)^2 \geq 1, \\ ds^2 &= y^2\{-\sin^2[x + f(t)]\,dt^2 + dx^2\} \\ &\quad + [1 - (2m/y) - \tfrac{1}{3}\Lambda y^2]^{-1}\,dy^2 \\ &\quad + [1 - (2m/y) - \tfrac{1}{3}\Lambda y^2]\,dz^2. \end{aligned} \tag{74}$$

Exactly analogous integrations of Eq. (72) for the remaining two permitted combinations of signs lead to

$$\begin{aligned} \text{II.} \qquad & \epsilon = -1, \qquad \epsilon' = 1, \qquad 0 \leq (y\phi)^2 \leq \infty, \\ ds^2 &= y^2\{-\sinh^2[x + f(t)]\,dt^2 + dx^2\} \\ &\quad + [-1 - (2m/y) - \tfrac{1}{3}\Lambda y^2]^{-1}\,dy^2 \\ &\quad + [-1 - (2m/y) - \tfrac{1}{3}\Lambda y^2]\,dz^2, \end{aligned} \tag{75}$$

and

$$\begin{aligned} \text{III.} \qquad & \epsilon = \epsilon' = -1, \qquad (y\phi)^2 \leq 1, \\ ds^2 &= y^2\{-\cosh^2[x + f(t)]\,dt^2 + dx^2\} \\ &\quad + [-1 - (2m/y) - \tfrac{1}{3}\Lambda y^2]^{-1}\,dy^2 \\ &\quad + [-1 - (2m/y) - \tfrac{1}{3}\Lambda y^2]\,dz^2. \end{aligned} \tag{76}$$

Returning now to the special case left behind, which was characterized precisely by $\mathbf{a}_\perp = \mathbf{a}\cdot\mathbf{v} = 0$, we see that it appears as a singular limit. After investigating some of its properties we can obtain the metrics for this situation. Reference to Eq. (50) shows that in this case $\nabla\mathbf{u} = 0$, so that $\mathbf{u}$ is a covariant constant vector in the quotient space. The integrability condition for this can be found using Eq. (5) of Sec. II;

$$\mathsf{E}\times\mathbf{u} = (-\mathsf{A} + \tfrac{1}{3}\Lambda\mathsf{I})\times\mathbf{u} = (\alpha^3 + \tfrac{1}{3}\Lambda)\mathbf{u}\times\mathsf{I} = 0,$$

where we have used Eqs. (10) and (46). Thus we require

$$\alpha^3 = -1/(ky)^3 = -\tfrac{1}{3}\Lambda.$$

The dyadics A and E become then

$$\mathsf{A} = -\tfrac{1}{3}\Lambda(3\mathbf{uu} - \mathsf{I}),$$

$$\mathsf{E} = \Lambda\mathbf{uu},$$

and we see that a nontrivial solution exists here only for $\Lambda \neq 0$. Following through a direct integration of the dyadic equations leads in one case to the space–time metric

IV. $\qquad \mathbf{a}_\perp = 0, \qquad \Lambda > 0,$

$$ds^2 = (1/\Lambda)\{-\sin^2[x + f(t)]\,dt^2 + dx^2 + d\theta^2 + \sin^2\theta\,d\phi^2\}. \tag{77}$$

However, this last may be more simply obtained as a singular limit at the coordinate singularity of metric I, Eq. (74). Note that the y coordinate becomes constant, $y = (3m/\Lambda)^{\frac{1}{3}}$; from Eq. (65) it follows that for $\mathbf{a}\cdot\mathbf{v} = 0$ we must further put $m = \frac{1}{3}(\epsilon/\Lambda)^{\frac{1}{3}}$ (requiring now $\epsilon/\Lambda > 0$), so $y = \epsilon(\epsilon/\Lambda)^{\frac{1}{3}}$ in the limit. This suggests inserting

$$m = \tfrac{1}{3}(\epsilon/\Lambda)^{\frac{1}{3}}(1 - \tfrac{3}{2}\epsilon\delta^2), \qquad y = \epsilon(\epsilon/\Lambda)^{\frac{1}{3}}(1 + \delta\bar{y}), \qquad z = (\epsilon/\Lambda)^{\frac{1}{3}}\,\delta^{-1}\bar{z},$$

and then letting the parameter δ approach zero. The result for all three metrics is

$$ds^2 = (\epsilon/\Lambda)\{-h^2[x + f(t)]\,dt^2 + dx^2 + (1 - \epsilon\bar{y}^2)^{-1}\,d\bar{y}^2 + (1 - \epsilon\bar{y}^2)\,d\bar{z}^2\},$$

where h is the appropriate circular or hyperbolic function in each case. The $(\bar{y}, \bar{z})$ 2-spaces have constant curvature Λ^{-1}. Thus, introducing polar coordinates, we see that metric I becomes IV, as in Eq. (77), and II and III become, respectively,

V. $\qquad \mathbf{a}_\perp = 0, \qquad \Lambda < 0,$

$$ds^2 = (1/-\Lambda)\{-\sinh^2[x + f(t)]\,dt^2 + dx^2 + d\theta^2 + \sinh^2\theta\,d\phi^2\}, \tag{78}$$

VI. $\qquad \mathbf{a}_\perp = 0, \qquad \Lambda < 0,$

$$ds^2 = (1/-\Lambda)\{-\cosh^2[x + f(t)]\,dt^2 + dx^2 + d\theta^2 + \cosh^2\theta\,d\phi^2\}. \tag{79}$$

The space–time metrics I, II, and III are well known. It is easily verified that the indefinite (x, t) 2-spaces in all cases have homogeneous, constant Riemannian curvatures, $\pm 1/y^2$. Thus, in the case of metric I, for instance, this 2-space could be written in terms of new, non co-moving coordinates $(\theta, \bar{t})$ as

$$y^2[-\sin^2\theta\,d\bar{t}^2 + d\theta^2],$$

and metric I then becomes

$$ds^2 = y^2\{-\sin^2\theta\,d\bar{t}^2 + d\theta^2\} + [1 - (2m/y) - \tfrac{1}{3}\Lambda y^2]^{-1}\,dy^2 + [1 - (2m/y) - \tfrac{1}{3}\Lambda y^2]\,dz^2. \tag{80}$$

Here the $\bar{t}$ lines are a timelike isometry of the manifold, and for $\Lambda = 0$ this is precisely the canonical form of a degenerate static vacuum metric of Class B-1 in the nomenclature of Ehlers and Kundt.[14] A similar rewriting of metrics II and III, bringing them to static form, shows that II is Class B-2 while III corresponds to the class of analytically extended metrics $\mathfrak{B}$-2, where the notation is again that of Ref. (14). Summarizing then we may state that, excluding spaces of constant curvature, the only Einstein spaces admitting nonisometric, irrotational, rigid congruences are the degenerate static metrics of Class B; and the special singular solutions IV, V, and VI, which require $\Lambda \neq 0$. This result is perfectly consistent with a general theorem recently presented by M. Trümper.[15]

In the co-moving coordinates, the world lines of the rigid frame for all these metrics are the lines t-varies and, so long as f is some function of t, the congruence is nonisometric. The process of direct construction we have followed demonstrates these to be the only nonisometric, irrotational, rigid congruences in Einstein spaces with nonvanishing Weyl tensor. In every case only one arbitrary function of time remains to be specified; in other words, irrotational rigid frames in exterior gravitational fields have no more than *one* degree of freedom. Anticipating the results of Sec. V, where it is shown that rotating rigid congruences allow no arbitrary time functions, we can actually generalize the preceding sentence by deleting the word "irrotational."

The relation between the co-moving coordinates of Eq. (74) and the static coordinates of Eq. (80) is interesting for its connection with the geometric construction for irrotational rigid motions in uniformly curved or flat space–time (as discussed in Refs. 9 and 16). Consider a 2-space of constant, unit curvature (but of hyperbolic metric, for the present cases). We *could* use polar coordinates θ, $\bar{t}$ in this 2-space, so that the metric would take the nonmaximal but simple form

$$dl^2 = -\sin^2\theta\,d\bar{t}^2 + d\theta^2. \tag{81}$$

---

[14] J. Ehlers and W. Kundt in *Gravitation: an introduction to current research*, edited by L. Witten (John Wiley & Sons, Inc., New York, 1962), Chap. 2.

[15] M. Trümper, J. Math. Phys. **6**, 584 (1965).

[16] A. D. Fokker, Rev. Mod. Phys. **21**, 406 (1949).

If we are given a spacelike curve $\mathcal{C}$ in this space, we may however prefer to introduce orthogonal coordinates $\bar{\theta}$, t based on $\mathcal{C}$ by the construction on the convex side of $\mathcal{C}$ shown in the figure. Tangent geodesics from $\mathcal{C}$ are taken as the lines t = constant. Orthogonal to them are the involutes of $\mathcal{C}$, the lines $\bar{\theta}$ = const.; the value of $\bar{\theta}$ labeling one of these is the proper distance to it along $\mathcal{C}$ from an origin 0 on $\mathcal{C}$. If the distance along $\mathcal{C}$ from 0 to the point of tangency of a curve t = const is $f(t)$, it follows by inspection that the form for interval ds^2 at any point P is being expressed in terms of "moving polar coordinates" based on a pole that migrates along $\mathcal{C}$ (see Fig. 1):

$$ds^2 = -\sin^2 [\bar{\theta} - f(t)]\, dt^2 + d\bar{\theta}^2. \tag{82}$$

In these coordinates, the equation for $\mathcal{C}$ itself is just $\bar{\theta} - f(t) = 0$. The congruence of involutes is clearly rigid, equidistant.

In Eq. (74) we have this construction if we identify $\bar{\theta}$ with $-x$. The curves $\mathcal{C}$ in the x, t pseudo-spherical subspaces form a spacelike 3-surface in space–time. All points in the rigid body achieve infinite acceleration on this 3-surface.

V. ROTATING RIGID FRAMES

We now inquire whether *any* nonisometric rotating rigid motions exist; and if so, the maximum number of degrees of freedom allowed, and the form of any constraints imposed on the time dependence of the acceleration, angular velocity or other variables.

Partial answers to some of these questions have been given in recent years.[7–9] Pirani and Williams[8] show how to construct (by means of a certain metric transformation) examples of space-time metrics admitting nonisometric rotating rigid congruences. Thus they have demonstrated that the Herglotz–Noether theorem is *not always* valid. Unfortunately, their discussion does not reveal explicitly the form of the Ricci tensor of the manifolds obtained, and it is apparently not at all clear whether the associated Einstein or stress tensor is physically reasonable. In particular, it seems not to be known whether any of their metrics include empty or Einstein spaces, nor whether the tangent vector of the rigid congruence is the timelike eigenvector of the stress tensor.

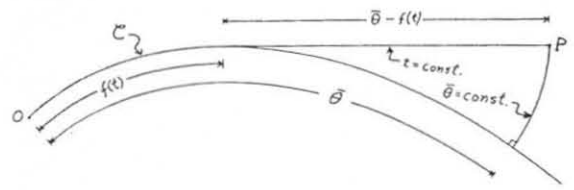

FIG. 1. Moving polar coordinates based on a curve $\mathcal{C}$.

For the case of rotating rigid frames in Einstein space, we have obtained a complete set of differential equations of first order in the local time derivatives, which govern the evolution of these motions. Thus we can state that in general Einstein space the entire history of rotating rigid motions is determined by initial conditions on any one spatial hypersurface, so that no arbitrary functions of time (degrees of freedom) are permitted. In a sense this incorporates much of the physical significance of the Herglotz–Noether theorem. Actually, as we shall see, the total set of equations is over-complete, since four subsidiary algebraic equations are obtained in addition to the differential equations. Isometric motions, for which all quantities are time independent, constitute a trivial solution of the total set; it may be that these are the *only* solutions consistent with the subsidiary equations, in which case the usual statement of the Herglotz–Noether theorem would hold. Unfortunately, the algebraic complexity of the equations has so far prohibited either reducing them to the conditions for isometries, or, conversely, demonstrating one or more particular nonisometric solutions.

The derivation of the equations depends crucially on the commutation of the space- and time-derivative operators for time-independent quantities [Eq. (4)], which as remarked in Sec. II is a special property of rigid motions. We start again with the Gauss equation (10) for rigid motion in Einstein space

$$\mathsf{E} = -\mathsf{A} - 3\boldsymbol{\Omega}\boldsymbol{\Omega} + \tfrac{1}{3}\Lambda\mathsf{I},$$

and the constraint $\dot{\mathsf{E}} = 0$. The trace gives

$$\mathrm{Tr}\,\mathsf{E} = \Lambda - 3\Omega^2,$$

and time-differentiating we find the important result of Rayner[7] that

$$\tfrac{1}{2}(\Omega^2)^{\cdot} = \boldsymbol{\Omega}\cdot\dot{\boldsymbol{\Omega}} = 0. \tag{83}$$

In view of this orthogonality property, we can define for later use a vector $\boldsymbol{\delta}$, by setting

$$(\Omega^2)\boldsymbol{\delta} \equiv \boldsymbol{\Omega} \times \dot{\boldsymbol{\Omega}}, \tag{84}$$

and in consequence express $\dot{\boldsymbol{\Omega}}$ as

$$\dot{\boldsymbol{\Omega}} = \boldsymbol{\delta} \times \boldsymbol{\Omega}, \tag{85}$$

where by definition

$$\boldsymbol{\delta}\cdot\dot{\boldsymbol{\Omega}} = \boldsymbol{\delta}\cdot\boldsymbol{\Omega} = 0. \tag{86}$$

Differentiating this last equation we have

$$\dot{\mathsf{o}}\cdot\Omega = 0. \tag{87}$$

The procedure now is to construct additional time-constant dyadics from the spatial derivatives of E. The existence of a chain of such quantities resulting from repeated spatial differentiation was previously pointed out by Pirani and Williams,[8] though not exploited by them. The success of this procedure depends on the fact that the differentiation process may be selectively applied to obtain only equations which do not contain spatial derivatives of time-dependent quantities; furthermore, with this criterion, the procedure terminates naturally, so that finally a complete set results—all first-order differential equations *in time only*.

We first define the dyadic F by

$$\mathsf{F} \equiv -\nabla\times\mathsf{E} = \mathsf{F}^s - \mathsf{f}\times\mathsf{I}, \tag{88}$$

where F^s and f are, respectively, symmetric and antisymmetric parts of F, and by commutation ($\dot{\mathsf{E}} = 0$) it follows that

$$\dot{\mathsf{F}} = \dot{\mathsf{F}}^s = \mathsf{f} = 0. \tag{89}$$

Some of the other properties of F can be obtained before calculating its explicit form. Since E is symmetric, we have $\mathrm{Tr}\,\mathsf{F} = \mathrm{Tr}\,\mathsf{F}^s = 0$, and using the fact that $\nabla\cdot\mathsf{E} = 0$ we find

$$\mathsf{f} = \tfrac{1}{2}\nabla(\mathrm{Tr}\,\mathsf{E}) = -\tfrac{3}{2}\nabla(\Omega^2). \tag{90}$$

From this it follows that

$$\nabla\times\mathsf{f} = 0. \tag{91}$$

Now taking the divergence of Eq. (88) gives

$$\nabla\cdot\mathsf{F} = -\nabla\cdot(\nabla\times\mathsf{E}) = \nabla\cdot\mathsf{F}^s - \nabla\times\mathsf{f},$$

but from the commutation relation Eq. (D57) of Paper I we have

$$\nabla\cdot\nabla\times\mathsf{E} = -\mathsf{E}\dot{\times}\mathsf{E} + 2\Omega\cdot\dot{\mathsf{E}} = 0,$$

so that finally also

$$\nabla\cdot\mathsf{F}^s = 0. \tag{92}$$

We obtain the explicit expression for F by substituting from Eqs. (15), (17), and (19) into

$$\begin{aligned}\mathsf{F} = -\nabla\times\mathsf{E} &= \nabla\times(\mathsf{A} + 3\Omega\Omega - \tfrac{1}{3}\Lambda\mathsf{I})\\ &= \nabla\times\mathsf{A} + 3(\nabla\times\Omega)\Omega - 3\Omega\times\nabla\Omega,\end{aligned}$$

and upon resolving find for F^s and f:

$$\begin{aligned}\mathsf{F}^s = \dot{\mathsf{B}} &- \Omega\times\mathsf{B} + \mathsf{B}\times\Omega + \mathsf{a}\times\mathsf{E} - \mathsf{E}\times\mathsf{a}\\ &+ 3\Omega\Omega\times\mathsf{a} - 3\mathsf{a}\times\Omega\Omega,\end{aligned} \tag{93}$$

$$\tfrac{1}{3}\mathsf{f} = 2(\Omega^2)\mathsf{a} - (\mathsf{a}\cdot\Omega)\Omega - \Omega\cdot\mathsf{B}. \tag{94}$$

Proceeding to the next level of spatial differentiation we define the dyadic G by

$$\mathsf{G} \equiv \nabla\times[\mathsf{F}^s + \tfrac{1}{3}\mathsf{f}\times\mathsf{I}] = \mathsf{G}^s - \mathsf{g}\times\mathsf{I}, \tag{95}$$

and as before, of course, have

$$\dot{\mathsf{G}} = \dot{\mathsf{G}}^s = \dot{\mathsf{g}} = 0. \tag{96}$$

Indeed from Eq. (95) we can show that

$$\mathsf{g} = \tfrac{1}{2}[\nabla\cdot\mathsf{F}^s - \tfrac{1}{3}\nabla\times\mathsf{f} - \nabla(\mathrm{Tr}\,\mathsf{F}^s)] = 0 \tag{97}$$

so that G is symmetric; and tracing Eq. (95) we have

$$\mathrm{Tr}\,\mathsf{G} = \mathrm{Tr}\,\mathsf{G}^s = -\tfrac{2}{3}\nabla\cdot\mathsf{f}. \tag{98}$$

To calculate the explicit form of G^s, Eqs. (93) and (94) are inserted into (95), and then, using virtually all the dyadic equations of Sec. II, it is possible to eliminate the spatial derivatives of all time-dependent quantities. This was the criterion for the particular choice of G. To assist in the elimination, it is convenient to introduce a new, symmetric, time-constant, dyadic G^k by

$$\begin{aligned}\mathsf{G}^k \equiv\ &\mathsf{G}^s + 3\mathsf{E}\cdot\mathsf{E} - 3(\Lambda - \Omega^2)\mathsf{E}\\ &- (\mathsf{E}:\mathsf{E} - 10\Omega^4 + \tfrac{23}{3}\Lambda\Omega^2 - \Lambda^2)\mathsf{I}.\end{aligned} \tag{99}$$

The lengthy result of all the substitutions then becomes

$$\begin{aligned}\mathsf{G}^k =\ &3\dot{\mathsf{o}}\times\Omega\Omega - 3\Omega\Omega\times\dot{\mathsf{o}} - \dot{\mathsf{a}}\times\mathsf{B} + \mathsf{B}\times\dot{\mathsf{a}}\\ &+ (\mathsf{o}\times\Omega)\times\mathsf{E} - \mathsf{E}\times(\mathsf{o}\times\Omega) + 6\mathsf{o}\times\Omega\mathsf{o}\times\Omega - 6\sigma^2\Omega\Omega\\ &- 2\Omega^2(\mathsf{o}\Omega + \Omega\mathsf{o}) - 5\Omega^2\mathsf{a}\mathsf{a} + (2\Omega^2 + \tfrac{4}{3}a^2 - \tfrac{8}{3}\Lambda)\Omega\Omega\\ &+ [\mathsf{E} + \tfrac{4}{3}\Omega\Omega - \tfrac{1}{3}\Lambda\mathsf{I}]\overset{\times}{\times}[3\mathsf{a}\mathsf{a} + 2\Omega\Omega + a^2\mathsf{I}]\\ &- 2[\mathsf{B} + (\mathsf{a}\cdot\Omega)\mathsf{I} - 2\mathsf{a}\Omega]\overset{\times}{\times}[\mathsf{B} + (\mathsf{a}\cdot\Omega)\mathsf{I} - 2\mathsf{a}\Omega]\\ &- \mathsf{a}\cdot\mathcal{E} - 2\mathsf{a}\times[\mathsf{F}^s + \tfrac{1}{3}\mathsf{f}\times\mathsf{I}] + 2[\mathsf{F}^s - \tfrac{1}{3}\mathsf{f}\times\mathsf{I}]\times\mathsf{a}\\ &+ \tfrac{1}{3}[20a^2\Omega^2 - \Lambda a^2 - 9\Omega\cdot\mathsf{E}\cdot\Omega + 3\mathsf{a}\cdot\mathsf{E}\cdot\mathsf{a}]\mathsf{I},\end{aligned} \tag{100}$$

where the symbol $\mathcal{E}$ denotes the time-constant triadic $\mathcal{E} \equiv \nabla\mathsf{E}$. Taking the trace of this equation, we have

$$\mathrm{Tr}\,\mathsf{G}^k = 2\mathsf{B}:\mathsf{B} - 8\Omega\cdot\mathsf{E}\cdot\Omega + \tfrac{10}{3}\mathsf{a}\cdot\mathsf{f} + 5\Omega^2(\Omega^2 - \Lambda). \tag{101}$$

This completes the derivation of the equations. They are all expressed in terms of the following two sets of quantities: (1) a, Ω, o, B; and (2) E, F^s, f, G^k, $\mathcal{E}$. The second set contains only quantities whose time derivatives vanish—they can therefore be treated as arbitrary constants whose values may be selected at any point on each world line of the congruence. The first set contains 14 scalars whose time dependence is governed, either explicitly or

implicitly, by the 14 scalar equations included in Eqs. (85), (87), (93), and the trace-free part of (100). These equations are linear in, and could in principle all be explicitly solved for, all the first time derivatives. This is indeed already accomplished for $\dot{\Omega}$ and $\dot{\mathsf{B}}$. Further differentiation and substitution back would then generate the time derivatives of all order in terms of undifferentiated quantities, *so that the entire evolution on each line is determined by conditions at one point.*

We have four algebraic constraints left over: Eqs. (94) and (101). These must hold throughout time, so that by differentiating them and substituting back for all time derivatives, we should obtain a further set of four, in general new, algebraic constraints. These in turn could be treated in identical fashion, and so on and on. If this process continues to generate a chain of independent equations, we should soon be able to solve algebraically for a, Ω, σ, and B in terms of the second, time-constant, set of quantities; so *all* the quantities would be time constant, and the reduction to isometric motions accomplished. This, however, remains to be done, and so we cannot exclude the possibility that a very few, very special, solutions might exist, for which the above described chain of algebraic manipulation would stop, self-consistent, but short of the complete reduction to a timelike isometry.

Determination of the Parameters of Model Field Theories from Phase Shifts and Bound-States Energies*

LUCIANO FONDA AND GIAN' CARLO GHIRARDI
Istituto di Fisica Teorica dell'Università, Trieste, Italy
(Received 22 November 1965)

The inversion problem is taken into account and solved in the framework of some models of quantum field theories. Both cases of one and multichannels models are considered. The conditions that must be satisfied by the phase shifts in order that they can be reproduced by the considered models are discussed and relations are derived among the scattering phase shifts, the energies of the bound states and the parameters of the models. It turns out that in certain cases the parameters are completely determined while in other cases some freedom is left.

1. INTRODUCTION

THE problem of finding the structure of the underlying interparticle interaction from knowledge of the phase shift and the energies of the bound states, has been extensively discussed in the framework of potential scattering.[1]

In this paper we show that this can also be done for two classes of quantum field theoretic models, provided that the scattering phase shifts meets certain conditions necessary to make the models reasonable.

We first consider a one-channel model[2] which describes the interaction of a spinless particle with a fixed scatterer, the mechanism of the interaction being such that compound states can be formed. In this case the inversion problem can be solved, the solution may in general depend on some free parameters. The interest in such a model arises from the fact that the situation it describes is not reproducible in potential scattering.

After having discussed the Lee model as a particular case, we take into account a multichannel model in which a separable direct interaction is assumed between the fixed scatterer and a spinless particle capable of a certain number of excited states. Also in this case the inversion problem can be solved, the solution now being completely determined by knowledge of the phase shifts and the bound-state energies.

2. THE DYSON MODEL

Let us consider a spinless particle ϑ of mass μ interacting with a fixed scatterer. The system incident particle plus scatterer gives rise to a certain number N_0 of compound states. The Hamiltonian for this model is (natural units $\hbar = c = 1$ are used)

$$H = H_0 + H_I,$$

$$H_0 = m_N \psi_N^\dagger \psi_N + \sum_{i=1}^{N_0} (m_N + \omega_i^{(0)}) \psi_i^\dagger \psi_i + \int \omega(\mathbf{k}) a^\dagger(\mathbf{k}) a(\mathbf{k})\, d^3\mathbf{k}, \qquad (2.1)$$

$$H_I = (2\pi)^{-\frac{3}{2}} \int \frac{f(\omega)}{(2\omega)^{\frac{1}{2}}} \left[\sum_{i=1}^{N_0} g_i \psi_i^\dagger \psi_N a(\mathbf{k}) + \text{h.c.} \right] d^3\mathbf{k}.$$

Here m_N and $m_N + \omega_i^{(0)}$ are the energies of the ground state of the scatterer and of the ith compound state, respectively; $\omega = (k^2 + \mu^2)^{\frac{1}{2}}$; ψ_N and ψ_i are annihilation operators for the scatterer in its ground state and for the ith compound state, respectively; $a(\mathbf{k})$ is the destruction operator for the ϑ particle. The form factor $f(\omega)$ is the Fourier transform of $U(r)$, the real spherically symmetric function of the position furnishing the spatial shape of the scatterer. We assume that $f(\mu) = 1$, as it follows from the normalization condition $\int U(r)\, d^3r = 1$. The quantity g_i is the coupling constant for the transition ground state $\rightarrow$ ith compound state and is real, as it follows from time reversal invariance and Hermiticity.

We consider here only the scattering of the ϑ particle on the scatterer in its ground state. As pointed out by Dyson, during the scattering process transitions occur to the various compound states almost in the same way that in nuclear reactions the compound nucleus states are formed at intermediate times of the collision process. Since in the model the compound states are not directly coupled among each other, all the scattering is elastic; besides, only the S wave is scattered.

The discrete eigenvalues of H in the considered $N - \vartheta$ sector are the roots of the equation:

* Work supported in part by the Istituto Nazionale di Fisica Nucleare.

[1] See the review article by R. G. Newton, J. Math. Phys. 1, 319 (1960) and references contained therein.

[2] F. J. Dyson, Phys. Rev. 106, 157 (1956).

$$\det \Big| \delta_{ij}(\omega - \omega_i^{(0)}) + \lim_{\epsilon \to 0} \int_\mu^\infty \frac{(y^2 - \mu^2)^{\frac{1}{2}} f^2(y)}{4\pi^2 (y - \omega - i\epsilon)} \, dy \, g_i g_j \Big| = 0. \tag{2.2}$$

Note that $f^2(\omega)$ must vanish at infinity more rapidly than ω^{-1} in order that the integral appearing in (2.2) be convergent. The solutions of (2.2) are all real. A necessary condition to have a solution of (2.2) for $\omega > \mu$ is that $f^2(\omega) = 0$ at that point. We assume in the following that

$$f^2(\omega) > 0, \qquad \omega > \mu, \tag{2.3}$$

which forbids the occurrence of bound states embedded in the continuum. Equation (2.2) will have, therefore, solutions only for $\omega < \mu$. By induction it can easily be seen that (2.2) can be put in the form

$$\prod_{i=1}^{N_0} (\omega - \omega_i^{(0)}) + \int_\mu^\infty \frac{(y^2 - \mu^2)^{\frac{1}{2}} f^2(y)}{4\pi^2 (y - \omega)} \, dy \times \sum_{i=1}^{N_0} g_i^2 \prod_{\substack{j=1 \\ j \neq i}}^{N_0} (\omega - \omega_i^{(0)}) = 0. \tag{2.4}$$

The roots of (2.4) are all different from the $\omega_i^{(0)}$. There follows that the roots of (2.4) are also real zeros of the analytic function

$$D(z) = \left(\sum_{i=1}^{N_0} \frac{g_i^2}{z - \omega_i^{(0)}} \right)^{-1} + \frac{1}{\pi} \int_\mu^\infty \frac{\varphi(y)}{y - z} \, dy, \tag{2.5}$$

where

$$\varphi(\omega) = \frac{(\omega^2 - \mu^2)^{\frac{1}{2}}}{4\pi} f^2(\omega). \tag{2.6}$$

Note, however, that $D(z)$ does not have other zeros besides the (real) roots of (2.4). So, we have a one to one correspondence between the zeros of $D(z)$ and the discrete eigenvalues of H.

The resolvent operator $G(z) = (z - H)^{-1}$ satisfies

$$\langle 0 | \, \psi_i G(z) \psi_i^\dagger \, | 0 \rangle = \delta_{ij}(z - \omega_i^{(0)})^{-1} - (z - \omega_i^{(0)})^{-1} \frac{1}{\pi} \int_\mu^\infty \frac{\varphi(y)}{y - z} \, dy \, g_i \times \sum_{l=1}^{N_0} g_l \langle 0 | \, \psi_l G(z) \psi_i^\dagger \, | 0 \rangle,$$

$|0\rangle$ being the vacuum state. Using this equation the scattering amplitude is immediately obtained on the energy shell

$$T(\omega) = \frac{f^2(\omega)}{16\pi^3 \omega} \sum_{i,j=1}^{N_0} g_i g_j \langle 0 | \, \psi_i G(\omega + i0) \psi_j^\dagger \, | 0 \rangle = \frac{f^2(\omega)}{16\pi^3 \omega} [D(\omega + i0)]^{-1}. \tag{2.7}$$

The S matrix is then given by

$$S(\omega) \equiv 1 - 8i\pi^2 \omega (\omega^2 - \mu^2)^{\frac{1}{2}} T(\omega) = D(\omega - i0)/D(\omega + i0) = e^{2i\delta(\omega)}. \tag{2.8}$$

Under the above hypotheses, the Levinson theorem is immediately derived. Let us consider the following integral:

$$\frac{1}{2i} \int_C \frac{[D(z)]'}{D(z)} \, dz = \frac{1}{2i} \int_{C_\infty} \frac{[D(z)]'}{D(z)} + \frac{1}{2i} \int_{C_1 + C_2} d \log D(z),$$

where $C = C_1 + C_2 + C_3 + C_\infty$, and the contours C_1, C_2, C_3, C_∞ are shown in Fig. 1. The small semicircle C_2 does not give any contribution since we exclude the occurrence of a zero energy resonance and the accidental coincidence of a pole of $D(z)$ with μ. From the behavior of $D(z)$ for $z \to \infty$:

$$D(z) \underset{z \to \infty}{\sim} z \Big/ \sum_{i=1}^{N_0} g_i^2, \tag{2.9}$$

the first term at the right-hand side can be evaluated giving the value π. To evaluate the second term we note that $D(z)$ has poles on the real axis. These are the so called *CDD* poles for the considered model.[3] The function $D(z)$ can be written as follows:

$$D(z) = \frac{\prod_{i=1}^{N_0} (z - \omega_i^{(0)})}{\sum_{i=1}^{N_0} g_i^2 \cdot \prod_{j=1}^{N_0 - 1} (z - z_j^{(0)})} + \frac{1}{\pi} \int_\mu^\infty \frac{\varphi(y)}{y - z} \, dy. \tag{2.10}$$

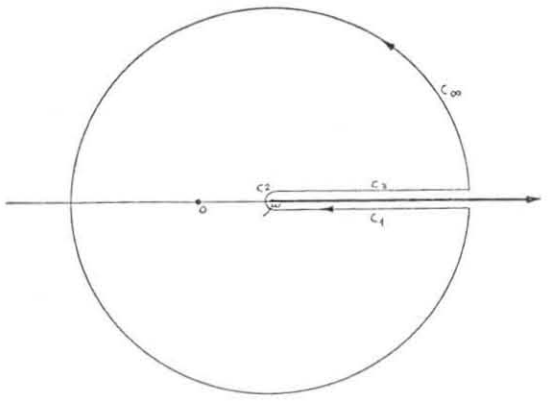

FIG. 1. The path of integration used to derive the Levinson theorem.

