JOURNAL OF
MATHEMATICAL
PHYSICS

VOLUME 5, NUMBER 6 JUNE 1964

Segal's Quantization Procedure

REESE T. PROSSER
Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts*
(Received 1 November 1963)

The procedure proposed by I. E. Segal for the quantization of nonlinear problems of field theory is here applied to the one-dimensional oscillator. It is shown that for the linear oscillator Segal's procedure is equivalent to the canonical procedure, but that for the nonlinear oscillator the two procedures lead to quite different results. The differences are reflected in the equations of motion, the energy spectra, and the scattering cross sections.

INTRODUCTION

A QUANTIZATION procedure has been proposed by I. E. Segal for use in nonlinear problems in quantum field theory.[1] For such problems the conventional quantization procedure breaks down, and Segal's procedure appears to hold out hope for a rigorous alternative. Roughly speaking, the conventional procedure tries to quantize the canonical coordinates of the system, and then solve the (quantized) equations of motion, while Segal's procedure sets out to solve the (classical) equations of motion and then quantize the solutions. Since neither procedure has yet been successfully carried through for any nontrivial problem in field theory, a comparison of methods and results must wait upon a good deal of unfinished work.

It is quite possible, however, to make a comparison for problems involving only a finite number of particles. For such problems, both procedures are rigorously defined and amenable to calculation. We consider here a class of problems arising from the motion of a single particle moving along a single axis under the influence of an arbitrary potential. We show that in all cases where the equations of motion are linear the two quantization procedures are, apart from multiplicities, unitarily equivalent. In certain nonlinear cases, however, the two procedures are quite different, and lead to different results. Our methods extend readily to problems involving several particles moving in three dimensions, though the calculations are correspondingly more complicated.

1. CLASSICAL MECHANICS

We consider a single point particle of unit mass moving along a single axis under the influence of a potential V. The position coordinate for this system we denote by q, the momentum coordinate by p; they form the fundamental canonically conjugate coordinates of the system. The Hamiltonian of the system is then

$$H = \tfrac{1}{2}p^2 + V(q). \tag{1.1}$$

Note that for a free particle $V(q) \equiv 0$, while for a harmonic-oscillator particle, $V(q) = \tfrac{1}{2}q^2$.

The equations of motion become

$$\begin{aligned} dq/dt &= \{H, q\} = p, \\ dp/dt &= \{H, p\} = -V'(q), \end{aligned} \tag{1.2}$$

where $\{\ ,\ \}$ denotes the Poisson bracket. These equations imply

$$dq^2/dt^2 = -V'(q). \tag{1.3}$$

* Operated with support from the U. S. Army, Navy, and Air Force.

[1] I. E. Segal, J. Math. Phys. 2, 468 (1960); 5, 269 (1964).

The invariant 2-form for this system is just

$$\Omega = dp\, dq, \tag{1.4}$$

which is already an invariant volume element on E_2.

Thus the motion in time generated by (1.2) is a one-parameter group of mappings $T(t)$ of E_2 onto itself which preserve the volume element.

This motion induces a one-parameter group of mappings $W(t)$ of $\mathfrak{L}_2(E_2)$ onto itself. Indeed, if $f(x)$ is any function in $\mathfrak{L}_2(E_2)$, where $x = (p, q)$, then

$$[W(t)f](x) = f(T(t)x). \tag{1.5}$$

Since $T(t)$ preserves volume, it follows that $W(t)$ preserves norms, and hence must have the form

$$W(t) = \exp(iXt) \tag{1.6}$$

for some self-adjoint generator X on $\mathfrak{L}_2(E_2)$.

To determine X, recall that if f is any differentiable function of p and q, then the behavior in time of f as p and q move is given by

$$\begin{aligned} df/dt &= \{H, f\} \\ &= \frac{\partial H}{\partial p}\frac{\partial f}{\partial q} - \frac{\partial H}{\partial q}\frac{\partial f}{\partial p} \\ &= iXf, \end{aligned} \tag{1.7}$$

where

$$\begin{aligned} X &= \frac{1}{i}\left(\frac{\partial H}{\partial p}\frac{\partial}{\partial q} - \frac{\partial H}{\partial q}\frac{\partial}{\partial p}\right) \\ &= \frac{1}{i}\left(p\frac{\partial}{\partial q} - V'(q)\frac{\partial}{\partial p}\right). \end{aligned} \tag{1.8}$$

It follows that (1.8) is the generator of (1.6).

2. THE CONVENTIONAL QUANTIZATION PROCEDURE

In the conventional procedure, we "replace" q by Q and p by P, where Q and P are self-adjoint operators on $\mathfrak{L}_2(E_1)$, defined by

$$\begin{aligned} Q &= q \text{ (i.e., multiplication by } q), \\ P &= (1/i)\,\partial/\partial q. \end{aligned} \tag{2.1}$$

These operators satisfy the commutation relations

$$[P, Q] = -iI. \tag{2.2}$$

The Hamiltonian becomes

$$H = \tfrac{1}{2}P^2 + V(Q), \tag{2.3}$$

where $V(Q)$ is the operation of multiplication by $V(q)$. The equations of motion (1.2) are "replaced" by

$$dQ/dt = i[H, Q], \qquad dP/dt = i[H, P]. \tag{2.4}$$

Hence the motion is determined by the one-parameter family $U(t) = \exp(iHt)$ via

$$Q(t) = U(t)QU(t)^{-1}, \quad P(t) = U(t)PU(t)^{-1}. \tag{2.5}$$

In particular, it follows from (2.4) that

$$dQ/dt = P, \qquad dP/dt = -V'(Q), \tag{2.6}$$

or

$$d^2Q/dt^2 = -V'(Q). \tag{2.7}$$

3. SEGAL'S QUANTIZATION PROCEDURE

Segal defines the operators

$$R = \frac{1}{2i}\frac{\partial}{\partial p} + q, \qquad R' = -\frac{1}{2i}\frac{\partial}{\partial q} + p \tag{3.1}$$

acting on $\mathfrak{L}_2(E_2)$. He then shows that

$$[R, R'] = -iI. \tag{3.2}$$

He then requires that the behavior in time of the operators R and R' be induced by the underlying motion $T(t)$ of E_2. This amounts to requiring

$$R(t) = W(t)RW(t)^{-1}, \quad R'(t) = W(t)R'W(t)^{-1}, \tag{3.3}$$

with $W(t)$ given by (1.5). These equations imply

$$dR/dt = i[X, R], \qquad dR'/dt = i[X, R']. \tag{3.4}$$

A glance at (1.8) shows that

$$\begin{aligned} \frac{dR}{dt} &= -\frac{1}{2i}\frac{\partial}{\partial q} + p = R', \\ \frac{dR'}{dt} &= -V''(q)\frac{1}{2i}\frac{\partial}{\partial p} - V'(q). \end{aligned} \tag{3.5}$$

Now consider cases. If $V(q) = c$ (constant), we get $dR'/dt = 0$. If $V(q) = cq$, we have $dR'/dt = cI$. If $V(q) = cq^2$, we find $dR'/dt = -(c/i)(\partial/\partial p) - 2cq = -2cR$. In each of these cases, the operator $R(t)$ satisfies the equation

$$(d^2/dt^2)R(t) = -V'[R(t)], \tag{3.6}$$

which corresponds to (2.7). Now it can be shown that any pair of operators which satisfy the commutation relations (3.2) and the (linear) equations of motion (3.6) are, apart from multiplicities, unitarily equivalent with the pair P, Q of Sec. 2. Indeed, Segal has shown[2] that in these cases the operators R and R' can be represented as acting on certain holomorphic functions on the complex plane, and the equivalence of this representation with the conventional one of Sec. 2 follows easily from its properties.

[2] I. E. Segal, Can. J. Math. **13**, 1 (1961).

Now consider the nonlinear cases. When the potential function $V(q)$ contains terms of power higher than quadratic in q, then the equations of motion (1.2) are no longer linear. In this case (3.5) does not reduce to (3.6). To see this in detail, consider the case $V(q) = \frac{1}{2}q^2 + \frac{1}{4}q^4$, which corresponds to the Thirring model in field theory. Equation (3.5) gives

$$\frac{dR'}{dt} = -\frac{(3q^2+1)}{2i}\frac{\partial}{\partial p} - (q^3 + q), \tag{3.7}$$

while (3.6) requires

$$\begin{aligned}\frac{dR'}{dt} &= -R - R^3 = -\left(\frac{1}{2i}\frac{\partial}{\partial p} + q\right) - \left(\frac{1}{2i}\frac{\partial}{\partial p} + q\right)^3 \\ &= -\frac{1}{2i}\frac{\partial}{\partial p} - q - \left(\frac{1}{2i}\frac{\partial}{\partial p}\right)^3 - 3\left(\frac{1}{2i}\frac{\partial}{\partial p}\right)^2 q \\ &\quad - 3\left(\frac{1}{2i}\frac{\partial}{\partial p}\right)q^2 - q^3. \end{aligned}\tag{3.8}$$

Now (3.7) and (3.8) differ by

$$\Delta = \left(\frac{1}{2i}\frac{\partial}{\partial p}\right)^3 + 3\left(\frac{1}{2i}\frac{\partial}{\partial p}\right)^2 q, \tag{3.9}$$

and this difference cannot vanish on any function in $\mathfrak{L}_2(E_2)$.

We must conclude in this case that the pairs R, R', and P, Q are not unitarily equivalent via the same unitary operator for all times t. Presumably this is true for all the nonlinear cases, but the possibility of a coincidental equivalence has not been ruled out. It would be interesting to see just what the relation is between R, R' and P, Q in the nonlinear cases. Does a knowledge of R, R' determine P, Q?

4. A COMPARISON OF THE HAMILTONIANS

It is conceivable that, even though Segal's operators R, R' fail to satisfy the equations of motion, perhaps the Hamiltonian operator X which generates the motion in time in his formulation is equivalent with the Hamiltonian operator H which generates the motion in time in the conventional formulation. Of course this is true for the linear problems where the two formulations are equivalent. We shall show here that it is no longer true for a large class of nonlinear problems, by computing the spectra of both X and H and showing that the spectrum of X is (apart from 0) continuous, whereas the spectrum of H is discrete.

We shall consider the class of problems for which the potential function $V(q)$ is twice differentiable and satisfies

$$V(0) = 0, \qquad V(q) > 0 \quad \text{if} \quad q > 0, \tag{4.1}$$

$$V(-q) = V(q), \tag{4.2}$$

$$V''(q) > 0. \tag{4.3}$$

It is known that under these conditions the solutions of the classical equations of motion (1.3) are all periodic, and that their orbits in the p–q plane are smooth and symmetric about both the x and y axes. Moreover, since $H(p, q)$ is a constant of the motion, the orbits are determined by the equation

$$H(p, q) = \tfrac{1}{2}p^2 + V(q) = \text{const.} \tag{4.4}$$

We shall label each orbit by its (unique) q intercept. In terms of this label, the energy of the orbit with q intercept a is $H(0, a) = V(a)$.

The period of this orbit we determine as follows: From

$$H(p, q) = \tfrac{1}{2}p^2 + V(q) = V(a) \tag{4.5}$$

we get

$$dq/dt = p = \sqrt{2}\,[V(a) - V(q)]^{\frac{1}{2}} \tag{4.6}$$

or

$$dt = \tfrac{1}{2}\sqrt{2}\,\{dq/[V(a) - V(q)]^{\frac{1}{2}}\}. \tag{4.7}$$

Hence the time required to transverse the orbit once is just

$$\tau(a) = \frac{4\sqrt{2}}{2}\int_0^a \frac{dq}{[V(a) - V(q)]^{\frac{1}{2}}}. \tag{4.8}$$

Similarly, the area enclosed by this orbit is

$$\begin{aligned} A(a) &= 4\int_0^a p\,dq \\ &= 4\sqrt{2}\int_0^a [V(a) - V(q)]^{\frac{1}{2}}\,dq. \end{aligned}\tag{4.9}$$

The dependence of the area on the intercept parameter a is illuminated by

$$\frac{dA}{da} = \frac{4\sqrt{2}}{2}\int_0^a \frac{V'(a)\,dq}{[V(a) - V(q)]^{\frac{1}{2}}} = \tau(a)V'(a). \tag{4.10}$$

We now introduce a new coordinate system (a, φ) into the p–q plane by assigning to each point (p, q) the coordinates (a, φ), where a is the q intercept of the (unique) orbit passing through (p, q), and φ is $2\pi/\tau(a)$ times the length of time it takes a point to reach (p, q) starting from $(0, a)$. This new coordinate system may be thought of as a sort of generalized polar system, in which the lines of constant a are the orbits, and the lines of constant φ are those passing through fixed fractions of the periods. The action of the motion $T(t)$ upon these

coordinates is simply

$$T(t)(a, \varphi) = [a, \varphi + 2\pi t/\tau(a)]. \tag{4.11}$$

Now we consider the subspace $\mathfrak{K}_n$ of $\mathfrak{K} = \mathfrak{L}_2(E_2)$ consisting of all those square-integrable functions of the special form

$$f(a, \varphi) = g(a)e^{in\varphi}. \tag{4.12}$$

The action of the group $W(t)$ induced by the motion upon these functions is

$$(W(t)f)(a, \varphi) = g(a)e^{in[\varphi + 2\pi t/\tau(a)]}. \tag{4.13}$$

Hence the subspace $\mathfrak{K}_n$ remains invariant under $W(t)$, and on $\mathfrak{K}_n$, $W(t)$ acts by multiplication by $\exp[2n\pi it/\tau(a)]$. Since $W(t) = e^{iXt}$, it follows that $\mathfrak{K}_n$ remains invariant under X, and on $\mathfrak{K}_n X$ acts by multiplication by $2n\pi/\tau(a)$. Hence the spectrum of X must include all numbers of the form $2n\pi/\tau(a)$.

Since X must be self-adjoint, we know that the $\mathfrak{K}_n$ are mutually orthogonal subspace of $\mathfrak{K}$. We now show that the $\mathfrak{K}_n$ actually span $\mathfrak{K}$, so that no other subspaces contribute to the spectrum of X.

For this purpose we split $V(q)$ into two parts,

$$V(q) = V_0(q) + V_1(q), \tag{4.14}$$

where

$$V_0(q) = \tfrac{1}{2}q^2, \tag{4.15}$$

and

$$V_1(q) = V(q) - V_0(q). \tag{4.16}$$

We now consider the linear problem whose Hamiltonian function is

$$H_0(p, q) = \tfrac{1}{2}p^2 + V_0(q) = \tfrac{1}{2}(p^2 + q^2). \tag{4.17}$$

All of the preceding analysis applies, of course, to this problem as well, but now the orbits are simply concentric circles. In this case it is well-known that the subspaces $\mathfrak{K}_n^0$ together span $\mathfrak{K}$.

We now proceed to define a nonsingular mapping M of the plane onto itself which carries the orbits of the linear problem onto orbits of the nonlinear problem. If $P(r, \theta)$ is any point of the plane described in ordinary polar coordinates, then we set

$$M: P(r, \theta) \to P(a, \varphi), \tag{4.18}$$

where $P(a, \varphi)$ is the image point described in generalized polar coordinates with

$$\left.\begin{aligned} a &= r \\ \varphi &= \theta \end{aligned}\right\}. \tag{4.19}$$

M is obviously one–one, invertible, and bicontinuous. Its Jacobian J may be obtained as follows: An area element enclosed by the increments dr and $d\theta$ is just $dA_0 = d\theta/2\pi \times 2\pi r\,dr$, where $2\pi r\,dr$ is the area enclosed in the annulus of width dr. This area increment is mapped by M into an area increment of the form $d\varphi/2\pi \times (dA/da)da$, where $(dA/da)da$ is the area enclosed in the annular region between nearby orbits whose q intercepts differ by da. Here we have used the fact that the fraction of the annular area swept out by an increment $d\varphi$ in φ is just $(d\varphi/2\pi)$ times the total annular area, which follows from the invariance of area under the motion. Combining these observations, we find that the Jacobian of M is given by

$$J(r, \theta) = \frac{dA}{dA_0} = \frac{\tau(r)V'(r)}{\tau_0(r)V_0'(r)} = \frac{\tau(r)V'(r)}{2\pi r}. \tag{4.20}$$

We see that under our hypotheses the Jacobian vanishes nowhere except perhaps at the origin. A similar result holds, of course, for the inverse M^{-1} of M, with Jacobian $J(a, \varphi) = 2\pi a/\tau(a)V'(a)$.

In terms of M we define a unitary mapping W of $\mathfrak{K}$ onto itself via

$$(Wf)(a, \varphi) = J(a, \varphi)^{\frac{1}{2}} f[M^{-1}(a, \varphi)]. \tag{4.21}$$

It is easy to check that W is unitary. Moreover, W carries the invariant subspaces $\mathfrak{K}_n^0$ of the linear problem onto the corresponding invariant subspaces $\mathfrak{K}_n$ of the nonlinear problem. Since the subspaces $\mathfrak{K}_n^0$ span $\mathfrak{K}$, we conclude that the subspaces $\mathfrak{K}_n$ also span $\mathfrak{K}$.

It is worth noting that we have established here a sort of completeness relation for the class of nonlinear problems under consideration which reduces to the well-known Parseval relation for the linear case.

We have shown that the spectrum of X on $\mathfrak{K}$ consists precisely of all numbers of the form $2n\pi/\tau(a)$. It is easy to see that, if $2n\pi/\tau(a)$ is constant on any set of positive measure on the q axis, then that constant lies in the point spectrum of X, with infinite multiplicity; all other numbers of this form must lie in the continuous spectrum. When $n = 0$ we obviously get the number 0 in the point spectrum. For other values of n the situation depends on the function $\tau(a)$.

In the linear case $\tau(a) \equiv$ const and the spectrum of X is discrete. Conversely, a recent paper of Levin and Schatz[3] shows that, if $\tau(a)$ is constant in any interval $0 < a < A$ under our hypothesis, then the problem must be linear. In the nonlinear case the work of Loud[4] shows that if we assume

[3] J. J. Levin and S. S. Schatz, "Nonlinear Oscillations of Fixed Period," J. Math. Anal. Appl. (to be published).

[4] W. S. Loud, Mem. Am. Math. Soc. 31 (1959).

$V''''(q) > 0$, then we have $\tau'(a) < 0$ for all $a > 0$, so that $\tau(a)$ decreases monotomically as $a \to \infty$. In the particular case that $V(q) = \frac{1}{2}q^2 + \frac{1}{4}q^4$, this conclusion can be verified directly, as follows.

$$\begin{aligned}\tau(a) &= 2\sqrt{2}\int_0^a \left(\frac{a^2 - x^2}{2} + \frac{a^4 - x^4}{4}\right)^{-\frac{1}{2}} dx \\ &= 4\int_0^a (a^2 - x^2)^{-\frac{1}{2}}\left(1 + \frac{a^2 + x^2}{2}\right)^{-\frac{1}{2}} dx \\ &= 4(1 + a^2)^{-\frac{1}{2}}\int_0^a (a^2 - x^2)^{-\frac{1}{2}} dx \\ &\quad + 4\int_0^a \left[\left(1 + \frac{a^2 + x^2}{2}\right)^{-\frac{1}{2}}\right. \\ &\quad \left. - (1 + a^2)^{-\frac{1}{2}}\right](a^2 - x^2)^{-\frac{1}{2}} dx, \qquad (4.22)\end{aligned}$$

or

$$\begin{aligned}\tau(a) &= 2\pi(1 + a^2)^{-\frac{1}{2}} + 4\int_0^a \left[\left(1 + \frac{a^2 + x^2}{2}\right)^{-\frac{1}{2}}\right. \\ &\quad \left. - (1 + a^2)^{-\frac{1}{2}}\right](a^2 - x^2)^{-\frac{1}{2}} dx. \qquad (4.23)\end{aligned}$$

Thus we see that

$$\tau(a) \to \begin{cases} 2\pi & \text{as } a \to 0, \\ 0 & \text{as } a \to \infty. \end{cases} \qquad (4.24)$$

Now the integrand in the second term in (4.23) vanishes as $x \to a$. Hence

$$\begin{aligned}\tau'(a) &= -\pi(1 + a^2)^{-\frac{1}{2}} \\ &+ 4\int_0^a \frac{\partial}{\partial a}\left\{\frac{[1 + \frac{1}{2}(a^2 + x^2)^{-\frac{1}{2}}] - (1 + a^2)^{-\frac{1}{2}}}{(a^2 - x^2)^{\frac{1}{2}}}\right\}. \qquad (4.25)\end{aligned}$$

Now

$$\begin{aligned}&\frac{\partial}{\partial a}\left\{\frac{[1 + \frac{1}{2}(a^2 + x^2)]^{-\frac{1}{2}} - (1 + a^2)^{-\frac{1}{2}}}{(a^2 - x^2)^{\frac{1}{2}}}\right\} \\ &= -a\left\{\frac{[1 + \frac{1}{2}(a^2 + x^2)]^{-\frac{1}{2}} - (1 + a^2)^{-\frac{1}{2}}}{(a^2 - x^2)^{\frac{1}{2}}}\right\} \\ &\quad - a\left\{\frac{[1 + \frac{1}{2}(a^2 + x^2)]^{-\frac{1}{2}} - (1 + a^2)^{-\frac{1}{2}}}{(a^2 - x^2)^{\frac{1}{2}}}\right\}, \qquad (4.26)\end{aligned}$$

which is < 0 for all $0 < x < a$. Hence $\tau'(a) < 0$ for all $a > 0$, and the periods of different orbits are different.

The corresponding problem in the conventional formulation requires that we determine the spectrum of the operator $H = \frac{1}{2}P^2 + V(Q)$ acting on $\mathfrak{L}_2(E_1)$. Here H is a second-order self-adjoint differential operator acting on the twice-differentiable functions in $\mathfrak{L}_2(E_1)$, and we are dealing with a singular form of the Sturm–Liouville problem. Our hypotheses on $V(q)$ imply that the operator H is in the limit-point case at $+\infty$ and $-\infty$,[5] and admits a complete set of eigenfunctions in $\mathfrak{L}_2(E_1)$. The spectrum of H is therefore always discrete, and consists of an increasing sequence of positive eigenvalues, each of multiplicity one. Under these circumstances, no part of H can be unitarily equivalent with any part of X.

5. A COMPARISON OF SCATTERING OPERATORS

It is also conceivable that, even though the Hamiltonians in the two formulations are not equivalent, perhaps the associated scattering operators are. Since all physically interesting properties of a system are reflected in the scattering operator as defined in the conventional formulation, an equivalence of scattering operators would be enough to validate Segal's procedure. We shall show here, however, that for a large class of scattering problems the two scattering operators cannot be equivalent, again by computing their spectra and showing that they are quite distinct.

We shall consider here the class of problems for which the potential function $V(q)$ is twice differentiable and satisfies

$$V(q) \geq 0; \qquad V(q) = 0 \quad \text{if} \quad |q| > 1, \qquad (5.1)$$

$$V(q) = V(-q), \qquad (5.2)$$

$$qV'(q) \leq 0 \quad \text{if} \quad q \neq 0. \qquad (5.3)$$

It is known that under these conditions the orbits of the classical equations of motion (1.3) are smooth and symmetric about both the p and q axes in the p–q plane, and that they run parallel to the q axes when $|q| > 1$. Again, these orbits are determined by (4.4).

To compute the scattering operator in Segal's formulation for a system with a potential of this type, we must compare it with the reference system obtained by setting the potential equal to zero. The orbits of the actual system coincide with the orbits of the reference system in the region $|q| > 1$, and the scattering operator essentially measures the difference in position of a point traversing the actual orbit through the region $|q| < 1$ and a point starting from the same position and traversing the corresponding reference orbit through the same region.

To be precise, we denote by $T(t)$ the group of mappings of the p–q plane onto itself determined by the development in time of the actual system,

[5] E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations* (McGraw-Hill Book Company, Inc., New York, 1955), Chap. 9.

and by $T_0(t)$ the corresponding group of mappings determined by the development of the reference system. We then form the combination

$$M(s, t) = T_0(-s)T(s - t)T_0(t). \tag{5.4}$$

If s is large and positive and t is large and negative, then $M(s, t)$ moves a point backwards along a reference orbit for a time $(-t)$, then forwards along an actual orbit for a time $(s - t)$, and then backwards along a reference orbit for a time (s). Since the actual and reference orbits coincide for $|q| > 1$, it is clear that the effect of this mapping on given point is independent of s and t if s and t are large enough. It follows that the limit

$$M = \lim_{s\to+\infty} \lim_{t\to-\infty} M(s, t) \tag{5.5}$$

exists, and determines a well-defined mapping of the p–q plane onto itself which preserves the invariant measure. Hence the mapping Z induced by M on $\mathfrak{L}_2(E_2)$,

$$(Zf)(p, q) = f[M^{-1}(p, q)], \tag{5.6}$$

is unitary. Moreover, it is easy to verify that

$$Z = \lim_{s\to+\infty} \lim_{t\to-\infty} W_0(-s)W(s - t)W_0(t), \tag{5.7}$$

where $W_0(t)$ and $W(t)$ are the unitary operators induced on $\mathfrak{L}_2(E_2)$ by $T_0(t)$ and $T(t)$, respectively, and the limits are taken in the sense of strong convergence. From (5.7) we conclude that Z is the scattering operator of the system as defined in Segal's formulation, and from (5.6) we see that Z is determined by the mapping M.

To determine M we argue as follows: We first observe that, for all s, t, and t',

$$T_0(t')M(s, t) = M(s - t', t - t')T_0(t'), \tag{5.8}$$

from which it follows that

$$T_0(t')M = MT_0(t'), \tag{5.9}$$

or

$$T_0(t')MT_0(-t') = M. \tag{5.10}$$

Hence given any point (p, q) in E_2, the effect of M on (p, q) can be computed by first applying $T_0(-t')$, then M, and then $T_0(t')$. Now $T_0(-t')$ moves every point parallel to the q axis with a velocity p. By a suitable choice of t' we can arrange it so that

$$T_0(-t')(p, q) = \begin{cases} (p, -1) & \text{if } p > 0, \\ (p, +1) & \text{if } p < 0. \end{cases} \tag{5.11}$$

Hence it is sufficient to determine the effect of M on points of the form $(p, -1)$, with $p > 0$. For such points we have

$$\begin{aligned} M(p, -1) &= M(s, t)(p, -1), \qquad t \ll 0 \ll s, \\ &= T_0(-s)T(s)(p, -1), \quad 0 \ll s, \end{aligned} \tag{5.12}$$

since the combination $T(-t)T_0(t)$ reduces to the identity for all points of the form $(p, -1)$, whenever $t < 0$.

The effect of $T_0(-s)T(s)$ on $(p, -1)$ depends on the value of p. If $\frac{1}{2}p^2 > V(0)$, then the actual orbit containing $(p, -1)$ as determined by (4.4) lies above the q axis, and we find that

$$T_0(-s)T(s)(p, -1) = T_0(-\sigma)T(\sigma)(p, -1), \tag{5.13}$$

where σ is chosen so that $T(\sigma)(p, -1) = (p, +1)$. This value of σ is easily computed from the formula (4.7),

$$\sigma = \frac{\sqrt{2}}{2} \int_{-1}^{+1} \frac{dq}{[\frac{1}{2}p^2 - V(q)]^{\frac{1}{2}}}. \tag{5.14}$$

Thus we have for $p > 0$ and $\frac{1}{2}p^2 > V(0)$,

$$\begin{aligned} M(p, -1) &= T_0(-\sigma)T(\sigma)(p, -1) \\ &= T_0(-\sigma)(p, +1) \\ &= (p, 1 - \sigma p). \end{aligned} \tag{5.15}$$

Similarly, if $p > 0$ and $\frac{1}{2}p^2 < V(0)$, then the actual orbit containing $(p, -1)$ lies to the left of the p axis, and we have

$$\begin{aligned} M(p, -1) &= T_0(-\sigma)T(\sigma)(p, -1) \\ &= T_0(-\sigma)(-p, -1) \\ &= (-p, -1 + \sigma p), \end{aligned} \tag{5.16}$$

where now σ is chosen so that $T(\sigma)(p, -1) = (-p, -1)$:

$$\sigma = \sqrt{2} \int_{-1}^{a} \frac{dq}{[\frac{1}{2}p^2 - V(q)]^{\frac{1}{2}}}. \tag{5.17}$$

Here a is the q intercept of the actual orbit. For $p < 0$ we make use of the symmetry about the q axis to obtain the same results.

In summary, we find that the effect of M on (p, q), with $p > 0$ and $q = -1 + pt'$, is given by

$$\begin{aligned} &M(p, q) \\ &= M(p, -1 + pt') \\ &= T_0(t')MT_0(-t')(p, -1 + pt') \\ &= T_0(t')M(p, -1); \\ &= \begin{cases} T_0(t')(p, 1 - \sigma p), & \text{if } \frac{1}{2}p^2 > V(0), \\ T_0(t')(-p, -1 + \sigma p), & \text{if } \frac{1}{2}p^2 < V(0); \end{cases} \end{aligned} \tag{5.18}$$

$$= \begin{cases} (p, -1 + pt' + 2 - \sigma p), & \text{if } \tfrac{1}{2}p^2 > V(0), \\ (-p, 1 - pt' - 2 + \sigma p), & \text{if } \tfrac{1}{2}p^2 < V(0); \end{cases}$$

$$= \begin{cases} (p, q + 2 - \sigma p), & \text{if } \tfrac{1}{2}p^2 > V(0), \\ (-p, -q - 2 + \sigma p), & \text{if } \tfrac{1}{2}p^2 < V(0); \end{cases}$$

From (5.18) it is easy to compute the spectrum of the scattering operator Z. For this purpose we define a unitary transformation G on $\mathfrak{L}_2(E_2)$ as follows:

$$(Gf)(p, k) = \begin{cases} (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{+\infty} e^{iqk} f(p, q)\, dq, & \tfrac{1}{2}p^2 > V(0), \\ (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{+\infty} e^{iqk} f(-p, -q)\, dq, & \tfrac{1}{2}p^2 < V(0). \end{cases} \tag{5.19}$$

A straightforward computation then shows that

$$GZG^{-1}f(p, k) = e^{i(2-\sigma p)k} f(p, k). \tag{5.20}$$

Hence Z is unitarily equivalent with the operation of multiplication by the function $e^{i(2-\sigma p)k}$. The spectrum of Z is therefore identical with the range of this function as (p, k) varies over E_2. In fact, for fixed p, the range of $e^{i(2-\sigma p)k}$, as k varies over E, is the whole unit circle. In this sense the phase shifts associated with a fixed energy in the reference system include all possible values from $-\infty$ to $+\infty$.

In the conventional formulation, the scattering operator is defined as the strong limit

$$S = \lim_{s \to +\infty} \lim_{t \to -\infty} U_0(-s) U(s, t) U_0(t), \tag{5.21}$$

where $U_0(t) = \exp(iH_0 t)$ and $U(t) = \exp(iHt)$, with $H_0 = \frac{1}{2}P^2$ and $H = \frac{1}{2}P^2 + V(Q)$. It is known that S commutes with H_0, and hence with P. Now under our hypothesis, P generates a maximal commutative set of operators, and so S is expressible as a function of P. If F denotes the ordinary Fourier transform on $\mathfrak{L}_2(E_1)$,

$$(Ff)(p) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{+\infty} e^{ipq} f(q)\, dq, \tag{5.22}$$

then $FPF^{-1}f(p) = pf(p)$, $FQF^{-1}f(p) = (1/i)(\partial/\partial p)f(p)$, and

$$FSF^{-1}f(p) = e^{2i\delta(p)} f(p), \tag{5.23}$$

where the phase shift $\delta(p)$ is a single-valued function of p satisfying $\delta(-p) = \delta(p)$. In particular, the phase shift associated with a fixed energy in the reference system is determined uniquely by the potential. A more detailed analysis, which we shall forego here, shows that $0 \leq \delta(p) \leq \frac{1}{2}\pi$, so that the spectrum of S consists of the upper half of the unit circle. Under these circumstances the scattering operators cannot be equivalent under any unitary operator which preserves the energies of the reference systems.

6. CONCLUSIONS

We have shown here that, for a large class of one-dimensional problems, the quantization procedure proposed by I. E. Segal, if taken literally, does not reduce to the conventional procedure, either in the equations of motion, or in the Hamiltonians, or in the scattering operators. The precise connection between the two procedures remains obscure. Nevertheless, the possible role of Segal's procedure in a divergenceless reformulation of the foundations of quantum field theory has not been ruled out, and further study of its possibilities seems worthwhile.

Convergent Perturbation Expansions for Certain Wave Operators

REESE T. PROSSER
Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts*
(Received 24 January 1964)

This paper establishes rigorously the validity of Dyson's perturbation expansion for the Møller wave operators under suitable restrictive assumptions on the interaction potential.

INTRODUCTION

LET H_0 and $H = H_0 + V$ be self-adjoint operators defined on a common dense domain $\mathfrak{D}$ in a Hilbert space $\mathfrak{H}$. What conditions on H_0 and V guarantee that H_0 and H are unitarily equivalent?

This question has commanded a considerable interest in recent times because of its applications in scattering theory, and partial results have been obtained by sundry methods.[1-7] In a fundamental paper on quantum electrodynamics, F. J. Dyson proposed a formal procedure for defining in terms of H_0 and V a unitary operator which implements the equivalence between H_0 and H. This procedure is based on Volterra's concept of the product integral which may be expanded in a formal perturbation series.

As far as I know, no attempt has been made to determine conditions on H_0 and V which render Dyson's procedure valid or ensure that the perturbation series converges. This paper establishes the validity of Dyson's procedure under rather stringent conditions, which nonetheless are sufficient to include some of the known results mentioned above. These conditions are not best possible; they are presented here to indicate the course of future developments rather than to encompass them.

1. THE DYSON EXPANSION

We shall assume throughout that H_0 is an essentially self-adjoint operator defined on a dense domain $\mathfrak{D}$ in a Hilbert space $\mathfrak{H}$, and that V is a bounded symmetric operator also defined on $\mathfrak{D}$. It is known that under these conditions the operator

$$H = H_0 + V \tag{1.1}$$

is also defined on $\mathfrak{D}$ and is essentially self-adjoint there.[8] Hence we may introduce the *transition operators*

$$U(s, t) = e^{-iH_0 s} e^{iH(s-t)} e^{iH_0 t}, \tag{1.2}$$

and state in terms of them our principal result:

Theorem 1.1. Suppose that V may be factored as a product AB of bounded operators A and B such that for each f in $\mathfrak{D}$ we have

(1) $||Be^{iH_0\tau}Af|| \leq K(\tau)\,||f||$,

where $K(\tau)$ is independent of f,

(2) $\int_{-\infty}^{+\infty} K(\tau)\,d\tau = \beta < 1.$

Then the transition operators $U(s, t)$ converge strongly to unitary operators $U(s, \pm\infty)$ as $t \to \pm\infty$, $U(\pm\infty, t)$ as $s \to \pm\infty$, and $U(\pm\infty, \pm\infty)$ as $s, t \to \pm\infty$.

Proof: First, observe that, for any f in $\mathfrak{H}$ and any real scalar ρ, the combination $U(s, t)e^{-iH_0\rho}Af$ is strongly differentiable in t, and that

$$(d/dt)U(s, t)e^{-iH_0\rho}Af = (-i)U(s, t)e^{-iH_0 t}ABe^{+iH_0(t-\rho)}Af. \tag{1.3}$$

The hypotheses imply that the operator on the right-hand side of (1.3) is bounded and strongly continuous in t. Hence, both sides are strongly integrable in t, and it follows that

$$U(s, t)e^{-iH_0\rho}Af = e^{-iH_0\rho}Af + (-i)\int_t^s U(s, \tau)e^{-iH_0\tau}ABe^{+iH_0(\tau-\rho)}Af\,d\tau. \tag{1.4}$$

The conditions under which (1.4) is obtained allow us to iterate this procedure, substituting for $U(s, t)$ on the right-hand side of (1.4) its value taken from the left. After a single iteration, we get

* Operated with support from the U. S. Air Force.

[1] J. M. Cook, J. Math. Phys. **36**, 82 (1957).
[2] K. O. Friedrichs, Commun. Pure Appl. Math. **1**, 361 (1948).
[3] T. Ikebe, Operators, Arch. Ratl. Mech. Anal. **5**, 1 (1960).
[4] T. Kato, Proc. Japan Acad. **33**, 260 (1957).
[5] S. T. Kuroda, Nuovo Cimento **12**, 431 (1959).
[6] S. T. Kuroda, J. Math. Soc. Japan **11** 247 (1959).
[7] J. Schwartz, Commun. Pure Appl. Math. **13**, 609 (1960).
[8] T. Kato, Trans. Amer. Math. Soc. **70**, 195 (1951).

$$U(s,t)e^{-iH_0\rho}Af = e^{-iH_0\rho}Af + (-i)\int_t^s e^{-iH_0\tau}ABe^{+iH_0(\tau-\rho)}Af\,d\tau + (-i)^2\int_t^s\int_{\tau_1}^s U(s,\tau_2)e^{-iH_0\tau_2} \times ABe^{+iH_0(\tau_2-\tau_1)}ABe^{+iH_0(\tau_1-\rho)}Af\,d\tau_2\,d\tau_1. \quad (1.5)$$

Similarly, after an n-fold iteration, we get

$$U(s,t)e^{-iH_0\rho}Af = e^{-iH_0\rho}Af + \sum_{k=1}^{n}(-i)^k\int_t^s\int_{\tau_1}^s\cdots\int_{\tau_{k-1}}^s e^{-iH_0\tau_k}ABe^{+iH_0(\tau_k-\tau_{k-1})} \times AB\cdots ABe^{+iH_0(\tau_1-\rho)}Af\,d\tau_k\cdots d\tau_1 + (-i)^{n+1}\int_t^s\int_{\tau_1}^s\cdots\int_{\tau_n}^s U(s,\tau_{n+1})e^{-iH_0\tau_{n+1}} \times ABe^{+iH_0(\tau_{n+1}-\tau_n)}AB\cdots \times ABe^{+iH_0(\tau_1-\rho)}Af\,d\tau_{n+1}\cdots d\tau_1. \quad (1.6)$$

Now observe that the integrand in the kth term is bounded,

$$\|e^{-iH_0\tau_k}ABe^{+iH_0(\tau_k-\tau_{k-1})}AB\cdots ABe^{+iH_0(\tau_1-\rho)}Af\| \le \|A\|\,K(\tau_k-\tau_{k-1}) \times K(\tau_{k-1}-\tau_{k-2})\cdots K(\tau_1-\rho)\,\|f\|, \quad (1.7)$$

and hence the kth term itself is bounded,

$$\|k\text{th term}\| \le \|A\|\int_t^s\int_{\tau_1}^s\cdots\int_{\tau_{k-1}}^s K(\tau_k-\tau_{k-1})\cdots K(\tau_1-\rho)\,d\tau_k\cdots d\tau_1\,\|f\| \le \|A\|\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}K(\tau_k-\tau_{k-1})\cdots K(\tau_1-\rho)\,d\tau_k\cdots d\tau_1\,\|f\| \le \|A\|\int_{-\infty}^{+\infty}(K*K*\cdots*K)(\tau_k-\rho)\,d\tau_k\,\|f\| \le \|A\|\,\beta^k\,\|f\|. \quad (1.8)$$

In the same way, the remainder term is bounded by $\|A\|\,\beta^{n+1}\,\|f\|$. Note that these bounds are independent of s and t.

From (1.8) it follows that the right-hand side of (1.6) without the remainder term converges strongly to the left-hand side as $n \to \infty$. Thus we may write

$$U(s,t)e^{-iH_0\rho}Af = e^{-iH_0\rho}Af + \sum_{k=1}^{\infty}(-i)^k\int_t^s\int_{\tau_1}^s\cdots\int_{\tau_{k-1}}^s e^{-iH_0\tau_k}ABe^{+iH_0(\tau_k-\tau_{k-1})} \times AB\cdots ABe^{+iH_0(\tau_1-\rho)}Af\,d\tau_k\cdots d\tau_1. \quad (1.9)$$

Moreover, the convergence is uniform in both s and t.

It also follows from (1.8) by standard arguments that the kth term converges strongly as $s \to \pm\infty$ or as $t \to \pm\infty$. It follows from this and the uniform convergence of (1.9) that $U(s,t)e^{-iH_0\rho}Af$ converges strongly as $s \to \pm\infty$ or as $t \to \pm\infty$, and that the limits may be computed term by term in the series in (1.9). Moreover, these results hold for any f in $\mathfrak{H}$.

Now denote by $\mathfrak{M}$ the closed linear subspace of $\mathfrak{H}$ spanned by all elements of the form $e^{iH_0\rho}ABf$, for any f in $\mathfrak{H}$ and any real scalar ρ. It follows from the arguments presented above that $U(s,t)$ converges strongly as $s \to \pm\infty$ or $t \to \pm\infty$ on a dense subset of $\mathfrak{M}$, and hence, by uniform continuity, on all of $\mathfrak{M}$. Moreover, if f lies in $\mathfrak{M}$, then clearly $e^{-iH_0\tau}f$ lies in $\mathfrak{M}$ for all real scalars τ, and so $\mathfrak{M}$ completely reduces $e^{-iH_0\tau}f$ for each τ, and hence $\mathfrak{M}$ completely reduces H_0. Similarly, if f lies in $\mathfrak{H}$, then Vf lies in $\mathfrak{M}$ and so $\mathfrak{M}$ completely reduces V, and on $\mathfrak{M}^\perp$, $V = 0$.

It follows that $\mathfrak{M}$ completely reduces H. On $\mathfrak{M}$ we have established the existence of the required limits, and on $\mathfrak{M}^\perp$ we have shown that $V = 0$ and hence that $H = H_0$. On $\mathfrak{M}^\perp$, then, $U(s,t) \equiv I$. We conclude that the required limits exist on all of $\mathfrak{H}$, and are given by the limiting forms of (1.9) on $\mathfrak{M}$ and by I on $\mathfrak{M}^\perp$.

It remains to show that these limits are all unitary. We first observe that as strong limits of unitary operators they are necessarily isometries. It remains to show that their ranges span all of $\mathfrak{H}$. For this purpose, consider the adjoint operators $U(s\ t)^* = U(t,s)$. Now the arguments that led to the expansion (1.9) of $U(s,t)$ apply equally well to $U(t,s)$ and yield a similar expansion with the roles of s and t interchanged. We conclude in the same way that adjoints $U(s,t)^* = U(t,s)$ converge strongly to isometries as $s \to \pm\infty$ or $t \to \pm\infty$. Moreover, these limits must be the adjoints of the corresponding limits of the operators $U(s,t)$. But if an operator and its adjoint are both isometries, then they must both be unitary. This argument concludes our proof.

The expansion (1.9) may now be rewritten as

$$U(s,t)f = f + \sum_{k=1}^{\infty}\frac{(-i)^k}{k!}\int_t^s\int_{\tau_1}^s\cdots\int_{\tau_{k-1}}^s e^{-iH_0\tau_k} \times Ve^{iH_0(\tau_k-\tau_{k-1})}V\cdots Ve^{iH_0\tau_1}f\,d\tau_k\,d\tau_{k-1}\cdots d\tau_1. \quad (1.10)$$

We have shown that this expansion converges strongly to $U(s,t)$ for any f in $\mathfrak{H}$, and that under

our assumptions on V the strong limits of $U(s, t)$ as $s, t \to \pm\infty$ may be computed from this expansion, term by term.

It is possible to rewrite this expansion in a more symmetrical form through the use of time-ordered products.[9] If we define

$$V(\tau) = e^{-iH_0\tau} V e^{iH_0\tau}, \tag{1.11}$$

and define the time-ordered product

$$T\{V(\tau_1)V(\tau_2)\cdots V(\tau_n)\} = V(\tau_{\pi(1)})V(\tau_{\pi(2)})\cdots V(\tau_{\pi(n)}), \tag{1.12}$$

where π is any permutation of the integers for which $\tau_{\pi(1)} \geq \tau_{\pi(2)} \geq \cdots \geq \tau_{\pi(n)}$, then we may symmetrize the domain of integration in each term of (1.9), and obtain the Dyson expansion[9]

$$U(s, t)f = f + \sum_{k=1}^{\infty} \frac{(-i)^k}{k!} \int_t^s \int_t^s \cdots \int_t^s T\{V(\tau_k) \times V(\tau_{k-1})\cdots V(\tau_1)\} f\, d\tau_k \cdots d\tau_1. \tag{1.13}$$

In this form the exponential character of the transition operators is placed in evidence.

Corollary 1.2. The limiting operators $U(0, \pm\infty)$ *and* $U(\pm\infty, \pm\infty)$ *satisfy the following relations:*

$$U(0, \pm\infty)H_0 = HU(0, \pm\infty), \tag{1.14}$$

$$U(\pm\infty, 0)H = H_0U(\pm\infty, 0), \tag{1.15}$$

$$U(\pm\infty, \pm\infty)H_0 = H_0U(\pm\infty, \pm\infty). \tag{1.16}$$

Proof: From (1.2) it follows that for all real scalars τ we have $U(0, t)e^{iH_0\tau} = e^{iH\tau}U(0, t + \tau)$. Hence $U(0, \pm\infty)e^{iH_0\tau} = e^{iH\tau}U(0, \pm\infty)$, which implies (1.14). Similar computations will establish (1.15) and (1.16).

Theorem 1.1 and its corollary give a precise formulation of the existence and unitary properties of the wave and scattering operators under the restrictive assumptions made on the potential operator V.

2. THE MØLLER EXPANSION

The computation of the individual terms of the Dyson expansion (1.10) is best carried out in terms of the eigenfunctions of H_0. This computation leads to an alternate form of the Dyson expansion first proposed by Møller.[10] We derive here a precise formulation of Møller's expansion under the same assumptions made in the previous section on the potential V.

[9] F. J. Dyson, Phys. Rev. **75**, 486 (1949).

[10] C. Møller, Kgl. Danske Videnskab. Selskab. Mat.-Fys. Medd. **23**, No. 1. (1945).

We first recall that the unitary operator e^{iH_0t} may always be expressed in the form

$$e^{iH_0t} = \int_{-\infty}^{+\infty} e^{i\lambda t}\, dE_0(\lambda), \tag{2.1}$$

where $dE_0(\lambda)$ is the spectral measure associated with the generator H_0.[11] This spectral measure plays the role of an eigenfunction expansion for H_0, and may be represented in any of the following ways:[12]

$$\begin{aligned} dE_0(\lambda) &= \lim_{\eta\to 0} \frac{1}{2\pi i}\left[\frac{1}{H_0 - \lambda + i\eta} - \frac{1}{H_0 - \lambda - i\eta}\right] d\lambda \\ &= \lim_{\eta\to 0} \frac{1}{\pi}\left[\frac{\eta}{(H_0 - \lambda)^2 + \eta^2}\right] d\lambda \\ &= \delta(H_0 - \lambda)\, d\lambda. \end{aligned} \tag{2.2}$$

We shall use the form $\delta(H_0 - \lambda)d\lambda$ as a shorthand notation for any of the forms in (2.2).

We now set out to evaluate the Dyson expansion (1.10) in terms of the spectral measure (2.2). Introducing (2.1) into the left-hand side of (1.10), we obtain for the wave operator $U(0, t)$ the form

$$U(0, t)f = \int_{-\infty}^{+\infty} e^{-i(H-\lambda)t}\delta(H_0 - \lambda)f\, d\lambda. \tag{2.3}$$

Multiplying by $\epsilon e^{\epsilon t}$ and integrating from $-\infty$ to 0, we obtain

$$\begin{aligned} &\epsilon \int_{-\infty}^{0} U(0, t)fe^{\epsilon t}\, dt \\ &= \epsilon \int_{-\infty}^{0}\int_{-\infty}^{+\infty} e^{-i(H-\lambda)t}\delta(H_0 - \lambda)f\, d\lambda e^{\epsilon t}\, dt. \end{aligned} \tag{2.4}$$

Since the measures $\delta(H_0 - \lambda)d\lambda$ and $e^{\epsilon t}dt$ are both finite, and the integrand is jointly continuous in λ and t, the order of integration may be interchanged. Thus we obtain

$$\begin{aligned} &\epsilon \int_{-\infty}^{0} U(0, t)fe^{\epsilon t}\, dt \\ &= \epsilon \int_{-\infty}^{+\infty}\int_{-\infty}^{0} e^{-i(H-\lambda+i\epsilon)t}\, dt\, \delta(H_0 - \lambda)f\, d\lambda \\ &= \int_{-\infty}^{+\infty} \frac{i\epsilon}{H - \lambda + i\epsilon}\, \delta(H_0 - \lambda)f\, d\lambda. \end{aligned} \tag{2.5}$$

If we introduce (2.1) into the right-hand side of (1.10), multiply by $\epsilon e^{\epsilon t}$ and integrate from $-\infty$ to 0, we find, after a laborious term-by-term com-

[11] M. H. Stone, *Linear Transformations in Hilbert Space* (American Mathematics Society, Providence, Rhode Island, 1932).

[12] N. Dunford and J. Schwartz, *Linear Operators* (Interscience Publishers, Inc., New York, 1958), Vol. I.

putation, the form

$$\epsilon \int_{-\infty}^{0} U(0, t)fe^{\epsilon t}\, dt = f + \int_{-\infty}^{+\infty} \frac{1}{H_0 - \lambda + i\epsilon} \times V \frac{i\epsilon}{H_0 - \lambda + i\epsilon}\, \delta(H_0 - \lambda)f\, d\lambda + \cdots + \int_{-\infty}^{+\infty} \frac{1}{H_0 - \lambda + i\epsilon} V \frac{1}{H_0 - \lambda + i\epsilon} V \cdots \times V \frac{i\epsilon}{H_0 - \lambda + i\epsilon}\, \delta(H_0 - \lambda)f\, d\lambda + \cdots . \quad (2.6)$$

This form may also be obtained directly from (2.5) by expanding the integrand there in powers of V.

These formulas all hold for all $\epsilon > 0$. Moreover, we have shown in Sec. 1 that the right-hand side of the expansion (1.10) converges strongly at a rate which depends on f, but not on s or t. It follows that the expansion (2.6) converges strongly at a rate which depends on f, but not on ϵ.

What happens as $\epsilon \to 0$? We recall that, if $U(0, t)$ converges strongly as $t \to -\infty$, then $\epsilon \int_{-\infty}^{0} U(0, t)e^{\epsilon t}\, dt$ also converges strongly as $\epsilon \to 0$, and both expressions converge to the same limit.[12] Thus we have

$$\lim_{\epsilon \to 0} \epsilon \int_{-\infty}^{0} U(0, t)fe^{\epsilon t}\, dt = U(0, -\infty)f. \quad (2.7)$$

Combining (2.7) with (2.5), we see that $U(0, -\infty)$ may be expressed as

$$U(0, -\infty)f = \lim_{\epsilon \to 0} \int_{-\infty}^{+\infty} \frac{i\epsilon}{H - \lambda + i\epsilon}\, \delta(H_0 - \lambda)f\, d\lambda. \quad (2.8)$$

This formula for the wave operator $U(0, -\infty)$ was first proposed by Lippman and Schwinger.[13]

Combining (2.7) with (2.6), and putting $V = AB$, we obtain

$$U(0, -\infty)f = f + \lim_{\epsilon \to 0} \int_{-\infty}^{+\infty} \frac{1}{H_0 - \lambda + i\epsilon} \times AB \frac{i\epsilon}{H_0 - \lambda + i\epsilon}\, \delta(H_0 - \lambda)f\, d\lambda + \cdots + \lim_{\epsilon \to 0} \int_{-\infty}^{+\infty} \frac{1}{H_0 - \lambda + i\epsilon} AB \frac{1}{H_0 - \lambda + i\epsilon} AB \cdots \times AB \frac{i\epsilon}{H_0 - \lambda + i\epsilon}\, \delta(H_0 - \lambda)f\, d\lambda + \cdots . \quad (2.9)$$

Here we have used the fact that the rate of convergence of the expansion (2.6) is independent of ϵ to compute the limit, term by term.

We now observe that, under our assumptions on V, the combination $B[1/(H_0 - \lambda + i\epsilon)]A$ converges uniformly as $\epsilon \to 0$. In fact, we have

$$\lim_{\epsilon \to 0} B \frac{1}{H_0 - \lambda + i\epsilon} Af = \lim_{\epsilon \to 0} i \int_0^{\infty} Be^{i(H_0 - \lambda)t} Afe^{-\epsilon t}\, dt = i \int_0^{\infty} Be^{i(H_0 - \lambda)t} Af\, dt, \quad (2.10)$$

with the convergence guaranteed by the first hypothesis of Theorem 1.1.

We also note that the combination

$$[i\epsilon/(H_0 - \lambda + i\epsilon)]\delta(H_0 - \lambda)f$$

reduces to $\delta(H_0 - \lambda)f$. This result follows immediately from the properties of the spectral measure given in (2.2).

Combining now (2.9) and (2.10), we finally obtain the expansion

$$U(0, -\infty)f = f + \int_{-\infty}^{+\infty} \frac{1}{H_0 - \lambda + i0} V\delta(H_0 - \lambda)f\, d\lambda + \cdots + \int_{-\infty}^{+\infty} \frac{1}{H_0 - \lambda + i0} V \frac{1}{H_0 - \lambda + i0} V \cdots \times V\delta(H_0 - \lambda)f\, d\lambda + \cdots . \quad (2.11)$$

This expansion for the wave operator $U(0, -\infty)$ was first proposed by Møller in Ref. 10. We have now established its validity under our assumptions on the potential V.

Similar arguments will also establish the following expansions for $U(0, +\infty)$ and $U(+\infty, -\infty)$:

$$U(0, +\infty)f = f + \int_{-\infty}^{+\infty} \frac{1}{H_0 - \lambda - i0} V\delta(H_0 - \lambda)f\, d\lambda + \cdots + \int_{-\infty}^{+\infty} \frac{1}{H_0 - \lambda - i0} V \frac{1}{H_0 - \lambda - i0} V \cdots \times V\delta(H_0 - \lambda)f\, d\lambda + \cdots , \quad (2.12)$$

$$U(+\infty, -\infty)f = f + \int_{-\infty}^{+\infty} 2\pi\, \delta(H_0 - \lambda)V\, \delta(H_0 - \lambda)f\, d\lambda + \cdots + \int_{-\infty}^{+\infty} 2\pi\, \delta(H_0 - \lambda)V \frac{1}{H_0 - \lambda + i0} V \cdots \times V\, \delta(H_0 - \lambda)f\, d\lambda + \cdots . \quad (2.13)$$

[13] B. A. Lippman and J. Schwinger, Phys. Rev. 79, 469 (1950).

3. AN EXAMPLE

To show that the hypotheses of Theorem 1.1 are applicable, we present here a simple example.

Let $\mathfrak{H}$ be $L_2(E_3)$, the space of square-integrable functions on E_3, and let $\mathfrak{D}$ be the subspace of all functions $f \in L_2$ such that $\nabla^2 f$ exists in a mean-square sense and lies in L_2. Define H_0 and V on $\mathfrak{D}$ by

$$\begin{aligned}(H_0 f)(x) &= (-\nabla^2 f)(x),\\ (Vf)(x) &= V(x)f(x),\end{aligned} \tag{3.1}$$

for all f in $\mathfrak{D}$, where $V(x)$ is a real-valued, bounded, measurable function subject to conditions described below. It is known that under these conditions both H_0 and $H = H_0 + V$ are essentially self-adjoint in $\mathfrak{D}$.[8]

Theorem 3.1. Let $M = max\ \{||V||_1, ||V||_\infty\}$. *Then the hypotheses of theorem* 1 *hold for* H_0 *and* V *of* (3.1) *provided that* $M < \frac{2}{3}\pi$.

We begin with a Lemma describing the unperturbed unitary operator $e^{iH_0\tau}$.

Lemma 3.2. If $f \in L_1$ *and* L_2, *then*

$$(e^{iH_0\tau}f)(x) = (4\pi\tau)^{-\frac{3}{2}} \int_{E_3} \exp^{-i(x-y)^2/4\tau} f(y)\, dy. \tag{3.2}$$

Proof: Let us put for a moment $t = \tau + i\epsilon$, and recall that $e^{iH_0 t}f$ is defined for all $f \in L_2$, and is determined via the Plancherel relations by the formula

$$(e^{iH_0 t}f)\hat{}(y) = e^{it y^2} \hat{f}(y), \tag{3.3}$$

where $\hat{f}$ denotes the Fourier transform of f. Taking inverse Fourier transforms, we find

$$e^{iH_0 t}f = G_t * f, \tag{3.4}$$

with G_t given by

$$G_t(x) = (4\pi t)^{-\frac{3}{2}} e^{-ix^2/4t}. \tag{3.5}$$

For $\epsilon > 0$, $G_t(x)$ is integrable, and the convolution in (3.4) is well-defined. A standard bounded convergence argument in L_2 now allows us to take limits on both sides of (3.4) as $\epsilon \to 0$. The result is (3.2).

Next we turn to the perturbation V. Defined

$$A(x) = |V(x)|^{\frac{1}{2}}, \tag{3.6}$$

$$B(x) = \operatorname{sgn} V(x)\, |V(x)|^{\frac{1}{2}}. \tag{3.7}$$

Then $V(x) = A(x)B(x)$, $A(x)$ and $B(x)$ are all bounded, and $M = \max\{||A||_2\, ||B||_2, ||A||_\infty\, ||B||_\infty\}$.

Proof of Theorem 3.1: We have factored V as a product AB, with A and B given by multiplication by $A(x)$ and $B(x)$, respectively. We must now verify that conditions (1) and (2) of Theorem 1.1 hold. Consider any $f \in L_2$ and note that $A(x)f(x) \in L_1$ and L_2. Hence Lemma 3.2 applies, and

$$\begin{aligned}&(e^{iH_0\tau}Af)(x)\\ &\quad = (4\pi\tau)^{-\frac{3}{2}} \int_{E_3} e^{-i(x-y)^2/4\tau} A(y)f(y)\, dy.\end{aligned} \tag{3.8}$$

It follows that $e^{iH_0\tau}Af$ is in both L_∞ and L_2, and

$$||e^{iH_0\tau}Af||_2 \le ||A||_\infty\, ||f||_2, \tag{3.9a}$$

$$||e^{iH_0\tau}Af||_\infty \le (4\pi\tau)^{-\frac{3}{2}}\, ||A||_2\, ||f||_2. \tag{3.9b}$$

Thus $Be^{iH_0\tau}Af$ is well-defined and in L_2, and

$$||Be^{iH_0\tau}Af||_2 \le ||B||_\infty\, ||A||_\infty\, ||f||_2, \tag{3.10a}$$

$$||Be^{iH_0\tau}Af||_2 \le (4\pi\tau)^{-\frac{3}{2}}\, ||B||_2\, ||A||_2\, ||f||_2. \tag{3.10b}$$

Combining (3.10a) and (3.10b), we obtain

$$||Be^{iH_0\tau}Af||_2 \le K(\tau)\, ||f||_2, \tag{3.11}$$

where

$$K(\tau) = \begin{cases} M & \text{if } |\tau| \le 1/4\pi,\\ M(4\pi\tau)^{-\frac{3}{2}} & \text{if } |\tau| \ge 1/4\pi.\end{cases} \tag{3.12}$$

In particular, $K(\tau)$ is independent of f.