[3] L. Castillejo, R. H. Dalitz, and F. J. Dyson, Phys. Rev. 101, 453 (1956).

The first term at the right-hand side has exactly $N_0 - 1$ poles; on the other hand, it is equal to

$$\left[\sum_{i=1}^{N_0} \frac{g_i^2}{z - \omega_i^{(0)}}\right]^{-1},$$

which has $N_0 - 1$ real poles and N_0 real zeros. We conclude that the $z_i^{(0)}$ are all real. Let us assume that P_0 of them fall in the interval $\mu < \omega < +\infty$. We then get

$$\frac{1}{2i}\int_{C_1+C_0} d \log D(z) = \delta(\mu) - \delta(\infty) + \frac{1}{2i}\int_{C_0} \frac{[D(z)]'}{D(z)}\, dz,$$

where we have indicated by C_0, P_0 small clockwise circles surrounding the poles on the real axis in the interval $\mu < \omega < +\infty$. Each of these gives a contribution equal to π, so we finally get

$$\frac{1}{2i}\int_C \frac{[D(z)]'}{D(z)}\, dz = \delta(\mu) - \delta(\infty) + (P_0 + 1)\pi.$$

On the other hand, the same integral can be evaluated by means of the Cauchy integral formula giving

$$\frac{1}{2i}\int_C \frac{[D(z)]'}{D(z)}\, dz = (P_0 + 1 + n - N_0)\pi,$$

where n is the number of zeros of $D(z)$, i.e., the number of the discrete eigenvalues of H. Collecting the results we finally have

$$\delta(\mu) - \delta(\infty) = \pi(n - N_0).$$

We choose the phase shift in such a way that $\delta(\infty) = 0$, getting, therefore,

$$\delta(\mu) = \pi(n - N_0). \tag{2.11}$$

3. THE INVERSION PROBLEM IN THE DYSON MODEL

From the considerations of Sec. 2 it follows that $D(z)$ is analytic in the cut complex z plane, it has n real simple zeros for $\omega < \mu$ at the positions of the discrete eigenvalues ω_i of H, $N_0 - P_0 - 1$ real simple poles for $\omega < \mu$ and P_0 real simple poles on the branch cut. Moreover, Eq. (2.9) gives the behavior of $D(z)$ for $z \to \infty$.

We take now into account the function

$$\bar{D}(z) = \exp\left[-\frac{1}{\pi}\int_\mu^\infty \frac{\delta(y)}{y - z}\, dy\right]. \tag{3.1}$$

It is an analytic function in the cut z plane, which goes to one for $z \to \infty$, it never vanishes and its behavior for $z \to \mu$ is easily obtained by means of an integration by parts:

$$\bar{D}(z) = \exp\left[\frac{1}{\pi}\delta(\mu)\log(\mu - z) + \frac{1}{\pi}\int_\mu^\infty \log(y - z)\, d\delta(y)\right]$$

By making use of (2.11) we therefore have that

$$\bar{D}(z) \underset{z\to\mu}{\sim} \text{const}\,(\mu - z)^{n-N_0} \tag{3.2}$$

Moreover, we have

$$\frac{\bar{D}(\omega - i0)}{\bar{D}(\omega + i0)} = e^{2i\delta(\omega)} \tag{3.3}$$

Comparing this equation with Eq. (2.8) we can conclude that the function

$$R(z) = D(z)/\bar{D}(z). \tag{3.4}$$

has no branch cut, has only isolated polar singularities and zeros, and for $z \to \infty$ behaves like $z/\sum_{i=1}^{N_0} g_i^2$. It is therefore the following rational function

$$R(z) = \frac{\left[\prod_{i=1}^{n}(z - \omega_i)\right](z - \mu)^{N_0 - n}}{\sum_{i=1}^{N_0} g_i^2 \prod_{j=1}^{N_0-1}(z - z_j^{(0)})}. \tag{3.5}$$

We finally get for the function $D(z)$ the following result:

$$D(z) = R(z)\exp\left[-\frac{1}{\pi}\int_\mu^\infty \frac{\delta(y)}{y - z}\, dy\right]. \tag{3.6}$$

Observing now that

$$2i\varphi(\omega) = D(\omega + i0) - D(\omega - i0)$$

we finally get

$$f^2(\omega) = -\frac{4\pi R(\omega)}{(\omega^2 - \mu^2)^{\frac{1}{2}}} \times \exp\left[-\frac{P}{\pi}\int_\mu^\infty \frac{\delta(y)}{y - \omega}\, dy\right]\sin\delta(\omega), \tag{3.7}$$

as the expression of the square of the form factor $f^2(\omega)$ in terms of the scattering phase shift $\delta(\omega)$, the number and the energies of the bound states and the N_0 parameters $z_j^{(0)}$, $\sum_{i=1}^{N_0} g_i^2$. P means that the Cauchy principal value of the integral has to be taken.

In order to see whether the form factor so obtained meets the requirements which are necessary to make it reasonable, we pay some attention to the bound-state equation which can be written as

$$\sum_{i=1}^{N_0} \frac{g_i^2}{\omega - \omega_i^{(0)}} = -[\phi(\omega)]^{-1}, \tag{3.8}$$

where

$$\phi(\omega) = \frac{P}{\pi}\int_{\mu}^{\infty}\frac{\varphi(y)}{y-\omega}\,dy. \tag{3.9}$$

The derivative of the left-hand side of Eq. (3.8) is always negative, while the derivative of $-[\phi(\omega)]^{-1}$ is positive for $\omega < \mu$. Calling M_0 the number of $\omega_i^{(0)}$ which are less than μ, we have only the two following possibilities

(A) $$n = M_0 + 1,$$

(B) $$n = M_0.$$

In case B, we moreover distinguish two cases according to the number P_0 of $z_i^{(0)}$, which fall at the right of $\omega = \mu$:

(B′) $$P_0 = N_0 - M_0 - 1,$$

(B″) $$P_0 = N_0 - M_0.$$

Note that in case A we certainly have $P_0 = N_0 - M_0 - 1$. The behavior of the two sides of Eq. (3.8) is shown in Fig. 2 for the three cases now considered. The sign of $R(\omega)$ for ω in a right neighborhood of μ is governed by the number P_0, since for $\omega > z_{N_0-1}^{(0)}$, $R(\omega)$ is positive, and crossing each $z_i^{(0)}$, $R(\omega)$ changes its sign.

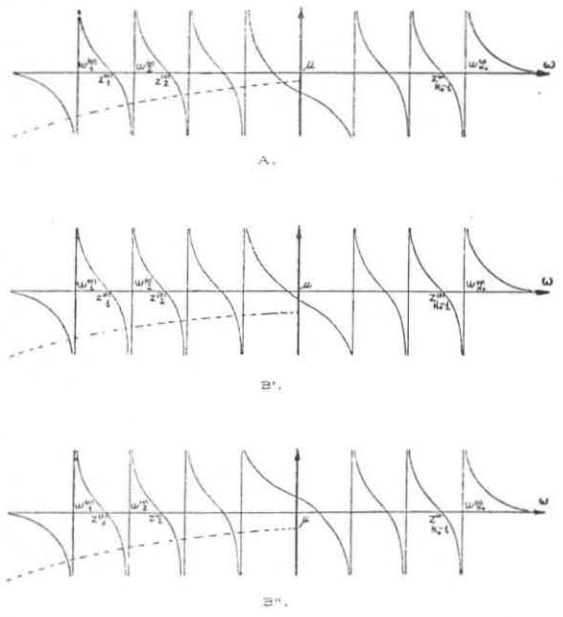

FIG. 2. The behavior of the functions $\sum_{i=1}^{N_0} g_i^2/(\omega - \omega_i^{(0)})$ (solid line) and $-[\phi(\omega)]^{-1}$ (dashed line) of Eq. (3.8) in the three cases A, B′, and B″. The intersections of the two curves give the energies of the normalizable eigenstates of H. The $\omega_i^{(0)}$ are the energies of the discrete eigenstates of H_0 and the $z_i^{(0)}$ the positions of the *CDD* poles for the model.

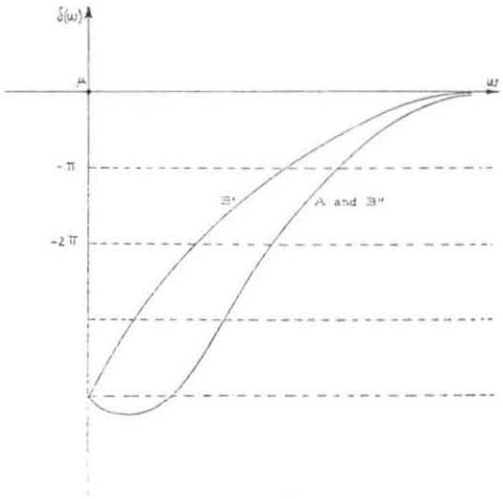

FIG. 3. The behavior of the phase shift $\delta(\omega)$ in the Dyson model for the three cases A, B′, and B″ considered in Sec. 3.

We discuss now the behavior of the phase shift in the three above cases.

(A) $n = M_0 + 1$; $P_0 = N_0 - M_0 - 1$. The Levinson theorem tells us that in this case $\delta(\mu) = \pi(M_0 + 1 - N_0)$. The phase shift therefore starts from a negative multiple of π or from zero and goes to zero for $\omega \to +\infty$. If P_0 is even, $R(\omega) > 0$ for ω in a right neighborhood of μ, while if P_0 is odd $R(\omega) < 0$ there. In order that $\varphi(\omega)$ be positive sin $\delta(\omega)$ must be negative for P_0 even and positive for P_0 odd in the same neighborhood $\omega \gtrsim \mu$. This implies that the phase shift starts decreasing. Since $\delta(\infty) = 0$ and $\delta(\omega)$ is a continuous function of ω, sin $\delta(\omega)$ vanishes at least $N_0 - M_0 - 1$ times in the interval $\mu < \omega < +\infty$. This number equals the P_0 number of times which $R(\omega)$ changes sign, therefore, in order to have a positive definite $\varphi(\omega)$ the behavior of the phase shift must be the one indicated in Fig. 3. The P_0 poles $z_i^{(0)}$ of $R(\omega)$ for $\omega > \mu$ are therefore uniquely determined since they must coincide with the zeros of sin $\delta(\omega)$.

(B′) $n = M_0$; $P_0 = N_0 - M_0 - 1$. From the Levinson theorem $\delta(\mu) = \pi(M_0 - N_0)$. Also in this case $R(\omega) > 0$ if P_0 is even and $R(\omega) < 0$ if P_0 is odd for $\omega \gtrsim \mu$. The requirement that $\varphi(\omega)$ be positive definite for $\omega > \mu$ implies then that the phase shift starts increasing. In the interval $\mu < \omega < +\infty$ sin $\delta(\omega)$ vanishes at least P_0 times, i.e., as many times as $R(\omega)$ changes its sign. It follows that the phase shift must have the behavior indicated in Fig. 3 and that all the $z_i^{(0)}$ for $\omega > \mu$ are determined.

(B″) $n = M_0$; $P_0 = N_0 - M_0$. From the Levinson theorem $\delta(\mu) = \pi(M_0 - N_0)$. Repetition of the

same argument as before shows that the phase shift must start decreasing and must have the behavior shown in Fig. 3. The $z_i^{(0)}$ for $\omega > \mu$ are therefore also in this case uniquely determined by the zeros of $\sin \delta(\omega)$.

We are now in the position to discuss all the details of the inversion problem in the considered model. We assume that the masses m_N and μ, the behavior of the phase shift $\delta(\omega)$ from μ to ∞ and the number and energies ω_i of the bound states of H are known. Moreover we must suppose that the phase shift has one of the behaviors shown in Fig. 3. If this is not the case it is not possible to describe the assumed situation in the framework of this model. Besides, note that the situations described in Fig. 3 cannot be reproduced in potential scattering. The parameters of the model are: the number N_0 and the energies $m_N + \omega_i^{(0)}$ of the normalizable eigenstates of H_0, the coupling constants g_i and the form factor $f(\omega)$. We can, in place of the g_i, consider the *CDD* poles, which are real and in number equal to $N_0 - 1$, and the sum $\sum_{i=1}^{N_0} g_i^2$.

As can be seen from Fig. 2 in cases A and B'' between the energies of two contiguous bound states, between two contiguous zeros of the phase shift and between the energy of the lightest bound state and the first zero of the phase shift falls one and only one $\omega_i^{(0)}$. In case B', instead, we have an additional discrete eigenvalue of H_0 between the energy of the lightest bound state and the first zero of the phase shift.

We observe that the behavior of the phase shift allows us to decide if we are in one of the cases A or B'' (phase shift decreasing departing from $\omega = \mu$) or in the case B' (phase shift increasing). Therefore the assignment of the phase shift, satisfying the above-mentioned requirements, and of the number of the bound states, uniquely determines N_0, more precisely

$N_0 = n + P_0$, the phase shift

starts decreasing, cases A and B''.

$N_0 = n + P_0 + 1$, the phase shift

starts increasing, case B'.

Moreover, we know approximately the positions of the eigenvalues $\omega_i^{(0)}$ of H_0, and, as we have already remarked, the *CDD* poles $z_i^{(0)}$ falling above $\omega = \mu$ which are determined by the zeros of the phase shift.

Let us now see which other relations we have at our disposal to determine the parameters of the model which are still free, i.e., the exact value of the N_0 quantities $\omega_i^{(0)}$, the $N_0 - P_0 - 1$ *CDD* poles $z_i^{(0)}$ which fall below $\omega = \mu$, the constant $\sum_{i=1}^{N_0} g_i^2$ and the form factor $f(\omega)$. First of all we recall that the form factor is given by Eq. (3.7) in terms of the parameters contained in $R(\omega)$. We must, moreover, take into account the n relations satisfied by the known eigenvalues ω_i of H:

$$1 + \phi(\omega_i) \frac{\sum_{j=1}^{N_0} g_j^2 \left[\prod_{j=1}^{N_0 - 1} (\omega_i - z_j^{(0)}) \right]}{\prod_{k=1}^{N_0} (\omega_i - \omega_k^{(0)})} = 0,$$

$$i = 1, \cdots, n. \qquad (3.10)$$

Note that in $\phi(\omega_i)$ are contained free parameters to be determined, exactly those which are necessary to fix the form factor. Note, moreover, that the constant $\sum_{i=1}^{N_0} g_i^2$ disappears from Eq. (3.10) since its reciprocal is contained as a multiplying factor in $\phi(\omega_i)$.

Other relations which must be taken into account are the equations giving the energies ω_R at which the resonances occur, which are determined by the phase shift

$$1 + \phi(\omega_R) \frac{\sum_{j=1}^{N_0} g_j^2 \prod_{j=1}^{N_0 - 1} (\omega_R - z_j^{(0)})}{\prod_{k=1}^{N_0} (\omega_R - \omega_k^{(0)})} = 0. \qquad (3.11)$$

Again $\sum_{i=1}^{N_0} g_i^2$ drops out. There are only $N_0 - n$ independent relations of the type (3.11).

This can be seen in general; it is instructive to show it explicitly in the case $N_0 = 2$ with no bound states. Let us suppose that in this case we have more than 2 relations of the type (2.11). We consider only 3 of them which we call α, β, and γ. We replace now β and γ by the differences $\alpha - \beta$ and $(\alpha - \beta) - (\alpha - \gamma)$, getting

$$(\omega_\alpha - \omega_1^{(0)})(\omega_\alpha - \omega_2^{(0)}) = (z_1^{(0)} - \omega_\alpha) \times \frac{P}{\pi} \int_\mu^\infty \frac{(\omega - \mu)^2 \bar{D}^P(\omega) \sin \delta(\omega)}{(z_1^{(0)} - \omega)(\omega - \omega_\alpha)} d\omega, \qquad (3.12)$$

$$\omega_\alpha + \omega_\beta - \omega_1^{(0)} - \omega_2^{(0)} = \frac{P}{\pi} \int_\mu^\infty \frac{(\omega - \mu)^2 \bar{D}^P(\omega) \sin \delta(\omega)}{(\omega - \omega_\alpha)(\omega - \omega_\beta)} d\omega, \qquad (3.13)$$

$$1 \equiv \frac{P}{\pi} \int_\mu^\infty \frac{(\omega - \mu)^2 \bar{D}^P(\omega) \sin \delta(\omega)}{(\omega - \omega_\alpha)(\omega - \omega_\beta)(\omega - \omega_\gamma)} d\omega \equiv \mathfrak{I}^P, \qquad (3.14)$$

where $\bar{D}^P(\omega)$ is defined for real ω by (3.1) with the principal value prescription for the integral appearing in the exponential. We immediately see that

the third of these equations is independent of the parameters $\omega_1^{(0)}$, $\omega_2^{(0)}$, and $z_1^{(0)}$. Therefore, the number of independent equations of the type (3.11) cannot exceed the minimum number of resonances, in this case 2. We have only to convince ourselves that Eq. (3.14) is an identity.

Using now the relation

$$-2i \sin \delta(\omega)\bar{D}^P(\omega) = \bar{D}(\omega + i0) - \bar{D}(\omega - i0), \tag{3.15}$$

we get for the right-hand side of Eq. (3.14)

$$\mathfrak{I}^P = 1 + \frac{i}{4\pi}\left[\int_{C_1+C_2+C_3+C_\infty} dz \frac{(z-\mu)^2\bar{D}(z)}{(z-\omega_\alpha - i\epsilon)(z-\omega_\beta - i\epsilon)(z-\omega_\gamma - i\epsilon)} - \text{c.c.}\right],$$

where the limit $\epsilon \to 0$ has to be taken after having performed the integration on the contour $C_1 + C_2 + C_3 + C_\infty$ shown in Fig. 1. We have added the contour $C_2 + C_\infty$ which gives the global contribution (-1). The integral is immediately evaluated by means of the Cauchy integral formula; we get

$$\mathfrak{I}^P = 1 - \frac{1}{2}\sum_{\substack{\alpha\beta\gamma \\ (\text{cyclic})}} \frac{(\omega_\alpha - \mu)^2[\bar{D}(\omega_\alpha + i0) + \bar{D}(\omega_\alpha - i0)]}{(\omega_\alpha - \omega_\beta)(\omega_\alpha - \omega_\gamma)}.$$

However, at the energies ω_R of the resonances from Eq. (3.3), we have

$$\bar{D}(\omega_R - i0)/\bar{D}(\omega_R + i0) = e^{2i\delta(\omega_R)} = -1. \tag{3.16}$$

So that $\mathfrak{I}^P$ is equal to one. It then follows that Eq. (3.14) is identically satisfied.

Coming back to our general problem, we see that the equations of definition of the bound states and of the resonances provide us with N_0 independent equations from which the $\omega_i^{(0)}$ are obtained in terms of the $2N_0 - P_0 - 1$ *CDD* poles falling below $\omega = \mu$ and of $\sum_{i=1}^{N_0} g_i^2$. We can finally use the normalization condition $f(\mu) = 1$ to express $\sum_{i=1}^{N_0} g_i^2$ in terms of the free parameters.

We have therefore a total of $N_0 + 1$ relations; from them $N_0 + 1$ of the $2N_0 - P_0$ free parameters can be expressed as a function of the remaining ones. Therefore, in general, free parameters are left in the model.

4. THE LEE MODEL AS A PARTICULAR CASE

As an example, we now consider the Lee model[4] with form factor which is a particular case of the Dyson model.

We can now have, for the behavior of the phase shift, one of the two situations sketched in Fig. 4, moreover in the case in which the phase shift starts decreasing we have one bound state of energy ω_V, while if the phase shift starts increasing we have no bound states. This situation must be assumed in order to have $N_0 = 1$, i.e., the Lee model. Being $N_0 = 1$ we do not have *CDD* poles.

For the form factor we get the expressions

$$f^2(\omega) = \frac{4\pi}{g_1^2(\omega^2 - \mu^2)^{\frac{1}{2}}} \times \exp\left[-\frac{P}{\pi}\int_\mu^\infty \frac{\delta(y)}{y-\omega}dy\right]\sin\delta(\omega) \times \begin{cases}(\omega_V - \omega) & \text{normalizable state present,} \\ (\mu - \omega) & \text{no normalizable state.}\end{cases} \tag{4.1}$$

The parameters to be determined are the constants g_1^2 and $\omega_1^{(0)}$. In both cases g_1^2 is immediately determined by the condition $f(\mu) = 1$.

In the first case the relation satisfied by the known eigenvalue ω_V of H furnishes the parameter $\omega_1^{(0)}$:

$$\omega_1^{(0)} = \omega_V + g_1^2\phi(\omega_V). \tag{4.2}$$

Note that $\phi(\omega_V)$ does not now contain any free

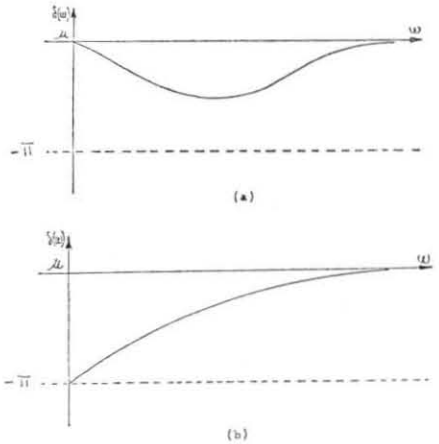

FIG. 4. The behavior of the scattering phase shift in the Lee model in the case of one bound state (a) and of no bound states (b).

[4] T. D. Lee, Phys. Rev. 95, 1329 (1954).

parameter. In this case the eventual occurrence of resonances does not give any additional information, since we know that the completely determined form factor reproduces the assigned phase shift and therefore also the resonances.

In the second case, on the contrary, the position of the resonance, which is certainly present, furnishes the value of the parameter $\omega_1^{(0)}$:

$$\omega_1^{(0)} = \omega_R + g_1^2\phi(\omega_R). \tag{4.3}$$

Again the occurrence of more than one resonance does not give any additional information.

The model is therefore in any case completely determined by the assignment of the phase shift and the energy of the eventual bound state.

5. MULTICHANNEL MODEL

We now consider the possibility that the spinless particle ϑ be capable of excited states ϑ_α, while the fixed scatterer will be supposed to remain in its ground state N. For simplicity a direct reaction mechanism is postulated for the elastic and inelastic scattering processes.

The allowed transitions are $N + \vartheta_\alpha \to N + \vartheta_\beta$; $\alpha, \beta = 1, \cdots, N_0$ and the interaction is assumed to be separable both in the channels and in momentum space, and to satisfy the requirements of Hermiticity and time reversal invariance. We have

$$H = H_0 + H_I,$$

$$H_0 = m_N\psi_N^\dagger\psi_N + \sum_{\alpha=1}^{N_0}\int \omega_\alpha a_\alpha^\dagger(\mathbf{k})a_\alpha(\mathbf{k})\,d^3k,$$

$$H_I = (2\pi)^{-3}\sum_{\alpha,\beta=1}^{N_0}\iint \frac{f_\alpha(\omega_\alpha)f_\beta(\omega_\beta')}{2(\omega_\alpha\omega_\beta')^{\frac{1}{2}}} \times \psi_N^\dagger\psi_N a_\alpha^\dagger(\mathbf{k})a_\beta(\mathbf{k}')\,d^3k\,d^3k', \tag{5.1}$$

where m_N and $\mu_\alpha = +(\omega_\alpha^2 - k^2)^{\frac{1}{2}}$ are the masses of the fixed scatterer and of the αth excited state of the ϑ particle. The operators $a_\alpha(\mathbf{k})$ and ψ_N are annihilation operators for the particle ϑ_α and N, respectively, while f_α is a form factor expressing the strength and the extension of the separable interaction. The functions f_α can be either all real or all purely imaginary. We assume that $f_\alpha(\omega)$ is different from zero and continuous in the interval $\mu_\alpha \leq \omega < +\infty$. Even though (5.1) is able to describe inelastic scattering, only S waves are involved in the collision process.

The discrete eigenvalues for $\omega < \mu_1$ are the roots of the equation

$$1 + \sum_{\alpha=1}^{N_0}\frac{1}{4\pi^2}\int_{\mu_\alpha}^{\infty}\frac{(y^2 - \mu_\alpha^2)^{\frac{1}{2}}f_\alpha^2(y)}{y - \omega}\,dy = 0. \tag{5.2}$$

Also in this case $f_\alpha^2(\omega)$ must go to zero at infinity more rapidly than ω^{-1}. A necessary condition to have a discrete eigenvalue falling on the continuous spectrum is that all the f_α^2 which refer to open channels vanish at that energy. Since we have assumed that f_α^2 never vanishes for $\omega > \mu_\alpha$, we do not have bound states embedded in the continuum in our theory.

Let us introduce the function

$$D(z) = 1 + \sum_{\alpha=1}^{N_0}\frac{1}{4\pi^2}\int_{\mu_\alpha}^{\infty}\frac{(y^2 - \mu_\alpha^2)^{\frac{1}{2}}f_\alpha^2(y)}{y - z}\,dy. \tag{5.3}$$

The zeros of $D(z)$ are all real and fall in the energy region $\omega < \mu_1$. We have a one to one correspondence between the zeros of $D(z)$ and the solutions of Eq. (5.2), i.e., the discrete eigenvalues of H. The function $D(z)$ has branch cuts extending from the lowest threshold μ_1 to $+\infty$.

The resolvent operator $G(z) = (z - H)^{-1}$ satisfies

$$\{1 + \text{Trace}\,[H_IG(z)]\}\{1 - \text{Trace}\,[H_IG_0(z)]\} = 1, \tag{5.4}$$

where $G_0(z) = (z - H_0)^{-1}$. The T matrix on the energy shell is

$$T_{\alpha\beta}(\mathbf{k}, \mathbf{k}') = f_\alpha(\omega)f_\beta(\omega)\{1 + \text{Trace}\,[H_IG(\omega + i0)]\}.$$

Using Eq. (5.4) we obtain

$$T_{\alpha\beta}(\mathbf{k}, \mathbf{k}') = \frac{f_\alpha(\omega)f_\beta(\omega)}{16\pi^3\omega}\,[D(\omega + i0)]^{-1}. \tag{5.5}$$

The relation between the T and the S matrix is

$$S_{\alpha\beta}(\omega) = \delta_{\alpha\beta} - 8i\pi^2\omega(k_\alpha k_\beta)^{\frac{1}{2}}T_{\alpha\beta}(\omega),$$

which gives

$$S_{\alpha\beta}(\omega) = \delta_{\alpha\beta} - \frac{2i\varphi_\alpha(\omega)\varphi_\beta(\omega)}{1 + \lim_{\epsilon\to 0^+}\sum_{\alpha=1}^{N_0}\frac{1}{\pi}\int_{\mu_\alpha}^{\infty}\frac{\varphi_\alpha^2(y)}{y - \omega - i\epsilon}\,dy}, \tag{5.6}$$

where

$$\varphi_\alpha(\omega) = (k_\alpha/4\pi)^{\frac{1}{2}}f_\alpha(\omega). \tag{5.7}$$

It is also convenient to introduce the complex phase of $S_{\alpha\beta}(\omega)$:

$$\delta_{\alpha\beta}(\omega) = (1/2i)\,\log S_{\alpha\beta}(\omega). \tag{5.8}$$

The denominator appearing in the second term at the right-hand side of Eq. (5.6) is $D(\omega + i0)$ written in terms of the functions φ_α. From Eq. (5.6) it can be seen that the determinant of the S matrix is equal to

$$\det S = D(\omega - i0)/D(\omega + i0) = e^{2i\Delta(\omega)}, \tag{5.9}$$

where $\Delta(\omega)$ is the phase of the function $D(\omega - i0)$:

$$D(\omega - i0) = |D(\omega - i0)|\, e^{i\Delta(\omega)}. \qquad (5.10)$$

The phase $\Delta(\omega)$ will play a key role in the inversion problem.

It is now possible to derive a Levinson-type relation for the phase $\Delta(\omega)$ which can be taken to be continuous in the interval $\mu_1 \leq \omega < +\infty$. Let us consider, as usual, the integral

$$\frac{1}{2i}\int_C \frac{[D(z)]'}{D(z)}\,dz = \frac{1}{2i}\int_{C_\infty} \frac{[D(z)]'}{D(z)}\,dz + \frac{1}{2i}\int_{\mu_1}^{\infty} d\log\frac{D(\omega + i0)}{D(\omega - i0)},$$

where the paths of integration $C = C_1 + C_2 + C_3 + C_\infty$ and C_∞ are shown in Fig. 1. Again we have excluded the possibility of a zero energy resonance. The first term at the right-hand side vanishes, the second one can be evaluated in the usual way furnishing

$$\frac{1}{2i}\int_C \frac{[D(z)]'}{D(z)}\,dz = \Delta(\mu_1) - \Delta(\infty).$$

The same integral can be evaluated by means of the Cauchy integral formula giving $n\pi$, where n is the number of zeros of $D(z)$, i.e., the number of bound states of H. We have then, assuming $\Delta(\infty) = 0$,

$$\Delta(\mu_1) = n\pi. \qquad (5.11)$$

Let us pay some attention to the bound-state equation (5.2). Introducing the function

$$\phi_\alpha(\omega) = \frac{P}{\pi}\int_{\mu_\alpha}^{\infty} \frac{\varphi_\alpha^2(y)}{y - \omega}\,dy,$$

Eq. (5.2) becomes

$$1 = -\sum_{\alpha=1}^{N_0} \phi_\alpha(\omega), \qquad \omega < \mu_1. \qquad (5.13)$$

As previously remarked, the functions φ_α^2 are all positive or all negative definite in their intervals of definition. There follows that the function $-\sum_{\alpha=1}^{N_0}\phi_\alpha(\omega)$, which is zero for $\omega = -\infty$ is monotonically decreasing in the first case and monotonically increasing in the second case. Therefore, if the f_α's are all real no bound state can be present, while if they are all purely imaginary we can have at most one bound state. It then follows that n in Eq. (5.11) is certainly zero in the first case and may be zero or one in the second case.

6. THE INVERSION PROBLEM IN THE MULTICHANNEL MODEL

The function $D(z)$ defined in Sec. 5 is an analytic function in the complex z plane cut from μ_1 to $+\infty$, it has a real simple zero for $\omega < \mu_1$ at the position of the bound state, if any, and no poles. At infinity it tends to one.