Finally, $K(\tau)$ is integrable in τ, and

$$\int_{-\infty}^{+\infty} K(\tau)\, d\tau = \tfrac{3}{2}(M/\pi) < 1 \quad \text{if} \quad M < \tfrac{2}{3}\pi. \tag{3.13}$$

We have shown that conditions (1) and (2) of Theorem 1.1 hold in our example, and this concludes the proof of Theorem 3.1.

Minor changes in proof will show that it is enough to assume that $V \in L_p \cap L_\infty$ for any $p < 3$, provided that $M = \max\{||V||_p, ||V||_\infty\} < \text{const}$, where the constant now depends on p.

4. DISCUSSION

The operators $U(0, -\infty)$, $U(0, +\infty)$, and $U(\infty, -\infty)$ are interpreted in scattering theory as the incoming wave operator, the outgoing wave operator, and the scattering operator, respectively.[9] Theorem 1.1 and its corollary give conditions on H_0 and V which ensure that these operators are mathematically well-defined and are unitary. Theorem 3.1 shows that these conditions are verifiable in certain simple cases, but they are not generally applicable in this form to the scattering problems of quantum field theory. The principal advantage of Dyson's procedure, of course, is that the expansion formulas for the wave and scattering

operators involve only e^{iH_0t} and V, which are relatively easy to evaluate in applications. The principal difficulty, on the other hand, is that this procedure requires the convergence of the expansions, and hence requires that the perturbation V be small.

It is known, for instance, that in some cases the wave operators are well-defined isometries, but are not unitary because of the presence of extraneous eigenfunctions of the perturbed operator H (i.e., "bound states"). This situation occurs in our example if the perturbation V is everywhere negative and sufficiently large.[5] These cases are excluded in Theorem 1.1 by the second hypothesis, which requires that the perturbation be small enough so that no bound states can occur. In the presence of bound states, Dyson's perturbation expansion for the wave operators cannot converge strongly on $\mathfrak{H}$.

The idea of splitting the potential V into two factors was introduced by J. Schwartz,[7] who used it to obtain the spectral decomposition for the total Hamiltonian H. Our results are equivalent to his when the hypotheses of both are satisfied. The same idea has been employed by Grossman and Wu[14] to study the analytic properties of the matrix elements in the Møller expansion. We are indebted to the referee for calling their paper[14] to our attention.

[14] A. Grossman and T. T. Wu, J. Math. Phys. **2**, 710 (1961); **3**, 684 (1962).

The Necessary and Sufficient Condition in Terms of Wightman Functions for a Field to be a Generalized Free Field*

W. GARCZYNSKI†
Department of Physics, McGill University, Montreal, Canada
(Received 10 January 1964)

The description of the generalized free field in terms of the Wightman functions is given. The necessary and sufficient condition for a field to be a generalized free field is established.

1. INTRODUCTION

RECENTLY Greenberg and Licht[1] have shown, within the framework of the L.S.Z. formalism,[2] that a quantum field theoretical model in which the truncated Wightman functions vanish beyond some order does not contain any interaction, i.e., $S = 1$, $A_{in}(x) = A_{out}(x)$.

Within the framework of the Wightman scheme[3] it is possible to get such a result, namely, that the necessary and sufficient condition for a field to be a generalized free field is the vanishing of all truncated Wightman functions except for the two-point function.

In this paper we give a description of the generalized free field in terms of the corresponding Wightman functions. This complements the analysis, in terms of the field operators, by Greenberg[4] and others.[5]

For simplicity, we shall consider only the theory of a single, scalar field. The generalization to the case of several fields is straightforward.

2. POSTULATES AND DEFINITIONS

We consider the theory which satisfies the following axioms[6]:

(1) The manifold of states of the system is represented by a separable Hilbert space H with a positive-definite metric. There exists in H a linear, countable and dense set D; $\bar{D} = H$.

(2) There exists in D the unitary representation $U(a, \Lambda)$ of the inhomogeneous proper Lorentz group such that $U(a, \Lambda)D \subset D$.

(3) The spectrum of the translation operators P_μ, defined by $U(a, 1) = \exp i(P_0 a_0 - \mathbf{P}\cdot\mathbf{a})$,[7] is contained in the forward light cone.

(4) In D there exists a one-dimensional subspace $\{\alpha\psi_0\}$ which is invariant under $U(a, \Lambda)$; ψ_0 is the vacuum state.

(5) In D the field $A(f)$ is given, i.e., is given an operator-valued tempered distribution and $A(f)D \subset D$ and also $A(f)^+ = A(f^*)$ is given in D.

(6) The field operator has the transformation property on D,

$$U(a, \Lambda)A(f)U^\dagger(a, \Lambda) = A(f'),$$

$$f'(x) = \begin{cases} f[\Lambda^{-1}(x - a)] & \text{for} \quad \Lambda \in L_+^+(R), \\ f^*[\Lambda^{-1}(x - a)] & \text{for} \quad \Lambda \in L_+^-(R). \end{cases}$$

[8]

(7) The field operator $A(f)$ is local on D.

(8) The field $A(f)$ is complete, i.e., each vector from D is cyclic with respect to an algebra generated by $A(f)$.

Definition[9]: We call a real, scalar, local field $A(f)$ the generalized free field if it satisfies the postulates (1)–(8) and has the property

$$[A(f), A(g)]_- = i\int_0^\infty \rho(u)\,du \int d^4x\, d^4y f(x) \times \Delta(x - y; u)g(y)\cdot\mathbb{1} \quad \text{on} \quad D; \qquad f, g \in S_4 \tag{1}$$

$$\Delta(x; u) = \Delta^-(x; u) - \Delta^-(-x; u), \tag{2}$$

---

* Supported by the National Research Council of Canada and the World University Service Council of Canada.

† On leave of absence from the Institute of Theoretical Physics, University of Wroclaw, Wroclaw, Poland.

[1] O. W. Greenberg and A. L. Licht, J. Math. Phys. **4**, 613 (1963).

[2] H. Lehmann, K. Symanzik, and W. Zimmermann, Nuovo Cimento **1**, 205 (1955); **6**, 319 (1957).
V. Glaser, H. Lehmann and W. Zimmermann, Nuovo Cimento **6**, 1122 (1957).

[3] A. S. Wightman, Phys. Rev. **101**, 860 (1956); Lectures presented at a Theoretical Physics summer seminar sponsored by International Atomic Energy Agency, Trieste, 1962 (International Atomic Energy Agency, Vienna, 1963). See also, H. Araki, Progr. Theoret. Phys. (Kyoto) Suppl. **18**, 83 (1961).

[4] O. W. Greenberg, Ann. Phys. (N. Y.) **16**, 158 (1961).

[5] A. L. Licht and J. S. Toll, Nuovo Cimento **21**, 346 (1961); G. F. Dell'Antonio, J. Math. Phys. **2**, 759 (1961), see also Chap. IV in Ref. 6.

[6] R. Haag and B. Schroer, J. Math. Phys. **3**, 248 (1962).

[7] The signature of the metric is $(+ - - -)$.

[8] The symbol $L_+^+(R)$ means the real (R) Lorentz transformation (L) with positive determinant $(_+)$ and preserving the direction of the time axis $(^+)$ $L_+^-(R)$ means the transformation which inverts $(^-)$ the direction of the time axis.

[9] See Ref. 4, Sec. B.

$$\Delta^-(x;u) = -\frac{i}{(2\pi)^3}\int \exp(-ikx) \times \theta(k_0)\delta(k^2-u^2)\,d^4k. \quad (3)$$

The function $\Delta^-(x;u)$ is a tempered distribution, continuous with respect to the parameter u, and may be understood as an improper limit[10] of the regularized, according to the Pauli–Villars function reg $\Delta^-(x;u)$, when the regularizing masses go to infinity.[11] The function $\rho(u)$ is the Källen–Lehmann measure, about which we assume

$$\rho(u) = \delta(u-m^2) + \sigma(u), \quad (4)$$

where $\sigma(u)$ is nonnegative and vanishes below $u = (2m)^2$. That is, we assume that there exists a positive lowest mass m in the theory. Then the irreducible zero-mass representations of the inhomogeneous proper Lorentz group does not occur in H.[12] The spectrum of P_μ is contained in $\overline{V_m^+}$[13] apart from the eigenvalue $P = 0$ corresponding to the vacuum.

Integration over the parameter u with measure $\rho(u)$ gives again a temperate distribution if $\sigma(u)$ does not increase too quickly.

Without loss of generality we can adopt the convention that the vacuum expectation value of the field $A(f)$ is zero.

3. CONDITION UNDER WHICH WIGHTMAN FUNCTIONS DESCRIBE A GENERALIZED FREE FIELD

Theorem. The necessary and sufficient condition for the field $A(f)$ to be generalized free field is the vanishing of all the truncated Wightman functions of this field except for the two-point function $\langle x_1, x_2\rangle$,[14] which is of the form

$$\langle x_1, x_2\rangle_T = i\int_0^\infty \rho(u)\,du\Delta^-(x_1-x_2;u). \quad (5)$$

Proof: It follows from the definition of the generalized free field and the postulates that[15]

$$\langle x_1\rangle = 0,$$

$$\langle x_1, x_2\rangle = i\int_0^\infty \rho(u)\,du\Delta^-(x_1-x_2;u), \quad (6)$$

$$\langle x_1, \cdots, x_n\rangle = \sum_{j=2}^{n}\langle x_1, x_j\rangle\langle x_2, \cdots, x_{j-1}, x_{j+1}, \cdots, x_n\rangle.$$

Hence, the Wightman functions of such field are

$$\langle x_1, \cdots, x_n\rangle = \begin{cases} 0 \quad \text{for } n \text{ odd}, \\ \sum\limits_{\substack{i<j \\ k<l}} \langle x_1, x_k\rangle \cdots \langle x_i, x_j\rangle \cdots \\ \quad \times \langle x_k, x_l\rangle \quad \text{for } n \text{ even}. \end{cases} \quad (7)$$

The general definition of truncated Wightman functions for scalar fields is[16]

$$\langle x_1, \cdots, x_n\rangle = \langle x_1, \cdots, x_n\rangle_T + \sum_{\text{div}} \langle x_1, \cdots\rangle_T \cdots \times \langle x_k, \cdots\rangle_T; \quad (n = 1, 2, \cdots), \quad (8)$$

where summation goes over all divisions of the set $(1, \cdots, n)$ into several subsets in such a way that the order of subscripts within each subset is the natural one. The above formulae together with the condition that $\langle x_1\rangle_T = \langle x_1\rangle$ define the truncated Wightman functions for all n.

The application of this definition to our particular case gives

$$\langle x_1\rangle_T = \langle x_1\rangle = 0,$$

$$\langle x_1, x_2\rangle_T = \langle x_1, x_2\rangle \neq 0, \quad \text{[see Eq. (6)]}$$

$$\langle x_1, \cdots, x_n\rangle_T + \sum_{\text{div}} \langle x_1, \cdots\rangle \cdots \langle x_k, \cdots\rangle_T = 0 \quad \text{for } n \text{ odd}, \quad (9)$$

$$\langle x_1, \cdots, x_n\rangle_T + \sum_{\text{div}}{}' \langle x_1, \cdots\rangle \cdots \langle x_k, \cdots\rangle_T = 0 \quad \text{for } n \text{ even},$$

where the prime stands for the elimination of all divisions which correspond to products of two-point functions only.

From the first condition we deduce that all truncated Wightman functions of odd order vanish identically. Then the second condition implies that, apart from the two-point functions, all functions of even order vanish.

Thus we get the necessary condition for a generalized free field

$$\langle x_1, \cdots, x_n\rangle_T = 0 \qquad n \neq 2,$$
$$\langle x_1, x_2\rangle_T = i\int_0^\infty \rho(u)\,du\Delta^-(x_1-x_2;u). \quad (10)$$

Sufficiency of these conditions follows directly by expressing the Wightman functions in terms of truncated vacuum expectation values, [see Eq. (8)], which in our case yields (7), and from Greenberg's work[4] which shows that such Wightman functions follow from the definition of the generalized free field.

[10] The limit, in the sense of the distribution theory.

[11] N. N. Bogoliubov and D. V. Shirkov, *Introduction to the Theory of Quantized Fields* (Interscience Publishers, Inc., New York, 1959), Sec. 16.4.

[12] A. S. Wightman, Nuovo Cimento Suppl. **14**, 81 (1959).

[13] $\overline{V}_m^\pm$ is the closure of $V_m^\pm = \{p|^\pm p_0 > 0,\ p^2 > m^2\}$.

[14] We use the notation

$$(\psi_0, A(x_1)\ldots A(x_n)\psi_0) = \langle x_1, \ldots, x_n\rangle;$$

the letter T stands for the truncated expectation value.

[15] See Ref. 4, Sec. B.

[16] R. Haag, Phys. Rev. **112**, 669 (1958).

However we would like to give a direct proof of the sufficiency condition in terms of the Wightman functions only. It may be useful to have the description of the generalized free field in terms of the corresponding Wightman functions, and hence, to accomplish the analysis of this field as given, in terms of fields operators by Greenberg and others.[5,6]

A. Condition that the Wightman Functions are Tempered Distributions

First of all, we must check that the functions given by (7) are tempered distributions. For $n = 2$ it is clear because $\Delta^-(x - y; u)$ is a tempered distribution,[17] which depends continuously on the parameter u. That the integral over u with the positive measure $\rho(u)$ be a tempered distribution implies that $\rho(u)$ be a function which has at most a polynomial increase.[18]

Our Wightman functions of higher order are expressed as the products of distributions $\langle x, y\rangle$ and, as such, may not, in general, be distributions. However, it was shown by Bogoliubov[19] that it is possible to determine the product of distributions $\Delta^-(x; u)$ as the improper limit of the corresponding product of regularized functions reg $\Delta^-(x; u)$ in the limit of the regularizing masses M_i going to infinity. As the result of such a procedure we obtain the tempered distribution again.

This procedure works because we only multiply distributions which have a given sign of the frequency (in our case only negative frequencies are present). A product of distributions $\Delta^{\pm}(x; u)$, with mixed frequencies, may not be a distribution at all; in this case we used to say that the product contained divergences.[20]

B. Lorentz Invariance Condition

The distribution $\Delta^-(x; u)$ has the elementary property

$$\Delta^-(\Lambda x; u) = \begin{cases} \Delta^-(x; u) & \text{for } \Lambda \in L_+^+(R), \\ -\Delta^{*-}(x; u) & \text{for } \Lambda \in L_+^-(R). \end{cases} \tag{11}$$

From this we get for two-point Wightman function the following property

$$\langle \Lambda x + a, \Lambda y + a\rangle = \begin{cases} \langle x, y\rangle & \text{for } \Lambda \in L_+^+(R), \\ \langle x, y\rangle^* & \text{for } \Lambda \in L_+^-(R). \end{cases} \tag{12}$$

[17] See Ref. 11, Sec. 16.3.
[18] See Ref. 11, Sec. 48.4.
[19] See Ref. 11, Sec. 16.5.
[20] A more elegant way to see it was communicated to me by Wightman, and is based on the observation that the distributions which are Fourier transforms of tempered distributions with support in a convex cone form a ring because they are boundary values of holomorphic functions which form a ring.

Using the definition of the higher-order Wightman functions we obtain

$$\langle \Lambda x_1 + a, \cdots, \Lambda x_n + a\rangle = \begin{cases} \langle x_1, \cdots, x_n\rangle & \text{for } \Lambda \in L_+^+(R), \\ \langle x_1, \cdots, x_n\rangle^* & \text{for } \Lambda \in L_+^-(R). \end{cases} \tag{13}$$

Thus the Lorentz invariance condition is satisfied.

C. Spectral Condition

The Wightman functions $\langle x_1, \cdots, x_n\rangle$ are tempered distributions, and hence, they have Fourier transforms which are also tempered distributions,[21]

$$\begin{aligned} \langle q_1, \cdots, q_{2m}\rangle &= \int \exp\left(i \sum_{s=1}^{2m} q_s x_s\right) \\ &\times \langle x_1, \cdots, x_{2m}\rangle \, d^4x_1 \cdots d^4x_{2m} \\ &= \int \prod_{j=1}^{m} \frac{-i}{(2\pi)^3} \rho(u_j)\theta(P_j^0)\delta(P_j^2 - u_j^2)\, du_j\, d^4P_j \\ &\times \exp\left(i \sum_{s=1}^{2m} q_s x_s\right) \sum_{i<j} \exp\left(-i[P_i(x_1 - x_h) + \cdots\right. \\ &\left. + P(x_i - x_j) + \cdots + P_m(x_k - x_i)]\right) d^4x_1 \cdots d^4x_{2m}. \end{aligned} \tag{14}$$

After integration over x and P we get

$$\begin{aligned} \langle q_1, \cdots, q_{2m}\rangle &= (2\pi)^{8m} \sum_{i_1<j_1\cdots i_m<j_m} \int \prod_{s=1}^{m} \frac{-i}{(2\pi)^3} \\ &\times \rho(u_s)\theta(q_{i_s}^0)\delta(q_{i_s}^2 - u_s^2)\delta(q_{i_s} + q_{j_s})\, du_s. \end{aligned} \tag{15}$$

The summation goes over all permutations

$$\begin{bmatrix} 1, 2, \cdots, 2m-1, 2m \\ i_1, j_1, \cdots, i_m, j_m \end{bmatrix}$$

such that $i_1 < j_1, \cdots, i_m < j_m$.

From the general term of the sum we get the following conditions on q:

$$\begin{aligned} &q_{i_1} \in V_m^+, \quad q_{j_1} = -q_{i_1} \in V_m^-; \quad j_1 > i_1 = 1, \\ &\cdots \\ &q_{i_m} \in V_m^+, \quad q_{j_m} = -q_{i_m} \in V_m^-; \quad j_m > i_m. \end{aligned} \tag{16}$$

By addition of the second column we obtain

$$q_1 + \cdots + q_{2m} = 0 \quad \text{(translation invariance)}. \tag{17}$$

For the partial sums of the vectors q we get

$$\begin{aligned} &q_1 \in V_m^+, \\ &q_1 + \cdots + q_i \in V_m^+ \cup 0, \\ &(i = 2, \cdots, 2m-1). \end{aligned} \tag{18}$$

[21] L. Gårding and J. L. Lions, Nuovo Cimento Suppl. 14, 9 (1959).

This last condition follows because for any vector $q_s \in V_m^-$ occurring in the sum there is a vector $q_r \in V_m^+$ $(r < s)$ which cancels it.

The condition $q_1 + \cdots + q_i = 0$ can not be satisfied for all the terms of the sum, since there always exists a permutation such that this partial sum does not vanish.

Finally we obtain that the support of $\langle q_1, \cdots, q_n \rangle$ is contained in the following set[22]:

$$q_1 + \cdots + q_i \in \overline{V_m^+} \quad (i = 1, \cdots, n-1), \qquad q_1 + \cdots + q_n = 0. \tag{19}$$

From this we can deduce the well-known analytic properties of Wightman functions in coordinate space.[3]

D. Hermiticity Condition

Under the operation of complex conjugation, the functions $\Delta^-(x; u)$ behaves like

$$[\Delta^-(x; u)]^* = -\Delta^-(-x; u), \tag{20}$$

which yields for $\langle x, y \rangle$ the following property:

$$\langle x, y \rangle^* = \langle y, x \rangle. \tag{21}$$

Taking into account the general formula (7) for Wightman functions, we have

$$\langle x_1, \cdots, x_n \rangle^* = \langle x_n, \cdots, x_1 \rangle. \tag{22}$$

Thus the Hermiticity condition is satisfied.

E. Locality Condition

It may be easily checked that our Wightman functions have the remarkable property

$$\langle x_1, \cdots, x_{i-1}, x, y, x_{i+1}, \cdots, x_n \rangle - \langle x_1, \cdots, x_{i-1}, y, x, x_{i+1}, \cdots, x_n \rangle = (\langle x, y \rangle - \langle y, x \rangle)\langle x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n \rangle, \tag{23}$$

where

$$\langle x, y \rangle - \langle y, x \rangle = i \int_0^\infty \rho(u)\, du \Delta(x - y; u). \tag{24}$$

This function vanishes for spacelike separation of the arguments because of the property of the Pauli–Jordan function[23] and hence, the locality condition is satisfied.

F. Cluster Decomposition Property[24]

In order to prove that our Wightman functions have the cluster-decomposition property, it is enough to check this only for the two-point function, since higher truncated Wightman functions vanish.

This can be easily done because the function falls off exponentially when the argument x proceeds to infinity in a spacelike direction.[25] Thus we have

$$\lim_{|a| \to \infty} |a|^N \Delta^-(x - a; u) = 0, \tag{25}$$

for any integer number N and $a = (0, \mathbf{a})$, $u \geq m^2 > 0$.

We obtain the cluster-decomposition property for higher-order Wightman functions immediately as

$$\lim_{a \to \infty} \langle x_1, \cdots, x_{k-1}, x_k + a, \cdots, x_n + a \rangle = \langle x_1, \cdots, x_{k-1} \rangle \langle x_k, \cdots, x_n \rangle, \tag{26}$$

where a goes to infinity in a spacelike direction.

G. Positive-definiteness Condition

Now we are going to prove the positive-definiteness condition which may be written as follows[3]:

$$\sum_{i,j=0}^{N} \alpha_i^* \alpha_j \int f_{i1}^*(x_1) \cdots f_{ii}^*(x_i) \langle x_i, \cdots, x_1, y_1, \cdots, y_j \rangle \times f_{j1}(y_1) \cdots f_{jj}(y_j)\, d^4x_1 \cdots d^4x_i\, d^4y_1 \cdots d^4y_j \geq 0, \tag{27}$$

for arbitrary functions $f_i(x) \in S_4$, an arbitrary set of complex numbers α_i and for each integer number N.

First of all we would like to remark that only the positive-frequency parts of the functions $f(y)$ and the negative-frequency parts of $f^*(x)$ contribute to the integrals. It is a consequence of a general formula for our Wightman functions (7) and the following property of the two-point function:

$$\int \langle z, y \rangle f(y)\, d^4y = \int \langle z, y \rangle f^{(+)}(y)\, d^4y, \qquad \int f^*(x) \langle x, z \rangle\, d^4x = \int f^{*(-)}(x) \langle x, z \rangle\, d^4x, \tag{28}$$

where

$$f^{(\pm)}(x) = \int \exp(-iPx) f(P) \theta(\pm P_0)\, d^4P, \qquad f^{*(\pm)}(x) = \int \exp(-iPx) f^*(P) \theta(\pm P_0)\, d^4P. \tag{29}$$

Without loss of generality we can assume that the functions $f_i(y)$ have positive frequencies only, and hence $f^*(x)$ have only negative ones.

It may be easily verified that the distribution $\langle x, y \rangle$ has zero value on functions $f(x)$, $g(y)$ if they contain the same frequencies,

$$\int d^4x\, d^4y f^{(\pm)}(x) \langle x, y \rangle g^{(\pm)}(y) = 0; \quad f, g \in S_4. \tag{30}$$

[22] See, e.g., H. Araki, Ref. 3, p. 92.
[23] See, e.g., Ref. 11, Eq. (15.13).
[24] D. Ruelle, Helv. Phys. Acta 35, 147 (1962), see also Ref. 16.
[25] See Ref. 11, Sec. 15.2.

Now, in the sum (27), only terms with $i + j$ even contribute since odd-order Wightman functions vanish. From this it follows that the terms with $i \neq j$ contain an even number of "unbalanced" variables x or y which are contracted among themselves [see Eq. (7)], but such contractions vanish in the sense of (30). Therefore only terms with $i = j$ survive and the positive-definiteness condition becomes

$$\sum_{i=0}^{N} |\alpha_i|^2 \int f_{i1}^*(x_1) \cdots f_{ii}^*(x_i)\langle x_i, \cdots, x_1, y_1, \cdots, y_i\rangle \times f_{i1}(y_1) \cdots f_{ii}(y_i)\, d^4x_1 \cdots d^4x_i\, d^4y_1 \cdots d^4y_i \geq 0. \tag{31}$$

Therefore all we have to check is the positiveness of integrals of the form

$$A_n(f_1, \cdots, f_n) = \int f_1^*(x_1) \cdots f_n^*(x_n)\langle x_n, \cdots, x_1, y_1, \cdots, y_n\rangle \times f_1(y_1) \cdots f_n(y_n)\, d^4x_1 \cdots d^4y_n. \tag{32}$$

Using the general expression for Wightman functions (7) and condition (30) we obtain

$$A_n(f_1, \cdots, f_n) = \sum_{(i_1, \cdots, i_n)} \int d^4x f_1^*(x)\langle x, y\rangle f_{i_1}(y)\, d^4y \cdots \times \int d^4x f_n^*(x)\langle x, y\rangle f_{i_n}(y)\, d^4y. \tag{33}$$

Here the summation extends over all permutations $(i_1, \cdots, i_n)$ of numbers $(1, \cdots, n)$.

Now we introduce the quantity

$$(f, g) \equiv \int d^4x f^*(x)\langle x, y\rangle f(y)\, d^4y. \tag{34}$$

It is not difficult to check that (f, g) satisfies all axioms of the scalar product namely

(i) $(f, g)^* = (g, f)$, it follows from (21)

(ii) $(f, \alpha_1 g_1 + \alpha_2 g_2) = \alpha_1(f, g_1) + \alpha_2(f, g_2)$, (35)

(iii) $(f, f) = (2\pi)^5 \int |f(P)|^2\, \theta(P_0) \times \delta(P^2 - u^2)\rho(u)\, du\, d^4P \geq 0.$

The last inequality follows from the assumption of the positiveness of the Källen–Lehmann measure $\rho(u)$. The equality $(f, f) = 0$ leads us to $f(P)\,|_{P \in V_m^+} \equiv 0$. Thus the positiveness of $\rho(u)$ is sufficient to make the quantity (f, g) a scalar product, and as we will show below this is sufficient for satisfying the positive definiteness condition of Wightman functions in general.[26]

Using notation (34) we may write down the integral A_n in the following form:

$$A_n(f_1, \cdots, f_n) = \sum_{(i_1, \cdots, i_n)} (f_1, f_{i_1}) \cdots (f_n, f_{i_n}). \tag{36}$$

On the right-hand side we have so called permanent[27] of the matrix $||(f_i, f_j)||_1^n$.

It has been proved[28] that for a positive semi-definite Hermitian matrix the following inequality holds:

$$\text{per } A \geq \det A. \tag{37}$$

Equality is reached if and only if A has a zero row or if A is diagonal.

Our matrix $||(f_i, f_j)||_1^n$ obviously satisfies all these conditions since the Gramm determinants of the vectors $f_1, \cdots, f_k$; $(k = 1, \cdots, n)$ are nonnegative.[29]

Thus finally, we obtain the inequality

$$A_n(f_i, \cdots, f_n) = \text{per } ||(f_i, f_j)||_1^n \geq \det ||(f_i, f_j)||_1^n \geq 0. \tag{38}$$

It is well-known that the Gramm determinant $\Gamma(f_1, \cdots, f_n)$ vanishes if and only if the vectors $f_1, \cdots, f_n$ are linearly dependent. From this and the previous conditions on equality in (37), as well as from the positive-definiteness of the scalar product, we conclude that the equality to zero for the integral $A_n(f_1, \cdots, f_n)$ can be reached if and only if some function

$$f_i(P)|_{P \in V_m^+}; \qquad (i = 1, \cdots, n)$$

vanishes identically.

The condition of positive-definiteness is then satisfied by our Wightman functions.

Now we want to prove that our Wightman functions describe the theory which is equivalent to the theory of generalized free field.

---

[26] Another proof of positiveness of $\rho(u)$ being sufficient to make a metric in Hilbert space positive, was given by Greenberg; see Ref. 4, Sec. E.

[27] The general definition of the permanent of n-square matrix $A = ||a_{ij}||_1^n$ is

$$\text{per } A = \sum_{\sigma \in S_n} \prod_{i=1}^{n} a_{i\sigma(i)},$$

where S_n is the full symmetric group. For details see Ref. 27.

[28] M. Marcus and M. Newman, Ann. Math. 75, 47 (1962), Theorem 8.

[29] See, e.g., N. I. Akhiezer and I. M. Glazman, *The Theory of Linear Operators in Hilbert Space* (Frederick Ungar Publishing Company, New York, 1961), Sec. 8.

4. RECONSTRUCTION OF THE THEORY

In the usual manner[3] we construct the Hilbert space H of the states. It was shown[30] that this space is separable, i.e., there exists in H a linear, countable and dense set D.[31]

The metric in D is positive definite as it follows from G.

In D there exists a continuous unitary representation of the inhomogeneous proper Lorentz group (see condition B). The spectrum of the generators of the translation group is contained in V_m^+ apart from the zero point which corresponds to the vacuum, (follows from condition C).

The field operator $A(f)$, which is a tempered operator-valued distribution, is given in condition D (it is a consequence of A).[32]

This operator is Hermitian [in the sense $A(f)^+ = A(f^*)$], local, and possesses the demanded transformation properties under $U(a, \Lambda)$, (see conditions D and E]. Moreover, it has been proved[33] that $A(f)$ is essentially self-adjoint, i.e., the closure $\overline{(Af)}$ is self adjoint on D if f is an arbitrary real test function with compact support. It follows immediately from the vanishing of all higher truncated Wightman functions for this field.

Due to the cluster decomposition property (see condition F), the field $A(f)$ is irreducible[34] and there exists only one vector in D, apart from a trivial multiplication factor, which is invariant under this is the vacuum state.[35]

That the field $A(f)$ is then the generalized free field is a consequence of the identity [see (23)].

$$\langle \cdots, x, y, z, \cdots \rangle - \langle \cdots, y, x, z, \cdots \rangle - \langle \cdots, z, x, y, \cdots \rangle + \langle \cdots, z, y, x, \cdots \rangle \equiv 0, \tag{39}$$

which holds for all our Wightman functions.

As a result of this equality we get that $\{[A(f), A(g)], A(h)\}\psi$ vanishes for all $f, g, h \in S_4$ and for all states ψ belonging to D. Thus from the density of D and the irreducibility of the field we conclude that $[A(f), A(g)] = C(f, g)\cdot 1$ is a multiple of unit operator. The value of this number is

$$C(f, g) = (\psi_0, [A(f), A(g)]\psi_0) = \int d^4x f(x)(\langle x, y\rangle - \langle y, x\rangle)g(y)\, d^4y, \tag{40}$$

and hence $A(f)$ is a generalized free field.

ACKNOWLEDGMENTS

The author would like to thank Professor P. R. Wallace and other members of the Theoretical Physics Group of the Department of Physics of the McGill University for their generous hospitality. The author is pleased to express his indebtedness to Dr. J. Lambek and Dr. N. Pullman for stimulating discussions, and to Professor A. S. Wightman for helpful remarks.

[30] H. J. Borchers, Nuovo Cimento **24**, 214 (1962). See also Ref. 11, Appendix.

[31] The set D may be taken as the linear manifold of the elements obtained by applying any polynomial in $A(f)$ to the vacuum state.

[32] See Ref. 11, Appendix.

[33] R. Jost, Lecture notes, Princeton University (1963), and also H. J. Borchers and W. Zimmermann, "On the Self-Adjointness of the Field Operators," (preprint).

[34] R. Haag and B. Schroer (see Ref. 6, Sec. IV), gives another proof of the irreducibility using some results from the spectral theory of operators. Our proof is more straightforward. Note: The irreducibility of $A(x)$ is understood in the sense of x varying over all Minkowski space; it is reducible, provided $\sigma \not\equiv 0$, when x is restricted to a timeslice.

[35] See Ref. 3, Theorems 7 and 8, and also Ref. 34.

On Dirac's Wave Equation in a Gravitational Field*

ASHER PERES
Department of Physics, Israel Institute of Technology, Haifa, Israel
(Received 3 December 1963)

The Dirac equation is investigated in the combined electric and gravitational field of a point charge in General Relativity. The wavefunctions are weakly singular at the origin, but still normalizable for a *continuous* range of the energy. The Hamiltonian, however, is not self-adjoint over the manifold of its own "eigenstates," though it can be made self-adjoint by a suitable choice of its domain of definition. The theory, however, is unable to decide how the Hamiltonian should be defined, and what are the bound states.

IT was shown some time ago by Callaway[1] that when account is taken of the gravitational field of a point charge, the Dirac equation for a charged particle in a Coulomb field does not possess solutions which can be expanded as a series of terms proportional to positive integral powers of the gravitational constant. In other words, the gravitational field, however weak, cannot be treated by standard perturbation methods. The purpose of this note is to show that actually in this case, the Hamiltonian is not self-adjoint over the manifold of its "eigenstates," though it can be made self-adjoint by a suitable choice of its domain of definition. The theory, however, is unable to decide how the Hamiltonian should be defined, and what are the bound states.

This problem may seem purely academic, as there are no point charges in nature, but only extended charges whose size is much larger than the charge itself.[2] Therefore the peculiar geometric properties of the Nordström metric[3] (which are here the cause of the trouble) do not really preclude the existence of hydrogen atoms. It is nevertheless worthwhile to discuss this problem, in view of the recent revival of interest about singular potentials in quantum mechanics.[4,5]

The Nordström metric corresponding to an electrostatic potential e/r due to a point source of charge e and mass M is[3]

$$ds^2 = C^2\,dt^2 - C^{-2}\,dr^2 - r^2(d\theta^2 + \sin^2\theta\,d\phi^2), \quad (1)$$

where[2]

$$C = [1 - (2M/r) + (e^2/r^2)]^{\frac{1}{2}} \quad (2)$$

is the local velocity of light. We shall consider here only the case $e > M$, so that the hypersurfaces $t = \text{const}$ are spacelike.[6]

The form of the Dirac equation for the metric (1) has been given by Brill and Wheeler[7] and can be written, after separation of the time and angular variables, as

$$Hu = Eu, \quad (3)$$

where H is the "radial Hamiltonian"

$$H = iC^2\sigma_y(\partial/\partial r) + C\sigma_x(k/r) - Cm\sigma_z - (e^2/r), \quad (4)$$

and

$$u \equiv \begin{pmatrix} f \\ g \end{pmatrix} = AC^{\frac{1}{2}}r\Psi. \quad (5)$$

Here, σ_x, σ_y, and σ_z are the familiar Pauli matrices, m is the electron rest mass, k is a positive or negative integer related to the angular momentum, and A is a 4×2 matrix depending only on the angular variables and relating u to Dirac's original Ψ.

As the scalar product of two Dirac states is[8]

$$(\Psi_1, \Psi_2) = \int \Psi_1^\dagger \beta\gamma^0 \Psi_2 r^2 \sin\theta\,dr\,d\theta\,d\phi \quad (6)$$

$$= \int \Psi_1^\dagger \Psi_2 C^{-1} r^2 \sin\theta\,dr\,d\theta\,d\phi, \quad (7)$$

* This research was supported in part by the United States Air Force under Grant No. AF-EOAR-63-107 and monitored by the European Office, Office of Aerospace Research.

[1] J. Callaway, Phys. Rev. **112**, 290 (1958).

[2] Natural units are used: $c = G = \hbar = 1$.

[3] G. Nordström, Verslag Gewone Vergader. Afdel. Natuurk. Koninkl. Ned. Akad. Wetenschap. **20**, 10 (1918).

[4] B. W. Roos and W. C. Sangren, J. Math. Phys. **3**, 882 (1962), **4**, 999 (1963). B. H. Armstrong and E. A. Power, Am. J. Phys. **31**, 262 (1963). B. H. Armstrong, Phys. Rev. **130**, 2506 (1963), where further references are given.

[5] P. Teitelbaum (preprint, 1963) has suggested to cut off the potential at some small radius, and to investigate how the wavefunction behaves when the cutoff radius tends to zero. This method is not applicable here, because it would violate the Einstein gravitational equations.

[6] The case $e < M$ was discussed by the author in Phys. Rev. **120**, 1044 (1960), where it is shown that the Schwarzschild singularity prevents the normalization of any bound-state solution. Roughly speaking, the particle has a tendency to escape through the Einstein–Rosen "bridge" [Phys. Rev. **48**, 73 (1935)].

[7] D. R. Brill and J. A. Wheeler, Rev. Mod. Phys. **29**, 265 (1957). For a shorter derivation of the Dirac equation in General Relativity, see A. Peres, Nuovo Cimento **28**, 865 (1963), Eq. (18).

[8] A. Peres, Nuovo Cimento Suppl. **24**, 389 (1962).

that of the radial wavefunctions is (within a constant factor which can be absorbed in A)

$$(u_1, u_2) = \int_0^\infty u_1^\dagger u_2 C^{-2}\, dr, \tag{8}$$

in virtue of (5).

In the vicinity of the origin, $C \to e/r$, and the leading terms in (3) are

$$iC^2\sigma_y(\partial u/\partial r) + (C\sigma_x k/r)u = O(r^{-1})u, \tag{9}$$

whence it follows that $f \to f(0)\exp(kr/e)$ and $g \to g(0)\exp(-kr/e)$ are both finite at the origin, and *a fortiori* normalizable there. [Special cases occur when either $f(0)$ or $g(0)$ is zero. It is easily shown that if $f(0) = 0$, then $f \to -\frac{1}{2}g(0)(1 - m/e)r^2$, while if $g(0) = 0$, then $g \to \frac{1}{2}f(0)(1 + m/e)r^2$.]

On the other hand, for $r \to \infty$, the leading terms in (3) are

$$i\sigma_y(\partial u/\partial r) - m\sigma_z u + O(r^{-1})u = Eu, \tag{10}$$

so that, asymptotically,

$$u \to \begin{bmatrix} -(m - E)^{\frac{1}{2}} \\ (m + E)^{\frac{1}{2}} \end{bmatrix} \exp\left[-(m^2 - E^2)^{\frac{1}{2}}r\right]. \tag{11}$$

Thus, if we start from a solution like (11), with $E < m$, and progress towards $r = 0$ by means of the first-order differential equations (3) and (4), then we always obtain normalizable eigenfunctions of H.[9] This is clearly incompatible with the basic principles of quantum mechanics, according to which normalizable eigenfunctions must always correspond to a discrete spectrum.

It is not difficult to find the origin of this inconsistency. The proof of the orthogonality of the Hamiltonian eigenfunctions requires the Hamiltonian to be self-adjoint, i.e.,

$$(u_1, Hu_2) = (Hu_1, u_2), \tag{12}$$

where the scalar product is given by (8). An elementary calculation then shows that (12) will be satisfied by the Hamiltonian (4) if, and only if

$$f_1^* g_2 - f_2 g_1^* = 0 \quad \text{for} \quad r = 0. \tag{13}$$

This may be satisfied, e.g., by defining H only over wavefunctions such that $f(0) = g(0) = 0$. This definition, however, is too restrictive, because such a H is not maximal[10] and actually has an infinity of distinct maximal extensions. Indeed, (13) still holds if the domain of H is extended to wavefunctions such that

$$f(0)/g(0) = R, \tag{14}$$

where R is any real number (possibly zero or infinity) the choice of which is arbitrary, but has to be made once for all.

Thus, once R has been chosen, the eigenvalues of H (if any) are determined. Unfortunately, there is nothing in this theory which can make some value of R better than others, so that the theory is conspicuously incomplete.[11]

Note that the above result is independent of the choice of both the coordinate system and the spin frame, because the basic equations of the theory are generally covariant.[8]

The reader himself should decide whether to consider this result as an amusing curiosity, or as an indication of the inadequacy of our present physical theories.[12]

[9] With the ordinary (special relativistic) Dirac equation, the wavefunctions obtained in this way diverge at the origin and are not normalizable, unless E belongs to the spectrum. See B. H. Armstrong, Ref. 4.

[10] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, New Jersey, 1955), p. 153. The example considered by von Neumann is that of the operator $i\partial/\partial q$, with $0 \leq q \leq 2\pi$, originally defined over $f(q)$ such that $f(0) = f(2\pi) = 0$. This operator is not maximal, but has an infinity of distinct maximal extensions, namely for $f(q)$ such that $f(0)/f(2\pi) = \exp(i\alpha)$, with α real. Here, α plays the same role as R in the present paper.

[11] The same also holds for the ordinary Dirac equation, but does not cause any serious trouble, because the Hamiltonian eigenfunctions satisfy $f(0) = g(0) = 0$.

[12] Similar results also hold for the Klein–Gordon equation, but will not be discussed in detail, since the Klein–Gordon equation is not a well-behaved quantum mechanical equation anyhow. See H. Feshbach and F. Villars, Rev. Mod. Phys. **30**, 24 (1958), especially p. 29.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 5, NUMBER 6 JUNE 1964

Properties of a Class of Nonlocal Solvable Interactions

G. C. GHIRARDI AND A. RIMINI
Istituto di Fisica dell'Università, Trieste, Italy
(Received 10 December 1963)

A generalization of factorable interactions is taken into account. The direct and inverse problems for the relative Schrödinger equation are investigated and they turn out to be workable. It is shown that such interactions can produce several bound states. The solution of the direct problem is given. A class of interactions of the considered type is constructed which produces an arbitrarily assigned finite set of bound states.

1. INTRODUCTION

VARIOUS authors[1,2] have recently considered, both to discuss quantitatively concrete problems and to investigate the analytic behavior of the scattering amplitudes, an interesting class of nonlocal interactions: the so-called factorable or separable interactions. The term factorable indicates the fact that the interaction, which is assumed to be Hermitian, has a representative of the form

$$V(\omega, \omega') = cv(\omega)v^*(\omega'). \tag{1.1}$$

It has been also considered a larger class of solvable interactions, often referred as "completely separable".[1] More precisely, the class of the nonlocal, rotationally invariant interactions, for which the equations arising from the Schrödinger equation when the separation in the various waves is accomplished, are eigenvalue equations with an interaction of type (1.1).

We shall use in this paper the term factorable to indicate the property expressed by (1.1). On the other hand, we shall speak of reducible interactions to indicate the property of every interaction H_I which commutes with a commuting set of observables (which we shall indicate with M) commuting also with the free Hamiltonian H_0 of the system. For such an H_I the Schrödinger equation can be separated in various equations in the eigenmanifolds of the considered set M.

The interesting feature of factorable interactions is that the corresponding Schrödinger equation is easily solvable for both the discrete and continuous spectra. The properties of factorable interactions are well known; one of the most characteristic ones being the fact that they can give at most only one bound state. This fact may seriously limit their applications to physical problems.

In this paper, starting from the observation that a factorable interaction is a constant times a projection operator of a one-dimensional linear manifold of the Hilbert space, we deduce all the known properties of such interactions in a simple way (Sec. 2). Then we take into consideration a class of interactions which is a natural generalization of factorable ones, i.e., the class of their finite linear combinations. The Schrödinger equation of such a class of interactions is still solvable in the sense that the determination of eigenvalues, eigenfunctions, and all scattering properties, reduces to the resolution of systems of linear algebraic equations. This fact is almost obvious as the considered interaction is an integral operator whose kernel belongs to the class of the degenerate (or Pincherle Goursat) kernels of the theory of integral equations. We shall show that the properties of the determinants of these systems, which in general are not workable, can be, in our case, exhaustively studied owing to the positive-definiteness of H_0. It results that the Schrödinger equation can have several bound states. The dependence of the number of these bound states on the interaction is investigated (Sec. 3). In Sec. 4 we briefly sketch the well known formalism of a reducible problem, assuming that the reduced problems to which it gives rise are of the class introduced in Sec. 3. We then impose to this class the requirements of rotational and time-reversal invariance and we obtain a class of solvable interactions which can have several bound states in each wave. In the concluding Section (Sec. 5), the inverse problem is investigated, i.e., the problem of constructing an interaction of the considered class which gives an assigned number of bound states with assigned eigenvalues and eigenfunctions.

[1] Y. Yamaguchi, Phys. Rev. **95**, 1628 (1954).
[2] A. N. Mitra and V. L. Narasimham, Nucl. Phys. **14**, 407 (1959); J. T. Cushing, Nuovo Cimento **28**, 819, (1963); A. N. Mitra and J. D. Anand, Phys., Rev. **130**, 2117 (1963). R. Alzetta, G. C. Ghirardi, and A. Rimini, Phys. Rev. **131**, 1740 (1963). Further references can be found in V. L. Narasimham, S. K. Shah, and S. P. Pandya, Nucl. Phys. **33**, 529 (1962).

It is found that there exists a unique interaction which gives n assigned bound states in the class of the linear combinations of n terms like (1.1). If we do not restrict ourselves to the combinations of n terms, there is a large arbitrariness which allows us to modify the scattering properties to a great extent.

We have confined in an Appendix the proof of the completeness of the eigenstates of the Hamilton operator considered in Sec. 2.

2. FACTORABLE INTERACTIONS

An Hermitian operator V, which has the ω representative

$$V(\omega, \omega') = cv(\omega)v^*(\omega'),$$

where c is real and $v(\omega)$ is such that $\int d\omega\ |v(\omega)|^2$ exists, will be called factorable in the ω representation. It is easily seen that if an operator is factorable in one representation it is factorable in every representation. If we call $|V\rangle$ the ket whose ω representative is $v(\omega)$, we have

$$V = c\ |V\rangle\langle V|,$$

and this shows that a factorable operator is a constant times a projection operator belonging to a one-dimensional linear manifold of the Hilbert space. Evidently it is also true the converse: any operator which is a constant times a projection operator belonging to a one-dimensional linear manifold is factorable.

If we assume V as the interaction operator of a quantum mechanical system, the respective Schrödinger equation reads

$$(E - H_0)\ |\psi_E\rangle = c\ |V\rangle\langle V\ |\ \psi_E\rangle. \tag{2.1}$$

If $E < 0$, $(E - H_0)$ is a definite negative operator[3] and therefore it has an inverse $(E - H_0)^{-1} = G^{(0)}(E)$. Thus Eq. (2.1) admits (apart from an arbitrary multiplicative factor) the unique solution

$$|\psi_E\rangle = G^{(0)}(E)\ |V\rangle, \tag{2.2}$$

if and only if E satisfies

$$c\langle V|\ G^{(0)}(E)\ |V\rangle = 1. \tag{2.3}$$

Equation (2.1), and therefore (2.3), can have at most one solution. In fact, let us assume that there are two orthogonal eigenfunctions $|\psi_1\rangle$, $|\psi_2\rangle$ belonging, respectively, to the eigenvalues E_1 and E_2. They are of the form

$$|\psi_1\rangle = G^{(0)}(E_1)\ |V\rangle; \qquad |\psi_2\rangle = G^{(0)}(E_2)\ |V\rangle.$$

Orthogonality implies

$$\langle V|\ G^{(0)}(E_1)G^{(0)}(E_2)\ |V\rangle = 0.$$

Since $G^{(0)}(E_1)G^{(0)}(E_2)$ is a positive-definite operator, this would imply $|V\rangle = 0$. Owing to the definite negativeness of $G^{(0)}(E)$, Eq. (2.3) certainly admits no solution if $c > 0$. Moreover $\langle V|\ G^{(0)}(E)\ |V\rangle$ is a decreasing function of E for $E < 0$. Concluding, (2.3) has one solution, if and only if

$$c < 0; \qquad \lim_{E\to 0^-} \langle V|\ G^{(0)}(E)\ |V\rangle < \frac{1}{c}.$$

If $E \geq 0$ we introduce the operators

$$G^{(0)+}(E) = \lim_{\epsilon\to 0^+} (E + i\epsilon - H_0)^{-1}$$

and

$$G^{+}(E) = \lim_{\epsilon\to 0^+} (E + i\epsilon - H)^{-1}$$

which satisfy

$$G^{+}(E) = G^{(0)+}(E) + G^{+}(E)VG^{(0)+}(E). \tag{2.4}$$

The T operator is expressed by:

$$T(E) = V + VG^{+}(E)V$$
$$= c\ |V\rangle\langle V|\ (1 + c\langle V|\ G^{+}(E)\ |V\rangle).$$

Owing to (2.4) we have

$$\langle V|\ G^{+}(E)\ |V\rangle$$
$$= \langle V|\ G^{(0)+}(E)\ |V\rangle(1 - c\langle V|\ G^{(0)+}(E)\ |V\rangle)^{-1},$$

and then

$$T(E) = \frac{c}{1 - c\langle V|\ G^{(0)+}(E)\ |V\rangle}\ |V\rangle\langle V|.$$

In the Appendix we have taken into account the scattering states and we have shown the completeness of the system of the eigenstates of H for our problem.

3. A GENERALIZATION OF FACTORABLE INTERACTIONS

Now we generalize the class of factorable interactions with two aims in mind. First of all we want that the Schrödinger equation still hold solvable, and secondly to have the possibility of producing several bound states. A natural generalization is to assume that

$$V = \sum_{i=1}^{n} c_i\ |V_i\rangle\langle V_i|, \tag{3.1}$$

[3] We shall call a positive- (negative-) definite operator any Hermitian operator D for which $\langle f|D|f\rangle \geq 0$ (≤ 0) for any $|f\rangle$, and $= 0$ if and only if $|f\rangle = 0$.

where the $|V_i\rangle$ are n linearly independent vectors, and the c_i's are real constants. The Schrödinger equation is solvable as may easily be shown. It runs as follows:

$$(E - H_0)\,|\psi_E\rangle = \sum_{i=1}^{n} c_i\,|V_i\rangle\langle V_i\,|\,\psi_E\rangle. \tag{3.2}$$

For $E < 0$, Eq. (3.2) admits only solutions of the form

$$|\psi_E\rangle = \sum_{i=1}^{n} k_i G^{(0)}(E)\,|V_i\rangle. \tag{3.3}$$

Equation (3.3) implies that the eigenvalues may have at most a degeneracy of order n. $|\psi_E\rangle$ is actually a solution if and only if the k_i's satisfy the system

$$\sum_{j=1}^{n} (\delta_{ij} - c_i\langle V_i|\,G^{(0)}(E)\,|V_j\rangle)k_j = 0. \tag{3.4}$$

System (3.4) has nontrivial solutions if and only if

$$\det|\delta_{ij} - c_i\langle V_i|\,G^{(0)}(E)\,|V_j\rangle| = 0. \tag{3.5}$$

The just-accomplished reduction of the problem to a system of linear algebraic equations is almost trivial, and is quite analogous to the standard method for the solution of integral equations with degenerate kernel.[4] But the general theory does not give any suggestion for the discussion of Eq. (3.5), the transcendental equation which determines the eigenvalues. We shall show that in our case, however, owing to the positive-definiteness of H_0, the eigenvalue problem can be studied. First of all, let us show that there are at most n distinct eigenvectors of H.

We consider m orthogonal vectors $|\psi_k\rangle$ such that

$$\langle\psi_i\,|\,\psi_k\rangle = -(1/E_k)\delta_{ik}, \tag{3.6}$$

where the E_k's are m negative real numbers. Putting

$$|W_k\rangle = (E_k - H_0)\,|\psi_k\rangle, \tag{3.7}$$

it is easily shown that from (3.6) it follows that the $|W_k\rangle$ are m linearly independent vectors. In fact from

$$\sum_{k=1}^{m} \alpha_k\,|W_k\rangle = 0,$$

i.e.,

$$\sum_{k=1}^{m} \alpha_k(E_k - H_0)\,|\psi_k\rangle = 0,$$

it follows that

$$\sum_{k=1}^{m} (\delta_{ik} + \langle\psi_i|\,H_0\,|\psi_k\rangle)\alpha_k = 0. \tag{3.8}$$

Writing

$$N_{ik} = \delta_{ik} + \langle\psi_i|\,H_0\,|\psi_k\rangle,$$

the $m \times m$ matrix N_{ik} is Hermitian and can be diagonalized. Let U_{ik} be the unitary matrix such that

$$\sum_{i,k=1}^{m} U_{ii}N_{ik}U_{lk}^{*} = \delta_{il}N_i.$$

Then

$$N_i = 1 + \Big(\sum_{j=1}^{m} U_{ij}\,\langle\psi_j|\Big)H_0\Big(\sum_{k=1}^{m} U_{ik}^{*}\,|\psi_k\rangle\Big).$$

This is positive owing to the positive-definiteness of H_0. In particular, as $\det|N_{ik}| = \prod_{i=1}^{m} N_i$, the determinant of N_{ik} is positive and therefore (3.8) admits only the zero solution.[5]

Recalling that the eventual solutions of Eq. (3.2) are of the form (3.3) the result of the preceding paragraph proofs that there are at most n eigenvectors of H, and therefore that Eq. (3.5) has at most n solutions.

Moreover it can be seen that the number of the possible bound states is governed by the number of negative coefficients in the interaction (3.1). In fact let us write V as

$$V = \sum_{i=1}^{l} c_i\,|V_i\rangle\langle V_i| + \sum_{i=l+1}^{n} c_i\,|V_i\rangle\langle V_i|,$$

[4] In fact, if we consider, for example, the case of a spinless particle with rotationally invariant interaction of our kind, the Schrödinger equation for the l wave is, in q representation (see Sec. 4),

$$K(q)\psi_l(q) + \sum_{k=1}^{m_l} c_k \int_0^{\infty} q'^2\,dq' v_{lk}(q) \times v_{lk}(q')\psi_l(q') = E\psi_l(q),$$

and can be written

$$\psi_l(q) = \lambda_l \int_0^{\infty} R(q, q'; \lambda_l)\psi_l(q')\,dq',$$

where

$$\lambda_l = \frac{1}{E_l}; \qquad R(q, q'; \lambda_l) = \frac{\sum_{k=1}^{m_l} c_k v_{lk}(q)v_{lk}(q')q'^2}{1 - \lambda_l K(q)}.$$

This is an integral equation with degenerate kernel depending on the eigenvalue λ_l. For such an equation the standard method of solution reduces it to an algebraic system of linear equations. In general however owing to the dependence of the kernel from the eigenvalue, nothing can be said on the eigenvalues of such an equation.

[5] The theorem now proved asserts that among the totality of vectors of the kind

$$|f(E, |W\rangle)\rangle = G^{(0)}(E)\,|W\rangle,$$

where E varies from $-\infty$ to 0 and $|W\rangle$ runs over an m-dimensional linear manifold, there cannot be found $m + 1$ orthogonal vectors. This result is not trivial if we observe that among these vectors there are infinite linearly independent ones.

where $c_i < 0$ for $i \leq l$ and $c_i > 0$ for $l < i \leq n$. If one supposes that there are m eigenvectors $|\psi_1\rangle$, $|\psi_2\rangle$, $\cdots$, $|\psi_m\rangle$ pertaining to the eigenvalues E_1, E_2, $\cdots$, E_m, respectively, there follows the validity of the equations

$$(E_k - H_0)\,|\psi_k\rangle = \sum_{i=1}^{l} c_i\,|V_i\rangle\langle V_i \mid \psi_k\rangle + \sum_{i=l+1}^{n} c_i\,|V_i\rangle\langle V_i \mid \psi_k\rangle. \quad (k = 1, 2, \cdots, m).$$

Multiplying from the left these equations by $\langle\psi_i|$ we get

$$-\sum_{i=1}^{l} c_i\langle\psi_j \mid V_i\rangle\langle V_i \mid \psi_k\rangle = N_{jk} + \sum_{i=l+1}^{n} c_i\langle\psi_j \mid V_i\rangle\langle V_i \mid \psi_k\rangle \quad (j, k = 1, 2, \cdots, m)$$

if Eq. (3.6) holds.

The determinant of the matrix at the right hand side is easily seen to be positive, while the determinant of the matrix at the left is zero if $l < m$. Indeed, let l be less than m and A_{ji} and B_{ik} be two rectangular $m \times l$ and $l \times m$ matrices, respectively. Putting $\gamma_i = 1$ for $i \leq l$, $\gamma_i = 0$ for $l < i \leq m$ we have

$$\sum_{i=1}^{l} A_{ji}B_{ik} = \sum_{i=1}^{m}\sum_{r=1}^{m} A_{ji}\gamma_i\delta_{ir}B_{rk},$$

anyhow we choose A_{ji} and B_{ik} for $l < i \leq m$. Since $\gamma_i\delta_{ir}$ has the determinant equal to zero the matrix at the left hand side too has this property. Thus it is proved that the number of bound states is at most equal to the number of negative coefficients in the interaction (3.1).

When $E \geq 0$ we consider the operators $G^{(0)+}(E)$, $G^{+}(E)$. Owing to the fact that they satisfy (2.4) we have

$$\sum_{l=1}^{n} (\delta_{il} - c_l\langle V_l|\, G^{(0)+}(E)\,|V_i\rangle)\langle V_i|\, G^{+}(E)\,|V_l\rangle = \langle V_i|\, G^{(0)+}(E)\,|V_i\rangle. \quad (3.9)$$

Eqs. (3.9) are n systems of order n. They allow us to have the numbers $\langle V_i|\, G^{+}(E)\,|V_l\rangle$ expressed in terms of the quantities $\langle V_i|\, G^{(0)+}(E)\,|V_l\rangle$. The T operator then results,

$$T(E) = V + VG^{+}(E)V = \sum_{i,l=1}^{n} c_i(\delta_{il} + c_l\langle V_i|\, G^{+}(E)\,|V_l\rangle)\,|V_i\rangle\langle V_l|.$$

4. FURTHER GENERALIZATION

The above obtained class of solvable interactions can be easily generalized, recalling that if we have a quantum mechanical system that has some invariance property, and consequently some conservation law, the eigenvalue problem of such a system can be separated in various simpler problems in the eigenmanifolds of the operators which commute with H as requested by the existence of the conservation law.

Let us now assume that $H = H_0 + H_I$ be the Hamilton operator for our system and that there exists an observable M with pure discrete spectrum which commutes with both terms in H.

We introduce the projection operators P_{M_i} belonging to the eigenmanifolds M_i of M, which satisfy

$$[H_0, P_{M_i}] = [H_I, P_{M_i}] = 0, \quad (4.1a)$$

$$P_{M_i}P_{M_j} = \delta_{ij}P_{M_i}, \quad (4.1b)$$

$$\sum_i P_{M_i} = 1. \quad (4.1c)$$

Owing to (4.1c), the eigenvalue equation

$$(H_0 + H_I)\,|\psi_E\rangle = E\,|\psi_E\rangle \quad (4.2)$$

may be written

$$H_0\sum_i P_{M_i}\,|\psi_E\rangle + H_I\sum_i P_{M_i}\,|\psi_E\rangle = E\sum_i P_{M_i}\,|\psi_E\rangle.$$

Multiplying this equation from the left by P_{M_j} and using (4.1a), and (4.1b), we get the equations

$$H_0P_{M_j}\,|\psi_E\rangle + H_IP_{M_j}\,|\psi_E\rangle = EP_{M_j}\,|\psi_E\rangle, \quad (4.3)$$

which, in turn, owing to (4.1c), imply the eigenvalue Eq. (4.2). With the aid of (4.1b), Eqs. (4.3) can be written:

$$(P_{M_j}H_0 + P_{M_j}H_I)P_{M_j}\,|\psi_E\rangle = EP_{M_j}\,|\psi_E\rangle. \quad (4.4)$$

The problem has so been reduced to various problems in the eigenmanifolds M_j.

If we assume that the operators $P_{M_j}H_I$ belong to the class of Sec. 3, the eigenvalue problem of H has been reduced to the solutions of systems of a finite number of algebraic linear equations.

The most general interaction H_I, which commutes with an observable M and is solvable along the lines sketched in this paper, is therefore of the type

$$H_I = \sum_i \sum_{k=1}^{m_i} c_{ik}Q_{ik}, \quad (4.5)$$

where the Q_{ik}'s, $(k = 1, 2, \cdots, m_i)$ are projection

operators belonging to finite-dimensional manifolds (which we can always assume to be orthogonal without loss of generality) which are contained in the manifold M_i. Formally,

$$\begin{aligned} Q_{ik}Q_{il} &= Q_{ik}\delta_{il}, \\ P_{M_i}Q_{ik} &= \delta_{ij}Q_{ik}. \end{aligned} \tag{4.6}$$

Let us now take into account a spinless particle of mass μ; we shall call $\mathbf{q}$ its momentum. In momentum representation, H_0 is the kinetic energy operator $K(q)$. Rotational invariance allows the simultaneous diagonalization of H_I, $\mathbf{L}^2$, L_z,

$$(H_I)_{lm|q|,l'm'|q'|} = h_{lm}(q, q')\delta_{ll'}\delta_{mm'},$$

and implies and is implied by the fact that in the representation in which H_0, $\mathbf{L}^2$ and L_z are diagonal, $h_{lm}(q, q') = h_l(q, q')$ does not depend on m.

In the momentum representation we have

$$(H_I)_{\mathbf{q},\mathbf{q}'} = \sum_{l,m} h_l(q, q')Y_l^m(\hat{\mathbf{q}})Y_l^{m*}(\hat{\mathbf{q}}') = \sum_l \frac{2l+1}{4\pi} h_l(q, q')P_l(\cos \widehat{\mathbf{q}\mathbf{q}}').$$

Time-reversal invariance implies that $h_l(q, q')$ is a real function of q and q'. Let us now identify the above mentioned operator M with the commuting set $\mathbf{L}^2$, L_z. Then

$$(P_{lm}H_I)_{\mathbf{q},\mathbf{q}'} = h_l(q, q')Y_l^m(\hat{\mathbf{q}})Y_l^{m*}(\hat{\mathbf{q}}').$$

In this case the P_{lm} are projection operators belonging to closed linear manifolds of infinite dimensionality.

The equations of the separate problems still are integral equations; for example, in q representation,

$$K(q)\psi_l(q) + \int q'^2\, dq' h_l(q, q')\psi_l(q') = E_l\psi_l(q),$$

we can now assume that

$$h_l(q, q') = \sum_{i=1}^{m_l} c_i v_{li}(q)v_{li}(q')$$

and we can solve the problem completely.

The most general reducible, factorable, rotationally and time-invariant interaction is therefore, in $\mathbf{q}$ representation,

$$(H_I)_{\mathbf{q},\mathbf{q}'} = \sum_l \sum_{i=1}^{m_l} c_i v_{li}(q)v_{li}(q')P_l(\cos \widehat{\mathbf{q}\mathbf{q}}').$$

This can have at most m_l different eigenvalues in the l wave. Note however that each eigenvalue is $(2l + 1)$ times degenerate owing to rotational invariance.

5. THE INVERSE PROBLEM

In this section we wish to take into account the problem of constructing the interactions of the previously considered kinds which make the Schrödinger equation to have an assigned set of eigenvectors belonging to assigned eigenvalues.[6]

Let us start with the case of only one eigenvector. We look for a pure factorable interaction. Let $|\psi\rangle$ be the eigenvector and E the eigenvalue. It is immediate to check that the operator $V = c\,|V\rangle\langle V|$, where $c^{-1} = \langle\psi|\,(E-H_0)\,|\psi\rangle$ and $|V\rangle = (E-H_0)\,|\psi\rangle$, is the required operator.

If we don't restrict ourselves to the pure factorable interactions, the problem becomes highly indeterminate. In particular, we observe that we can add, to the previously determined interaction, the term

$$V' = \sum_{k=1}^{N} c_k\,|V_k\rangle\langle V_k| \tag{5.1}$$

provided

$$\langle\psi\mid V_k\rangle = 0 \qquad (k = 1, 2, \cdots, N) \tag{5.2}$$

is satisfied, obtaining an interaction which still admits the assigned bound state and eigenvalue. The further requirement that no more bound states are introduced still leaves clearly a large arbitrariness in the choice of the number N, of the $|V_k\rangle$'s, and of the c_k's. For example, as shown in Sec. 3, the conditions $c_k > 0$ for the coefficients in the interaction (5.1) is sufficient to guarantee the fullfillment of such a requirement. This arbitrariness can be used to modify the resulting T operator.

If n orthogonal bound states $|\psi_1\rangle$, $|\psi_2\rangle$, $\cdots$, $|\psi_n\rangle$, with negative energies E_1, E_2, $\cdots$, E_n, are to be reproduced, restricting ourselves to the superpositions of n factorable interactions, we can proceed as follows. Let us consider the interaction

$$V = -\sum_{i,k=1}^{n} X_{ik}\,|W_i\rangle\langle W_k|, \tag{5.3}$$

where

$$|W_i\rangle = (E_i - H_0)\,|\psi_i\rangle. \tag{5.4}$$

The unknown matrix X_{ik} will be determined in such a way that

$$(E_i - H_0)\,|\psi_i\rangle = -\sum_{i,k=1}^{n} X_{ik}\,|W_i\rangle\langle W_k\mid\psi_i\rangle, \tag{5.5}$$

[6] Only for the case of a pure factorable interaction, a different kind of inversion problem has been considered by M. Gourdin and A. Martin [Nuovo Cimento 8, 699 (1958)]. They showed that the assignment of a phase-shift function satisfying Levinson's theorem, together with the energy of the eventual bound state, allows to construct a unique pure factorable interaction. On the same line of thought Kh. Chadan [Nuovo Cimento 10, 892 (1958)] takes into account the case of a local interaction plus a pure factorable one.

i.e., owing to Eq. (5.4),

$$\sum_{j=1}^{n} \delta_{ij} |W_j\rangle = \sum_{j=1}^{n} \left[-\sum_{k=1}^{n} X_{jk} \langle W_k| G^{(0)}(E_i) |W_i\rangle \right] |W_j\rangle.$$

The linear independence of the $|W_j\rangle$'s, which follows from the orthogonality of the $|\psi_j\rangle$'s, as proved in Sec. 3, implies

$$\delta_{ij} = -\sum_{k=1}^{n} X_{jk} \langle W_k| G^{(0)}(E_i) |W_i\rangle. \tag{5.6}$$

Let us now show that, owing to the orthogonality of the $|\psi_i\rangle$'s, it holds the equality

$$\langle W_k| G^{(0)}(E_i) |W_i\rangle = \langle W_k| G^{(0)}(E_k) |W_i\rangle.$$

In fact,

$$\begin{aligned} &\langle W_k| G^{(0)}(E_i) |W_i\rangle - \langle W_k| G^{(0)}(E_k) |W_i\rangle \\ &\quad = \langle W_k| \frac{(E_k - H_0) - (E_i - H_0)}{(E_k - H_0)(E_i - H_0)} |W_i\rangle \\ &\quad = (E_k - E_i)\langle W_k| G^{(0)}(E_k) G^{(0)}(E_i) |W_i\rangle \\ &\quad = (E_k - E_i)\langle \psi_k \mid \psi_i\rangle = 0. \end{aligned}$$

Then we have

$$\begin{aligned} &-\langle W_k| G^{(0)}(E_i) |W_i\rangle \\ &\quad = -\tfrac{1}{2}\langle W_k| [G^{(0)}(E_k) + G^{(0)}(E_i)] |W_i\rangle \\ &\quad = (-E_i)\langle \psi_i \mid \psi_i\rangle \delta_{ik} + \langle \psi_k| H_0 |\psi_i\rangle \\ &\quad = \delta_{ik} + \langle \psi_k| H_0 |\psi_i\rangle = N_{ki} \end{aligned}$$

if the $|\psi_i\rangle$'s are chosen such that $\langle \psi_i | \psi_i\rangle = -1/E_i$. As already shown in Sec. 3, the diagonal elements of N_{ki}, when this is put in diagonal form, are all real and positive. In particular, N_{ki} admits the inverse matrix. Owing to (5.6), the unknown matrix X_{jk} is just this inverse matrix. The now-constructed interaction (5.4), which gives the assigned bound states, is not yet of the required type. It is easy to show that it can be written in the desired form in infinite ways. In fact we can write

$$V = \sum_{i=1}^{n} c_i |V_i\rangle\langle V_i|$$

with

$$|V_i\rangle = \sum_{j=1}^{n} a_{ij} |W_j\rangle,$$

if the c_i's and a_{ij}'s are such that

$$X_{jk} = \sum_{i=1}^{n} (-c_i) a_{ik}^* a_{ij} \tag{5.7}$$

Owing to the fact that the diagonal elements of X_{jk} too, when this is put in diagonal form, are all real and positive, Eqs. (5.7) can be solved in infinitely many ways. For example, if all the coefficients c_i, which, as proved in Section 3 must be negative, are chosen to be equal to -1, the matrix a_{ij} can be chosen to be the positive square root of the matrix X_{jk}. Note that the now obtained result proves that there are linear combinations of n factorable interactions which actually have n bound states.

The considerations made in the case of one assigned bound state can now be repeated for n assigned bound states. If we do not restrict ourselves to the linear combinations of n factorable interactions we can add to the otherwise unique interaction a term like (5.2) where the $|V_k\rangle$'s satisfy the conditions

$$\langle \psi_j \mid V_k\rangle = 0 \quad (j = 1, 2, \cdots, n; k = 1, 2, \cdots, N)$$

Here also the requirement that no further bound state is introduced still leaves a large arbitrariness which allows to modify the resulting T operator.

ACKNOWLEDGMENT

We thank Professor L. Fonda for his interest in this work.

APPENDIX

We wish to show here the completeness of the set of eigenstates belonging to the discrete and continuous spectrum of the operator $H = H_0 + c |V\rangle \langle V|$ introduced in Sec. 2. Let us consider for simplicity the case of a spinless particle. Let $|\mathbf{q}\rangle$ be the simultaneous eigenstate of the momentum operator and of H_0, belonging to the eigenvalues $\mathbf{q}$ and E, respectively. We shall normalize this state according to $\langle \mathbf{q} | \mathbf{q}'\rangle = \delta^3(\mathbf{q} - \mathbf{q}')$. If we call $|\mathbf{q}^+\rangle$ the scattering state of energy E containing an incident plane wave $|\mathbf{q}\rangle$ plus outgoing waves only, we have

$$|\mathbf{q}^+\rangle = |\mathbf{q}\rangle + G^{(0)+}(E) V |\mathbf{q}^+\rangle.$$

Projection on $|V\rangle$ gives

$$\langle V \mid \mathbf{q}^+\rangle = \langle V \mid \mathbf{q}\rangle + c\langle V| G^{(0)+}(E) |V\rangle\langle V \mid \mathbf{q}^+\rangle,$$

and thus

$$\langle V \mid \mathbf{q}^+\rangle = \langle V \mid \mathbf{q}\rangle / [1 - c\langle V| G^{(0)+}(E) |V\rangle].$$

On the other hand, projections on $|\mathbf{q}'\rangle$ give

$$\langle \mathbf{q}' \mid \mathbf{q}^+\rangle = \delta^3(\mathbf{q} - \mathbf{q}') + c\langle \mathbf{q}'| G^{(0)+}(E) |V\rangle\langle V \mid \mathbf{q}^+\rangle$$

or

$$\langle \mathbf{q}' \mid \mathbf{q}^+ \rangle = \delta^3(\mathbf{q} - \mathbf{q}') + \{c\langle \mathbf{q}' \mid V\rangle\langle V \mid \mathbf{q}\rangle/(E + i\epsilon - E') \times [1 - c\langle V| G^{(0)+}(E) |V\rangle]\}.$$

We can therefore write

$$|\mathbf{q}^+\rangle = |\mathbf{q}\rangle + \int a(q, q') |\mathbf{q}'\rangle \, d^3q', \qquad \text{(A1)}$$

where

$$a(q, q') = \frac{cv(q)v(q')}{\left[1 - c\int \frac{v^2(q'')\, d^3q''}{E + i\epsilon - E''}\right]} (E + i\epsilon - E'). \qquad \text{(A2)}$$

In (A2) we have called $v(q) \equiv \langle V|\mathbf{q}\rangle$. $v(q)$ is real from time-reversal invariance.

Let us now take into account $\int |\mathbf{q}^+\rangle \langle \mathbf{q}^+| \, d^3q$. We have

$$\int |\mathbf{q}^+\rangle\langle \mathbf{q}^+| \, d^3q = \int |\mathbf{q}\rangle\langle \mathbf{q}| \, d^3q + \iint d^3q \, d^3q' \, |\mathbf{q}'\rangle\langle \mathbf{q}| \left[a(q, q') + a^*(q', q) + \int a(q'', q')a^*(q'', q) \, d^3q''\right]. \qquad \text{(A3)}$$

We define

$$D(z) = 1 - c\int v^2(q') \, d^3q'/(z - E').$$

Assuming nonrelativistic kinematics $E = q^2/2\mu$ (the generalization to relativistic energies is trivial), after having integrated over angles we get

$$\int a(q'', q')a^*(q'', q) \, d^3q'' = 4\pi\mu c^2 v(q)v(q') \times \int_0^\infty \frac{v^2(q'')q'' \, dE''}{(E'' + i\epsilon - E')(E'' - i\epsilon - E)D(E'' + i\epsilon)D(E'' - i\epsilon)}.$$

It is easy to check that it holds the equality

$$\int a(q'', q')a^*(q'', q) \, d^3q'' = -\frac{cv(q)v(q')}{2\pi i} \times \int_C \frac{dz}{D(z)(z + i\epsilon - E')(z - i\epsilon - E)}, \qquad \text{(A4)}$$

where C is a clockwise contour which encloses the cut of $D(z)$, excluding the eventual zero of $D(z)$ on the negative real axis. The contour C of integration can be closed with the aid of a very large circle because the integral over the circle vanishes. The integral along the closed path is easily evaluated by means of the Cauchy integral formula. We must distinguish two cases.

A. No bound state.

In this case we immediately get

$$\int a(q'', q')a^*(q'', q) \, d^3q'' = -\frac{cv(q)v(q')}{(E + 2i\epsilon - E')\left[1 - c\int \frac{v^2(q'')\, d^3q''}{E + i\epsilon - E''}\right]} - \frac{cv(q)v(q')}{(E' - 2i\epsilon - E)\left[1 - c\int \frac{v^2(q'')\, d^3q''}{E' - i\epsilon - E''}\right]} = -a(q, q') - a^*(q', q).$$

Equation (A3) therefore gives

$$\int |\mathbf{q}^+\rangle\langle \mathbf{q}^+| \, d^3q = \int |\mathbf{q}\rangle\langle \mathbf{q}| \, d^3q,$$

which means completeness.

B. One bound state of energy E_b.

We have in this case

$$\int a(q'', q')a^*(q'', q) \, d^3q'' = -a(q, q') - a^*(q', q) - \frac{v(q)v(q')N^2}{(E_b - E')(E_b - E)},$$

where

$$N^{-2} = c^{-1}\left[\frac{dD(z)}{dz}\right]_{z=E_b} = \int \frac{v^2(q') \, d^3q'}{(E_b - E')^2}.$$

Equation (A3) gives

$$\int |\mathbf{q}^+\rangle\langle \mathbf{q}^+| \, d^3q = \int |\mathbf{q}\rangle\langle \mathbf{q}| \, d^3q - \iint |\mathbf{q}'\rangle\langle \mathbf{q}| \, d^3q \, d^3q' \frac{N^2 v(q)v(q')}{(E_b - E')(E_b - E)} = \int |\mathbf{q}\rangle\langle \mathbf{q}| \, d^3q - \left[N\int d^3q' \frac{v(q') \, |\mathbf{q}'\rangle}{E_b - E'}\right] \times \left[N\int d^3q \langle \mathbf{q}| \frac{v(q)}{E_b - E}\right] = \int |\mathbf{q}\rangle\langle \mathbf{q}| \, d^3q - |E_b\rangle\langle E_b|,$$

where $|E_b\rangle$ is the normalized state of energy E_b. Finally,

$$|E_b\rangle\langle E_b| + \int |\mathbf{q}^+\rangle\langle \mathbf{q}^+| \, d^3q = \int |\mathbf{q}\rangle\langle \mathbf{q}| \, d^3q.$$

The completeness has thus been proved.

The above arguments can be easily generalized to the class of interactions of Sec. 3.

Asymptotic Convergence and the Coulomb Interaction*†

JOHN D. DOLLARD
Department of Physics, Princeton University, Princeton, New Jersey
(Received 26 November 1963)

A definition of asymptotic convergence is given for nonrelativistic time-dependent scattering problems involving Coulomb potentials. Convergence proofs have been found both for potential and for n-body multichannel scattering. For pure Coulomb potential scattering, the Møller wave matrix is computed explicitly and found to have its usual meaning.

INTRODUCTION

IN the nonrelativistic time-dependent theory of scattering, an important role is played by the concept of asymptotic convergence. This concept arises in a natural way when we try to define transition amplitudes and the S operator in the time-dependent theory. Mathematically, asymptotic convergence problems typically take the following form: we are given a "free Hamiltonian" H_0 and a "Hamiltonian with interaction" H. We think of these operators as acting in a space L^2 of square-integrable functions. This space is viewed as a Hilbert space with the scalar product

$$(f, g) = \int d\mathbf{x}\bar{f}(\mathbf{x})g(\mathbf{x}),$$

and the norm

$$||f|| = (f, f)^{\frac{1}{2}}.$$

We define the (unitary) operators

$$U(t) = \exp(-iH_0t), \qquad W(t) = \exp(-iHt), \tag{1}$$

and

$$\Omega(t) = W^{-1}(t)U(t) = W(-t)U(t). \tag{2}$$

Then the typical asymptotic convergence problem is to show that there exist operators $\Omega^{\pm}$ (Møller wave matrices) such that for a certain class of functions $f \in L^2$

$$\lim_{t\to\pm\infty} ||\Omega(t)f - \Omega^{\pm}f|| = 0. \tag{3}$$

If Eq. (3) holds, we say that $\Omega(t)$ *converges strongly* to $\Omega^{\pm}$ on the class of functions f as $t \to \pm\infty$.

To see how such asymptotic convergence problems arise in collision theory we briefly discuss the scattering of a particle by a potential. For this purpose we take as our Hamiltonians

$$H_0 = -\Delta_x/2m \tag{4}$$

and

$$H = H_0 + V(\mathbf{x}). \tag{5}$$

The Schrödinger equation for the scattering problem is (we take $\hbar = 1$ throughout)

$$i(\partial f/\partial t)(\mathbf{x}, t) = Hf(\mathbf{x}, t). \tag{6}$$

At large negative times t we imagine that there is at hand a "freely moving" wave packet $U(t)f_1$, $f_1 \in L^2$, representing a particle that has just left a collimator (we have suppressed the dependence on $\mathbf{x}$, which is of no interest at present). We wish to calculate the probability amplitude that at large positive times t a freely moving wave packet $U(t)f_2$, $f_2 \in L^2$ will be found, provided that in the meantime the particle was moving with a time dependence governed by the Hamiltonian H. To obtain this probability amplitude, choose a large negative time t_1 and a large positive time t_2 and define two solutions $g_1(t)$, $g_2(t)$ of the Schrödinger equation such that $g_1(t)$ agrees with $U(t)f_1$ at $t = t_1$ and $g_2(t)$ agrees with $U(t)f_2$ at $t = t_2$. Clearly

$$\begin{aligned} g_1(t) &= W(t - t_1)U(t_1)f_1, \\ g_2(t) &= W(t - t_2)U(t_2)f_2. \end{aligned} \tag{7}$$

Then the required probability amplitude is given by the scalar product of g_1 and g_2:

$$\begin{aligned} T(t_2, t_1) &= (g_2(t), g_1(t)) \\ &= (W(-t_2)U(t_2)f_2, W(-t_1)U(t_1)f_1) \\ &= (\Omega(t_2)f_2, \Omega(t_1)f_1). \end{aligned} \tag{8}$$

The assumption that t_1 is large and negative and t_2 is large and positive can now be replaced by the condition $t_1 \to -\infty$, $t_2 \to +\infty$, provided that $T(t_2, t_1)$ converges as these limits are taken. Now $T(t_2, t_1)$ will surely converge if the operators $\Omega(t)$

* This work was supported in part by the National Science Foundation.

† The material in this paper is contained in a dissertation submitted to the faculty of Princeton University in partial fulfillment of the requirements for the Ph.D. degree.

converge strongly to operators $\Omega^{\pm}$ on every function $f \in L^2$. In this case we have

$$\lim_{\substack{t_1 \to -\infty \\ t_2 \to +\infty}} T(t_2, t_1) = (\Omega^+ f_2, \Omega^- f_1) \equiv T_{21}. \tag{9}$$

In addition [still assuming strong convergence of the $\Omega(t)$], we can define the S-operator[1]

$$S = \Omega^- \Omega^{+*} \tag{10}$$

(Ω^{+*} is the adjoint of Ω^+) which maps the set of states of the form $\Omega^+ f$ onto the set of states of the form $\Omega^- f$:

$$S\Omega^+ f = \Omega^- f. \tag{11}$$

Equation (11) holds because of the relation

$$\Omega^{+*}\Omega^+ = 1,$$

which follows from the fact that Ω^+ is the strong limit of a unitary operator. The expression T_{21} of Eq. (9) can then be viewed as the matrix element of S between $\Omega^+ f_2$ and $\Omega^+ f_1$. Thus, in order to establish the existence of the scattering amplitude and the S operator, it is natural to consider the asymptotic behavior of the operators $\Omega(t)$. In the multichannel case it is similarly desirable to define operators analogous to the $\Omega(t)$ given above and to study their asymptotic behavior. A number of authors[2] have discussed these operators under various assumptions on the potentials appearing in the Hamiltonian H. Convergence proofs have, for instance, been given in the multichannel case for potentials which are square-integrable. However, no convergence proofs have been given for problems involving Coulomb potentials.

The purpose of this paper is to supply asymptotic convergence theorems for problems involving Coulomb interactions, both for potential and for n-body multichannel scattering.[3] [At this point, the author wishes to mention the independent work of M. Saffouri (to be published) who has investigated the asymptotic behavior of wave packets describing a particle moving under the influence of a Coulomb potential. His results are similar to some of those presented here for potential or two-body scattering.][4] We shall find that when Coulomb potentials are involved in a scattering problem, it is necessary to modify the definition of asymptotic convergence in order to obtain a convergence proof. Essentially, this is a result of the long range of the Coulomb potentials. This modification of the definition of asymptotic convergence will necessitate a modification of the picture of the scattering process presented above. In order to make clear the nature of this modification we shall first deal in some detail with scattering by a short-range (square-integrable) potential. A proof of asymptotic convergence will be given because the reader may not be familiar with such proofs and because the author feels that the proof to be presented, while modeled on the work of previous authors[2], is simpler and more easily extendible to n-body problems.

I. SCATTERING BY A SHORT-RANGE POTENTIAL

We first give a convergence proof for the operator $\Omega(t)$ with H_0 and H given by (4) and (5), respectively, under the assumption that $V(\mathbf{x})$ is a square-integrable function. In this case (Kato[5]), the operator H is essentially self-adjoint on the space L^2, and both $U(t)$ and $W(t)$ are unitary operators. We begin by showing that the sequence $\Omega(t)f$ converges when f belongs to the space $\mathcal{S}$ of functions discussed by Schwartz[6]: $f(\mathbf{x})$ belongs to $\mathcal{S}$ if and only if

(1) f is infinitely differentiable with respect to each of the components x_1, x_2, x_3 of $\mathbf{x}$.

(2) f and all of its derivatives decrease rapidly for large $|\mathbf{x}|$. Explicitly, for any integers m_1, m_2, m_3 and n, the expression

$$|\mathbf{x}|^n \left| \frac{\partial^{m_1+m_2+m_3}}{\partial x_1^{m_1}\, \partial x_2^{m_2}\, \partial x_3^{m_3}} f(\mathbf{x}) \right|$$

is bounded. (Not necessarily the same bound for all m_1, m_2, m_3, n.) If a function $f(\mathbf{x})$ has properties (1) and (2), its Fourier transform $\hat{f}(\mathbf{K})$ has the same properties.

To prove that $\Omega(t)f$ converges for $f \in \mathcal{S}$ as $t \to +\infty$, say, we use the following lemma:

Lemma 1. Suppose that the derivatives $(\partial/\partial t)(\Omega(t)f)$ exist, belong to L^2, and are continuous with respect to t in the L^2 norm. In order that the sequence $\Omega(t)f$ converge strongly as $t \to +\infty$, it suffices that the integral

$$I(f, t_0) = \int_{t_0}^{\infty} \left\| \frac{\partial}{\partial t} (\Omega(t)f) \right\| dt \tag{12}$$

exist for some value of t_0.

[1] W. Brenig and R. Haag, Fortschr. Physik **7**, 183 (1959).

[2] J. M. Cook, J. Math. and Phys. **36**, 82 (1957); M. N. Hack, Nuovo Cimento **13**, 231 (1959); J. M. Jauch and I. Zinnes, *ibid.* **11**, 553 (1959).

[3] To conserve space, it was necessary to omit some of the mathematical details in this paper. These are given in full in the author's thesis, Princeton University, 1963.

[4] I wish to thank Professor R. Haag of the University of Illinois for communicating to me the results of M. Saffouri's investigation.

[5] T. Kato, Trans. Am. Math. Soc. **70**, 195 (1951).

[6] L. Schwartz, *Theorie des Distributions* (Hermann & Cie., Paris, 1957), Vol. II, p. 89ff.

Proof: We have

$$||\Omega(t_1)f - \Omega(t_2)f|| = \left|\left|\int_{t_1}^{t_2} dt\, \frac{\partial}{\partial t}(\Omega(t)f)\right|\right| \leq \int_{t_1}^{t_2} dt \left|\left|\frac{\partial}{\partial t}(\Omega(t)f)\right|\right|, \tag{13}$$

and the convergence of the integral in (12) implies that the right-hand side of (13) can be made arbitrarily small by choosing t_1 and t_2 large enough. Thus $\Omega(t)f$ is a Cauchy sequence and is convergent.

If $f \in \mathcal{S}$ and $\Omega(t)$ is given by (2) it is easy to show, under our assumptions on the potential $V(\mathbf{x})$, that $(\partial/\partial t)(\Omega(t)f)$ exists and is continuous with respect to t in the L^2 norm. To establish the existence of $I(f, t_0)$ for $t_0 > 0$ we compute

$$\begin{aligned}\frac{\partial}{\partial t}(\Omega(t)f) &= \frac{\partial}{\partial t}(W^{-1}(t)U(t)f) \\ &= iW^{-1}(t)(H - H_0)U(t)f \\ &= iW^{-1}(t)VU(t)f, \end{aligned} \tag{14}$$

and

$$||(\partial/\partial t)(\Omega(t)f)|| = ||VU(t)f||, \tag{15}$$

since $W^{-1}(t)$ is unitary. We now show that $I(f, t_0) = \int_{t_0}^{\infty} dt\, ||VU(t)f||$ exists by exploiting the following two facts:

(1) $V(\mathbf{x})$ is square-integrable. (Note that this is false for a Coulomb potential e^2/x, and that the proof therefore does not hold for such potentials.)

(2) $(U(t)f)(\mathbf{x})$ vanishes at any point $\mathbf{x}$ as $1/t^{\frac{3}{2}}$.

The second fact follows easily from the well-known representation[7] ($t \neq 0$)

$$(U(t)f)(\mathbf{x}) = \left(\frac{m}{2\pi i t}\right)^{\frac{3}{2}} \int d\mathbf{x}' e^{im(\mathbf{x}-\mathbf{x}')^2/2t} f(\mathbf{x}'), \tag{16}$$

which immediately yields

$$|(U(t)f)(\mathbf{x})| \leq c/t^{\frac{3}{2}} \qquad (t > 0). \tag{17}$$

We can now estimate

$$||VU(t)f|| \leq (c/t^{\frac{3}{2}})\, ||V||, \tag{18}$$

showing that $I(f, t_0)$ exists for $t_0 > 0$ and $\Omega(t)f$ converges as $t \to +\infty$ when $f \in \mathcal{S}$. The proof as $t \to -\infty$ is no different.

We can now prove that $\Omega(t)g$ converges for any $g \in L^2$ by using the fact that the space $\mathcal{S}$ is dense in the space L^2, i.e., for any $g \in L^2$ and any $\epsilon > 0$ there is an $f \in \mathcal{S}$ such that $||f - g|| \leq \epsilon$. Using this fact and the unitarity of $\Omega(t)$ we have for any $g \in L^2$ and any $\epsilon > 0$

$$\begin{aligned}||\Omega(t_1)g - \Omega(t_2)g|| &\leq ||\Omega(t_1)g - \Omega(t_1)f|| \\ &\quad + ||\Omega(t_1)f - \Omega(t_2)f|| + ||\Omega(t_2)f - \Omega(t_2)g|| \\ &\leq ||\Omega(t_1)f - \Omega(t_2)f|| + 2\epsilon. \end{aligned} \tag{19}$$

The fact that $\Omega(t)f$ is a Cauchy sequence now implies the same for $\Omega(t)g$, proving the convergence of the latter sequence. Thus the sequence of operators $\Omega(t)$ converges strongly on the entire space L^2:

$$\lim_{t \to \pm\infty} \Omega(t) = \Omega^{\pm}. \tag{20}$$

We now mention a result due to Ikebe[8] which will aid in the interpretation of our later results. Under somewhat more restrictive assumptions on the potential $V(\mathbf{x})$, Ikebe proves the following: Let $\psi_{\mathbf{K}}^{(-)}(\mathbf{x})$ and $\psi_{\mathbf{K}}^{(+)}(\mathbf{x})$ be, respectively, the stationary-state scattering solutions for our problem corresponding to incoming and outgoing momentum $\mathbf{K}$. [We choose the normalization so that $\psi_{\mathbf{K}}^{(\pm)}(\mathbf{x})$ has leading term $e^{i\mathbf{K}\cdot\mathbf{x}}$.] Then either of the sets $\{\psi_{\mathbf{K}}^{(-)}(\mathbf{x})\}$ or $\{\psi_{\mathbf{K}}^{(+)}(\mathbf{x})\}$, taken together with the bound states for our problem, is complete in the space L^2. Further, if $f \in L^2$ and

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{3}{2}}} \int d\mathbf{K} e^{i\mathbf{K}\cdot\mathbf{x}} \hat{f}(\mathbf{K}), \tag{21}$$

then

$$(\Omega^{\pm}f)(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{3}{2}}} \int d\mathbf{K} \psi_{\mathbf{K}}^{(\pm)}(\mathbf{x}) \hat{f}(\mathbf{K}). \tag{22}$$

Equation (22) is of interest to us for two reasons: first, in view of the symbolic relations

$$\begin{aligned}(\psi_{\mathbf{K}}^{(+)}, \psi_{\mathbf{K}}^{(-)}) &= (2\pi)^3 \delta(\mathbf{K} - \mathbf{K}') \\ &\quad + (2\pi)^3 \frac{i}{\pi} \delta(K^2 - K'^2) f(\mathbf{K}, \mathbf{K}'), \end{aligned} \tag{23}$$

where $f(\mathbf{K}, \mathbf{K}')$ is the scattering amplitude from momentum $\mathbf{K}$ to momentum $\mathbf{K}'$, (22) shows that the transition amplitude calculated using the definition (9) agrees with the transition amplitude calculated from the time-independent theory. Second, (22) shows that any function $g \in L^2$ which has no bound-state components can be written in the form $g = \Omega^{\pm}h^{\pm}$, $h^{\pm} \in L^2$, since any such g can be written in the form given on the right-hand side of (22). We can use this fact to show that any scattering solution (no bound components) $W(t)g$ of the Schrödinger equation becomes "asymptotically free" in the sense that, for large t, $W(t)g$ approaches a solution of the free Schrödinger equa-

[7] P. J. Redmond and J. L. Uretsky, Ann. Phys. (N. Y.) 9, 106 (1960).

[8] T. Ikebe, Arch. Ratl. Mech. Anal. 5, 1 (1960). A slight modification of Kato's argument is necessary since we do not assume that V is bounded for large $|\mathbf{x}|$.

tion. In fact, writing $g = \Omega^{\pm}h^{\pm}$, we have

$$\lim_{t\to\pm\infty} ||W(t)g - U(t)h^{\pm}|| = \lim_{t\to\pm\infty} ||W(t)\Omega^{\pm}h^{\pm} - U(t)h^{\pm}|| = \lim_{t\to\pm\infty} ||\Omega^{\pm}h^{\pm} - W^{-1}(t)U(t)h^{\pm}|| = 0. \quad (24)$$

The asymptotic freedom of the scattering solution $W(t)g$ establishes the correctness of the picture of scattering outlined in the introduction, since it justifies the assumption that "at large negative times a wave packet $U(t)f_1$ is at hand" and the similar statement for large positive times. This may seem to be a trivial point, but it has been stressed here because the above-mentioned assumptions are *not* justified when a Coulomb potential is present. In fact we shall see that in a Coulomb field the asymptotic form of a wave-packet is given by $U_c(t)f$ where $U_c(t)$ differs from $U(t)$ as defined above. Since this is a departure from the notion of asymptotic freedom in the usual quantum-mechanical sense, we shall examine this notion somewhat more closely before turning to the case of Coulomb scattering: Let $f(\mathbf{x}) \in L^2$ and let $\hat{f}(\mathbf{K})$ be the Fourier transform of $f(\mathbf{x})$. As is well-known, if $|f(\mathbf{x})|^2$ determines the probability density for the coordinates of a particle, $|\hat{f}(\mathbf{K})|^2$ determines the probability density for its momenta. We wish to call attention to the following property of the function $U(t)f$: at large positive and negative times, the probability density determined by $U(t)f$ becomes the probability density for a free classical particle which at time $t = 0$ is located at the origin and has a momentum spectrum given by $|\hat{f}(\mathbf{K})|^2$. Indeed, using the representation (16) for $f \in \mathcal{S}$ and extending the results to all of L^2, it is not difficult to show that

$$\lim_{t\to\pm\infty} \left|\left| U(t)f - \left(\frac{m}{it}\right)^{\frac{3}{2}} e^{imx^2/2t} \hat{f}\left(\frac{m\mathbf{x}}{t}\right) \right|\right| = 0. \quad (25)$$

The strong convergence to 0 in (25) indicates that the probability density $|U(t)f|^2$ can be replaced asymptotically by $(m/|t|)^3\ |\hat{f}(m\mathbf{x}/t)|^2$, which is the classical probability distribution mentioned above. This asymptotic behavior of $U(t)f$ provides a criterion of asymptotic freedom which is physically intuitive and which embodies the essential requirements that we want to impose on incoming and outgoing wave packets. We shall see that wave packets in a Coulomb field have the proper "classical" asymptotic behavior although they do not have the asymptotic form $U(t)f$. We now turn to the study of Coulomb scattering.

II. SCATTERING BY A COULOMB POTENTIAL

We first state and prove the theorem on asymptotic convergence in a Coulomb field.

Theorem 1. Let

$$H_c = -\Delta/2m + e_1e_2/x \quad (26)$$

be the Hamiltonian governing the scattering of a particle by a Coulomb potential. Define the operator

$$W_c(t) = \exp(-iH_ct) \quad (27)$$

and the "distorted free propagation operator"

$$U_c(t) = \exp(-iH_{0c}(t)) \quad (28)$$

with

$$H_{0c}(t) = H_0t + \frac{\epsilon(t)me_1e_2}{(-\Delta)^{\frac{1}{2}}} \log\left(-\frac{2\,|t|\,\Delta}{m}\right) = H_0t + H'_{0c}(t), \quad (29)$$

where $H_0 = -\Delta/2m$ and

$$\epsilon(t) = \begin{cases} 1 & t > 0, \\ -1 & t < 0, \end{cases} \quad (30)$$

and the action of $H'_{0c}(t)$ is defined by passing to momentum space: if

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{3}{2}}} \int d\mathbf{K} e^{i\mathbf{K}\cdot\mathbf{x}} \hat{f}(\mathbf{K}), \quad (31)$$

then

$$H'_{0c}(t)f(\mathbf{x}) = \frac{\epsilon(t)me_1e_2}{(2\pi)^{\frac{3}{2}}} \int d\mathbf{K} e^{i\mathbf{K}\cdot\mathbf{x}} \log\left(\frac{2K^2\,|t|}{m}\right) \frac{\hat{f}(\mathbf{K})}{K}, \quad (32)$$

provided that the integral in (32) exists.

Then $W_c^{-1}(t)U_c(t) = \Omega_c(t)$ is a unitary operator and the strong limits

$$\lim_{t\to\pm\infty} \Omega_c(t) = \Omega_c^{\pm} \quad (33)$$

exist on all of L^2.

Before giving the proof of this theorem, we remark that, for $t > 0$ or $t < 0$, $H'_{0c}(t)$ can be written as the sum of a time-dependent and a time-independent part. For $t > 0$, for instance,

$$H'_{0c}(t) = \frac{me_1e_2}{(-\Delta)^{\frac{1}{2}}} \log t + \frac{me_1e_2}{(-\Delta)^{\frac{1}{2}}} \log\left(-\frac{2\Delta}{m}\right) \quad (t > 0). \quad (34)$$

Clearly the second, time-independent term can play no role in a convergence proof. This term is included in order that the analog of Eq. (22) should hold

in the present case. We shall return later to a discussion of this point. At present, we give the proof of Theorem 1.

Proof of Theorem 1: The unitarity of $W_c^{-1}(t)$ follows from the essential self-adjointness of H_c (Kato[5]). The unitarity of $U_c(t)$ is easily established using the representation of $U_c(t)$ in momentum space. Thus $\Omega_c'(t)$ is also unitary. We will give the convergence proof for $\Omega_c(t)$ as $t \to +\infty$. We first prove convergence on the space of *C functions*, defined as follows: $f(\mathbf{x})$ is a C function if $f \in \mathcal{S}$ and its Fourier transform $\hat{f}(\mathbf{K})$ vanishes in a neighborhood of $\mathbf{K} = 0$. It is not hard to show that the C functions are dense in L^2 and that if f is a C function then $U_c(t)f$ and $[1/(-\Delta)^{\frac{1}{2}}]U_c(t)f = U_c(t)[1/(-\Delta)^{\frac{1}{2}}]f$ are C functions. Proceeding as in the previous convergence proof, we let f be a C function and estimate the norm of $(\partial/\partial t)(\Omega_c(t)f)$:

$$\begin{aligned}\frac{\partial}{\partial t}(\Omega_c(t)f) &= iW_c^{-1}(t)\left[H - \frac{\partial H_{0c}(t)}{\partial t}\right]U_c(t)f \\ &= iW_c^{-1}(t)\left[\frac{e_1e_2}{x} - \frac{me_1e_2}{t(-\Delta)^{\frac{1}{2}}}\right]U_c(t)f\end{aligned} \qquad (t > 0), \tag{35}$$

so that $(t > 0)$

$$\left\|\frac{\partial}{\partial t}(\Omega_c(t)f)\right\| = \left\|\left[\frac{e_1e_2}{x} - \frac{me_1e_2}{t(-\Delta)^{\frac{1}{2}}}\right]U_c(t)f\right\|. \tag{36}$$

Comparing (36) with (15) we see that the effect of using the operator $U_c(t)$ instead of $U(t)$ is to subtract a term from the potential. Essentially, the rest of the analysis consists in showing that the subtracted term tends to cancel against e_1e_2/x. In order to do this we prove the following Lemma:

Lemma 2. Let $h(\mathbf{x})$ be a C function and let $\hat{h}(\mathbf{K})$ be its Fourier transform. Then for $|t| \geq t_0 > 1$

$$\begin{aligned}U_c(t)h(\mathbf{x}) &= \left(\frac{m}{it}\right)^{\frac{3}{2}}\varphi_c(\mathbf{x})\hat{h}\left(\frac{m\mathbf{x}}{t}\right) \\ &\quad + \left(\frac{m}{2\pi it}\right)^{\frac{3}{2}}e^{imx^2/2t}R_h(\mathbf{x}, t),\end{aligned} \tag{37}$$

where

$$\varphi_c(\mathbf{x}) = e^{imx^2/2t}e^{-ie(t)me_1e_2t/x\log(2mx^2/|t|)}, \tag{38}$$

and for any integer n there exists a number μ and a constant C such that

$$|R_h(\mathbf{x}, t)| \leq \frac{C(\log|t|)^\mu}{|t|^{\frac{3}{2}}\left[1 + \left(\frac{x}{t}\right)^2\right]^n}. \tag{39}$$

Lemma 2 implies, among other things, the relation

$$\lim_{t\to\pm\infty}\left\|U_c(t)h(\mathbf{x}) - \left(\frac{m}{it}\right)^{\frac{3}{2}}\varphi_c(\mathbf{x})\hat{h}\left(\frac{m\mathbf{x}}{t}\right)\right\| = 0, \tag{40}$$

which is the present analog of the relation (25) established earlier.

To prove Lemma 2, we write

$$U_c(t)h(\mathbf{x}) = U(t)h_c'(\mathbf{x}, t) \tag{41}$$

with

$$\begin{aligned}h_c'(\mathbf{x}, t) &= \exp(-iH_{0c}'(t))h(\mathbf{x}) \\ &= \frac{1}{(2\pi)^{\frac{3}{2}}}\int d\mathbf{K}e^{i\mathbf{K}\cdot\mathbf{x}}e^{-ie(t)me_1e_2/K\log(2K^2|t|/m)}\hat{h}(\mathbf{K}).\end{aligned} \tag{42}$$

Integrating by parts in (42), remembering that $\hat{h}(\mathbf{K}) \in \mathcal{S}$ and $\hat{h}(\mathbf{K})$ vanishes in a neighborhood of $\mathbf{K} = 0$, we find that for any integers m_1, m_2, m_3, n there exist constants C and μ such that

$$\left|(1 + x^2)^n\frac{\partial^{m_1+m_2+m_3}}{\partial x_1^{m_1}\,\partial x_2^{m_2}\,\partial x_3^{m_3}}h_c'(\mathbf{x}, t)\right| \leq C(\log|t|)^\mu. \tag{43}$$

(Note $\log|t| > 0$ since $|t| > 1$). From (41) we find [compare (16)]

$$\begin{aligned}U_c(t)h(\mathbf{x}) &= \left(\frac{m}{2\pi it}\right)^{\frac{3}{2}}\int d\mathbf{x}'e^{im(\mathbf{x}-\mathbf{x}')^2/2t}h_c'(\mathbf{x}', t) \\ &= \left(\frac{m}{2\pi it}\right)^{\frac{3}{2}}e^{imx^2/2t}\int d\mathbf{x}'e^{-im\mathbf{x}\cdot\mathbf{x}'/t}h_c'(\mathbf{x}', t) \\ &\quad + \left(\frac{m}{2\pi it}\right)^{\frac{3}{2}}e^{imx^2/2t}\int d\mathbf{x}'e^{-im\mathbf{x}\cdot\mathbf{x}'/t} \\ &\quad \times (e^{imx'^2/2t} - 1)h_c'(\mathbf{x}', t).\end{aligned} \tag{44}$$

Using the Fourier transform of $h_c'(\mathbf{x}, t)$ from (42) we find that (44) is the same as (37) with

$$R_h(\mathbf{x}, t) = \int d\mathbf{x}'e^{-im\mathbf{x}\cdot\mathbf{x}'/t}(e^{imx'^2/2t} - 1)h_c'(\mathbf{x}', t). \tag{45}$$

Using in Eq. (45) the estimate

$$|e^{imx'^2/2t} - 1| \leq \left|\frac{mx'^2}{2t}\right| \tag{46}$$

and one of the bounds (43) on $h_c'(\mathbf{x}', t)$, we can find a C and a μ such that

$$|R_h(\mathbf{x}, t)| \leq C(\log|t|)^\mu/|t|. \tag{47}$$

On the other hand, integrating by parts in (45) and using the bounds (43), given any n we can find C and μ such that

$$|R_h(\mathbf{x}, t)| \leq \frac{C(\log|t|)^\mu}{[1 + (x/t)^2]^{2n}}. \tag{48}$$

The bound (39) is now obtained by multiplying

the square roots of the bounds in (47) and (48). This completes the proof of Lemma 2. We now use Lemma 2 on the C functions occurring in (36). For convenience in writing, we shall denote $1/(-\Delta)^{\frac{1}{2}}f$ by f'. Note that f' is a C function and that

$$\hat{f}'(\mathbf{K}) = \hat{f}(\mathbf{K})/K. \tag{49}$$

Using the fact that $1/(-\Delta)^{\frac{1}{2}}$ commutes with U_c and applying Lemma 2 to f and f', we have ($t \geq t_0$)

$$\begin{aligned}\left\|\frac{\partial}{\partial t}(\Omega_c(t)f)\right\| &= \left\|\frac{e_1e_2}{x}U_c(t)f - \frac{me_1e_2}{t}U_c(t)f'\right\| \\ &= \left\|\left(\frac{m}{it}\right)^{\frac{3}{2}}\varphi_c(\mathbf{x})\left[\frac{e_1e_2}{x}\hat{f}\left(\frac{m\mathbf{x}}{t}\right) - \frac{me_1e_2}{t}\hat{f}'\left(\frac{m\mathbf{x}}{t}\right)\right]\right. \\ &\quad \left. + \left(\frac{m}{2\pi it}\right)^{\frac{3}{2}}e^{imx^2/2t}\left[\frac{e_1e_2}{x}R_f(\mathbf{x}, t) - \frac{me_1e_2}{t}R_{f'}(\mathbf{x}, t)\right]\right\|.\end{aligned} \tag{50}$$

Now the terms involving $\hat{f}$ and $\hat{f}'$ on the right-hand side of (50) cancel because of (49). The remaining terms can be estimated separately using (39). For instance,

$$\begin{aligned}&\left|\frac{m}{2\pi it}\right|^{\frac{3}{2}}\frac{|me_1e_2|}{t}\,||R_{f'}(\mathbf{x}, t)|| \\ &\leq \frac{C(\log t)^{\mu}}{t^{\frac{3}{2}}}\frac{1}{t^{\frac{3}{2}}}\left\|\frac{1}{\left[1 + \left(\frac{x}{t}\right)^2\right]^n}\right\| = \frac{C'(\log t)^{\mu}}{t^{\frac{3}{2}}},\end{aligned} \tag{51}$$

and similarly for the term in R_f. All told, we obtain

$$||(\partial/\partial t)(\Omega_c(t)f)|| \leq C(\log t)^{\mu}/t^{\frac{3}{2}} \tag{52}$$

for some constants C and μ. Since the right-hand side of (52) is integrable with respect to t over the interval from $t_0 > 1$ to ∞, Lemma 1 now shows that $\Omega_c(t)$ converges strongly on any C function. [With regard to the hypotheses of this lemma it should perhaps be mentioned that there is no difficulty in showing that when f is a C function $(\partial/\partial t)(\Omega_c(t)f)$ is continuous with respect to t in the L^2 norm for $t \neq 0$.] We can now extend our convergence proof to all of L^2 as before, using the fact that the C functions are dense in L^2. Thus we have proved (33) for the case $t \to +\infty$. The proof as $t \to -\infty$ is no different.

We now state the analog of (22) in the present case:

Theorem 2. Let $\psi_{c\mathbf{K}}^{(-)}(\mathbf{x})(\psi_{c\mathbf{K}}^{(+)}(\mathbf{x}))$ be, respectively, the stationary state scattering solutions in the pure Coulomb field e_1e_2/x corresponding to incoming (outgoing) momentum $\mathbf{K}$,[9]

$$\begin{aligned}\psi_{c\mathbf{K}}^{(-)}(\mathbf{x}) &= e^{-\frac{1}{2}\pi n}\Gamma(1 + in)\;{}_1F_1(-in, 1, i(Kx - \mathbf{K}\cdot\mathbf{x})) \\ \psi_{c\mathbf{K}}^{(+)}(\mathbf{x}) &= \overline{(\psi_{c,-\mathbf{K}}^{(-)}(\mathbf{x}))},\end{aligned} \tag{53}$$

where $n = me_1e_2/K$. Then if $\hat{f}(\mathbf{K})$ is the Fourier transform of $f(\mathbf{x})$,

$$(\Omega_c^{\pm}f)(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{3}{2}}}\int d\mathbf{K}\psi_{c\mathbf{K}}^{(\pm)}(\mathbf{x})\hat{f}(\mathbf{K}). \tag{54}$$

This shows that the new Møller wave matrices $\Omega_c^{\pm}$ have their customary meaning, as in (22).

The proof of Theorem 2 will not be given here, as the only proof which the author has been able to find is extremely long and tedious, although very simple in principle.[3] We shall now use Theorems 1 and 2 to outline the scattering process in a Coulomb field.

We first study the behavior of the scattering solutions of the Coulomb Schrödinger equation. These have the form $W_c(t)g$ where g is orthogonal to all the Coulomb bound states. In view of the known orthogonality and completeness theorems for the Coulomb wavefunctions,[10] g can be expanded in terms of the $\psi_{c\mathbf{K}}^{(+)}$ or the $\psi_{c\mathbf{K}}^{(-)}$. Thus according to Theorem 2 we have $g = \Omega_c^{\pm}h^{\pm}$. But then Theorem 1 implies that

$$\lim_{t\to\pm\infty}||W_c(t)g - U_c(t)h^{\pm}|| = 0, \tag{55}$$

so that asymptotically the time behavior of such a wave packet is governed by $U_c(t)$ rather than $U(t)$. This shows that the only possible asymptotic form for a wave packet in a Coulomb field is $U_c(t)h^{\pm}(\mathbf{x})$. The logarithmic phase distortion in $U_c(t)$ can be shown to be a result of the logarithmic phase distortion in the asymptotic (large $|\mathbf{x}|$) expansion of $\psi_{c\mathbf{K}}^{(\pm)}(\mathbf{x})$ if we multiply both sides of (54) by $W_c(t)$ and take limits as $t \to \pm\infty$.

We can now construct the following picture of scattering in a Coulomb field: at large negative times a wave packet $U_c(t)f_1$ is present. Although this packet is not "free" in the usual quantum-mechanical sense, in view of (40) (which can easily be extended from C functions to any $h \in L^2$), we find that at large negative times $(U_c(t)f_1)(\mathbf{x})$ converges strongly to $(m/it)^{\frac{3}{2}}\varphi_c(\mathbf{x})\hat{f}_1(m\mathbf{x}/t)$, and that the probability density determined by $U_c(t)f$ can therefore be replaced by the "classical" probability density $|m/t|^3\;|\hat{f}_1(m\mathbf{x}/t)|^2$, so that asymptotically $|U_c(t)f_1|^2$ satisfies our intuitive physical criterion of "freedom."

We now ask for the probability amplitude T_{21}

[9] L. Schiff, *Quantum Mechanics* (McGraw-Hill Book Company, Inc., New York, 1955), p. 117.

[10] E. C. Titchmarsh, *Eigenfunction Expansions* (Clarendon Press, Oxford, England, 1946), Vol. I.

that at large positive times a wave packet $U_c(t)f_2$ will be found. Proceeding as in the introduction we find

$$T_{21} = \lim_{\substack{t_2\to+\infty \\ t_1\to-\infty}} (W_c^{-1}(t_2)U_c(t_2)f_2,\ W_c^{-1}(t_1)U_c(t_1)f_1) = (\Omega_c^+ f_2,\ \Omega_c^- f_1). \tag{56}$$

The interpretation of T_{21} as the desired transition amplitude is corroborated by Eq. (54) [see the discussion following (22)]. The S operator can now be defined in complete analogy with what was done in the Introduction.

This completes our discussion of potential scattering in a pure Coulomb field.

III. n-BODY PROBLEMS

Since the time-dependent formalism for dealing with n-body problems does not seem to be generally known, we give a sketch of the formalism applicable when short-range potentials are involved before stating the corresponding results for problems involving Coulomb interactions.

We consider n particles of masses $m_1 \cdots m_n$ interacting with each other and with static potentials as described by the Hamiltonian

$$H = -\sum_{i=1}^{n} \frac{\Delta_i}{2m_i} + \sum_{i=1}^{n} V_{0i}(\mathbf{x}_i) + \sum_{1\le i<j\le n} V_{ij}(\mathbf{x}_i - \mathbf{x}_j). \tag{57}$$

It may happen that the part of H describing the mutual interaction of certain particles $i_1, i_2, \cdots i_j$ allows bound states to be formed. In this case we may imagine that, at large negative times, particles $i_1, \cdots i_j$ can enter the scattering picture in various ways—either as j free particles, or bound together in any of the ways allowed by their mutual interaction. In case these particles are bound, we shall refer to them collectively as a *composite particle.* The wavefunction describing such a composite particle at large negative times can be written as $U_j(t)f\psi_j$, where f is a square-integrable wave packet in the center-of-mass coordinate, ψ_j is a bound-state wavefunction of energy E_j, written in terms of the internal coordinates of the composite particle, and

$$U_j(t) = \exp\,[-i(-\Delta_y/2M + E_j)t], \tag{58}$$

where $\mathbf{y}$ is the center-of-mass coordinate and $M = m_{i_1} + \cdots + m_{i_j}$ is the mass of the composite particle. It should be clear that $U_j(t)$ provides the usual time-dependence for the bound state ψ_j and propagates the wave packet f according to the free Schrödinger equation with mass M.

It may also happen that initially particles $i_1 \cdots i_j$ are bound by the static potentials and thus localized near the origin. This situation is described by the wavefunction $e^{-iE_B t}\psi_B$, where ψ_B is the bound-state wavefunction (of energy E_B) describing the configuration of the particles.

Taking all possible subsets $i_1 \cdots i_j$, $j = 1, \cdots n$ of particles and enumerating the bound states which they can form, we find all the ways in which particles can enter the scattering situation.[11] The most general situation at large negative times is that the n particles under consideration are grouped as l simple particles and m composite particles, plus possibly a number of particles bound near the origin. Any such grouping, along with a specification of the bound-state wavefunctions describing the composite particles and the particles bound near the origin, constitutes a *channel* of the system, and we shall distinguish various channels by means of a subscript α. By relabeling the particles if necessary, we may assume that the simple particles have numbers $1, \cdots l$. We denote the bound states describing the composite particles occurring in channel α by $\psi_{\alpha_1} \cdots \psi_{\alpha_m}$, and the associated energies and masses by $E_{\alpha_1} \cdots E_{\alpha_m}$ and $M_1 \cdots M_m$. Each bound state ψ_{α_j} depends only on the internal coordinates appropriate to the set of bound simple particles which it describes. We denote by $\mathbf{Y}_1 \cdots \mathbf{Y}_m$ the center-of-mass coordinates of the composite particles. The bound-state wavefunction describing the particles (if any) bound near the origin is denoted by ψ_B, and its energy by E_B. We suppress entirely the coordinates of ψ_B, which are of no interest in the following formulae. (If there are no particles bound at the origin ψ_B should be replaced by unity and E_B by zero in (59) and (61).) Then the wavefunction corresponding to the situation described at large negative times is given by $U_\alpha(t)\varphi_\alpha$, where

$$\varphi_\alpha = f(\mathbf{x}_1 \cdots \mathbf{x}_l;\ \mathbf{Y}_1 \cdots \mathbf{Y}_m)\psi_{\alpha_1} \cdots \psi_{\alpha_m}\psi_B \tag{59}$$

(f is any function which is square-integrable in all its arguments), and

$$U_\alpha(t) = \exp\,(-iH_\alpha t), \tag{60}$$

with

$$H_\alpha = -\sum_{j=1}^{l} \frac{\Delta_{x_j}}{2m_j} - \sum_{j=1}^{m} \frac{\Delta_{Y_j}}{2M_j} + \sum_{j=1}^{m} E_{\alpha_j} + E_B. \tag{61}$$

H_α is called the *channel Hamiltonian* for the configuration described, and $U_\alpha(t)$ is called the *channel operator* for this configuration. Note that the defini-

[11] By convention we exclude from consideration those cases in which *all* particles are bound near the origin (this would correspond to taking for ψ_B a normalizable eigenfunction of the full Hamiltonian H) since such cases do not lead to any scattering processes.

tions above include as a special case the free Hamiltonian H_0 and the free propagation operator $U_0(t)$

$$H_0 = -\sum_{i=1}^{n} \frac{\Delta_i}{2m_i} \qquad U_0(t) = \exp(-iH_0t) \tag{62}$$

corresponding to $l = n$, $m = 0$ above (i.e., all particles asymptotically free) and in this case φ_α is any square-integrable function of $\mathbf{x}_1 \cdots \mathbf{x}_n$.

It is easily seen that there are at most a countable number of channels, since the number of ways in which n particles can be grouped is finite and the number of energies with which any collection of particles can be bound is at most countable. We can thus assume that α takes the values 0, 1, 2, $\cdots$. We make the convention that, as in (62), the $\alpha = 0$ channel refers to n asymptotically free particles. Let us define

$$W(t) = \exp(-iHt), \qquad \Omega_\alpha(t) = W^{-1}(t)U_\alpha(t). \tag{63}$$

Then arguing as in the Introduction, we can see that the transition amplitude from an initial state $U_\alpha(t)\varphi_\alpha$ to a final state $U_\beta(t)\varphi_\beta$ is given by

$$T_{\beta\alpha} = \lim_{\substack{t_1\to-\infty \\ t_2\to+\infty}} (\Omega_\beta(t_2)\varphi_\beta, \Omega_\alpha(t_1)\varphi_\alpha), \tag{64}$$

provided the limit exists. The limit will certainly exist if the operators $\Omega_\alpha(t)$ and $\Omega_\beta(t)$ converge strongly when applied to φ_α and φ_β, respectively. Now it can be proved[12] that, under certain conditions on the potentials V_{ij} occurring in H (it suffices that all of them be square-integrable), the expressions $\Omega_\alpha(t)\varphi_\alpha$ converge strongly as $t \to \pm\infty$ to functions $\Omega_\alpha^\pm\varphi_\alpha$, for any $\alpha = 0, 1, 2 \cdots$ and any φ_α of the form (59). [Note that since the f in (59) is an arbitrary square-integrable function, there are infinitely many such φ_α for each choice of $\psi_{\alpha_1} \cdots \psi_{\alpha_m}$.]

We can now define a "partial S operator" S_α describing the results of scattering initiated in the channel α. In analogy with the operator given in (10) we write

$$S_\alpha = \Omega_\alpha^- \Omega_\alpha^{+*}. \tag{65}$$

S_α maps functions $\Omega_\alpha^+\varphi_\alpha$ onto functions $\Omega_\alpha^-\varphi_\alpha$,

$$S_\alpha \Omega_\alpha^+ \varphi_\alpha = \Omega_\alpha^- \varphi_\alpha, \tag{66}$$

and we can view the transition amplitude $T_{\beta\alpha}$ as the matrix element of S_α between $\Omega_\beta^+\varphi_\beta$ and $\Omega_\alpha^+\varphi_\alpha$. Now it would be convenient to be able to add up the effects of all the partial S operators into a total S operator which would describe scattering initiated in any channel. The way in which to do this has been found by Jauch.[13] We describe his technique briefly, omitting most of the reasoning behind it, and trusting that the interested reader will consult his paper: A total S operator will have to be applicable to any function in L^2. Now S_α of (65) is constructed using the operators $\Omega_\alpha^\pm$ which are defined only on the subspace of L^2 on which $\Omega_\alpha(t)$ converges strongly. [This subspace contains all functions of the form (59).] We can, however, extend the definition of $\Omega_\alpha^\pm$ by setting $\Omega_\alpha^\pm$ equal to zero outside this subspace. Then $\Omega_\alpha^\pm$ as well as $\Omega_\alpha^{\pm*}$ and S_α are defined on all of L^2 and we can write the full S matrix as

$$S = \sum_{\alpha=0}^{\infty} S_\alpha. \tag{67}$$

Jauch proves that the sum in (67) converges strongly and that it reproduces correctly the transition amplitudes $T_{\beta\alpha}$ when taken between the states $\Omega_\beta^+\varphi_\beta$ and $\Omega_\alpha^+\varphi_\alpha$. The matrix given by (67) is unitary in the usual sense if both the sets of states $\Omega_\alpha^- f$ and $\Omega_\alpha^+ f$, $f \in L^2$, $\alpha = 0, 1, 2, \cdots$, span the orthogonal complement of the set of "true bound states" (normalizable eigenfunctions of the full Hamiltonian H).