We now take into account the function

$$\bar{D}(z) = \exp\left[-\frac{1}{\pi}\int_{\mu_1}^{\infty} \frac{\Delta(y)}{y - z}\,dy\right]. \qquad (6.1)$$

Owing to the fact that $\Delta(\omega)$ is a continuous function of ω which tends to zero as $\omega \to +\infty$, $\bar{D}(z)$ is an analytic function of z in the z plane cut from μ_1 to $+\infty$ which goes to zero as z tends to ∞ and never vanishes for complex z. Its behavior for $z \to \mu$ follows from Eq. (5.11):

$$\bar{D}(z) \underset{z\to\mu}{\sim} \text{const}\,(\mu - z)^n.$$

Moreover, $\bar{D}(z)$ satisfies

$$\bar{D}(\omega - i0)/\bar{D}(\omega + i0) = e^{2i\Delta(\omega)}. \qquad (6.2)$$

Comparing this equation with Eq. (5.9) we can conclude that the function

$$R(z) = D(z)/\bar{D}(z) \qquad (6.3)$$

does not have the branch cut, has only isolated simple zeros at the points of the discrete eigenvalues of H, and an isolated pole of order n at $z = \mu$. From the behavior of $\bar{D}(z)$ and $D(z)$ for $z \to \infty$ we obtain that $R(z) \to 1$ for $z \to \infty$. It is therefore the following rational function:

$$R(z) = \begin{cases} (z - \omega_1)/(z - \mu_1) & \text{bound state present,} \\ 1 & \text{no bound state.} \end{cases} \qquad (6.4)$$

We finally get for the function $D(z)$ the following result:

$$D(z) = R(z)\exp\left[-\frac{1}{\pi}\int_{\mu_1}^{\infty} \frac{\Delta(y)}{y - z}\,dy\right]. \qquad (6.5)$$

We want now to see how to relate the parameters of the model with the scattering and reaction complex phase shifts and the energies of the bound state.

Let us consider the discontinuity of the function $D(z)$ across its branch cut; in terms of $\Delta(\omega)$ we get

$$\sum_{\alpha=1}^{N_\gamma} \varphi_\alpha^2(\omega) = -|D(\omega + i0)|\sin\Delta(\omega), \quad \omega > \mu_\gamma, \qquad (6.6)$$

the sum extending to the channels which are open at the energy ω. On the other hand, from the relation

defining the S matrix element for the reaction $N + \vartheta_\alpha \to N + \vartheta_\beta$, $\alpha \neq \beta$, we get

$$2\varphi_\alpha(\omega)\varphi_\beta(\omega) = |D(\omega + i0)|\, e^{-2\delta_{\alpha\beta}{}^I} \times \exp\,[i(2\delta^R_{\alpha\beta}(\omega) - \Delta(\omega) + \tfrac{1}{2}\pi)], \quad \alpha \neq \beta. \tag{6.7}$$

We have introduced the real and the imaginary parts $\delta^R_{\alpha\beta}(\omega)$ and $\delta^I_{\alpha\beta}(\omega)$ of the complex phase shift $\delta_{\alpha\beta}$.

Note that, being the φ_α's either all real or all purely imaginary, we must have

$$\Delta(\omega) = 2\delta^R_{\alpha\beta}(\omega) + (m_{\alpha\beta} + \tfrac{1}{2})\pi \tag{6.8}$$

with $m_{\alpha\beta}$ integer. The behavior at infinity of $\delta^R_{\alpha\beta}(\omega)$ is

$$\delta^R_{\alpha\beta}(\omega) \underset{\omega\to\infty}{\sim} \tfrac{1}{4}(2k + 1)\pi, \quad \text{integer } k, \tag{6.9}$$

so that the condition that $\Delta(\infty) = 0$ can be satisfied and determines $m_{\alpha\beta}$ once $\delta^R_{\alpha\beta}(\infty)$ is given. Note, moreover, that Eq. (6.8) gives a simple relation between all the real parts of the reaction phase shifts, namely,

$$\delta^R_{\alpha\beta}(\omega) = \delta^R_{\gamma\epsilon}(\omega) + \tfrac{1}{2}(m_{\gamma\epsilon} - m_{\alpha\beta})\pi. \tag{6.10}$$

Equation (6.8), together with the continuity requirement for $\Delta(\omega)$ and the condition $\Delta(\infty) = 0$, enables us to determine uniquely $\Delta(\omega)$ in terms of $\delta^R_{12}(\omega)$ for $\omega > \mu_2$. On the other hand, below the threshold μ_2, $\Delta(\omega)$ is equal to $\delta_{11}(\omega)$, a part from multiples of π.

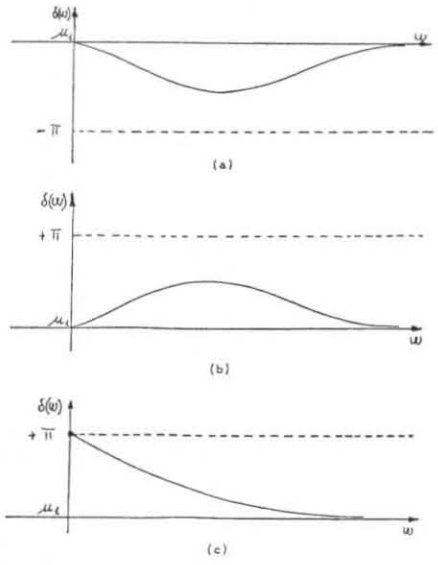

FIG. 5. The behavior of the phase of the determinant of the S matrix of the multichannel model considered in Secs. 5 and 6. The three curves correspond, respectively, to real form factors $f_\alpha(\omega)$ (a), to purely imaginary form factors when no bound state is present (b), and to purely imaginary form factors when there is one bound state (c).

From Eq. (6.6) it follows that $\sin\Delta(\omega)$ is always different from zero for $\omega > \mu_1$, is negative when all the functions f_α are real, and positive if all the functions f_α are purely imaginary. The function $\Delta(\omega)$ therefore certainly has one of the three behaviors (a), (b), and (c) shown in Fig. 5. Case (a) corresponds to real f_α's in which case no bound state is present, case (b) to purely imaginary f_α's without bound state, and case (c) to purely imaginary f_α's with one bound state. Note that the knowledge of $\delta_{11}(\omega)$ for $\mu_1 \leq \omega \leq \mu_2$ determines in which of the two cases (a) and (b) we are, and therefore also if real or purely imaginary f_α's have to be taken. If there is one bound state, then we certainly are in case (c). Owing to Eq. (6.8), Eq. (6.7) can be written

$$2\varphi_\alpha(\omega)\varphi_\beta(\omega) = (-)^{m_{\alpha\beta}}\, |D(\omega + i0)|\, e^{-2\delta_{\alpha\beta}{}^I}, \quad \alpha \neq \beta. \tag{6.11}$$

We can now determine the form factors in the intervals $\mu_1 \leq \omega < \mu_2$, $\mu_2 \leq \omega < \mu_3$, and so on, by making use of Eqs. (6.8) and (6.11), provided we know the imaginary parts of the reaction phase-shifts. We get first of all[5]

$$\varphi_1(\omega) = [-\,|D(\omega + i0)|\sin\Delta(\omega)]^{\frac{1}{2}}, \quad \mu_1 \leq \omega < \mu_2. \tag{6.12}$$

For the interval $\mu_2 \leq \omega < \mu_3$ we proceed as follows. From Eqs. (6.8) and (6.11) we get

$$[\varphi_1(\omega) \pm \varphi_2(\omega)]^2 = |D(\omega + i0)|\,[-\sin\Delta(\omega) \pm (-)^{m_{12}} e^{-2\delta_{12}{}^I(\omega)}]. \tag{6.13}$$

From these two equations it follows that

$$\varphi_1(\omega) = \tfrac{1}{2}\,|D(\omega + i0)|^{\frac{1}{2}} \times \{[-\sin\Delta(\omega) + (-)^{m_{12}} e^{-2\delta_{12}{}^I(\omega)}]^{\frac{1}{2}} + [-\sin\Delta(\omega) - (-)^{m_{12}} e^{-2\delta_{12}{}^I(\omega)}]^{\frac{1}{2}}\}, \quad \mu_2 \leq \omega < \mu_3 \tag{6.14}$$

$$\varphi_2(\omega) = \tfrac{1}{2}\,|D(\omega + i0)|^{\frac{1}{2}} \times \{[-\sin\Delta(\omega) + (-)^{m_{12}} e^{-2\delta_{12}{}^I(\omega)}]^{\frac{1}{2}} - [-\sin\Delta(\omega) - (-)^{m_{12}} e^{-2\delta_{12}{}^I(\omega)}]^{\frac{1}{2}}\},$$

having used continuity across $\omega = \mu_2$ for φ_1 and the fact that $\varphi_2(\mu_2) = 0$. Proceeding in this way we can determine the form factors step by step.

From Eq. (6.13) we see that $\delta^I_{12}(\omega)$ must satisfy

[5] The inversion problem has been considered for the one-channel separable case by M. Gourdin and A. Martin, Nuovo Cimento **8**, 699 (1958). Our procedure applied to this particular case yields the same result in an extremely simplified way.

the condition

$$e^{-2\delta_{12}{}^{I}(\omega)} < |\sin \Delta(\omega)|; \qquad \mu_2 \leq \omega < \mu_3. \tag{6.15}$$

Slightly more complicated conditions are obtained among the $\delta^I_{\alpha\beta}(\omega)$ for more than two open channels.

Summarizing, the assignment of the phase shift $\delta_{11}(\omega)$ in the interval $\mu_1 \leq \omega < \mu_2$, of $\delta^R_{12}(\omega)$ in the interval $\omega \geq \mu_2$, of all the $\delta^I_{\alpha\beta}(\omega)$'s in the interval $\omega \geq \mu_\beta$ (β being greater than α), and of the energy of the eventual bound state, completely determines the model. If we arbitrarily assign the complex phase shifts (in agreement with the unitarity of the S matrix) for a multichannel process, we can describe the situation by means of our model if and only if they satisfy all the above-mentioned conditions.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 7, NUMBER 5 MAY 1966

Gauge Invariant Formulation of the BCS Model*

GÉRARD EMCH† AND MARCEL GUENIN‡
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey
(Received 12 April 1965)

A generalized Valatin–Bogoliubov transformation is exhibited, which preserves the gauge invariance of the theory. Some consequences of this treatment of the BCS-model are analyzed in the framework of nonrelativistic axiomatic field theory.

1. INTRODUCTION

THE interest of the BCS model[1] in connection with the (nonrelativistic) axiomatic approach of field theory has been recognized by Haag.[2] He introduced some of the algebraic methods, which are, by now in current use, to show how and why this model is exactly soluble in the limit of infinite volume. Later on his arguments have still been improved by Ezawa.[3] We therefore do not intend to focus on this point. Rather, we try to present a detailed discussion of those aspects of the model which are related to its symmetry with respect to the gauge group. In particular, we want to point out that this model indeed exhibits peculiar features whose nature has to be taken into account in the general axioms of quantum field theory.[4]

There is *one* paradox in the BCS model for superconductivity and this paradox shows itself in full light in Haag's elegant treatment of the problem. The derivation starts from a gauge-invariant theory and ends up with an Hamiltonian which is no more invariant under the original particle gauge group; this Hamilitonian, however, is invariant under a new symmetry group (the quasi-particle gauge group) which is itself *not* a symmetry group for the original theory. Haag already pointed out that this strange feature is linked with the seemingly innocent assumption that the algebra generated by the field operators is irreducible. This can hardly be considered as a completely satisfactory explanation, especially if one notes that this paradox is not proper to Haag's derivation and already appears in the most early and traditional account of the modern theory of superconductivity.[5] Indeed, the articulation of all derivations is the Valatin–Bogoliubov transformation from particles to quasi-particles and we show that it is there that the apparent "symmetry breaking" has to be traced. As an illustration of this point, one remarks that, in the orthodox treatment, the quasi-particle vacuum[5] is not left invariant by any of the particle gauge transformation, nor is it even multiplied by a phase factor. This last fact has however not to be taken too seriously since one can interpret it in terms of relative phase change between superselection spaces. We therefore concentrate mainly on the lack of invariance of the Hamiltonian itself, a fact which is really more disturbing.

* Research supported in part by U. S. Air Force Office Research, ARDC.
† Partial support of this work by the Swiss NSF is acknowledged. Present address: Institute for Applied Mathematics, University of Maryland.
‡ This author is indebted to the Fonds Birkigt University of Geneva, for a travel grant.

[1] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. 108, 1175 (1957).
[2] R. Haag, Nuovo Cimento 25, 287 (1962).
[3] H. Ezawa, J. Math. Phys. 5, 1078 (1964).
[4] M. Guenin, *Axiomatic Foundations of Quantum Theories* (to be published).
[5] J. R. Schrieffer, *Theory of Superconductivity* (W. A. Benjamin Company, Inc., New York, 1964).

the condition

$$e^{-2\delta_{12}{}^{I}(\omega)} < |\sin \Delta(\omega)|; \qquad \mu_2 \leq \omega < \mu_3. \tag{6.15}$$

Slightly more complicated conditions are obtained among the $\delta^I_{\alpha\beta}(\omega)$ for more than two open channels.

Summarizing, the assignment of the phase shift $\delta_{11}(\omega)$ in the interval $\mu_1 \leq \omega < \mu_2$, of $\delta^R_{12}(\omega)$ in the interval $\omega \geq \mu_2$, of all the $\delta^I_{\alpha\beta}(\omega)$'s in the interval $\omega \geq \mu_\beta$ (β being greater than α), and of the energy of the eventual bound state, completely determines the model. If we arbitrarily assign the complex phase shifts (in agreement with the unitarity of the S matrix) for a multichannel process, we can describe the situation by means of our model if and only if they satisfy all the above-mentioned conditions.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 7, NUMBER 5 MAY 1966

Gauge Invariant Formulation of the BCS Model*

GÉRARD EMCH† AND MARCEL GUENIN‡
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey
(Received 12 April 1965)

A generalized Valatin–Bogoliubov transformation is exhibited, which preserves the gauge invariance of the theory. Some consequences of this treatment of the BCS-model are analyzed in the framework of nonrelativistic axiomatic field theory.

1. INTRODUCTION

THE interest of the BCS model[1] in connection with the (nonrelativistic) axiomatic approach of field theory has been recognized by Haag.[2] He introduced some of the algebraic methods, which are, by now in current use, to show how and why this model is exactly soluble in the limit of infinite volume. Later on his arguments have still been improved by Ezawa.[3] We therefore do not intend to focus on this point. Rather, we try to present a detailed discussion of those aspects of the model which are related to its symmetry with respect to the gauge group. In particular, we want to point out that this model indeed exhibits peculiar features whose nature has to be taken into account in the general axioms of quantum field theory.[4]

There is *one* paradox in the BCS model for superconductivity and this paradox shows itself in full light in Haag's elegant treatment of the problem. The derivation starts from a gauge-invariant theory and ends up with an Hamiltonian which is no more invariant under the original particle gauge group; this Hamilitonian, however, is invariant under a new symmetry group (the quasi-particle gauge group) which is itself *not* a symmetry group for the original theory. Haag already pointed out that this strange feature is linked with the seemingly innocent assumption that the algebra generated by the field operators is irreducible. This can hardly be considered as a completely satisfactory explanation, especially if one notes that this paradox is not proper to Haag's derivation and already appears in the most early and traditional account of the modern theory of superconductivity.[5] Indeed, the articulation of all derivations is the Valatin–Bogoliubov transformation from particles to quasi-particles and we show that it is there that the apparent "symmetry breaking" has to be traced. As an illustration of this point, one remarks that, in the orthodox treatment, the quasi-particle vacuum[5] is not left invariant by any of the particle gauge transformation, nor is it even multiplied by a phase factor. This last fact has however not to be taken too seriously since one can interpret it in terms of relative phase change between superselection spaces. We therefore concentrate mainly on the lack of invariance of the Hamiltonian itself, a fact which is really more disturbing.

* Research supported in part by U. S. Air Force Office Research, ARDC.
† Partial support of this work by the Swiss NSF is acknowledged. Present address: Institute for Applied Mathematics, University of Maryland.
‡ This author is indebted to the Fonds Birkigt University of Geneva, for a travel grant.

[1] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. 108, 1175 (1957).
[2] R. Haag, Nuovo Cimento 25, 287 (1962).
[3] H. Ezawa, J. Math. Phys. 5, 1078 (1964).
[4] M. Guenin, *Axiomatic Foundations of Quantum Theories* (to be published).
[5] J. R. Schrieffer, *Theory of Superconductivity* (W. A. Benjamin Company, Inc., New York, 1964).

Because of widely spread prejudices and in order to keep the reader from misunderstanding what follows, we found it necessary to spell out already in this Introduction the restriction we do *not* impose to our scheme:

(i) We do not assume that the algebra generated by the fields is irreducible.

(ii) We do not assume that a symmetry leaves the center of this algebra element-wise invariant.

(iii) We do not identify "elementary description of a physical system" with "irreducible representation of its associated algebras." (This idea is already present in the familiar case of a Lorentz invariant theory, where one does not restrict the reflections to map each coherent subspace onto itself[6]; the new feature is that, instead of a discrete symmetry group like the reflections group, we are dealing here with the gauge group, which is *continuous*.)

It turns out that each of the above extended postulates is necessary for a consistent treatment of the gauge-invariant BCS model.

2. THE MODEL

We define in the usual way the nonrelativistic, free Fermi fields Ψ_i (hereafter referred to as "electron fields" or "particle fields") as linear mappings defined on the Hilbert space $\mathfrak{L}^2(\mathbf{R}^3)$ of the square-integrable functions on $\mathbf{R}^3$, and taking their values in an abstract B^* algebra $\mathfrak{B}$ (see the next-to-last remark of Sec. 4):

$$\Psi_i : \mathfrak{L}^2(\mathbf{R}^3) \to \mathfrak{B}. \tag{2.1}$$

The $\Psi_i(f)$ are referred to as "smeared-out" fields.

The usual Fermi anticommutation rules are assumed:

$$\begin{aligned} \{\Psi_i(f), \Psi_j(g)\} &= 0, \\ \{\Psi_i(f), \Psi_j^*(g)\} &= \delta_{ij}(f, g)I. \end{aligned} \tag{2.2}$$

The relevant object for the theory is the B^* subalgebra $\mathfrak{A}$ of $\mathfrak{B}$ generated by the smeared out fields (and their adjoints, i.e., the closure in the norm of $\mathfrak{B}$ of the set algebraically generated by them).

It is appropriate for this model to imagine the system as enclosed in a finite volume V with periodic boundary conditions, and then to consider the infinite volume as a limiting case (in a sense to be made more precise). For the sake of convenience let us consider V as a cube of edge $2l$, centered at the origin of $\mathbf{R}^3$. The corresponding smeared out fields are defined on the subspace $\mathfrak{L}^2(V)$ of $\mathfrak{L}^2(\mathbf{R}^3)$ consisting of the square-integrable functions with support in V.

[6] G. Emch and C. Piron, J. Math. Phys. 4, 469 (1963).

These particular smeared out fields generate (in the same sense as above) a B^* subalgebra $\mathfrak{A}(V)$ of $\mathfrak{A}$.

We finally define the elements $a_i(p)$ of $\mathfrak{A}(V)$

$$a_i(p) = \Psi_i(\hat{f}_p), \tag{2.3}$$

where

$$\hat{f}_p(x) = \begin{cases} (1/V^{\frac{1}{2}})\, e^{ipx}, & x \in V, \\ 0, & \text{elsewhere}, \end{cases} \tag{2.4}$$

with

$$p = n\pi/l \quad \text{with} \quad n \in E^3 \tag{2.5}$$

(E is the group of all integers).

The following elements of $\mathfrak{A}(V)$ play an important role in the model:

$$b(p) = a_2(-p)a_1(p). \tag{2.6}$$

The above definitions allow a rigorous treatment which can parallel the more traditional accounts of the model.[1,5]

The original idea of the BCS model was to describe a superconductor, enclosed in a finite box with periodic boundary conditions, by a "reduced" Hamiltonian of the form

$$\hat{H} = \hat{H}_0 + \hat{H}_1, \tag{2.7}$$

where

$$\begin{aligned} \hat{H}_0 &= \sum_p \hat{H}_0(p) \\ &= \sum_{p,i} \epsilon(p) a_i^*(p) a_i(p) \qquad (i = 1, 2), \end{aligned} \tag{2.8}$$

with

$$\epsilon(p) = -\mu + (p^2/2m), \tag{2.9}$$

and

$$\begin{aligned} \hat{H}_1 &= \sum_p \hat{H}_1(p) \\ &= \sum_p b^*(p) \sum_q \hat{v}(p, q) b(q), \end{aligned} \tag{2.10}$$

with

$$\hat{v}(p, q) = \int_V d\xi\, d\eta\, \hat{f}_p(\xi) v(\xi, \eta) \hat{f}_q^*(\eta). \tag{2.11}$$

Let us finally introduce

$$\hat{\Delta}(p) = \sum_q \hat{v}(p, q) b(q). \tag{2.12}$$

These objects (which are closely related to the "energy gap") also play a central role in the arguments presented in this paper.

Within the usual Hilbert space formalism, the free Hamiltonian is an unbounded operator, and there-

fore, it cannot be an element of $\mathfrak{A}(V)$. We take care of this fact later on. The same remark could also apply to the interaction Hamiltonian. However, it turns out that this is not at all a drastic obstacle to our approach, provided that we make more specific assumptions on the interaction; the following properties are assumed on the nonlocal potential $v(\xi, \eta)$:

$$v \equiv \int d\xi\, d\eta\, |v(\xi, \eta)| < \infty, \tag{2.13}$$

$$v(\xi, \eta)^* = v(\eta, \xi), \tag{2.14}$$

$$\sum_q |\hat{v}(p, q)| < \infty. \tag{2.15}$$

The condition (2.13) is in the heart of all our subsequent arguments; indeed it implies that

$$\lim_{V\to\infty} |\hat{v}(p, q)| = 0 \tag{2.16}$$

since [from (2.4), (2.11), and (2.13)] we have

$$|\hat{v}(p, q)| \leq v/V. \tag{2.17}$$

Our condition (2.15) ensures that the $\hat{\Delta}(p)$ belong to $\mathfrak{A}(V)$. As a consequence, the $\hat{H}_1(p)$ too, belong to $\mathfrak{A}(V)$, as well as the $\hat{H}_0(p)$; this is almost all that we really need in the following.

From now on we represent our algebra $\mathfrak{A}$ as a ring of operators acting on a Hilbert space $\mathfrak{H}$. We also denote this ring by $\mathfrak{A}$ since we always work in the same representation. We require later that this representation satisfies a certain continuity property. For the time being we impose that there exists for each p an element $\Delta(p)$ of $\mathfrak{A}''$ (the weak closure of $\mathfrak{A}$ in $\mathfrak{H}$) such that

$$\lim_{V\to\infty} (\psi_1, [\hat{\Delta}(p) - \Delta(p)]\psi_2) = 0 \text{ for all } \psi_1, \psi_2 \in \mathfrak{H}. \tag{2.18}$$

As usual the limit $V \to \infty$ is taken over a sequence of volumes $V_i(p)$ compatible with the given p. This sequence of course changes with p but this is of no importance since we do not require any uniformity in p.

The reader may remark that all the above assumptions are already present (although not always explicitly stated) in the previous papers on this model and that they are satisfied in some cases of practical importance.

3. THE DIAGONALIZATION

Let us begin this section by a preliminary remark. From the definition of $\hat{\Delta}(p)$ we have that

$$[\hat{\Delta}(p), a_i(p')] = 0, \quad \text{for all } p \text{ and } p'. \tag{3.1}$$

As a consequence of (2.16),

$$\lim_{V\to\infty} (\psi_1, [\Delta(p), a_i^*(p')]\psi_2) = 0 \quad \text{for all } \psi_1, \psi_2 \in \mathfrak{H}, \tag{3.2}$$

and with a barely more sophisticated evaluation, one obtains the general result

$$\lim_{V\to\infty} (\psi_1, [\Delta(p), \Psi_i^{(*)}(f)]\psi_2) = 0$$

$$\text{for all } f \in \mathfrak{L}^2(\mathbf{R}^3) \quad \text{and all } \psi_1, \psi_2 \in \mathfrak{H}. \tag{3.3}$$

Together with our assumption (2.18), this leads to the conclusion that $\Delta(p)$ belongs to the *center* $\mathfrak{Z}(\mathfrak{A}) = \mathfrak{A}' \cap \mathfrak{A}''$ of $\mathfrak{A}''$

$$\Delta(p) \in \mathfrak{Z}(\mathfrak{A}), \quad \forall p. \tag{3.4}$$

Since we do not want to restrict our attention to the case where the algebra generated by the Fermi fields has only a trivial center, we cannot conclude as Haag did, that $\Delta(p)$ is a c number. Our main point in this section is to emphasize that *even without this restriction, one can justify the usual affirmation that "the BCS model becomes exactly soluble in the limit of infinite volume."*

Haag's recipe can now be followed *cum grano salis*: we define a new Hamiltonian by the substitution:

$$\hat{H}_1(p) \to \hat{H}_1'(p) = b^*(p)\, \Delta(p) + \text{h.c.} \tag{3.5}$$

[see (2.7), (2.10), (2.12)]. It is not the purpose of this note to discuss in which sense this new Hamiltonian is a good approximation of the original Hamiltonian. [The fact that the commutators of the $a_i^{(*)}(p)$ with the two Hamiltonians are identical, up to a substitution $\hat{\Delta} \to \Delta$, can be regarded only as an indication that the two Hamiltonians could differ at most by an element of the center of $\mathfrak{A}''$]. We simply want to mention here that

$$\hat{H}'(p) = \hat{H}_0(p) + \hat{H}_1'(p) + (\textit{idem} \text{ with } p \to -p) \tag{3.6}$$

can be *exactly* "diagonalized" and rewritten as

$$\hat{H}'(p) = E(p)[\gamma_1^*(p)\gamma_1(p) + \gamma_2^*(p)\gamma_2(p)] + (\textit{idem} \text{ with } p \to -p) \tag{3.7}$$

with

$$\begin{aligned} \gamma_1(p) &= u(p)a_1(p) + v(p)a_2^*(-p), \\ \gamma_2(p) &= -v(-p)a_1^*(-p) + u(-p)a_2(p), \end{aligned} \tag{3.8}$$

where

$$E(p) = \{\epsilon^2(p) + [\Delta(p)\, \Delta^*(p)]\}^{\frac{1}{2}}, \tag{3.9}$$

$$u(p) = \Delta^*(p)/D(p), \tag{3.10}$$

$$v(p) = [E(p) - \epsilon(p)]/D(p), \tag{3.11}$$

$$D(p) = \{[E(p) - \epsilon(p)]^2 + \Delta(p)\,\Delta^*(p)\}^{\frac{1}{2}}. \tag{3.12}$$

We recall that *here* $\Delta(p)$, and consequently the quantities defined by (3.9)–(3.12), are not necessarily c numbers, but are in general elements of the center $\mathfrak{Z}(\mathfrak{A})$. For this reason, the relations (3.8) have to be regarded as a slight, but important *generalization of the Bogoliubov–Valatin transformations*. The $\gamma_i^{(*)}(p)$ have the usual Fermi-commutation rules and are referred hereafter as the "quasi-particle" smeared-out fields [by opposition to the $a_i^{*}(p)$ referred to as the "particles" smeared-out free fields].

The interest of the substitution (3.5) is

(i) It determines (almost) uniquely $E(p)$, $u(p)$, and $v(p)$.

(ii) By a straightforward application of Valatin's method (for an account of this method, see for instance Ref. 5), one sees that the quasi-particles diagonalize effectively $\hat{H}$ (as well as $\hat{H}'$) when the volume approaches infinity.

4. GAUGE INVARIANCE

From general principles,[6,4] one can identify a symmetry of a physical system as a *-automorphism of the algebra associated with the *observables* of the physical system under consideration. We moreover *assume* here that (when one gives a field-theoretic description of the system) the same is true for the algebra associated with the *fields*. By this assumption we precisely mean that, if β is a symmetry of the theory, it has to satisfy at least the following axioms:

β is a bijective mapping of $\mathfrak{A}$ onto itself which preserves the *-algebraic structure of $\mathfrak{A}$, i.e., for any $A, B, \cdots$ in $\mathfrak{A}$ and any $\lambda, \cdots$ in $\mathbf{C}$, one must have:

$$\begin{aligned}
&\text{(i)} && \beta[A + B] = \beta[A] + \beta[B],\\
&\text{(ii)} && \beta[AB] = \beta[A]\beta[B],\\
&\text{(iii)} && \beta[\lambda A] = \lambda\beta[A],\\
&\text{(iv)} && \beta[A^*] = \beta[A]^*,
\end{aligned} \tag{4.1}$$

(v) β is continuous in the (|norm|-) topology of $\mathfrak{A}$.

The condition (v) is in fact implied by (i)–(iv) since $\mathfrak{A}$ is a B^* ring (see Ref. 7, Th. 2.5.16). However, whereas this continuity condition is sufficient when one is dealing with the abstract properties of the ring considered, it is often convenient to require stronger continuity properties for symmetries. These stronger continuity requirements are in general representation-dependent, and to formulate them properly one has to specify the representation. We then suppose that:

(vi) The representation $\mathfrak{A}$ in $\mathfrak{H}$ is such that the automorphism β initially defined on $\mathfrak{A}$ can be extented to the weak closure $\mathfrak{A}''$ of $\mathfrak{A}$.

One can prove quite generally that this restriction does not in fact diminish the generality of the theory. We should, however, emphasize that the property (vi) is essential in the discussion of the transformation properties of the Hamiltonian, since the spectral projectors of H belong to $\mathfrak{A}''$ (and not to $\mathfrak{A}$ in general). The continuity of β in both the strong and the weak operator topologies relative to the considered representation can then be proved from (vi).[8] From (i)–(vi) it also follows that β leaves invariant the center $\mathfrak{Z}(\mathfrak{A})$ of $\mathfrak{A}''$. Nothing however indicates that $\mathfrak{Z}(\mathfrak{A})$ is left *element-wise* invariant by β. As we see later, the present model exhibits an example where this strong supplementary property is *not* satisfied.

We hereafter consider the symmetry group (referred to as the gauge group) of all the *-automorphisms β of $\mathfrak{A}''$ (referred to as group transformations) defined by

$$\begin{aligned}
&\beta[\Psi_i(f)] = e^{i\beta}\Psi_i(f), \qquad \forall f \in \mathfrak{L}^2(\mathbf{R}^3),\\
&\text{with} \qquad \beta \in [0, 2\pi), \qquad i = 1, 2.
\end{aligned} \tag{4.2}$$

From their very definitions, it follows that the effect of β on $\tilde{\Delta}(p)$ is

$$\beta[\tilde{\Delta}(p)] = e^{2i\beta}\,\tilde{\Delta}(p), \tag{4.3}$$

and therefore

$$\beta[\Delta(p)] = e^{2i\beta}\,\Delta(p). \tag{4.4}$$

Since we proved that $\Delta(p)$ is one of the elements of the center of $\mathfrak{A}''$, we have exhibited here explicitly a symmetry which *does not* leave the center of $\mathfrak{A}''$ *element-wise* invariant.

Before going further, and discussing some more involved consequences of the above property, we would like to point out some of its direct implications which contrast the usual treatment of the BCS model. First of all we have the following transformation laws [see Def. (3.8) to (3.12)]:

$$\begin{aligned}
\beta[E(p)] &= E(p),\\
\beta[D(p)] &= D(p),\\
\beta[u(p)] &= e^{-2i\beta}u(p),\\
\beta[v(p)] &= v(p),
\end{aligned} \tag{4.5}$$

which follow from (4.4), and therefore

$$\beta[\gamma_i(p)] = e^{-i\beta}\gamma_i(p). \tag{4.6}$$

[7] L. Rickart, *Banach Algebras* (Princeton University Press, Princeton, New Jersey, 1960).

[8] J. Dixmier, *Les Algèbres d'opérateurs dans l' espace Hilbertien* (Gauthier-Villars, Paris, 1957), Cor. 1, p. 253.