This completes our discussion of the formalism for dealing with n-body multichannel scattering when only short-range interactions occur.

When Coulomb potentials are involved we can carry through a scattering formalism precisely analogous to the one given above, provided that we change the definition of the channel operators. To carry through the formalism we must change this definition in such a way that the analogs of the operators $\Omega_\alpha(t)$ of (63) converge strongly on the appropriate subspaces. The expression for the new channel operators is rather complicated in the case of a general channel. For this reason we shall write down the new channel operator for the $\alpha = 0$ channel in which all particles are free asymptotically, and indicate what changes must be made in order to obtain convergence for a general channel.

Theorem 3. Let H_c be given by

$$H_c = -\sum_{r=1}^{n} \frac{\Delta_r}{2m_r} + \sum_{1\le r<s\le n} \frac{e_r e_s}{|\mathbf{x}_r - \mathbf{x}_s|} + \sum_{1\le r<s\le n} V_{rs}(\mathbf{x}_r - \mathbf{x}_s) + \sum_{r=1}^{n} \frac{e_r \mathcal{E}}{x_r} + \sum_{s=1}^{n} V_{0s}(\mathbf{x}_r), \tag{68}$$

where the V_{rs} are square-integrable ($0 \le r < s \le n$).

[12] M. N. Hack, Ref. 2; J. Dollard, Ref. 3. (Note: In the latter reference, no proof is given for the case in which a subset of particles is bound near the origin. However, it is trivial to extend the proofs presented to this case.)

[13] J. M. Jauch, Helv. Phys. Acta **31**, 661 (1958). Note: the example of a nonrelativistic channel operator given in this paper is in error. The center-of-mass motion should be separated out of the first equation on p. 665, and the nonrelativistic H_0 in the third equation of this page should be $H_0 = 1/2(m_1 + m_2)\mathbf{P}^2 + E$.

Define the operators

$$D_{rs} = (m_s\nabla_r - m_r\nabla_s)^2, \tag{69}$$

and the distorted channel operator $U_{0c}(t)$ for the $\alpha = 0$ channel,

$$U_{0c}(t) = \exp(-iH_{0c}(t)), \tag{70}$$

with

$$H_{0c}(t) = -\sum_{i=1}^{n}\frac{\Delta_i}{2m_i}t + \epsilon(t) \times\left\{\sum_{1\le r<s\le n}\frac{m_rm_se_re_s}{(-D_{rs})^{\frac{1}{2}}}\log\left[\frac{-2\,|t|\,D_{rs}}{m_rm_s(m_r+m_s)}\right] + \sum_{r=1}^{n}\frac{m_re_r\mathcal{E}}{(-\Delta_r)^{\frac{1}{2}}}\log\left[\frac{-2\,|t|\,\Delta_r}{m_r}\right]\right\}. \tag{71}$$

Define

$$W_c(t) = \exp(-iH_ct). \tag{72}$$

Then $\Omega_{0c}(t) = W_c^{-1}(t)U_{0c}(t)$ converges strongly, as $t \to \pm\infty$, on the entire space of square-integrable functions in n variables, to operators $\Omega_{0c}^{\pm}$.[3]

It should be clear that $H_{0c}(t)$ of (71) is a straightforward generalization of $H_{0c}(t)$ of (29). In (29), $m/(-\Delta)^{\frac{1}{2}}$ is the reciprocal of the velocity operator for the particle of interest. In (71), in the sum over r and s, this is replaced by $m_rm_s/(-D_{rs})^{\frac{1}{2}}$, the reciprocal of the relative velocity operator for particles r and s. The expressions $\log\{-2D_{rs}/[m_rm_s(m_r+m_s)]\}$ are irrelevant in the convergence proof [see Eq. (34)*ff*]. The expression $-2D_{rs}/[m_rm_s(m_r+m_s)]$ is the operator for four times the relative energy of particles r and s, and was chosen to correspond to the factor $-2\Delta/m$ in (29). With this choice, correct results are obtained for the pure Coulomb two-body problem, which reduces to the case of potential scattering. However, since the n-body stationary scattering solutions are not known, it is not certain that this choice produces the n-body analog of (54). Instead, it is possible that an additional phase factor appears in the integral corresponding to the right-hand side of (54).

It should be clear from Theorem 3 that the scattering picture for the $\alpha = 0$ channel is now as follows: at large negative times a wave packet $U_{0c}(t)f_1$ is present, with $f_1 \in L^2$, and at large positive times this becomes a wave packet $U_{0c}(t)f_2$. The transition amplitude for the process is $(\Omega_{0c}^{+}f_2, \Omega_{0c}^{-}f_1)$. As in Sec. 2, we do not have "asymptotic freedom" in the usual quantum-mechanical sense, but we can use the n-body analog of (40) to show that our intuitive criterion of asymptotic freedom is still satisfied. Similar statements hold for the $\alpha \neq 0$ channels, which we now discuss.

We now outline the process for obtaining the distorted channel operators for channels other than $\alpha = 0$. First notice that the distorted channel operator $U_{0c}(t)$ can be written as the product of the usual channel operator $U_0(t)$ of (62) with another operator $U'_{0c}(t)$ which is the exponential of $-i\{H_{0c}(t) + \sum_{i=1}^{n}\Delta_i/2m_i\}$,

$$U_{0c}(t) = U_0(t)U'_{0c}(t). \tag{73}$$

To construct the operator $U'_{0c}(t)$, we used the charges, masses, and coordinates of the n particles under consideration. Now the distorted channel operator $U_{\alpha c}(t)$ for the channel α containing l simple and m composite particles and some particles bound near the origin factors as above:

$$U_{\alpha c}(t) = U_\alpha(t)U'_{\alpha c}(t), \tag{74}$$

with U_α defined by (60), (61); $U'_{\alpha c}(t)$ is obtained as follows: the jth composite particle has a total mass M_j, a total charge $\mathcal{E}_j$, and a center-of-mass coordinate $\mathbf{Y}_j$. We imagine that for each j ($j = 1\cdots m$) the jth composite particle has been replaced by a simple particle of mass M_j, charge $\mathcal{E}_j$, and position coordinate $\mathbf{Y}_j$, and that we are describing the scattering of $l + m$ simple particles each of which is asymptotically free. We replace $\mathcal{E}$, the fixed charge at the origin, by $\mathcal{E} + \mathcal{E}'$, where $\mathcal{E}'$ is the sum of the charges of the particles bound near the origin. Aside from making this replacement, we ignore these particles entirely. We write the appropriate operator $U'_{0c}(t)$ for the scattering of the $l + m$ simple particles. [This operator is obtained using (71) with $\mathcal{E}$ replaced by $\mathcal{E} + \mathcal{E}'$, $n = l + m$.] Then this operator $U'_{0c}(t)$ is the $U'_{\alpha c}(t)$ required in (74). In other words, as far as the distorting operator $U'_{\alpha c}(t)$ is concerned, the jth composite particle behaves like a simple particle of mass M_j, charge $\mathcal{E}_j$, and coordinate $\mathbf{Y}_j$. One can then show that $W^{-1}(t)U_{\alpha c}(t)$ converges on the appropriate subspace. (Note: the proof of this given by the author requires the hypothesis that for each bound-state wavefunction ψ_{α_j} describing a composite particle or a set of particles bound near the origin, and for each coordinate $\mathbf{x}$ of the bound state, there should exist some number $\epsilon > 0$ such that $\int d\mathbf{x}\overline{\psi_{\alpha_j}}\,|\mathbf{x}|^{\epsilon}\,\psi_{\alpha_j} < \infty$. In view of the usual exponential damping of bound states, this does not seem a very restrictive hypothesis.)

As mentioned above, the formalism for n-body scattering can now be carried through in its entirety for problems involving Coulomb potentials. This

does not follow immediately from the multichannel convergence proofs and the work of Jauch, because the operators $U_\alpha(t)$ are not a set of channel operators in his sense. Nonetheless, Jauch's final results remain valid and his program for defining the S operator can be carried out.[3]

ACKNOWLEDGMENT

I wish to thank Professor Arthur Strong Wightman for suggesting to me the topic investigated in this paper, and for his encouragement and the many valuable suggestions which he made in the course of the work.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 5, NUMBER 6 JUNE 1964

Eigenfunctions of the Electron Spin Operators

DON SECREST AND LLOYD MARTIN HOLM*†
Department of Chemistry, University of Illinois, Urbana, Illinois
(Received 11 October 1963)

The problem of finding linear combinations of Slater determinants which are eigenfunctions of both the S^2 and S_z operators is considered. The method of projection operators is used in forming eigenfunctions of S^2. A recursion relation is given for the coefficients in the linear combination. Rules are given for selecting a complete linearly independent set of eigenfunctions and an explicit formula is given for the members of the set.

INTRODUCTION

THE use of projection operators[1] is a practical method for the computation of spin eigenfunctions. This is especially true when there are a large number of singly occupied orbitals. The results of applying a spin projection operator to a Slater determinant is known explicitly in many special cases.[2-5] We shall give a recursion relation for finding the coefficients of the Slater determinants in the projection of a single Slater determinant to the space of eigenfunctions of the S^2 operator for any eigenvalue s. The problem of finding eigenfunctions of the S^2 operator is not difficult. However, finding a complete linearly independent set of eigenfunctions is not so simple. Pratt[6] has solved this problem for the singlet case by a rather complicated method. Löwdin[3] gave a rule for picking a set of Slater determinants whose projections are linearly independent which works for a small number of spins, but this rule breaks down for the many spin case (more than six).

For a set of n electrons in singly occupied orbitals it is well known that we may form

$$N = \binom{n}{\frac{1}{2}n + s} - \binom{n}{\frac{1}{2}n + s + 1} \qquad (1)$$

independent eigenfunctions of the S^2 and S_z operator with eigenvalue $s(s + 1)$ and s_z, respectively. In forming spin eigenfunctions from Slater determinants it is only necessary to consider singly occupied orbitals. We can obtain a complete set of Slater determinants for any particular spacial configuration by permuting the spin parts in all possible ways. We shall abbreviate a Slater determinant by writing only the spin part, the α's and β's, of the singly occupied orbitals of the diagonal element. We may order all these Slater determinants in an alphabetical order. That is, we may put them in the order in which they would appear in a Greek dictionary. A rule for finding N independent eigenfunctions of S^2 with eigenvalue s when there are less than 6 singly occupied orbitals is as follows: The projections of the first N Slater determinants in alphabetical order with $s_z = s$ will be linearly independent.[3] These eigenfunctions will of course have eigenvalue $s_z = s$. If eigenfunctions with a different s_z eigenvalue are desired, they may be obtained from these by repeated application of the

* Participant in the National Science Foundation Undergraduate Science Education Program.
† Present address: Theoretical Chemistry Institute, University of Wisconsin, Madison, Wisconsin.
[1] P. O. Löwdin, Advan. Phys. 5, 1 (1956).
[2] J. K. Percus and A. Rotenberg, J. Math. Phys. 3, 928 (1962).
[3] P. O. Löwdin, Coll. Intern. Centre Natl. Rech. Sci. (Paris) 82, 23 (1958).
[4] P. O. Löwdin, Phys. Rev. 97, 1509 (1955).
[5] James W. Cooley, "Some Computational Methods for the Study of Diatomic Molecules," NYO Rept. No. 9490 (1961).
[6] G. W. Pratt, Jr., Phys. Rev. 92, 278 (1953).

does not follow immediately from the multichannel convergence proofs and the work of Jauch, because the operators $U_\alpha(t)$ are not a set of channel operators in his sense. Nonetheless, Jauch's final results remain valid and his program for defining the S operator can be carried out.[3]

ACKNOWLEDGMENT

I wish to thank Professor Arthur Strong Wightman for suggesting to me the topic investigated in this paper, and for his encouragement and the many valuable suggestions which he made in the course of the work.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 5, NUMBER 6 JUNE 1964

Eigenfunctions of the Electron Spin Operators

DON SECREST AND LLOYD MARTIN HOLM*†
Department of Chemistry, University of Illinois, Urbana, Illinois
(Received 11 October 1963)

The problem of finding linear combinations of Slater determinants which are eigenfunctions of both the S^2 and S_z operators is considered. The method of projection operators is used in forming eigenfunctions of S^2. A recursion relation is given for the coefficients in the linear combination. Rules are given for selecting a complete linearly independent set of eigenfunctions and an explicit formula is given for the members of the set.

INTRODUCTION

THE use of projection operators[1] is a practical method for the computation of spin eigenfunctions. This is especially true when there are a large number of singly occupied orbitals. The results of applying a spin projection operator to a Slater determinant is known explicitly in many special cases.[2-5] We shall give a recursion relation for finding the coefficients of the Slater determinants in the projection of a single Slater determinant to the space of eigenfunctions of the S^2 operator for any eigenvalue s. The problem of finding eigenfunctions of the S^2 operator is not difficult. However, finding a complete linearly independent set of eigenfunctions is not so simple. Pratt[6] has solved this problem for the singlet case by a rather complicated method. Löwdin[3] gave a rule for picking a set of Slater determinants whose projections are linearly independent which works for a small number of spins, but this rule breaks down for the many spin case (more than six).

For a set of n electrons in singly occupied orbitals it is well known that we may form

$$N = \binom{n}{\frac{1}{2}n + s} - \binom{n}{\frac{1}{2}n + s + 1} \qquad (1)$$

independent eigenfunctions of the S^2 and S_z operator with eigenvalue $s(s + 1)$ and s_z, respectively. In forming spin eigenfunctions from Slater determinants it is only necessary to consider singly occupied orbitals. We can obtain a complete set of Slater determinants for any particular spacial configuration by permuting the spin parts in all possible ways. We shall abbreviate a Slater determinant by writing only the spin part, the α's and β's, of the singly occupied orbitals of the diagonal element. We may order all these Slater determinants in an alphabetical order. That is, we may put them in the order in which they would appear in a Greek dictionary. A rule for finding N independent eigenfunctions of S^2 with eigenvalue s when there are less than 6 singly occupied orbitals is as follows: The projections of the first N Slater determinants in alphabetical order with $s_z = s$ will be linearly independent.[3] These eigenfunctions will of course have eigenvalue $s_z = s$. If eigenfunctions with a different s_z eigenvalue are desired, they may be obtained from these by repeated application of the

* Participant in the National Science Foundation Undergraduate Science Education Program.
† Present address: Theoretical Chemistry Institute, University of Wisconsin, Madison, Wisconsin.
[1] P. O. Löwdin, Advan. Phys. 5, 1 (1956).
[2] J. K. Percus and A. Rotenberg, J. Math. Phys. 3, 928 (1962).
[3] P. O. Löwdin, Coll. Intern. Centre Natl. Rech. Sci. (Paris) 82, 23 (1958).
[4] P. O. Löwdin, Phys. Rev. 97, 1509 (1955).
[5] James W. Cooley, "Some Computational Methods for the Study of Diatomic Molecules," NYO Rept. No. 9490 (1961).
[6] G. W. Pratt, Jr., Phys. Rev. 92, 278 (1953).

lowering operator

$$S_- = S_x - iS_y. \tag{2}$$

Here and in the following we use units in which $\hbar = 1$. In this paper we extend this rule to arbitrarily many singly occupied orbitals and give an explicit formula for the eigenfunctions with arbitrary s and s_z eigenvalues.

Spin Projection Operator

Let $(\alpha^{\frac{1}{2}n+s_z} \mid \beta^{\frac{1}{2}n-s_z})$ be any Slater determinant with $\frac{1}{2}n + s_z$ α's and $\frac{1}{2}n - s_z$ β's. A projection operator, which, acting on this Slater determinant, produces an eigenfunction of S^2 with eigenvalue $s(s+1)$, and $S_z \geq 0$ may be written

$$O_s = \prod_{\substack{i=s_z \\ i\neq s}}^{\frac{1}{2}n} \left[\frac{S^2 - i(i+1)}{s(s+1) - i(i+1)}\right] = \prod_{\substack{i=s_z \\ i\neq s}}^{\frac{1}{2}n} \left[1 - \frac{s(s+1) - S^2}{s(s+1) - i(i+1)}\right]. \tag{3}$$

It can be seen easily[7] that each factor of this operator will produce new Slater determinants with one interchange of α's and β's in all possible ways and that all new Slater determinants will have the same coefficients. Since there are $\frac{1}{2}n - s_z$ factors in the operator, the projection will contain, besides the original Slater determinant, all possible Slater determinants with 1, 2, $\cdots$ $\frac{1}{2}n - s_z$ interchanges. Since there are only $\frac{1}{2}n - s_z$ β's in the original Slater determinant, the projection is a linear combination of all possible Slater determinants. The projection may be written

$$O_s(\alpha^{\frac{1}{2}n+s_z} \mid \beta^{\frac{1}{2}n-s_z}) = \sum_{p=0}^{\frac{1}{2}n-s_z} q_p^{n,s}(\alpha^{\frac{1}{2}n+s_z-p}\beta^p \mid \beta^{\frac{1}{2}n-s_z-p}\alpha^p), \tag{4}$$

where we define $(\alpha^{n-p}\beta^p \mid \beta^{m-p}\alpha^p)$ to be the sum of all possible Slater determinants which disagree with the original by p interchanges of α's and β's. For example, if

$$(\alpha^3 \mid \beta^2) = (\alpha\beta\alpha\beta\alpha), \tag{5}$$

then

$$(\alpha^2\beta \mid \beta\alpha) = (\beta\alpha\alpha\beta\alpha) + (\beta\beta\alpha\alpha\alpha) + (\alpha\alpha\beta\beta\alpha) + (\alpha\beta\beta\alpha\alpha) + (\alpha\alpha\alpha\beta\beta) + (\alpha\beta\alpha\alpha\beta). \tag{6}$$

[7] This is obvious if we note that S^2 may be written

$$S^2 = -\tfrac{1}{4}n(n-4) + \tfrac{1}{2}\sum_{ij}{}' P_{ij},$$

where P_{ij} is a permutation operator interchanging spin variables only. See P. O. Löwdin, *Advances in Chemical Physics* (Interscience Publishers, Inc., New York, 1959), Vol. II.

Since Eq. (4) is an eigenfunction of S^2 with eigenvalue $s(s+1)$, the operator $S^2 - s(s+1)$ will annihilate it. This operator may be written[8]

$$S^2 - s(s+1) = S_-S_+ + S_z(S_z+1) - s(s+1). \tag{7}$$

If we allow this operator to act on each term of the linear combination in Eq. (4) and collect terms, we obtain the coefficient of each Slater determinant in the expansion. Let us consider one particular Slater determinant d which differs from the original, $(\alpha^{\frac{1}{2}n+s_z} \mid \beta^{\frac{1}{2}n-s_z})$, by r interchanges of α's and β's. This determinant, d, has a coefficient $q_r^{n,s}$ in the expansion. When the operator, Eq. (7), operates on d we get d back again multiplied by $\frac{1}{2}n - s_z + s_z(s_z+1) - s(s+1)$ along with some other determinants which differ from d by one interchange. There are $(\frac{1}{2}n - s_z - r)r + (\frac{1}{2}n + s_z - r)r$ determinants which differ from d by one interchange and from the original by r interchanges. These also carry coefficient $q_r^{n,s}$. Each of these contributes one determinant d when acted on by the operator, Eq. (7). Similarly there are $(\frac{1}{2}n-s_z-r)(\frac{1}{2}n+s_z-r)$ which differ from d by one interchange and from the original by $r+1$ interchanges, and r^2 which differ from the original by $r-1$ interchanges. These carry coefficients $q_{r+1}^{n,s}$ and $q_{r-1}^{n,s}$, respectively. The sum of these contributions gives the coefficient of d after the operator Eq. (7) has acted. But since this operator annihilates the eigenfunction, this coefficient must be zero and collecting terms we obtain

$$[nr - 2r^2 - s(s+1) + s_z^2 + \tfrac{1}{2}n]q_r^{n,s} + r^2 q_{r-1}^{n,s} + (\tfrac{1}{2}n - s_z - r)(\tfrac{1}{2}n + s_z - r)q_{r+1}^{n,s} = 0. \tag{8}$$

This is a recursion relation by the use of which we may compute all the $q_i^{n,s}$ in terms of $q_0^{n,s}$.

The eigenfunction produced by projection is not normalized even though the original was normalized. Thus, the value of $q_0^{n,s}$ is of little interest and we may arbitrarily set $q_0^{n,s} = 1$. Percus and Rotenberg[2] have given an expression for $q_r^{n,s}$ for the case $s_z = 0$. In the numerical computation of this quantity, however, Eq. (8) is easier to use than their rather complicated sum. Our recursion relation, Eq. (8), reduces to that given by Löwdin[4] for the case $s_z = 0$. It is easy to verify that for $s = s_z$ the well-known solution

$$q_r^{s,n} = (-1)^r \frac{2s+1}{\frac{1}{2}n + s + 1}\binom{\frac{1}{2}n + s}{r}^{-1} \tag{9}$$

satisfies the recursion relation Eq. (8).

[8] H. Eyring, J. Walter and G. E. Kimball, *Quantum Chemistry* (John Wiley & Sons, Inc., New York, 1944).

A LINEARLY INDEPENDENT SET OF SPIN EIGENFUNCTIONS

There are many ways of picking a linearly independent set of eigenfunctions. There is no unique set. Choosing a set is not a trivial problem, however. We give here a set of rules for picking a set of Slater determinants whose projection to an eigenfunction with eigenvalue $s = s_z$ is linearly independent. Then, using the stepping-down operator, Eq. (2), we obtain a set with arbitrary eigenvalue s_z. There are $\frac{1}{2}n - s$ rules, and if we take all possible Slater determinants with $s_z = s$ which obey *all* of the rules, their projections to eigenfunctions with eigenvalue s will form a complete and linearly independent set. The jth rule may be stated.

Rule j. The determinant must not have j β's in the first 2j−1 positions. Here j runs from 1 to $\frac{1}{2}n-s$.[9]

As an example, let us consider the six spin case with $s = 0$. The set obtained by our rules is $(\alpha\alpha\alpha\beta\beta\beta)$, $(\alpha\alpha\beta\alpha\beta\beta)$, $(\alpha\alpha\beta\beta\alpha\beta)$, $(\alpha\beta\alpha\alpha\beta\beta)$, and $(\alpha\beta\alpha\beta\alpha\beta)$. There are exactly five functions, which is precisely the number required for this case by Eq. (1). These rules are very easy to implement in a machine computation. A proof that these rules always give a complete linearly independent set of projections is given in the appendix.

The rule mentioned earlier, which states that the determinants should be arranged in alphabetical order and the first N taken, breaks down in this case. By that rule, the set $(\alpha\alpha\alpha\beta\beta\beta)$, $(\alpha\alpha\beta\alpha\beta\beta)$, $(\alpha\alpha\beta\beta\alpha\beta)$, $(\alpha\alpha\beta\beta\beta\alpha)$, and $(\alpha\beta\alpha\alpha\beta\beta)$ should have a linearly independent projection. Our rules eliminate $(\alpha\alpha\beta\beta\beta\alpha)$ because it breaks rule 3, i.e., 3 β's in the first 5 positions. In fact,

$$O_0(\alpha\alpha\beta\beta\beta\alpha) = -O_0(\alpha\alpha\alpha\beta\beta\beta) - O_0(\alpha\alpha\beta\alpha\beta\beta) - O_0(\alpha\alpha\beta\beta\alpha\beta), \quad (10)$$

and so the set chosen by the alphabet rule is *not* linearly independent.

Once a set of Slater determinants has been chosen in this manner, the projection of each individual determinant is given by substituting Eq. (9) into Eq. (4),

$$O_s(\alpha^{\frac{1}{2}n+s} \mid \beta^{\frac{1}{2}n-s}) = \sum_{p=0}^{\frac{1}{2}n-s} (-1)^p \frac{2s+1}{\frac{1}{2}n+s+1} \binom{\frac{1}{2}n+s}{p}^{-1} \times (\alpha^{\frac{1}{2}n+s-p}\beta^p \mid \beta^{\frac{1}{2}n-s-p}\alpha^p). \quad (11)$$

These are eigenfunctions of both S^2 and S_z with eigenvalues $s = s_z$.

[9] Notice that each allowed configuration corresponds to a path in the branching diagram sometimes used to describe spin functions. See, for example, E. M. Corson, *Perturbation Methods in the Quantum Mechanics of n-Electron Systems* (Hafner Publishing Co., Inc., New York, 1951).

To obtain eigenfunctions with other s_z eigenvalues, we apply the operator S_- $(s - s_z)$ times. It is easy to see that all determinants which have p α's where the original had β's will have a common coefficient after application of the lowering operator. Thus, after k applications, we may write

$$S_-^k O_s(\alpha^{\frac{1}{2}n+s} \mid \beta^{\frac{1}{2}n-s}) = \sum_{p=0}^{\frac{1}{2}n-s} Q_p^{n,s,k}(\alpha^{\frac{1}{2}n+s-p-k}\beta^{p+k} \mid \beta^{\frac{1}{2}n-s-p}\alpha^p), \quad (12)$$

where the $Q_p^{n,s,k}$ are constants to be determined. Application of the operator S_- one more time gives determinants in the group $(\alpha^{\frac{1}{2}n+s-p-k-1}\beta^{p+k+1} \mid \beta^{\frac{1}{2}n-s-p}\alpha^p)$. The coefficient of each member in this sum may be computed by noting that it is produced $p + k + 1$ times by the set $(\alpha^{\frac{1}{2}n+s-p-k}\beta^{p+k} \mid \beta^{\frac{1}{2}n-s-p}\alpha^p)$, and $\frac{1}{2}n - s - p$ times by the set

$$(\alpha^{\frac{1}{2}n+s-p-k-1}\beta^{p+k+1} \mid \beta^{\frac{1}{2}n-s-p-1}\alpha^{p+1}).$$

Thus

$$Q_p^{n,s,k+1} = (p + k + 1)Q_p^{n,s,k} + (\tfrac{1}{2}n - s - p)Q_{p+1}^{n,s,k}. \quad (13)$$

This is a partial difference equation with boundary condition

$$Q_p^{n,s,0} = (-1)^p \frac{2s+1}{\frac{1}{2}n+s+1} \binom{\frac{1}{2}n+s}{p}^{-1}. \quad (14)$$

It is easy to verify that

$$Q_p^{n,s,k} = \frac{2s+1}{\frac{1}{2}n+s+1} \frac{(2s)^{(k)}(-1)^p}{\binom{\frac{1}{2}n+s}{p+k}} \quad (15)$$

is a solution to this equation. Here we use the notation

$$n^{(r)} = \prod_{i=1}^{r} (n + 1 - i). \quad (16)$$

Thus

$$S_-^{s-s_z} O_s(\alpha^{\frac{1}{2}n+s} \mid \beta^{\frac{1}{2}n-s}) = \frac{2s+1}{\frac{1}{2}n+s+1} \times \sum_{p=0}^{\frac{1}{2}n-s} \frac{(2s)^{(s-s_z)}(-1)^p}{\binom{\frac{1}{2}n+s}{s-s_z+p}} (\alpha^{\frac{1}{2}n+s_z-p}\beta^{p+s-s_z} \mid \beta^{\frac{1}{2}n-s-p}\alpha^p). \quad (17)$$

With this formula and the rules given above, we can immediately write down a complete set of spin eigenfunctions with arbitrary eigenvalues s and s_z.

APPENDIX

Our proof that the rules give a complete linearly independent set consists of two parts. First we prove the rules give exactly the right number of functions, and next, that any other function may be expanded in terms of the set obeying the rules.

We prove that the number is correct by induction. We show that it is right in the two-electron case. Then we show that, if the rules give the correct number of functions in the n electron case, they do for $n + 1$ electrons. The correct number[10] of functions N is given by Eq. (1). For the two-electron case with $s = 0$,

$$N = \binom{2}{1} - \binom{2}{2} = 1 \tag{18}$$

eigenfunction, and by the rules, $O_0(\alpha\beta)$ is the only function. If the rules give the correct result in the n-electron case, we get

$$N = \binom{n}{\frac{1}{2}n + s} - \binom{n}{\frac{1}{2}n + s + 1} \tag{19}$$

functions which obey the rules. In the $n + 1$-spin case for $\frac{1}{2}n > s > 0$, we obtain all functions which obey the rules by adding an α to the right of all n spin functions which obey the rules with $s_z = s - \frac{1}{2}$ and β to all configurations with $s_z = s + \frac{1}{2}$. This number is

$$N' = \binom{n}{\frac{1}{2}n + s - \frac{1}{2}} - \binom{n}{\frac{1}{2}n + s + \frac{1}{2}} + \binom{n}{\frac{1}{2}n + s + \frac{1}{2}} - \binom{n}{\frac{1}{2}n + s + \frac{3}{2}}. \tag{20}$$

Using the well-known recursion relation for binomial coefficients

$$\binom{n+1}{s+1} = \binom{n}{s} + \binom{n}{s+1}, \tag{21}$$

we see that

$$N' = \binom{n+1}{\frac{1}{2}(n+1) + s} - \binom{n+1}{\frac{1}{2}(n+1) + s + 1}. \tag{22}$$

When $s = \frac{1}{2}n$ there is only one function, all α's, so there is no problem. This leaves only the case $s = 0$ to test. If $s = 0$, there are the same number of α's and β's, and the rule which states that there must not be $\frac{1}{2}n$ β's in the first $n - 1$ positions implies that there must be no α in the last position. Thus we may obtain all determinants which obey the rules for the $n + 1$-electron case and $s = 0$ from the n-electron case by adding a β to the right of the set for $s = \frac{1}{2}$, and we obtain

$$N' = \binom{n}{\frac{1}{2}n + \frac{1}{2}} - \binom{n}{\frac{1}{2}n + \frac{1}{2} + 1}. \tag{23}$$

It is easy to verify that

$$\binom{n}{\frac{1}{2}n + \frac{1}{2}} - \binom{n}{\frac{1}{2}n + \frac{3}{2}} = \binom{n+1}{\frac{1}{2}(n+1)} - \binom{n+1}{\frac{1}{2}(n+1) + 1}. \tag{24}$$

This proves that our rules give the correct number of functions.

If we now show that every projection can be written as a sum of members of the set, we have proved completeness and, since the set has been shown to contain no more than the proper number, it must also be linearly independent.

Let us take any Slater determinant δ which breaks one or more of our rules. Let the ith rule be the one of lowest cardinality which it breaks. The determinant has i β's in the first $2i - 1$ positions. We choose the set of all Slater determinants with the same s_z which have their first $i - 1$ α's the same as δ and all $\frac{1}{2}n - s_z - i$ β's not in the first $2i - 1$ positions the same as δ. No member of this set breaks rule i, nor any rule of lower cardinality, though some may break higher rules. We shall show that the projection of δ is minus the sum of the projections of the elements in this set. In a similar manner, all elements of the set which break rules higher than i may be expanded in terms of elements which may break still higher rules until no more rules are broken.

Let γ be any particular Slater determinant. We shall show that when the sum of the projection of every member in the set including δ is formed the coefficient of the determinant γ is zero. Let us assume that γ disagrees with δ, and thus with every member of the set on ℓ of the first $i - 1$ α's and k of the β's which occur in positions *beyond* the first $2i - 1$. Hence γ must have at least k interchanges with respect to every member of the set. It is easy to see that there are

$$\binom{i + k - \ell}{i - j}\binom{\frac{1}{2}n + s - i + 1 + \ell - k}{j}$$

configurations in the set which disagree with γ on exactly $k + j$ β's, and thus have exactly $k + j$ interchanges with respect to it. So when all the coefficients of γ are added, we obtain the coefficient

[10] For a proof of this statement, see E. M. Corson, footnote 9, p. 190.

$$C_\gamma = \sum_{j=0}^{i} \binom{i+k-\ell}{i-j} \times \binom{\frac{1}{2}n+s-i+1+\ell-k}{j} q^{n,s}_{i+k}. \quad (25)$$

Using the value of q_{i+k} given by Eq. (9), we obtain

$$C_\gamma = \frac{2s+1}{\frac{1}{2}n+s+1} \cdot \frac{(i+k-\ell)!\,(\frac{1}{2}n+s-i+1+\ell-k)!}{(\frac{1}{2}n+s)!\,i!} \times \sum_{j=0}^{i} \binom{i}{j} (j+k)^{(\ell)} (\tfrac{1}{2}n+s-k-j)^{(i-\ell-1)} (-1)^j. \quad (26)$$

But the factor $(j+k)^{(\ell)}(\frac{1}{2}n+s-k-j)^{(i-\ell-1)}$ is a polynomial in j of order $i-1$. It is easy to show that

$$\sum_{j=0}^{i} \binom{i}{j} (-1)^j P_n(j) = 0 \quad (27)$$

whenever $P_n(j)$ is a polynomial in j of order $n < i$; thus

$$C_\gamma = 0 \quad (28)$$

for all n, s, $k < \frac{1}{2}n - s - i$ and $\ell < i - 1$. But this includes all Slater determinants, and we have proved what we set out to prove.

Perturbation Theory for Strong Repulsive Potentials*

STEVEN WEINBERG†
Department of Physics, University of California, Berkeley, California
(Received 19 November 1963)

A conformal mapping of the coupling-constant plane is used to rearrange the Born series. The new series is guaranteed to converge for any decent repulsive potential. The first few terms do well in actual calculations of the scattering length.

I. REARRANGEMENT OF THE BORN SERIES

THIS article is intended as a complement to our current series[1-3] on the quasiparticle method. It has been shown there that nonrelativistic scattering problems may always be solved by perturbation theory, provided that we first introduce an appropriate set of fictitious elementary particles. For attractive potentials, these "quasiparticles" must be chosen as replacements for the real bound states and resonances that prevent the convergence of the ordinary Born series.[4] But for repulsive potentials the quasiparticle method is much less appealing, because the quasiparticles must correspond to unreal composite particles which would exist for an interaction of opposite sign. Our purpose here is to describe an alternative method of using perturbation theory to solve scattering problems, which is particularly well suited to the case of strong repulsive potentials.

To explain our approach in very general terms, let us consider any scattering amplitude, which can be regarded as an analytic function $A(\lambda)$ of a complex variable coupling parameter, λ, with singularities at points λ_ν. For example:

(1) The partial-wave amplitude[2] $A_\ell(E;\lambda)$ for potential $\lambda V(r)$ is meromorphic in λ, provided that

$$\int^\infty V^2(r)\,dr < \infty \quad \text{and} \quad \int_0 V^2(r)r^2\,dr < \infty. \tag{1.1}$$

The poles $\lambda_\nu(E,\ell)$ are at the reciprocals of the eigenvalues of the scattering kernel. That is, they are the λ values for which there exists a solution of Schrödinger's equation

$$\left[\frac{-1}{2m}\frac{d^2}{dr^2} + \frac{\ell(\ell+1)}{2mr^2} + \lambda_\nu(E,\ell)V(r) - E\right] \times \psi_\nu(r;E,\ell) = 0, \tag{1.2}$$

with boundary conditions

$$\psi_\nu(r;E,\ell) \propto r^{\ell+1} \qquad (r\to 0) \tag{1.3}$$

$$\propto e^{ikr}[r\to\infty;\, k = +(2mE)^{\frac{1}{2}}]. \tag{1.4}$$

[A power of r may be allowed in (1.4) if $V(r)$ is of long range.] All $\lambda_\nu(E,\ell)$ are complex, becoming real only for $E = 0$. They form a denumerable set whose only limit point is at $\lambda = \infty$.

(2) The three-dimensional scattering amplitude[2] $A(E,\theta;\lambda)$ for potential $\lambda V(r)$ is also meromorphic in λ, under a slightly stronger restriction on V:

$$\int |V(r)|^2\,d^3\mathbf{r} < \infty. \tag{1.5}$$

The poles consist of the $\lambda_\nu(E,\ell)$, for all ℓ.

(3) In multiparticle scattering problems[3] the various reaction amplitudes are not meromorphic in the usual coupling constant, since there appear cuts as well as poles. But they are meromorphic in the magnitudes of the "irreducible" kernels, provided that the potentials satisfy (1.5).

The physical scattering amplitude $A(g)$ is the value of $A(\lambda)$ at the actual couping constant g. If there are no singularities λ_ν in the circle $|\lambda| \le |g|$, then $A(g)$ can be calculated by a Taylor series expansion in g. But even if there are singularities in this circle, it is still always possible to find some other connected region $D(g)$ which contains the origin $\lambda = 0$ as an interior point, which contains the actual coupling constant $\lambda = g$ on its boundary, but which does not contain any of the singularities λ_ν. [See, e.g., Fig. 1.]

It follows from Riemann's theorem[5] that there exists a one-to-one conformal mapping of the unit circle onto the region $D(g)$, such that the center

* This research was supported in part by the Air Force Office of Scientific Research, Grant No. AF-AFOSR-232-63.
† Alfred P. Sloan Foundation Fellow.
[1] S. Weinberg, Phys. Rev. **130**, 776 (1963).
[2] S. Weinberg, Phys. Rev. **131**, 440 (1963). See, also, M. Scadron, S. Weinberg, and J. Wright, to be published in Phys. Rev.
[3] S. Weinberg, Phys. Rev. **133**, B232 (1964).
[4] The quasiparticle method is tested in practical calculations with attractive potentials by M. Scadron and S. Weinberg, Phys. Rev. **133**, B1589 (1964).
[5] E. C. Titchmarsh, *The Theory of Functions* (Oxford University Press, London, 1939), 2nd ed., p. 207.

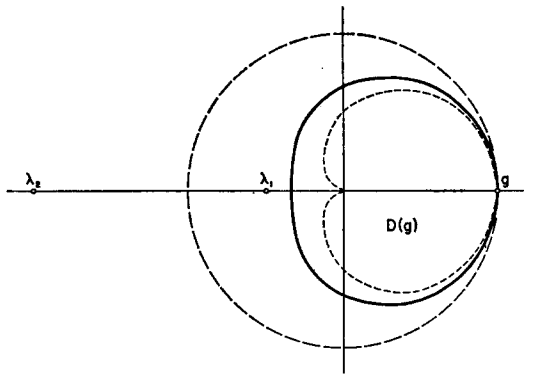

FIG. 1. The complex λ plane. The solid curve is the limaçon of Pascal (2.15), and encloses the region $D(g)$ corresponding to $|z| \leqq 1$. The outer dashed curve is the circle with radius equal to the actual coupling constant g. Observe that the pole λ_1 lies inside this circle but outside $D(g)$, so in this case the ordinary Born series diverges but the rearranged Born series (2.16) is convergent. (We have drawn these curves with $g_0 = 1$, $g = 2$, and have placed the poles as appropriate for zero-energy scattering by the Hulthén potential. See the last line of Table I.) The inner dashed curve is a cardiod (the limiting limaçon for $g_0 = 0$) surrounding the smallest region $D(g)$ that can be obtained with the mapping (2.10).

maps into $\lambda = 0$, and the point at unity maps into $\lambda = g$. That is, there exists a function $\Lambda(z; g)$ with the properties

(i) $\Lambda(z; g)$ is analytic and single-valued for $|z| \leq 1$, taking each value $\lambda \in D(g)$ just once;

(ii) $\Lambda(0; g) = 0$; (1.6)

(iii) $\Lambda(1; g) = g$; (1.7)

(iv) $\Lambda(z; g)$ does not become equal to any of the singular values λ_r for any z with $|z| \leq 1$.

[In fact, the function $\Lambda(z; g)$ is uniquely determined by the region $D(g)$.]

These properties of Λ ensure that the scattering amplitude can be regarded as an analytic function $A(\Lambda(z; g))$ of a modified coupling parameter z in the unit circle $|z| \leq 1$. It can therefore be calculated by the Taylor series expansion in z:

$$A(\Lambda(z; g)) = \sum_{n=0}^{\infty} z^n \mathfrak{A}_n(g), \tag{1.8}$$

where

$$\begin{aligned} \mathfrak{A}_0(g) &= A_0, \qquad \mathfrak{A}_1(g) = A_1\Lambda_1(g), \\ \mathfrak{A}_2(g) &= A_2\Lambda_1(g)^2 + A_1\Lambda_2(g), \\ \mathfrak{A}_3(g) &= A_3\Lambda_1(g)^3 + 2A_2\Lambda_1(g)\Lambda_2(g) + A_1\Lambda_3(g), \end{aligned} \tag{1.9}$$

and so on, with

$$A_n \equiv \frac{1}{n!}\left[\frac{d^n}{d\lambda^n} A(\lambda)\right]_{\lambda=0}, \tag{1.10}$$

$$\Lambda_n(g) \equiv \frac{1}{n!}\left[\frac{d^n}{dz^n} \Lambda(z; g)\right]_{z=0}. \tag{1.11}$$

The series (1.8) converges absolutely for $|z| \leq 1$. In particular, at the physical point $z = 1$, it give for all g

$$A(g) = \sum_{n=0}^{\infty} \mathfrak{A}_n(g). \tag{1.12}$$

This is of course just a rearrangement[6] of the ordinary Born series $A_0 + A_1 + A_2 + \cdots$.

The calculation of $A(g)$ from (1.12) and (1.9) is perfectly straightforward, once we have chosen a function $\Lambda(z; g)$ with the properties (i)–(iv). This is not always a trivial task, since we don't usually know the precise locations λ_r of the singularities of $A(\lambda)$. But for purely repulsive potentials these singularities generally lie so far from the physical value $\lambda = g$ that it is not hard to make sure that none of them lie in $D(g)$.

Some progress can even be made towards optimizing the choice of $\Lambda(z; g)$. If the singularity of $A(\Lambda(z; g))$ closest to $z = 0$ is at a distance $\rho > 1$ then the radius of convergence of (1.8) will be ρ, and we may guess that (1.12) will converge roughly as fast as

$$1 + \rho^{-1} + \rho^{-2} + \cdots. \tag{1.13}$$

Hence we want ρ as large as possible. It seems intuitively reasonable that in the optimum case all singularities of $A(\Lambda(z; g))$ will be *equally* far from $z = 0$, lying on some circle $|z| = \rho$. We will adopt this as a guide in our choice of $\Lambda(z; g)$.

This method will be demonstrated in Sec. II by using it to rearrange the Born series for the scattering length in the case of an arbitrary strong repulsive short-range potential. This rearranged Born series is calculated up to third order for the Hulthén potential in Sec. III, and converges quite rapidly to the correct answer.

II. SCATTERING LENGTHS

We consider here an arbitrarily strong potential $gV(r)$ which is purely repulsive,

$$g > 0; \qquad V(r) \geq 0, \tag{2.1}$$

and is of finite range, in the sense that

[6] The idea of rearranging a perturbation series is not new. See, e.g., J. Blair, N. C. Metropolis, J. von Neumann, A. H. Taub, and M. Tsingou, Math. Tables and other Aids to Computation **13**, 166 (1956), and M. Rotenberg, Ann. of Phys. (N. Y.) **21**, 579 (1963). We hope that the analytic approach presented here may help to illuminate this general idea. It is interesting that the rearrangement defined by Rotenberg's Eq. (18) is almost the same as generated here by taking Λ as the Möbius transformation (3.17), differing in that we would get Rotenberg's series if we multiplied the amplitude by a simple function of z before expanding in z.

$$\int_0^\infty rV(r)e^{\alpha r}\,dr < \infty \quad (\text{some } \alpha > 0), \tag{2.2}$$

$$\int_0^\infty r^2V(r)^2\,dr < \infty. \tag{2.3}$$

The scattering length in the ℓth partial wave is defined by the zero-energy behavior of the phase shift,

$$\delta_\ell(k; g) \to -a_\ell(g)k^{2\ell+1}/(2\ell+1)!!^2. \tag{2.4}$$

The label ℓ will be dropped henceforth.

It follows[7] from (2.2) and (2.3) that $a(g)$ may be extended to a meromorphic function $a(\lambda)$ of a complex coupling parameter λ. The poles λ_ν are given by the reciprocal eigenvalues of the zero-energy scattering kernel; in other words, they are the λ values for which there just barely exist ℓ-wave bound states at zero energy. These poles all lie on the negative real axis with

$$0 > \lambda_1 \geq \lambda_2 \geq \cdots \geq -\infty. \tag{2.5}$$

The λ_ν must not vanish, and must tend to $-\infty$ faster than $-\nu$, because they obey a sum rule[8]

$$\sum_\nu \lambda_\nu^{-1} = \frac{-2m}{2\ell+1}\int_0^\infty rV(r)\,dr. \tag{2.6}$$

For $|\lambda| < |\lambda_1|$, the scattering length can be calculated by the ordinary Born series

$$a(\lambda) = a_1\lambda + a_2\lambda^2 + a_3\lambda^3 + \cdots, \tag{2.7}$$

where

$$a_1 = 2m\int_0^\infty dr\, r^{2\ell+2}V(r), \tag{2.8}$$

$$a_2 = -\frac{8m^2}{2\ell+1}\int_0^\infty dr\int_0^\infty dr'\, r^{\ell+2}r'^{\ell+2} \times V(r)V(r')r_<^\ell r_>^{-\ell-1}, \tag{2.9}$$

and so on. But (2.7) is useless for calculation of the actual scattering length $a(g)$ if $g \geq |\lambda_1|$.

We will rearrange the Born series by the method described in Sec. I, taking the mapping Λ as

$$\Lambda(z; g) = \frac{4z\rho(g)}{[\rho(g) - z]^2}\,g_0, \tag{2.10}$$

where g_0 is a parameter at our disposal, and $\rho(g)$ is determined by the condition that $\Lambda(1; g) = g$, which gives

$$\rho(g) = \frac{[1 + g/g_0]^{\frac{1}{2}} + 1}{[1 + g/g_0]^{\frac{1}{2}} - 1}. \tag{2.11}$$

[7] The properties of $a(\lambda)$ quoted here are derived in Ref. 2.
[8] Reference 2, Eq. (96). The poles λ_ν are the reciprocals of the eigenvalue η_ν.

The singularities of $a(\Lambda(z; g))$ are at

$$z_\nu = \rho(g)\exp(i\theta_\nu), \quad \sin^2(\tfrac{1}{2}\theta_\nu) = -g_0/\lambda_\nu. \tag{2.12}$$

In addition there is an essential singularity at $z_\infty = \rho(g)$. In order that all z_ν lie equally far from $z = 0$ (as suggested in Sec. I) we want all θ_ν real, so that g_0 is to be chosen to be real and to satisfy

$$0 < g_0 \leq |\lambda_1|. \tag{2.13}$$

It is easy to check that this $\Lambda(z; g)$ has the properties (i)–(iv) listed in Sec. I. In particular, (2.13) and (2.11) give

$$\rho(g) > 1 \tag{2.14}$$

for all g, so that the singularities z_ν always lie outside the unit circle $|z| \leq 1$. This can also be seen from the fact that the region $D(g)$ in the λ plane corresponding to $|z| \leq 1$ has as its boundary the limaçon of Pascal:

$$\lambda(\theta) = \frac{ge^{i\theta}}{[\rho(g)+1]^2}[1 + 2\rho(g)\cos\theta + \rho(g)^2] \quad (0 \leq \theta \leq 2\pi), \tag{2.15}$$

and we can see in Fig. 1 that all λ_ν lie outside $D(g)$.

The rearranged Born series is now

$$a(g) = \sum_{n=1}^\infty a_n(g), \tag{2.16}$$

where

$$a_1(g) = \left(\frac{4g_0}{\rho(g)}\right)a_1,$$

$$a_2(g) = \left(\frac{4g_0}{\rho(g)}\right)^2[a_2 + (a_1/2g_0)], \tag{2.17}$$

$$a_3(g) = \left(\frac{4g_0}{\rho(g)}\right)^3[a_3 + (a_2/g_0) + (3a_1/16g_0^2)],$$

and in general

$$a_n(g) = \frac{1}{n!}\left[\frac{d^n}{dz^n}\sum_{m=1}^\infty a_m\Lambda^m(z; g)\right]_{z=0} = \rho(g)^{-n}\sum_{m=1}^n \begin{bmatrix} m+n-1 \\ 2m-1 \end{bmatrix}(4g_0)^m a_m. \tag{2.18}$$

The series (2.16) is guaranteed to converge for any $g > 0$.

In order to maximize the rate of convergence of (2.16), we must make $\rho(g)$ as large as possible. The optimum choice of the parameter g_0 subject to the constraint (2.13) is therefore

$$g_0 = |\lambda_1|, \tag{2.19}$$

which gives

$$\rho(g) = \frac{[1 + g/|\lambda_1|]^{\frac{1}{2}} + 1}{[1 + g/|\lambda_1|]^{\frac{1}{2}} - 1}. \tag{2.20}$$

As remarked in Sec. I, we expect the series (2.16) to converge more or less like $1 + \rho(g)^{-1} + \rho(g)^{-2} + \cdots$. In the weak coupling case $g \ll |\lambda_1|$, when the ordinary Born series already converges, we have

$$\rho(g) \cong 4\ |\lambda_1|/g, \tag{2.21}$$

so we may expect (2.16) to converge about as fast as the ordinary Born series for a potential *one-quarter* as strong. In the strong coupling case $g \gg |\lambda_1|$, in which the ordinary Born series diverges, we have

$$\rho(g) \cong 1 + 2(|\lambda_1|/g)^{\frac{1}{2}}, \tag{2.22}$$

so the series starts converging after about $\frac{1}{2}(g/|\lambda_1|)^{\frac{1}{2}}$ terms.

It only remains to describe how we are to guess at the value of λ_1, so that we can choose a $g_0 > 0$ which we are certain is less than $|\lambda_1|$ but as close to it as possible. It is very fortunate that (2.6) provides an inequality

$$|\lambda_1| > \left[\frac{2m}{2\ell + 1}\int_0^\infty rV(r)\,dr\right]^{-1}. \tag{2.23}$$

Hence an acceptable strategy is to take

$$g_0 = \left[\frac{2m}{2\ell + 1}\int_0^\infty rV(r)\,dr\right]^{-1}. \tag{2.24}$$

This will not be very far below the optimum value $|\lambda_1|$ if the sum rule (2.6) is largely saturated by the $\nu = 1$ term, as is usually the case. There are other sum rules[9] which provide a larger lower bound on $|\lambda_1|$, but (2.23) is good enough.

If the potential $V(r)$ is generally positive but becomes slightly negative for some r, then the poles λ_ν will lie at two infinite sets of points $\lambda_\nu^{(-)} < 0$ and $\lambda_\nu^{(+)} > 0$, with

$$|\lambda_\nu^{(+)}| \gg |\lambda_\nu^{(-)}|. \tag{2.25}$$

The positive poles $\lambda_\nu^{(+)}$ will certainly then lie far outside the region $D(g)$ bounded by the limaçon (2.15), and the series (2.16) will converge with almost the same speed. As the attractive part of $V(r)$ is deepened, the poles at $\lambda_\nu^{(+)}$ will move in towards the origin, and (2.16) will break down when $\lambda_1^{(+)}$ reaches g. For such potentials with strong attractive parts it is still possible to find some other mapping Λ which makes (2.16) converge, but the search for Λ requires accurate knowledge of the $\lambda_\nu^{(+)}$, and the series will probably not converge very fast. In this attractive case the quasiparticle method seems more appropriate.

In scattering problems at finite energy, the poles λ_ν are complex. It is still possible to use the rearranged Born series (2.16) for some value of g_0, provided that the limaçon (2.15) can be made small enough to exclude all the λ_ν. The smallest limaçon is obtained for $g_0 = 0$, $\rho(g) = 1$, in which case (2.15) reduces to the cardiod (the inner dashed curve in Fig. 1),

$$\lambda(\theta) = \tfrac{1}{2}ge^{i\theta}[1 + \cos\theta]. \tag{2.26}$$

Hence (2.16) can always be used for some g_0 if all λ_ν lie outside this curve. At very low energy the λ_ν for repulsive potentials lie close to the negative real axis, and hence outside (2.26). At very high energy the λ_ν move to infinity, and hence certainly lie outside (2.26). At intermediate energies for strong repulsive potentials, it is possible that some λ_ν enter the cardiod (2.26); in this case some other mapping, like (3.17), must be employed.

III. AN EXAMPLE

We will test the rate of convergence of the rearranged Born series (2.16) by comparison with the known results[10] for s-wave scattering by a Hulthén potential

$$gV(r) = (g/2ma^2)[\exp(r/a) - 1]^{-1}. \tag{3.1}$$

The exact $\ell = 0$ scattering length here is

$$a(g) = 2ga\sum_{\nu=1}^{\infty}\frac{1}{\nu(\nu^2 + g)}. \tag{3.2}$$

Hence the nth-order term in the Born series (2.7) is given by

$$a_n = 2a(-)^{n+1}\zeta(2n + 1), \tag{3.3}$$

where ζ is the Riemann zeta function. In particular,

$$a_1 = 2a\zeta(3) = 2.4041a, \tag{3.4}$$

$$a_2 = -2a\zeta(5) = -2.0738a, \tag{3.5}$$

$$a_3 = 2a\zeta(7) = 2.0167a. \tag{3.6}$$

Of course in a calculation with a less convenient potential we would not know the general formula (3.3), and we would have to calculate $a_1, a_2, \cdots$ by actually doing the integrals in (2.8), (2.9), etc.

The poles of $a(\lambda)$ are at

$$\lambda_\nu = -\nu^2, \tag{3.7}$$

[9] Reference 2, Eq. (79).

[10] R. Jost and A. Pais, Phys. Rev. 82, 840 (1951).

so the optimum choice of the parameter g_0 would be

$$g_0 = |\lambda_1| = 1. \tag{3.8}$$

This may be compared with the suggested choice (2.24), which would give

$$g_0 = 1/\zeta(2) = 0.61. \tag{3.9}$$

Our rearranged series will of course converge to the correct answer for either g_0, but it will be more convenient to do our numerical calculations with $g_0 = 1$. If we didn't know the pole positions (3.7), we could use (3.9) or an estimate based on better sum rules,[9] with little change in final results.

The first, second, and third rearranged Born approximations (RBA) suggested by (2.16) are

$$(1\text{ RBA}) \quad a(g) \cong a_1(g), \tag{3.10}$$

$$(2\text{ RBA}) \quad a(g) \cong a_1(g) + a_2(g), \tag{3.11}$$

$$(3\text{ RBA}) \quad a(g) \cong a_1(g) + a_2(g) + a_3(g), \tag{3.12}$$

with

$$a_1(g) = a_1[4/\rho(g)] = 2.4041a[4/\rho(g)] \tag{3.13}$$

$$a_2(g) = (a_2 + \tfrac{1}{2}a_1)[4/\rho(g)]^2 = -0.8717a[4/\rho(g)]^2, \tag{3.14}$$

$$a_3(g) = (a_3 + a_2 + \tfrac{3}{16}a_1)[4/\rho(g)]^3 = 0.3947a[4/\rho(g)]^3, \tag{3.15}$$

and

$$\rho(g) = [(1 + g)^{\frac{1}{2}} - 1]/[(1 + g)^{\frac{1}{2}} - 1]. \tag{3.16}$$

The scattering lengths are presented in Table I for the ordinary and rearranged first, second, and third Born approximations, taking $g = 0.1$, 0.5, 1, and 2. At $g = 0.1$ all approximations give good results, but the rearranged Born approximations are noticeably better than the ordinary ones. At $g = 0.5$ the ordinary Born series is still converging, but rather slowly, while the 2 RBA and 3 RBA are still excellent. At $g = 1.0$ and $g = 2.0$ the ordinary Born series is completely useless (recall that $\lambda_1 = -1$), while the rearranged series is still converging rapidly enough to give very good results in third order.

TABLE I. Scattering lengths for s-wave scattering by the Hulthén potential $V(r) = (g/2ma^2)[\exp(r/a) - 1]^{-1}$ in units of the range a. The different columns list values obtained by the ordinary first, second, and third Born approximations [1BA, 2BA, 3BA], the corresponding rearranged Born approximations [1RBA, 2RBA, 3RBA; see Eqs. (3.10)–(3.16)] and the exact values (EX) calculated from Eq. (3.2). These results are discussed at the end of Sec. III.

g	1BA	2BA	3BA	1RBA	2RBA	3RBA	EX
0.1	0.2404	0.2197	0.2218	0.2289	0.2210	0.2214	0.2215
0.5	1.202	0.684	0.936	0.971	0.829	0.855	0.852
1.0	2.40	0.33	2.34	1.65	1.24	1.36	1.34
2.0	4.81	−3.49	12.66	2.58	1.58	1.97	1.94

We have also calculated $a(g)$ in the 1 RBA and 2 RBA, using instead of (2.10) the Möbius transformation

$$\Lambda(z; g) = gg_0z/[g_0 + g(1 - z)] \tag{3.17}$$

which satisfies conditions (i)–(iv) of Sec. I for all g if

$$0 < g_0 \leq 2\,|\lambda_1|. \tag{3.18}$$

The optimum choice of g_0 here is

$$g_0 = 2\,|\lambda_1|, \tag{3.19}$$

which puts the smallest singularity of $a(\Lambda(z; g))$ at

$$z_\infty = z_1 = 1 + 2\,|\lambda_1|/g. \tag{3.20}$$

This is closer than the singularities (2.12), and indeed it turns out that the 1 RBA and 2 RBA derived from (3.17) are somewhat worse than the corresponding approximations based on (2.10). The Möbius transformation (3.17) maps the unit circle $|z| \leq 1$ into a region $D(g)$ given by

$$\left|\lambda - \frac{g^2}{g_0 + 2g}\right| \leq \frac{g(g_0 + g)}{g_0 + 2g}. \tag{3.21}$$

This circle can always be made small enough to exclude any real *or* complex singularities λ, lying in the left half-plane.

Quantum Kinetic Equations for Electrons in a Periodic System*

AMIRAM RON†
University of California, San Diego, La Jolla, California
(Received 1 November 1963)

The approach to equilibrium of a system of electrons in a periodic system is studied by deriving a kinetic equation in the self-consistent field approximation. The derivation based on the introduction of an hierarchy of equations for the s-body density matrices and employing a truncation scheme, valid in the self-consistent field approximation. The irreversibility is introduced by the adiabatic hypothesis of Bogoliubov. The kinetic equation takes proper account of the collective aspects of the electron system and includes the dynamic shielding.

I. INTRODUCTION

In recent years a number of investigations have been devoted to the problem of the approach to equilibrium and the description of nonequilibrium phenomena of Coulombic many-body systems. The case of the classical plasma was studied thoroughly by several authors[1,2] and a kinetic equation, valid in the limit that the number of electrons in the Debye sphere is large, was derived. This work was later extended to include quantum systems[3] and phonons.[4,5] In the quantum electron system the electron–electron interaction is treated in the so-called random phase approximation (RPA)[6] which is apparently equivalent to the self-consistent field approach (SCF)[7] and the quantum kinetic equations are valid as far as the RPA or the SCF approaches are valid. All of this work was concerned with spatially homogeneous systems only.

In the present paper we attempt to include the effect of a periodic system, e.g., a crystal, on the kinetic equation for the electron system. We consider a system of *interacting* electrons in a perfect periodic lattice under the simplifying assumption that local field correction and hence electron–umklapp processes may be neglected. Starting from the Hamiltonian of the system we derive an *hierarchy* of equations for the quantized density operators by a method developed in Ref. (5), where instead of employing the Liouville equation we use the Heisenberg equations of motion for the second quantized *operators*. Taking into account the statistics of the electrons, distribution functions and correlation functions are introduced. The hierarchy of equations is truncated in a *systematic* manner consistent with the SCF approach and a *closed* set of equations is obtained. The irreversibility is then introduced by assuming with Bogoliubov[8] that the two-electron correlation function *relaxes* and comes into equilibrium *with* the one-electron distribution function in a time much shorter than the relaxation time of the distribution function itself.

II. THE GENERAL FORMALISM

We consider a system of N electrons with a mass m and a charge $+e$ in a crystal occupying a unit volume. The electrons are interacting via the Coulomb potential $v(r) = e^2/r$ and moving in the periodic potential $V(\mathbf{r})$ of the crystal. In the second quantization representation the Hamiltonian of the system reads (with $\hbar = 1$)

$$H = \sum_{\sigma} \int d\mathbf{r}\, \psi_\sigma^\dagger(\mathbf{r}) \left[-\frac{1}{2m}\frac{\partial^2}{\partial \mathbf{r}^2} + V(\mathbf{r}) \right] \psi_\sigma(\mathbf{r}) + \tfrac{1}{2} \sum_{\sigma,\sigma'} \int d\mathbf{r}\, d\mathbf{r}'\, \psi_\sigma^\dagger(\mathbf{r}) \psi_{\sigma'}^\dagger(\mathbf{r}') v(\mathbf{r}-\mathbf{r}') \psi_{\sigma'}(\mathbf{r}') \psi_\sigma(\mathbf{r}), \quad (1)$$

where $\psi_\sigma^\dagger(\mathbf{r})$ and $\psi_\sigma(\mathbf{r})$ are, respectively, the creation and annihilation operators of the electrons with spin σ and position $\mathbf{r}$, satisfying the usual anti-commutation relations

$$[\psi_\sigma(\mathbf{r}),\, \psi_{\sigma'}^\dagger(\mathbf{r}')]_+ = \delta_{\sigma\sigma'}\delta(\mathbf{r}-\mathbf{r}'), \qquad [\psi_\sigma(\mathbf{r}),\, \psi_{\sigma'}(\mathbf{r}')]_+ = [\psi_\sigma^\dagger(\mathbf{r}),\, \psi_{\sigma'}^\dagger(\mathbf{r}')]_+ = 0, \quad (2)$$

* This work was supported by the U. S. Atomic Energy Commission.

† On leave of absence from Technion, Israel Institute of Technology, Haifa, Israel.

[1] M. N. Rosenbluth and N. Rostoker, Phys. Fluids **3**, 1 (1960); R. Balescu, Phys. Fluids **3**, 52 (1960); A. Lenard, Ann. Phys. (N. Y.) **10**, 390 (1960); R. L. Guernsey, Phys. Fluids **5**, 322 (1962).

[2] T. H. Dupree, Phys. Fluids **4**, 696 (1961).

[3] R. L. Guernsey, Phys. Rev. **127**, 1446 (1962); R. Balescu, Phys. Fluids **4**, 94 (1960); V. P. Silin, Zh. Eksperim, i Teor. Fiz. **40**, 1768 (1961) [English transl.: Soviet Phys.—JETP **13**, 1244 (1961)]; H. W. Wyld, Jr. and D. Pines, Phys. Rev. **127**, 1851 (1962); H. W. Wyld, Jr. and B. D. Fried, Ann. of Phys. **23**, 374 (1963).

[4] D. Pines and J. R. Schrieffer, Phys. Rev. **125**, 804 (1962).

[5] A. Ron, J. Math. Phys. **4**, 1182 (1963).

[6] D. Bohm and D. Pines, Phys. Rev. **92**, 609 (1953).

[7] H. Ehrenreich and M. H. Cohen, Phys. Rev. **115**, 786 (1959).

[8] N. H. Bogoliubov, *Studies of Statistical Mechanics*, edited by J. de Boer and G. E. Uhlenbeck (North-Holland Publishing Company, Amsterdam, 1962).

with the usual notations for δ Kronecker and δ Dirac.

Making use of the fact that in the Heisenberg representation an operator $A(t)$ satisfies

$$i(\partial A/\partial t) = [A, H] = AH - HA, \tag{3}$$

we obtain for the creation and annihilation operators

$$i\frac{\partial \psi_\sigma(\mathbf{r})}{\partial t} = \left[-\frac{1}{2m}\frac{\partial^2}{\partial \mathbf{r}^2} + V(\mathbf{r})\right]\psi_\sigma(\mathbf{r}) + \sum_{\sigma'}\int d\mathbf{r}' v(\mathbf{r}-\mathbf{r}')\psi^\dagger_{\sigma'}(\mathbf{r}')\psi_{\sigma'}(\mathbf{r}')\psi_\sigma(\mathbf{r}), \tag{4a}$$

and

$$i\frac{\partial \psi^\dagger_\sigma(\mathbf{r})}{\partial t} = -\left[-\frac{1}{2m}\frac{\partial^2}{\partial \mathbf{r}^2} + V(\mathbf{r})\right]\psi^\dagger_\sigma(\mathbf{r}) - \sum_{\sigma'}\int d\mathbf{r}' \psi^\dagger_\sigma(\mathbf{r}) v(\mathbf{r}-\mathbf{r}')\psi^\dagger_{\sigma'}(\mathbf{r}')\psi_{\sigma'}(\mathbf{r}'), \tag{4b}$$

with the time t suppressed for simplicity.

We now introduce the one-electron density matrix

$$\begin{aligned} F_{\sigma_1}(1, 1') &\equiv F_{\sigma_1}(\mathbf{r}_1, \mathbf{r}_1', t) \\ &\equiv \mathrm{Tr}\,\{D\psi^\dagger_{\sigma_1}(\mathbf{r}_1', t)\psi_{\sigma_1}(\mathbf{r}_1, t)\} \\ &\equiv \langle \psi^\dagger_{\sigma_1}(\mathbf{r}_1')\psi_{\sigma_1}(\mathbf{r}_1)\rangle, \end{aligned} \tag{5}$$

where D is the density matrix of the whole system. In the same manner we define the two-electron density matrix

$$F_{\sigma_1\sigma_2}(1, 2; 1', 2') = \langle \psi^\dagger_{\sigma_1}(\mathbf{r}_1')\psi^\dagger_{\sigma_2}(\mathbf{r}_2')\psi_{\sigma_2}(\mathbf{r}_2)\psi_{\sigma_1}(\mathbf{r}_1)\rangle, \tag{6}$$

and so forth. In the Heisenberg representation D is time-independent, and the equations of motion of the F's are determined by Eq. (4); that is if we introduce the shorthand

$$T_i = \frac{1}{2m}\left(\frac{\partial^2}{\partial \mathbf{r}_i^2} - \frac{\partial^2}{\partial \mathbf{r}_i'^2}\right) - [V(\mathbf{r}_i) - V(\mathbf{r}_i')], \tag{7}$$

$$W_{ij} = v(\mathbf{r}_i - \mathbf{r}_j) - v(\mathbf{r}_i' - \mathbf{r}_j'), \tag{8a}$$

and

$$\bar{W}_{ij} = v(\mathbf{r}_i - \mathbf{r}_j) - v(\mathbf{r}_i' - \mathbf{r}_j), \tag{8b}$$

we obtain the hierarchy of equations

$$\left(i\frac{\partial}{\partial t} + T_1\right)F_{\sigma_1}(1, 1') = \sum_{\sigma_2}\int d\mathbf{r}_2 \bar{W}_{12} F_{\sigma_1\sigma_2}(1, 2; 1'; 2), \tag{9}$$

$$\left(i\frac{\partial}{\partial t} + T_1 + T_2 - W_{12}\right)F_{\sigma_1\sigma_2}(1, 2; 1', 2') = \sum_{\sigma_3}\int d\mathbf{r}_3(\bar{W}_{13} + \bar{W}_{23})F_{\sigma_1\sigma_2\sigma_3}(1, 2, 3; 1', 2', 3), \tag{10}$$

and so on. Equations (9) and (10) represent the lowest members of the hierarchy. The exact solution of the equations of the hierarchy is a hopeless task, and thus we shall resort to approximate methods of truncating the hierarchy.

Before we turn to the approximations, it is convenient to introduce an antisymmetrization operator

$$\gamma_n = \prod_{i=2}^{n}\left[1 - \sum_{k=1}^{i-1}\delta_{\sigma i,\sigma k}P_{i,k}\right], \tag{11}$$

where P_{ik} permutes the $\mathbf{r}_i$ and $\mathbf{r}_k$. Now if we define new functions by means of

$$\begin{aligned} F_{\sigma_1}(1, 1') &= \gamma_1 f_{\sigma_1}(1, 1') = f_{\sigma_1}(1, 1'), \\ F_{\sigma_1\sigma_2}(1, 2; 1', 2') &= \gamma_2 f_{\sigma_1\sigma_2}(1, 2; 1', 2'), \quad \text{etc.}, \end{aligned} \tag{12}$$

and use the fact that γ_n satisfies

$$\gamma_{n+1} = \gamma_n\left(1 - \sum_{i=1}^{n}\delta_{\sigma i,\sigma_{n+1}}P_{i,n+1}\right) \tag{13}$$

and commutes with the operators T and W, we obtain from Eqs. (9) and (10)

$$\left(i\frac{\partial}{\partial t} + T_1\right)f_{\sigma_1}(1, 1') = \sum_{\sigma_2}\int d\mathbf{r}_2 \bar{W}_{12}(1 - \delta_{\sigma_1,\sigma_2}P_{1,2}) \times f_{\sigma_1\sigma_2}(1, 2; 1', 2), \tag{14}$$

and

$$\left(i\frac{\partial}{\partial t} + T_1 + T_2 - W_{12}\right)f_{\sigma_1\sigma_2}(1, 2; 1', 2') = \sum_{\sigma_3}\int d\mathbf{r}_3(\bar{W}_{13} + \bar{W}_{23})(1 - \delta_{\sigma_1,\sigma_3}P_{1,3} - \delta_{\sigma_2,\sigma_3}P_{2,3})f_{\sigma_1\sigma_2\sigma_3}(1, 2, 3; 1', 2', 3). \tag{15}$$

To truncate the hierarchy we assume with Bogoliubov[8] that the intrinsic correlation functions between particles can be treated as small, compared to the product of one-particle functions. For example, if we write

$$f_{\sigma_1\sigma_2}(1, 2; 1', 2') = f_{\sigma_1}(1, 1')f_{\sigma_2}(2, 2') + g_{\sigma_1\sigma_2}(1, 2; 1', 2'), \tag{16}$$

we can treat g as small compared to ff. This approximation rests on the observation that, roughly speaking, g is proportional to the ratio (average potential energy)/(average kinetic energy) per particle, and dies out when the particles are far away from each other. The same argument applies to the higher correlation functions, and thus we have

$$f_{\sigma_1\sigma_2\sigma_3}(1, 2, 3; 1', 2', 3') = f_{\sigma_1}(1, 1')f_{\sigma_2}(2, 2')f_{\sigma_3}(3, 3') + f_{\sigma_1}(1, 1')g_{\sigma_2\sigma_3}(2, 3; 2', 3') + f_{\sigma_2}(2, 2')g_{\sigma_1\sigma_3}(1, 3; 1', 3') + f_{\sigma_3}(3, 3')g_{\sigma_1\sigma_2}(1, 2; 1', 2') + h_{\sigma_1\sigma_2\sigma_3}(1, 2, 3; 1', 2', 3'), \quad (17)$$

where h is small compared to fg, etc.

It should be pointed out that Eqs. (16) and (17) do not carry the symmetry properties of the system, which is taken care of by the equations of motion [Eqs. (14) and (15)] themselves.

If we now substitute Eqs. (16) and (17) into Eqs. (14) and (15), we obtain to first order

$$\left[i\frac{\partial}{\partial t} + T_1 - \sum_{\sigma_2}\int d\mathbf{r}_2 \bar{W}_{12} f_{\sigma_2}(2, 2)\right] f_{\sigma_1}(1, 1') + \int d\mathbf{r}_2 \bar{W}_{12} f_{\sigma_1}(2, 1') f_{\sigma_1}(1, 2) = \sum_{\sigma_2}\int d\mathbf{r}_2 \bar{W}_{12} g_{\sigma_1\sigma_2}(1, 2; 1', 2), \quad (18)$$

and

$$\left[i\frac{\partial}{\partial t} + T_1 + T_2 - \sum_{\sigma_3}\int d\mathbf{r}_3(\bar{W}_{13} + \bar{W}_{23}) \times f_{\sigma_3}(3, 3)\right] g_{\sigma_1\sigma_2}(1, 2; 1', 2') - \sum_{\sigma_3}\int d\mathbf{r}_3[\bar{W}_{13} f_{\sigma_1}(1, 1') g_{\sigma_2\sigma_3}(2, 3; 2', 3) + \bar{W}_{23} f_{\sigma_2}(2, 2') g_{\sigma_1\sigma_3}(1, 3; 1', 3)] = W_{12} f_{\sigma_1}(1, 1') f_{\sigma_2}(2, 2') - \int d\mathbf{r}_3[\bar{W}_{13} f_{\sigma_1}(1, 1') f_{\sigma_2}(2, 3) f_{\sigma_3}(3, 2') + \bar{W}_{23} f_{\sigma_1}(1, 3) f_{\sigma_2}(2, 2') f_{\sigma_3}(3, 1')]. \quad (19)$$

It is to be noticed that, although we assume the average potential energy per particle is small compared to the average kinetic energy per particle, the *long range* of the Coulomb potential produces "shielding" effects and thus terms including an *integration* over the Coulomb potential are *not* small. This does not apply, of course, to the "exchange" terms, where the range of the integration is *reduced* to the order of the de Broglie wavelength. This scheme is in the spirit of the SCF approach.

Equations (18) and (19) are not yet suitable for further investigation, being integrodifferential equations in the configuration space, and one should seek for better representation for them. A "natural" representation for the electrons in the periodic system seems to be the "Bloch wave" representation. We, thus, introduce the Bloch functions $\phi_{n,\mathbf{p}}(\mathbf{r})$ by the equation

$$\left[-\frac{1}{2m}\frac{\partial^2}{\partial \mathbf{r}^2} + V(\mathbf{r})\right]\phi_{n,\mathbf{p}}(\mathbf{r}) = E(n, \mathbf{p})\phi_{n,\mathbf{p}}(\mathbf{r}), \quad (21)$$

with the solution

$$\phi_{n,\mathbf{p}}(\mathbf{r}) = e^{i\mathbf{p}\cdot\mathbf{r}} u_{n,\mathbf{p}}(\mathbf{r}), \quad (22)$$

where, as usual, $\mathbf{p}$ is the quasimomentum, n represents the band, and $u_{n,\mathbf{p}}(\mathbf{r})$ has the periodicity of the crystal.

In terms of the Bloch states we introduce the one-electron distribution function

$$f_\sigma(n, \mathbf{p}) = \int d\mathbf{r}_1\, d\mathbf{r}_1' \phi_{n,\mathbf{p}}(\mathbf{r}')\phi^*_{n,\mathbf{p}}(\mathbf{r}) f_\sigma(1, 1'), \quad (23)$$

which can be written as

$$f_\sigma(n, \mathbf{p}) = \langle a^\dagger_{n,\mathbf{p},\sigma} a_{n,\mathbf{p},\sigma}\rangle \quad (24)$$

with the normalization

$$N = \sum_{n,\mathbf{p},\sigma} f_\sigma(n, \mathbf{p}). \quad (25)$$

In Eq. (24), $a^\dagger_{n,\mathbf{p},\sigma}$ and $a_{n,\mathbf{p},\sigma}$ are, respectively, the creation and annihilation operators for the Bloch electrons. Furthermore, we assume

$$f_\sigma(1, 1') = \sum_{n,\mathbf{p}} \phi^*_{n,\mathbf{p}}(\mathbf{r}_1')\phi_{n,\mathbf{p}}(\mathbf{r}) f_\sigma(n, \mathbf{p}), \quad (26)$$

that is, the f_σ of Eqs. (16) and (17) corresponds to the *diagonal* terms of the density matrix in the Bloch representation, while the nondiagonal terms are incorporated into the correlation functions. This assumption may be argued to be in the spirit of our truncation scheme as a generalization to the periodic system case, of the usual treatment of homogeneous systems. It is justified *a posteriori* by the explicit dependence of the correlation function on the strength of the interaction. [Actually, it is proportional to the ratio (potential energy)/(kinetic energy) which is small for the systems under consideration.]

We now introduce the shorthand $\alpha \equiv n, \mathbf{p}$ to represent the Bloch electron with the quasimomentum $\mathbf{p}$ and band number n, and write

$$g_{\sigma_1\sigma_2}(\alpha_1; \alpha_2; \alpha_1', \alpha_2') = \int d\mathbf{r}_1\, d\mathbf{r}_1'\, d\mathbf{r}_2\, d\mathbf{r}_2' \phi_{\alpha_1'}(\mathbf{r}_1') \times \phi_{\alpha_2'}(\mathbf{r}_2')\phi^*_{\alpha_2}(\mathbf{r}_2)\phi^*_{\alpha_1}(r_1) g_{\sigma_1\sigma_2}(1, 2; 1', 2') \quad (27)$$

with its proper inversion. It should be noticed that we have not omitted any "off diagonal" terms in g.

Further, we denote by v_k the Fourier transform of $v(r)$ and by η the matrix element

$$\eta(\alpha, \alpha', \mathbf{k}) = \int d\mathbf{r}\phi^*_{\alpha'}(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}\phi_\alpha(\mathbf{r}). \tag{28}$$

By neglecting local field corrections the latter reads

$$\begin{aligned}\eta(\alpha, \alpha', \mathbf{k}) &= \eta(n, \mathbf{p}, n', \mathbf{p}+\mathbf{k}) \\ &= \delta_{\mathbf{p}', \mathbf{p}+\mathbf{k}}(n', \mathbf{p}+\mathbf{k} \mid n, \mathbf{p}),\end{aligned} \tag{29}$$

with

$$(n', \mathbf{p}+\mathbf{k} \mid n, \mathbf{p}) = \frac{1}{V_0}\int_0 d\mathbf{r}u^*_{n', \mathbf{p}+\mathbf{k}}(\mathbf{r})u_{n,\mathbf{p}}(\mathbf{r}), \tag{30}$$

where the integral extends over the unit cell occupying the volume V_0. Equation (29) rests on the assumption that the core states of the atoms composing the crystal are sufficiently tightly bound, while the valence or conduction bands are sufficiently broad.[9] Using Eqs. (23), (26), (27), and (28) we obtain for Eq. (18)

$$\frac{\partial}{\partial t} f_{\sigma_1}(\alpha_1) = -2\sum_{\mathbf{k}} v_k \sum_{\alpha_1'} \mathrm{Im}\,\{G_{\sigma_1}(\alpha_1, \alpha_1', \mathbf{k})\}, \tag{31}$$

where

$$\begin{aligned}G_{\sigma_1}(\alpha_1, \alpha_1', \mathbf{k}) = &\sum_{\sigma_2, \alpha_2, \alpha_2'} \eta(\alpha_1, \alpha_1', \mathbf{k}) \\ &\times \eta(\alpha_2, \alpha_2', -\mathbf{k})g_{\sigma_1\sigma_2}(\alpha_1, \alpha_2; \alpha_1', \alpha_2').\end{aligned} \tag{32}$$

In deriving Eq. (31), use has been made of the fact that

$$g^*_{\sigma_1\sigma_2}(\alpha_1', \alpha_2'; \alpha_1, \alpha_2) = g_{\sigma_1\sigma_2}(\alpha_1, \alpha_2; \alpha_1', \alpha_2'). \tag{33}$$

We notice that in order to obtain the kinetic equation [Eq. (31)] one does not attempt to find the correlation function, but rather a "moment" of it, given by Eq. (32). This feature is common to kinetic equations of other systems.[1,3,5]

To proceed, we transform Eq. (19) to the Bloch representation and obtain

$$\begin{aligned}&\left[i\frac{\partial}{\partial t} + E(\alpha_1') - E(\alpha_1) + E(\alpha_2') - E(\alpha_2)\right] \\ &\times g_{\sigma_1\sigma_2}(\alpha_1, \alpha_2; \alpha_1', \alpha_2') \\ &- \sum_{\mathbf{q}} v_q\eta(\alpha_1', \alpha_1, \mathbf{q})[f_{\sigma_1}(\alpha_1') - f_{\sigma_1}(\alpha_1)] \\ &\times \sum_{\sigma_3, \alpha_3\alpha_3'} \eta(\alpha_3, \alpha_3', -\mathbf{q})g_{\sigma_2\sigma_2}(\alpha_2, \alpha_3; \alpha_2', \alpha_3') \\ &- \sum_{\mathbf{q}} v_q\eta(\alpha_2', \alpha_2, \mathbf{q})[f_{\sigma_2}(\alpha_2') - f_{\sigma_2}(\alpha_2)] \\ &\times \sum_{\sigma_3, \alpha_3, \alpha_3'} \eta(\alpha_3, \alpha_3', -\mathbf{q})g_{\sigma_1\sigma_2}(\alpha_1, \alpha_3; \alpha_1', \alpha_3') \\ &= \sum_{\mathbf{q}} v_q\eta(\alpha_1', \alpha_1, \mathbf{q})\eta(\alpha_2', \alpha_2, -\mathbf{q}) \\ &\times [F_{\sigma_1}(\alpha_1, \alpha_1')F_{\sigma_2}(\alpha_2, \alpha_2') \\ &- F_{\sigma_1}(\alpha_1', \alpha_1)F_{\sigma_2}(\alpha_2', \alpha_2)],\end{aligned} \tag{34}$$

where

$$F_\sigma(\alpha, \alpha') = f_\sigma(\alpha')[1 - f_\sigma(\alpha)]. \tag{35}$$

The remainder of the report is devoted to the solution of Eq. (34), which leads to the kinetic equation for the electrons in the crystal.

III. THE ADIABATIC HYPOTHESIS

By inspection one sees that Eq. (34) is a very complicated integral equation in the q, α, α' space and it is almost a hopeless task to solve it in general. Fortunately, only the combination

$$\begin{aligned}\mathcal{G}(\beta_1^+, \beta_2^-) = &\eta(\alpha_1, \alpha_1', \mathbf{k}) \\ &\times \eta(\alpha_2, \alpha_2', -\mathbf{k})g_{\sigma_1\sigma_2}(\alpha_1, \alpha_2; \alpha_1', \alpha_2')\end{aligned} \tag{36}$$

is of interest in order to obtain the kinetic equation, and we may proceed, having this in mind. For the sake of shorthand writing we have introduced in Eq. (36) the notation $\beta^\pm$ to stand for the spin σ and the Bloch states α, α', with $\alpha \equiv n, \mathbf{p}$ and $\alpha' \equiv n', p \pm k$,

$$\beta^\pm \equiv \sigma; n, \mathbf{p}; n', \mathbf{p} \pm \mathbf{k}. \tag{37}$$

If we now multiply Eq. (34) by

$$\eta(\alpha, \alpha_1', \mathbf{k})\eta(\alpha_2, \alpha_2', -\mathbf{k}),$$

we can cast this equation to the form (compare with Ref. 2)

$$\begin{aligned}&[i(\partial/\partial t) + H(\beta_1^+) + H(\beta_2^-)] \\ &\times \mathcal{G}(\beta_1^+, \beta_2^-) = S(\beta_1^+, \beta_2^-),\end{aligned} \tag{38}$$

where $H(\beta^\pm)$ is the integral operator

$$\begin{aligned}H(\beta^\pm) = &E(n', \mathbf{p} \pm \mathbf{k}) - E(n, \mathbf{p}) \\ &- v_k\,|(n', \mathbf{p} \pm \mathbf{k} \mid n, \mathbf{p})|^2\,[f_\sigma(n', \mathbf{p} \pm \mathbf{k}) \\ &- f_\sigma(n, \mathbf{p})]\sum_{\sigma, n, n', \mathbf{p}, \mathbf{p}'} \delta_{\mathbf{p}', \mathbf{p}\pm\mathbf{k}},\end{aligned} \tag{39}$$

and the source term

$$\begin{aligned}&S(\beta_1^+, \beta_2^-) \\ &= v_k\,|(n_1', \mathbf{p}_1 + \mathbf{k} \mid n_1, \mathbf{p}_1)|^2\,|(n_2', \mathbf{p}_2 - \mathbf{k} \mid n_2, \mathbf{p}_2)|^2 \\ &\times [F_{\sigma_1}(n_1, \mathbf{p}_1; n_1', \mathbf{p}_1 + \mathbf{k})F_{\sigma_2}(n_2, \mathbf{p}_2; n_2', \mathbf{p}_2 - \mathbf{k}) \\ &- F_{\sigma_1}(n_1', \mathbf{p}_1 + \mathbf{k}; n_1, \mathbf{p}_1)F_{\sigma_2}(n_2', \mathbf{p}_2 - \mathbf{k}; n_2, \mathbf{p}_2)].\end{aligned} \tag{40}$$

In obtaining Eq. (38) we have made use of the fact

[9] See, e. g., Ref. 7 and N. Wiser, Phys. Rev. 129, 62 (1963).

that

$$g_{\sigma_1\sigma_2}(\alpha_1, \alpha_2; \alpha_1', \alpha_2') = g_{\sigma_2\sigma_1}(\alpha_2, \alpha_1; \alpha_2', \alpha_1'). \quad (41)$$

The method of solution of Eq. (38), which is originally due to Dupree,[2] rests on the observation that the inversion of the operator $H(\beta)$ is a much simpler problem, then the solution of the full integral equation for $\mathcal{G}$, and that $H(\beta_1^+)$, $H(\beta_2^-)$ commute with each other, being functions of different variables. We give here a brief account of the method. It is recommended that the reader refers to the original elegant paper of Dupree[2,10] for more details and interpretation.

We define an operator $P(\beta^\pm, t)$ by the equation

$$[i(\partial/\partial t) + H(\beta^\pm, t)]P(\beta^\pm, t) = 0, \quad (42)$$

with the initial condition $P(\beta^\pm, 0) = I$, where I is the unity operator. It should be pointed out that the operator H depends on the time only via the distribution function f. It is now evident by inspection that

$$\mathcal{G}(\beta_1^+, \beta_2^-, t) = -i \int_0^t dt' P(\beta_1^+, t - t')P(\beta_2^-, t - t')S(t'), \quad (43)$$

where the initial conditions for $\mathcal{G}$ were left out for convenience, being of no relevance to what follows. Thus the problem has been reduced to the solution of Eq. (42), which is of lower dimension and easier to treat, although not trivial. Once this has been accomplished, the right-hand side of Eq. (43) is a known functional of the distribution function f, and with Eq. (32), Eq. (31) is, in principle, an equation for f alone, as required for a kinetic equation.

Unfortunately, we are not able to solve Eq. (42) in general. However invoking the *physical* assumption, which is due to Bogoliubov,[8] that the correlation function of two electrons reaches an asymptotic value in a time short compared with the time in which the distribution function changes appreciably, Eq. (42) can be solved. This is the so called adiabatic hypothesis, which leads to an irreversible approach to equilibrium of the distribution function. Thus, in solving Eq. (34) or Eq. (38) we assume f to be time-independent and look for the asymptotic solution for g with the initial correlations washed out. To this end Eq. (43) is written as

$$\mathcal{G}(\beta_1^+, \beta_2^-) = -i \int_0^\infty d\tau e^{-\epsilon\tau} P(\beta_1^+, \tau)P(\beta_2^-, \tau)S, \quad (44)$$

where the source term S is time-independent [see Eq. (40)], and ϵ is a small positive convergence factor, introduced to ensure the vanishing of the integrand at infinity, and the limit $\epsilon \to 0$ is understood.

To proceed we employ the Laplace transform methods. Introducing the notations

$$\Delta(\beta^\pm) = E(n', \mathbf{p} \pm \mathbf{k}) - E(n, \mathbf{p}), \quad (45)$$

and

$$A(\beta^\pm) = v_k \, |(n', \mathbf{p} \pm \mathbf{k} \mid n, \mathbf{p})|^2 \times [f_\sigma(n', \mathbf{p} \pm \mathbf{k}) - f_\sigma(n, \mathbf{p})] \quad (46)$$

in Eq. (39), we write Eq. (42) as

$$[is + \Delta(\beta)]P(\beta, s) - A(\beta) \sum_{\beta'} P(\beta', s) = iI, \quad (47)$$

with the formal solution

$$\sum_{\beta^\pm} P(\beta^\pm, s) = \frac{1}{\epsilon_\pm(s)} \sum_{\beta^\pm} \frac{iI}{is + \Delta(\beta^\pm)}, \quad (48)$$

$$P(\beta^\pm, s) = \frac{1}{is + \Delta(\beta^\pm)} \times \left[iI + \frac{A(\beta^\pm)}{\epsilon_\pm(s)} \sum_{\beta^\pm} \frac{iI}{is + \Delta(\beta^\pm)} \right], \quad (49)$$

where the dielectric functions, $\epsilon_\pm(s)$, are given by

$$\epsilon_\pm(s) = 1 - \sum_{\beta^\pm} \frac{A(\beta^\pm)}{is + \Delta(\beta^\pm)}. \quad (50)$$

The correlation function now reads

$$\mathcal{G}(\beta_1^+, \beta_2^-) = -i \int_0^\infty dt e^{-\epsilon t} \int_{C_1} \frac{ds_1}{2\pi i} e^{s_1 t} P(\beta_1^+, s_1) \times \int_{C_2} \frac{ds_2}{2\pi i} e^{s_2 t} P(\beta_2^-, s_2) S(\beta_1^+, \beta_2^-), \quad (51)$$

where the contours C_1 and C_2 are the usual Laplace-transforms inversion contours, located to the right of the imaginary axis, to ensure the convergence of the integrals.

It is now convenient to rotate s_1 and s_2 by $\pm\pi$, respectively, by choosing $s_1 = iz_1$ and $s_2 = -iz_2$, and to fix C_1^* and C_2^* to be straight lines parallel to the real axis in a distance $-\delta_1$, $+\delta_2$ from this axis, respectively,

$$\mathcal{G}(\beta_1^+, \beta_2^-) = -i \int_{C_1^*} \frac{dz_1}{2\pi i} \int_{C_2^*} \frac{dz_2}{2\pi i} P(\beta_1^+, iz_1) \times P(\beta_2^-, -iz_2) \int_0^\infty dt e^{-i(z_2 - z_1 - i\epsilon)t} S(\beta_1^+, \beta_2^-). \quad (52)$$

Performing the t integration under the requirement

$$\epsilon > \mathrm{Im}\,\{z_2 - z_1\}, \quad (53)$$

[10] See also P. A. Wolff, Phys. Fluids 5, 316 (1962).

we have

$$\mathcal{G}(\beta_1^+, \beta_2^-) = -\int_{C_1^*} \frac{dz_1}{2\pi i} \int_{C_2^*} \frac{dz_2}{2\pi i} P(\beta_1^+, iz_1) \times P(\beta_2^-, -iz_2) \frac{1}{z_2 - z_1 - i\epsilon} S(\beta_1^+, \beta_2^-). \quad (54)$$

If we now assume that the distribution function f is stable against small perturbation [thus $\epsilon_-(-iz)$ does not vanish in the upper half-plane], the expression $P(\beta_2^-, -iz_2)S(\beta_1^+, \beta_2^-)$ is analytic, at least in the upper half-plane above the real axis, and vanishes as $1/z_2$ at infinity. We thus may close the z_2 contour by a large semicircle at the upper half-plane and perform the z_2 integration. The only contribution to this integral comes from the pole where $z_2 = z_1 + i\epsilon$ [with Eq. (53)], and yields

$$\mathcal{G}(\beta_1^+, \beta_2^-) = -\int_{C_1^*} \frac{dz_1}{2\pi i} P(\beta_1^+, iz_1) \times P[\beta_2^-, -i(z_1 + i\epsilon)] S(\beta_1^+, \beta_2^-). \quad (55)$$

Substituting Eqs. (49), (40), and (36) one obtains now, in principle, the correlation function.

Being interested in the kinetic equation, Eq. (31) tells us that we have to find only the imaginary part of

$$G(\beta_1^+) = \sum_{\beta_2^-} \mathcal{G}(\beta_1^+, \beta_2^-), \quad (56)$$

the problem is somewhat simplified, and the integration over z_1 may be carried out explicitly. To proceed we use Eqs. (48), (49), and (40), and make the change of variables $z_1 = \omega - i\delta$ with $\delta < \epsilon$; we obtain

$$\begin{aligned} \text{Im}\,\{G(\beta_1^+)\} = \text{Im} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi i} \frac{1}{\Delta(\beta_1^+) - \omega + i\delta} \\ \times \left\{1 + \frac{A(\beta_1^+)}{\epsilon(k, \omega)} \sum_{\beta_2^+} \frac{1}{\Delta(\beta_1^+) - \omega + i\delta}\right\} \\ \times \frac{1}{\epsilon^*(k, \omega)} \sum_{\beta_2^-} \frac{1}{\Delta(\beta_2^-) + \omega + i\delta} v_k \{\beta_1^+\}\{\beta_2^-\} \\ \times [F(\beta_1^+)F(\beta_2^-) - \tilde{F}(\beta_1^+)\tilde{F}(\beta_2^-)], \end{aligned} \quad (57)$$

where the limit $\delta \to +0$ is understood. In Eq. (57) we have introduced the following notations:

$$\{\beta^\pm\} = |(n', \mathbf{p} \pm \mathbf{k} \mid n, \mathbf{p})|^2, \quad (58)$$

$$F(\beta^\pm) = f_\sigma(n', \mathbf{p} \pm \mathbf{k})[1 - f_\sigma(n, \mathbf{p})], \quad (59)$$

and

$$\tilde{F}(\beta^\pm) = f_\sigma(n, \mathbf{p})[1 - f_\sigma(n', \mathbf{p} \pm \mathbf{k})]. \quad (60)$$

We have also denoted by $\epsilon(k_1\omega)$

$$\epsilon(k, \omega) = \epsilon_+(i\omega + \delta) = 1 - \sum_{\beta^+} \frac{A(\beta^+)}{\Delta(\beta^+) - \omega + i\delta}, \quad (61)$$

which leads, by explicit evaluation, to

$$\epsilon_-(-i\omega + \delta) = \epsilon^*(k, \omega), \quad (62)$$

where the star denotes the complex conjugate. If we exhibit the explicit expression for the dielectric function

$$\epsilon(k, \omega) = 1 - v_k \sum_{\sigma, n, n', \mathbf{p}} |(n', \mathbf{p} + \mathbf{k} \mid n, \mathbf{p})|^2 \times \frac{f_\sigma(n', \mathbf{p} + \mathbf{k}) - f_\sigma(n, \mathbf{p})}{E(n', \mathbf{p} + \mathbf{k}) - E(n, \mathbf{p}) - \omega + i\delta}, \quad (63)$$

we recognize that $\epsilon(k, \omega)$ of Eq. (63) may be considered as a generalization, to the nonequilibrium case, of the longitudinal frequency and wave-vector-dependent dielectric function for electrons in a crystal, which was obtained by Eherenreich and Cohen.[7] It may be thus interpreted as the linear response of the electron system, while described by the eventual distribution function $f_\sigma(n, \mathbf{p})$ to an external test charge with wave vector $\mathbf{k}$ and frequency ω.

The evaluation of the ω integration in Eq. (57) is somewhat tedious, but straightforward, as shown in the Appendix. It yields

$$\begin{aligned} \text{Im}\, G(\beta_1^+) = -\pi v_k \{\beta_1^+\} \, |\epsilon[k, \Delta(\beta_1^+)]|^{-2} \\ \times \sum_{\beta_2^-} \{\beta_2^-\} \delta[\Delta(\beta_2^-) - \Delta(\beta_1^+)] \\ \times [F(\beta_1^+)F(\beta_2^-) - \tilde{F}(\beta_1^+)\tilde{F}(\beta_2^-)], \end{aligned} \quad (64)$$

and our final kinetic equation takes the form

$$\begin{aligned} \frac{\partial}{\partial t} f_\sigma(n, \mathbf{p}) = -2\pi \sum_{\mathbf{k}} \sum_{\mathbf{p}'} \sum_{n_1, n_2, n_3} \sum_{\sigma'} \frac{v_k^2 \, |(n_1, \mathbf{p} + \mathbf{k} \mid n, \mathbf{p})|^2 \, |(n_3, \mathbf{p}' - \mathbf{k} \mid n_2, \mathbf{p}')|^2}{|\epsilon[k, E(n_1, \mathbf{p} + \mathbf{k}) - E(n, \mathbf{p})]|^2} \\ \times \delta[E(n_1, \mathbf{p} + \mathbf{k}) - E(n, \mathbf{p}) + E(n_3, \mathbf{p}' - \mathbf{k}) - E(n_2, \mathbf{p}')] \\ \times \{f_\sigma(n_1, \mathbf{p} + \mathbf{k}) f_{\sigma'}(n_3, \mathbf{p}' - \mathbf{k})[1 - f_\sigma(n, \mathbf{p})][1 - f_{\sigma'}(n_2, \mathbf{p}')] \\ - f_\sigma(n, \mathbf{p}) f_{\sigma'}(n_2, \mathbf{p}')[1 - f_\sigma(n_1, \mathbf{p} + \mathbf{k})][1 - f_{\sigma'}(n_3, \mathbf{p}' - \mathbf{k})]\}. \end{aligned} \quad (65)$$

Equation (65) is a generalization of the kinetic equation of electrons in spatially uniform system, as was obtained by several authors,[3] to the case of electrons in periodic systems like crystals. It includes both the *individual* and *collective* aspects of the electron interactions, while embedded in a crystal. The dielectric function in the denominator represents a *dynamic shielding* of the electrons, due to their simultaneous motion in the periodic system. It may readily be seen that the kinetic equation has the equilibrium solution

$$f_\sigma(n, \mathbf{p}) = \{\exp[\beta E(n, \mathbf{p}) - \beta\mu] + 1\}^{-1}, \qquad (66)$$

and thus governs the approach to equilibrium of the system. In Eq. (66), β and μ are, respectively, the inverse temperature in energy units, and the chemical potential of the electrons.

IV. DISCUSSION

The present paper concerns the problem of extending known theories of the approach to equilibrium of an homogeneous electron system to the case of periodic systems. A kinetic equation for the electrons was derived employing a truncation scheme for the hierarchy of coupled s-body density matrices and using the adiabatic hypothesis to introduce irreversibility. The crucial mathematical problem is the solution of the integral equation for the correlation function in a periodic system. The kinetic equation, which takes into account both the individual and the collective aspects of the electron interactions, governs the dynamics of the system while approaching to equilibrium, i.e., to Fermi distribution for the electrons.

Although one is able to derive Eq. (65) using an *ad hoc* method of transition probability with an *effective* potential of interaction[3] (to include the dynamic shielding), we believe that the present method is more rigorous and may also be regarded as a *justification* of the former method. Moreover, the present method can be used to calculate the correlation function of two electrons,[10] which is of great interest by its own merit (see, e.g., the Guernsey[3] treatment of the average Coulomb energy). Furthermore, the adiabatic hypothesis, which is equivalent to the transition probability approach, may be relaxed for some cases of small perturbations around thermal equilibrium, when the validity of the hypothesis is questionable. Thus, a linearized transport equation can be developed, which does not carry the assumption of different time scales for the distribution and the correlation functions (see, e.g., Guernsey[3]).

The kinetic equation we have obtained provides a physical description of the way equilibrium is established by both interband and intraband transitions of the electrons in a crystal. It may also be applied to problems of transport theory in crystals, when the electron–electron scattering and band transitions are important. The generalization of the method used here to electron–hole systems (e.g., in semiconductors and semimetals) is straightforward.

APPENDIX

We wish to represent here a brief account of the derivation of Eq. (64) from Eq. (57). If we use Eq. (46), with $A(\beta^\pm)$ of Eq. (46) rewritten as

$$A(\beta^\pm) = v_k\{\beta^\pm\}[F(\beta^\pm) - \tilde{F}(\beta^\pm)], \qquad (A1)$$

we can write Eq. (57) as

$$\begin{aligned}
\operatorname{Im} G(\beta_1^+) = \operatorname{Im} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi i} \\
\times \frac{1}{\epsilon(k, \omega)\epsilon^*(k_1\omega)} \frac{1}{\Delta(\beta_1^+) - \omega + i\delta} v_k \\
\times \Bigg\{\Bigg(1 - v_k \sum_{\beta_3^+} \{\beta_3^+\} \frac{F(\beta_3^+) - \tilde{F}(\beta_3^+)}{\Delta(\beta_3^+) - \omega + i\delta} \\
+ v_k\{\beta_1^+\}[F(\beta_1^+) - \tilde{F}(\beta_1^+)] \sum_{\beta_1^+} \frac{1}{\Delta(\beta_1^+) - \omega + i\delta}\Bigg) \\
\times \{\beta_1^+\} \sum_{\beta_2^-} \{\beta_2^-\} \frac{F(\beta_1^+)F(\beta_2^-) - \tilde{F}(\beta_1^+)\tilde{F}(\beta_2^-)}{\Delta(\beta_2^-) - \omega + i\delta}\Bigg\}. \qquad (A2)
\end{aligned}$$

Now applying the summation operators in Eq. (A2), and collecting the terms left out after some cancellations occur, we obtain in the curly brackets

$$\begin{aligned}
\{\beta_1^+\}F(\beta_1^+)\Bigg\{\sum_{\beta_2^-} \{\beta_2^-\} \frac{F(\beta_2^-)}{\Delta(\beta_2^-) + \omega + i\delta} \\
+ \sum_{\beta_3^+} \{\beta_3^+\} \frac{\tilde{F}(\beta_3^+)}{\Delta(\beta_3^+) - \omega + i\delta}[1 - \epsilon^*(k, \omega)]\Bigg\} \\
+ \text{term } (F \leftrightarrows \tilde{F}). \qquad (A3)
\end{aligned}$$

If we substitute Eq. (A3) into (A2) we can carry out the integration of the terms with $\epsilon^*(k, \omega)$ by closing the contour in the lower half-plane. Since the integrals have no residues, we can discard the terms with $\epsilon^*(k, \omega)$ from Eq. (A3). Next we notice that

$$\begin{aligned}
\sum_{\beta^+} \{\beta^+\} \frac{\tilde{F}(\beta^+)}{\Delta(\beta^+) - \omega + i\delta} \\
= -\sum_{\beta^-} \{\beta^-\} \frac{F(\beta^-)}{\Delta(\beta^-) + \omega - i\delta}, \qquad (A4)
\end{aligned}$$

and Eq. (A3) reads

$$\{\beta_1^+\}F(\beta_1^+)\sum_{\beta_2^-}\{\beta_2^-\}F(\beta_2^-)\left[\frac{1}{\Delta(\beta_2^-)+\omega-i\delta}-\frac{1}{\Delta(\beta_2^-)+\omega+i\delta}\right]+\text{term}\ (F\leftrightarrows\tilde{F})$$
$$=2\pi i\sum_{\beta_2^-}\{\beta_1^+\}\{\beta_2^-\}[F(\beta_1^+)F(\beta_2^-)-\tilde{F}(\beta_1^+)\tilde{F}(\beta_2^-)]\delta[\Delta(\beta_2^-)+\omega]. \tag{A5}$$

Finally, we substitute Eq. (A5) into (A2) and obtain

$$\text{Im}\,G(\beta_1^+)=v_k\sum_{\beta^-}\{\beta_1^+\}\{\beta_2^-\}[F(\beta_1^+)F(\beta_2^-)-\tilde{F}(\beta_1^+)\tilde{F}(\beta_2^-)]\times\int d\omega\,\delta[\omega+\Delta(\beta_2^-)]\frac{1}{|\epsilon(k,\omega)|^2}\times\text{Im}\frac{1}{\Delta(\beta_1^+)-\omega+i\delta}, \tag{A6}$$

which, upon integration over ω, leads to Eq. (64).

Systematic Characterization of mth-Order Energy-Level Spacing Distributions*†

HARVEY S. LEFF‡§
State University of Iowa, Iowa City, Iowa
(Received 17 January 1964)

The mth-order energy-level spacing distributions $p^{(m)}(x)$ for complex spectra are defined in terms of a general joint probability distribution $P_N(\lambda_1, \cdots \lambda_N)$ for N consecutive eigenvalues. The precise limiting processes involved are explained, and are subsequently used to obtain two formal representations of $p^{(m)}(x)$. Both representations yield $p^{(m)}(x) = x^m/m! \exp(-x)$ for statistically independent eigenvalues. One of the representations, which is an extension of Dyson's method for $m = 0, 1$, is applied to the superposition of n independent sequences of levels. General asymptotic results are found for the mth-order distributions for (a) small x, arbitrary n, and (b) arbitrary x with $n \to \infty$.

I. INTRODUCTION

HEAVY and intermediate-weight nuclei and atoms have energy-level spectra which are quite complex. For such systems, statistical methods have been used to describe the level structure away from the ground state. Theoretical investigations of the statistical properties of energy-level spectra have been based on the concept of ensembles of matrices.[1] A matrix ensemble gives rise to a joint probability distribution function $P_N(\lambda_1, \cdots \lambda_N)$ for N consecutive energy eigenvalues, $\lambda_1, \cdots \lambda_N$ (not necessarily ordered with respect to labeling). From this fundamental function, one can, in principle, determine (as $N \to \infty$) all statistical properties of spectra.

The statistical properties of interest here are the mth-order spacing distributions. In particular, the present work is devoted to (1) the formal definition of the mth-order spacing distributions for all nonnegative integers m, (2) obtaining methods through which analytical calculations of these distributions can be realized, and (3) a discussion of the properties of these distributions for spectra consisting of n independently superposed sequences of levels.

This work differs from prior discussions in several respects. First, we do not specify a *particular* matrix ensemble, and thus do not work with a *particular* function $P_N(\lambda_1, \cdots \lambda_N)$. Second, we define the mth-order spacing distributions for all integers $m \geq 0$. Previous efforts have been restricted to $m \leq 1$, and even for these cases, definitions of the present type do not appear in the published literature. Third, we explain and make explicit use of the various necessary limiting processes in a formal way. This leads to a simple explanation of why the distributions of spacings relative to the mean nearest-neighbor spacing, and not of the spacings themselves, are important. Fourth, we derive, from first principles, two formalisms for practical calculations of the mth-order distributions. One of these is an extension of Wigner's infinite series representation for $m = 0$.[2] The other derivation gives rise to a very natural extension of Dyson's methods for $m = 0$ and $m = 1$.[3] Fifth, we show that the case of n superposed sequences can be treated as a simple application of the latter formalism, without the necessity of detailed combinatorial analyses.

II. DEFINITION OF mth-ORDER SPACING DISTRIBUTIONS

For generality, we assume $P_N(\lambda_1, \cdots \lambda_N)$ to be defined for $a \leq \lambda_i \leq b$, $i = 1, 2, \cdots N$. The real numbers a and b may be finite or infinite, with the restriction that $a < 0 < b$.[4] $P_N(\lambda_1, \cdots \lambda_N)$ is assumed to be a normalized distribution function,

$$\int_a^b \cdots \int_a^b P_N(\lambda_1, \cdots \lambda_N)\, d\lambda_1 \cdots d\lambda_N = 1. \quad (1)$$

The density of levels $D^{-1}(\lambda)$ at the point λ is just

$$D^{-1}(\lambda) = N \int_a^b \cdots \int_a^b P_N(\lambda, \lambda_2, \lambda_3, \cdots \lambda_N)\, d\lambda_2 \cdots d\lambda_N. \quad (2)$$

* This work is based on portions of a Ph.D. thesis submitted before the faculty of the State University of Iowa, July, 1963.
† This work was supported in part by the Air Force Office of Scientific Research.
‡ State University of Iowa Predoctoral Fellow 1962–63.
§ Present address: Department of Physics, Case Institute of Technology, Cleveland 6, Ohio.
[1] For a review of matrix ensembles and an extensive bibliography, see N. Rosenzweig, "Statistical Mechanics of Equally Likely Quantum Systems" in *Statistical Physics* (W. A. Benjamin Company, Inc., New York, 1963), pp. 91–158.
[2] E. P. Wigner, "Distribution Laws for the Roots of a Random Hermitian Matrix," 1962 (unpublished).
[3] F. J. Dyson, J. Math. Phys. 3, 166 (1962).
[4] This restriction is made for later (mathematical) convenience when we discuss spacing distributions in the neighborhood of zero energy. Since the reference level of the energy scale is arbitrary, the restriction does not lessen the generality of what follows.

The spacing distributions of interest physically apply to groupings of a *finite number* of levels in some domain R such that for all λ in R, $D^{-1}(\lambda)$ is essentially a constant D^{-1}. However, matrix ensemble formulations require that the limit of infinite matrix dimensionality ($N \to \infty$) be taken. In this limit, every point in $[a, b]$ will have infinite level density. This is a consequence of the normalization condition, and is true for all reasonable ensembles. The only way which we can mathematically obtain a spacing distribution for a finite number of levels and still let $N \to \infty$ is to let the measure of R shrink to zero as N increases. A consequence of this double limiting process is that it is not sensible to talk about the distribution of the spacings themselves. It is however possible and, in fact, highly suggestive to examine the distributions of spacings relative to the mean nearest-neighbor spacing D in R.

For finite N, we define the conditional distributions $f_N^{(m)}(\sigma \mid \lambda)$: Given a level at λ, the probability to find a level in $(\lambda + \sigma, \lambda + \sigma + d\sigma)$ and m levels in $(\lambda, \lambda + \sigma)$ is $f_N^{(m)}(\sigma \mid \lambda)d\sigma$ $(\sigma \geq 0)$. The corresponding conditional mth-order distribution for spacings relative to $D(\lambda)$ is denoted by $p_N^{(m)}(x \mid \lambda)$:

$$p_N^{(m)}(x \mid \lambda) = D(\lambda) f_N^{(m)}(\sigma \mid \lambda), \tag{3}$$

$$x \equiv \sigma D^{-1}(\lambda). \tag{4}$$

Finally, the entities of physical interest, which are referred to simply as the mth-order spacing distributions, are

$$p_\lambda^{(m)}(x) \equiv \lim_{N\to\infty,\,\sigma\to 0} D(\lambda) f_N^{(m)}(x \mid \lambda). \tag{5}$$

The double limit is to be taken keeping x finite and fixed. It is usually convenient to examine the spacing distributions in a domain R_0 centered about the origin. If this is done, the λ dependence of $p_\lambda^{(m)}(x)$ "disappears" and we call the resulting function $p^{(m)}(x)$:

$$p_0^{(m)}(x) \equiv p^{(m)}(x). \tag{6}$$

The philosophy used is that "suitable" matrix ensembles will give rise to level densities which are slowly varying functions of λ near the origin. Adopting this restriction, the spacing distributions of particular interest are given by (5), with $\lambda = 0$.

It now remains to express $p^{(m)}(x)$ in terms of the fundamental function $P_N(\lambda_1, \cdots \lambda_N)$. For this purpose, we introduce the following integral operators:

$$S_k^{(x,y)} \equiv -I_k^{(x,y)} \equiv \int_y^x d\lambda_k, \tag{7}$$

$$O_k^{(x,y)} \equiv \int_a^x d\lambda_k + \int_y^b d\lambda_k, \tag{8}$$

$$T_k \equiv O_k^{(x,y)} - S_k^{(x,y)}. \tag{9}$$

Since the variables λ_i are not ordered with respect to labeling, $P_N(\lambda_1, \cdots \lambda_N)$ is symmetric under the transposition of any two variables λ_i and λ_k, $i \neq k$. Therefore

$$[T_i S_k^{(x,y)} - T_k S_i^{(x,y)}] P_N(\lambda_1, \cdots \lambda_N) = 0 \tag{10}$$

for all $i \neq k$. Furthermore, since all integrals are interchangeable, the commutator between any two of the above operators (with different subscripts) acting on $P_N(\lambda_1, \cdots \lambda_N)$ must vanish.

The conditional distribution $f_N^{(0)}(\sigma \mid \lambda)$ is obtained from $P_N(\lambda_1, \cdots \lambda_N)$ by the relation

$$P(\lambda) f_N^{(0)}(\sigma \mid \lambda) = (N-1)\left\{\prod_{k=1}^{(N-2)} O_k^{(\lambda,\lambda+\sigma)}\right\} \times P_N(\lambda_1, \cdots \lambda_{N-2}, \lambda, \lambda + \sigma), \tag{11}$$

where

$$P(\lambda) \equiv N^{-1} D^{-1}(\lambda). \tag{12}$$

The factor $(N-1)$ arises from the fact that the level at λ can be any one of the first $(N-1)$ *ordered* eigenvalues (the largest eigenvalue has no level to its right). For $f^{(m)}(\sigma \mid \lambda)$, the direct generalization of (11) is

$$P(\lambda) f_N^{(m)}(\sigma \mid \lambda) = (N-1)\binom{N-2}{m}\left\{\prod_{k=1}^{(N-m-2)} O_k^{(\lambda,\lambda+\sigma)}\right\} \times \left\{\prod_{l=N-m-1}^{(N-2)} I_l^{(\lambda,\lambda+\sigma)}\right\} P_N(\lambda_1, \cdots \lambda_{N-2}, \lambda, \lambda + \sigma). \tag{13}$$

The factor $(N-1)$ has the same meaning as before, while the binomial coefficient $\binom{N-2}{m}$ corresponds to the number of ways the m integration variables $\lambda_{N-m-1}, \lambda_{N-m}, \cdots \lambda_{N-2}$ can be picked out of the $(N-2)$ variables $\lambda_1, \cdots \lambda_{N-2}$. Combining Eqs. (3), (4), (5), (6), and (13), we have

$$p^{(m)}(x) = \lim_4 N(N-1) \times D^2(\lambda)\binom{N-2}{m}\left\{\prod_{k=1}^{(N-m-2)} O_k^{(\lambda,\lambda+\sigma)}\right\} \times \left\{\prod_{l=N-m-1}^{(N-2)} I_l^{(\lambda,\lambda+\sigma)}\right\} P_N(\lambda_1, \cdots \lambda_{N-2}, \lambda, \lambda + \sigma). \tag{14}$$

Here, $\lim_4$ is an abbreviation for the multiple limit operation $N \to \infty$, $\sigma \to 0$, $\lambda \to 0$, $x =$ finite. Equations (14) for $m = 0, 1, \cdots$, etc. define the hierarchy of spacing distributions $p^{(0)}(x)$, $p^{(1)}(x)$, $\cdots$, etc. explicitly in terms of the fundamental function $P_N(\lambda_1, \cdots \lambda_N)$.

III. INFINITE SERIES REPRESENTATION OF $p^{(m)}(x)$

It is possible to examine (14) for large N so as to arrive at an infinite series representation of the mth-order spacing distributions. The method used is a direct extension of a procedure given by Wigner[5] for the case $m = 0$. We introduce the following two functions:

$$J_k^{(m)} \equiv I_1 \cdots I_k O_{k+1} \cdots O_{N-m-2} \left\{ \prod_{l=N-m-1}^{(N-2)} I_l \right\} \times P_N(\lambda_1, \cdots \lambda_{N-2}, \lambda, \lambda + \sigma), \quad (15)$$

$$\mathfrak{L}_k^{(m)} \equiv \left\{ \prod_{j=1}^{k} I_j \right\} \left\{ \prod_{l=k+1}^{(N-m-2)} T_l \right\} \left\{ \prod_{s=N-m-1}^{(N-2)} I_s \right\} \times P_N(\lambda_1, \cdots \lambda_{N-2}, \lambda, \lambda + \sigma). \quad (16)$$

The label $(\lambda, \lambda + \sigma)$ is left off operators I_k and O_k for convenience. For large N, (14) can be rewritten as

$$p^{(m)}(x) = \lim_4 D^2(\lambda)[N^{m+2}/m!]J_0^{(m)}. \quad (17)$$

Due to property (10), we may expand products of T_k operators as follows:

$$\left\{ \prod_{k=1}^{s} T_k \right\} P_N(\lambda_1, \cdots \lambda_s, \lambda_{s+1}, \cdots \lambda_N) = \left\{ \sum_{j=0}^{s} \binom{s}{j} I_1 \cdots I_j O_{j+1} \cdots O_s \right\} \times P_N(\lambda_1, \cdots \lambda_s, \lambda_{s+1}, \cdots \lambda_N), \quad (18)$$

whereupon $\mathfrak{L}_k^{(m)}$ can be written

$$\mathfrak{L}_k^{(m)} = \sum_{j=0}^{N-2-(m+k)} \binom{N-2-(m+k)}{j} J_{j+k}^{(m)}. \quad (19)$$

For large N, the binonial coefficient in (19) can be replaced by $N^j/j!$. If, then, we multiply (19) by $(-N)^k/k!$ and sum both sides over k from zero to infinity, we find

$$J_0^{(m)} \approx \sum_{k=0}^{\infty} [(-N)^k/k!]\mathfrak{L}_k^{(m)}. \quad (20)$$

In light of (17),

$$p^{(m)}(x) = \lim_4 D^2(\lambda)[N^{m+2}/m!] \times \sum_{k=0}^{\infty} [(-N)^k/k!]\mathfrak{L}_k^{(m)}. \quad (21)$$

Expression (21) is the desired infinite series representation for $p^{(m)}(x)$. It is the direct extension of Wigner's Eq. (52).[2]

In general, one cannot evaluate *all* the $\mathfrak{L}_k^{(m)}$ for arbitrary $P_N(\lambda_1, \cdots \lambda_N)$, and (21) is generally useful only when a few terms of the series are needed. For the very special case of statistically independent levels, however, (21) can be evaluated explicitly. In this case,

$$P_N(\lambda_1, \cdots \lambda_N) = \prod_{i=1}^{N} P(\lambda_i), \quad (22)$$

$$p^{(m)}(x) = x^m/m! \exp(-x). \quad (23)$$

The mth-order functions (23) are the Poisson distributions, which are already known to represent the differential probabilities that m points randomly placed on a line occur in an interval of length x.[6] Here they arise as a special case of a general formalism in the context of energy-level spectra. We show in Sec. V that the Poisson distributions are the limiting forms for a well-defined physical problem, the superposition of many sequences of energy levels, each with a different symmetry character.

IV. SECOND-DERIVATIVE FORM

The purpose of this section is to obtain a generalization of Eqs. (98) and (106) of Dyson.[3] These equations express $p^{(0)}(x)$ and $p^{(1)}(x)$ as second derivatives of well-defined probability functions. Here, we are interested in similar equations for $p^{(m)}(x)$, $0 \leq m < \infty$. Our results will be seen to be useful in the treatment of superposed sequences in Sec. V.