For the sake of completeness, we could mention here that, because of the ambiguity remaining in the definition of the u and v, these transformation laws are not quite unique; for instance, we could slightly modify the definition of the u and v in such a way that the γ have exactly the same transformation laws as the a. This is, however, of no importance for the present analysis. In any case, the transformation laws (4.5)–(4.6) are qualitatively quite different from those usually obtained when one imposes that the center of $\mathfrak{A}''$ is trivial. In one sentence one can summarize the situation and say that the *generalized* Valatin–Bogoliubov transformations introduced in the previous section preserves the gauge invariance of the theory, whereas this is *not* the case with the usual Valatin–Bogoliubov transformation. More precisely, one remarks that, for any volume V and any compatible p [see (2.5)], we have not only

$$\beta[\hat{H}(p)] = \hat{H}(p) \tag{4.7}$$

but also [see (3.6) or (3.7)]

$$\beta[\hat{H}'(p)] = \hat{H}'(p). \tag{4.8}$$

This last invariance is *not* present within the usual formalism (see, for instance, Ref. 2).

We incidentally recall that the invariance of the total Hamiltonian under a symmetry transformation is not at all related to the very definition of a symmetry, but rather to its dynamical interpretation. We can thus assert that within the present formalism, the diagonalization (and the related limiting procedure) not only does not break the symmetry, but also does not change its dynamical interpretation.

We now want to analyze the deeper (although quite elementary) consequences of the fact that the algebra $\mathfrak{A}$ possesses a symmetry group G which does not leave the center element-wise invariant. We hereafter suppose for simplicity that $\mathfrak{B}(\mathfrak{A}) \subset \mathfrak{A}$.

We define a representation φ of of the B^* algebra $\mathfrak{A}$ as a continuous *-homomorphism of $\mathfrak{A}$ into the C^* algebra $\mathfrak{B}(\mathfrak{H})$ of all bounded operators of a Hilbert space $\mathfrak{H}$ (the continuity of φ is already implied by the fact that it is a *homomorphism, see Ref. 7, Th. 4.1.20):

$$\varphi : \mathfrak{A} \to \varphi(\mathfrak{A}) \subseteq \mathfrak{B}(\mathfrak{H}). \tag{4.9}$$

We say that this representation is *compatible*[4] with the symmetry group G of $\mathfrak{A}$ if it is possible to find for every $\beta \in G$ a *-automorphism β' of $\varphi[\mathfrak{A}]$ such that

$$\varphi \circ \beta = \beta' \circ \varphi, \tag{4.10}$$

i.e. that there exists a commutative diagram

$$\begin{array}{ccc} \mathfrak{A} & \xrightarrow{\beta} & \mathfrak{A} \\ \varphi\downarrow & & \downarrow\varphi \\ \varphi[\mathfrak{A}] & \xrightarrow{\beta'} & \varphi[\mathfrak{A}]. \end{array}$$

If this is the case, we say that (φ, β') is a representation of the pair $(\mathfrak{A}, G)$.

The first consequence of the fact that G does not leave the center of $\mathfrak{A}$ element-wise invariant [see in particular (4.4)] is that every primary representation φ of $\mathfrak{A}$, compatible with the gauge group, leads to a trivial energy gap. (By a primary representation we mean that

$$\varphi[\mathfrak{Z}(\mathfrak{A})] = \{\lambda I\}, \tag{4.11}$$

i.e., that $\varphi[\mathfrak{A}]$ is a factor; in particular, we note that every *irreducible* representation is also primary.) Let us prove our assertion.

Because of the fact that β' is an algebraic automorphism we have

$$\beta' \circ \varphi[z] = \varphi[z], \qquad \forall \beta \in G, \qquad \forall z \in \mathfrak{Z}.$$

Now using (4.10):

$$\varphi[\beta[z]] = \varphi[z], \qquad \forall \beta \in G, \qquad \forall z \in \mathfrak{Z},$$

in particular.

$$\varphi[\beta[\Delta(p)]] = \varphi[\Delta(p)]$$

and since φ is an algebriac homomorphism from (4.4) one gets

$$e^{2i\beta}\varphi[\Delta(p)] = \varphi[\Delta(p)], \qquad \forall \beta \in S^1, \qquad \forall p \in \mathbf{R},$$

which is only possible if

$$\varphi[\Delta(p)] = 0, \qquad \forall p \in \mathbf{R},$$

i.e., since φ is a *-homomorphism:

$$\varphi[\Delta^*(p)\ \Delta(p)] = 0, \qquad \forall p \in \mathbf{R}, \tag{4.12}$$

which proves our assertion.

The fact that $\Delta(p)$ is *not* left invariant under the gauge group and *belongs to the center of* $\mathfrak{A}$ is therefore the very reason why Haag loosed the gauge invariance of the theory when he restricted himself to consider $\mathfrak{A}$ as an irreducible algebra.

In $\mathfrak{B}(\mathfrak{H})$ we can now define a family of primary representations φ_β of $\mathfrak{A}$:

$$\{\varphi_\beta \mid \varphi_\beta \equiv \varphi \circ \beta, \qquad \forall \beta \in G\}. \tag{4.13}$$

That these representations are all inequivalent among themselves (in conformity with Haag's result) is again a consequence of the fact that G does not leave the center element-wise invariant. Because

of the group property in G, it is sufficient to prove that each of the φ_β is inequivalent to φ; to prove this, let us suppose on the contrary that φ_β is equivalent to φ, i.e., that there exists a unitary mapping of $\mathfrak{H}$ onto itself such that

$$U\varphi[A]U^{-1} = \varphi_\beta[A], \qquad \forall A \in \mathfrak{A}. \tag{4.14}$$

Using then the definition (4.13) and the conditions (4.11), one obtains that (4.14) implies again (4.12). The next thing to do is to decide which substitute to irreducibility of the field algebra one has to consider in order to have a "minimal" and "complete" description of the systems. We propose to call *elementary*, a representation $\varphi[\mathfrak{A}, G]$ which satisfies the condition

$$\varphi[\mathfrak{Z}_G] = \{\lambda I\}, \tag{4.15}$$

where

$$\mathfrak{Z}_G = \{z \in \mathfrak{Z}(\mathfrak{A}) \mid \beta(z) = z, \qquad \forall \beta \in G\}. \tag{4.16}$$

When the center is left element-wise invariant under G, this condition reduces to the assumption that φ is a primary (rather than an irreducible) representation. One reason for introducing this definition is that, if one wishes to transfer it on the abstract algebra $\mathfrak{A}$ itself, it has a definite meaning, whereas this is not the case for irreducibility).

As a particular consequence of this definition, we would like to remark that, within an elementary representation, the "energy gap" $[\Delta^*(p)\Delta(p)]^{\frac{1}{2}}$ and consequently the "energy" $E(p)$ [see (3.7) and (3.9)] are *c numbers*. Therefore, for elementary representations, our generalized Valatin–Bogoliubov transformation (3.8)–(3.12) (which even in this restricted case does not reduce to the ordinary Valatin–Bogoliubov transformation) provides a true diagonalization of the total Hamiltonian.

From the information we have about our model, we want now to illustrate by an *explicit* construction the concept of *elementary* representation. The interest of this construction not only lies in the exhibition of an example, but could also help the reader to pass the bridge between our formalism (see also Ref. 4) and the usual treatment of the model.[5,2,3]

Let $\mathfrak{H}$ be the Hilbert space attached to a primary representation φ of $\mathfrak{A}$. Consider now the family of inequivalent primary representations (4.13) and form for each (positive or negative) integer N the representation

$$\phi_N = \int_{S^1} d\mu_\alpha \, \varphi_{\alpha^N} \tag{4.17}$$

acting in the space

$$\mathfrak{K} = \int_{S^1} d\mu_\alpha \, \mathfrak{H}_{(\alpha^N)}, \tag{4.18}$$

where $d\mu_\alpha = (1/2\pi)d\alpha$ is the Haar measure on the considered symmetry group G, i.e., in this model the gauge group $G \equiv S^1$. [In (4.18), the indice (α^N) to $\mathfrak{H}$ is simply supposed to recall that, in the direct integral space, the representation φ_{α^N} acts in that space, but the $\mathfrak{H}_\alpha$ are replicae of $\mathfrak{H}$.]

To show that each of the ϕ_N is compatible with G, it suffices to exhibit a representation $\mathrm{B}_N(G)$ for each N; let us then define for each $\beta \in G$:

$$\mathrm{B}_N(\beta) \circ \phi_N = \int_{S^1} d\mu_\alpha \, \varphi_{\alpha^N \cdot \beta}. \tag{4.19}$$

One immediately sees that $[\phi_N(\mathfrak{A}), \mathrm{B}_N(G)]$ satisfies

$$\phi_N \circ \beta[A] = \mathrm{B}_N(\beta) \circ \phi_N[A], \quad \forall A \in \mathfrak{A}, \quad \forall \beta \in G, \tag{4.20}$$

and

$$\phi_N[\mathfrak{Z}_G] = \{\lambda I\}. \tag{4.21}$$

Starting from any primary representation φ of $\mathfrak{A}$, we have therefore exhibited, for each integer N, an elementary representation of $(\mathfrak{A}, G)$.

(The reader who prefers to start from irreducibility rather than primarity, can obviously reproduce the above construction for his own.)

We still want to give as an indication, the image of the smeared out field and their transformation laws under the above representations.

$$\phi_N[\Psi_i(f)] = \int_{S^1} d\mu_\alpha \, e^{iN\alpha} \varphi[\Psi_i(f)], \tag{4.22}$$

$$\mathrm{B}_N(\beta) \circ \phi_N[\Psi_i(f)] = e^{i\beta} \phi_N[\Psi_i(f)], \tag{4.23}$$

$$\phi_N[\gamma_i(p)] = \int_{S^1} d\mu_\alpha \, e^{-iN\alpha} \varphi[\gamma_i(p)], \tag{4.24}$$

$$\mathrm{B}_N(\beta) \circ \phi_N[\gamma_i(p)] = e^{-i\beta} \phi_N[\gamma_i(p)]. \tag{4.25}$$

The reader is urged here to remember ϕ_N acts in $\mathfrak{K}$ and *not* in $\mathfrak{H}$. [An oversight of this would lead to a catastrophically trivial result in (4.22) or (4.24) for instance!]

We want next to mention, in connection with the above treatment of the BCS model, the question of continuous versus discrete superselection rules.

Because of the structure of the symmetry group G, we used for each elementary representation (4.17) a continuous family of replicae $\mathfrak{H}_{(\alpha^N)}$ of the same Hilbert space $\mathfrak{H}$ [see (4.18)]. The field operators, and consequently the observables constructed out of them, map each of these spaces $\mathfrak{H}_{(\alpha^N)}$ onto itself, so that we can really consider them as superselection

sectors (i.e., "coherent subspaces"). The fact that we have here an explicit example of a theory with *continuous* family of nontrivial superselection sectors has some interesting consequences.

For instance, a well-known theorem asserts that every connected symmetry-group maps every superselection sector onto itself, provided that the sequence of such sectors is at most countable. (For a proof of this assertion, in terms of the lattice structure of a physical theory, see Ref. 9, corollary to Th. 1.2; as already mentioned in Ref. 9, the proof, however, breaks down if the later assumption is not satisfied.) In the present model, although the gauge transformations obviously form a connected group, they exhibit a counterexample. In fact, classical mechanics also exhibits such a counterexample; for some reason, however, people are in general not very impressed by it, having probably in mind the prejudice that classical mechanics is "so different" from quantum mechanics that there cannot be any common formalism for both of them, and therefore no meaningful analogy between them; this idea is misleading; in fact, in the general frame of proposition calculus (see for instance Refs. 10, 11, or 4), a quantum mechanical theory with a continuous family of nontrivial superselection sectors has to be considered as a special intermediary case between the two extremes constituted respectively by classical and ordinary quantum mechanics.

The so-called infinite degeneracy of the vacua, their transformations under the gauge group, etc. are another, and more or less trivial aspect of the facts exposed up to here. We would therefore like to leave to the reader the translation in our formalism of Schrieffer's interesting comments (see Secs. 2.4, 2.5, and paragraph 6 of Sec. 2.7).

We only want to remark that, among all the vacua present in an elementary representation, there exists one vacuum state which is gauge invariant (and not pure!). The existence of this state presents some interest in our context. Because of a very general theorem (see Ref. 12, Theorem 2.12.11, for instance) the existence of an invariant cyclic state ensures that β can be implemented by a unitary operator U_β. Existence of an invariant cyclic state ensures that β can be implemented by a unitary operator U_β. However, U_β does not belong to $\mathfrak{A}''$ [otherwise it would leave $\mathfrak{B}(\mathfrak{A})$ element-wise invariant!] and can therefore not be constructed out of the field operators. Its generator is consequently deprivated of any physical significance.

One could further argue that even if it is convenient to have some nice continuity properties for the symmetries, it might be even better to "break" the symmetry and go over to an irreducible representation. We do not share this point of view. The main advantage (in a quantum field theory) of an irreducible representation is the existence of cluster properties. However, the cluster property reflects the local nature of a theory whereas the BCS model is a highly nonlocal theory because of the very existence of the "pairons."

We thus think that there is no compelling physical reason to prefer an irreducible representation in the case of the BCS model. In last analysis, which representation is chosen is mainly a matter of convenience. In this line we conclude that a reducible representation is mostly suitable to the discussion of the global properties of the model, whereas an irreducible representation (corresponding to a fixed gauge) is probably more adapted to certain practical calculations.

Note added in proof. Similar techniques have also been applied to other systems exhibiting phase transitions. See: G. Emch, Tech. notes BN-433 and BN-437, University of Maryland, 1966.

ACKNOWLEDGMENTS

We would like to thank here Professor V. Bargmann and Professor A. S. Wightman for discussions on the matter presented in this paper, and for their warm interest during the completion of this work.

[9] G. Emch, Helv. Phys. Acta **36**, 739 (1963). For a statement of this theorem only, see Ref. 10.

[10] G. Emch, Proceedings of the Symposium on the Lorentz Group, 7th Annual Summer Institute for theoretical physics, University of Colorado, Boulder, Colorado, June 1964,

[11] C. Piron, thesis, Lausanne, 1964.

[12] Ref. 8, p. 57.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 7, NUMBER 5 MAY 1966

Some Properties of the Long-Time Values of the Probability Densities for Moderately Dense Gases

ROBERT A. PICCIRELLI
National Bureau of Standards, Washington, D. C.
(Received 13 August 1965)

It has been argued that, for sufficiently large times, the n-particle probability densities of a moderately dense, simple gas become time-independent functionals of the one-particle probability densities. Proofs are given for several properties of the power series representation of these functionals. In particular, it is shown that the equilibrium value of the n-particle functional is identical to the usual equilibrium probability density term by term, and that the corresponding generalized Boltzmann collision integral vanishes as it should. The Green–Cohen form of the functional is shown to be a formal power series solution of Bogoliubov's functional differential hierarchy. Moreover, a proof is given that the Bogoliubov and Green–Cohen forms of the functional are formally identical term by term. It is argued that the higher terms of these two series probably diverge *together*.

In the course of the discussion, several new properties of the coefficient operators of the power series for the functional are derived. Moreover, an integral equation for the n-particle functional is derived which may have solutions not representable as functional power series in the one-particle probability density.

I. INTRODUCTION

THERE have been several studies of the long-time behavior of the many-particle probability densities for the case of large, dilute systems.[1–5] In particular, it has been argued[6,7] that for such systems, consisting of particles interacting through short-range, repulsive forces, if the initial values of the probability densities satisfy the product condition, then after long enough times they become time-independent functionals of the one-particle probability density.

Adopt the convention that the symbol $[n]$ denotes a set of n integers and that an integer appearing as the argument of a function or operator represents the position and momentum of the particle named by the integer. Then the asymptotic functional can be denoted by $f([n] \mid f_1(t))$ and has been represented in the following way:[6]

$$f([n] \mid f_1(t)) = \sum_{l \geq 0} \frac{1}{l!} \int d([l]) \tau^{(n)}([n]; [l]) \times \prod_{\alpha \in [n]+[l]} f_1(\alpha; t). \quad (1)$$

In this equation $f_1(\alpha, t)$ is the singlet probability density, and the "coefficient operator" $\tau_l^{(n)}$ is a sum of products of time-independent substitution operators each of which uniquely maps a given phase point into another.[8]

The purpose of the present paper is to provide *formal* proofs for several commonly supposed properties of this functional and its coefficient operators. Moreover, the intention is to indicate the non-algebraic steps in the proofs.

The first property is that the n-particle functional given in Eq. (1) evaluated "at" the equilibrium singlet density is identical with the density expansion of the equilibrium n-particle density. This property provides a *check* on the validity of the functional as the asymptotic solution and incidentally provides a re-expression of equilibrium theory.

This property has been studied before for the pair functional, $f(12|f_1)$. In particular, Bogoliubov[1] and Choh and Uhlenbeck[3] established it for the first two terms of Bogoliubov's form while Green established it for the first two terms of his form [Ref. 2, Eq. (1)]. In order to generalize these previous results to all the n-particle densities, a new recursion property of the coefficient operators is introduced which allows us to make the proofs algebraically.

Some results related to this one are given. In particular, it is shown that the generalized Boltzmann collision integral implied by Eq. (1) vanishes when evaluated at the local equilibrium singlet density for some fixed space point.[9] This result is

[1] N. N. Bogoliubov, J. Phys. (USSR) **10**, 265 (1946). See also *Studies in Statistical Mechanics*, J. De Boer and G. E. Uhlenbeck, Eds. (North-Holland Publishing Company, Amsterdam, 1962) Vol. 51.
[2] M. S. Green, Physica **24**, 393 (1958).
[3] S. T. Choh and G. E. Uhlenbeck, *The Kinetic Theory of Phenomena in Dense Cases*, Navy Contract Rept: Nonr. 1224 (1958).
[4] E. D. G. Cohen, *Fundamental Problems in Statistical Mechanics* (North-Holland Publishing Company, Amsterdam, 1962) p. 110.
[5] E. G. D. Cohen, J. Math. Phys. **4**, 183 (1963).
[6] M. S. Green and R. A. Piccirelli, Phys. Rev. **132**, 1388 (1963), hereafter referred to as I.
[7] E. G. D. Cohen, Physica **28**, 1025 (1962).
[8] The coefficient operators are more fully defined in the sequal by, for example, Eq. (10).
[9] The meaning of this last qualification is clarified in the text.

required in developing a generalized Chapman–Enskog theory of the Boltzmann equation.[10] For the first two terms of the pair functional in Bogoliubov's form, this result has been established by Choh and Uhlenbeck.[3] The generalization to all orders is a trivial consequence of the first property.

A second property of the functional given by Eq. (1) is that it is a formal series solution of the hierarchy for finite times. This verifies a necessary property of the form of the functional and at the same time is one part of a proof that Eq. (1) represents Bogoliubov's functional. The proof depends on a new expression for the commutator of the coefficient operators $\tau_l^{(n)}$ with the free-particle Liouville operator.

A third property of the series given in Eq. (1) is that it is term-by-term *formally* identical with the terms of the series implied by Bogoliubov's method. Obviously, such identity is necessary to ensure concurrence of the two lines of development. Proofs of particular cases of this result have been given before.[11] More generally, a proof of the equivalence of the two forms of the pair functional has been given by Cohen by using a uniqueness theorem.[12] In the course of generalizing these results, a modified and more compact form of Bogoliubov's derivation is given; namely, his series is shown to be the solution of a functional integral equation. Furthermore, the proof process yields forms which may be useful in analyzing the convergence of integrals.

Essentially, then, some previous work on the properties of the pair functional is generalized to all the n-particle densities and to arbitrary orders in the number density and some new results on the coefficient operators $\tau_l^{(n)}$ are obtained. The methods are algebraical.

Incidentally, the functional integral equation whose iteration solution is Bogoliubov's series is given. This result may be of some interest in view of the growing evidence that higher terms of the series in Eq. (1) are divergent.[13–15] It may be, for example, that the integral equation has a solution which is not a power series in f_1. In general, it is felt that, making the formal aspects of the proofs compact and overseeable (for example, by using recursion formulas instead of graphs), provides a useful point of departure for further analysis.

II. VALUE OF THE FUNCTIONAL AT EQUILIBRIUM

A proof will be given that the equilibrium value of the functional is the equilibrium probability density. The method will be to establish term-by-term equality with the equilibrium density series written in a suitable form. Several related results are also given.

A useful form of the equilibrium series can be derived in the following way. The activity series for the n-particle density may be given by

$$^{\circ}f([n]) = \sum_{l\geq 0} \frac{z^{n+l}}{l!} \times \int d([l])U^{(n)}([n];[l]) \prod_{\alpha\in[n]+[l]} {}^{\circ}\varphi_1(\alpha), \qquad (2)$$

where z is the absolute activity,

$$^{\circ}\varphi_1(\alpha) = (\beta/2\pi m)^{\frac{3}{2}} \exp(-\beta p_\alpha^2/2m), \qquad (3)$$

and $U^{(n)}([n];[l])$ is the usual modified Ursell function.[16] By analogy with the dynamical ones of Ref. 6, these may be defined recursively by:

$$e^{-\beta V([n]+[l])} = \sum_{[h]+[k]=[l]} U^{(n)}([n];[h])e^{-\beta V([k])}, \qquad (4)$$

where the summation is over all distinct partitions of $[l]$ into two disjoint parts, $[h]$ and $[k]$, either of which may be empty.[17] $V([m])$ is the potential energy of m particles in the configuration corresponding to $[m]$. It should be mentioned that, although $U^{(n)}$ depends only on positions, the integration in Eq. (2) is over positions and momenta.

The equilibrium density series can be derived by introducing a set of functions $H^{(n)}$ defined by

$$U^{(n)}([n];[l]) = \sum_{[q]+\sum_\alpha[r_\alpha]=[l]} H^{(n)}([n];[q]) \times \prod_{\alpha\in[n]+[q]} U^{(1)}(\alpha;[r_\alpha]), \qquad (5)$$

where the summation is over all partitions of $[l]$ into $[q]$, $(n+q)$, and other disjoint parts $[r_\alpha]$, any of which may be empty.[18] If one inserts this

[10] M. S. Green and L. Garcia-Colin, Physica (in press).
[11] M. S. Green (private communication).
[12] E. G. D. Cohen, Physica **28**, 1045 (1962).
[13] J. Weinstock, Phys. Rev. **132**, 454 (1963) (and another to be published).
[14] J. R. Dorfman and E. G. D. Cohen, Phys. Letters **16**, 124 (1965).
[15] J. V. Sengers, Phys. Rev. Letters **15**, 515 (1965), where full references are given.
[16] J. O. Hierschfelder, C. F. Curtiss, and R. B. Bird, *Molecular Theory of Gases and Liquids* (John Wiley & Sons, Inc., New York, 1954) p. 145 *et. seq.* where references to original papers are given. Moreover, Eq. (2) is the equilibrium value of Eq. (2.6) of Ref. 6.
[17] The symbolism for this type of summing is adhered to in the sequel.
[18] The functions $H_m^{(n)}$ are a generalization of Husimi's functions [J. Chem. Phys. **18**, 682 (1950)]. E. G. D. Cohen, Physica **28**, 1060 (1962), mentions this method of deriving the equilibrium series. It does not assume pair-wise forces.

expression into Eq. (2) and uses the dummy character of the integration variables, one finds that

$$^{\circ}f([n]) = \sum_{q\geq 0}\frac{1}{q!}\int d([q])H^{(n)}([n];[q]) \times \prod_{\alpha\in[n]+[q]}\sum_{r_\alpha\geq 0}\frac{z^{1+r_\alpha}}{r_\alpha!}\int d([r_\alpha])U^{(1)}(\alpha;[r_\alpha]) \times \prod_{\beta\in[\alpha,[r_\alpha]]}{}^{\circ}\varphi_1(p_\beta). \quad (6)$$

But, according to Eq. (2) for the case $n = 1$ the factors appearing in Eq. (6) are just the activity series for $^{\circ}f_1(\alpha) = c^{\circ}\varphi_1(\alpha)$. Thus, introducing for brevity

$$^{\circ}\phi([m]) = \prod_{\alpha\in[m]}{}^{\circ}f_1(\alpha), \quad (7)$$

one has that

$$^{\circ}f([n]) = \sum_{l\geq 0}\frac{1}{l!}\int d([l])H^{(n)}([n];[q])^{\circ}\phi([n]+[q]) \quad (8)$$

is a form of the equilibrium density series.

According to Eq. (1), the equilibrium value of the functional is given by

$$f([n]\mid {}^{\circ}f_1) = \sum_{l\geq 0}\frac{1}{l!}\int d([l])\tau^{(n)}([n];[l]) \times {}^{\circ}\phi([n]+[l]). \quad (9)$$

At first sight one wants to prove equality of the integrands in Eqs. (8) and (9), but they are, in fact, not equal. The integrals, however, are equal and this will be proved by showing that they obey the same recursion relation.

To accomplish this, we will use the result that $\tau^{(n)}$ is the solution of a recursion relation analogous to Eq. (5); namely,

$$\mathfrak{U}^{(n)}([n];[l]) = \sum_{[q]+\sum_\alpha[r_\alpha]=[l]}\tau^{(n)}([n];[q]) \times \prod_{\alpha\in[n]+[q]}\mathfrak{U}^{(1)}(\alpha;[r_\alpha]). \quad (10)$$

The Ursell operators $\mathfrak{U}^{(n)}$ are defined in terms of $\mathfrak{S}$-operators $\mathfrak{S}([m])$ in the same way as the $U^{(n)}$ are defined in terms of the Boltzmann factors by Eq. (4); that is, by

$$\mathfrak{S}([n]+[l]) = \sum_{[h]+[k]=[l]}\mathfrak{U}^{(n)}([n];[h])\mathfrak{S}([k]). \quad (11)$$

Finally, the $\mathfrak{S}$-operators are defined as the long-time limit of a product of streaming operators by

$$\mathfrak{S}([m]) = \lim_{\tau\to\infty}S([m];-\tau)S_0([m];\tau), \quad (12)$$

where $S([m];-\tau)$ translates the particles backwards in time according to the full m-body dynamics, while $S_0([m];\tau)$ translates them forward according to free-particle dynamics.[19]

In Appendix A, the recursion relation for $\tau^{(n)}$ given by Eq. (10) is established as a consequence of the original one given in Ref. 6.

The important step in the proof is to realize that

$$\mathfrak{U}^{(n)}([n];[l])^{\circ}\phi([n]+[l]) = U^{(n)}([n];[l])^{\circ}\phi([n]+[l]). \quad (13)$$

This result is demonstrated in Appendix B beginning with a basic property of the $\mathfrak{S}$-operators first used by Bogoliubov;[1] namely, that

$$\mathfrak{S}([m])\sum_{\alpha\in[n]}p_\alpha^2/2m = \sum_{\alpha\in[n]}p_\alpha^2/2m + V([m]). \quad (14)$$

This is simply a restatement of conservation of energy and is valid for any point. It implies that

$$\mathfrak{S}([m])^{\circ}\phi([m]) = e^{-\beta V([m])}{}^{\circ}\phi([m]), \quad (15)$$

from which Eq. (13) follows by algebra.[20]

With the result of operating with $\mathfrak{U}^{(n)}$ given by Eq. (13), one can derive a recursion relation for the terms in Eq. (9). Thus, multiplying Eq. (10) on the right by $^{\circ}\phi([n]+[l])$, using Eq. (13), and integrating over the points $[l]$, one finds that

$$\frac{1}{l!}\int d([l])U^{(n)}([n];[l])^{\circ}\phi([n]+[l]) = \sum_{q+\sum_\alpha r_\alpha=l}\frac{1}{q!}\int d([q])\tau^{(n)}([n];[q])^{\circ}\phi([n]+[q]) \times \prod_{\alpha\in[n]+[q]}\frac{1}{r_\alpha!}\int d([r_\alpha])U^{(1)}(\alpha;[r_\alpha])^{\circ}\phi([r_\alpha]). \quad (16)$$

Introducing the modified cluster integrals by

$$c^l b_l^{(n)}([n]) = \frac{1}{l!}\int d([l])U^{(n)}([n];[l])^{\circ}\phi([l]), \quad (17)$$

and also defining integrals $\tilde{B}_l^{(n)}$ by

$$c^l{}^{\circ}\phi([n])\tilde{B}_l^{(n)}([n]) = \frac{1}{l!}\int d([l])\tau^{(n)}([n];[l])^{\circ}\phi([n]+[l]), \quad (18)$$

Eq. (16) implies that

$$b_l^{(n)}([n]) = \sum_{q+\sum_\alpha r_\alpha=l}\tilde{B}_q^{(n)}([n])\prod_{1\leq\alpha\leq n+q}b_{r_\alpha}. \quad (19)$$

An important point to notice is that the last factors in Eq. (16) are actually equal to $b_{r_\alpha}^{(1)}(\alpha)$, but that for the infinite system with pairwise forces these

[19] See Ref. 6 for a discussion of these operators.

[20] Incidentally, one can view this equation as establishing the $\mathfrak{S}$-operator to be the analog of the Boltzmann factor (instead of S_{-t}, for example).

are independent of the position of particle α and are, in fact, equal to b_{r_α}, the usual cluster integrals. Thus, the terms in the $\tau^{(n)}$ are operating only on $^\circ\phi$ and not on the remaining factors.

Now, if one defines integrals $B_l^{(n)}$ by

$$^lB_l^{(n)}([n]) = \frac{1}{l!}\int d([l])H^{(n)}([n];\,[l])^\circ\phi([l]), \qquad (20)$$

one finds from Eq. (5) [multiplying by $^\circ\phi([l])$ and integrating] exactly the *same* relation for $B_l^{(n)}$ as Eq. (19) gives for $\tilde{B}_l^{(n)}$. But, according to Eqs. (18) and (20),

$$\tilde{B}_0^{(n)}([n]) = B_0^{(n)}([n]) = \exp\{-\beta V([n])\}, \qquad (21)$$

where we have used the fact that $\tau_0^{(n)}([n]) = \mathcal{S}([n])$. Moreover, it seems clear that once one has begun the solution of the recursion relation with a given $l = 0$ values, the solution is unique. Therefore, $\tilde{B}_l^{(n)} = B_l^{(n)}$ for any l and the term-by-term equality of the equilibrium value of the functional with the equilibrium density series has been established.

Several related results are now easily established. First, recall that the terms of the equilibrium density series are known to converge for appropriate potentials. Since we have just proved equality of these with the terms of the equilibrium value of the functional without any switching of limits or other possibly nonrigorous steps, one can conclude that the terms of the equilibrium value of the functional converge.

A second result is that the equilibrium value of the functional satisfies Bogoliubov's boundary condition. In the present notation this condition states that the functional must satisfy

$$\lim_{\tau\to\infty} S([n];\,-\tau)f([n]\mid S_1(\tau)f_1) = S([n])\prod_{\alpha\in[n]} f_1(\alpha;\,t). \qquad (22)$$

For the equilibrium value of $f([n]|f_1)$, then, one needs

$$\lim_{\tau\to\infty} S([n];\,-\tau)e^{-\beta V([n])}{}^\circ\phi([n]) = S([n])^\circ\phi([n]), \qquad (23)$$

and, for the higher-order terms, one needs

$$c^l \lim_{\tau\to\infty} S([n];\,-\tau)^\circ\phi([n])B_l^{(n)}([n]) = 0. \qquad (24)$$

But since $S([n];\,-\tau)$ conserves energy and one can use Eq. (15), Eq. (23) is an identity for any point $[n]$. Moreover, $B_l^{(n)}([n])$, being the term of the usual equilibrium density series, vanishes whenever there are widely separated clusters. But for purely repulsive forces any point eventually becomes complete; that is, $S([n];\,-\tau)$ produces a point in which all the particles are widely separated and moving with the asymptotic values of their momenta. Then the first factor in Eq. (24) becomes a constant while the second vanishes.

The third and more useful result which can be established by similar techniques is that the generalized Boltzmann collision integral vanishes when evaluated at the local equilibrium density $f_0(\mathbf{r}t)$ for the fixed point $\mathbf{r}$. More precisely, the generalized collision integral may be expressed by

$$C(x_1 \mid f_1(t)) = \int d(2)L'(12)f(12 \mid f_1(t)), \qquad (25)$$

where the pair functional is given in Eq. (1).

The pair-interaction Liouville operator is defined by

$$L'(ij) = \mathbf{F}_{ij}\cdot\nabla_{\mathbf{p}i} + \mathbf{F}_{ji}\cdot\nabla_{\mathbf{p}j}, \qquad (26)$$

where $\mathbf{F}_{ij}$ is the interparticle force. Also, the local equilibrium density is defined by

$$f_0(\mathbf{r}\mathbf{p},\,t) = c(\beta/2\pi,\,m)^{\frac{3}{2}} \times \exp\{-\beta(\mathbf{p}_\alpha - m\mathbf{v})^2/2m\}, \qquad (27)$$

where c, β, and $\mathbf{v}$ are the local density, inverse temperature (times Boltzmann's constant) and local velocity considered as functions of position and time. Then the statement is that

$$C(x_1 \mid f_0(\mathbf{r}_1,\,t)) = 0, \qquad (28)$$

where the instruction is to put $f_0(\mathbf{r}_1\mathbf{p}'t)$ for $f_1(\mathbf{r}'\mathbf{p}'t)$ in the definition of the functional.[21]

The result given by Eq. (28) follows by proving that

$$f(12 \mid f_0(\mathbf{r}_1,\,t)) = {}^\circ f(\mathbf{r}_1\mathbf{p}_1^*\mathbf{r}_2\mathbf{p}_2^*), \qquad (29)$$

where $\mathbf{p}_i^* = \mathbf{p}_i - m\mathbf{v}$. For, we know that the equilibrium pair function $^\circ f$ (12) is *even* in r_{12} while L' (12) is *odd*

To prove that Eq. (29) holds, one has only to notice that

$$\mathcal{S}([m])\prod_{\alpha\in[m]} g_1(\mathbf{r}_\alpha\mathbf{p}_\alpha)\big|_{g_1(\alpha)=f_0(\mathbf{r}_1\mathbf{p}_\alpha)} = e^{-\beta V([m])}\prod_{\alpha\in[m]} f_0(\mathbf{r}_1\mathbf{p}_\alpha) \qquad (30)$$

in complete analogy with Eq. (15). This is because one has not only Eq. (14) but also

$$\mathcal{S}([m])\sum_{\alpha\in[m]}\mathbf{p}_\alpha = \sum_{\alpha\in[m]}\mathbf{p}_\alpha \qquad (31)$$

by conservation of momentum. Thus, beginning with Eq. (30), the parallel to the whole previous proof is obvious and Eq. (29) is established.

[21] This means, in other words, that we consider only the zero gradient approximation to the functional. The point $(\mathbf{r}', \mathbf{p}')$ here denotes, for example, the dummy integration variables in the definition of the functional. Also, see Eq. (30).

III. A SOLUTION OF THE FUNCTIONAL HIERARCHY

The purpose here is to assure ourselves that, regardless of previous methods of derivation, the Green–Cohen form of the functional given in Eq. (1) is indeed a series solution of the "functional hierarchy" first given by Bogoliubov. Even if the terms in the series do not exist, assurance that at least one does have a representation of the functional is a necessary preliminary to further work such as resummation. In addition, the commutator of the coefficient operators $\tau_l^{(n)}$ with the free-particle Liouville operator is evaluated. This result (which is suggested by presupposing that the Green–Cohen form is a solution) is of some interest in itself.

That the functional must satisfy an appropriate hierarchy is a consequence of the fact that the probability densities for an infinite system satisfy the hierarchy demanded by Liouville's theorem. In fact, the renormalized, generic, n-particle, probability densities are solutions of the system

$$\left\{\frac{\partial}{\partial t} + L([n])\right\} f([n]; t) + \epsilon \int d(\beta) \times \sum_{\alpha \in [n]} L'(\alpha\beta) f([n] + \beta; t) = 0, \tag{32}$$

where the renormalized n-particle density is related to the usual one by a factor ϵ^n, and ϵ is the density times the cube of the length parameter of the potential.[22] This technique is simply to remind us of orders of magnitude. The n-particle Liouville operator is defined by

$$L([n]) = L_0([n]) + \sum_{\substack{\alpha,\beta \in [n] \\ \alpha < \beta}} L'(\alpha\beta), \tag{33}$$

where $L'(\alpha\beta)$ is the pair-interaction Liouville operator defined in Eq. (26) and $L_0([n])$ is the free-particle Liouville operator defined by

$$L_0([n]) = \sum_{\alpha \in [n]} \frac{\mathbf{p}_\alpha}{m} \cdot \nabla_{\mathbf{r}_\alpha}. \tag{34}$$

Suppose one seeks a solution of Eq. (32) which is a time-independent functional of f_1. By definition

$$\frac{\partial f([n] \mid f_1(t))}{\partial t} = \int d(\gamma) \frac{\delta f([n] \mid f_1(t))}{\delta f_1(\gamma; t)} \frac{\partial f_1(\gamma t)}{\partial t}, \tag{35}$$

and the time derivative of f_1 can be re-expressed by using the first of Eqs. (32). Thus, in order for such a functional to be a solution of the hierarchy, it must satisfy the following hierachy of intergro-differential equations:

$$- \int d(\gamma) \frac{\delta f([n] \mid f_1)}{\delta f_1(\gamma)} L_0(\gamma) f_1(\gamma) + L([n]) f([n] \mid f_1) + \epsilon \Phi([n] | f_1(t)) = 0. \tag{36}$$

The functional Φ which has been introduced is defined by

$$\Phi([n] \mid f_1(t)) = \int d(\beta) \left\{ \sum_{\alpha \in [n]} L'(\alpha\beta) f([n], \beta \mid f_1(t)) - \int d(\gamma) \frac{\delta f([n] \mid f_1(t))}{\delta f_1(\gamma, t)} L'(\beta\gamma) f(\beta\gamma \mid f_1(t)) \right\}. \tag{37}$$

This differential functional hierarchy was first derived by Bogoliubov.[1]

Consider the series solution of Eq. (36) of the form

$$f([n] \mid f_1) = \sum_{l \geq 0} \epsilon^l f^{(l)}([n] \mid f_1), \tag{38}$$

where the terms are ordered according to increasing powers of the density. Then, equating equal powers of ϵ, one finds a recursion relation for the terms of the functional; namely,

$$- \int d(\gamma) \frac{\delta f^{(l)}([n] \mid f_1)}{\delta f_1(\gamma)} L_0(\gamma) f_1(\gamma) + L([n]) f^{(l)}([n] \mid f_1) + \Phi^{(l)}([n] \mid f_1(t)) = 0. \tag{39}$$

The functional $\Phi^{(l)}$ is the term of $0(\epsilon^l)$ in the expansion of $\epsilon\Phi$; that is, it vanishes for $l = 0$ and for $l \geq 1$,

$$\Phi^{(l)}([n] \mid f_1) = \int d(\beta) \left\{ \sum_{\alpha \in [n]} L'(\alpha\beta) f^{(l-1)}([n], \beta \mid f_1) - \sum_{\substack{q,r \geq 0 \\ q+r=l-1}} \int d(\gamma) \frac{\delta f^{(q)}([n] \mid f_1)}{\delta f_1(\gamma)} L'(\beta\gamma) f^{(r)}(\beta\gamma \mid f_1) \right\}. \tag{40}$$

Now, according to Eq. (1),

$$f_G^{(l)}([n] \mid f_1) = \frac{1}{l!} \int d([l]) \tau^{(n)}([n]; [l]) \phi([n] + [l]), \tag{41}$$

where, for brevity,

$$\phi([m]; t) = \prod_{\alpha \in [m]} f_1(\alpha, t) \tag{42}$$

has been introduced. The series given by Eq. (1) solves the functional hierarchy if $f_G^{(l)}$ is a solution of Eq. (39).

To show that this is the case, notice that by Eq. (41)

$$\frac{\delta f_G^{(l)}([n] \mid f_1)}{\delta f_1(\gamma)} = \frac{1}{l!} \int d([l]) \tau^{(n)}([n]; [l]) \times \sum_{\lambda \in [n]+[l]} \delta(x_\gamma - x_\lambda) \prod_{\substack{\alpha \in [n]+[l] \\ \alpha \neq \lambda}} f_1(\gamma). \tag{43}$$

[22] The parameter ϵ can also be considered as a purely formal counting parameter eventually to be set equal to unity.

Inserting Eqs. (41) and (43) into the system Eq. (39), and doing the integrations over the δ-functions yields the equation

$$\frac{1}{l!}\Big\{L([n])\int d([l])\tau^{(n)}([n];[l])\phi([n]+[l]) - \int d([l])\tau^{(n)}([n];[l])L_0([n]+[l])\phi([n]+[l])\Big\} + \Phi_G^{(l)}([n]\mid f_1) = 0, \tag{44}$$

and the subscript G on $\Phi^{(l)}$ denotes that it has been evaluated using Eq. (41). It is given by

$$\Phi_G^{(l)}([n]\mid f_1) = \frac{1}{(l-1)!}\int d(\beta)\sum_{\alpha\in[n]}L'(\alpha\beta) \times \int d([l-1])\tau^{(n+1)}([n]+\beta;[l-1]) \times \phi([n]+\beta+[l-1]) - \sum_{q+r=l-1}\frac{1}{q!\,r!}\int d(\beta) \times \int d([q])\tau^{(n)}([n];[q])\sum_{\gamma\in[n]+[q]}L'(\beta\gamma) \times \int d([r])\tau^{(2)}(\beta\gamma;[r])\phi([n]+\beta+[q]+[r]). \tag{45}$$

Care has been taken in deriving Eq. (44) so as not to interchange differentiation with the phase-space integrations (which are over infinite regions). Such an interchange is the only nonalgebraic step in the proof, and it requires that the integrals converge sufficiently strongly in some sense. Having remarked on this point we proceed formally by letting the Liouville operators operate before integration. Thus, taking advantage of the dummy character of integration variables to transform away the combinatorial factors, Eq. (44) becomes

$$\frac{1}{l!}\int d([l])\{L([n])\tau^{(n)}([n];[l]) - \tau^{(n)}([n];[l])L_0([n]+[l]) + \Delta\Gamma_1^{(n)}([n];[l])\}\phi([n]+[l]) = 0. \tag{46}$$

The operator $\Delta\Gamma_1^{(n)}([n];[l])$ vanishes for $l = 0$, and for $l \geq 1$ is defined by

$$\Delta\Gamma_1^{(n)}([n];[l]) = +\sum_{\substack{\alpha\in[n]\\ \beta+[q]=[l]}}L'(\alpha\beta)\tau^{(n+1)}([n],\beta;[q]) - \sum_{\beta+[q]+[r]=[l]}\tau^{(n)}([n];[q])\sum_{\alpha\in[n]+[q]}L'(\alpha\beta)\tau^{(2)}(\alpha\beta;[r]), \tag{47}$$

where the summation is over all distinct partitions of $[l]$ into the single element β and two other disjoint parts $[q]$ and $[r]$, either of which may be empty.

That Eq. (46) is indeed an identity is a direct consequence of the following identity for the commutator of $\tau_l^{(n)}$ with L_0:

$$[L_0;\tau^{(n)}([n];[l])] = -\{L'([n])+L'([l])\}\tau^{(n)}([n];[l]) - \Delta\Gamma_1^{(n)}([n];[l]) - \Delta\Gamma_2^{(n)}([n];[l]), \tag{48}$$

where $L'([n])$ is the interaction part of the n-particle Liouville operator [the second sum in Eq. (33)]. The new operator $\Delta\Gamma_2^{(n)}([n];[l])$ which appears vanishes for $l = 0$, and for $l > 1$ is a sum of terms each of which contains interaction operators which operate only on particles of the set $[l]$. Although a detailed form of this operator will not be needed here, for completeness such a form is given in Appendix D.

The proof of the result given by Eq. (48) is based on the analogous result for the scattering operator; namely, that

$$[L_0([m]);\mathcal{S}([m])] = -L'([m])\mathcal{S}([m]). \tag{50}$$

This identity has been given before.[23] An alternate proof which perhaps shows more directly what is involved is given in Appendix C. From Eq. (50) the commutator of $\mathfrak{U}^{(n)}$ with L_0 can be computed (see Appendix C) and this result together with the defining Eq. (10) for $\tau^{(n)}$ enables one to compute the commutator of $\tau^{(n)}$. This calculation is carried out in Appendix D. Granting Eq. (50), the expressions for the other two commutators are derived by algebraic manipulation.

Now, according to Eq. (48) one has for the left-hand side of Eq. (46) that

$$\int d([l])\{L([n])\tau^{(n)}([n];[l]) - \tau^{(n)}([n];[l])L_0([n]+[l]) + \Delta\Gamma_1^{(n)}([n];[l])\}\phi([n]+[l]) = -\int d([l])\{L'([l])\tau^{(n)}([n];[l]) + \Delta\Gamma_2^{(n)}([n];[l])\} \times \phi([n]+[l]). \tag{51}$$

But the right-hand side of this equation vanishes identically because every term is either a space gradient or a momentum gradient with respect to one of the variables of integration. Therefore Eq. (46) holds.

Since the validity of Eq. (46) implies that $f_G^{(l)}([n]|f_1)$ is a solution of the recursion relation, Eq. (39), one concludes that the Green–Cohen form

[23] M. S. Green, Phys. Rev. **136** A905 (1964).

of the functional is indeed a series solution of Bogoliubov's functional hierarchy.

It should be remarked that if the terms of the functional are divergent, the proof becomes formal but it still establishes that the series is a solution. More precisely, it would be a *formal* power series solution and as such would still contain useful information about the functional.

If one could now also show that this form of the functional satisfied Bogoliubov's boundary condition, one could invoke Cohen's uniqueness theorem,[12] thereby proving equivalence between the two forms. Since the satisfaction of the boundary condition is conditional on the convergence of the terms of the functional, however, a separate discussion of it is given. An alternate equivalence proof given next is more direct and seems more useful.

IV. BOGOLIUBOV'S FORM OF THE FUNCTIONAL AND COMPARISON

Bogoliubov's functional is also a series solution of Eq. (36) and yet the terms have a different form from those given in Eq. (1). As a preliminary to discussing the equivalence of the two forms, a compact but general version of his derivation will be given which shows that his solution is the solution of a hierarchy of functional integral equations. The term-by-term equivalence proof is presented in this context.

In his work Bogoliubov first assumes a series solution of the differential hierarchy Eq. (36) so that he is immediately led to Eqs. (39) which he then solves. It seems more instructive, however, to invert these steps.

To do this first notice that

$$\frac{\partial f([n] \mid S_1(T)f_1(t))}{\partial T} = \int d(\gamma)\, \frac{\delta f([n] \mid S_1(T)f_1(t))}{\delta S_1(T)f_1(t)}\, L_0(\gamma)S_1(\gamma; T)f_1(\gamma; T), \tag{52}$$

where we have used the defining relation

$$\frac{\partial S_1(\gamma T)}{\partial T} = L_0(\gamma)\cdot S_1(\gamma; T). \tag{53}$$

Since Eq. (36) is a functional equation, one can make a change of variable from $f_1(t)$ to $S_1(-T)f_1(t)$, and then, using Eq. (52), one derives that the functional has to satisfy

$$\frac{\partial f([n] \mid S_1(T)f_1(t))}{\partial T} - L([n])f([n] \mid S_1(T)f_1(t)) = +\epsilon\Phi([n] \mid S_1(T)f_1(t)), \tag{54}$$

where, it may be recalled, Φ is defined by Eq. (37). Now, treating the functional Φ as an inhomogeneous term, one can write down the Green's function solution of Eq. (54). If, in turn, one operates on this solution with $S([n]; -T)$, one finds that

$$f([n] \mid f_1(t)) = S([n]; -T)f([n] \mid S_1(T)f_1(t)) - \epsilon \int_0^T d\tau\, S([n]; -\tau)\Phi([n] \mid S_1(\tau)f_1(t)). \tag{55}$$

This integral functional hierarchy is *equivalent* to the original Eq. (36) and must be satisfied at least for any *finite* time T.

At this point one supposes with Bogoliubov two things: (a) that the boundary condition given by Eq. (22) holds, thus imposing a weakening of correlations; and (b) Eqs. (55) continue to hold in the limit as $T \to \infty$. Under these conditions the functional must be a solution of a hierarchy of integral equations, namely,

$$f([n] \mid f_1(t)) = \mathfrak{s}([n])\phi([n]; t) - \epsilon \int_0^\infty d\tau\, S([n]; -\tau)\Phi([n] \mid S_1(\tau)f(t)). \tag{56}$$

This system of equations is the functional integral hierarchy already announced and perhaps two remarks should be made. The first is that, because of the limiting process, it is not entirely clear how the solutions of Eq. (56) are related to those of Eq. (36). The second is that it is certainly possible that this system has solutions which are not expandable as functional power series.[24]

Continuing with the derivation of Bogoliubov's form, however, assume a series solution as in Eq. (38). A recursion relation for the terms emerges by substituting this series into Eq. (56) and equating like powers of ϵ. Using the subscript B to indicate that these are Bogoliubov's form of the terms, one finds that

$$f_B^{(0)}([n] \mid f_1(t)) = \mathfrak{s}([n])\phi([n]; t), \tag{57}$$

and, for $l \geq 1$,

$$f_B^{(l)}([n] \mid f_1(t)) = -\int_0^\infty d\tau\, S([n]; -\tau)\Phi_B^{(l)}([n] \mid S_1(\tau)f_1(t)). \tag{58}$$

The functional $\Phi_B^{(l)}$ is given by Eq. (40) except that the particular terms $f_B^{(l)}$ are used. Equation (58) is implicit in Bogoliubov's work and his explicit expression for $f^{(1)}$ is easily recovered by using Eq. (57) to evaluate Eq. (58) for $l = 1$.

Equation (41) gives the explicit expression for

[24] For example, it may be representable as a sequence of nonanalytic functionals.

the Green–Cohen form of $f^{(l)}$. To show *directly* that here one has two equivalent representations of the same functional, an inductive proof of term-by-term equality will be given.

The implication which will be established for any $[n]$ is the following: if the terms of the two functionals are equal up to, say, the $(l-1)$th, then the lth terms are also equal. Since the first terms are equal by definition, a proof of this implication establishes the equivalence.

Proceed by first noticing that terms of $\Phi_B^{(l)}$ appearing in Eq. (58) involve $f_B^{(r)}$ only for $0 \leq r \leq l-1$, so that by hypothesis each of these can be evaluated using $f_G^{(r)}$. The result of evaluating $\Phi^{(l)}$ by using $f_G^{(r)}$ is already given by Eq. (45). If one again proceeds formally by interchanging the Liouville operators with the phase space integrals, one then has by hypothesis that

$$\Phi_B^{(l)}([n] \mid f_1) = \frac{1}{l!}\int d([l])\Delta\Gamma_1^{(n)}([n];[l])\phi([n]+[l]), \quad (59)$$

where $\Delta\Gamma_1^{(n)}$ is defined by Eq. (47). Using this result in Eq. (58), one finds that the hypothesis of equivalence for all $r \leq l-1$ implies that

$$f_B^{(l)}([n] \mid f_1) = -\frac{1}{l!}\int_0^\infty d\tau \int d([l])S([n];-\tau) \times \Delta\Gamma_1^{(n)}([n];[l])S_0([n]+[l];\tau)\phi([n]+[l]). \quad (60)$$

No additional assumption is needed in allowing the S-operator to operate before the integration since the operator simply instructs one to evaluate at a certain point.

To continue the proof of the implication statement, one uses the expression for $\Delta\Gamma_1^{(n)}$ given by Eq. (48) to rewrite Eq. (60) and finds that

$$\begin{aligned} f_B^{(l)}([n] \mid f_1) = {} & +\frac{1}{l!}\int_0^\infty d\tau \int d([l])S([n];-\tau) \\ & \times \{[L([n])+L_0([l])]\tau^{(n)}([n];[l]) \\ & - \tau^{(n)}([n];[l])L_0([n]+[l])\} \\ & \times S_0([n]+[l];+\tau)\phi([n]+[l]) \\ & + \frac{1}{l!}\int_0^\infty d\tau\, S([n],-\tau)\int d([l])\{L'([l])\tau^{(n)}([n];[l]) \\ & + \Delta\Gamma_2^{(n)}([n];[l])\}S_0([n]+[l];+\tau)\phi([n]+[l]). \end{aligned} \quad (61)$$

The second term vanishes, however, because every term in the integrand contains the pair operator $L'(ij)$ for arguments which are contained in $[l]$; that is, each term contains a momentum gradient over which one must integrate. Moreover, it is formally obvious that

$$\begin{aligned} & -(\partial/\partial\tau)\{S([n];-\tau)S_0([l];-\tau) \\ & \qquad \times \tau^{(n)}([n];[l])S_0([n]+[l];\tau)\} \\ & = S([n];-\tau)S_0([l];-\tau) \\ & \quad \times \{[L([n])+L_0([l])]\tau^{(n)}([n];[l]) \\ & \quad - \tau^{(n)}([n];[l])L_0([n]+[l])\}S_0([n]+[l];\tau). \end{aligned} \quad (62)$$

Since, in the first term in Eq. (61) one is free to transform coordinates from x_i to $S([l],-\tau)x_i$, one can use Eq. (62) to rewrite Eq. (61). The result is that at this stage one has, by hypothesis, that

$$\begin{aligned} f_B^{(l)}([n] \mid f_1) = {} & -\lim_{T\to\infty}\frac{1}{l!}\int_0^T d\tau \int d([l]) \\ & \times \frac{\partial}{\partial\tau}\{S([n];-\tau)S_0([l],-\tau)\tau^{(n)}([n];[l]) \\ & \times S_0([n]+[l];\tau)\}\phi([n]+[l];t). \end{aligned} \quad (63)$$

The final steps of the proof are not algebraic, nor is the justification of them trivial. In order to complete a formal proof without interruption, questions of justification are ignored for the moment. Obviously, one wants to perform the time-integral in Eq. (63), but to do so one must either interchange the two integrations or interchange the time derivative with the phase space integration. Following the first course[25] and then doing the time integral, one gets formally that

$$\begin{aligned} f_B^{(l)}([n] \mid f_1) = {} & -\lim_{\tau\to\infty}\frac{1}{l!}\int d([l]) \\ & \times \{T^{(n)}([n];[l],\tau) - \tau^{(n)}([n];[l])\}\phi([n]+[l];t), \end{aligned} \quad (64)$$

where we have defined the operator $T^{(n)}$ by

$$\begin{aligned} T^{(n)}([n];[l],\tau) = {} & S([n];-\tau)S_0([l];-\tau) \\ & \times \tau^{(n)}([n];[l])S_0([n]+[l];\tau). \end{aligned} \quad (65)$$

But the vanishing of the first term is assured if the Green–Cohen form of the functional satisfies Bogoliubov's boundary condition. For one has by inspection that

$$\begin{aligned} & \lim_{\tau\to\infty}\frac{1}{l!}\int d([l])T^{(n)}([n];[l],\tau)\phi([n]+[l];t) \\ & \qquad = \lim_{\tau\to\infty} S([n];-\tau)f_G^{(l)}([n] \mid S_1(\tau)f_1(t)), \end{aligned} \quad (66)$$

and the vanishing of the right-hand side of this equation is Bogoliubov's boundary condition. [See

[25] The conditions for either are fairly strong because the integrals are improper.

Eq. (22).] Thus, the argument finally yields that by hypothesis

$$f_B^{(l)}([n] \mid f_1) = \frac{1}{l!}\int d([l])\tau^{(n)}([n]; [l])\phi([n] + [l])$$
$$= f_G^{(l)}([n] \mid f_1). \tag{67}$$

Thus, the implication is established. This result can be summarized in the following way: The formal term-by-term equivalence between the two forms of the functional has been established granting that both forms satisfy the Bogoliubov boundary condition. A complete proof requires: (a) the existence of the terms in each form; (b) a proof that the Green–Cohen form does satisfy Bogoliubov's boundary condition; and (c) some conditions for the inversion of limit processes must be met.

V. DISCUSSION

It has been shown by Dorfman and Cohen that (contrary to an earlier conjecture in Ref. 6) in three dimensions the volume of the phase space for which one has contributions to the 4-particle term of the pair functional is infinite.[14] Weinstock also arrives at an infinite phase volume for a related integral.[13] Furthermore, Sengers has shown by explicit calculation that, for hard disks in two dimensions, the 3-particle term in the solution of the generalized Boltzmann equation diverges.[15] With this evidence, one should conjecture that all the higher terms of the functional series do not exist (at least for some kinds of f_1).[26] Without giving detailed estimates, a number of general remarks can be made giving the bearing of convergence questions on the previous results.

First, the various properties of the S-operators and their derivates are completely independent of any convergence questions. Thus, properties such as those in Eqs. (13) and (14), and the results for commutators [i.e., Eqs. (48) and (50) as well as Eq. (C10)] are rigorous rather than merely formal. Second, since only such properties are used to prove that the Green–Cohen form of the functional, when evaluated at the equilibrium singlet density, is equal to the usual equilibrium series, then this result is also considered rigorous along with the result that the generalized Boltzmann collision integral evaluated at local equilibrium vanishes.

As we have already remarked, the proof of Sec. III formally establishes the Green–Cohen form of the functional as a series solution of the differential hierarchy. Clearly, the proof cannot be made rigorous if the terms of the series are nonexistent. One can only conclude that the Green–Cohen form is a *formal* series representation of the functional. Even such formal series embody useful information.

A fourth remark is that, as one expects, $f_G^{(l)}$ will not satisfy Bogoliubov's boundary condition for those f_1 for which it is infinite. In the spatially homogeneous case, $f_G^{(l)}$ evaluated at $S_1(\tau)f_1$ is equal to its value at f_1 which value is supposed infinite for any point $[n]$. Thus, the required limit on the right-hand side of Eq. (66) does not exist. In the general case, there is a qualitative argument that the integral on the left-hand side of Eq. (66) is nonexistent. For example, consider $T^{(2)}$ (12; 3, τ) as a function of time, say t_2, of the last collision of particle 3 with particle 1 or 2. For points which contribute to the divergence of the integral (e.g., the recollision event), it is not difficult to see that, considered as a function of t_2, $T^{(2)}(12; 3, \tau)$ increases *sharply* at time τ from very small values to $\tau^{(2)}(12; 3)$. Using t_2 as a variable, the integration over particle 3 then contains two parts corresponding to these two values of $T^{(2)}(12; 3, \tau)$. The part for which it is equal to $\tau^{(2)}(12; 3)$ is a piece of $f_G^{(1)}$ (12) and diverges for the same reason. In its general form, this argument implies that the limit required by Bogoliubov's boundary condition does not exist.

One concludes that the uniqueness of proof of equivalence mentioned at the end of Sec. III is no longer valid if $f_G^{(l)}$ does not exist. This opens up the possibility that Bogoliubov's terms $f_B^{(l)}$ might exist even if the $f_G^{(l)}$ do not. However, our fifth remark is that the term-by-term method of Sec. IV suggests that Bogoliubov's terms fail to exist whenever the Green–Cohen terms do not exist and for the same reason. In other words, one has, so to speak, two *equivalent forms* of the same infinite quantity.

One has the following qualitative argument: The essential reason for the divergence of $f_G^{(l)}$ is that the volume of space corresponding to tightly connected points does not decrease with the size of the diagram sufficiently rapidly. Consider a cut-off integral of the Green–Cohen form in which the integration over the "size" of the diagram is truncated at some finite value. Using such a cut-off $f_G^{(l)}$, the entire calculation of Sec. IV can be repeated and the cut-off allowed to become infinite

[26] Preliminary calculation by Sengers (private communication) on the 3-particle term of the pair functional for hard disks in two dimensions indicates that this term diverges for, say, the Hermite polynomials (to be published in Phys. Fluids.)

in the *last* step. Our suggestion is that all the steps which lead to Eq. (67) are then justified and one arrives at this same equation except that the integration over the particles $[l]$ is truncated. This would indicate that $f_B^{(l)}$ becomes infinite with $f_G^{(l)}$.

Two final remarks may be of some interest. There is no divergence of the functional evaluated at equilibrium and one can speculate that the functional should have a kind of continuity in the functions f_1. Thus, one feels that there should be some class of perturbations about equilibrium (i.e., functions f_1) for which the terms exist in spite of the fact that they do not exist for, say, the Hermite polynomials.[27] In particular, they might exist for functions corresponding to perturbations which are local in configuration space. Then one might further speculate that the equivalence proofs *can* be established for this class of functions.

Such a situation seems difficult to verify, however, and in any case, the result would be of no value for calculating transport coefficients. It seems more useful to speculate that a time-independent functional exists and is a solution of the functional integral hierarchy given by Eq. (56). One must represent it, however, not by the present power series but by some resummed version of it.[28] In this connection one can observe that according to Eq. (64) the divergence in Bogoliubov's terms might arise at the upper limit of the time integrations. This suggests that a convergent representation of the functional might be achieved if the series were resummed so as to replace the ordinary streaming operator by a modified one which decayed with time.

ACKNOWLEDGMENTS

I would like to thank Dr. M. S. Green for his continued encouragement and support in this work, and Dr. J. V. Sengers for his helpful criticism.

APPENDIX A. RECURSION RELATION FOR $\tau^{(n)}$

Equation (10) is entirely equivalent to the original defining relation given in Ref. 6 [Eq. (4.2)]; namely,

$$\mathcal{S}([n]+[l]) = \sum_{[q]+[r]+\sum_\alpha [s_\alpha]=[l]} \tau^{(n)}([n];[q]) \times \prod_{\alpha\in[n]+[q]} \mathcal{S}(\alpha;[q_\alpha])\mathcal{G}_{n+q-2}([r]), \quad \text{(A1)}$$

where

$$\mathcal{G}_\sigma([r]) = \sum_{p=1}^{r} (-1)^p \frac{(\sigma+p)!}{\sigma!} \sum_{\substack{\mathcal{P}:[r] \\ into\ p}} \prod_{i=1}^{p} \mathcal{S}([r_i]). \quad \text{(A2)}$$

The summation in Eq. (A2) is over all distinct partitions of $[r]$ into p *nonempty* parts.

In complete analogy with the definition of the equilibrium modified Ursell function [see Eq. (B1)] one defines

$$\mathfrak{U}^{(n)}([n];[l]) = \sum_{[k]+[\bar{k}]=[l]} \mathcal{S}([n]+[h])\mathcal{G}_0([k]), \quad \text{(A3)}$$

where $\mathcal{G}_0$ is given by Eq. (A2) for $\sigma = 0$. Equation (A3) is precisely the solution of the recursion relation Eq. (11). This is quickly established by direct substitution if one notes the identity:

$$\sum_{[\mu]+[\nu]=[w]} \mathcal{S}([u])\mathcal{G}_0([v]) = \delta_{0,w}, \quad \text{(A4)}$$

where $\delta_{0,w}$ is the Kronecker symbol. This result, in turn, is already established by the combinatorial argument given at the end of Appendix A of paper I.