We introduce a function $Q_N^{(m)}(s, t)$, which represents the probability of finding m levels in the closed interval $[s, t]$ and $(N - m)$ levels outside this interval. In terms of $P_N(\lambda_1, \cdots \lambda_N)$ we have

$$Q_N^{(m)}(s, t) = \binom{N}{m} \left\{ \prod_{j=1}^{m} I_j^{(s,t)} \right\} \left\{ \prod_{k=m+1}^{N} O_k^{(s,t)} \right\} \times P_N(\lambda_1, \cdots \lambda_N). \quad (24)$$

Clearly,

$$\sum_{m=0}^{N} Q_N^{(m)}(s, t) = 1. \quad (25)$$

Our ultimate aim is to relate (24) to (14). The desired result is obtained by taking the second partial derivative of (24) with respect to s and t and then letting $s \to \lambda \to 0$, $t \to \lambda + \sigma \to 0$. The differentiations can be carried out by expanding the I_j and O_k in series, as done in (18).[7] After some

[5] Wigner (Ref. 2) did not introduce integral operators, and performed the combinatorial analysis without their aid.

[6] P. M. Morse, *Thermal Physics* (W. A. Benjamin Company, Inc., New York, 1962), p. 91.

[7] The differentiation procedure is outlined in some detail by H. S. Leff, "Statistical Theory of Energy Level Spacing Distributions for Complex Spectra," Ph.D. thesis, State University of Iowa, 1963. Also printed as S.U.I. Physics Res. Rept. 63-23.

algebra, one finds

$$\frac{\partial^2}{\partial s\,\partial t}Q_N^{(m)}(s,t) = -\binom{N}{m}[(N-m)(N-m-1)\tilde{J}_0^{(m)} - 2(N-m)m\tilde{J}_0^{(m-1)} + m(m-1)\tilde{J}_0^{(m-2)}]. \quad (26)$$

Here $\tilde{J}_0^{(m)}$ corresponds to $J_0^{(m)}$ as defined by (15), but with $\lambda = s$, $\lambda + \sigma = t$. In the limit of large N, (31) becomes (replacing m by k)

$$(\partial^2/\partial s\,\partial t)Q_N^{(k)}(s,t) \approx -(N^k/k!) \times [N^2\tilde{J}_0^{(k)} - 2Nk\tilde{J}_0^{(k-1)} + k(k-1)\tilde{J}_0^{(k-2)}]. \quad (27)$$

This expression can be effectively iterated by operating on both sides with the summation operator $\sum_{l=0}^{m}\sum_{k=0}^{l}$, yielding

$$J_0^{(m)} \approx -[m!/N^{m+2}]\frac{\partial^2}{\partial s\,\partial t} \times \sum_{j=0}^{m}(j+1)Q_N^{(m-j)}(s,t). \quad (28)$$

Using (17), it follows that

$$p^{(m)}(x) = \lim_4 D^2(s)\frac{\partial^2}{\partial s\,\partial t} \times \sum_{j=0}^{m}(j+1)Q_N^{(m-j)}(s,t). \quad (29)$$

At this point it is convenient to transform variables, letting $t - s = u$, $t + s = v$, and

$$Q_N^{(k)}(s,t) \equiv M_N^{(k)}(u,v). \quad (30)$$

Equation (29) can then be rewritten as

$$p^{(m)}(x) = \lim_4 D^2(\lambda)\sum_{j=0}^{m}(j+1) \times \left[\frac{\partial^2}{\partial u^2}M_N^{(m-j)}(u,v) - \frac{\partial^2}{\partial v^2}M_N^{(m-j)}(u,v)\right]. \quad (31)$$

Clearly, (31) and (29) are equivalent. We have obtained the form (31) for the following reason. For actual complex spectra, it is expected that as $t \to s$, $Q_N^{(m)}(s,t)$ will be a function of $(t-s)$ only,[1] and that "suitable" ensembles will give rise to functions with this property. If, in fact, this situation occurs, the second term of (31) gives zero contribution, and we have[8]

$$p^{(m)}(x) = d^2/dx^2\sum_{j=0}^{m}(j+1)\psi^{(m-j)}(x) = d^2/dx^2\sum_{l=0}^{m}\sum_{k=0}^{l}\psi^{(k)}(x). \quad (32)$$

[8] Strictly speaking, this argument assumes that the limit and derivative operations commute.

$\psi^{(k)}(x)$ is defined by

$$\psi^{(k)}(x) = \lim_4 M_N^{(k)}(u,v). \quad (33)$$

The expressions (32) are much simpler than (29), but are applicable only if translational invariance exists in the proposed limit. It is expected that all ensembles of physical interest share this property.[9] For such ensembles, it is only necessary to find $Q_N^{(m)}(-s, +s)$, which simplifies calculations significantly.

Expressions (32) are equivalent statements of what we call the second-derivative form. For $m = 0$ and $m = 1$, they reduce to Dyson's Eqs. (98) and (106).[3,10] For all nonnegative integers m, (32) has the simple interpretation of being the second derivative of a sum of well-defined probabilities $S_l(x)$,

$$S_l(x) \equiv \sum_{k=0}^{l}\psi^{(k)}(x), \quad (34)$$

$S_l(x)$ = probability to have an interval of length u in a region of density x/u (centered about zero) containing no more than l levels.

It is understood that the zero point has been singled out for reasons of convenience. The derivation can be repeated without doing this. Physically, we expect the spacing distributions to be the same for all spectral regions which satisfy the requirements of complexity and slowly varying level density, and which contain a sufficient number of levels for statistical analyses. Hopefully, Eqs. (32) will prove useful in analytic calculations of $p^{(m)}(x)$ for $m \geq 2$, as they have for $m = 0$ and $m = 1$. For the case of statistically independent levels, the second-derivative form also yields the Poisson distributions, since for large N,

$$Q_N^{(m)}(s,t) = \binom{N}{m}\left[\int_s^t P(\lambda)\,d\lambda\right]^m\left[1 - \int_s^t P(s)\,ds\right]^{N-m}, \quad (35)$$

which yields

[9] It can be shown that $Q_N{}^{(m)}(s,t)$ is translationally invariant for finite N, if (a) $P_N(\lambda_1+\Delta, \lambda_2+\Delta, \cdots \lambda_N+\Delta) = P_N(\lambda_1, \lambda_2, \cdots \lambda_N)$ for all real Δ, (b) $b = -a$, and (c) $P_N(\lambda_1 \cdots \lambda_i + 2b, \cdots \lambda_N) = P_N(\lambda_1, \cdots \lambda_i, \cdots \lambda_N)$. These three properties are satisfied by Dyson's ensembles. Other ensembles, such as the Gaussian ensemble, evidently possess the translational-invariance property only in the full limit $s \to \lambda \to 0$, $t \to \lambda + \sigma \to 0$, $N \to \infty$, x being finite.

[10] It was learned after the completion of this work that an alternate derivation of (32) has been given by P. B. Kahn, Proceedings of the Symposium on Statistical Aspects of Energy Levels of Complex Systems, State University of New York, Stony Brook, New York, 1963. However, this derivation does not explicitly utilize $P_N(\lambda_1, \cdots \lambda_N)$, and does not make use of the multiple limiting processes. A similar derivation for $m = 0$ can also be found in Ref. 1. These derivations are considered to be less basic than the present work.

$$\psi^{(m)}(x) = x^m/m! \exp(-x). \tag{36}$$

Application of the second-derivative form to (36) shows that, for the case of statistically independent levels, $\psi^{(m)}(x)$ and $p^{(m)}(x)$ are identical. Therefore this result is identical to (23).

V. APPLICATION TO SUPERPOSITION OF INDEPENDENT SEQUENCES OF ENERGY LEVELS

The basic formalisms of Secs. III and IV presumably allow us to calculate spacing distributions for any type of complex spectra, if the functions $P_N(\lambda_1, \cdots \lambda_N)$ are known for the particular spectra of interest. In matrix ensemble theories, one applies the function $P_N(\lambda_1, \cdots \lambda_N)$ to sequences of levels of the same symmetry character.[11] Levels of different symmetry character are statistically independent,[1,3,12] and each particular sequence presumably has the same spacing distributions $p^{(0)}(x)$, $p^{(1)}(x)$, $\cdots$, etc. Here, we use these facts to develop a formalism for finding the mth-order spacing distributions for spectra containing two or more sequences, each of a different symmetry character. This problem was originally investigated in detail by Rosenzweig and Porter for $m = 0$.[12] The present work is valid for integers $m \geq 0$. The method is different from that of R–P, and is believed to be more transparent.[13]

It is not necessary to specify $P_N(\lambda_1, \cdots \lambda_N)$ for spectra with mixed sequences. One need only work with the functions $\psi^{(m)}(x)$ and the second-derivative form (32). Consider the superposition of n independent sequences, each sequence labeled with the index i: $i = 1, 2, \cdots, n$. As before, we are interested in a region of the spectrum where the level density is essentially constant. Let the mean level density for sequence i be denoted by D_i^{-1}. The mean density for the mixture of sequences is just

$$D^{-1} = \sum_{i=1}^{n} D_i^{-1}. \tag{37}$$

The fractional level densities a_i are defined by

$$a_i = D/D_i. \tag{38}$$

Clearly, from (37), the sum of the fractional densities is unity. In this section, the symbol x will refer to spacings relative to the total mean nearest neighbor spacing D. Therefore, the mth-order spacing distributions for sequence i are determined via (32) by the functions $\psi^{(0)}(a_i x)$, $\psi^{(1)}(a_i x)$, $\cdots$ $\psi^{(m)}(a_i x)$. The object here is to explicitly use the statistical independence of the sequences to determine the set of functions $\Psi_n^{(m)}(x)$ which replace the $\psi^{(m)}(x)$ in (32) when two or more sequences are present. By definition, $\Psi_1^{(m)}(x) = \psi^{(m)}(x)$. The spacing distributions for the mixture will be denoted by $P_n^{(m)}(x)$. The second-derivative form is just

$$P_n^{(m)}(x) = \frac{d^2}{dx^2} \sum_{l=0}^{m} \sum_{k=0}^{l} \Psi_n^{(k)}(x). \tag{39}$$

One can verify by a simple probability argument that

$$\Psi_n^{(0)}(x) = \prod_{i=1}^{n} \psi^{(0)}(a_i x), \tag{40}$$

$$\Psi_n^{(1)}(x) = \sum_{i=1}^{n} \psi^{(1)}(a_i x) \left\{ \prod_{j \neq i}^{n} \psi^{(0)}(a_i x) \right\}, \tag{41}$$

$$\Psi_n^{(2)}(x) = \tfrac{1}{2} \sum_{i \neq j}^{n} \psi^{(1)}(a_i x) \psi^{(1)}(a_i x) \left\{ \prod_{k \neq i,j}^{n} \psi^{(0)}(a_k x) \right\} + \sum_{i=1}^{n} \psi^{(2)}(a_i x) \{ \prod_{j \neq i} \psi^{(0)}(a_i x) \}, \tag{42}$$

$$\Psi_n^{(3)} = \tfrac{1}{6} \sum_{i \neq j \neq k}^{n} \psi^{(1)}(a_i x) \times \psi^{(1)}(a_i x) \psi^{(1)}(a_k x) \{ \prod_{l \neq i,j,k} \psi^{(0)}(a_l x) \} + \tfrac{1}{2} \sum_{i \neq j}^{n} \psi^{(2)}(a_i x) \psi^{(1)}(a_i x) \left\{ \prod_{k \neq i,j}^{n} \psi^{(0)}(a_k x) \right\} + \sum_{i=1}^{n} \psi^{(3)}(a_i x) \left\{ \prod_{j \neq i}^{n} \psi^{(0)}(a_i x) \right\}, \tag{43}$$

$\cdots$, etc. The prescription given by (39)–(43) and their natural extensions allows one to find $P_n^{(m)}(x)$ for arbitrary n and m, given

$$a_1, \cdots a_n, \psi^{(0)}(x), \cdots \psi^{(m)}(x).$$

Using several very general properties of the functions $\psi^{(m)}(x)$, one can obtain the behavior of $P_n^{(m)}(x)$ for (i) arbitrary n and m as $x \to 0$, and (ii) arbitrary m and x as $n \to \infty$. This is now done.

From the definitions of $Q_N^{(m)}(s, t)$ and $\psi^{(m)}(x)$, it is clear that

$$\psi^{(m)}(0) = \delta_{m0}. \tag{44}$$

Furthermore, $\psi^{(m)}(x)$ must approach zero for all finite m as $x \to \infty$, which implies that

$$(d/dx)\psi^{(m)}(x)|_{x=\infty} = 0. \tag{45}$$

We can obtain $d/dx\psi^{(m)}(x)\,|_{x=0}$ by inverting (32) to express each $d^2/dx^2\psi^{(m)}(x)$ as a linear combination

[11] A sequence of a given symmetry character in nuclear spectra, for example, means a set of levels from a given isotope with a given spin and parity.

[12] N. Rosenzweig and C. E. Porter, Phys. Rev. 120, 1698 (1960).

[13] The author is indebted to N. Rosenzweig for suggesting this procedure.

of the $p^{(i)}(x)$, $i \leq m$. The inversion of (32) is accomplished by treating each distribution $p^{(i)}(x)$ and each function $(d^2/dx^2)\psi^{(i)}(x)$ as components of vectors $\mathbf{p}$ and $\mathbf{q}$, respectively. The dimensionality of these vectors is infinite, but we only need concern ourselves with finite-dimensional subspaces. (32) takes on the form $\mathbf{p} = A\mathbf{q}$, where

$$\mathbf{p} = (p^{(0)}, p^{(1)}, \cdots p^{(m)}, \cdots),$$

$$\mathbf{q} = (d^2/dx^2\psi^{(0)}, d^2/dx^2\psi^{(1)}, \cdots d^2/dx^2\psi^{(m)}, \cdots),$$

$$A = \begin{bmatrix} 1 & 0 & 0 & & \cdots \\ 2 & 1 & 0 & 0 & \cdots \\ 3 & 2 & 1 & 0 & 0 \cdots \\ 4 & 3 & 2 & 1 & 0 \cdots \\ 5 & & & & \cdots \\ \vdots & & & & \end{bmatrix}.$$

By induction, one finds that the inverse of A has elements

$$(A^{-1})_{ij} = \delta_{i,j} - 2\delta_{i,j+1} + \delta_{i,j+2}; \qquad i, j = 0, 1, 2, \cdots.$$

Therefore, we may write

$$(d^2/dx^2)\psi^{(m)}(x) = p^{(m)}(x) - 2p^{(m-1)}(x) + p^{(m-2)}(x) \tag{46}$$

for all m, if we define $p^{(m)}(x)$ to be zero for $m < 0$. Integrating (46) and applying (45), we find

$$(d/dx)\psi^{(m)}(x)|_{x=0} = -\delta_{m0} + \delta_{m1}. \tag{47}$$

Furthermore, since $p^{(m)}(0)$ necessarily vanishes for $m > 0$,

$$(d^2/dx^2)\psi^{(m)}(x)|_{x=0} = (\delta_{m0} - 2\delta_{m1} + \delta_{m2})p^{(0)}(0). \tag{48}$$

From (44), (47) and (48) we deduce

$$\psi^{(m)}(x) = \delta_{m0} + (\delta_{m1} - \delta_{m0})x + \tfrac{1}{2}p^{(0)}(0)(\delta_{m0} - 2\delta_{m1} + \delta_{m2})x^2 + b_m x^3 + \cdots. \tag{49}$$

For the cases of physical interest, where matrix ensembles are used to obtain $P_N(\lambda_1, \cdots \lambda_N)$, it is well-known that the probability to find two (or more) equal eigenvalues is zero.[1] Therefore, for our purposes, $p^{(0)}(0) = 0$, and the functions $\psi^{(m)}(x)$ do not contain terms proportional to x^2. The coefficients b_m can be expressed, using (46), as

$$b_m = \tfrac{1}{6}[(d/dx)p^{(m)} - 2(d/dx)p^{(m-1)} + (d/dx)p^{(m-2)}]_{x=0}. \tag{50}$$

For small x, using (40)–(43), we find

$$\Psi_n^{(0)}(x) = 1 - x + \sum_{i<j}^{n} a_i a_j x^2 - \left[\sum_{i<j<k}^{n} a_i a_j a_k - \sum_{i=1}^{n} b_0 a_i^3\right]x^3 + \cdots, \tag{51}$$

$$\Psi_n^{(1)}(x) = x - 2\sum_{i<j}^{n} a_i a_j x^2 + \left[3\sum_{i<j<k}^{n} a_i a_j a_k + \sum_{i=1}^{n} b_1 a_i^3\right]x^3 + \cdots, \tag{52}$$

$$\Psi_n^{(2)}(x) = \sum_{i<j}^{n} a_i a_j x^2 + \left[\sum_{i=1}^{n} b_2 a_i^3 - 3\sum_{i<j<k}^{n} a_i a_j a_k\right]x^3 + \cdots, \tag{53}$$

$$\Psi_n^{(3)}(x) = \left[\sum_{i<j<k}^{n} a_i a_j a_k + \sum_{i=1}^{n} b_3 a_i^3\right]x^3 + \cdots, \tag{54}$$

$$\Psi_n^{(j)}(x) = \sum_{i=1}^{n} b_j a_i^3 x^3 + \cdots \quad \text{for} \quad j \geq 4. \tag{55}$$

Therefore, to *first order in* x,

$$P_n^{(m)}(x) = 2\left[\sum_{i<j}^{n} a_i a_j\right]\delta_{m0} - 6x\left[\sum_{i<j<k}^{n} a_i a_j a_k\right][\delta_{m0} - \delta_{m1}] + 6x\left[\sum_{i=1}^{n} a_i^3\right]\sum_{j=0}^{m}(j+1)b_{m-j}. \tag{56}$$

From (56) one easily obtains $P_n^{(m)}(0)$ and

$$d/dxP_n^{(m)}(x)|_{x=0}.$$

For $m = 0$, the results are in exact agreement with Rosenzweig and Porter.[12] For $m > 0$, the results are new and there exists no basis of comparison at present.

If a large number n of sequences are superposed, the fractional densities $a_1, \cdots a_n$ will generally be very small. In the extreme case, as $n \to \infty$, each a_i will almost always approach zero. To analyze the situation, we assume that all fractional densities are equal; i.e., $a_i = 1/n$ for $i = 1, 2, \cdots, n$ (this restriction can be relaxed to some extent). As n becomes large the functions $\Psi_n^{(m)}(x)$ behave in the following way:

$$\Psi_n^{(m)}(x) \to \binom{n}{m}(x/n)^m(1 - x/n)^{n-m} \underset{n\to\infty}{\to} x^m/m!\exp(-x). \tag{57}$$

Using (39), the spacing distributions are the Poisson

distributions which we encountered in connection with statistically independent eigenvalues. Here, as we superpose an infinite number of sequences, we essentially force an infinite number of uncorrelated levels into each region of the spectrum. The condition $a_i = 1/n$ can be relaxed to allow a *finite* number j of sequences with different fractional densities (each of which approaches zero) and $(n - j)$ sequences with equal fractional densities, with the same result. Evidently, if the sequences are such that one a_i doesn't approach zero, the levels of this sequence "compete" with the uncorrelated levels with sufficient strength to prevent the Poisson distributions from occurring.

Although we have only applied the present formalism for superposed sequences to special cases, it is amenable to numerical calculations of $P_n^{(m)}(x)$ for extended ranges of x, provided that $\psi^{(0)}(x)$, $\psi^{(1)}(x)$, $\cdots$, $\psi^{(m)}(x)$ are known. For $m = 0$, Rosenzweig and Porter accomplished this with their formalism, using an approximation for $\psi^{(0)}(x)$.[12] Similar calculations for $m = 1$ have been made recently with the present formalism,[14] and the higher-order cases will undoubtedly be explored in the future.[15]

ACKNOWLEDGMENTS

I would like to express my gratitude to Professor Max Dresden for his encouragement and guidance during the course of this work. I am also indebted to Dr. Norbert Rosenzweig for many informative discussions at Argonne National Laboratory. Finally, I should like to thank Daniel J. Tambasco for useful conversations concerning several technical points.

[14] H. S. Leff, Bull. Am. Phys. Soc. **8**, 31 (1963).

[15] Recently, numerical computations of $p^{(m)}(x)$ for $m = 0, 1, 2, \cdots 8$ have been made using a 10×10 matrix ensemble. See C. E. Porter, J. Math. Phys. **4**, 1039 (1963). This should provide impetus for analytical calculations of these functions in the infinite-dimensional limit.

Class of Ensembles in the Statistical Theory of Energy-Level Spectra*†

HARVEY S. LEFF‡§
State University of Iowa, Iowa City, Iowa
(Received 17 January 1964)

The investigation of a large class of ensembles in the statistical theory of energy-level spectra is initiated. Each member of this class is characterized by a joint probability density for N consecutive eigenvalues of the form

$$P_{N\beta}(\lambda_1, \cdots \lambda_N) = \Omega_{N\beta}^{-1}\left\{\prod_{i=1}^{N} f(\lambda_i)\right\} \prod_{k<l} |\lambda_k - \lambda_l|^{\beta},$$

where $a \leq \lambda_i \leq b$, and β may be 1, 2, or 4. Formal calculations of the nearest-neighbor spacing distribution and the level density are made for $\beta = 2$. Results are in terms of asymptotic properties of orthogonal polynomials. It is conjectured that spacing distributions are relatively insensitive to the function $f(\lambda)$ and the interval $[a, b]$. When $f(\lambda) = 1$ and $b = -a = 1$, the resulting (Legendre) ensemble has the same spacing distribution as the Gaussian and Dyson ensembles. The level density is concave upward and rapidly increasing for $\lambda \geq 0$, qualitatively resembling actual nuclear and atomic densities. This feature is not present in previously investigated ensembles. Certain invariant matrix ensembles introduced by Dyson, which are of the above type, have the same level density and nearest-neighbor spacing distribution as the Legendre ensemble.

I. INTRODUCTION

THEORIES involving ensembles of matrices have been used extensively in attempts to explain various statistical properties of atomic and nuclear spectra.[1] In particular, analytic calculations of spacing distributions and level densities have been made for both the Gaussian and Dyson ensembles.[2-4] These have shown that while the latter two ensembles give rise to the same nearest-neighbor spacing distributions, their level densities are markedly different. The essence of the present paper is to initiate the analytical investigation of still other ensembles in order to observe the sensitivity of the level density and spacing distributions on the specific ensemble which is used.

The Gaussian ensembles, which were first proposed by Wigner,[5] are characterized by the joint probability density

$$G_{N\beta}(\lambda_1, \cdots \lambda_N) = K_{N\beta}^{-1} \exp\left(-\sum_{i=1}^{N} \lambda_i^2/\sigma^2\right) \prod_{k<l} |\lambda_k - \lambda_l|^{\beta} \qquad (1)$$

for N consecutive eigenvalues $\lambda_1, \lambda_2, \cdots \lambda_N$ (not necessarily ordered with respect to labeling). $K_{N\beta}$ is a normalization constant, σ is a free parameter and $-\infty < \lambda_i < +\infty$ for $i = 1, 2, \cdots, N$. The parameter β may take on the values 1, 2, and 4. The case $\beta = 1$ corresponds to physical situations where time-reversal invariance exists, and where either the total angular momentum is conserved or the total number of half-integer spin particles is even. $\beta = 4$ corresponds to time-reversal invariant systems with an odd number of half-integer spin particles and nonconserved total angular momentum. $\beta = 2$ corresponds to systems which are *not* time-reversal invariant. The Dyson ensembles are characterized by the joint probability density

$$W_{N\beta}(\theta_1, \cdots \theta_N) = C_{N\beta}^{-1} \prod_{i<j} |\exp(i\theta_i) - \exp(i\theta_j)|^{\beta}. \qquad (2)$$

Here the θ_i are identified with the energy eigenvalues although they originate in the eigenvalues $\exp(i\theta_i)$ of unitary matrices. $C_{N\beta}$ is a normalization constant, β has the same meaning as above, and $-\pi \leq \theta_i \leq \pi$.

---

* This work is based on portions of a Ph.D. thesis submitted before the faculty of the State University of Iowa, July, 1963.

† This work was supported in part by the Air Force Office of Scientific Research.

‡ State University of Iowa Predoctoral Fellow 1962-63.

§ Present address: Department of Physics, Case Institute of Technology, Cleveland, Ohio.

[1] For a review of matrix ensembles and an extensive bibliography, see N. Rosenzweig, "Statistical Mechanics of Equally Likely Quantum Systems" in *Statistical Physics* (W. A. Benjamin, Company, Inc., New York, 1963), pp. 91-158.

[2] E. P. Wigner, "Proceedings Gatlinburg Conference on Neutron Physics by Time of Flight," ORNL 2309, 67 (1956), (Unpublished). Also, see *Proceedings of the Fourth Canadian Mathematics Congress* (Toronto University Press, Toronto, Canada, 1959), p. 174.

[3] M. L. Mehta, Nucl. Phys. **18**, 395 (1960); M. L. Mehta and M. Gaudin, *ibid.* **18**, 420 (1960); M. Gaudin, *ibid.* **25**, 447 (1961).

[4] F. J. Dyson, J. Math. Phys. **3**, 140, 157, 166 (1962).

[5] To be precise, Wigner proposed only the case $\beta = 1$. See footnote 2.

In the present paper we propose the investigation of a class of ensembles which have joint probability densities of the form

$$P_{N\beta}(\lambda_1, \cdots \lambda_N) = \Omega_{N\beta}^{-1}\left\{\prod_{i=1}^{N} f(\lambda_i)\right\} \prod_{k<l} |\lambda_k - \lambda_l|^\beta. \quad (3)$$

Here, $a \leq \lambda_i \leq b$, where a and b need not be specified. $\Omega_{N\beta}$ is the normalization constant and β is 1, 2, or 4, as before. The functions $f(\lambda)$ are required to be integrable and to possess finite moments over $[a, b]$. The purpose of examining this particular class of ensembles is simply to investigate the dependence of the spacing distributions and level density on the function $f(\lambda)$ and the interval $[a, b]$. It is well to point out that while (1) and (2) are obtained from invariant matrix ensembles, the matrix ensembles giving rise to (3) are unknown in general.[6]

The choice of the functional form (3) is prompted by several considerations which shall now be enumerated:

1. The factors $\prod_{k<l} |\lambda_k - \lambda_l|^\beta$ are a general consequence of Hermitian matrix ensemble considerations[7] and therefore must be included.

2. In general, $P_{N\beta}(\lambda_1, \cdots \lambda_N)$ will contain a factor $F(\lambda_1, \cdots \lambda_N)$ rather than $\prod_i f(\lambda_i)$. However, ensembles of this very general type *cannot* be treated analytically via the Mehta–Gaudin methods of determinantal integration. The special cases for which $F(\lambda_1, \cdots \lambda_N) = \prod_i f(\lambda_i)$ are the only known ones which are amenable to analytical investigation.

3. A completely different way to arrive at (3) is to relate Hermitian matrix ensembles to the Dyson ensembles given by (2). The latter arise from considerations of ensembles of unitary matrices S which are *unspecified* functions of the Hermitian Hamiltonian matrices H. As Dyson states, no functional form $S(H)$ is specified since such a function cannot be correct globally.[4] However, spacing distribution calculations are dependent on local rather than global properties.[8] Therefore, one might attempt a specification of $S(H)$ which allows one to obtain a formal relationship between the distribution of Dyson's θ_i variables and the eigenvalues λ_i. When this is done, one finds that for small λ_i, Dyson's distribution (2) is tantamount to a distribution of the λ_i which is of the form[9]

$$\left\{\prod_{i=1}^{N} f_{N\beta}(\lambda_i)\right\} \prod_{k<l} |\lambda_k - \lambda_l|^\beta.$$

The consideration of small λ_i arises because spacing distributions are ordinarily calculated in shrinkingly small intervals centered about $\lambda = 0$. $f_{N\beta}(\lambda_i)$ is dependent on the particular functional form $S(H)$ as well as on N and β. The N and β dependence of $f_{N\beta}(\lambda_i)$ makes the above expression untractable, in general. However, the form is *highly suggestive* of (3). Unless $f_{N\beta}(\lambda_i)$ exhibits pathological behavior near the origin, it is expected that the replacement of some function $f(\lambda_i)$ for $f_{N\beta}(\lambda_i)$ will not affect the spacing distributions. This means of arriving at (3) is admittedly crude and largely heuristic, but is retained for the following reasons. Dyson's ensembles allow (in principle) all possible functional forms $S(H)$. The overwhelming majority of the ensemble members evidently have the same nearest-neighbor spacing distributions. Since the factors $\prod_{k<l} |\lambda_k - \lambda_l|^\beta$ are present for all choices of $S(H)$, they are suspected of being mainly responsible for these "universal" distributions. This argument minimizes the importance of the functions $f_{N\beta}(\lambda_i)$ and indeed could be carried to the point of replacing these functions by constants. (This is done in fact in Sec. III.) Here, we go only to the point of replacing the $f_{N\beta}(\lambda_i)$ by $f(\lambda_i)$ and ask the question: How does changing $f(\lambda)$ affect the spacing distributions and the level density? Throughout, we do not consider the interval $[a, b]$ to be of importance in spacing distribution calculations because the latter are performed locally about $\lambda = 0$. We need only require that $a < 0 < b$.

4. Finally, we point out that the choice of (3) on the above grounds does not allow a conclusion that the resulting global level density will be at all meaningful. Our attitude here is that *if* $f(\lambda)$ and $[a, b]$ do not significantly affect the spacing distributions, we can search for particular ensembles which yield qualitatively correct *global* level densities. The Gaussian and Dyson ensembles are both lacking in this respect.

The special case $\beta = 2$ allows a particularly simple formal investigation of the nearest-neighbor spacing distribution and level density for (3). This investigation constitutes Sec. II. We find that all results are in terms of certain asymptotic properties of orthogonal polynomials which are uniquely defined with respect to the weight function $f(\lambda)$ and interval $[a, b]$.

It is tempting to begin the proposed investigation by picking specific ensembles which give rise to

---

[6] Some specific cases are known; for example, (1). Others are discussed in Sec. IV.

[7] H. S. Leff, "Statistical Theory of Energy Level Spacing Distributions for Complex Spectra," Ph.D. Thesis, State University of Iowa, 1963, and S. U. I. Physics Res. Rept. 63-23. Here, the class of ensembles (3) are called "generalized."

[8] H. S. Leff, J. Math. Phys. 5, 756 (1964) (previous paper).

[9] The details are contained in Ref. 7.

well-known orthogonal polynomials. This allows the exploitation of known asymptotic properties of these functions without unnecessarily detailed mathematical analyses. Since the classical polynomials are the most well known, these should undoubtedly be examined first. The Hermite polynomials arise in the study of the Gaussian ensembles, which have already been investigated extensively. The simplest of the remaining classical polynomials are the Legendre functions. These appear if we choose $f(\lambda) = 1$ and $b = -a = 1$. With this choice the ensemble represented by (3) is given the name Legendre. The Legendre ensemble is investigated in Sec. III.

Some results for Jacobi polynomials[10] are contained in Sec. IV, which consists of a discussion and evaluation of our main results. The Laguerre polynomials are not treated.[11]

II. FORMAL CALCULATIONS FOR $\beta = 2$

In this section we follow the general methods of Mehta and Gaudin[3] in formal calculations of the level density and nearest-neighbor spacing distribution for ensembles characterized by (3), with $\beta = 2$. The specific Mehta–Gaudin techniques which we use consist of three main parts. First, the product $\prod_{k<l} |\lambda_k - \lambda_l|^2$ can be written as a determinant which is named after Vandermonde,[12]

$$\prod_{k<l} (\lambda_l - \lambda_k) = |\lambda_1^i, \lambda_2^i, \cdots \lambda_N^i|. \tag{4}$$

The right-hand side is a convenient way to depict the determinant in terms of its $(j + 1)$th typical row, $j = 0, 1, \cdots (N - 1)$. Second, each power of λ_i can be expressed as a linear combination of suitably chosen orthogonal polynomials. Third, the necessary integrals can be carried out using an extension of a theorem due to Gram.[13]

Before proceeding to the level density calculation, we recapitulate several well-established facts in the theory of orthogonal polynomials.[14] Let e_k denote the kth moment with respect to the weight function $f(x)$ over $a \leq x \leq b$. It is possible to construct nth-degree polynomials $p_n(x)$ which are orthogonal with respect to $f(x)$ over the interval $[a, b]$. This construction, which is not elaborated upon here (see Ref. 14), involves only $e_0, \cdots e_{2n-1}$ and k_n, the coefficient of x^n in the polynomial $p_n(x)$. The $p_n(x)$ are unique within a normalization constant, and obey the orthogonality condition

$$\int_a^b p_m(x)p_n(x)f(x)dx = h_n\delta_{mn}, \tag{5}$$

h_n being expressible in terms of the e_k and k_n. From the foregoing, one can derive a general recurrence formula for the $p_n(x)$, from which the so-called Christoffel–Darboux formula results,

$$\sum_{i=0}^{(n-1)} h_i^{-1}p_i(x)p_i(y) = [k_{n-1}/(k_nh_{n-1})](x - y)^{-1} \times [p_n(x)p_{n-1}(y) - p_{n-1}(x)p_n(y)]. \tag{6}$$

Finally, we point out that one may write

$$p_n(x) = \sum_{j=0}^{n} d_j^{(n)}x^j \tag{7}$$

where $d_n^{(n)} = k_n$, and the $d_j^{(n)}$ for $0 \leq j \leq (n - 1)$ are obtainable in terms of the moments e_k.

In order to apply the second Gaudin–Mehta technique to (4), we multiply row one by $d_0^{(0)}$, row two by $d_1^{(1)}, \cdots$, row N by $d_{N-1}^{(N-1)}$, and divide the entire determinant by $\prod_{k=0}^{N-1} d_k^{(k)}$. Multiply (mentally) row one by $(d_0^{(1)}/d_0^{(0)})$ and add it to row two. The resulting first and second rows have the typical elements $p_0(\lambda_i)$ and $p_1(\lambda_i)$, respectively. Continuing similar manipulations with rows 3 to N, we find that (4) can be replaced by

$$\prod_{k<l} (\lambda_l - \lambda_k) = \left\{\prod_{i=1}^{N} k_{i-1}^{-1}\right\} |p_i(\lambda_1), p_i(\lambda_2), \cdots p_i(\lambda_N)|. \tag{8}$$

Having obtained (8), we now proceed to calculate the density of levels, which is equal to N multiplied by $P(\lambda)$—the probability density for finding a level at λ. Clearly, we may write

$$NP(\lambda) = \sum_{k=1}^{N} w_k(\lambda), \tag{9}$$

where $w_k(\lambda)$ is the probability density for finding $\lambda_k = \lambda$.

$$w_k(\lambda) = N! \int_{a\leq\lambda_1\leq\cdots\lambda_N\leq b} \cdots \int P_{N2}(\lambda_1, \cdots \lambda_N) \times \delta(\lambda_k - \lambda)\, d\lambda_1 \cdots d\lambda_N. \tag{10}$$

[10] After completion of this work, investigation of Jacobi and Laguerre ensembles were made by P. B. Kahn, C. E. Porter, and Y. C. Tang (to be published). A preliminary report is in the Proceedings of the Eastern Conference in Theoretical Physics, University of North Carolina (1963) (unpublished).

[11] Our previous requirement $a < 0 < b$ does not allow the use of the Laguerre polynomials without some modification.

[12] See R. Bellman, *Introduction to Matrix Analysis*, (McGraw-Hill Book Company, Inc., New York, 1960), p. 186.

[13] Ref. 1, pp. 149, 150 contains a discussion of this.

[14] A. Erdélyi, W. Magnus, F. Oberhettinger, F. G. Tricomi, in the Bateman manuscript project, *Higher Transcendental Functions*, edited by A. Erdélyi (McGraw-Hill Book Company, Inc., New York, 1953), Vol. 2, pp. 153–227.

Combining (3) and (8) one notes that the integrand of (9) is symmetric under the interchange of any two of the λ_k. Therefore, the ordering of the integrations can be removed (dividing by $N!$). Expansion of the squared determinant allows the direct application of (5) to yield

$$NP(\lambda) = N!\,\Omega_{N2}^{-1} f(\lambda) \times \left\{\prod_{i=0}^{(N-1)} (h_i k_i^{-2})\right\} \sum_{k=0}^{(N-1)} h_k^{-1} p_k^2(\lambda). \quad (11)$$

The normalization constant Ω_{N2} is easily evaluated by demanding that the integral of $NP(\lambda)$ over $[a, b]$ be equal to N, the total number of levels. This yields

$$\Omega_{N2} = N! \prod_{i=0}^{(N-1)} (h_i k_i^{-2}), \quad (12)$$

from which it follows that

$$NP(\lambda) = (k_{N-1}/k_N h_{N-1}) f(\lambda) \times [p_{N-1}(\lambda) p'_N(\lambda) - p_N(\lambda) p'_{N-1}(\lambda)]. \quad (13)$$

$p'_N(\lambda)$ represents the derivative of $p_N(\lambda)$ with respect to λ. In order to find $NP(\lambda)$ for large N, the asymptotic forms for $p_N(\lambda)$ and $p'_N(\lambda)$ usually must be known. Specific application of (13) is made in Sec. III.

We now carry out a formal calculation of the nearest-neighbor spacing distributions resulting from (3) for $\beta = 2$. We use the known fact[15] that the nearest-neighbor spacing distribution $p^{(0)}(x)$ is given by

$$p^{(0)}(x) = d^2/dx^2 \psi^{(0)}(x). \quad (14)$$

Here, $\psi^{(0)}(x)$ is the probability to find an interval of length u in a region of density x/u (centered about $\lambda = 0$) containing *no* levels. x, therefore, represents the nearest-neighbor spacing divided by the mean nearest-neighbor spacing at $\lambda = 0$. In order to obtain $\psi^{(0)}(x)$ from (3) we examine

$$Q_N^{(0)}(-\sigma, \sigma) = \int_a^b \cdots \int_a^b \left\{\prod_{k=1}^{N} \Phi_\sigma(\lambda_k)\right\} \times P_{N2}(\lambda_1, \cdots \lambda_N)\, d\lambda_1 \cdots d\lambda_N \quad (15)$$

in the limit as $N \to \infty$, $\sigma \to 0$, and $\sigma NP(0)$ is fixed. $\Phi_\sigma(\lambda)$ is defined to be zero for $-\sigma \leq \lambda \leq \sigma$ and is equal to unity otherwise. Obviously, $\Phi_\sigma^2(\lambda) = \Phi_\sigma(\lambda)$. Using (8) and expanding the squared determinant (this is the extension of Gram's theorem mentioned previously), we find that

$$Q_N^{(0)}(-\sigma, \sigma) = 1/N! \int_a^b \cdots \int_a^b |u_i(\lambda_1), \times u_i(\lambda_2), \cdots u_i(\lambda_N)|^2\, d\lambda_1 \cdots d\lambda_N. \quad (16)$$

A further technique, separate from the three already mentioned, is due to Gaudin.[16] We notice that (16) is the Fredholm determinant of the integral equation

$$\epsilon(N)\varphi(z) = \int_{-\sigma}^{\sigma} K_N(z, y)\varphi(y)\, dy, \quad (17)$$

if $\epsilon(N)$ is set equal to unity and the kernel is

$$K_N(z, y) = \sum_{i=0}^{N-1} f^{\frac{1}{2}}(z) f^{\frac{1}{2}}(y) h_i^{-1} p_i(z) p_i(y). \quad (18)$$

Therefore, we may write

$$Q_N^{(0)}(-\sigma, \sigma) = \prod_{i=1}^{N} [1 - \epsilon_i(N)]. \quad (19)$$

The limit of $Q_N^{(0)}(-\sigma, \sigma)$ which is necessary to obtain $\psi^{(0)}(x)$ is found by examining the corresponding limit of the discrete eigenvalues $\epsilon_i(N)$. Toward this end, we examine (17) in the proposed limit.

Let $u = zNP(0)$ and $v = yNP(0)$. The problem is now transformed to the analysis of

$$\epsilon(N)\varphi(u/NP(0)) = \int_{-x/2}^{x/2} (NP(0))^{-1} \times K_N(u/NP(0), v/NP(0))\varphi(v/NP(0))\, dv \quad (20)$$

in the single limit $N \to \infty$. If we let $\epsilon = \mathrm{Lim}_{N\to\infty}\, \epsilon(N)$, $w(u) = \mathrm{Lim}_{N\to\infty}\, \varphi(u/NP(0))$, and

$$M(u, v) = \lim_{N\to\infty} [NP(0)]^{-1} K_N(u/NP(0), v/NP(0)),$$

we must then find the eigenvalues ϵ_i of

$$\epsilon w(u) = \int_{-x/2}^{x/2} M(u, v) w(v)\, dv. \quad (21)$$

Finally,

$$\psi^{(0)}(x) = \prod_{i=1}^{\infty} (1 - \epsilon_i). \quad (22)$$

$p^{(0)}(x)$ is obtained by applying (14) to (22). For the Dyson and Gaussian ensembles, (21) turns out to be the integral equation for the prolate spheroidal functions. For this case, the infinite product (22) is rapidly converging and $\psi^{(0)}(x)$ has been obtained numerically by Gaudin[16] for $\beta = 1$ and by Kahn[17] for $\beta = 2$.

[15] For a proof of this and its generalization for the mth-order spacing distributions see Ref. 8. Also see Ref. 1 and the remarks by P. B. Kahn, Proceedings of the Symposium on Statistical Properties of Atomic and Nuclear Spectra, State University of New York at Stony Brook, (1963).

[16] See the third reference of footnote 3.

[17] P. B. Kahn, Nucl. Phys. **41**, 159 (1963).

III. THE LEGENDRE ENSEMBLE

We now apply the results of the preceding section to the specific ensemble for which $f(\lambda) = 1$ for all $-1 \leq \lambda \leq 1$, and $b = -a = 1$. The Legendre polynomials are *uniquely* defined with respect to this weight function and interval. This accounts for the name Legendre. The following relations hold for the Legendre functions $P_j(x)$:

$$h_j = (j + \tfrac{1}{2})^{-1}, \quad k_j = (2j)!/2^j(j!)^2. \tag{23}$$

To obtain the asymptotic form of $NP(\lambda)$ we use the following formulas,[18] which hold for $0 < \theta < \pi$ in the limit of large N:

$$P_N(\cos\theta) = [2/(N\pi \sin\theta)]^{\frac{1}{2}} \times \cos[(N + \tfrac{1}{2})\theta - \tfrac{1}{4}\pi] + O(N^{-\frac{3}{2}}), \tag{24}$$

$$d/d(\cos\theta)P_N(\cos\theta) = [2N/(\pi \sin^3\theta)] \times \{\sin[(N + \tfrac{1}{2})\theta - \tfrac{1}{4}\pi] - O(1)/N \sin\theta\}. \tag{25}$$

Substitution of these in (13) yields the asymptotic level density

$$NP(\lambda) \underset{\text{large } N}{\sim} (N/\pi)(1 - \lambda^2)^{-\frac{1}{2}}. \tag{26}$$

(26) is properly normalized as it stands. (24), (25), and (26) break down at the end points, $\lambda = \pm 1$, where the density is actually $\frac{1}{2}N^2$. A discussion of (26) is given in Sec. IV.

Noticing that $NP(0) = N/\pi$, we proceed to obtain $M(u, v)$,

$$M(u, v) = \lim_{N\to\infty} (\tfrac{1}{2}N)[u - v]^{-1}[P_N(\pi u/N) \times P_{N-1}(\pi v/N) - P_{N-1}(\pi u/N)P_N(\pi v/N)]. \tag{27}$$

Here, the necessary asymptotic formulas do not seem to appear in the standard literature. However, examination of the coefficients for the Legendre polynomials of large M shows that

$$P_{2M}(\pi u/2M) \underset{\text{large } M}{\longrightarrow} (-1)^M(\pi M)^{-\frac{1}{2}} \cos(\pi u), \tag{28}$$

$$P_{2M-1}(\pi u/2M) \underset{\text{large } M}{\longrightarrow} (-1)^M(\pi M)^{-\frac{1}{2}} \sin(\pi u). \tag{29}$$

We may simplify matters without loss of generality by looking at the case $N = 2M$, whereby $M(u, v)$ is quickly found to be

$$M(u, v) = [\pi(u - v)]^{-1} \sin(u - v). \tag{30}$$

By a series of transformations analogous to those of Gaudin,[16] one transforms the integral equation

$$\epsilon w(u) = \int_{-\frac{1}{2}x}^{\frac{1}{2}x} [\pi(u - v)]^{-1} \sin\pi(u - v)w(v)\, dv \tag{31}$$

to

$$\nu F(r) = \tfrac{1}{2}\int_{-1}^{1} \exp\{\tfrac{1}{2}i\pi xrs\}F(s)\, ds. \tag{32}$$

Here, $\nu^2 = \eta\epsilon/x$ and η is $(1, -1)$ according as the corresponding eigenfunction $F(r)$ or $w(u)$ is (even, odd). We emphasize that $F(r)$ and ν are both dependent on x, although this dependence is suppressed by our shorthand notation. (32) is the integral equation for the prolate spheroidal functions, which have been investigated extensively.[19] (22) can be rewritten as

$$\psi^{(0)}(x) = \prod_{i=1}^{\infty} (1 - \eta_i\nu_i^2 x). \tag{33}$$

Application of (14) yields $p^{(0)}(x)$, the nearest-neighbor spacing distribution.

IV. DISCUSSION

We now discuss the foregoing results. The level density (26) should be compared with the corresponding level densities for the Gaussian and Dyson ensembles (both for $\beta = 2$). The Gaussian ensemble yields[20] zero level density for $|\lambda| > (2N\sigma^2)^{\frac{1}{2}}$, and

$$NP_G(\lambda) = (\pi\sigma^2)^{-1}[2N\sigma^2 - \lambda^2]^{\frac{1}{2}} \tag{34}$$

for $|\lambda| < (2N\sigma^2)^{\frac{1}{2}}$, while the Dyson ensemble yields a constant level density equal to $N/2\pi$. The latter two densities bear little resemblance, qualitative or quantitative, to actual nuclear or atomic *global* level densities. On the other hand, for $\lambda > 0$, (26) does *qualitatively* resemble such global densities; i.e., it is concave upward and rapidly increasing. The significance of this feature, if any, is not completely understood. Nevertheless, it is certainly aesthetically pleasing and should be investigated further.

In general, matrix ensembles corresponding to the distributions (3) are not known. Special examples which *are* known arise in the study of "formally invariant" ensembles of real orthogonal matrices.[21] The quantities $\lambda_j = \cos\theta_j$, where the θ_j are the rotation angles in a random real orthogonal matrix, are distributed according to (3). The weight function $f(\lambda)$ may be $[1 - \lambda^2]^{-\frac{1}{2}}$, $[1 - \lambda^2]^{\frac{1}{2}}$, $[(1 - \lambda)/(1 + \lambda)]^{\frac{1}{2}}$,

[18] G. Szego, *Orthogonal Polynomials* (American Mathematical Society Publications, New York, 1959), 2nd ed., pp. 192 and 234.

[19] J. A. Stratton, P. M. Morse, L. J. Chew, J. D. C. Little, and F. J. Corbato, *Spheroidal Wave Functions* (The Technology Press, Cambridge, Massachusetts and John Wiley & Sons, Inc., New York, 1956).

[20] E. P. Wigner, "Distribution Laws for the Roots of a Random Hermitean Matrix" (1962) (Unpublished).

[21] F. J. Dyson, J. Math. Phys. 3, 1199 (1962).

or $[(1 + \lambda)/(1 - \lambda)]^{\frac{1}{2}}$, depending on the dimension and parity of the matrices. The corresponding orthogonal polynomials which naturally arise are the Jacobi polynomials $P_n^{(-\frac{1}{2},-\frac{1}{2})}(\lambda)$, $P_n^{(\frac{1}{2},\frac{1}{2})}(\lambda)$, $P_n^{(\frac{1}{2},-\frac{1}{2})}(\lambda)$, and $P_n^{(-\frac{1}{2},\frac{1}{2})}(\lambda)$, respectively (see Ref. 18, Chap. IV). It is not difficult to show that these ensembles each have level densities identical to (26). Further investigation shows that this feature is shared by all "Jacobi ensembles"; i.e., ensembles with $f(\lambda) = (1 - \lambda)^{\alpha}(1 + \lambda)^{\beta}$, $b = -a = 1$, and $\alpha > -1$, $\beta > -1$.[10]

Since the level density (26) is different from (34), we may conclude that the Legendre eigenvalue distribution *cannot* originate from a so-called Wigner matrix ensemble.[22] If a Hermetian matrix ensemble exists which corresponds to the Legendre case, it will evidently correlate different matrix elements, unlike the Wigner ensembles. This is not surprising if one accepts the heuristic derivation of (3) from the Dyson ensembles, where no such correlation need exist.

Several remarks are in order concerning the method used by Mehta[3] and Wigner[20] to obtain the Gaussian ensemble level density for $\beta = 1$ and 2, respectively. These authors used the fact that the Hermite functions (multiplied by their weight function) can be interpreted as the probabilities to find a quantum-mechanical linear harmonic oscillator in particular energy eigenstates. Their analyses used a semiclassical argument to evaluate (11) for large N. This method depended critically on the fact that the Hermite functions multiplied by $f^{\frac{1}{2}}(x)$ satisfy a Schrödinger equation. No such argument can be made for the Legendre ensemble. It is fortunate that the asymptotic formulas (28) and (29) exist in view of the fact that a semiclassical argument does not. Furthermore, it is noteworthy that for the Hermite case no asymptotic forms have been found in the literature which yield the semicircle law directly from (11), although such forms must be derivable in principle.

The results (39) and (41) for the nearest-neighbor spacing distribution are identical to Dyson's equations (114) and (115), respectively.[23] Also, Mehta and Dyson[24] have pointed out that the same results are found for the Gaussian distribution (1) with $\beta = 2$. The methods employed in Sec. III reveal that, in fact, *all* Jacobi ensembles with $\beta = 2$ also yield the same results, the Legendre ensemble being a special case.[10] A numerical calculation of the spacing distribution has been made by Kahn.[17]

The above results suggest that perhaps a *large* class of functions $f(\lambda)$ and intervals $[a, b]$ yield the same nearest-neighbor spacing distributions but, in general, different level densities. This, in turn, suggests that many orthogonal polynomials give rise to the same kernel $M(u, v)$. Wigner[20] has pointed out that, for the Hermite case, *all* the n-level joint distribution functions are expressible as functions of the $K_N(x, y)$. This fact is true for the ensembles given by (3) as well. *If* it is also true that the higher-order spacing distributions are expressible as functionals of the $K_N(x, y)$, then we expect the mth-order spacing distributions ($m > 0$) to be the same for all ensembles whose nearest-neighbor distributions are identical. This conjecture rests on the fact that double limits similar to those in Sec. II are involved in calculations of the mth-order distributions.[8] We suspect that the higher-order distributions may be expressible in terms of the $M(u, v)$ alone.

The cases $\beta = 1$ and $\beta = 4$ and the higher-order spacing distributions have yet to be investigated for the class of ensembles given by (3).

ACKNOWLEDGMENT

I would like to express my gratitude to Professor Max Dresden for his encouragement and guidance during the course of this work.

[22] See Ref. 1, p. 115. Strictly speaking, the Wigner ensembles are defined for $\beta = 1$. However, we expect their generalizations for $\beta = 2$ to have semicircular level densities.

[23] These equations refer to the third reference of footnote 4.

[24] M. L. Mehta and F. J. Dyson, J. Math. Phys. 4, 713 (1963).

Some Dynamical Properties of an Impurity in the Hard-Sphere Bose Gas*

TOSHIO SODA
The Enrico Fermi Institute for Nuclear Studies, The University of Chicago, Chicago, Illinois
(Received 3 December 1963)

The dynamical properties of an impurity in the hard-sphere Bose gas are investigated by extending the pseudopotential method and Hamiltonian of the author's previous work. The interaction potential between two impurities is calculated up to one quasiparticle exchange and is shown to have a Yukawa form of attraction with a range of $(16\pi\rho a)^{-1/2}$ in addition to the hard-core interaction at the origin. (ρ and a are the density of the boson gas and hard-core diameter of the boson, respectively.) Necessary conditions for impurities to have a bound state are also obtained. It is found that the mass of the impurity must be greater than approximately nine times the boson mass, and the density of the boson gas must lie in some range with a maximum and minimum. Then the differential cross section for Compton-type scattering and one quasiparticle production on an impurity are calculated. The Čerenkov radiation of sound quanta is studied when the impurity is traveling very fast in the hard-sphere Bose gas. It is found that there is no radiation unless the velocity v of the impurity is faster than $[k^2 + (16\pi\rho a)]^{1/2}$ and the radiation has a maximum for the angle θ_0 such that $\cos\theta_0 = (k^2 + 16\pi\rho a)^{\frac{1}{2}}/v$. The cross section for Čerenkov radiation is also calculated. (k is the momentum of the quasiparticle emitted in the Čerenkov radiation.)

1. INTRODUCTION

In a previous paper (to be referred to as I),[1] we have investigated the excitation spectrum of an impurity in the hard-sphere Bose gas, using the pseudopotential method of Lee, Huang, and Yang. There we have studied only the self energy of an impurity. But there are several other interesting properties which can be investigated using the methods of I. Examples are: (a) the force between two impurities due to the exchange of a quasiparticle, (b) scattering of an impurity by a quasiparticle, (c) production of a quasiparticle in the scattering of a quasiparticle by an impurity, and (d) the Čerenkov radiation of sound quanta by a fast impurity. This paper is concerned with the understanding of these dynamical properties of an impurity. In Sec. 2, we will discuss the kind of interaction potential which works between two impurities and check if the resultant potential can produce a bound state of two impurities. In Sec. 3, we calculate the differential cross section for Compton-type scattering and for production of one quasiparticle by scattering. In Sec. 4, we study the Čerenkov radiation of sound quanta when an impurity travels faster than the sound velocity in the hard Bose gas medium.

2. INTERACTION POTENTIAL BETWEEN TWO IMPURITIES

We assume a few (n) impurity particles among N identical bosons in a volume V. We use the same notation and unit system of $\hbar = 2m_a = 1$ as in I, except for the introduction of the symbol V for the volume. We have the following equation for the total number of "b" particles, instead of Eq. (8) of I:

$$\Sigma_p b_p^* b_p = n. \tag{1}$$

There is an additional direct impurity–impurity interaction given in the pseudopotential form,

$$\frac{4\pi a_i r}{V} \Sigma_{p,p',q} b_{p+q}^* b_{p'}^* b_{p'+q} b_p, \tag{2}$$

where a_i is the hard-core diameter of an impurity and r is the ratio of the boson to the impurity masses ($r = m_a/m_b$).

We work here for a general impurity case in which the impurity has a different mass and interaction than the medium bosons. We use the same approximation for the Hamiltonian as in I, that is, keeping terms which are coupled to the zero momentum state of the boson medium. By adding an interaction term (2) to the Hamiltonian defined Eq. (27) of I and the use of Eq. (1) and a Bogoliubov transformation on the "a" particle part, we have the following Hamiltonian form for our system[2]:

$$H = \frac{1}{2}\frac{N(N-1)}{V} 8\pi a + \Sigma_p p(p^2 + 16\pi\rho a)^{\frac{1}{2}} \alpha_p^* \alpha_p$$

$$+ \tfrac{1}{2}\Sigma_p [p(p^2 + 16\pi\rho a)^{\frac{1}{2}} - (p^2 + 8\pi\rho a)]$$

* This work was supported by the U. S. Atomic Energy Commission.

[1] T. Soda, J. Math. Phys. 5, 142 (1964). We refer to this as I.

[2] *Note added in proof.* The last term of Eq. (3) is obtained from the impurity-boson interaction

$$\Sigma_{p,p',q} V_{ab} a_{p+q}^* b_{p'}^* b_{p'+q} a_p$$

by replacing one of the "a" operators with $(N)^{\frac{1}{2}}$ and the other "a" operator with a quasiparticle representation:

$$a_p = u_p \alpha_p - v_p \alpha_{-p}^*.$$

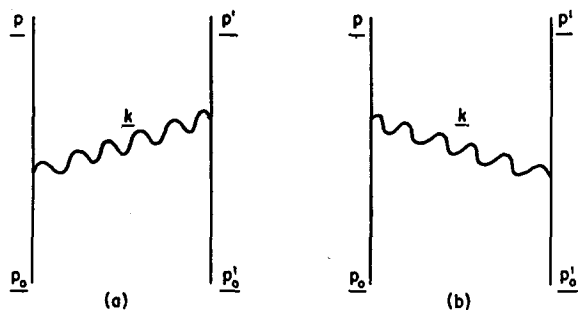

FIG. 1. The diagram for the one-quasiparticle exchange potential between two impurities.

$$+\frac{N}{V}8\pi c\left(\frac{1+r}{2}\right)n+\Sigma_p rp^2 b_p^* b_p$$

$$+\frac{4\pi a_i r}{V}\Sigma_{p,p',q} b_{p+q}^* b_{p'}^* b_{p'+q} b_p+\frac{N^{\frac{1}{2}}}{V}8\pi c\left(\frac{1+r}{2}\right)$$

$$\times\{\Sigma_p(u_p-v_p)[(\alpha_{-p}^*+\alpha_p)b_0^* b_{-p}+b_{-p}^* b_0(\alpha_{-p}+\alpha_p^*)]$$

$$+\sum_{\substack{p\neq 0,\\ q\neq 0}}(u_{p+q}\alpha_{p+q}^*-v_{p+q}\alpha_{-p-q})b_{-p}^* b_q$$

$$+b_q^* b_{-p}(u_{p+q}\alpha_{p+q}-v_{p+q}\alpha_{-p-q}^*)\}. \tag{3}$$

The last three lines of Hamiltonian contain an interaction term linear in the quasiparticle operator and quadratic in the "b" particle operators in addition to the direct "b" particle interaction given in (2). Methods of the treatment of this linear term in the quasiparticle operators are well known in the construction of nuclear potentials.[3] There are several methods used to derive the results, but all of them give the same form for lowest order in one-quantum exchange. The lowest-order potential is given by the Fourier transform of the second-order perturbation energy given by

$$\langle H_i(1/-H_0)H_i\rangle_{\text{quasiparticle vacuum}} \tag{4}$$

where H_0 is the free Hamiltonian and H_i is the interaction Hamiltonian between the quasiparticles and the impurities. The lowest-order diagram in the one-quasiparticle exchange potential is that given in Fig. 1(a) and (b).