Substituting the expression for $\mathcal{S}([n]+[h])$ given by Eq. (A1) and rearranging the order of summations one finds that

$$\mathfrak{U}^{(n)}([n];[l]) = \sum_{[q]+[r]+\sum_\alpha [s_\alpha]=[l]} \tau^{(n)}([n];[q]) \prod_{\alpha\in[n]+[q]} \mathcal{S}(\alpha;[s_\alpha]) \sum_{[r_1]+[r_2]=[r]} \mathcal{G}_{n+q-2}([r_1])I_0([r_2]). \quad \text{(A5)}$$

But it has already been shown in Ref. 6 [i.e., Eq. (A5)] that

$$\sum_{[r_1]+[r_2]=[r]} I_{n+q-2}([r_1])I_0([r_2]) = \sum_{\sum_\alpha [r_\alpha]=[r]} \prod_{\alpha\in[r]} \mathcal{G}_0([r_\alpha]). \quad \text{(A6)}$$

Putting this result into Eq. (A5) and again rearranging the order of summations yields

$$\mathfrak{U}^{(n)}([n];[l]) = \sum_{[q]+\sum_\alpha [r_\alpha]=[l]} \tau^{(n)}([n];[q]) \times \prod_{\alpha\in[n]+[q]} \sum_{[s_\alpha]+[t_\alpha]=[r_\alpha]} \mathcal{S}(\alpha,[s_\alpha])I_0([t_\alpha]). \quad \text{(A7)}$$

But, according to Eq. (A3) the product of sums is just the product of $U^{(1)}(\alpha;[r_\alpha])$, so Eq. (A7) is identical with Eq. (10).

APPENDIX B: RELATION BETWEEN URSELL OPERATORS AND FUNCTIONS.

Equation (13), which expresses the result of operating with the Ursell operator on a product of equilibrium densities, can be established directly from Eq. (15).

[27] We wish to acknowledge private communication with E. G. D. Cohen on this point. See also Ref. 26.

[28] Such a resummed series might involve the modified $\mathcal{S}$-operator implicit in the work of K. Kawasaki and I. Oppenheim, Phys. Rev. **139**, A1763 (1965).

One has that the equilibrium Ursell function is given in terms of the equilibrium Boltzmann factors by

$$U^{(n)}([n];[l]) = \sum_{[h]+[k]=[l]} W([n]+[h])I_0([k]), \quad \text{(B1)}$$

where

$$W([m]) \equiv \exp\{-\beta V([m])\}. \quad \text{(B2)}$$

The sum I_0 corresponds to I_0 and is defined by

$$I_0([k]) = \sum_{p=1}^{k} (-1)^p p! \sum_{\substack{\mathcal{P}:[k] \\ \text{into } p}} \prod_{i=1}^{p} W([k_i]). \quad \text{(B3)}$$

Equation (B1) is the solution of the recursion relation, Eq. (4), and is the usual definition; namely, that $U^{(n)}$ is the sum of all distinct products of Boltzmann factors, the arguments of which partition $[n]+[l]$ into disjoint parts, one of which contains $[n]$.

Now, multiplying Eq. (B3) by ${}^{\circ}\phi([k])$, using the fact that the $[k_i]$ partition $[k]$, and then using Eq. (15), one finds that

$$\mathcal{J}_0([k])^{\circ}\phi([k]) = \mathcal{J}_0([k])^{\circ}\phi([k]), \quad \text{(B4)}$$

where $\mathcal{J}_0$ is the sum of $\mathcal{S}$-operators defined by Eq. (A2).

Similarly, multiplying Eq. (B1) by ${}^{\circ}\phi([n]+[l])$, one has that

$$\begin{aligned} &U^{(n)}([n];[l])^{\circ}\phi([n]+[l]) \\ &= \sum_{[h]+[k]=[l]} W([n]+[h])^{\circ}\phi([n]+[h])I_0([k])^{\circ}\phi([k]) \\ &= \sum_{[h]+[k]=[l]} \mathcal{S}([n]+[h])\mathcal{J}_0([k])^{\circ}\phi([n]+[l]). \end{aligned} \quad \text{(B5)}$$

Since $\mathfrak{U}^{(n)}$ is defined by Eq. (A3), Eq. (13) is established.

APPENDIX C: COMMUTATORS OF $\mathcal{S}$ AND $\mathfrak{U}^{(n)}$

The expression for the commutator of L_0 with an $\mathcal{S}$-operator [Eq. (50)] is discussed and then used to calculate the commutator of L_0 with $\mathfrak{U}^{(n)}$

Consider the operator whose limit is the $\mathcal{S}$-operator. According to Eq. (12) it is defined by

$$\mathcal{S}([m];\tau) = S([m];-\tau)S_0([m];\tau), \quad \text{(C1)}$$

and its time derivative is

$$\begin{aligned} &\dot{\mathcal{S}}([m];\tau) \\ &= -L([m])\mathcal{S}([m];\tau) + \mathcal{S}([m];\tau)L_0([m]). \end{aligned} \quad \text{(C2)}$$

Equation (50) is just the infinite time limit of this equation.

For example, for the case of two particles, $\mathcal{S}(12;\tau)$ operating on any two-particle function goes from the value of the function at $\tau = 0$ through a region of change about the time of closest approach to an asymptotic value equal to $\mathcal{S}(12)$ operating on the function. Thus, $\dot{\mathcal{S}}$ vanishes both when there is no collision at any earlier time and when a collision occurs at some finite time.

One wants to use the commutator of L_0 and $\mathcal{S}$ given by Eq. (50) to evaluate

$$\begin{aligned} &\gamma^{(n)}([n];[m]) \\ &= [L_0([n]+[m]);\mathfrak{U}^{(n)}([n];[m])], \end{aligned} \quad \text{(C3)}$$

where $\mathfrak{U}^{(n)}([n];[m])$ is the modified Ursell operator defined by Eq. (11) [or Eq. (A3)]. Equation (A3) yields directly for $\gamma^{(n)}$ that

$$\begin{aligned} &\gamma^{(n)}([n];[m]) \\ &= \sum_{[h]+[l]=[m]} \{-L'([n]+[h])\mathcal{S}([n]+[h]) \\ &\quad \times I_0([l]) + \mathcal{S}([n]+[h])[L_0([l]);I_0([l])]\}, \end{aligned} \quad \text{(C4)}$$

where in the first term the commutator of $\mathcal{S}$ has been evaluated using Eq. (50). Using Eq. (11) to re-express the $\mathcal{S}$-operators and rearranging summations one finds that

$$\begin{aligned} &\gamma^{(n)}([n];[m]) \\ &= -\sum_{[q]+[r]+[s]=[m]} L'([n]+[q]+[r]) \\ &\quad \times \mathfrak{U}^{(n)}([n];[q])\mathcal{S}([r])\mathcal{J}_0([s]) \\ &\quad + \sum_{[q]+[r]=[m]} \mathfrak{U}^{(n)}([n];[q]) \\ &\quad \times \sum_{[u]+[s]=[r]} \mathcal{S}([u])[L_0([v]);\mathcal{J}_0([v])]. \end{aligned} \quad \text{(C5)}$$

According to Eq. (A4)

$$\begin{aligned} &\sum_{[u]+[s]=[r]} \mathcal{S}([u])[L_0([v]);\mathcal{J}_0([v])] \\ &\quad = \sum_{[u]+[s]=[r]} L'([u])\mathcal{S}([u])\mathcal{J}_0([v]), \end{aligned} \quad \text{(C6)}$$

where $[L_0;\mathcal{S}]$ has been evaluated by using Eq. (50). Furthermore, one has by definition

$$L'([u]+[v]) = L'([u]) + L'([v]) + \sum_{\substack{i\in[u] \\ j\in[s]}} L'(ij). \quad \text{(C7)}$$

Using these two results in Eq. (C5), the second group of terms is seen to cancel so that rearranging orders of summation one finds that

$$\begin{aligned} &\gamma^{(n)}([n];[m]) \\ &= -\sum_{[q]+[v]=[m]} L'([n]+[q])\mathfrak{U}^{(n)}([n];[q]) \\ &\quad \times \sum_{[u]+[s]=[r]} \mathcal{S}([u])\mathcal{J}_0([v]) \\ &\quad - \sum_{\substack{j+[q]+[r]=[m] \\ i\in[n]+[q]}} L'(ij)\mathfrak{U}^{(n)}([n];[q]) \\ &\quad \times \sum_{[u]+[v]=[r]} \mathcal{S}(j+[u])\mathcal{J}_0([v]). \end{aligned} \quad \text{(C8)}$$

The first group of terms immediately simplifies according to Eq. (A4), while in the second group of terms one can use

$$\sum_{[u]+[v]=[r]} \mathcal{S}(j+[u])\mathcal{J}_0([v]) = \sum_{[h]+[l]=[r]} \mathfrak{U}^{(1)}(j;[h]) \times \sum_{[u]+[v]=[l]} \mathcal{S}([u])\mathcal{J}_0([v]) = \mathfrak{U}^{(1)}(j;[r]). \quad \text{(C9)}$$

Hence, one has established that

$$\gamma^{(n)}([n];[m]) = -L'([n]+[m])\mathfrak{U}^{(n)}([n];[m]) - \sum_{\substack{j+[q]+[r]=[m] \\ i\in[n]+[q]}} L'(ij)\mathfrak{U}^{(n)}([n];[q])\mathfrak{U}^{(1)}(j;[r]). \quad \text{(C10)}$$

For our purpose an alternate form of Eq. (C10), which explicitly separates the terms containing interaction operators only for members of $[m]$, is more useful. If one again uses Eq. (C7), one can write that

$$\gamma^{(n)}([n];[m]) = -(L'([n])+L'([m]))\mathfrak{U}^{(n)}([n];[m]) - \sum_{\substack{i\in[n] \\ j\in[m]}} L'(ij)\{\mathfrak{U}^{(n)}([n];[m]) + \sum_{[q]+[r]\in[m]-j} \mathfrak{U}^{(n)}([n];[q])\mathfrak{U}^{(1)}(j,[r])\} - \sum_{\substack{j+[q]+[r]=[m] \\ i\in[q]}} L'(ij)\mathfrak{U}^{(n)}([n];[q])\mathfrak{U}^{(1)}(j;[r]), \quad \text{(C11)}$$

where the second term of Eq. (C10) has been split into the two groups, one for which particle i is in $[n]$.

The operand of $L'(ij)$ in the second term is, in fact, just $\mathfrak{U}^{(n+1)}([n]+j;[m]-j)$. To see this, notice that, according to Eq. (A3), if one splits the sum into its two parts depending on which set contains particle j,

$$\mathfrak{U}^{(n)}([n];j+[m]) = \mathfrak{U}^{(n+1)}([n]+j;[m]) + \sum_{[q]+[r]=[m]} \mathfrak{U}^{(n)}([n];[q]) \times \sum_{[u]+[v]=[r]} \mathcal{S}([u])I_0(j+[v]). \quad \text{(C12)}$$

That the sum in the second term is precisely $-\mathfrak{U}^{(1)}(j;[r])$ follows immediately from Eq. (A4) with $[w] = j+[r]$ (so that $[l]$ is never empty).

Thus, one has that

$$\gamma^{(n)}([n];[m]) = -[L'([n])+L'([m])]\mathfrak{U}^{(n)}([n];[m]) - \sum_{\substack{i\in[n] \\ j\in[m]}} L'(ij)\mathfrak{U}^{(n+1)}([n]+j;[m]-j) - \sum_{\substack{i+j+[q]+[r]=[m] \\ i\neq j}} L'(ij)\mathfrak{U}^{(n)}([n];i+[q])\mathfrak{U}^{(1)}(j;[r]). \quad \text{(C13)}$$

APPENDIX D: COMMUTATOR OF $\tau^{(n)}$

To derive the expression for the commutator of L_0 and $\tau^{(n)}$ given by Eq. (48), one can again proceed in a way analogous to that in Appendix C. Thus, if one defines

$$B([n]+[q];[r]) = \sum_{\sum_\alpha [r_\alpha]=[r]} \prod_{\alpha\in[n]+[q]} \mathfrak{U}^{(1)}(\alpha;[r_\alpha]), \quad \text{(D1)}$$

Eq. (10) which was established in Appendix A can be written:

$$\mathfrak{U}^{(n)}([n];[l]) = \sum_{[q]+[r]=[l]} \tau^{(n)}([n];[q])B([n]+[q];[r]). \quad \text{(D2)}$$

To evaluate the commutator $\Gamma^{(n)}$ defined by

$$\Gamma^{(n)}([n];[m]) = [L_0([n]+[m]);\tau^{(n)}([n];[m])], \quad \text{(D3)}$$

one needs the solution to this recursion relation. The solution can be written

$$\tau^{(n)}([n];[m]) = \sum_{[h]+[l]=[m]} \mathfrak{U}^{(n)}([n];[h])B^{-1}([n]+[h];[l]). \quad \text{(D4)}$$

The quantity B^{-1} is a sum of products of the operators $\mathfrak{U}^{(1)}$ and, in fact, it will be shown elsewhere that it is just such that each summand in Eq. (D4) is a connected tree rooted on the set $[n]+[h]$. For the present, one only needs the property which B^{-1} must have if Eq. (D4) solves Eq. (D2); namely, that

$$\sum_{[u]+[v]=[r]} B([q];[u])B^{-1}([q]+[u];[v]) = \delta_{0,r}. \quad \text{(D5)}$$

According to Eq. (D4),

$$\Gamma^{(n)}([n];[m]) = \sum_{[h]+[l]=[m]} \{\gamma^{(n)}([n];[h])B^{-1}([n]+[h];[l]) + \mathfrak{U}^{(n)}([n];[h])[L_0;B^{-1}([n]+[h];[l])]\}. \quad \text{(D6)}$$

Since one does not want to evaluate the commutator of B^{-1}, first eliminate it in favor of the commutator of B by using the identity:

$$\sum_{[u]+[v]=[r]} B([u]+[q];[u]) \times [L_0;B^{-1}([n]+[q]+[u];[v])] = -\sum_{[u]+[v]=[r]} [L_0;B([n]+[q];[u])] \times B^{-1}([n]+[q]+[u];[v]). \quad \text{(D7)}$$

The result follows directly from Eq. (D5). Moreover, one has from the definition of $B([n] + [q]; [r])$ that

$$[L_0; B([n] + [q]; [r])] = \sum_{\substack{i \in [n]+[q] \\ [s]+[t]=[r]}} \gamma^{(1)}(i; [s]) B([n] + [q] - i; [t]). \quad \text{(D8)}$$

Now, if the second group of terms in Eq. (D6) is rewritten by using Eq. (D2) to evaluate $\mathfrak{U}^{(n)}$, the sums can be rearranged so that the left-hand side of Eq. (D7) appears. Thus, also using Eq. (D8), one can rewrite Eq. (D6) in the form:

$$\begin{aligned}\Gamma^{(n)}([n]; [m]) = \sum_{[h]+[l]=[m]} \{ & \gamma^{(n)}([n]; [h]) \\ & - \sum_{\substack{[q]+[r]+[s]=[h] \\ i \in [n]+[q]}} \tau^{(n)}([n]; [q]) \\ & \times \gamma^{(1)}(i; [r]) B([n] + [q] - i; [s]) \} \\ & \times B^{-1}([n] + [h]; [l]). \end{aligned} \quad \text{(D9)}$$

To continue the evaluation, divide $\gamma^{(n)}([n]; [h])$ into two parts, $\gamma_1^{(n)}$ and $\gamma_2^{(n)}$, $\gamma_1^{(n)}$ containing all the terms for which the interaction operators operate on members of $[n]$, and $\gamma_2^{(n)}$ the rest; that is, from Eq. (C13):

$$\begin{aligned}\gamma_1^{(n)}([n]; [h]) \equiv & -L'([n]) \mathfrak{U}^{(n)}([n]; [h]) \\ & - \sum_{\substack{i \in [n] \\ j \in [h]}} L'(ij) \mathfrak{U}^{(n+1)}([n] + j; [h] - j), \end{aligned} \quad \text{(D10)}$$

and

$$\begin{aligned}\gamma_2^{(n)}([n]; [h]) = & -L'([h]) \mathfrak{U}^{(n)}([n]; [h]) \\ & - \sum_{\substack{i+j+[q]+[r]=[h] \\ i \neq j}} L'(ij) \mathfrak{U}^{(n)}([n]; i + [q]) \mathfrak{U}^{(1)}(j; [r]). \end{aligned} \quad \text{(D11)}$$

If one now uses Eq. (D2) to evaluate $\mathfrak{U}^{(n)}$ and $\mathfrak{U}^{(n+1)}$ in Eq. (D10) one finds after rearranging sums that

$$\begin{aligned}\gamma_1^{(n)}([n]; [h]) = & - \sum_{[q]+[r]=[h]} \{ L'([n]) \tau^{(n)}([n]; [q]) \\ & + \sum_{\substack{i \in [n] \\ j \in [q]}} L'(ij) \tau^{(n+1)}([n] + j; [q] - j) \} \\ & \times B([n] + [q]; [r]), \end{aligned} \quad \text{(D12)}$$

and, in particular, since $L'(\alpha) = 0$,

$$\gamma_1^{(1)}(i; [r]) = - \sum_{\substack{[u]+[v]=[r] \\ j \in [u]}} L'(ij) \tau^{(2)}(y; [u] - j) B(i + [u]; [v]). \quad \text{(D13)}$$

Equation (D13) enables one to further evaluate the coefficient of B^{-1} in the second term of Eq. (D9). Thus, since by definition, if $[n_1] + [n_2] = [n]$, then

$$B([n]; [m]) = \sum_{[m_1]+[m_2]=[m]} B([n_1]; [m_1]) B([n_2]; [m_2]), \quad \text{(D14)}$$

one has

$$\begin{aligned} & - \sum_{\substack{[q]+[r]+[s]=[h] \\ i \in [n]+[q]}} \tau^{(n)}([n]; [q]) \\ & \quad \times \gamma_1^{(1)}(i; [r]) B([n] + [q] - i; [s]) \\ & = \sum_{[q]+[r]+[s]=[h]} \tau^{(n)}([n]; [q]) \sum_{\substack{i \in [n]+[q] \\ j \in [r]}} L'(ij) \\ & \quad \times \tau^{(2)}(ij; [r] - j) B([n] + [q] + [r]; [s]). \end{aligned} \quad \text{(D15)}$$

Taken together, Eqs. (D12) and (D15) yield an explicit expression for $\Gamma_1^{(n)}$, the part of $\Gamma^{(n)}$ arising from $\gamma_1^{(n)}$ and $\gamma_1^{(1)}$. Thus, substituting these two results into Eq. (D9), rearranging summations, and using Eq. (D5) yields

$$\Gamma_1^{(n)}([n]; [m]) = -L'([n]) \tau^{(n)}([n]; [m]) - \Delta\Gamma_1^{(n)}([n]; [m]), \quad \text{(D16)}$$

where $\Delta\Gamma_1^{(n)}$ has been defined by Eq. (47).

An expression for the remaining part of $\Gamma^{(n)}$ can be derived in a similar way. If one uses Eq. (D11) to evaluate $\Gamma_2^{(n)}$ as given by Eq. (D9), one sees directly that $\Gamma_2^{(n)}$ consists entirely of terms in which the interaction operator works only on members of $[m]$. This is the only result used in the text. For completeness we only quote the following result for $\Gamma_2^{(n)}$ in terms of previously defined quantities:

$$\Gamma_2^{(n)}([n]; [m]) = -L'([m]) \tau^{(n)}([n]; [m]) - \Delta\Gamma_2^{(n)}([n]; [m]), \quad \text{(D17)}$$

where

$$\begin{aligned}\Delta\Gamma_2^{(n)}([n]; [m]) = & \sum_{\substack{i+j+[q]+[r]=[m] \\ i \neq j}} L'(ij) \{ \tau^{(n)}([n]; [q] + i) \\ & \times [\sum_{\alpha \in [n]+[q]} \Delta\tau^{(2)}(\alpha j; [r]) + \tau^{(2)}(ij; [r])] \\ & + \tau^{(n)}([n]; [q]) \\ & \times \sum_{[u]+[v]=[r]} \sum_{\substack{\alpha, \beta \in [n]+[q] \\ \alpha < \beta}} \Delta\tau^{(2)}(\alpha i; [u]) \Delta\tau^{(2)}(\beta j; [v]) \}, \end{aligned} \quad \text{(D18)}$$

and, in turn,

$$\Delta\tau^{(2)}(ij; [m]) \equiv \tau^{(2)}(ij; [m]) - \delta_{0,m}. \quad \text{(D19)}$$

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 7, NUMBER 5 MAY 1966

On Irreducible Tensor Operators for Finite Groups

P. RUDRA
Saha Institute of Nuclear Physics, Calcutta, India
(Received 14 September 1965)

A method is derived for obtaining the tensor covariants of the finite groups, belonging to any given factor system, starting from an arbitrary tensor of rank r in a N-dimensional vector space. The transformation properties of the irreducible tensor operators for the projective representations of the group have been discussed. The Wigner–Eckart theorem for these representations has also been studied.

1. INTRODUCTION

IN a previous paper[1] we have studied the projective representations of finite group belonging to any factor system. There we found out the algebraic method of obtaining the characters of the inequivalent irreducible projective representations of finite groups. We also deduced, among others, the transformation properties and the projection operators for the basis functions, the Kronecker (inner) direct product representation and the Clebsch–Gordan coefficients.

In this paper, we study the transformation properties of tensor operators belonging to irreducible projective representations of any particular factor system. We also derive a method for obtaining the tensor basis of the irreducible projective representations. Wigner–Eckart theorem can again be written as the product of the Clebsch–Gordan coefficient and a reduced matrix element, the expression for the latter term, differing from that in the case of vector representation, by a multiplicative term which is a function of the factors.

Sirotin[2] and Smith[3] have studied the tensor invariants for anisotropic tensors. Smith has also discussed different methods for obtaining tensor invariants. Our method of obtaining the tensor invariants (forming the basis of the identical representation) and tensor covariants (forming bases of other irreducible representations) depends on the generalization of the concept of projection operators for tensor operators belonging to the projective representations.

Erdös'[4] work of finding the tensor invariants for crystals runs in a line similar to ours, as applied to the irreducible vector representations.

The general properties of the factors of the projective representations have been deduced in Ref. 1. We shall copiously use these results. We also want to point out a change of notation from that used in Ref. 1. There Γ_μ denotes the μth irreducible projective representation; here $D^{(\mu)}$ denotes the same, μ except being used as a superscript.

2. TENSOR OPERATORS FOR FINITE GROUPS: TRANSFORMATION RULES FOR PROJECTIVE REPRESENTATIONS

Tensor operators have been used in quantum mechanics for a long time in the form of spherical tensors,[5] which transform as different irreducible representations of the full rotation group. Similarly, tensor operators may be defined[6–9] as belonging to different irreducible representations of the finite group.

If we define the Wigner operators[9] O_R for all $R \in G$, where G is the group of order g, then the transformation properties of the irreducible tensor operator T_i^μ forming the ith basis of the μth irreducible projective representation $D^{(\mu)}$ of the group G will be as follows.[9] If the coordinate basis in the vector space V is changed by R, i.e.,

$$\mathbf{u}' = \mathbf{u}R, \tag{1}$$

then any function $\psi(\mathbf{u})$ in V will become[1]

$$\psi' = O_R\psi,$$

where

$$\psi'(\mathbf{u}) = O_R\psi(\mathbf{u}) = [\prod_{P\in G} \omega_{R,P}\omega_{P,R}]^{1/2g}\psi(\mathbf{u}R^{-1}), \tag{2}$$

[1] P. Rudra, J. Math. Phys. 6, 1273 (1965). We shall refer to this paper as I.
[2] Iu. I. Sirotin, Dokl. Akad. Nauk SSSR 133, 321 (1960) [English transl.: Soviet Phys.—Doklady 5, 774 (1961)].
[3] G. F. Smith, J. Math. Phys. 5, 1612 (1964).
[4] P. Erdös, Helv. Phys. Acta 37, 493 (1964).
[5] M. E. Rose, *Elementary Theory of Angular Momentum* (John Wiley & Sons, Inc., New York, 1957).
[6] U. Fano and G. Racah, *Irreducible Tensorial Set* (Academic Press, Inc., New York, 1959).
[7] H. Weyl, *Classical Groups* (Princeton University Press, Princeton, New Jersey, 1946).
[8] H. Weyl, *The Theory of Groups and Quantum Mechanics* (Dover Publications, Inc., New York) (English transl.)
[9] E. P. Wigner, *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra* (Academic Press, Inc., New York, 1959) (English transl.).

and the tensor operator T_i^μ will transform as

$$T_i^{\mu\prime} = O_R T_i^\mu O_{R^{-1}} = \sum_j \frac{1}{\omega_{R,R^{-1}}} D_{ij}^{(\mu)}(R^{-1}) T_j^\mu, \quad (3)$$

where $\omega_{P,Q}$ for $P, Q \in G$ are the factors of the projective representation.[1,10] The rank of the tensor is suppressed in the notation.

We now show that, with respect to the tensor bases T_i^μ's ,the O_R's behave as the projective representation of G belonging to the complex conjugate of the factor systems to which $D^{(\mu)}$'s also belong, i.e., to $\omega^*_{P,Q} = 1/\omega_{P,Q}$ for all $P, Q \in G$. Thus

$$O_S O_R T_i^\mu O_{R^{-1}} O_{S^{-1}} = \omega^*_{S,R} O_{SR} T_i^\mu O_{R^{-1}S^{-1}} \quad (4)$$

for all μ and i.

Proof.

$$O_S O_R T_i^\mu O_{R^{-1}} O_{S^{-1}}$$

$$= \frac{1}{\omega_{R,R^{-1}}} \sum_j D_{ij}^{(\mu)}(R^{-1}) O_S T_j^\mu O_{S^{-1}}$$

$$= \frac{1}{\omega_{R,R^{-1}}\omega_{S,S^{-1}}} \sum_{j,k} D_{ij}^{(\mu)}(R^{-1})\, D_{jk}^{(\mu)}(S^{-1}) T_k^\mu$$

$$= \frac{\omega_{R^{-1},S^{-1}}\omega_{SR,R^{-1}S^{-1}}}{\omega_{R,R^{-1}}\omega_{S,S^{-1}}} \cdot \frac{1}{\omega_{SR,R^{-1}S^{-1}}} \sum_k D_{ik}^{(\mu)}(R^{-1}S^{-1}) T_k^\mu$$

$$= \omega^*_{S,R} O_{SR} T_i^\mu O_{R^{-1}S^{-1}},$$

since

$$\omega_{R^{-1},S^{-1}}\omega_{SR,R^{-1}S^{-1}}/\omega_{R,R^{-1}}\omega_{S,S^{-1}}$$

$$= \omega_{S,R}\omega_{SR,R^{-1}}\omega_{S,S^{-1}}/\omega_{S,R}\omega_{R,R^{-1}}\omega_{S,S^{-1}}$$

$$= \omega_{S,E}\omega_{R,R^{-1}}/\omega_{S,R}\omega_{R,R^{-1}} = 1/\omega_{S,R} = \omega^*_{S,R}.$$

3. PROJECTION OPERATORS FOR TENSOR OPERATORS

In this section we show how to obtain the irreducible tensor basis T_i^μ from any tensor T of rank r in an N-dimensional vector space V. Our method is a generalization of Wigner projection operator $P_i^{(\mu)}$, which, operating on any function ψ, gives the ith functional basis of the μth irreducible representation.[1,9,10]

We define a corresponding projection operator $\mathsf{P}_i^{(\mu)}$ whose action on any arbitrary tensor T gives a new tensor T_i^μ which, as is shown, transforms according to Eq. (3):

$$T_i^\mu = \mathsf{P}_i^{(\mu)}\mathsf{T} = \frac{n_\mu}{g} \sum_{R\in G} D_{ij}^{(\mu)}(R) O_R \mathsf{T} O_{R^{-1}}, \quad (5)$$

where n_μ is the dimension of the irreducible projective representation $D^{(\mu)}$. We first explain the meaning of $O_R\mathsf{T}O_{R^{-1}}$. The components of the tensor T of rank r in a N-Dimensional vector space V may be looked upon as spanning the space V^r of dimension N^r. This space will in general be reducible under the operations of the group G. The components of $O_R\mathsf{T}O_{R^{-1}}$ may be interpreted as the transformed components of T under a basis change in V by the operation $R \in G$. With this interpretation, Eq. (4) is also valid for T. Now we show that T_i^μ's obtained from (5) form the ith basis of the μth irreducible projective representation. We have to utilize, in this process, the Eq. (4) and the basic property of the factors of the projective representation.[1]

Proof:

$$O_S T_i^\mu O_{S^{-1}} = \frac{n_\mu}{g} \sum_{R\in G} D_{ij}^{(\mu)}(R) O_S O_R \mathsf{T} O_{R^{-1}} O_{S^{-1}}$$

$$= \frac{n_\mu}{g} \sum_{R\in G} \omega^*_{S,R}\, D_{ij}^{(\mu)}(S^{-1}SR) O_{SR} \mathsf{T} O_{R^{-1}S^{-1}}$$

$$= \sum_k D_{ik}^{(\mu)}(S^{-1}) \frac{n_\mu}{g} \cdot \sum_{R\in G} \frac{1}{\omega_{S,R}\omega_{S^{-1},SR}} D_{kj}^{(\mu)}(SR) \times O_{SR}\mathsf{T}O_{R^{-1}S^{-1}}$$

$$= \sum_k \frac{1}{\omega_{S,S^{-1}}} D_{ik}^{(\mu)}(S^{-1}) \frac{n_\mu}{g} \sum_{SR\in G} D_{kj}^{(\mu)}(SR) \times O_{SR}\mathsf{T}O_{R^{-1}S^{-1}}$$

$$= \sum_k \frac{1}{\omega_{S,S^{-1}}} D_{ik}^{(\mu)}(S^{-1}) T_k^\mu.$$

4. WIGNER-ECKART THEOREM FOR PROJECTIVE REPRESENTATIONS

In Ref. 1, we have shown that, under a coordinate basis transformation in the vector space V, the ith functional basis of the μth irreducible projective representation transforms as

$$|O_R\psi_i^\mu\rangle = O_R\,|\psi_i^\mu\rangle = \sum_j D_{ji}^{(\mu)}(R)\,|\psi_j^\mu\rangle. \quad (6)$$

The basis functions $\langle\psi_k^\mu|$ of the dual space V^* will transform as

$$\langle O_R\psi_k^\mu| = \sum_l \frac{1}{\omega_{R^{-1},R}} D_{kl}^{(\mu)}(R^{-1})\langle\psi_l^\mu|, \quad (7)$$

so that the orthonormalization of the functions is maintained:

$$\langle O_R\psi_k^\mu\,|\,O_R\psi_i^\mu\rangle = \sum_{j,l} \frac{1}{\omega_{R^{-1},R}} D_{kl}^{(\mu)}(R^{-1})\, D_{ji}^{(\mu)}(R)\langle\psi_l^\mu\,|\,\psi_j^\mu\rangle$$

$$= \sum_j \frac{1}{\omega_{R^{-1},R}} D_{kj}^{(\mu)}(R^{-1})\, D_{ji}^{(\mu)}(R)$$

$$= D_{ki}^{(\mu)}(E) = \delta_{ki}.$$

[10] M. Hamermesh, *Group Theory and its Application to Physical Problems* (Addison-Wesley Publishing Company, Inc. Reading, Massachusetts, 1962).

Combining Eqs. (3), (6), and (7) we get for the matrix element $\langle \tau'\alpha'i'|\ T^{\mu}_{m}\ |\tau\alpha i|$ (where α', μ, α are the irreducible projective representations, i', m, i are the components, and τ', τ are any other parameters characterizing the states):

$$\langle \tau'\alpha'i'|\ T^{\mu}_{m}\ |\tau\alpha i\rangle = \langle O_R(\tau'\alpha'i')|\ O_R T^{\mu}_{m} O_{R^{-1}}\ |O_R(\tau\alpha i)\rangle$$

$$= \sum_{j'nj} \frac{1}{\omega_{R,R^{-1}}} D^{(\tau'\alpha')}_{i'j'}(R^{-1}) \frac{1}{\omega_{R,R^{-1}}} D^{(\mu)}_{mn}(R^{-1})\ D^{(\tau\alpha)}_{ji}(R) \times \langle \tau'\alpha'j'|\ T^{\mu}_{n}\ |\tau\alpha j\rangle.$$

Using Eqs. (21) and (30) of Ref. 1 for the Kronecker inner direct product and the Clebsch–Gordan coefficients, we have

$$\langle \tau'\alpha'i'|\ T^{\mu}_{m}\ |\tau\alpha i\rangle \cdot \sum_{R\in G} \omega_{R,R^{-1}} [\prod_{P\in G} \omega_{R,P}\omega_{P,R}]^{-1/2g}$$

$$= \frac{g}{n_\alpha} \langle \tau'\alpha'i',\ \mu m\ |\ \tau\alpha i\rangle \times \sum_{j'nj} \langle \tau'\alpha'j'|\ T^{\mu}_{n}\ |\tau\alpha j\rangle\langle \tau\alpha j\ |\ \tau'\alpha'j',\ \mu n\rangle.$$

Now, since

$$\omega_{R,R^{-1}}\omega_{R^{-1},R}/\omega_{R,P}\omega_{P,R}$$

$$= \omega_{P,E}\omega_{R,R^{-1}}\omega_{R^{-1},R}\omega_{E,P}/\omega_{R,P}\omega_{P,R}$$

$$= \omega_{P,R}\omega_{PR,R^{-1}}\omega_{R^{-1},RP}\omega_{R,P}/\omega_{P,R}\omega_{R,P}$$

$$= \omega_{PR,R^{-1}}\omega_{R^{-1},RP}$$

for arbitrary $P \in G$, we obtain

$$\sum_{R\in G} \omega_{R,R^{-1}} [\prod_{P\in G} \omega_{R,P}\omega_{P,R}]^{-1/2g}$$

$$= \sum_{R\in G} \left[\frac{(\omega_{R,R^{-1}}\omega_{R^{-1},R})^{g}}{\prod_{P\in G} \omega_{R,P}\omega_{P,R}}\right]^{1/2g}$$

$$= \sum_{R\in G} [\prod_{P\in G} \omega_{PR,R^{-1}}\omega_{R^{-1},RP}]^{1/2g}$$

$$= \sum_{R\in G} [\prod_{P\in G} \omega_{R,P}\omega_{P,R}]^{1/2g}.$$

Thus we can write

$$\langle \tau'\alpha'i'|\ T^{\mu}_{m}\ |\tau\alpha i\rangle = \frac{\langle \tau'\alpha'||\ T^{\mu}\ ||\tau\alpha\rangle}{n_\alpha} \cdot \langle \tau'\alpha'i',\ \mu m\ |\ \tau\alpha i\rangle, \tag{8}$$

where the reduced matrix element

$$\langle \tau'\alpha'||\ T^{\mu}\ ||\tau\alpha\rangle = (g/\sum_{R\in G} [\prod_{P\in G} \omega_{R,P}\omega_{P,R}]^{1/2g}) \times \sum_{j'nj} \langle \tau'\alpha'j'|\ T^{\mu}_{n}\ |\tau\alpha j\rangle\langle \tau\alpha j\ |\ \tau'\alpha'j',\ \mu n\rangle \tag{9}$$

does not depend on the components i', m, and i, and differs from the similar expression for the vector representations only in the multiplicative term

$$g/\sum_{R\in G} [\prod_{P\in G} \omega_{R,P}\omega_{P,R}]^{1/2g}.$$

Here we shall like to point out that Eqs. (8) and (9) are valid only for simply reducible groups,[1,10] i.e., groups which satisfy the following two conditions:

(a) All classes are ambivalent (i.e., every element of the group is equivalent to its inverse), and

(b) the Kronecker inner direct product of any two irreducible representations of the group contains no irreducible representations more than once.

The generalization[11,12] of Wigner–Eckart theorem for non-simply reducible groups may be approached in the same line as is done by Ginibre[11] for simple Lie groups.

5. CONCLUSION

We have thus obtained the transformation properties of irreducible tensor operators for the projective representation of finite groups, belonging to any factor system. A method is also deduced for obtaining the tensor invariants and covariants of finite groups for these representations and the form of the Wigner–Eckart theorem is given in this case. All these formulas can be used for vector representations if we note that in this case $\omega_{P,Q} = 1$ for all $P, Q \in G$. The importance of the projective representations lies in the fact that Wigner antiunitary representations[9] are actually projective representations, belonging to a particular factor system.

---

[11] J. Ginibre, J. Math. Phys. 4, 720 (1963).

[12] G. E. Baird and L. C. Biedenharn, J. Math. Phys. 5, 1730 (1964).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 7, NUMBER 5 MAY 1966

Improved Method for Quantum Mechanical Three-Body Problems. II. Repulsive Potentials

LEONARD EYGES
Air Force Cambridge Research Laboratories, Office of Aerospace Research, Bedford, Massachusetts
(Received 23 September 1965)

We extend a method previously presented for the three-body problem with attractive interparticle potentials to the case of repulsive interparticle potentials, with periodic boundary conditions on each particle. As before, we decompose the wave function into three parts, and from the Schrödinger equation write equations for these parts. We observe that, in one-dimension and for δ-function potentials, these equations can be easily solved numerically, and we present these solutions for potential "strengths" ranging from zero to infinity. We then discuss a more general method for solving these equations, which method involves expansions in a certain set of two-body functions. This general method is not necessarily limited to either one dimension or δ-function potentials, but, as a check on it, we do apply it to that case and get good agreement with the previous numerical results. As a further exploration of the method, we apply it to square well potentials in one dimension. In an Appendix we discuss the set of two-body functions that we use.

INTRODUCTION

IN previous papers[1] we have considered the problem of finding the ground-state energy and eigenfunction for three identical particles bound by pairwise attractive potentials. It is then natural to try to extend the method of solution to repulsive potentials and it is this problem we consider here. For purely repulsive potentials there is, of course, no bound state, so to define the problem we must impose some kind of external boundary condition; we take this to be the conventional periodic boundary condition. The method is presented for one-dimensional problems, but there is no intrinsic limitation to one dimension in it. For simplicity, and to explain the method without extraneous details we have dealt mainly, but not exclusively, with δ-function repulsive potentials, but again there is no limitation in principle to these; thus in the last section we discuss the case of square potentials. These one-dimensional problems are still models in the sense that there are no physical systems to which the results apply directly; rather, we are interested in them as applications of our general method. We have, therefore, not presented profuse numerical results for them, but have gone only to the point required to show that the method works in their case. We have also tried to highlight the new features that repulsive potentials introduce and to discuss them in terms that can be transcribed to three-dimensional problems. One such feature, which is worth consideration in its own right, is the introduction of a complete set of two-body functions which serve as a set of basis functions for expansion in the three-body problem.

[1] L. Eyges, Phys. Rev. **121**, 1744 (1961); J. Math. Phys. **6**, 1320 (1965).

The outline of the paper is as follows. In Sec. I we present and discuss some general equations for the problem. In Sec. II we apply these to δ-function potentials and observe that they then reduce to a particularly simple form which can be solved by easy numerical calculation. We do this to provide a useful test case against which we can later check the results got from our more general method. In Sec. III this general method is presented. In Sec. IV we go back and, as a check, solve the δ-function case with it. The results agree well with those derived by the special method of Sec. II. Finally, in the same section we briefly treat the case of square well repulsive potentials.

I. GENERAL EQUATIONS

We consider the one-dimensional problem of three particles, identical and of mass m, with coordinates x_1, x_2, x_3, and with pairwise interaction forces, i.e., with total potential energy

$$V(|x_1 - x_2|) + V(|x_1 - x_3|) + V(|x_2 - x_3|). \quad (1)$$

The ground-state energy of the system is $\mathcal{E}$. With

$$E = \frac{2m\mathcal{E}}{\hbar^2}, \qquad v = \frac{2mV}{\hbar^2}, \qquad v_{ij} = v(|x_i - x_j|),$$

the Schrödinger equation for the wave function is

$$(-\nabla^2 + v_{12} + v_{13} + v_{23})\Psi = E\Psi, \quad (2)$$

where

$$\nabla^2 = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}.$$

As we have explained previously[1] and hence do not expand upon here, it is part of our method to work

simultaneously with three different sets of coordinates. These are redundant in that any single set would suffice to characterize the problem; nonetheless, it is very advantageous to use all three. We have already seen[1] some of the advantages of these sets for the problem with attractive potentials. It is interesting to note that Fadeev[2] has also found similar sets of coordinates advantageous in scattering problems and has used them to derive many results of formal scattering theory for three-body systems.

A typical set (x_{12}, y_3, X) is defined by

$$x_{12} = x_1 - x_2, \qquad y_3 = 2x_3 - x_1 - x_2, \qquad X = x_1 + x_2 + x_3, \tag{3}$$

with sets (x_{13}, y_2, X) and (x_{23}, y_1, X) defined analogously. The Laplacian becomes typically in these new coordinates

$$\nabla^2 = 2\frac{\partial^2}{\partial x_{12}^2} + 6\frac{\partial^2}{\partial y_3^2} + 3\frac{\partial^2}{\partial X^2}.$$

Following our previous work we now define a function[3] $\psi_{12}(x_{12}, y_3)$ which satisfies

$$-\nabla^2\psi_{12} + v_{12}\Psi = E\psi_{12}, \tag{4}$$

with similar equations for ψ_{13} and ψ_{23}. We sometimes refer to ψ_{12} as a "two-body orbital" for reasons we have discussed elsewhere.[1] ψ_{12} is related to Ψ by

$$\Psi = \psi_{12} + \psi_{13} + \psi_{23}, \tag{5}$$

since, if we add Eq. (4) to its two counterparts in the other sets of coordinates, we get back Eq. (2) for Ψ, using this last relation. Thus Eqs. (4) and (5) are equivalent to the Schrödinger equation for the problem. Moreover, as we have discussed previously, we know that for the ground state ψ_{12} will be the same function of its coordinates that the other two functions are of theirs, so that we have only a single function ψ_{12} to determine. Thus, it is sufficient to solve Eq. (4), which contains only the single potential v_{12} and which is therefore appreciably simpler than the complete Schrödinger equation. This advantage is especially marked if v_{12} is a potential of finite range, i.e., one for which $v_{12}(x) = 0$ for $x > x_0$. The solution of Eq. (4) for $x > x_0$ is then the free particle solution. To be sure, if we considered the ordinary Schrödinger equation, in which $v_{12} + v_{13} + v_{23}$ entered, there would be a region of x_1, x_2, x_3 space for which the potential is zero and hence for which the solution would also be a free particle solution. But this region, being a function of x_1, x_2, x_3, would be much more complicated to visualize and deal with in general than in the case above.

The boundary condition we want to use is that Ψ be periodic, of period L, in x_1, x_2, x_3 separately. This condition will be satisfied if each of the orbitals ψ_{12}, ψ_{13}, ψ_{23} is similarly periodic, and we now assume this. Let us look then at ψ_{12} which, with this assumption, can be expanded in a Fourier series,[4]

$$\psi_{12} = \sum_{s_1,s_2,s_3} H(s_1, s_2, s_3) \times \exp\left[\frac{2\pi i}{L}(s_1x_1 + s_2x_2 + s_3x_3)\right]. \tag{6}$$

If we use the transformations

$$x_1 = \tfrac{1}{3}X + \tfrac{1}{2}x_{12} - \tfrac{1}{6}y_3,$$
$$x_2 = \tfrac{1}{3}X - \tfrac{1}{2}x_{12} - \tfrac{1}{6}y_3,$$
$$x_3 = \tfrac{1}{3}(X + y_3),$$

this can be written in the form

$$\sum_{s_1,s_2,s_3} H(s_1, s_2, s_3) \exp\left\{\frac{2\pi i}{L}\left[\frac{X}{3}(s_1 + s_2 + s_3) + \frac{x_{12}}{2}(s_1 - s_2) + \frac{y_3}{6}(-s_1 - s_2 + 2s_3)\right]\right\}. \tag{7}$$

The factor $\exp[(2\pi i/L)\frac{1}{3}X(s_1 + s_2 + s_3)]$ in this expression refers to the center-of-mass motion. For the ground-state wavefunction with which we deal we can assume that the center of mass is "at rest," i.e., that $s_1 + s_2 + s_3 = 0$. If we put this condition into (7), we conclude that the most general function periodic in x_1, x_2, x_3 and with center of mass at rest is [with $H(s_1, s_2, -s_1 - s_2) \to G(s_1, s_2)$]

$$\sum_{s_1,s_2} G(s_1, s_2) \exp\left\{\frac{2\pi i}{L}\left[\frac{x_{12}}{2}(s_1 - s_2) - \frac{y_3}{2}(s_1 + s_2)\right]\right\}. \tag{8}$$

Note that this is periodic in x_{12} and y_3 not with period L, but with period $2L$. If we write $s_1 - s_2 = n_1$, $s_1 + s_2 = -n_2$, and let $G(s_1, s_2) \to C(n_1, n_2)$, we can now assume that the general expression for ψ_{12} is of the form[5]

$$\psi_{12}(x_{12}, y_3) = \sum C(n_1, n_2) \exp\left[\frac{2\pi i}{2L}(n_1x_{12} + n_2y_3)\right]. \tag{9}$$

[2] L. D. Fadeev, Zh. Eksperim. i Teor. Fiz. **39**, 1459 (1960) [English transl.: Soviet Phys.—JETP **12**, 1014, 1961]; Dokl. Akad. Nauk USSR **138**, 565 (1961) [English transl.: Soviet Phys.—Dokl. **6**, 384 (1961)]; Dokl. Akad. Nauk USSR **145**, 301 (1962) [English transl.: Soviet Phys.—Dokl. **7**, 600 (1963)].

[3] Henceforth, we neglect a trivial dependence of our functions on the center-of-mass coordinate X.

[4] A sum simply written as $\sum_l$ will mean $\sum_{l=-\infty}^{\infty}$ unless otherwise indicated, and analogously for double or triple sums.

[5] It will be sufficient to assume that the function $C(n_1, n_2)$ is real.

We have deliberately omitted from this last sum the range of values of the summation variables n_1 and n_2 since this requires a little discussion. At first sight it might seem that n_1 and n_2 take on arbitrary independent integral values. This is not quite true, however, since for arbitrary s_1 and s_2, if $s_1 - s_2$ is even (odd) then $s_1 + s_2$ is even (odd). We must then put in the condition in the expansion (9) that n_1 and n_2 be both even or both odd. Instead of trying to keep this even–odd condition always in mind, it will sometimes be more convenient to deal with it explicitly in the following way. We define $\Delta(n, m)$ by

$$\Delta(n, m) = \begin{cases} 1; & n, m \text{ both even or both odd,} \\ 0; & \text{otherwise.} \end{cases} \tag{10}$$

Then we can write (9) formally as a sum over all integral n_1 and n_2

$$\psi_{12}(x_{12}, y_3) = \sum_{n_1, n_2} C(n_1, n_2)\, \Delta(n_1, n_2) \times \exp\left[\frac{2\pi i}{2L}(n_1 x_{12} + n_2 y_3)\right]. \tag{11}$$

With this form for ψ_{12}, we now turn to the solution of Eq. (4).

The potential energy $v(x_{12})$ is defined for x_1 and x_2 between zero and L, say, and we seek a wavefunction which is periodic in these variables. Often when one has this periodicity condition on a wavefunction it is because the potential itself is periodic, as in the canonical solid state physics problem of the solution of the Schrödinger equation for a periodic lattice. It is convenient then to adopt the fiction that the potential $v(x_{12})$ is periodic, i.e., to replace $v(x_{12})$ by its periodic counterpart $v_p(x_{12})$ defined by

$$v_p(x_{12}) = \sum_{n=-\infty}^{\infty} v(x_{12} + nL), \tag{12}$$

so that we are closer to this problem. This is, of course, merely an artifice; it does *not* change the potential inside the basic domain of x_{12}, but merely adds "replicas" of this potential outside that domain. Nonetheless, working with v_p, we can use Fourier series

$$v_p(x_{12}) = \sum_l U(l) \exp\left[i2\pi l x_{12}/L\right]. \tag{13}$$

If then we put (11) and (13) into Eq. (4) we get

$$\sum_{n_1', n_2'} \left[E - \left(\frac{\pi}{L}\right)^2 (2n_1'^2 + 6n_2'^2)\right] C(n_1', n_2')\, \Delta(n_1', n_2') \times \exp\left[\frac{i\pi}{L}(n_1' x_{12} + n_2' y_3)\right]$$

$$= \sum_{l, n_1', n_2'} C(n_1', n_2') U(l)\, \Delta(n_1', n_2') \times \left\{\exp\left[\frac{i2\pi l x_{12}}{L} + \frac{i\pi}{L}(n_1' x_{12} + n_2' y_3)\right] + \exp\left[\frac{i2\pi l x_{12}}{L} + \frac{i\pi}{L}(n_1' x_{13} + n_2' y_2)\right] + \exp\left[\frac{i2\pi l x_{12}}{L} + \frac{i\pi}{L}(n_1' x_{23} + n_2' y_1)\right]\right\}.$$

We transform the second and third exponents in curly brackets using

$$x_{13} = -\tfrac{1}{2}(y_3 - x_{12}), \qquad y_2 = -\tfrac{1}{2}(3x_{12} + y_3),$$
$$x_{23} = -\tfrac{1}{2}(y_3 + x_{12}), \qquad y_1 = \tfrac{1}{2}(3x_{12} - y_3),$$

and then multiply by $\exp\left[-(2\pi i/2L)(n_1 x_{12} + n_2 y_3)\right]$ and integrate from zero to $2L$ to get

$$[E - 2(\pi/L)^2(n_1^2 + 3n_2^2)]C(n_1, n_2)\, \Delta(n_1, n_2) = \sum_{n_1'} \left\{U\left(\frac{n_1 - n_1'}{2}\right) C(n_1', n_2)\, \Delta(n_1', n_2) + U\left(\frac{n_1 - 2n_1' - 3n_2}{2}\right) C(n_1', -n_1' - 2n_2) \times \Delta(n_1', -n_1' - 2n_2) + U\left(\frac{n_1 + 2n_1' + 3n_2}{2}\right) C(n_1', -n_1' - 2n_2) \times \Delta(n_1', -n_1' - 2n_2)\right\}. \tag{14}$$

This is the basic equation, whose solution we now investigate.

II. SPECIAL METHOD OF SOLUTION FOR ONE-DIMENSIONAL δ FUNCTIONS

We first want to consider Eq. (14) for the special case of one-dimensional delta-function potentials. We are able to solve it by a special method, that is, one applicable only to these potentials; the solution, besides being interesting in its own right, then constitutes a very useful test case for the more general method we develop later.

For the δ-function potential,[6] $v(x) = t\delta(x)$, we have then

$$U(l) = t/L$$

and Eq. (14) becomes[7]

[6] Note that our definition of the δ-function strength differs by a factor of 2 from that of Lieb and Liniger, Ref. 10.

[7] The function $\Delta(n_1, n_2)$ is really superfluous except when it appears within a summation sign, so we drop it from the left-hand side of this equation.

TABLE I. Ground-state eigenvalues E' and corresponding eigenvector $F(n)$ for one dimensional three-body problem with repulsive δ-function interactions of strength t, and with periodic boundary conditions over length L. $F(0)$ is arbitrarily normalized to minus one: $F(0) = -1$. The figures (1), (2), (3) in the boxes for $F(n)$ mean that the number there given must be multiplied by 10^{-1}, 10^{-2}, 10^{-3}, respectively.

tL	E'	$F(1)$	$F(2)$	$F(3)$	$F(4)$	$F(5)$	$F(6)$	$F(7)$
2	0.2809	4.637(2)	1.097(2)	4.829(3)	2.707(3)	1.730(3)	1.200(3)	8.816(4)
5	0.6327	1.025(1)	2.246(2)	9.762(3)	5.449(3)	3.476(3)	2.410(3)	1.769(3)
10	1.091	1.71(1)	3.34(2)	1.43(2)	7.92(3)	5.04(3)	3.49(3)	2.56(3)
20	1.722	2.56(1)	4.09(2)	1.70(2)	9.39(2)	5.96(3)	4.12(3)	3.02(3)
40	2.426	3.40(1)	3.94(2)	1.60(2)	8.71(3)	5.50(3)	3.80(3)	2.79(3)
100	3.197	4.22(1)	2.64(2)	1.03(2)	5.50(3)	3.41(3)	2.30(3)	1.64(3)
200	3.561	4.58(1)	1.62(2)	6.24(3)	3.40(3)	2.18(3)	1.55(3)	1.18(3)

$$\left[E - 2\left(\frac{\pi}{L}\right)^2(n_1^2 + 3n_2^2)\right]C(n_1, n_2)$$

$$= \frac{t}{L}\sum_{n_1'}[C(n_1', n_2)\,\Delta(n_1', n_2)$$

$$+ 2C(n_1', -n_1' - 2n_2)\,\Delta(n_1', -n_1' - 2n_2)]. \qquad (15)$$

Since n_1 enters this equation explicitly only through its square, it is clear that $C(n_1, n_2) = C(-n_1, n_2)$. Moreover, under the transformation $n_2 \to -n_2$, the equation satisfied by $C(n_1, -n_2)$ turns out to be the same as that satisfied by $C(n_1, n_2)$, whence we conclude that $C(n_1, n_2) = C(n_1, -n_2)$. Now we observe that we can reduce this equation to one for a function, call it F, of the single variable[8] n_2 by writing

$$C(n_1, n_2) = \frac{F(n_2)}{E - 2(\pi/L)^2(n_1^2 + 3n_2^2)}. \qquad (16)$$

If we put (16) into (15) we get

$$F(n_2) = F(n_2)\frac{t}{L}\sum_{n_1'}\frac{\Delta(n_1', n_2)}{E - 2(\pi/L)^2(n_1'^2 + 3n_2^2)}$$

$$+ 2\frac{t}{L}\sum_{n_1'}\frac{F(n_1' + 2n_2)\,\Delta(n_1', n_1' + 2n_2)}{E - 2(\pi/L)^2[n_1'^2 + 3(n_1' + 2n_2)^2]}. \qquad (17)$$

We can simplify this equation by first noting that if n_1' is even (odd) then $n_1' + 2n_2$ is even (odd) so $\Delta(n_1', n_1' + 2n_2)$ is always unity in the second sum. Also in this sum we let $n_1' + 2n_2 = l$, where l ranges integrally from $-\infty$ to ∞. Finally we define E' by

$$E = 2E'(\pi/L)^2$$

and the equation becomes, on relabelling a summation variable,

$$2\pi^2F(n_2) = (tL)F(n_2)\sum_m\frac{\Delta(m, n_2)}{E' - m^2 - 3n_2^2}$$

$$+ 2(tL)\sum_l\frac{F(l)}{E' - [(l - 2n_2)^2 + 3l^2]},$$

$$n_2 = 0, 1, 2\cdots. \qquad (18)$$

The fact that $F(n_2) = F(-n_2)$ has enabled us to restrict the values of the n_2 we consider to $0, 1, 2\cdots$ as we have indicated. Note that the first sum in this equation can be done although the result makes the equation more cumbersome. But, for reference, we have

n even

$$\sum_m\frac{\Delta(m, n)}{E' - 3n^2 - m^2} = \sum_{s=-\infty}^{\infty}\frac{1}{E' - 3n^2 - 4s^2}$$

$$= \begin{cases} \dfrac{\pi}{2(E' - 3n^2)^{\frac{1}{2}}}\cot\dfrac{\pi}{2}(E' - 3n^2)^{\frac{1}{2}}, & E' > 3n^2, \\ -\dfrac{\pi}{2(3n^2 - E')^{\frac{1}{2}}}\coth\dfrac{\pi}{2}(3n^2 - E')^{\frac{1}{2}}, & E' < 3n^2. \end{cases}$$

n odd

$$\sum_m\frac{\Delta(m, n)}{E' - 3n^2 - m^2} = \sum_{s=-\infty}^{\infty}\frac{1}{E' - 3n^2 - (2s + 1)^2}$$

$$= \begin{cases} -\dfrac{\pi}{2(E' - 3n^2)^{\frac{1}{2}}}\tan\dfrac{\pi}{2}(E' - 3n^2)^{\frac{1}{2}}, & E' > 3n^2, \\ -\dfrac{\pi}{2(3n^2 - E')^{\frac{1}{2}}}\tanh\dfrac{\pi}{2}(3n^2 - E')^{\frac{1}{2}}, & E' < 3n^2. \end{cases}$$

The set of equations that (18) represents is then our basic set; since it is homogeneous, it has solutions only for certain eigenvalues of E', and the lowest one is the ground-state energy we seek.

We present in Table I results on the numerical solution of these equations. Before we discuss these, it is worth pointing out, however, that we know the answer to expect in two different limiting cases: $t \to 0$ and $t \to \infty$. For $t \to 0$ we expect the perturba-

[8] We see here quite concretely an advantage of our way of writing the wave function as a sum of two-body orbitals: it is only as a consequence of this form that we can reduce the problem to one for a function of only one variable.

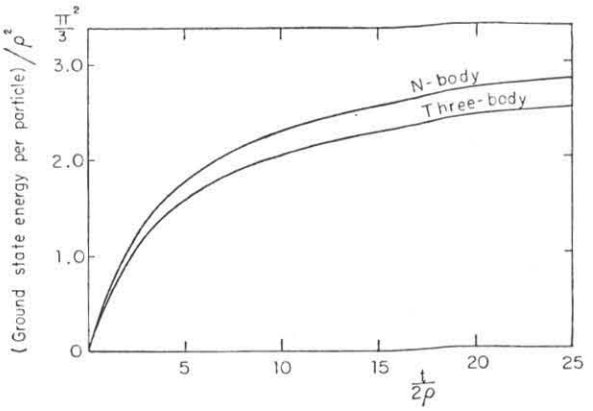

FIG. 1. The ground-state energy per particle, for a one-dimensional N-body system with repulsive δ-function interparticle potentials of strength t is compared with that for a three-body system of the same linear density ρ. For the three-body system $t/2\rho \rightarrow tL/6$. The N-body results are due to Lieb and Liniger,[10] and following them we set $\hbar = 2m = 1$ for the purposes of this plot.

tion theory result $E = 3t/L$, and expect moreover that $F(0)$ is of the order of unity and all other $F(n)$ are small and of order t. These results can be derived simply from the equations above; we omit the details. For $t \rightarrow \infty$ we also know from the work of Girardeau[9] that $E' = 4$ although Girardeau was not interested in and hence did not derive expressions for the orbitals, nor by the same token for the function $F(n)$. Our numerical results do seem to approach the value $E' = 4$ in this limit.

For arbitrary values of t the eigenvalues of the (infinite) set of homogeneous equations (18) are determined by the vanishing of the corresponding determinant. In practice we must of course truncate these equations so that their determinant becomes of finite order, and then test for the error introduced by this truncation. We have done this numerically for the lowest eigenvalue, by examining the sequence of results for this eigenvalue and corresponding "eigenvector" $F(n)$ as a function of the size of the truncated determinant. It turns out that for all values of t the ground-state eigenvector $F(n)$ decreases rapidly with n, so that generally one gets good convergence with truncated determinants of relatively small order. The convergence worsens as tL increases; this is reflected in the smaller number of significant figures presented for the larger values. The largest determinants used in deriving these results were of thirteenth order, although for the smaller values of tL accurate results were obtained with determinants of fourth, third, and sometimes second order.

[9] M. Girardeau, J. Math. Phys. 1, 516 (1960).

To our knowledge, there are for this problem no other results in the literature with which we might compare the above calculations. However, Lieb and Liniger[10] have solved, also numerically, the analogous N-body problem in the limit $N \rightarrow \infty$, $L \rightarrow \infty$, $N/L = \text{const}$. Now $N = 3$ is not, on the face of it, a large number, but it is nonetheless interesting to see how close the present results for the energy *per particle* are to those of Lieb and Liniger, for systems of the *same linear density*. For the moment then we denote by L_∞ the periodicity length for the N-body system so that the linear density ρ is[11]

$$\rho = N/L_\infty.$$

For the three-body system, on the other hand, the linear density is $3/L$. If then we let $N/L_\infty = 3/L$ the two systems will have the same density. Now Lieb and Liniger calculate energies as a function of the dimensionless parameter $t/2\rho = tL_\infty/2N$. We can then compare our results with theirs by also plotting energy against $t/2\rho$, which parameter for the three-body case is $tL/6$. Such a plot is shown in Fig. 1. The results for the two cases are, perhaps surprisingly, fairly close.

III. GENERAL METHOD OF SOLUTION: EXPANSION IN TWO-BODY FUNCTIONS

In this section we discuss a general method for the solution of Eqs. (14). An essential feature of it is the expansion of the function $C(n_1, n_2)$ in terms of a certain set of two-body functions that we describe below. The motivation for this is twofold. First, we have previously discussed reasons for believing that the orbital ψ_{12} should, as a function of x_{12}, resemble a two-body wavefunction and it is then very natural to try to expand it [or by the same token its Fourier counterpart $C(n_1, n_2)$] in a complete set of such functions. We have also found[1] that such an expansion works well for the three-body problem with attractive potentials. Secondly, the introduction of the two-body functions will enable us to eliminate from the three-body problem any explicit appearance of the two-body potential. This is most useful when we consider infinite repulsive potentials ("hard cores"), either in one or three dimensions. For it is generally true, with a Schrödinger equation involving a repulsive potential of strength v_0, that even in the limit $V_0 \rightarrow \infty$ the wavefunctions and energy levels remain unpathological. One has only to think of a particle between high repulsive walls, say, or the

[10] Elliott H. Lieb and Werner Liniger, Phys. Rev. 130, 1605 (1963).

[11] Strictly, ρ is not N/L_∞, but is the limit of this ratio as both N and L_∞ become infinite.

scattering by a repulsive sphere. As we see shortly, the expansion in two-body functions[12] enables us to eliminate from the three-body problem any explicit appearance of the two-body potential itself; instead there appears the complete set of eigenfunctions and corresponding eigenvalues of the solution to the two-body problem with this potential. There is then no difficulty in going to the limit of infinite potential since these eigenfunctions and energies are well behaved in that limit.