Suppose the two impurities in Fig. 1 to be in positions of $\mathbf{r}_1$ and $\mathbf{r}_2$ with initial momenta $\mathbf{p}_0$ and $\mathbf{p}_0'$ and final momenta $\mathbf{p}$ and $\mathbf{p}'$. Let one quasiparticle with momentum k be exchanged. Then the potential is given by

$$V(r_{12})=\left(8\pi c\frac{1+r}{2}\right)^2\frac{N}{V^2}\Sigma_k(u_k-v_k)^2 e^{i\mathbf{k}\cdot\mathbf{r}_{12}}$$
$$\times\left[\frac{1}{r(p_0^2-p^2)-\omega_k}+\frac{1}{r(p_0'^2-p'^2)-\omega_k}\right]$$
$$[\omega_k=k(k^2+16\pi\rho a)^{\frac{1}{2}}]. \tag{5}$$

Suppose we choose our coordinate system to be in the center-of-momentum system, that is $\mathbf{p}_0+\mathbf{p}=0$ and $\mathbf{p}_0'+\mathbf{p}'=0$. Then expression (5) can be evaluated as follows.

$$V(r_{12})$$
$$=-2\rho\left(c\frac{1+r}{2}\right)^2\frac{(8\pi)^2}{8\pi^3}\int d^3k\,\frac{e^{i\mathbf{k}\cdot\mathbf{r}_{12}}k^2}{k^2(k^2+16\pi\rho a)}$$
$$=-64\rho\left(c\frac{1+r}{2}\right)^2\frac{1}{r_{12}}\int_0^\infty dk\,\frac{k\sin kr_{12}}{k^2+16\pi\rho a}$$
$$=-32\pi\rho\left(c\frac{1+r}{2}\right)^2\frac{\exp[-(16\pi\rho a)^{\frac{1}{2}}r_{12}]}{r_{12}}. \tag{6}$$

We obtain an attractive potential of Yukawa or Debye–Hückel type between the two impurities.[4] The potential for two or more quasiparticle exchange is of a higher order in an expansion in powers of (ρa^3). For example, the two-quasiparticle potential has a factor of $(\rho a^2)^2$ in the coupling and can be neglected compared to $V(r_{12})$ given in Eq. (6) in a low-density expansion of the perturbation result. The resulting over-all potential between two impurities including the direct interaction appears in Fig. 2. It has a hard core in the center, and a Yukawa-shaped attractive tail with a range of $(16\pi\rho a)^{-\frac{1}{2}}$. This situation resembles a molecular interaction.

We check now whether this attractive interaction permits two impurities to form a bound state. For this purpose, an exact formula derived by Low[5] is used. This is a condition for a bound state,

$$-\int_0^\infty V(r_{12})u(r_{12})\sinh(\alpha r_{12})\,dr_{12}=2\alpha r, \tag{7}$$

where $\alpha=(-\epsilon/2r)^{\frac{1}{2}}$, $V(r_{12})$ is the potential including the hard core, and $u(r_{12})$ is the bound-state wavefunction, satisfying

$$\frac{d^2u(r_{12})}{dr_{12}^2}-\alpha^2u(r_{12})=\frac{1}{2r}V(r_{12})u(r_{12}). \tag{8}$$

Here the function $\sinh(\alpha r_{12})$ is a solution for free-state wavefunction, w, with an energy ϵ, which satisfies the equation

$$d^2w(r_{12})/dr_{12}^2-\alpha^2w(r_{12})=0. \tag{9}$$

[3] K. Nishijima, Prog. Theoret. Phys. (Kyoto) Suppl. 3, 138 (1956). See also other articles in the same issue.

[4] *Note added in proof.* The potential of Eq. (7) is also derived and ascertained using a different method by T. Usui and Y. Nagaoka, Bussei Kenkyu, (Kyoto) 1, 201 (1963). See also D. Pines, *Liquid Helium, Proceedings of the International School of Physics, Enrico Fermi*, (Academic Press, Inc., New York, 1963), discussion on p. 183.

[5] F. Low, Phys. Rev. 74, 188 (1948). See also J. M. Blatt and V. F. Weisskopf, *Theoretical Nuclear Physics* (John Wiley & Sons, Inc., New York, 1952), p. 54.

Using the following properties of the wave-function and potential,

$$\left.\begin{aligned} u(r_{12}) &= 0, \\ V(r_{12}) &= \infty \end{aligned}\right\} \quad \text{for} \quad r_{12} < a_i, \tag{10}$$

$$u(r_{12}) \leq e^{-\alpha r_{12}} \quad \text{for} \quad r_{12} > a_i,$$

and

$$(1 - e^{-2\alpha r_{12}})/2\alpha \leq r_{12},^{6}$$

we can derive the following necessary condition from Eq. (7),

$$-\int_{a_i}^{\infty} V(r_{12}) r_{12}\, dr_{12} \geq 2r. \tag{11}$$

Here $V(r_{12})$ is given by Eq. (6) and if we insert this in Eq. (11), we obtain

$$32\pi\left(c\,\frac{1+r}{2}\right)^2 \rho \int_{a_i}^{\infty} \exp\left[-(16\pi\rho a)^{\frac{1}{2}} r_{12}\right] dr_{12} \geq 2r. \tag{12}$$

By calculation, we have the following relationship:

$$\left(\frac{c^2}{aa_i}\right)(16\pi\rho a a_i^2)^{\frac{1}{2}} \exp\left[-(16\pi\rho a a_i^2)^{\frac{1}{2}}\right] \geq r\left(\frac{2}{1+r}\right)^2. \tag{13}$$

If we set $x = (16\pi\rho a\, a_i^2)^{\frac{1}{2}}$, $y = -\ln [xe^{-x}]$, and $s = a_i/a$, we obtain $[c = \frac{1}{2}(a + a_i)]$

$$y = x - \ln x \leq 2 \ln \tfrac{1}{2}(1 + r) - \ln r + \ln \tfrac{1}{2}(1 + s) + \ln \tfrac{1}{2}(1 + 1/s). \tag{14}$$

In Fig. 3, the function $y = x - \ln x$ is plotted. On the same figure we draw horizontal lines with y's corresponding to the values of the right-hand side of the inequality in Eq. (14) for different values of r. Here we need a value of the parameter s, the hard-core ratio. But we assume s to behave as a function of r, corresponding to real spherical atoms. And these r have been chosen to correspond to masses of impurities in helium for which the ratio of hard-core radius s is empirically known. These impurities are He, A, Ne, Kr, Xe, etc. This is a fine point, because at any rate the sum of the last two terms of Eq. (14), which contains s, is always very close to zero.

From this curve we can draw the following necessary condition for having a bound state for two impurities. The impurities must be heavier than approximately nine times the boson mass, because it is impossible to satisfy the inequality (14) for r greater than $\frac{1}{9}$. The exact value at which the curve $y = x - \ln x$ intercepts with the horizontal line is $r = 0.1142$, and ρa^3 takes the value 0.0136.

[6] *Note added in proof.* The excessive simplification of the last inequality may be justified by the smallness of α and our approximation of not taking into account higher quasi-particle exchange forces for the bound-state problem.

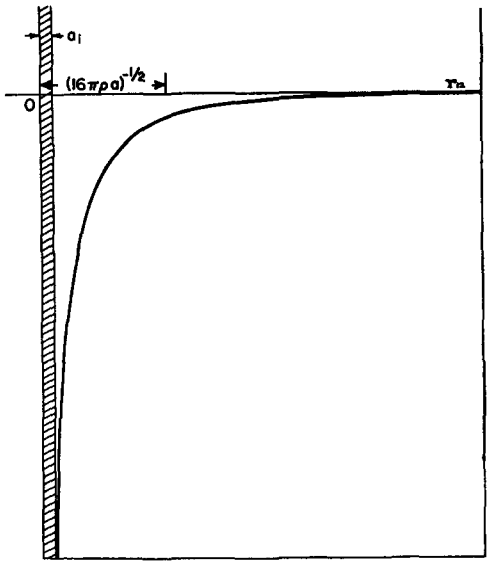

FIG. 2. Over-all potential between two impurities in the relative coordinate, r_{12}, of the two impurities.

When the bound states are allowed, the density of the hard-sphere boson gas must lie between some maximum and minimum values. The minimum density is that at which the strength of the attractive interaction becomes too weak to produce a bound state. The maximum exists, because the ratio of the hard-core radius to the attractive potential range becomes too large. The range of the allowed densities increases with impurity mass. Some examples for this maximum and minimum densities are given below.

$$\rho a^3 = 0.0136 \quad \text{for} \quad r = 0.1142;$$

$$0.0224 \geq \rho a^3 \geq 0.0116 \quad \text{for} \quad r = 1/9;$$

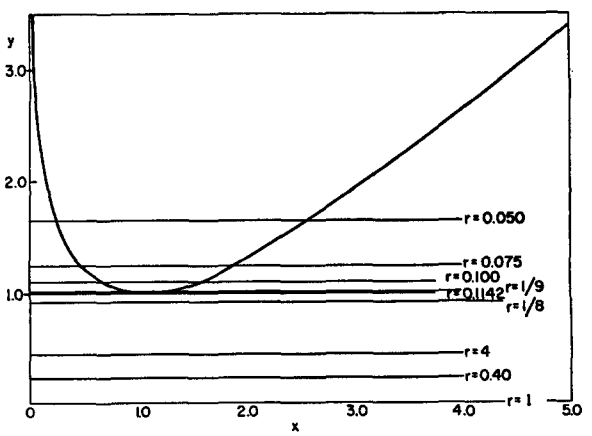

FIG. 3. Plot of $y = x - \ln x$ vs $x = (16\pi\rho a)^{\frac{1}{2}} a_i$, defining the lower bound for a bound state. The horizontal lines $y(r)$ gives the value y for particular mass ratios r. The intersections of these lines with the curve define the interval of x, within which binding is allowed.

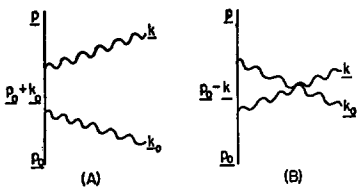

FIG. 4. The diagram of an impurity scattering by one quasiparticle.

$$0.0456 \geq \rho a^3 \geq 0.00316 \quad \text{for} \quad r = 7.5 \times 10^{-2};$$
$$0.0795 \geq \rho a^3 \geq 0.00076 \quad \text{for} \quad r = 5 \times 10^{-2}.$$

Thus we see that impurities in a low-density Bose gas attract each other, and we have found that the necessary conditions for these impurities to form bound states are that the mass of impurity is greater than approximately nine times the boson mass and that the density of the boson gas lies between some maximum and minimum values of the order mentioned above.[7]

3. SCATTERING AND EMISSION OF A QUASIPARTICLE BY AN IMPURITY

Let us consider the following scattering process. In the initial state, an impurity has a momentum $\mathbf{p}_0$ and is scattered by an incident quasiparticle with momentum $\mathbf{k}_0$ into the final state of momentum $\mathbf{p}$, changing the quasiparticle into the momentum state $\mathbf{k}$. This is an analogous case to the Compton scattering of an electron by a photon. The diagrams for this process are given by (A) and (B) of Fig. 4.

The transition matrix element R is given by

$$R_{fi} = \Sigma_c \frac{\langle f|H_i|c\rangle\langle c|H_i|i\rangle}{E_i - E_c + i\epsilon}, \tag{15}$$

where f, i, c denote the final, initial and intermediate states, and E_i and E_c are the energies of the initial and intermediate states.

The actual transition matrix elements for diagrams (A) and (B) are

$$R_A = \left(8\pi c \frac{1+r}{2}\right)^2 \left(\frac{N}{V}\right)^{\frac{1}{2}} \times \frac{(u_{k_0} - v_{k_0})(u_k - v_k)}{\omega_{k_0} + r[p_0^2 - (\mathbf{p}_0 + \mathbf{k}_0)^2]}, \tag{16}$$

and

$$R_B = \left(8\pi c \frac{1+r}{2}\right)^2 \left(\frac{N}{V}\right)^{\frac{1}{2}} \times \frac{(u_{k_0} - v_{k_0})(u_k - v_k)}{-\omega_k + r[p_0^2 - (\mathbf{p}_0 - \mathbf{k}_0)^2]}. \tag{17}$$

[7] *Note added in proof.* We have not gone beyond the one quasiparticle exchange potential here. However, if we take into account higher quasiparticle exchange potentials near the core region, and the large effective mass of an impurity as calculated in I, we might hope to obtain a qualitative explanation of the phase separation of He^3 in He^4 at $T = 0$, by improving the lowerbound of the mass of the impurity for the bound state to $\frac{3}{4}$, instead of the rather large value quoted in the above, so that He^3 in He^4 might form clusters by itself and not mix with He^4 liquid.

The transition probability for a quasiparticle to be scattered into the solid angle between Ω and $\Omega + d\Omega$ is then

$$dw = 2\pi \int |R_A + R_B|^2 \times \frac{Vk^2}{(2\pi)^3}\frac{dk}{dE_f} d\Omega \delta(E_f - E_i)\, dE_f \tag{18}$$

$$= 2\pi \frac{\rho^2}{V}\left(8\pi c \frac{1+r}{2}\right)^4 \times \frac{(u_{k_0} - v_{k_0})^2(u_k - v_k)^2 k^2\, d\Omega}{2r[k - (p_0 + k_0)\cos\theta] + \frac{k}{\omega_k}(16\pi\rho a + 2k^2)} \times \left\{\frac{1}{\omega_{k_0} + r[p_0^2 - (\mathbf{p}_0 + \mathbf{k})^2]} + \frac{1}{-\omega_k + r[p_0^2 - (\mathbf{p}_0 - \mathbf{k})^2]}\right\}^2. \tag{19}$$

$E_f = \omega_k + rp^2$ is the energy of the final state, and we have made use of the relationship

$$\frac{dE_f}{dk} = \frac{k(2k^2 + 16\pi\rho a)}{\omega_k} + 2r[k - (p_0 + k_0)\cos\theta]. \tag{20}$$

Here $\cos\theta$ is the angle between $\mathbf{k}$ and $\mathbf{p}_0 + \mathbf{k}_0$. The differential cross section $d\sigma$ is obtained by dividing the quantity dw in Eq. (19) by $V^{-1}(d\omega_{k_0}/dk_0)$ and we have

$$d\sigma = \frac{dwV}{(d\omega_{k_0}/dk_0)} = \frac{1}{(2\pi)^2}\frac{[8\pi c\tfrac{1}{2}(1+r)]^4}{[2k_0^2 + 16\pi\rho a]}\frac{k^2}{\omega_k} \times \frac{k_0 k^2\, d\Omega}{2r[k - (p_0 + k_0)\cos\theta] + k^3/\omega_k + \omega_k/k} \times \left[\frac{1}{\omega_{k_0} - r(k_0^2 + 2\mathbf{p}_0\cdot\mathbf{k}_0)} + \frac{1}{-\omega_k - r(k_0^2 - 2\mathbf{p}_0\cdot\mathbf{k}_0)}\right]^2. \tag{21}$$

We can see that the maximum scattering for the quasiparticle occurs near the forward direction.

Next we proceed to calculate the production process for one extra quasiparticle. There are five diagrams, shown in Fig. 5, which contribute to this process. In this process the impurity undergoes a change of momentum from $\mathbf{p}_0$ to $\mathbf{p}$, while the quasiparticle of momentum $\mathbf{k}_0$ is absorbed and two quasiparticles of momenta $\mathbf{k}_1$ and $\mathbf{k}_2$ are produced.

In Fig. 5 the processes (D) and (E) contribute $O(1/V)$ in the scattering cross section, as we will see later, and are vanishingly small. Therefore, the transition probability is due to the processes (A), (B), and (C). It is

$$R_{A+B+C} = \left(\frac{N^{\frac{1}{2}}}{V}\right)^3 \left(8\pi c \frac{1+r}{2}\right)^3$$
$$\times (u_{k_0} - v_{k_0})(u_{k_1} - v_{k_1})(u_{k_2} - v_{k_2})$$
$$\times \left\{ \frac{1}{-\omega_{k_1} + r[p_0^2 - (\mathbf{p}_0 - \mathbf{k}_1)^2]} \right.$$
$$\times \frac{1}{\omega_{k_0} - \omega_{k_1} + r[p_0^2 - (\mathbf{p}_0 + \mathbf{k}_0 - \mathbf{k}_1)^2]}$$
$$+ \frac{1}{-\omega_{k_1} + r[p_0^2 - (\mathbf{p}_0 - \mathbf{k}_1)^2]}$$
$$\times \frac{1}{-\omega_{k_1} - \omega_{k_2} + r[p_0^2 - (\mathbf{p}_0 - \mathbf{k}_1 - \mathbf{k}_2)^2]}$$
$$+ \frac{1}{\omega_{k_0} + r[p_0^2 - (\mathbf{p}_0 + \mathbf{k}_0)^2]}$$
$$\left. \times \frac{1}{\omega_{k_0} - \omega_{k_1} + r[p_0^2 - (\mathbf{p}_0 + \mathbf{k}_0 - \mathbf{k}_1)^2]} \right\}. \quad (22)$$

The differential cross section for these processes is

$$d\sigma_{A+B+C} = \frac{\rho^3}{(2\pi)^5} \left(8\pi c \frac{1+r}{2}\right)^6 \frac{k_0}{2k_0^2 + 16\pi\rho a}$$
$$\times \frac{k_1^2 k_2^2}{\omega_{k_1}\omega_{k_2}} k_1^2\, dk_1\, d\Omega_1 k_2^2\, dk_2\, d\Omega_2$$
$$\times [\delta(rp_0^2 + \omega_{k_0} - rp^2 - \omega_{k_1} - \omega_{k_2})]$$
$$\times \left\{ \frac{1}{-\omega_{k_1} + r[p_0^2 - (\mathbf{p}_0 - \mathbf{k}_1)^2]} \right.$$
$$\times \frac{1}{\omega_{k_0} - \omega_{k_1} + r[p_0^2 - (\mathbf{p}_0 + \mathbf{k}_0 - \mathbf{k}_1)^2]}$$
$$+ \frac{1}{-\omega_{k_1} + r[p_0^2 - (\mathbf{p}_0 - \mathbf{k}_1)^2]}$$
$$\times \frac{1}{-\omega_{k_1} - \omega_{k_2} + r[p_0^2 - (\mathbf{p}_0 - \mathbf{k}_1 - \mathbf{k}_2)^2}$$
$$+ \frac{1}{\omega_{k_0} + r[p_0^2 - (\mathbf{p}_0 + \mathbf{k}_0)^2]}$$
$$\left. \times \frac{1}{\omega_{k_0} - \omega_{k_1} + r[p_0^2 - (\mathbf{p}_0 + \mathbf{k}_0 - \mathbf{k}_1)^2}} \right\}^2, \quad (23)$$

where a quasiparticle is scattered into the solid angle $d\Omega_1$, in the momentum interval between k_1 and $k_1 + kd_1$, and another quasiparticle is produced into the solid angle $d\Omega_2$, having momentum in the interval between k_2 to $k_2 + dk_2$.

For the calculation of processes (D) and (E), we have to add a neglected part of the full Hamiltonian. That is,

$$H_4' = \frac{8\pi c^{\frac{1}{2}}(1+r)}{V} \Sigma_{\mathbf{p},\mathbf{p}',\mathbf{q}} (u_{p+q}\alpha_{\mathbf{p}+\mathbf{q}}^* - v_{p+q}\alpha_{-\mathbf{p}-\mathbf{q}})$$
$$\times b_{\mathbf{p}}^* b_{\mathbf{p}'+\mathbf{q}} (u_p \alpha_{\mathbf{p}} - v_p \alpha_{-\mathbf{p}}^*). \quad (24)$$

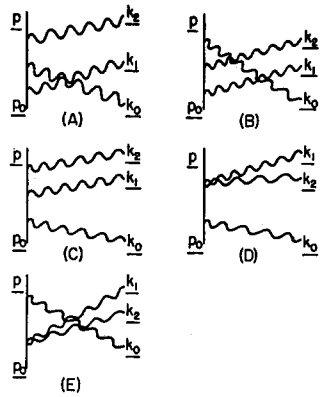

FIG. 5. The diagram of one quasiparticle production process by impurity–boson scattering.

The transition probabilities for the processes (D) and (E) are

$$R_D = \left(8\pi c \frac{1+r}{2}\right)^2 \frac{N^{\frac{1}{2}}}{V^2} \frac{(u_{k_0} - v_{k_0})u_{k_1}v_{k_2}}{\omega_{k_0} + r[p_0^2 - (\mathbf{k}_0 + \mathbf{p}_0)^2]}, \quad (25)$$

and

$$R_E = \left(8\pi c \frac{1+r}{2}\right)^2 \frac{N^{\frac{1}{2}}}{V^2}$$
$$\times \frac{(u_{k_0} - v_{k_0})u_{k_1}v_{k_2}}{-\omega_{k_1} + r[p_0^2 - (\mathbf{p}_0 - \mathbf{k}_1 - \mathbf{k}_2)^2]}. \quad (26)$$

The differential cross section for these is

$$d\sigma_{D+E} = \frac{1}{(2\pi)^2} \frac{V^2}{d\omega_{k_0}/dk_0} |R_D + R_E|^2$$
$$\times k_1^2\, dk_1\, d\Omega_1 k_2^2\, dk_2\, d\Omega_2 \delta(E_i - E_f) \simeq O\left(\frac{1}{V}\right), \quad (27)$$

which is inversely proportional to the volume.

The absorption process can also be calculated in a similar fashion.

4. ČERENKOV RADIATION OF SOUND QUANTA

When an impurity particle travels in the medium with a faster velocity than the speed of a quasiparticle, this impurity particle emits quasiparticles. This is the Čerenkov type of radiation associated with quasiparticles. We may call this quasiparticle a sound quantum, and thus this is the Čerenkov radiation of sound quanta.

Let us calculate this process. The diagram is given in Fig. 6. We suppose that an impurity undergoes a momentum change from $\mathbf{p}_0$ to $\mathbf{p}$, and the quasiparticle field of the medium balances this momentum change by emitting a quasiparticle of momentum $\mathbf{k}$. We apply time dependent perturbation theory to this process. From the formula[8] for

[8] L. I. Schiff, *Quantum Mechanics* (McGraw-Hill Book Company, Inc., New York, 1949), 1st ed., the equation after Eq. (29.9), p. 191. See, also, S. M. Neamtan, Phys. Rev. 92, 1362 (1953).

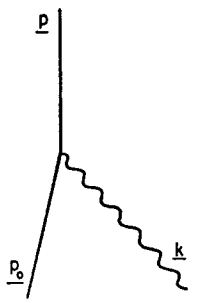

FIG. 6. The diagram of the Čerenkov radiation emitting a sound quantum with a momentum k.

the transition amplitude in Schiff, we obtain the following transition probability for the emission of one quantum into the solid angle $d\Omega$ and the momentum interval between k and $k + dk$, (t: time);

$$dw = \frac{4}{t} |\langle f |H_i| i\rangle|^2 \frac{k^2\, dk\, d\Omega}{(2\pi)^3} \frac{\sin^2 \frac{1}{2}\omega t}{\omega^2}, \qquad (28)$$

where ω is the difference in energy between the initial and final state and is given by

$$\begin{aligned}\omega &= r(p_0^2 - p^2) - \omega_k \\ &= kv \cos\theta - k(k^2 + 16\pi\rho a)^{\frac{1}{2}}, \end{aligned} \qquad (29)$$

where θ is the angle between $\mathbf{k}$ and $\mathbf{p}$. Here we have used $\mathbf{p}_0 - \mathbf{p} = \mathbf{k}$ and $r(\mathbf{p}_0 + \mathbf{p}) \simeq \mathbf{v}$, where $\mathbf{v}$ is the velocity of the impurity and the deflection of the high-momentum impurity is neglected. The matrix element is given by

$$\langle f |H_i| i\rangle = (N^{\frac{1}{2}}/V)[8\pi c\tfrac{1}{2}(1 + r)](u_k - v_k). \qquad (30)$$

Then dw is given by

$$dw = \frac{4}{t}\frac{\rho}{V}\left(8\pi c \frac{1+r}{2}\right)^2 \frac{1}{(2\pi)^3} \frac{k}{(k^2 + 16\pi\rho a)^{\frac{1}{2}}} k^2\, dk\, d\Omega \times \frac{\sin^2 \frac{1}{2}kt[v\cos\theta - (16\pi\rho a + k^2)^{\frac{1}{2}}]}{k^2[v\cos\theta - (16\pi\rho a + k^2)^{\frac{1}{2}}]^2}. \qquad (31)$$

From the expression (31), we can see that dw has a sharp maximum at θ_0, where

$$\cos\theta_0 = [16\pi\rho a + k^2]^{\frac{1}{2}}/v, \qquad (32)$$

and it also can be seen that there is no radiation unless $v > (16\pi\rho a + k^2)^{\frac{1}{2}}$. For small k this is the analogous result to that for ordinary radiation, because $(16\pi\rho a)^{\frac{1}{2}}$ is the magnitude of sound velocity. Let us now calculate the total transition probability of the radiation over the surface of a unit sphere. To facilitate the evaluation of the integral over the angle, we extend the upper and lower bound of the integral to infinity. This is allowed because the integral has the sharp maximum at $\theta = \theta_0$ given in Eq. (32). Then we have

$$2\pi \int_{-\infty}^{\infty} \frac{\sin^2 \frac{1}{2}kt[v\cos\theta - (16\pi\rho a + k^2)^{\frac{1}{2}}]}{k^2[v\cos\theta - (16\pi\rho a + k^2)^{\frac{1}{2}}]^2} d\cos\theta = \frac{\pi^2 t}{2kv}, \qquad (33)$$

and the total transition probability is given by

$$w = \frac{16\pi\rho}{V} \frac{[c\tfrac{1}{2}(1 + r)]^2}{v} \frac{k^2\, dk}{(k^2 + 16\pi\rho a)^{\frac{1}{2}}}. \qquad (34)$$

The total cross section σ is given by

$$\sigma = \frac{w}{v/V} = \frac{16\pi\rho[c\tfrac{1}{2}(1 + r)]^2}{v^2} \frac{k^2\, dk}{(k^2 + 16\pi\rho a)^{\frac{1}{2}}}. \qquad (35)$$

As an example, we calculate how many sound quanta will be produced per centimeter in the Bose gas. We assume the impurity particle to be a 5-MeV α particle, choose the density to be that of liquid helium and the hard-core diameter to be 2.6Å, and take the momentum of the sound quanta to be the cut-off momentum of a^{-1}. We have an energy loss of 2.8 eV/cm, and the production of approximately 1550 quanta per centimeter. Each sound quantum has an energy of 20°K Debye temperature or 1.8×10^{-3} eV. However, it should be remembered that our hard-sphere Bose gas model is far from real liquid helium, because our method is a low-density perturbation expansion.

Finally we remark that the neglected terms of the full Hamiltonian, which does not appear in Eq. (3), contribute either to the inverse of the volume or higher order in (ρa^3) in the calculation of the Čerenkov radiation of sound quanta, as well as in the calculation of the scattering and production processes of the previous section.

ACKNOWLEDGMENTS

The author acknowledges his gratitude to Professor Y. Nambu for his stimulating discussions and suggestions, and to Professor G. Wentzel for his helpful advice. He also thanks Dr. F. von Hippel for his helpful discussions. He appreciates the constructive comments of Professor K. Sawada and wishes to thank Dr. L. W. Bruch for helpful discussions.

Kramers-Wannier Duality for the 2-Dimensional Ising Model as an Instance of Poisson's Summation Formula

H. P. McKean, Jr.
Rockefeller Institute, New York, New York
(Received 21 January 1964)

The well-known Kramers–Wannier high-low-temperature duality for the 2-dimensional Ising model is derived by means of the Poisson summation formula for a commutative group.

1. POISSON'S SUMMATION FORMULA AND JACOBI'S IMAGINARY TRANSFORMATION

GIVEN a rapidly decreasing function $f \in C^\infty(R^1)$, its Fourier transform $f\hat{}(b) = \int \exp(2\pi iba) f\, da$ is also rapidly decreasing and belongs to $C^\infty(R^1)$, and Poisson's summation formula holds:

$$\sum_{m \in Z^1} f(m) = \sum_{n \in Z^1} f\hat{}(n)$$

Z^1 = the integers: $0, \pm 1, \pm 2$, etc.[1]

Jacobi's celebrated imaginary transformation of the theta function

$$Z(T) = \sum_{n \in Z^1} e^{-\pi n^2/T} \qquad (T > 0)$$

is an instance of this formula; in fact, putting $f = \exp(-\pi a^2/T)$ gives

$$Z(T) = T^{\frac{1}{2}} Z(1/T).$$

Z is the quantum-mechanical partition function of a 2-dimensional rotator with Boltzmann's constant replaced by $1/\pi$, and Jacobi's transformation can be viewed as a duality between high and low temperatures.

2. KRAMERS–WANNIER DUALITY

A second instance of a high–low-temperature duality is the Kramers–Wannier duality for the 2-dimensional Ising model.[2] This refers to an $n \times m$ square lattice Q with bounding edges identified in pairs so as to make it a torus; each pair of nearest-neighbor sites is joined by a *bond* and to each site is attached a *spin* $\sigma = \pm 1$; the spins interact along bonds in nearest-neighbor pairs according to the rule $\sigma'\sigma''$, and the (antiferromagnetic) partition function is

$$Z(C) = \sum_{\text{spins}} \exp\left(C \sum_{\text{bonds}} \sigma'\sigma''\right),$$

$C > 0$ being a constant divided by the Kelvin temperature T. Kramers–Wannier's duality states that

$$Z(C) = Z(C^*) \times (\tfrac{1}{2} \sinh 2C^*)^{-\frac{1}{2}B}.$$

B is the number of bonds of Q and C^* is the positive root of

$$\sinh 2C \sinh 2C^* = 1;$$

this relation between C and C^* interchanges high and low temperatures. Actually, the formula is not exact for an $n \times m$ latice, but it becomes so for an infinite lattice:

$$\mathfrak{Z}(C) = \lim_{\substack{n \uparrow \infty \\ m \uparrow \infty}} [Z(C)]^{1/nm} = (\tfrac{1}{2} \sinh 2C^*)^{-\frac{1}{2}} \mathfrak{Z}(C^*).$$

Here it is found that, like the Jacobi transformation, the Kramers–Wannier duality is an instance of a Poisson summation formula. This Poisson formula is stated in Sec. 3, a transformation of Z due to van der Waerden is outlined in Sec. 4, and then in Sec. 5 the Poisson formula is applied to the van der Waerden expression for Z and the Kramers–Wannier duality is obtained. A Kramers–Wannier duality holds between the triangular and honeycomb lattices also,[3] and can be obtained in the same manner.

3. POISSON'S SUMMATION FORMULA FOR A FINITE COMMUTATIVE GROUP

Given a finite commutative group G, let $G\hat{}$ be the class of all homomorphisms (characters) $b : G \to S^1$ of G into the circle group S^1 : $\exp(2\pi i\theta)(0 \leq \theta < 1)$. $G\hat{}$ is a group under multiplication $[(b_1b_2)(a) = b_1(a)b_2(a)]$. $G\hat{}$ is the *character* or *dual group* of G; it is isomorphic to G. Given a subgroup $H \subset G$, the dual $(G/H)\hat{}$ of the factor group G/H can be identified as the characters on the big group G that are constant ($\equiv 1$) on the little group H, and using this identification the Poisson summation formula can be stated as follows:

$$\sum_{a \in H} f(a) = \sum_{b \in (G/H)\hat{}} f\hat{}(b).$$

[1] R. Courant and D. Hilbert, Methods of Mathematical Physics (Interscience Publishers, Inc., New York, 1953, 1962), Vols. I and II.

[2] A. Münster, *Statistische Thermodynamik* (Springer-Verlag, Berlin, 1956).

[3] R. M. F. Houtappel, Physica **16**, 425 (1960).

$f\hat{}$ is the transform

$$f\hat{}(b) = \sum_{a\in G} b(a)f(a);$$

the inverse transform is

$$f(a) = \sum_{b\in G\hat{}} a(b)f\hat{}(b) \qquad [a(b) \equiv b(a)^*].$$

The reader will observe the similarity of this summation formula to the summation formula cited above in which R^1 plays the role of G, Z^1 the role of H, and $(R^1/Z^1)\hat{} = (S^1)\hat{} \simeq Z^1$ the role of $(G/H)\hat{}$; in fact, there is a general Poisson summation formula that includes both as special cases.[4]

G is going to be the group of *paths* on Q with coefficients from Z_2 (integers mod 2), i.e., the group of sums of bonds of Q under addition, counting a double bond as 0. H will be the subgroup of closed paths (*cycles*), and $(G/H)\hat{}$ will be the subgroup of *bounding cycles* on the dual lattice Q^*, but before the Poisson formula can be applied to Z, it is necessary to explain van der Waerden's expression for it.

4. VAN DER WAERDEN'S EXPRESSION FOR Z

Define C^* as the root of $\sinh 2C \sinh 2C^* = 1$; then

$$e^{\pm C} = (\tfrac{1}{2}\sinh 2C)^{\frac{1}{2}}(e^{C^*} \pm e^{-C^*}),$$

and following van der Waerden,

$$Z = \sum_{\text{spins}} \prod_{\text{bonds}} e^{C\sigma'\sigma''}$$
$$= (\tfrac{1}{2}\sinh 2C)^{\frac{1}{2}B} \sum_{\text{spins}} \prod_{\text{bonds}} (e^{C^*} + \sigma'\sigma''e^{-C^*}).$$

Develop the product into a sum over paths $a \in G$:

$$\prod_{\text{bonds}} (e^{C^*} + \sigma'\sigma''e^{-C^*}) = \sum_{a\in G} e^{(B-n)C^*-nC^*} \prod_{\substack{\text{bonds}\\ \text{of } a}} \sigma'\sigma'',$$

$n = n(a)$ being the number of bonds (length) of $a \in G$. Bringing in the sum over spins, it is evident that if an *odd* number of bonds of a path meet at a single site, then the corresponding piece of the sum over paths drops out since $(+1)^o+(-1)^o=0$ if o is odd; the surviving terms correspond to closed paths (*cycles*) and for such a path, the spin product under the path sum $\equiv +1$ so that the partition function is expressed by van der Waerden's formula

$$Z = (\tfrac{1}{2}\sinh 2C)^{\frac{1}{2}B} \sum_{\text{spins}} \sum_{a\in H} e^{(B-n)C^*-nC^*}$$
$$= 2^A(\tfrac{1}{2}\sinh 2C)^{\frac{1}{2}B} \sum_{a\in H} f(a),$$

A being the number of sites of the lattice and $f(a) = e^{(B-n)C^*-nC^*}$.

[4] L. H. Loomis, *Abstract Harmonic Analysis*, (D. Van Nostrand Company, Princeton, New Jersey, 1953).

5. APPLICATION OF POISSON'S FORMULA TO Z

Before Poisson's formula is applied to Z, $G\hat{}$, $f\hat{}$, and $(G/H)\hat{}$ have to be computed.

G is identified with $Z_2^B = Z_2 \times \cdots \times Z_2$ (B times), so that $a \in G$ can be expressed as $a = (a_1, \cdots, a_B) \in Z_2^B$, and $b = (b_1, \cdots, b_B) \in G\hat{} \simeq Z_2^B$ acts on G according to the rule

$$b(a) = e^{\pi i b\cdot a},$$

the inner product $b\cdot a$ being computed in Z_2^B. Regarding $G\hat{}$ as the group of paths with coefficients from Z_2 on the dual lattice Q^*, $b\cdot a$ can be interpreted as the number of intersections of b and a, and $(G/H)\hat{}$ can be identified as the class of paths b on Q^* that intersect each *closed* path a on Q an *even* number of times; in brief, $(G/H)\hat{}$ is identified as the *bounding cycles* on Q^*. $f\hat{}$ is now computed as follows: $f(a) = \exp(BC^* - 2a\cdot 1C^*)$, $1 \in G\hat{}$ being the sum of *all* the bonds of the dual lattice Q^*, and so

$$f\hat{}(b) = \sum_{a\in G} b(a)f(a)$$
$$= e^{BC^*} \sum_{a\in Z_2^B} \exp(\pi i b\cdot a - 2a\cdot 1C^*)$$
$$= (e^{C^*} + e^{\pi i b_1 - C^*}) \cdots (e^{C^*} + e^{\pi i b_B - C^*})$$
$$= (\tfrac{1}{2}\sinh 2C)^{-\frac{1}{2}B} \exp[(e^{\pi i b_1} + \cdots + e^{\pi i b_B})C]$$
$$= (\tfrac{1}{2}\sinh 2C)^{-\frac{1}{2}B} e^{(B-n^*)C-n^*C},$$

$n^* = n^*(b)$ being the number of bonds (length) of the dual path $b \in G\hat{}$.

Poisson's formula is now applied to obtain

$$Z = 2^A(\tfrac{1}{2}\sinh 2C)^{\frac{1}{2}B} \sum_{a\in H} f(a)$$
$$= 2^A(\tfrac{1}{2}\sinh 2C)^{\frac{1}{2}B} \sum_{b\in (G/H)\hat{}} f\hat{}(b)$$
$$= 2^A \sum_{b\in (G/H)\hat{}} f^*(b)$$
$$= 2^A \sum_{a\in K_1} f^*(a).$$

f^* is the function f formed with C^* in place of C and $K_1 \simeq (G/H)\hat{}$ is the subgroup of G of *bounding* cycles on the original lattice. Replacing K_1 by H, one finds the desired duality:

$$Z(C) = (\tfrac{1}{2}\sinh 2C^*)^{-\frac{1}{2}B}2^A(\tfrac{1}{2}\sinh 2C^*)^{\frac{1}{2}B} \sum_{a\in H} f^*(a)$$
$$= (\tfrac{1}{2}\sinh 2C^*)^{-\frac{1}{2}B}Z(C^*).$$

Although this is not exact ($H/K_1 \simeq Z_2$ contains 4 members so that only $\frac{1}{4}$ of the closed paths have been accounted for), it becomes exact for an infinite lattice as noted above.

A Remarkable Connection Between Kemmer Algebras and Unitary Groups*

P. ROMAN
Department of Physics, Boston University, Boston, Massachusetts
(Received 10 December 1963)

Starting from a Kemmer algebra generated by n elements, a new algebra generated by $\frac{1}{2}(n^2 + n)$ elements is derived. This algebra has only one irreducible representation which is of dimension $(n + 1)$. This representation is equivalent to the first elementary representation of the infinitesimal operators of the unitary group in a space of $(n + 1)$ dimensions. Omitting the identity element, one of course obtains the first elementary representation of the unitary unimodular group in a space of $(n + 1)$ dimensions. The procedure is first illustrated for the case of $n = 3$ and explicit representations of the infinitesimal operators of U_4 (SU_4) are obtained. Next, the proof for arbitrary n is outlined. Finally, the degenerate case of U_3 (SU_3) is discussed.

INTRODUCTION

IN an earlier stage of development, considerable attention has been paid to the algebraic aspects of particle equations. The aim of these investigations was to associate particular algebras with the space–time properties of sets of elementary particles. Apart from the Dirac equation, the first success along these lines was scored by Kemmer,[1] who succeeded to give compact field equations related to both spin-zero and spin-one mesons. The method has been generalized in many directions, and eventually has become a postulational basis for the theory of elementary particles.[2-4]

In more recent years, emphasis shifted to the study of "internal" properties of particle families, like isospin and strangeness. In these investigations algebraic methods (mainly group-theoretic) played again a decisive role.[5] Very few and only tentative efforts have been directed, however, toward the systematic incorporation of these internal degrees of freedom into the algebra of field equations. One interesting example was given by Votruba and Lokajíček[6] who, notably, advocated the use of general Kemmer and Dirac algebras.[6a]

With the discovery of a bewilderingly large number of new "particles" or "resonances" and the concomitant interest in higher symmetries, especially unitary groups,[7] the question may be asked whether it is not timely and even possible to incorporate the structure of higher symmetries into the algebraic structure of field equations. The first step in such a program would be, clearly, to identify algebras having representations isomorphic to certain representations of Lie groups which may be considered as likely candidates for higher symmetry groups. One example of a connection between a "particle algebra" and a semisimple Lie group is long known[8]: the Dirac algebra generated by n elements furnishes the elementary spinor representation of the orthogonal group in n dimensions.

The present paper demonstrates a new kind of such a connection. We shall show that there exists a unique connection between the Kemmer algebra generated by n elements and the unitary group in a space of $n + 1$ dimensions. A precise formulation of this connection will be given in Sec. III, after the theorem and its proof has been elucidated in Sec. II on the very simple example furnished by $n=3$. The case of $n=2$ (leading to U_3) is degenerate and will be discussed separately in Sec. IV.

I. FUNDAMENTAL NOTIONS

We start by recalling a few facts about Kemmer algebras[9] $\mathcal{K}(n)$ generated by n elements β_i ($i = 1, 2, \cdots, n$). The generators are taken to satisfy the relations

$$\beta_i\beta_k\beta_l + \beta_l\beta_k\beta_i = \beta_i\delta_{kl} + \beta_l\delta_{ki}, \qquad (1)$$

and the algebra $\mathcal{K}(n)$ consists of all linearly independent elements that can be formed of the β_i and their multiple products. Including the identity element I,

* The research reported in this paper has been sponsored by the U. S. Air Force, Office of Scientific Research, through Grant No. AF-AFOSR-385-63.

[1] N. Kemmer, Proc. Roy. Soc. (London) **A 173**, 91 (1939).
[2] H. J. Bhabha, Rev. Mod. Phys. **21**, 451 (1949).
[3] I. M. Gelfand and A. M. Yaglom, Zh. Eksperim. i Teor. Fys. **18**, 703 (1948).
[4] For a review, see also P. Roman, *Theory of Elementary Particles* (North-Holland Publishing Company, Amsterdam, 1960), Chap. II, Secs. 2 and 3.
[5] For a general review, up to 1960, see for example, P. Roman, Ref. 4, Chap. V.
[6] V. Votruba and M. Lokajíček, Joint Inst. of Nucl. Res. Rept. P-191 (Dubna, 1958).
[6a] *Note added in proof.* A more sophisticated scheme has been recently suggested by H. C. Corben, Phys. Rev. **131**, 2219 (1963).
[7] Much of recent developments in this field is well reviewed in *Theoretical Physics* (Trieste Seminar) (International Atomic Energy Agency, Vienna, 1963), Part II.
[8] R. Brauer and H. Weyl, Am. J. Math. **57**, 425 (1935).
[9] N. Kemmer, Proc. Camb. Phil. Soc. **39**, 189 (1943).

the number of elements of $\mathcal{K}(n)$ is found to be

$$N = \frac{1}{2}\binom{2n+2}{n+1}. \tag{2}$$

The algebra is finite and semisimple. Therefore, as a consequence of the generalized Frobenius–Schur lemma,[10]

$$N = a_1^2 + a_2^2 + \cdots + a_r^2. \tag{2a}$$

Here r is the number of elements that commute with every member of $\mathcal{K}(n)$, and a_s is the dimensionality of the sth irreducible representation. Kemmer[9] has also shown that

$$\left.\begin{aligned} r &= \tfrac{1}{2}n + 1 \quad \text{if } n \text{ even},\\ r &= \tfrac{1}{2}(n+5) \quad \text{if } n \text{ odd}. \end{aligned}\right\} \tag{2b}$$

An explicit construction of the irreducible representations of the β_i was given by Littlewood.[11]

It will be important for us to note that, as seen from the above formulas, the algebra $\mathcal{K}(n)$ always has an irreducible representation of dimensionality $n+1$. (The only exception is the degenerate case of $n = 1$.)

For our purposes, it will be useful to spell out the basic Kemmer relations (1) in detail. If, for convenience, we introduce the shorthand notation

$$\eta_i = 2\beta_i^2 - I, \tag{3}$$

then we can write instead of (1)

$$\left.\begin{aligned} \beta_i^3 &= \beta_i,\\ \beta_i\beta_k^2 + \beta_k^2\beta_i &= \beta_i, \qquad (i \neq k),\\ \beta_i\beta_k\beta_l + \beta_i\beta_k\beta_i &= 0, \quad (i \neq k,\ k \neq l,\ l \neq i),\\ \beta_i\beta_k\beta_i &= 0, \qquad (i \neq k), \end{aligned}\right\} \tag{4a}$$

$$\left.\begin{aligned} \eta_i^2 &= I,\\ \eta_i\eta_k - \eta_k\eta_i &= 0,\\ \eta_i\beta_k + \beta_k\eta_i &= 0, \quad (i \neq k),\\ \eta_i\beta_i &= \beta_i\eta_i = \beta_i. \end{aligned}\right\} \tag{4b}$$

It is perhaps interesting to mention that peculiar connections between specific *representations* of certain $\mathcal{K}(n)$ algebras and specific representations of Lie groups have been noted occasionally. For example,[12] there is a relation between the combination of the two 10-dimensional representations of $\mathcal{K}(5)$ and the 20-dimensional irreducible representation of the full six-dimensional orthogonal group. More recently, Kursunoglu[13] found a connection between the three-dimensional representation of $\mathcal{K}(3)$ and the three-dimensional representation of U_3. Incidentally, one three-dimensional representation of the *generators* of $\mathcal{K}(3)$ was also used in connection with the vectorial representation of the three-dimensional rotation group,[14] and further an algebra was derived from this representation which plays the same role in three dimensions as the Pauli algebra does in two dimensions.[15–15a]

To conclude this Section, we list some relevant properties of unitary groups and their representations.[16] The unitary group U_n can be factored as

$$U_n = U_1 \times SU_n, \tag{5}$$

where SU_n, the unitary unimodular group in n dimensions is a simple Lie group.[17] Hence, it will suffice to concentrate on SU_n. This group has $n^2 - 1$ real parameters and of course the same number of infinitesimal operators. The number of mutually commuting infinitesimal operators is $n-1$. All irreducible representations of SU_n can be obtained from only two, so-called *elementary representations*. The first elementary representation is realized by a set of all $n^2 - 1$ linearly independent *traceless* matrices of order n. That is, denoting the infinitesimal operators in this representation by I_i, we have

$$\left.\begin{aligned} &\sum_{i=1}^{n^2-1} c_i I_i \neq 0 \quad \text{unless} \quad c_i \equiv 0,\\ &\qquad \operatorname{Tr} I_i = 0. \end{aligned}\right\} \tag{6}$$

The second elementary representation of SU_n is obtained from the first by simply taking

$$I_i' = -\bar{I}_i. \tag{7}$$

All other irreducible representations are then obtained by taking first the direct products of the two elementary representations and isolating the irreducible component of highest weight, and then proceeding in the same manner.

If one wishes to obtain an irreducible representation of dimension d for U_n then, in view of (5), one merely has to adjoin the unit matrix of order d

[10] See, for example, N. Kemmer, Ref. 1, or P. Roman, Ref. 4, p. 35.

[11] D. E. Littlewood, Proc. Camb. Phil. Soc. **43**, 406 (1947).

[12] E. M. Corson, *Introduction to Tensors, Spinors, and Relativistic Wave Equations* (Blackie & Son Ltd., London, 1953), p. 185.

[13] B. Kursunoglu, Proceedings of the Second Eastern Theoretical Conference (University of North Carolina, Chapel Hill, 1963). See also J. Math. Phys. **2**, 22 (1961).

[14] V. Votruba, Phys. Rev. **85**, 141 (1951).

[15] P. Roman, Proc. Phys. Soc. (London) **74**, 649 (1959).

[15a] *Note added in proof.* Quite recently, a general connection between $\mathcal{K}(n)$ and the Lie algebra of rotation groups has been established, cf. C. Ryan and E. C. G. Sudarshan, Nucl. Phys. **47**, 207 (1963).

[16] A thorough treatise on Lie groups was given, for example, by G. Racah, CERN 61-8. A shorter, more recent review available is by A. Salam, in Ref. 7, p. 173.

[17] To avoid confusion, we should mention that in Cartan's classical classification scheme SU_n is denoted by A_{n-1}.

to the $n^2 - 1$ matrices of order d that provide the d-dimensional representation of SU_n.

Based on various philosophies, simple Lie groups and in particular SU_3, recently attracted tremendous interest.[18] Of course, SU_2, being locally isomorphic with the three-dimensional real rotation group, was essentially the basis of the less ambitious isobaric formalism. SU_3 has no such geometrical interpretation, but by accident, SU_4 happens to be locally isomorphic to the six-dimensional real rotation group. It may be that this circumstance has physical implications, so we now establish a new relation between SU_4 and $\mathcal{K}(3)$.

II. RELATION BETWEEN $\mathcal{K}(3)$ AND SU_4

Let us consider $\mathcal{K}(3)$, generated by β_1, β_2, and β_3, and let us introduce in place of these three new generators defined by

$$S_i = -i\epsilon_{ikl}\beta_k\beta_l. \tag{8}$$

Here ϵ_{ikl} is the completely antisymmetric Levi-Civita symbol with three indices, and the usual summation convention is employed. With the help of (4a, b) one easily verifies that the S_i also obey the basic Kemmer relations (1), i.e.,

$$S_iS_kS_l + S_lS_kS_i = S_i\delta_{kl} + S_l\delta_{ki}. \tag{9}$$

In addition, they also fulfill the commutator relations

$$[S_i, S_k] = i\epsilon_{ikl}S_l. \tag{10}$$

Next, let us introduce three further generators defined by

$$R_{ik} = (i/2\sqrt{2})[\eta_i, \beta_k] + (i/2\sqrt{2})[\eta_k, \beta_i], \qquad (i \neq k), \tag{11}$$

where η_i is given by (3). Using the relations (4a, b) one finds that these new generators satisfy the commutation law

$$[R_{ik}, R_{lm}] = \tfrac{1}{2}i(\epsilon_{ila} + \epsilon_{ima} + \epsilon_{kla} + \epsilon_{kma})S_a. \tag{12}$$

In addition, one obtains also the relations

$$[R_{ik}, S_m]S_l + [R_{ik}, S_l]S_m = R_{ik}\{S_l, S_m\}. \tag{13}$$

We now build up an algebra from the generators S_i and R_{ik}. As a consequence of the relations (9), (10), (12), and (13), the algebra will have much fewer elements than the original $\mathcal{K}(3)$ algebra where the generators were restricted only by (1). While, according to (2), $\mathcal{K}(3)$ has 35 elements, we now find that the new algebra generated by S_i and R_{ik} has only 16 elements. To write down these independent elements in an economical way, we first introduce the convenient abbreviation

$$\xi_i \equiv (1/\sqrt{2})(2S_i^2 - I). \tag{14}$$

We then find that the linearly independent elements of this new algebra, to be called henceforth the $\mathcal{R}(3)$ algebra, can be listed as follows:

Element	Comment	Number
I		1
S_i		3
$\{S_i, S_k\}$	$i \neq k$	3
ξ_i		3
R_{ik}	$i \neq k$	3
$i\sqrt{2}\,\xi_iR_{kl}$	$(ikl) = (123)$ cyclic	3
	Total:	16

We can, further, establish that the only element of $\mathcal{R}(3)$ which commutes with all other elements, is the identity I, and that $\mathcal{R}(3)$ does not contain any invariant Abelian subalgebra. Hence, according to the Frobenius–Schur lemma, Eq. (2a), the algebra $\mathcal{R}(3)$ has only one irreducible representation which must then be of dimension four.

This representation is easily written out explicitly if we utilize the well-known four-dimensional representation of $\mathcal{K}(3)$. We have[19]

$$\beta_1 = \begin{pmatrix} 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}, \quad \beta_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{pmatrix}, \quad \beta_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & -i & 0 \end{pmatrix};$$

so that, using our construction, we obtain

$$\rho_0 \equiv I = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \rho_1 \equiv S_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \rho_2 \equiv S_2 = \begin{pmatrix} 0 & 0 & i & 0 \\ 0 & 0 & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \rho_3 \equiv S_3 = \begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

[18] The three classical papers are by M. Gell-Mann, Phys. Rev. **125**, 1067 (1962); Y. Ne'eman, Nucl. Phys. **26**, 222 (1961); and M. Ikeda, S. Oyawa, and Y. Ohnuki, Progr. Theor. Phys. (Kyoto), **22**, 715 (1959). Some of the further developments are reviewed in Ref. 7.

[19] See, for example, P. Roman, Ref. 4, p. 428.

$$\rho_4 \equiv \{S_1, S_2\} = \begin{bmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \rho_5 \equiv \{S_2, S_3\} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \rho_6 \equiv \{S_1, S_3\} = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\rho_7 \equiv \xi_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}, \quad \rho_8 \equiv \xi_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}, \quad \rho_9 \equiv \xi_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}, \tag{15}$$

$$\rho_{10} \equiv R_{12} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}, \quad \rho_{11} \equiv R_{23} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}, \quad \rho_{12} \equiv R_{13} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix},$$

$$\rho_{13} \equiv i\sqrt{2}\,\xi_1 R_{23} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & 0 & i \\ 0 & -i & -i & 0 \end{bmatrix}, \quad \rho_{14} \equiv i\sqrt{2}\,\xi_2 R_{31} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ -i & 0 & -i & 0 \end{bmatrix},$$

$$\rho_{15} \equiv i\sqrt{2}\,\xi_3 R_{12} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 0 & 0 & i \\ 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ -i & -i & 0 & 0 \end{bmatrix}.$$

We see that

$$\mathrm{Tr}\ \rho_i = 0 \qquad (i \neq 0). \tag{16}$$

Thus, we have here *a set of* 15 *linearly independent traceless matrices of dimension four*. According to Sec. I, Eq. (6), *they provide the first elementary representation of* SU_4, which was the theorem we wanted to demonstrate.

In addition, in consequence of our choice of numerical factors, we have the convenient normalization

$$\mathrm{Tr}\ \rho_i \rho_k = 2\delta_{ik} \qquad (i, k \neq 0). \tag{17}$$

It is seen that the three mutually commuting generators of SU_4 are given by ξ_1, ξ_2, ξ_3. In our representation, they are already diagonal.

Counting also the identity $\rho_0 \equiv I$, we have, of course, a representation of U_4.

The results of this Section can be generalized for arbitrary $n > 3$. In fact, the few last statements which we made above by utilizing the explicit representation, can also be arrived at in general. We now turn to this problem.

III. THE GENERAL CASE

We consider the Kemmer algebra $\mathfrak{K}(n)$ generated by $\beta_1, \beta_2, \cdots, \beta_n$. Let us introduce n new generators by setting

$$S_i = -if_{ikl}\beta_k\beta_l, \tag{18}$$

where f_{ikl} is real and completely antisymmetric in all indices. For $n > 3$ this requirement does not determine, of course, the relative magnitudes of the $n!/3!(n-3)!$ independent coefficients. Therefore, we set the additional requirement

$$-\tfrac{1}{2}(n+1)f_{ikl} = f_{klm}f_{iab}f_{mcd}\,\mathrm{Tr}\,\beta_a\beta_b\beta_c\beta_d. \tag{19}$$

From the symmetry properties of product traces and the antisymmetry of the f's, it is not difficult to show that, apart from the trivial solution $f_{ikl} \equiv 0$ (any combination of i, k, l), the system (19) has essentially *one* solution only: all other solutions arise from a particular one by permuting indices.[20]

Using (18), (19), and (4a, b), one then can show, with considerable labor, that the generators S_i satisfy Kemmer relations with a changed metric. One finds

$$S_iS_kS_l + S_lS_kS_i = \tfrac{1}{4}(n+1)(S_i\delta_{kl} + S_l\delta_{ki}). \tag{20}$$

In addition, they satisfy the commutation relations

$$[S_i, S_k] = if_{ikl}S_l. \tag{21}$$

Next, we define $\frac{1}{2}(n^2 - n)$ further generators by setting

$$R_{ik} = (n-1)^{-\frac{1}{2}}\tfrac{1}{2}i([\eta_i, \beta_k] + [\eta_k, \beta_i]) \quad (i \neq k). \tag{22}$$

They are found to satisfy the relations

[20] Useful formulas for the traces of β products can be obtained from Littlewood's paper, Ref. 11. When actually solving Eqs. (19), care must be taken not to use any specific irreducible representation of the β_i, because in some of these multiple β products might vanish. The full, reducible representation of degree $\sum a_r$ must be utilized.

$$[R_{ik}, R_{lm}]$$
$$= [i/(n-1)](f_{ila} + f_{ima} + f_{kla} + f_{kma})S_a, \quad (23)$$

and also

$$[R_{ik}, S_m]S_l + [R_{ik}, S_l]S_m = R_{ik}\{S_l, S_m\}. \quad (24)$$

We are now prepared to construct the complete algebra from the generators S_i and R_{ik}. We find that the number of linearly independent elements of this algebra, to be called the $\mathfrak{R}(n)$ algebra, becomes

$$N' = (n+1)^2, \quad (25)$$

considerably less than the number of elements of $\mathfrak{K}(n)$, given by (2). If we introduce, for brevity, the notation

$$\xi_i \equiv (n-1)^{-\frac{1}{2}}(2S_i^2 - I), \quad (26)$$

then we can list the elements of $\mathfrak{R}(n)$ as follows:

Element	Comment	Number
I		1
S_i		n
$(2/n-1)^{\frac{1}{2}}\{S_i, S_k\}$	$i \neq k$	$\frac{1}{2}(n^2-n)$
ξ_i		n
R_{ik}	$i \neq k$	$\frac{1}{2}(n^2-n)$
$i(n-1)^{\frac{1}{2}}\xi_i R_{kl}$	$(ikl) = (123)$ cycl.	n
	Total:	$(n+1)^2$

(27)

Our next task is to determine the traces. To start with, from (4b) we obtain

$$\beta_k\beta_l = \beta_k\beta_l\eta_l = [\beta_k\beta_l, \eta_l].$$

Since the trace of a commutator vanishes, the definition (18) immediately yields

$$\mathrm{Tr}\, S_i = 0. \quad (28)$$

To continue, we first must make a little detour and determine $\mathrm{Tr}\, S_iS_k$. Multiplying (19) by $f_{rst}\, \mathrm{Tr}\, \beta_k\beta_l\beta_s\beta_t$ and using (18), we find that

$$\tfrac{1}{2}(n+1)\, \mathrm{Tr}\, S_iS_r = \mathrm{Tr}\, S_{im}\, \mathrm{Tr}\, S_{mr}.$$

The solution of this equation is

$$\mathrm{Tr}\, S_iS_k = \tfrac{1}{2}(n+1)\delta_{ik}. \quad (29)$$

Then it immediately follows that the traces of the elements in the third row of the array (27) vanish.

We proceed to the next row and find,[21] with (26) and (29),

$$\mathrm{Tr}\, \xi_i = (n-1)^{-\frac{1}{2}}\, \mathrm{Tr}\, (2S_i^2 - I) = 0. \quad (30)$$

[21] The $\mathfrak{R}(n)$ algebra (27), consisting of $(n+1)^2$ linearly independent elements, certainly possesses an $(n+1)$-dimensional representation. In this representation $\mathrm{Tr}\, I = n+1$, and (30) follows. We shall see shortly that this representation is the only one. Alternatively, we could have first established directly that $\mathfrak{R}(n)$ is semisimple.

Proceeding to the next row of (27), it is obvious from the definition (22) that $\mathrm{Tr}\, R_{ik} = 0$.

Finally, to show that the trace of the elements in the last row vanishes, we first establish the lemma

$$[S_i^2, S_j^2] = 0. \quad (31)$$

This can be seen by writing

$$[S_i^2, S_j^2] = \{S_i, \{S_j, [S_i, S_j]\}\}$$
$$= if_{ijk}\{S_i, \{S_j, S_k\}\} = 0,$$

where in the last but one step (21) was taken into account. Furthermore, S_i (and thus S_i^2) is Hermitian as follows from (18), the reality of f_{ikl}, and the fact that the β_i can always be chosen Hermitian. Then we know that the n elements S_i^2 can be simultaneously diagonalized. Now we are prepared to determine the trace of the remaining elements. We have

$$\mathrm{Tr}\, (i(n-1)^{\frac{1}{2}}\xi_i R_{kl}) = i\, \mathrm{Tr}\, (2S_i^2 R_{kl} - R_{kl})$$
$$= 2i \sum_\alpha s_\alpha^i (R_{kl})_{\alpha\alpha}, \quad (32)$$

where we denoted by s_α^i the (diagonal) elements of S_i^2. On the other hand,

$$R_{kl} \sim [\eta_k, \beta_l] + [\eta_l, \beta_k],$$

and it follows from (3) and (4b) that the η_i elements can be all diagonalized simultaneously. Consequently,

$$(R_{kl})_{\alpha\alpha} = 0,$$

so that the right side of (32) vanishes, Q.E.D.