Given the potential $v(x_{12})$ that enters into the three-body equation, let us consider the two-body problem for a potential of this *shape*, but differing in *strength* by a factor $1/\lambda$. As we show in the Appendix the two-body momentum space wavefunctions[13] $\tilde{C}_{lm}$ for this potential satisfy the equation [essentially Eq. (A12)]

$$\left[\tilde{E}_{lm} - 2\left(\frac{\pi}{L}\right)^2 (n_1^2 + n_2^2)\right]\tilde{C}_{lm}(n_1, n_2)\ \Delta(n_1, n_2)$$

$$= \frac{1}{\lambda}\sum_{n_1'} U\left(\frac{n_1 - n_1'}{2}\right)\tilde{C}_{lm}(n_1', n_2)\ \Delta(n_1', n_2) \qquad (19)$$

with the orthogonality relations

$$\sum_{n_1, n_2} \tilde{C}_{lm}(n_1, n_2)\tilde{C}_{l'm'}(n_1, n_2) = \delta_{ll'}\ \delta_{mm'}. \qquad (20)$$

Specifically, the $\tilde{C}_{lm}$ are of the form

$$\tilde{C}_{lm} = D_l(n_1)\ \delta_m(n_2). \qquad (21)$$

Now as we have seen, the function $C(n_1, n_2)$ satisfies the same "even–odd" condition on its variables as occurs in Eq. (19) for $\tilde{C}_{lm}$. We can then expand $C(n_1, n_2)$ in terms of these $\tilde{C}_{lm}$, and Eqs. (14) will become an infinite set of coupled equations for the expansion coefficients. We write then[14]

$$C(n_1, n_2) = \sum_{\langle l\rangle, m} A_{lm}\tilde{C}_{lm}(n_1, n_2) \qquad (22)$$

and put this into (14). We have then for the first term on the right-hand side

$$\sum_{\langle l\rangle, m, n_1'} U\left(\frac{n_1 - n_1'}{2}\right)A_{lm}\tilde{C}_{lm}(n_1', n_2)\ \Delta(n_1', n_2)$$

$$= \lambda\sum_{\langle l\rangle, m} A_{lm}\tilde{C}_{lm}\left[\tilde{E}_{lm} - 2\left(\frac{\pi}{L}\right)^2 (n_1^2 + n_2^2)\right]\Delta(n_1, n_2). \qquad (23)$$

We transpose this to the left-hand side, multiply the resulting equation by $\tilde{C}_{st}(n_1, n_2)$, and sum over n_1 and n_2. On the right-hand side we use a formula which is a variant of (19)

$$\sum_{n_1} \tilde{C}_{st}(n_1, n_2)U\left(\frac{p_1 - n_1}{2}\right)\Delta(n_1, n_2)$$

$$= \lambda\left[\tilde{E}_{st} - 2\left(\frac{\pi}{L}\right)^2 (n_2^2 + p_1^2)\tilde{C}_{st}(p_1, n_2)\right]\Delta(p_1, n_2).$$

We get

$$(E - \lambda\tilde{E}_{st})A_{st} - 2(\pi/L)^2(1 - \lambda)$$

$$\times \sum_{\langle l\rangle, m, n_1, n_2} n_1^2 A_{lm}\tilde{C}_{lm}(n_1, n_2)\tilde{C}_{st}(n_1, n_2)\ \Delta(n_1, n_2)$$

$$- 2(\pi/L)^2(3 - \lambda)$$

$$\times \sum_{\langle l\rangle, m, n_1, n_2} n_2^2 A_{lm}\tilde{C}_{lm}(n_1, n_2)\tilde{C}_{st}(n_1, n_2)\ \Delta(n_1, n_2)$$

$$= \lambda \sum_{\langle l\rangle, m, n_1', n_2} A_{lm}(\tilde{C}_{lm}(n_1', -n_1' - 2n_2)$$

$$\times \{\tilde{C}_{st}(2n_1' + 3n_2, n_2) + \tilde{C}_{st}[-2n_1' - 3n_2, n_2]\}$$

$$\times \{\tilde{E}_{st} - 2(\pi/L)^2[n_2^2 + (2n_1' + 3n_2)^2]\}).$$

Now we put in the more specific form (21) for $\tilde{C}_{lm}$. We have for use in the left-hand side

$$\sum_{\langle l\rangle, m, n_1, n_2} n_1^2 A_{lm}\ D_l(n_1)\ \delta_m(n_2)\ D_s(n_1)\ \delta_t(n_2)\ \Delta(n_1, n_2)$$

$$= \sum_{\langle l\rangle, n_1} n_1^2 A_{lt}\ D_l(n_1)\ D_s(n_1)\ \Delta(n_1, t), \qquad (24)$$

$$\sum_{\langle l\rangle, m, n_1, n_2} n_2^2 A_{lm}\ D_l(n_1)\ \delta_m(n_2)\ D_s(n_1)\ \delta_t(n_2) = t^2 A_{st}.$$

The first sum on the right-hand side is

$$\lambda \sum_{\langle l\rangle, m, n_1', n_2} A_{lm}\, D_l(n_1')\ \delta_m(-n_1' - 2n_2)\ D_s(2n_1' + 3n_2)$$

$$\times\ \delta_t(n_2)\left\{\tilde{E}_{st} - 2\left(\frac{\pi}{L}\right)^2 [N_2^2 + (2n_1' + 3n_2)^2]\right\}$$

$$= \lambda \sum_{\langle l\rangle, m} A_{lm}\left\{\tilde{E}_{st} - 2\left(\frac{\pi}{L}\right)^2 [(2m + t)^2 + t^2]\right\}$$

$$\times\ D_s(-2m - t)\ D_t(-2t - m) \qquad (25)$$

with a similar result for the second sum. Finally Eq. (14) becomes

$$(E - \lambda\tilde{E}_{st})A_{st} - 2\left(\frac{\pi}{L}\right)^2 [(1 - \lambda)\sum_{\langle l\rangle, n_1} n_1^2 A_{lt}\ D_l(n_1)$$

$$\times\ D_s(n_1)\ \Delta(n_1, t) + (3 - \lambda)t^2 A_{st}]$$

$$= \lambda \sum_{\langle l\rangle, m} A_{lm}\left\{\tilde{E}_{st} - 2\left(\frac{\pi}{L}\right)^2 [(2m + t)^2 + t^2]\right\}$$

$$\times\ D_t(-2t - m)[D_s(2m + t) + D_s(-2m - t)]. \qquad (26)$$

[12] Fadeev has used a similar expansion in his elegant discussion of the formal properties of the three-body scattering problem. Cf. Ref. 2.

[13] We generally use the tilde ($\sim$) to denote quantities that refer to the two-body problem.

[14] As we discuss in the Appendix, the index l which labels the eigenvalues takes on the values 0, 1, 2, $\cdots$, and hence does not follow the convention of footnote 4. We indicate this by enclosing it in angular brackets.

Once again we have an infinite homogeneous set of linear equations, whose vanishing determinant defines the eigenvalues. Moreover, we have achieved the objective of eliminating any explicit appearance of the two-body potential; it has been replaced by the complete set of two-body wavefunctions and energies. Also we note that the parameter λ in these equations is at our disposal. We can choose it arbitrarily, since taking different values for it merely corresponds to choosing different complete sets of basis functions. In principle then, the set of equations (26) has the same set of solutions whatever the value of λ. In practice, of course, we shall be able to solve the infinite set of equations only by truncating them, and we would like to get good accuracy with as small a truncated set as possible. To this end we can consider λ to be a parameter which we can try to choose to achieve this.

IV. APPLICATIONS OF THE GENERAL METHOD

In this section we first apply the general equations (26) to the problem of repulsive δ functions, and compare the results with the exact ones of Sec. II. Since we do have these exact results, our chief aim will not be to duplicate them, but only to check Eqs. (26) by actually using them and also to turn up any problems that may be involved in truncating them. We begin then by truncating them as severely as possible, and keep only the lowest order term; i.e., we assume that only A_{00} is different from zero, so that only the lowest order eigenfunction $D_0(n_1)$ appears. Also, to conform with the conventions of the Appendix, we add a plus superscript to $D_0(n_1)$ although for the present purpose it is really superfluous: $D_0(n_1) \rightarrow D_0^+(n_1)$. We get then the single simple equation,

$$(E - \lambda\tilde{E}_{00}) - 2\left(\frac{\pi}{L}\right)^2(1 - \lambda)\sum_{\substack{n_1 \\ \text{even}}} n_1^2[D_0^+(n_1)]^2 = 2\lambda\tilde{E}_{00}[D_0^+(0)]^2. \qquad (27)$$

We must now choose λ, which we do according to the following prescription. With E the ground-state energy of the three-body problem, we choose λ such that the potential strength t/λ for the *two-body* problem would also have ground-state eigenvalue E. That is, we choose λ so that $\tilde{E}_{00} = E$. This is, of course, the same prescription that we used for the three-body problem for attractive potentials[1] and the motivation for choosing it here is the same for having chosen it there in the first place, plus the *a posteriori* justification that it did turn out to work well. Of course we do not know the ground-state energy E in advance, but we can get λ as a *function* of this unknown E from the two-body solution. If we put this expression for λ as well as the expression for the corresponding $D_0^+(n_1)$ back into (27), we get a transcendental equation for E. Before we do this we must mention that there is one limitation here which is peculiar to repulsive potentials and in particular to repulsive potentials in one dimension. That is, one cannot fill the above prescription for λ for *all* values of t. For we know from the work of Girardeau[9] that in the limit $t \rightarrow \infty$ the three-body problem has ground-state energy $E = 8(\pi/L)^2$. Thus as t varies from zero to infinity, the ground-state energy varies from zero to this value. For the two-body problem, however, we know similarly that E varies from zero to $2(\pi/L)^2$. Thus it is only for such t where the three-body energy is between zero and $2(\pi/L)^2$ that we can find a λ for which the two-body problem with strength t/λ has the same ground-state energy. For attractive potentials, on the other hand, one can make the two-body energy vary between zero and minus infinity and thereby match any three-body ground-state energy. Also, for three-dimensional repulsive potentials with radius r_0 one can find a two-body problem with the same ground-state energy as the three-body one by taking the radius for the two-body problem to be r_0/λ, $\lambda < 1$. Since the limitation mentioned above does seem peculiar to the one dimensional repulsive potential problem, we do not worry further about it, but simply limit the comparison of our results to those values of t for which our prescription works.

With this, we now return to the evaluation of Eq. (27). For one-dimensional δ functions, we find from the Appendix that, for the two-body problem, λ and $D_0^+(n_1)$ are just (apart from a normalization factor)

$$\lambda = \frac{t}{2(2E)^{\frac{1}{2}}}\cot\frac{(E)^{\frac{1}{2}}L}{2\sqrt{2}} \qquad (28)$$

and

$$D_0^+(n_1) = \frac{1}{E - 2(\pi/L)^2 n_1^2}, \qquad n_1 \text{ even}. \qquad (29)$$

We must then put (28) and (29) into (27), and do the sum over n_1. This can be done directly, or more simply, we can use the fact that $D_0^+(n_1)$ satisfies the equation

$$\left[E - 2\left(\frac{\pi}{L}\right)^2 n_1^2\right] D_0^+(n_1) = \frac{t}{L\lambda}\sum_{\substack{n_1' \\ \text{even}}} D_0^+(n_1').$$

We multiply this by $D_0^+(n_1)$ and sum over n_1 to get an expression for $\sum n_1^2[D_0^+(n_1)]^2$ in terms of the simpler $\sum D_0^+(n_1)$. Putting this into (27), it becomes

$$(1-\lambda)\frac{t}{L\lambda}\Big[\sum_{\substack{n_1\\ \text{even}}} D_0^+(n_1)\Big]^2 = 2\lambda E[D_0^+(0)]^2. \tag{30}$$

The sum over n_1 is

$$\sum_{\substack{n_1\\ \text{even}}} D_0^+(n_1) = \frac{L\lambda}{t} = \frac{L}{2(2E)^{\frac{1}{2}}}\cot\frac{(E)^{\frac{1}{2}}L}{2\sqrt{2}},$$

and with $D_0^+(0) = 1/E$ we find that (30) becomes

$$\begin{aligned} 2\pi^2E' &= tL[\tfrac{1}{2}\pi(E')^{\frac{1}{2}}\cot\tfrac{1}{2}\pi(E')^{\frac{1}{2}} + 2],\\ E &= 2E'(\pi/L)^2. \end{aligned} \tag{31}$$

Equation (31) is easy enough to solve numerically, and the results are plotted in Fig. 2 along with the exact results of Sec II and the perturbation theory result $E = 3tL$. One sees there that, even when the potential is fairly strong (by which we mean even when the perturbation result begins to be off by a factor of 2), Eq. (31) still gives a fair approximation to the exact results.

We leave the δ-function problem with this. We could presumably get better numerical results by truncating Eqs. (26) less severely, but this is not to the present point, which has been to illustrate the simplest application of these equations. The agreement we get with the lowest order approximation is sufficient for the purpose of the moment.

We now consider in the same spirit as for δ-function potentials the truncated equation (27) as applied to the square well potential: $v(x) = v_0$, $|x| < b/2$, $v(x) = 0$, otherwise. We derive the analog of Eq. (31), just to see what it looks like. Now, if we were to use Eq. (27) directly we would have to find the function $D_0^+(n_1)$, but, just as above, it is more convenient to transform this equation into one for wavefunctions in position space. From the Appendix, and, in particular, from Eq. (A7) satisfied by $D_0^+(n_1)$, we see that $\{E - 2(\pi/L)^2\sum_{n_1} n_1^2[D_0^+(n_1)]^2\}$ represents the total energy minus the kinetic energy, i.e., it represents the expectation value of the potential energy, which we denote by $\langle$P.E.$\rangle$. Thus Eq. (27) can be written

$$[(1/\lambda) - 1]\langle\text{P.E.}\rangle = 2E[D_0^+(0)]^2.$$

Remembering that this potential energy refers to a potential of strength v_0/λ, this expectation value is, in terms of the position wavefunction $\bar{\psi}(x)$,

$$\langle\text{P.E.}\rangle = \frac{v_0}{\lambda}\int_{-\frac{1}{2}L}^{\frac{1}{2}L} |\bar{\psi}(x)|^2\,dx \tag{32}$$

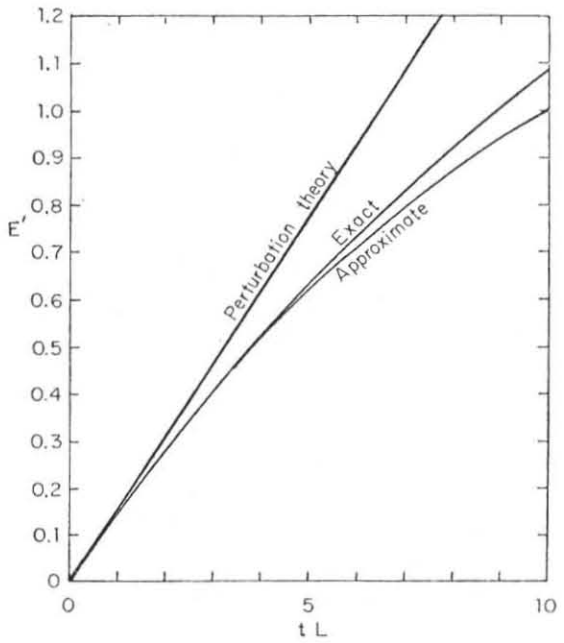

FIG. 2. Ground-state eigenvalue E' as a function of tL for one-dimensional δ-function potentials as given by: (a) perturbation theory, (b) approximate solution to lowest order according to Eq. (27), (c) exact numerical solution. $E = 2E'(\pi/L)^2$.

with

$$\bar{\psi}(x) = \sum_l D_0^+(l)\exp\left[\frac{i2\pi xl}{L}\right]. \tag{33}$$

From this last equation we also have

$$D_0^+(0) = \frac{1}{L}\int_{-\frac{1}{2}L}^{\frac{1}{2}L}\bar{\psi}(x)\,dx.$$

We now put (32) and (33) into (27) to get our final equation in terms of the two-body wavefunction $\bar{\psi}$, an expression for which is given in the Appendix. With this we get the result, the analog of Eq. (31) for δ-function potentials,

$$\begin{aligned}\left(\frac{1}{\lambda}-1\right)\frac{v_0}{\beta L\lambda}(e^{\beta b} - e^{-\beta b} + 2\beta b) &= \frac{32E}{(\cos^2\frac{1}{2}\alpha a)L^2}\\ \times\left(\frac{\cos\frac{1}{2}\alpha a\sinh\frac{1}{2}\beta b}{\beta} + \frac{\cosh\frac{1}{2}\beta b\sin\frac{1}{2}\alpha a}{\alpha}\right)^2,\end{aligned} \tag{34}$$

where

$$\beta\tanh\beta b = \alpha\tan\alpha a \tag{35}$$

and

$$\alpha = (E)^{\frac{1}{2}}, \qquad \beta = [(v_0/\lambda) - E]^{\frac{1}{2}}. \tag{36}$$

Equation (34) defines λ in terms of E, v_0, a, b so that, along with Eqs. (35) and (36), it becomes a transcendental equation for E as a function of these

parameters, which can be solved numerically. We do not present numerical results here, since our main point in this paper is to check its general viewpoint, and there are no independent results for this problem with which to compare. In this spirit, however, it is useful to note that the perturbation theory result can be derived from these equations. That is, in the absence of a potential the ground-state three-body wavefunction is a constant; we expect then that for v_0 small the energy eigenvalue E will be just the expectation value of the total potential for this wavefunction, and this perturbation energy, E_{pert}, is

$$E_{\text{pert}} = 3v_0 b/(a + b).$$

This result comes directly out of Eqs. (34), (35), (36).

ACKNOWLEDGMENTS

I would like to thank Mr. Joseph Martine of the Data Analysis Branch at this Laboratory for programming the solution of Eqs. (18).

APPENDIX. TWO-BODY WAVEFUNCTIONS

A. General Results

We want to solve the two-body Schrödinger equation

$$\left[-\frac{\partial^2}{\partial x_1^2} - \frac{\partial^2}{\partial x_2^2} + v(x_{12})\right]\bar{\psi} = \bar{E}\bar{\psi} \tag{A1}$$

with the boundary condition that $\bar{\psi}$ be periodic in x_1 and x_2 with period L. As in Sec. I, it is convenient to deal not with $v(x_{12})$ but with its periodic counterpart

$$v_p(x_{12}) = \sum_l U(l) \exp\left[\frac{i2\pi l x_{12}}{L}\right]. \tag{A2}$$

Since we work in relative and center-of-mass coordinates x and X,

$$x = x_1 - x_2, \qquad X = x_1 + x_2,$$

we must first consider the relation between periodicity in these variables and in x_1 and x_2. The general function periodic in x_1 and x_2 is

$$\sum_{s_1, s_2} (\text{function of } s_1, s_2) \exp\left[\frac{i2\pi}{L}(s_1 x_1 + s_2 x_2)\right]$$
$$= \sum_{s_1, s_2} (\text{function of } s_1, s_2) \times \exp\left[\frac{i2\pi}{2L}(s_1 - s_2)x + (s_1 + s_2)X\right]$$

with

$$s_1 - s_2 = n_1', \qquad s_1 + s_2 = n_2'.$$

We see that, as s_1 and s_2 independently range over all integers, n_1' and n_2' also range over the integers, although not independently, since they are subject to the condition that both must be even or both be odd. We conclude then that the general form for a function periodic in x_1 and x_2 is

$$\sum_{n_1', n_2'} \tilde{C}(n_1', n_2') \exp\left[\frac{i2\pi}{2L}(n_1' x + n_2' X)\right], \qquad n_1' \text{ and } n_2' \text{ both even or both odd,}$$

where $\tilde{C}(n_1', n_2')$ is an arbitrary function. As in Sec. I, we can conveniently keep track of the even–odd condition above by using the $\Delta(m, n)$ defined there. With it we write

$$\bar{\psi}(x, X) = \sum_{n_1', n_2'} \tilde{C}(n_1', n_2')\, \Delta(n_1', n_2') \times \exp\left[\frac{i2\pi}{2L}(n_1' x + n_2' X)\right]. \tag{A3}$$

If then we put (A3) and (A2) into (A1), multiply by $\exp[-(i2\pi/2L)(n_1 x + n_2 X)]$ and integrate from zero to $2L$ we get

$$\left[\bar{E} - 2\left(\frac{\pi}{L}\right)^2 (n_1^2 + n_2^2)\right]\tilde{C}(n_1, n_2)\, \Delta(n_1, n_2) = \sum_{n_1'} U\left(\frac{n_1 - n_1'}{2}\right)\tilde{C}(n_1', n_2)\, \Delta(n_1', n_2). \tag{A4}$$

We see immediately that the form of the solution of this equation is

$$\tilde{C}(n_1, n_2) = D(n_1)\, \delta_m(n_2), \tag{A5}$$

where $D(n_1)$ is for the moment an arbitrary function and $\delta_m(n_2)$ is just a Kronecker δ function, with m an integer. Given the form (A5), we note that $D(n_1)$ is either of two different types according to whether m is even or odd. For if m is even the only allowed value of n_2 is perforce even, and as a consequence, n_1 ranges over even integers. Similarly, for m odd n_1 ranges over odd integers. It is convenient to label these two types of solution[15] explicitly; we call the first one D^+ and the second D^-,

$$D(n_1) \to \begin{cases} D^+(n_1), & m \text{ even}, \\ D^-(n_1), & m \text{ odd}. \end{cases} \tag{A6}$$

We also define W by

$$W = \bar{E} - 2(\pi/L)^2 m^2.$$

We first consider m even. We put (A5) into (A4)

[15] These two solutions are equivalent to those derived by Lieb and Liniger for the same problem from another point of view. Cf. Ref. 10.

and add a superscript to $D(n_1)$ according to Eq. (A6). Equation (A4) then becomes

$$\left[W - 2\left(\frac{\pi}{L}\right)^2 n_1^2\right] D^+(n_1) = \sum_{\substack{n_1' \\ \text{even}}} U\left(\frac{n_1 - n_1'}{2}\right) D^+(n_1'),$$

$$n_1 \text{ even.} \qquad \text{(A7)}$$

Similarly, for m odd we have

$$\left[W - 2\left(\frac{\pi}{L}\right)^2 n_1^2\right] D^-(n_1) = \sum_{\substack{n_1' \\ \text{odd}}} U\left(\frac{n_1 - n_1'}{2}\right) D^-(n_1),$$

$$n_1 \text{ odd.} \qquad \text{(A8)}$$

Equations (A7) and (A8), derived from a two-body Schrödinger equation, can be related to the solution of a one-body equation, namely to the problem of a single particle, coordinate x, moving in the periodic potential $v_p(x)$. This is obviously a useful reduction, since there are standard methods for solving such one-body problems. Consider then the Schrödinger equation for the periodic potential $v_p(x)$ of Eq. (A2). If we call the wavefunction $\phi(x)$, and choose units properly, it can be written

$$-2\, d^2\phi/dx^2 + v_p(x)\phi = W\phi. \qquad \text{(A9)}$$

We know from Bloch's theorem that $\phi(x)$ can be written in the form $e^{ikx}u_k(x)$, where u_k is periodic in x; if we imagine this periodic function expanded in Fourier series,

$$\phi(x) = e^{ikx} \sum_s F(s) \exp\left[\frac{i2\pi sx}{L}\right],$$

and write $k = 2\pi\kappa/L$ then the equation for the coefficients $F(s)$ is

$$\left[W - 2\left(\frac{2\pi}{L}\right)^2 (s + \kappa)^2\right] F(s) = \sum_{s'} U(s - s')F(s'). \qquad \text{(A10)}$$

Now consider Eq. (A7) for $D^+(n_1)$. Since n_1 is even we write $n_1 = 2s$, with s integral, and define $G^+(s)$ by

$$D^+(n_1) \equiv D^+(2s) = G^+(s).$$

Equation (A7) for $D^+(n_1)$ then becomes, in terms of $G^+(s)$

$$\left[W - 2\left(\frac{2\pi}{L}\right)^2 s^2\right] G^+(s) = \sum_{s'} U(s - s')G^+(s').$$

Similarly, for $D^-(n_1)$, we let $n_1 = 2s + 1$, with s integral, and define $G^-(s)$ by

$$D^-(n_1) \equiv D^-(2s + 1) = G^-(s).$$

We have then for $G^-(s)$

$$\left[W - 2\left(\frac{2\pi}{L}\right)^2 (s + \tfrac{1}{2})^2\right] G^-(s) = \sum_{s'} U(s - s')G^-(s').$$

Now we observe that these equations for G^+ and G^- are identical to Eq. (A10), for $\kappa = 0$ and $\kappa = \frac{1}{2}$, respectively. The problem of solving these equations is thereby reduced to a canonical problem of solid state physics: calculation of the wavefunctions and energies for a particle in a periodic lattice for the points $\kappa = 0$ and $\kappa = \frac{1}{2}$ in the "Brillouin zone."

To discuss the orthogonality properties of the eigenfunctions we have found that we must first put some additional labels on them. Call $D_l^+(n_1)$ and $D_l^-(n_1)$ the eigenfunctions[16] corresponding to the lth eigenvalue of Eqs. (A7) and (A8), respectively. Then the function $\tilde{C}(n_1, n_2)$ can be similarly labeled and we define $\tilde{C}_{lm}(n_1, n_2)$ by

$$\tilde{C}(n_1, n_2) \to \tilde{C}_{lm}(n_1, n_2) \equiv D_l(n_1)\ \delta_m(n_2) \qquad \text{(A11)}$$

$$\to \begin{cases} D_l^+(n_1)\ \delta_m(n_2), & m = 0, \pm 2, \pm 4 \cdots \\ D_l^-(n_1)\ \delta_m(n_2), & m = \pm 1, \pm 3, \pm 5 \cdots. \end{cases}$$

If we call $\tilde{E}_{lm}$ the corresponding eigenvalues, we get Eq. (19) of the text by labeling Eq. (A4),

$$\left[\tilde{E}_{lm} - 2\left(\frac{\pi}{L}\right)^2 (n_1^2 + n_2^2)\right] \tilde{C}_{lm}(n_1, n_2)\ \Delta(n_1, n_2)$$

$$= \sum_{n_1'} U\left(\frac{n_1 - n_1'}{2}\right) \tilde{C}_{lm}(n_1', n_2)\ \Delta(n_1', n_2). \qquad \text{(A12)}$$

We want now to investigate the orthogonality properties of the $\tilde{C}_{lm}$, that is, the properties of the "orthogonality sum" defined by

$$\sum_{n_1, n_2} \tilde{C}_{lm}\tilde{C}_{l'm'}.$$

From Eq. (A11) we see that this sum contains $\delta_m(n_2)\delta_{m'}(n_2)$, and therefore vanishes unless $m = m'$. But if $m = m'$, D_l and $D_{l'}$ must both be of plus type or both of minus type. Thus the problem of the orthogonality properties of the sum above reduces to that for the sums

$$\sum_{\substack{n_1 \\ \text{even}}} D_l^+(n_1)\ D_{l'}^+(n_1)$$

and

$$\sum_{\substack{n_1 \\ \text{odd}}} D_l^-(n_1)\ D_{l'}^-(n_1).$$

But it is easy to show in a standard way directly

[16] We let l range through 0, 1, 2, $\cdots$ so the lowest eigenfunctions are D_0^+ and D_0^-.

from the equations defining D_l^+ and D_l^- that these sums are zero[17] unless $l = l'$, in which case they can be made unity by a proper choice of normalization constant. All this discussion then can be subsumed by the remark that the $\tilde{C}_{lm}$ form an orthonormal set with respect to summations over n_1 and n_2,

$$\sum_{n_1,n_2} \tilde{C}_{lm}(n_1, n_2)\tilde{C}_{l'm'}(n_1, n_2) = \delta_{ll'}\,\delta_{mm'}. \qquad \text{(A13)}$$

B. Delta Function and Square Well Potentials

We now specialize the preceding discussion to δ function and square well potentials, for application of the results to the corresponding three-body problems of the text.

For the δ-function potential, $v(x) = t\delta(x)$, we have

$$U(l) = t/L,$$

and Eq. (A7) becomes

$$\left[W - 2\left(\frac{\pi}{L}\right)^2 n_1^2\right] D^+(n_1) = \frac{t}{L} \sum_{\substack{n_1' \\ \text{even}}} D^+(n_1'), \quad n_1 \text{ even}. \qquad \text{(A14)}$$

The solution of this equation is immediate. We have, after adding a superscript to W to show that it is associated with D^+,

$$D^+(n_1) = \frac{N^+}{W^+ - 2(\pi/L)^2 n_1^2}, \qquad \text{(A15)}$$

where N^+ is a normalization constant. We have only to determine the allowed values W^+. This is simple enough, for, if we put (A15) back into (A14), we get a transcendental equation involving W^+ in terms of an infinite sum, which sum can be done by the Poisson summation formula:

$$1 = \frac{t}{L} \sum_{\substack{n_1 \\ \text{even}}} \frac{1}{W^+ - 2(\pi/L)^2 n_1^2} = \frac{t}{2(2W^+)^{\frac{1}{2}}} \cot \frac{(W^+)^{\frac{1}{2}}L}{2\sqrt{2}}. \qquad \text{(A16)}$$

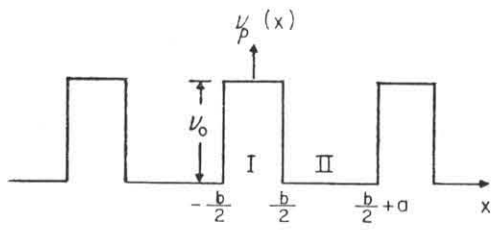

FIG. 3. The periodic potential $v_p(x)$ for square well repulsive potentials.

[17] We assume that there are no degenerate states.

Similarly, we have for D^-

$$D^-(n_1) = \frac{N^-}{W^- - 2(\pi/L)^2 n_1^2}, \qquad n_1 \text{ odd},$$

with the transcendental equation for determining W^-

$$1 = -\,[t/2(2W^-)^{\frac{1}{2}}] \tan\,[(W^-)^{\frac{1}{2}}L/2\sqrt{2}]. \qquad \text{(A17)}$$

The normalization factors turn out to be the same for the plus and minus solutions. Dropping the superscript on W for the moment, they are

$$(N^+)^2 = (N^-)^2 = \frac{t^2}{L^2[(t^2/16W) + (t/2WL) + \frac{1}{2}]}.$$

Now we turn to two-body square-well potentials, i.e., we take $v(x)$ to be

$$v(x) = \begin{cases} v_0, & -\frac{1}{2}b < x < \frac{1}{2}b, \\ 0, & \text{otherwise}. \end{cases}$$

$v_p(x)$, the periodic counterpart of $v(x)$, is then just the well-known Kronig–Penney potential of solid state physics. It is sketched in Fig. 3. The functions D^+ and D^- can therefore be found as the Fourier coefficients of the known coordinate space solutions in this potential. Since for present use in this paper we do not, however, need D^+ and D^- but only the lowest coordinate space solution [corresponding to $\kappa = 0$ in Eq. (A10)], we derive it directly here via the Wigner–Seitz method.[18] Referring to Fig. 3, the symmetric solution (for $\kappa = 0$) for region I, $|x| < \frac{1}{2}b$ is

$$C(e^{\beta x} + e^{-\beta x}), \qquad \beta = [\tfrac{1}{2}(v_0 - W)]^{\frac{1}{2}},$$

and the solution in region II is

$$A\,e^{i\alpha x} + B\,e^{-i\alpha x}, \qquad \alpha = (\tfrac{1}{2}W)^{\frac{1}{2}}.$$

The usual continuity conditions at $x = \frac{1}{2}b$ and the periodicity condition that $d\psi/dx$ be zero at $x = \frac{1}{2}L \equiv \frac{1}{2}(a + b)$ then yield the equations

$$B = A\,e^{i\alpha L},$$

$$C = \frac{A(e^{\frac{1}{2}i\alpha b} + e^{i\alpha(L-\frac{1}{2}b)})}{e^{\frac{1}{2}\beta b} + e^{-\frac{1}{2}\beta b}},$$

which define the wavefunction (except for normalization); the transcendental equation determining the energy turns out to be

$$\beta \tanh \tfrac{1}{2}\beta b = \alpha \tan \tfrac{1}{2}\alpha a.$$

[18] See almost any text on solid state physics: e.g., C. Kittel, *Introductory Solid State Physics*, (John Wiley & Sons, Inc., New York, 1956), 2nd ed.