Thus, we have shown that the traces of all elements of $\mathfrak{R}(n)$ vanish (excepting, of course, I). In addition, by utilizing repeatedly the relations (20), (21), (23), (24), and employing techniques similar to those adopted above, we also find, in a tedious way, that

$$\mathrm{Tr}\, \rho_i\rho_k = \frac{n+1}{2}\delta_{ik}. \quad (33)$$

Here we have used the generic symbol ρ_i for an arbitrary element of $\mathfrak{R}(n)$, excepting I.

It is now easy to establish that $\mathfrak{R}(n)$ is semisimple. We use the criterion quoted by Kemmer,[9] namely that we must have

$$\det G \neq 0, \quad (34)$$

where

$$G \equiv (G_{ik}) = \mathrm{Tr}\, (\rho_i\rho_k).$$

In consequence of (33), criterion (34) is obviously fulfilled.

Finally, it is not difficult to establish that the only element of $\mathfrak{R}(n)$ which commutes with all

other elements is the identity I. Thus, the Frobenius–Schur lemma tells us that $\mathfrak{R}(n)$ has precisely one irreducible representation which is of the dimension $(n+1)$. The n^2+2n matrices ρ_i of this representation we have shown to have vanishing trace; hence, in view of what was said in Sec. I, in connection with Eq. (6), *these matrices furnish the first elementary representation of* SU_{n+1}.

The n mutually commuting elements of the representation are precisely our operators ξ_i. For, we have by the definition (26) and the lemma (31)

$$[\xi_i, \xi_j] = [4/(n-1)][S_i^2, S_j^2] = 0. \tag{35}$$

They are Hermitian and can be diagonalized simultaneously.

In summary, our main result is the following. *From any Kemmer algebra* $\mathfrak{K}(n)$ *one can derive uniquely an* $\mathfrak{R}(n)$ *algebra, spelled out by Eq.* (27), *which has only one irreducible representation. The matrices of this representation yield the first elementary representation of* U_{n+1}.

IV. THE DEGENERATE CASE OF SU_3

The simplest unitary group, U_2 (or SU_2) cannot be associated with any Kemmer algebra. The reason for this is that $\mathfrak{K}(1)$ does not really exist, it has only the trivial one-dimensional representations. Or, in other words: there is no Kemmer algebra which has a two-dimensional representation.

Even the next case, U_3 (or SU_3) is degenerate. This is so because one cannot form a completely antisymmetric symbol f_{ikl} when the indices are capable of only two values 1 and 2. Thus, our construction of Sec. III cannot be applied in the literal sense. Nevertheless, we can give in this case an *ad hoc* construction.

We simply start with the two generators β_1 and β_2. According to the general method of Littlewood,[11] the only nontrivial irreducible representation is

$$\beta_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \beta_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}. \tag{36}$$

For convenience, we adjust a normalization constant and set

$$S_i = \tfrac{1}{2}\sqrt{3}\,\beta_i. \tag{37}$$

We find that these new generators satisfy the general relations (20) with $n = 2$. Equation (21) has no analog, of course. Next we introduce *two* more generators

$$R_i = \tfrac{1}{2}i(\tfrac{1}{2}\sqrt{3})[\eta_i, \beta_k] \qquad (i \neq k). \tag{38}$$

(This is a further departure from the general scheme.) Also, there are no direct analogs of (23) and (24). Nevertheless, when we now try to build up an algebra of S_i and R_i, we find that there are only nine linearly independent elements as contrasted with the ten elements of $\mathfrak{K}(2)$. To list the nine elements, we introduce, exactly as in the general case, the ξ_i as defined by (26). On the other hand, since (21) has no analog, we have to list $[S_1, S_2]$ as an independent element. But R_1 and R_2, when multiplied by ξ_i, do not give new elements. Thus the algebra $\mathfrak{R}(2)$ is as follows:

Element	Comment	Number
I		1
S_i		2
$(2/\sqrt{3})\{S_i, S_k\}$	$i \neq k$	1
$(2/\sqrt{3})i[S_i, S_k]$	$i \neq k$	1
ξ_i		2
R_i		2
	Total:	9

The numerical coefficients have been chosen in such a way as to ensure the uniform normalization as given by (33) which now reads

$$\mathrm{Tr}\ \rho_i\rho_k = \tfrac{3}{2}\delta_{ik}. \tag{39}$$

Thus, exactly as in the general case, it is seen that the algebra is semisimple; hence, it has only one irreducible representation which is of dimension three, and it is precisely the one furnished by the explicit forms (36). We find, in fact (apart from $\rho_0 \equiv I$),

$$\rho_1 \equiv S_1 = \frac{\sqrt{3}}{2}\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \rho_2 \equiv S_2 = \frac{\sqrt{3}}{2}\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad \rho_3 \equiv \frac{2}{\sqrt{3}}\{S_1, S_2\} = \frac{\sqrt{3}}{2}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix},$$

$$\rho_4 \equiv \frac{2}{\sqrt{3}}i[S_1, S_2] = \frac{\sqrt{3}}{2}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & i \\ 0 & -i & 0 \end{bmatrix}, \quad \rho_5 \equiv \xi_1 = \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}, \quad \rho_6 \equiv \xi_2 = \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{40}$$

$$\rho_7 \equiv R_1 = \frac{\sqrt{3}}{2}\begin{bmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{bmatrix}, \quad \rho_8 \equiv R_2 = \frac{\sqrt{3}}{2}\begin{bmatrix} 0 & i & 0 \\ -i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

These matrices are all manifestly traceless and hence provide the first elementary representation of SU_3. The representation used by Gell-Mann[18] differs from (40) firstly, by a common normalization factor and secondly, insofar as his λ_3 is a linear combination of our ξ_1 and ξ_2.

One might speculate that, since the connection between SU_3 and $\mathcal{K}(2)$ does not follow precisely the pattern valid for $n \geq 3$, the physical content of SU_3 is not sufficiently general. It would be, therefore, interesting to follow up a scheme based on SU_4. Investigations in this direction are in progress.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 5, NUMBER 6 JUNE 1964

Theory of Scattering in Solids*

JOSEPH CALLAWAY
Department of Physics, University of California, Riverside, California
(Received 10 December 1963)

The general theory of the scattering of excitations in solids by localized imperfections is discussed. The solid-state analog of the usual partial-wave expansion of the scattering amplitude is derived. In an appendix, the applicability of the general theory to phonons and spin waves as well as electrons is demonstrated.

I. INTRODUCTION

IN this paper, we discuss the scattering of wavelike excitations in solids by localized imperfections. This situation is analagous to potential scattering in ordinary quantum mechanics, and it is interesting to see the extent to which the familiar mathematical description is applicable. The present analysis may be applied to the scattering of phonons by mass defects,[1-11] spin waves by magnetic defects,[12-13] or electrons by foreign atoms.[14-18] Most of the discussion will be conducted using the terminology of energy-band theory, but in an appendix it will be shown how, with slight modifications of notation, the same basic equation may be used for the different situations. This paper contains approximately equal portions of review and original material. An attempt has been made to make it largely self-contained.

In contrast to the situation in nuclear or elementary-particle physics, scattering cross sections in solids are not directly observable quantities. They are, however, important inasmuch as a mean free path for the excitation can be determined from them if the concentration of imperfections is small enough so that multiple scattering can be neglected. It is readily shown that

$$l_D^{-1} = N_D \int \frac{d\sigma}{d\Omega} (1 - \cos\theta)\, d\Omega,$$

where l_D is the free path, N_D is the concentration of imperfections, and $d\sigma/d\Omega$ is the differential cross section.[19] Knowledge of the mean free path is important for transport calculations, and under some circumstances, experimental information concerning transport properties can be interpreted to obtain information concerning the cross sections of interest. In addition, an impurity potential of sufficient strength will cause localized modes to split off from the continuum of levels characterizing the perfect

* This work was supported by the Air Force Office of Scientific Research.

[1] I. M. Lifshitz, J. Phys. (U.S.S.R.) **7**, 211, 249 (1943).
[2] I. M. Lifshitz, Nuovo Cimento Suppl. **3**, 716 (1956).
[3] M. Lax, Phys. Rev. **94**, 1391 (1954).
[4] E. B. Montroll and R. B. Potts, Phys. Rev. **100**, 525 (1955), **102**, 72 (1956).
[5] A. A. Maradudin, P. Mazur, E. W. Montroll, and G. H. Weiss, Rev. Mod. Phys. **30**, 175 (1958).
[6] R. Brout and W. Visscher, Phys. Rev. Letters **9**, 54 (1962).
[7] W. M. Visscher, Phys. Rev. **129**, 28 (1963).
[8] J. Callaway, Nuovo Cimento **29**, 883 (1963).
[9] M. V. Klein, Phys. Rev. **131**, 1500 (1963).
[10] M. Wagner, Phys. Rev. **133**, A750 (1964).
[11] A. A. Maradudin, E. W. Montroll, and G. H. Weiss in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1963), Supplement 3.
[12] T. Wolfram and J. Callaway, Phys. Rev. **130**, 2207 (1963).
[13] J. Callaway, Phys. Rev. **132**, 2003 (1963).
[14] G. F. Koster and J. C. Slater, Phys. Rev. **95**, 1167 (1954).
[15] G. F. Koster, Phys. Rev. **95**, 1436 (1954).
[16] G. F. Koster and J. C. Slater, Phys. Rev. **96**, 1208 (1954).
[17] A. M. Clogston, Phys. Rev. **125**, 439 (1962).
[18] P. A. Wolff, Phys. Rev. **124**, 1030 (1961).
[19] J. M. Ziman, *Electrons and Phonons* (Oxford University Press, Oxford, England 1960), p. 306.

These matrices are all manifestly traceless and hence provide the first elementary representation of SU_3. The representation used by Gell-Mann[18] differs from (40) firstly, by a common normalization factor and secondly, insofar as his λ_3 is a linear combination of our ξ_1 and ξ_2.

One might speculate that, since the connection between SU_3 and $\mathcal{K}(2)$ does not follow precisely the pattern valid for $n \geq 3$, the physical content of SU_3 is not sufficiently general. It would be, therefore, interesting to follow up a scheme based on SU_4. Investigations in this direction are in progress.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 5, NUMBER 6 JUNE 1964

Theory of Scattering in Solids*

JOSEPH CALLAWAY
Department of Physics, University of California, Riverside, California
(Received 10 December 1963)

The general theory of the scattering of excitations in solids by localized imperfections is discussed. The solid-state analog of the usual partial-wave expansion of the scattering amplitude is derived. In an appendix, the applicability of the general theory to phonons and spin waves as well as electrons is demonstrated.

I. INTRODUCTION

IN this paper, we discuss the scattering of wavelike excitations in solids by localized imperfections. This situation is analagous to potential scattering in ordinary quantum mechanics, and it is interesting to see the extent to which the familiar mathematical description is applicable. The present analysis may be applied to the scattering of phonons by mass defects,[1-11] spin waves by magnetic defects,[12-13] or electrons by foreign atoms.[14-18] Most of the discussion will be conducted using the terminology of energy-band theory, but in an appendix it will be shown how, with slight modifications of notation, the same basic equation may be used for the different situations. This paper contains approximately equal portions of review and original material. An attempt has been made to make it largely self-contained.

In contrast to the situation in nuclear or elementary-particle physics, scattering cross sections in solids are not directly observable quantities. They are, however, important inasmuch as a mean free path for the excitation can be determined from them if the concentration of imperfections is small enough so that multiple scattering can be neglected. It is readily shown that

$$l_D^{-1} = N_D \int \frac{d\sigma}{d\Omega} (1 - \cos\theta)\, d\Omega,$$

where l_D is the free path, N_D is the concentration of imperfections, and $d\sigma/d\Omega$ is the differential cross section.[19] Knowledge of the mean free path is important for transport calculations, and under some circumstances, experimental information concerning transport properties can be interpreted to obtain information concerning the cross sections of interest. In addition, an impurity potential of sufficient strength will cause localized modes to split off from the continuum of levels characterizing the perfect

* This work was supported by the Air Force Office of Scientific Research.

[1] I. M. Lifshitz, J. Phys. (U.S.S.R.) **7**, 211, 249 (1943).
[2] I. M. Lifshitz, Nuovo Cimento Suppl. **3**, 716 (1956).
[3] M. Lax, Phys. Rev. **94**, 1391 (1954).
[4] E. B. Montroll and R. B. Potts, Phys. Rev. **100**, 525 (1955), **102**, 72 (1956).
[5] A. A. Maradudin, P. Mazur, E. W. Montroll, and G. H. Weiss, Rev. Mod. Phys. **30**, 175 (1958).
[6] R. Brout and W. Visscher, Phys. Rev. Letters **9**, 54 (1962).
[7] W. M. Visscher, Phys. Rev. **129**, 28 (1963).
[8] J. Callaway, Nuovo Cimento **29**, 883 (1963).
[9] M. V. Klein, Phys. Rev. **131**, 1500 (1963).
[10] M. Wagner, Phys. Rev. **133**, A750 (1964).
[11] A. A. Maradudin, E. W. Montroll, and G. H. Weiss in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1963), Supplement 3.
[12] T. Wolfram and J. Callaway, Phys. Rev. **130**, 2207 (1963).
[13] J. Callaway, Phys. Rev. **132**, 2003 (1963).
[14] G. F. Koster and J. C. Slater, Phys. Rev. **95**, 1167 (1954).
[15] G. F. Koster, Phys. Rev. **95**, 1436 (1954).
[16] G. F. Koster and J. C. Slater, Phys. Rev. **96**, 1208 (1954).
[17] A. M. Clogston, Phys. Rev. **125**, 439 (1962).
[18] P. A. Wolff, Phys. Rev. **124**, 1030 (1961).
[19] J. M. Ziman, *Electrons and Phonons* (Oxford University Press, Oxford, England 1960), p. 306.

crystal. The energies of these modes can be found in a natural way from the scattering calculation.

In this paper, we consider neither the application of the cross sections in transport theory, nor the influence of defects on the thermodynamic properties of the crystal. Attention is focused entirely on the calculation of the scattering amplitudes and of the energies of localized states. In addition, general formal properties such as the existence and unitarity of the S matrix which cannot differ in the solid-state problem will not be discussed. Emphasis will be given to the derivation and interpretation of the solid-state analog of the partial-wave formula of ordinary scattering theory.

The theory may be outlined in a formal way quite simply, One has an effective Hamiltonian, H_0, which characterizes the perfect crystal, and a localized perturbation V which describes the defect. The eigenfunctions of H_0 are running waves characterized by a continuous vector parameter $\mathbf{k}$ on account of Bloch's theorem, and a discrete band or polarization index s. We denote them by $\psi_s(\mathbf{k}, \mathbf{r})$:

$$H_0\psi_s(\mathbf{k}, \mathbf{r}) = E_s(\mathbf{k})\psi_s(\mathbf{k}, \mathbf{r}). \tag{1}$$

The full Schrödinger equation is then of the form

$$(H_0 + V)\Psi = E\Psi. \tag{2}$$

Our discussion here is conducted in the language appropriate to energy-band theory. We will see in Appendix A how it is also to be applied in the case of lattice vibrations or of spin waves. In order to make use of the fact that the perturbation is localized, it is convenient to introduce a set of localized functions $a_s(\mathbf{r} - \mathbf{R}_n)$ (Wannier functions),

$$a_s(\mathbf{r} - \mathbf{R}_n) = \frac{\Omega^{\frac{1}{2}}}{(2\pi)^3}\int e^{-i\mathbf{k}\cdot\mathbf{R}_n}\psi_s(\mathbf{k}, \mathbf{r})\, d^3k. \tag{3}$$

In this equation, $\mathbf{R}_n$ is a lattice vector, Ω is the volume of the unit cell (for simplicity we consider a crystal containing only one atom per cell), and the integration over $\mathbf{k}$ includes a single Brillouin zone.

The a_s can easily be shown to be a complete orthonormal set if the ψ_s are. The function Ψ which is a solution of (2) is expanded in the a_s,

$$\Psi(\mathbf{r}) = \sum_{n,s} B_s(\mathbf{R}_n)a_s(\mathbf{r} - \mathbf{R}_n). \tag{4}$$

Then it is easily seen that the coefficients B satisfy the equation

$$\sum_n (s, m\,|E - H_0|\, s, n)B_s(\mathbf{R}_n) = \sum_{n,t} (s, m\,|V|\, t, n)B_t(\mathbf{R}_n). \tag{5}$$

The matrix elements of any quantity Q on the basis of Wannier functions are denoted by

$$(s, m\,|Q|\, t, n) = \int a_s^*(\mathbf{r} - \mathbf{R}_m)Qa_t(\mathbf{r} - \mathbf{R}_n)\, d^3r \tag{6}$$

in which the integration includes the entire crystal. If one substitutes (3) into (6), it is easy to show that the matrix element of H_0 depends only on the difference of the site vectors, i.e., $\mathbf{R}_m - \mathbf{R}_n$. In addition, H_0 is diagonal with respect to the band index.

We now introduce the matrix inverse to the left side of (5)

$$\mathcal{G}_s(\mathbf{R}_m - \mathbf{R}_n) = (s, m\,|(E - H_0)^{-1}|\, s, n), \tag{7a}$$

which is conveniently described as a Green's function. The Green's function satisfies the equation

$$\sum_n (s, m\,|E - H_0|\, s, n)\mathcal{G}_s(\mathbf{R}_n - \mathbf{R}_l) = \delta_{m,l}. \tag{7b}$$

If the energy, E, coincides with one of the eigenvalues of H_0, there is a singular point in the integral, and it is necessary to specify how the singularity is to be treated. In general, we have the option of obtaining outgoing, incoming, or standing waves. In the scattering problem, we wish to consider outgoing waves, and therefore (7) is evaluated with the usual device in which the energy, E, is allowed to have an infinitesimal, positive, imaginary part which is set equal to zero after the integration is performed. It is then possible to rewrite (5) as follows:

$$B_s(\mathbf{R}_m) = B_s^{(0)}(\mathbf{R}_m) + \sum_{t,l,n} \mathcal{G}_s(\mathbf{R}_m - \mathbf{R}_l)(s, l\,|V|\, t, n)B_t(\mathbf{R}_n), \tag{8}$$

in which $B_s^{(0)}$ is a solution of the homogeneous equation for energy E,

$$\sum_n (s, m\,|E - H_0|\, s, n)B_s^{(0)}(\mathbf{R}_n) = 0. \tag{9}$$

Equation (8) is analogous to the Lippman–Schwinger equation of formal scattering theory.[20] It is easily shown with the use of the definition of the Wannier function that

$$B_s^{(0)}(\mathbf{R}_m) = [\Omega^{\frac{1}{2}}/(2\pi)^3]e^{i\mathbf{k}_0\cdot\mathbf{R}_m}, \tag{10}$$

in which the wave is incident in band s, and $\mathbf{k}_0$ is a solution of

$$E_s(\mathbf{k}_0) = E. \tag{11}$$

[20] B. A. Lippmann and J. Schwinger, Phys. Rev. **79**, 469 (1963).

It is apparent that Eq. (11) does not determine a unique $\mathbf{k}_0$ since the direction of $\mathbf{k}$ is largely arbitrary. Of course if E is such that no solution of (11) exists, the first term in (8) must be set equal to zero. If Eq. (10) is substituted into (4), one finds that $\Psi(\mathbf{r})$ is then just a Bloch function $\psi_s(\mathbf{k}_0, r)$.

II. THE KOSTER-SLATER MODEL

We now wish to investigate the solution of (8). A particularly simple case exists if the impurity potential extends over only one lattice site and in addition, is diagonal in the bands. In this situation,

$$(s, m\,|V|\,t, n) = V_0\delta_{s,t}\delta_{m,0}\delta_{n,0}. \tag{12}$$

We have placed the defect at the origin. V_0 may be a function of energy. This simple problem is actually of some practical importance, since it approximately describes an impurity of different mass (isotope scattering) in the lattice-vibration problem. We will return to a discussion of a more general potential after having analyzed this example.

Use of Eq. (12) enables (8) to be simplified to the form

$$B_s(\mathbf{R}_m) = B_s^{(0)}(\mathbf{R}_m) + V_0\mathcal{G}_s(\mathbf{R}_m)B_s(0). \tag{13}$$

This equation may be solved for $B(0)$ (since this potential is diagonal in the bands, we drop the band index),

$$B(0) = \frac{B^{(0)}(0)}{1 - V_0\mathcal{G}(0)} = \frac{\Omega^{\frac{1}{2}}}{(2\pi)^{\frac{3}{2}}}\frac{1}{1 - V_0\mathcal{G}(0)}. \tag{14}$$

This result is substituted into (13),

$$B(\mathbf{R}_m) = \frac{\Omega^{\frac{1}{2}}}{(2\pi)^{\frac{3}{2}}}\left[e^{i\mathbf{k}_0\cdot\mathbf{R}_m} + \frac{V_0\mathcal{G}(\mathbf{R}_m)}{1 - V_0\mathcal{G}(0)}\right]. \tag{15}$$

We must evaluate the Green's functions to proceed further. In the scattering problem we need $\mathcal{G}(\mathbf{R}_m)$ only for large values of $\mathbf{R}_m$. In general, we have from Eq. (7)

$$\begin{aligned}\mathcal{G}(\mathbf{R}_m - \mathbf{R}_n) &= \int a^*(\mathbf{r} - \mathbf{R}_m)(E - H_0)^{-1}a(\mathbf{r} - \mathbf{R}_n)\,d^3r \\ &= \frac{\Omega}{(2\pi)^3}\iiint e^{i\mathbf{k}\cdot\mathbf{R}_m}\psi^*(\mathbf{k}, \mathbf{r})(E - H_0)^{-1} \\ &\quad\times e^{-i\mathbf{k}'\cdot\mathbf{R}_n}\psi(\mathbf{k}', \mathbf{r})\,d^3k\,d^3k'\,d^3r.\end{aligned}$$

With the use of orthonomality of the Bloch functions,

$$\int \psi_s^*(\mathbf{k}, \mathbf{r})\psi_t(\mathbf{k}', \mathbf{r})\,d^3r = \delta_{s,t}\delta(\mathbf{k} - \mathbf{k}'), \tag{16}$$

the expression for $\mathcal{G}$ is simplified to

$$\mathcal{G}(\mathbf{R}_m - \mathbf{R}_n) = \frac{\Omega}{(2\pi)^3}\int \frac{\exp\,[i\mathbf{k}\cdot(\mathbf{R}_m - \mathbf{R}_n)]}{E - E(\mathbf{k})}\,d^3k. \tag{17}$$

This expression is quite general. Now we require the asymptotic form of $\mathcal{G}$ for large values of $|\mathbf{R}_m - \mathbf{R}_n|$.

The integration of (17) has been discussed in detail by Koster[15] and Lifshitz.[21] We will follow here the treatment of Koster, which is an application of the method of stationary phase. We use the fact that, in order to obtain outgoing wave, E has been given a small positive imaginary part; introduce an auxiliary variable t, and rewrite (17) as

$$\begin{aligned}\mathcal{G}_s(\mathbf{R}_m - \mathbf{R}_n) &= \frac{-i\Omega}{(2\pi)^3}\int_0^\infty dt\int d^3k \\ &\quad\times \exp\,\{i[E - E_s(\mathbf{k})]t + i\mathbf{k}\cdot(\mathbf{R}_m - \mathbf{R}_n)\}.\end{aligned} \tag{18}$$

A point of stationary phase $\mathbf{q}_0$, t_0 is determined by the conditions

$$E = E_s(\mathbf{q}_0);\quad t_0[\nabla E_s(\mathbf{k})]_{q_0} = \mathbf{R}_m - \mathbf{R}_n. \tag{19}$$

We define variables $\boldsymbol{\kappa}$, τ by $\mathbf{k} = \mathbf{q}_0 + \boldsymbol{\kappa}$, $t = t_0 + \tau$, and expand the argument of the exponential about $\mathbf{q}_0$, t_0. With the use of Eq. (19), we obtain for (18) when terms higher than second order are discarded,

$$\begin{aligned}\mathcal{G}_s(\mathbf{R}) &= \frac{-i\Omega}{(2\pi)^3}\sum e^{i\mathbf{q}_0\cdot\mathbf{R}}\int d^3\kappa\,d\tau \\ &\quad\times \exp\left\{-i\left[\tau\boldsymbol{\kappa}\cdot\nabla E_s + \frac{(\boldsymbol{\kappa}\cdot\nabla)^2}{2}E_s t_0\right]\right\},\end{aligned} \tag{20}$$

in which the summation includes all points q_0, t_0 which satisfy (19). The trick now is to choose the coordinate system in such a way that the quadratic form $[(\boldsymbol{\kappa}\cdot\nabla)^2E_s]_{q_0}$ is diagonal,

$$[(\boldsymbol{\kappa}\cdot\nabla)^2E_s]_{q_0} = \alpha_1\kappa_x^2 + \alpha_2\kappa_y^2 + \alpha_3\kappa_z^2. \tag{21}$$

Also, let

$$(\boldsymbol{\kappa}\cdot\nabla)E_s = \beta_1\kappa_x + \beta_2\kappa_y + \beta_3\kappa_z.$$

With this choice, the integral over $\boldsymbol{\kappa}$ splits up into a product of three one-dimensional integrals,

$$\begin{aligned}\frac{-i}{(2\pi)^3}\int d\tau\int d^3\kappa\,\exp\left\{-i\left[\tau\boldsymbol{\kappa}\cdot\nabla E_s + \frac{(\boldsymbol{\kappa}\cdot\nabla)^2}{2}E_s t_0\right]\right\} \\ = \frac{-i}{(2\pi)^3}\int d\tau I_1 I_2 I_3,\end{aligned} \tag{22}$$

where

$$\begin{aligned}I_1 &= \int_{-\infty}^{\infty} d\kappa_1\,\exp\,\{-i[\beta_1\kappa_1\tau + \alpha_1\kappa_1^2 t_0/2]\} \\ &= (\pi/\alpha_1 t_0)^{\frac{1}{2}}(1 - i)\,\exp\,(i\beta_1^2\tau^2/2\alpha_1 t_0).\end{aligned}$$

The limits of integration on κ_1 were extended to $\pm\infty$. The integral over τ can now be done. We obtain for (22)

[21] I. M. Lifshitz, Zh. Eksperim. i Teor. Fiz. 18, 293 (1948).

$$-[2\pi\,|t_0|\,(\alpha_2\alpha_3\beta_1^2+\alpha_1\alpha_3\beta_2^2+\alpha_1\alpha_2\beta_3^2)^{\frac{1}{2}}]^{-1}.$$

We determine $|t_0|$ from (19). Then we have

$$\mathcal{G}_s(\mathbf{R}) = \frac{-\Omega}{2\pi}\sum\frac{e^{i\mathbf{q}_0\cdot\mathbf{R}}}{|\mathbf{R}|} \times\left[\frac{\beta_1^2+\beta_2^2+\beta_3^2}{\alpha_2\alpha_3\beta_1^2+\alpha_1\alpha_3\beta_2^2+\alpha_1\alpha_2\beta_3^2}\right]^{\frac{1}{2}}. \quad (23)$$

Equation (23) takes a more familiar form if we consider a spherical band

$$E_s(\mathbf{q}) = \gamma_s q^2. \quad (24)$$

Then from (19), we see that $\mathbf{q}_0$ is parallel to $\mathbf{R}_m - \mathbf{R}_n = \mathbf{R}$, and there is only one value of $\mathbf{q}_0$. We find the familiar result of elementary quantum mechanics

$$\mathcal{G}_s(R) = (-\Omega/4\pi\gamma_s)(e^{iq_0R}/R). \quad (25)$$

In general, we notice that if the surface $E_s(\mathbf{q}_0) = E$ is everywhere convex, there will be only one solution to (19), and $\mathbf{q}_0\cdot\nabla E_s$ will be positive so that the scattered wave is outgoing. If the surface contains concave portions, there may be more than one solution. In this case, there will be waves characterized by different $\mathbf{q}_0$ proceeding in the same direction. Finally, it may happen for complex band structures, that $\mathbf{q}_0\cdot\nabla E_s$ is negative. In this case we apparently have an incoming wave, and it is necessary to change the sign of the imaginary part of E, which amounts to taking the complex conjugate of the first term in the exponential in Eq. (18).

There is one additional special case in which the evaluation of $\mathcal{G}$ can be done explicitly. Let us consider an energy band of the form

$$E(k) = \varepsilon[3 - \cos k_x a - \cos k_y a - \cos k_z a], \quad (26)$$

which is appropriate for an s band in a simple cubic lattice. Then (putting $\Omega = a^3$)

$$\mathcal{G}(\mathbf{R}_m - \mathbf{R}_n) = \frac{-a}{2\pi\varepsilon}\frac{e^{i\mathbf{q}_0\cdot(\mathbf{R}_m-\mathbf{R}_n)}}{|\mathbf{R}_m-\mathbf{R}_n|} \times\left[\frac{\sin^2 q_{0x}a+\sin^2 q_{0y}a+\sin^2 q_{0z}a}{\sin^2 q_{0x}a\cos q_{0y}a\cos q_{0z}a+\sin^2 q_{0y}a\cos q_{0x}a\cos q_{0z}a+\sin^2 q_{0z}a\cos q_{0x}a\cos q_{0y}a}\right]^{\frac{1}{2}}. \quad (27)$$

To keep matters reasonably simple, we will restrict our considerations henceforth to the situation in which there is only one $\mathbf{q}_0$ which is a solution of (19) and, more stringently, assume that the energy is a function of $|\mathbf{k}|$ only. This implies that $\mathbf{q}_0$ is parallel to $\mathbf{R}_m - \mathbf{R}_n$. We introduce a function $g(\mathbf{q}_0)$,

$$g_s(\mathbf{q}_0) = 2\left[\frac{\beta_1^2+\beta_2^2+\beta_3^2}{\alpha_1\alpha_2\beta_3^2+\alpha_1\alpha_3\beta_2^2+\alpha_2\alpha_3\beta_1^2}\right]^{\frac{1}{2}}, \quad (28)$$

so that (23) becomes

$$\mathcal{G}_s(\mathbf{R}_m - \mathbf{R}_n) = \frac{-\Omega}{4\pi}\,g_s(\mathbf{q}_0)\frac{e^{ik_0|\mathbf{R}_m-\mathbf{R}_n|}}{|\mathbf{R}_m-\mathbf{R}_n|}, \quad (29)$$

in which we have

$$\mathbf{q}_0 = k_0\frac{\mathbf{R}_m-\mathbf{R}_n}{|\mathbf{R}_m-\mathbf{R}_n|}. \quad (30)$$

This equation is substituted into (15), which yields

$$B(\mathbf{R}_m) = \frac{\Omega^{\frac{1}{2}}}{(2\pi)^{\frac{1}{2}}}\left[e^{i\mathbf{k}_0\cdot\mathbf{R}_m} - \frac{V_0\Omega}{4\pi}\frac{g(\mathbf{q}_0)}{1-V_0\mathcal{G}(0)}\frac{e^{ik_0R_m}}{R_m}\right]. \quad (31)$$

The scattering amplitude is the coefficient of the outgoing spherical wave in the bracket of Eq. (31). A more complete discussion of this identification is given in Appendix B.

$$f = -\frac{V_0\Omega}{4\pi}\frac{g(\mathbf{q}_0)}{1-V_0\mathcal{G}(0)}. \quad (32)$$

The differential cross section is

$$\sigma_D = \left(\frac{V_0\Omega}{4\pi}\right)^2\left|\frac{g(\mathbf{q}_0)}{1-V_0\mathcal{G}(0)}\right|^2. \quad (33)$$

Since the energy is a function of $|\mathbf{k}|$, no additional velocity factors appear in passing from (32) to (33). The scattering here is purely "s-wave."

To determine the general properties of the scattering amplitude, we must consider $\mathcal{G}_s(0)$. From (16) we have (writing the $i\epsilon$ explicitly)

$$\mathcal{G}_s(0) = \frac{\Omega}{(2\pi)^3}\lim_{\epsilon\to 0+}\int\frac{d^3k}{E-E_s(\mathbf{k})+i\epsilon} = \frac{\Omega}{(2\pi)^3}\lim_{\epsilon\to 0+}\iint\frac{1}{E-E_s'+i\epsilon}\frac{1}{|\nabla_k E_s'|}\,dS_{E'}\,dE_s'. \quad (34)$$

In the second line we integrate first over a surface of constant energy dS_E', and then over energies E_s'. We can recognize in (34) the density of states for the sth band $G_s(E')$, which is given by the expression

$$G_s(E') = \frac{\Omega}{(2\pi)^3}\int\frac{dS_{E'}}{|\nabla_k E_s'|}. \quad (35)$$

Hence (34) is

$$\mathcal{G}_s(0) = \lim_{\epsilon \to 0^+} \int \frac{G_s(E')\,dE'}{E - E' + i\epsilon}$$

$$= P \int \frac{G_s(E')\,dE'}{E - E'} - i\pi G_s(E). \qquad (36)$$

In the last step we have used the identity

$$1/(x + i\epsilon) = P(1/x) - i\pi\delta(x), \qquad (37)$$

where P stands for principal value.

We see from (36) that $\mathcal{G}_s(0)$ is real for energies outside the sth band. This implies that the quantity $1 - V_0\mathcal{G}_s(0)$ appearing in the denominator of the scattering amplitude can vanish only outside the band. In this case, the scattering amplitude is infinite and a real bound state exists. The condition for the bound state is

$$V_0 I(E) = 1, \qquad (38)$$

where

$$I(E) = P \int \frac{G(E')}{E - E'}\,dE'.$$

If V_0 is negative, the bound state must lie below the band, while if V_0 is positive, it can only occur above the band. In the case of a one-dimensional band, an impurity always produces a split-off state; but such is not the case here, since for sufficiently weak potentials, Eq. (38) will not have a solution.

It may happen that Eq. (38) is satisfied for some energy E_0 inside the band. In this case the real part of the expression $1 - V_0\mathcal{G}(0)$ vanishes. Then the cross section may have a maximum for an energy near E_0. It is convenient to define the energy of a resonant state as the location of a pole of the scattering amplitude. This energy will, in general, be complex since as we have seen, $1 - V_0\mathcal{G}(0)$ does not vanish for real energies inside the band. Let the pole of the amplitude occur for $E = E_R - i\Gamma/2$, where Γ is positive. We frequently refer to Γ as the width of the resonance. It must be a positive number since a state of energy E should decay exponentially with time. The quantities E_R and Γ can easily be determined for resonances which are near a point E_0 as defined above. We expand $\mathcal{G}(0)$ retaining only first-order terms,

$$\mathcal{G}(0) = I(E_0) - i\pi G(E_0) + (E - E_0)[I'(E_0) - i\pi G'(E_0)],$$

in which

$$I' = dI/dE, \quad \text{etc.}$$

Then, substituting for E, we have

$$1 - V_0\mathcal{G}(0) = 1 - V_0 I(E_0) - V_0(E_R - E_0 - i\Gamma/2)I'(E_0) + i\pi V_0[G(E_0) + (E_R - E_0 - i\Gamma/2)G'(E_0)] = 0.$$

We separate real and imaginary parts, and solve the resulting pair of equations for E_R and Γ:

$$E_R = E_0 - \frac{\pi^2 G(E_0)G'(E_0)}{I'^2(E_0) + \pi^2 G'^2(E_0)},$$
$$\Gamma = -2\pi G(E_0)I'(E_0)/[I'^2(E_0) + \pi^2 G'^2(E_0)]. \qquad (39a)$$

Since only first-order terms were retained in the expansion of $\mathcal{G}(0)$, these equations are valid only so long as the displacements from E_0 are small. The equation for Γ takes a more familiar form if we assume that

$$\pi G'(E_0)/I'(E_0) \ll 1.$$

We then have

$$E_R = E_0 - \pi^2 G(E_0)G'(E_0)/[I'^2(E_0)],$$
$$\Gamma = -2\pi G(E_0)/I'(E_0). \qquad (39b)$$

The significance of these results may be easily appreciated if we substitute the expanded form of the expression $1 - V_0\mathcal{G}(0)$ into Eq. (32). Since $1 - V_0 I(E_0) = 0$, we have

$$f = \Omega g(q_0)[4\pi I'(E_0)\{E - E_0 - i\pi G(E)/I'(E_0)\}]^{-1}$$
$$\approx -\Omega g(q_0)\Gamma[8\pi^2 G(E_0)(E - E_0 + i\Gamma/2)]^{-1}. \qquad (40a)$$

The differential cross section is

$$\sigma_D = \frac{\Omega^2}{64\pi^4 G^2(E_0)} \frac{|g(q_0)|^2\,\Gamma^2}{(E - E_0)^2 + \pi^2 G^2(E)/I'^2(E_0)}, \qquad (40b)$$

or approximately

$$\sigma_D = A(E_0)\Gamma^2/[(E - E_0)^2 + \tfrac{1}{4}\Gamma^2]. \qquad (40c)$$

On differentiating (40b) with respect to energy, one finds immediately that the maximum of the cross section occurs for $E = E_R$, where E_R is given by (39b), provided that the energy dependence of $g(q_0)^2$ may be neglected.

If we use the simple expression given for the energy in Eq. (24), we have

$$G(E) = \Omega k(E)/4\pi^2\gamma.$$

Then Eq. (40c) simplifies to

$$\sigma_D = \frac{1}{4k^2} \frac{\Gamma^2}{(E - E_0)^2 + \tfrac{1}{4}\Gamma^2}. \qquad (40d)$$

This is a standard result.[22]

[22] Ta-You Wu and Takashi Ohmura, *Quantum Theory of Scattering*, (Prentice Hall, Inc., Englewood Cliffs, New Jersey, 1962), p. 13.

As has been mentioned above, it is necessary that Γ be positive at a resonance. Since $G(E)$ is always positive, we see that a resonance can occur only when $I'(E_0)$ is negative. The significance of this condition can be seen somewhat more clearly through a slightly different argument which is presented below. Observe that near a resonance, the cross section does not depend explicitly on V_0 (the dependence is through E_0).

To see the significance of the negative sign in (39), we investigate the contribution of the impurity to the density of states for the crystal. The density of states $N(E)$ (including the impurity) may be written formally as[23]

$$N(E) = -(1/\pi)\ \mathrm{Im}\ [\mathrm{Tr}\ 1/(E^+ - H)] \qquad (41)$$

where H is the full Hamiltonian for the system including the defect and Tr denotes the trace. The symbol E^+ indicates that the energy E in (41) has an infinitesimal, positive, imaginary part. Equation (41) may be rewritten as

$$N(E) = -(1/\pi)\ \mathrm{Im}\ [\mathrm{Tr}\ (d/dE)\{\ln (E^+ - H)\}]$$
$$= -(1/\pi)\ \mathrm{Im}\ [d/dE\{\ln \det (E^+ - H)\}]. \qquad (42)$$

The contribution from the impurity state may be separated by writing

$$E^+ - H = (E^+ - H_0)[I - (E^+ - H_0)^{-1}V].$$

The density of states for the crystal with the impurity replaced by a normal atom is

$$G(E) = -(1/\pi)\ \mathrm{Im}\ [d/dE\{\ln \det (E^+ - H_0)\}]. \qquad (43)$$

Since the determinant of the product of two matrices is equal to the product of the determinants, the change in the density of states due to the defect is

$$\Delta N = N(E) - G(E)$$
$$= -\frac{1}{\pi}\ \mathrm{Im}\left[\frac{d}{dE}\ln \det \left\{I - \frac{1}{E^+ - H_0}\,V\right\}\right]. \qquad (44)$$

Equation (44) is general and exact. The matrix $I - (E - H_0)^{-1}V$ is constructed on the basis of Wannier functions. For a general potential, the nonzero portion of the matrix V has dimension $d \times d$, where d is determined by the range of the potential; and let there be $\mathfrak{N}$ lattice sites altogether. Then the matrix $I - [1/(E^+ - H_0)]V$ has the block form

$$\left[\begin{array}{c|c} = & 0 \\ = & \\ \hline = & I \\ = & \end{array}\right],$$

where the upper left portion is of dimension $d \times d$ and the whole matrix has dimension $\mathfrak{N} \times \mathfrak{N}$. One easily verifies that the determinant of the matrix is just the determinant of the upper $d \times d$ part. In the simple model we have been discussing, there is only one element to consider. This element has the value

$$1 - V_0\mathcal{G}(0) = 1 - V_0 I(E) + i\pi V_0 G(E).$$

Then we have

$$\Delta N = -\frac{1}{\pi}\ \mathrm{Im}\ \frac{d}{dE}\ln\,[1 - V_0 I(E) + i\pi V_0 G(E)]$$
$$= \frac{-V_0[I'(E)V_0 G(E) + G'(E)(1 - V_0 I(E))]}{[1 - V_0 I(E)]^2 + \pi^2 V_0^2 G^2(E)}. \qquad (45)$$

We investigate (45) near a resonance, and expand near that point. We have approximately

$$\Delta N \approx -G(E_0)/[I'(E_0)$$
$$\times \{(E - E_0)^2 + \pi^2 G^2(E)/I'^2(E_0)\}]$$
$$\approx \Gamma/2\pi[(E - E_0)^2 + \tfrac{1}{4}\Gamma^2]. \qquad (46)$$

In the last line of (46), we have introduced Eq. (39b). In particular, at the resonance $E = E_0$, and

$$\Delta N(E_0) = 2/\pi\Gamma \qquad (47)$$

It is necessary that the density of states be increased near a resonance, which means that Γ must be positive, in agreement with the previous discussion.

An interesting expression is obtained if we attempt to write Eq. (32), which gives the scattering amplitude, in a form equivalent to the ordinary partial-wave expression for s-wave scattering. This is

$$f = (1/2ik)(e^{2i\delta_0} - 1), \qquad (48)$$

where δ_0 is the s-wave phase shift. In order to make (32) look like (48), it is necessary to assume that the energy is given by (24). With the use of the appropriate density of states, we find that the s-wave phase shift δ_0 is given by

$$\tan \delta_0 = -\pi V_0 G(E)/[1 - V_0 I(E)]. \qquad (49)$$

Let us now examine the results we have obtained. In general, $I(E)$ is negative for energies below the bottom of the band, is positive above the band, and goes through zero somewhere in the band. We assume here that it has only one maximum and one minimum. This behavior is shown schematically in Fig. 1. In addition, we suppose that V_0 is negative, corresponding to an attractive potential.

(1) V_0 is sufficiently small so that the $1/V_0 \neq I(E)$ for any E. In this case there is neither a bound

[23] J. R. Klauder, Ann. Phys. (N. Y.) 14, 43 (1961). See also M. Baker, Ann. Phys. (N. Y.) 4, 271 (1958).

state nor a resonance. The phase shift δ_0 starts at 0^0 at the bottom of the band, rises to a maximum, and decays to 0^0 again at the top of the band.

(2) A resonance occurs. Note that the equation $1/V_0 = I(E)$ must have either two solutions or none. If the solution of lowest energy occurs within the band, there may be a resonance near that energy. The solution of high energy does not correspond to a resonance. The phase shift starts at zero at the bottom of the band, passes through $\frac{1}{2}\pi$ at the resonance, reaches a maximum (less than π), and decreases. It passes again through $\frac{1}{2}\pi$ at the energy corresponding to the second intersection of $1/V_0$ with $I(E)$ and goes to zero at the top of the band. The density of states is increased near the resonance and diminished near the second intersection.

(3) A bound state exists below the band. In this case, the intersection of $1/V_0$ with $I(E)$ has already occurred before the band begins. Therefore $1 - V_0 I$ is negative at the bottom of the band, and the phase shift starts out at π, decreasing as a function of energy. It goes through $\frac{1}{2}\pi$ at the second intersection of $1/V_0$ with $I(E)$, and thereafter decreases to zero at the top of the band. Note that only one bound state can exist since I' is always negative below the band. Also, for usual band shapes, if there is a bound state, there will not be a resonance, since that would require the negative portion of $I(E)$ to have two minima.

III. AN EXTENDED PERTURBATION

Next, we consider the situation in which the perturbation extends over more than one lattice site. In this case, we must develop the analog of the partial wave analysis of the scattering amplitude. We will still keep two of the restrictions previously imposed in that the potential will not have matrix elements between bands and the energy considered is low so that the band is spherical: $E = E(|\mathbf{k}|)$. Instead of Eq. (13), we now have to consider

$$B_s(\mathbf{R}_m) = B_s^{(0)}(\mathbf{R}_m) + \sum_{pn} \mathcal{G}_s(\mathbf{R}_m - \mathbf{R}_p) V_{pn} B_s(\mathbf{R}_n), \quad (50)$$

in which we have abbreviated the matrix elements of V as

$$V_{pn} = (s, p \,|V|\, s, n).$$

The impurity potential is assumed to have the point symmetry of the crystal lattice in which it is placed, and to vanish beyond some specified distance from the origin:

$$V_{pn} = 0 \quad \text{if} \quad R_p > R_0, \quad \text{or} \quad R_n > R_0. \quad (51)$$

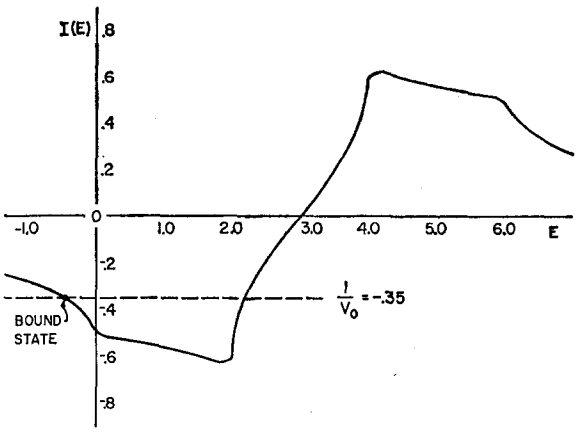

FIG. 1. The function $I(E) = \int [G(E')/(E - E')]\, dE'$ is shown as a function of energy for a simple cubic lattice in which $E(k) = 3 - \cos k_x a - \cos k_y a - \cos k_z a$. The straight line is drawn for $1/V_0 = -0.35$ to illustrate the determination of bound states. A bound state occurs at an energy determined by the left-hand intersection of $1/V_0$ with $I(E)$.

It follows that the functions $B_s(\mathbf{R}_m)$ must be linear combinations of functions transforming according to the irreducible representations of the crystal point group.

Let us introduce a set of symmetrized linear combinations of plane waves

$$C_\alpha^{(0)}(R_m) = \sum_m{}' U(\alpha, \mathbf{R}_m) B^{(0)}(\mathbf{R}_m), \quad (52)$$

in which the index α denotes an irreducible representation of the point group (when necessary to designate a row of a degenerate representation, we place a subscript ν on α). The sum over m runs over all the different vectors $\mathbf{R}_m$ which can be found from any one of them by applying the operators of the point group. All these vectors have the same length. A prime on the summation sign indicates such a restricted sum. The matrix $U(\alpha, \mathbf{R}_m)$ is unitary in that

$$\sum_m{}' U(\alpha, \mathbf{R}_m) U^+(\mathbf{R}_m, \beta) = \delta_{\alpha\beta},$$
$$\sum_\alpha U^+(\mathbf{R}_m, \alpha) U(\alpha, \mathbf{R}_{m'}) = \delta_{m,m'}, \quad (53)$$

where in the second of Eq. (52), the sum over α includes each row and each representation as many times as it may occur, and

$$U^+(\mathbf{R}_m, \alpha) = U^*(\alpha, \mathbf{R}_m). \quad (54)$$

Next, we apply the unitary transformation to (50). We thereby form symmetrized linear combinations of scattering functions

$$C_\alpha(R_m) = \sum_m{}' U(\alpha, \mathbf{R}_m) B(\mathbf{R}_m), \quad (55)$$

which satisfy the equations

$$C_\alpha(R_m) = C_\alpha^{(0)}(R_m) + \sum_m{}' \sum_{pn} U(\alpha, \mathbf{R}_m)\mathcal{G}(\mathbf{R}_m - \mathbf{R}_p)V_{pn}B_s(\mathbf{R}_n) = C_\alpha^{(0)}(\mathbf{R}_m) + \sum\!\!\!\!\!\!- \sum_{pn} \Gamma_{\alpha,mp}V_{\alpha,pn}C_\alpha(R_n), \tag{56}$$

in which

$$\Gamma_{\alpha,mp}\delta_{\alpha\beta} = \sum_{mp}{}' U(\alpha, \mathbf{R}_m)\mathcal{G}(\mathbf{R}_m - \mathbf{R}_p)U^+(\mathbf{R}_p, \beta),$$
$$V_{\alpha,pn}\delta_{\alpha\beta} = \sum_{pn}{}' U(\alpha, \mathbf{R}_p)V_{pn}U^+(\mathbf{R}_n, \beta). \tag{57}$$

In the second line of Eq. (56), the sums over p and n include all the direct lattice vector groups. This is denoted by a bar over the summation sign. For some sets of lattice vectors, more than one symmetrized function belonging to a given row of a degenerate irreducible representation exists, and each must be included in the sum. The indices m, n, p, etc. are to be generalized as required. According to the principles of group theory, the operators Γ and V do not connect states belonging to different irreducible representations.

We must now solve Eq. (56). To do this, we consider first values of m such that $R_m < R_0$. From Eq. (51), it follows that functions $C_\alpha(R_m)$, with R_m satisfying the above criterion, are connected through the right side of Eq. (56) only with other functions $C_\alpha(R_n)$ for which $R_n < R_0$. Therefore such values of R may be considered separately.

We therefore define the matrix

$$Q_{\alpha,mn} = \delta_{mn} - \sum\!\!\!\!\!\!- \sum_p \Gamma_{\alpha,mp}V_{\alpha,pn}, \tag{58}$$

so that Eq. (56) can be written as

$$\sum\!\!\!\!\!\!- \sum_n Q_{\alpha,mn}C_\alpha(R_n) = C_\alpha^{(0)}(R_m). \tag{59}$$

Therefore

$$C_\alpha(R_m) = \sum\!\!\!\!\!\!- \sum_n [Q_\alpha^{-1}]_{mn}C_\alpha^{(0)}(R_n). \tag{60}$$

We choose to write the inverse matrix Q^{-1} as follows:

$$[Q_\alpha^{-1}]_{mn} = P_{\alpha,mn}/D_\alpha, \tag{61}$$

where

$$D_\alpha = \det(Q_\alpha) = \det[I - \Gamma_\alpha V_\alpha]. \tag{62}$$

Thus we have

$$C_\alpha(R_m) = D_\alpha^{-1} \sum_n P_{\alpha,mn}C_\alpha^{(0)}(R_n) \qquad (R_m, R_n < R_0). \tag{63}$$

It is desirable to separate D_α from Q^{-1} since it is D which determines the locations of possible resonances and bound states. We substitute (63) into (56) and obtain for $R_m > R_0$

$$C_\alpha(R_m) = C_\alpha^{(0)}(R_m) + D_\alpha^{-1} \sum\!\!\!\!\!\!- \sum_{pnl} \Gamma_{\alpha,mp}V_{\alpha,pn}P_{\alpha,nl}C_\alpha^{(0)}(R_l). \tag{64}$$

If, as occasionally happens, $V_{\alpha,pn}$ has only one element, say $V_{\alpha,11}$, different from zero, Eq. (64) reduces to

$$C_\alpha(R_m) = C_\alpha^{(0)}(R_m) + D_\alpha^{-1}\Gamma_{\alpha,m1}V_{\alpha,11}C_\alpha^{(0)}(R_1) = C_\alpha^{(0)}(R_m) + \frac{\Gamma_{\alpha,m1}V_{\alpha,11}}{1 - \Gamma_{\alpha,11}V_{\alpha,11}} C_\alpha^{(0)}(R_1). \tag{65}$$

To obtain the scattering amplitude, we transform back to the functions B. From (64) and (55), we have that

$$B(\mathbf{R}_m) = \sum_\alpha U^+(\mathbf{R}_m, \alpha)C_\alpha(R_m) = B^{(0)}(\mathbf{R}_m) + \sum_\alpha \sum\!\!\!\!\!\!- \sum_{pnl} U^+(\mathbf{R}_m, \alpha)\Gamma_{\alpha,mp}D_\alpha^{-1}V_{\alpha,pn}P_{\alpha,nl}C_\alpha^{(0)}(R_l). \tag{66}$$

To simplify this expression and exhibit the scattering amplitude, we replace the symmetrized Green's function Γ by $\mathcal{G}$ through (57),

$$B(\mathbf{R}_m) = B^{(0)}(\mathbf{R}_m) + \sum_\beta \sum_p \sum\!\!\!\!\!\!- \sum_{nl} \mathcal{G}(\mathbf{R}_m - \mathbf{R}_p)U^+(\mathbf{R}_p, \beta) \times D_\beta^{-1}V_{\beta,pn}P_{\beta,nl}C_\beta^{(0)}(R_l). \tag{67}$$

Let us examine the asymptotic form of $\mathcal{G}$. From (29) in the limit $R_m \gg R_p$, we find

$$\mathcal{G}(\mathbf{R}_m - \mathbf{R}_p) = \frac{-\Omega g(\mathbf{k}')}{4\pi R_m} \exp\{ik_0R_m(1 - \mathbf{R}_m\cdot\mathbf{R}_p/R_m^2)\} = \frac{-\Omega g(\mathbf{k}')}{4\pi R_m} e^{-i\mathbf{k}'\cdot\mathbf{R}_p}e^{ik_0R_m}, \tag{68}$$

where $\mathbf{k}'$ is a vector in the direction of $\mathbf{R}_m$,

$$\mathbf{k}' = k_0\mathbf{R}_m/R_m.$$

Next, observe that

$$\sum_p{}' e^{-i\mathbf{k}'\cdot\mathbf{R}_p}U^+(\mathbf{R}_p, \beta) = \frac{(2\pi)^{\frac{3}{2}}}{\Omega^{\frac{1}{2}}} \sum_p{}' B^{(0)*}(\mathbf{k}', \mathbf{R}_p)U^+(\mathbf{R}_p, \beta) = \frac{(2\pi)^{\frac{3}{2}}}{\Omega^{\frac{1}{2}}} \sum_p{}' U^*(\beta, \mathbf{R}_p)B^{(0)*}(\mathbf{k}', \mathbf{R}_p)$$

$$= \frac{(2\pi)^{\frac{1}{2}}}{\Omega^{\frac{1}{2}}} C_\beta^{(0)*}(k', R_p). \tag{69}$$

We have explicitly indicated that the symmetrized function $C_\beta^{(0)}$ depends on the wave vector k' which characterizes the outgoing wave. Similarly the function $C_\beta^{(0)}(R_l)$ which is present in (67) is really $C_\beta^{(0)}(k_0, R_l)$ since it depends on the wave vector of the incident wave.

With the use of (69) and (10), Eq. (67) may be written as

$$B(\mathbf{R}_m) = \frac{\Omega^{\frac{1}{2}}}{(2\pi)^{\frac{1}{2}}}\left[e^{i\mathbf{k}_0\cdot\mathbf{R}_m} - 2\pi^2 g(\mathbf{k}')\frac{e^{ik_0R_m}}{R_m}\right.$$
$$\left.\times \sum_\beta \sum_{pnl}{}^{\prime\prime} D_\beta^{-1} V_{\beta,pn} P_{\beta,nl} C_\beta^{(0)*}(k', R_p) C_\beta^{(0)}(k_0, R_l)\right]. \tag{70}$$

The scattering amplitude is

$$f = -2\pi^2 g(k_0) \sum_\beta D_\beta^{-1} \sum_{pnl}{}^{\prime\prime} V_{\beta,pn} \times P_{\beta,nl} C_\beta^{(0)*}(k', R_p) C_\beta^{(0)}(k_0, R_l). \tag{71}$$

This equation expresses the scattering amplitude as a sum of contributions from the irreducible representations of the symmetry group of the potential and is therefore the analog of the usual partial-wave formula. An analysis of the significance of the scattering amplitude in the solid-state problem will be found in Appendix B.

If we write

$$f = \sum_\beta f_\beta,$$

where now the sum over β includes only irreducible representations themselves, and not their rows, we have

$$f_\beta = -2\pi^2 g(k_0) D_\beta^{-1} \sum_{pnl}{}^{\prime\prime} V_{\beta,pn} P_{\beta,nl} \times \sum_\nu C_{\beta,\nu}^{(0)*}(k', R_p) C_{\beta,\nu}^{(0)}(k_0, R_l), \tag{72}$$

in which the index ν designates a row of the βth irreducible representation.

Only those symmetrized combinations of plane waves which can be formed from $B^{(0)}$ for R_m within the range of the impurity potential can be scattered. This means that only those partial waves which belong to the irreducible representations of the point group which can be formed in this manner can appear in (71). For a sufficiently extended perturbation, all of the irreducible representations will appear. If, however, the potential includes only one cell, only one term (the s wave, Γ_1) appears as has already been seen. For potentials extending to first neighbors only, we have, in cubic systems, the following sets of partial waves to consider:

Simple cubic		Body-centered cubic		Face-centered cubic	
Γ_1	(1, s)	Γ_1	(1, s)	Γ_1	(1, s)
Γ_{15}	(3, p)	Γ_{15}	(3, p)	Γ_{15}	(3, p)
Γ_{12}	(2, d)	$\Gamma_{25'}$	(3, d)	Γ_{12}	(2, d)
		$\Gamma_{2'}$	(1, f)	$\Gamma_{25'}$	(3, d)
				Γ_{25}	(3, f)

The representations are designated according to the notation of Bouckaert, Smoulochowski, and Wigner.[24] The degeneracy of the representation and the approximate symmetry in terms of spherical, or more properly Kubic, harmonics is given in parenthesis for each representation.

Much of the general analysis of the scattering amplitude which was developed for the model of Eq. (12) is also applicable here. In particular, bound states or localized modes are found at energies E outside the continuum for which

$$D_\beta(E) = 0, \tag{73}$$

whereas resonances occur for energies, E_R, which are the real parts of complex energy solutions of (73), as has been discussed above (Of course, E_R must lie in the band). The imaginary part of the energy defines a level width as has been discussed previously.

Although the order, n, of D_β is finite and is given by the number of symmetrized linear combinations of plane waves which can be formed from the $B^{(0)}(\mathbf{R}_m)$ that belong to a particular row of the βth irreducible representation, D_β is not a polynomial function of E. However, for energies below the bottom of the band, or above the top of the band, the Green's functions, which are real for such energies, may be expanded in powers of $1/E$ (we measure energies from the midpoint of the band). Hence, in this region, we may write:

$$D_\beta(E) = E^{-n}[d_\beta^{(n)}(E) + O(1/E) + \cdots], \tag{74}$$

where $d_\beta^{(n)}(E)$ is an nth order polynomial in E, and such that the coefficient of E^n is unity. The only circumstances in which such a statement will not hold occur if there should be "accidental" cancellations of the terms. Hence, for energies well-removed from the band, the resonance condition is

$$d_\beta^{(n)}(E) = 0,$$

and this equation cannot have more than n roots. In addition, one observes that the algebraic sign

[24] L. P. Bouckaert, R. Smoluchowski, and E. Wigner, Phys. Rev. 50, 58 (1936).

of D_β at the bottom (or the top) of the band will be $(-1)^l$ where l is the number of bound states, below (or above) the band. This result has a certain resemblance to Levinson's theorem.

The change in the density of states produced by the imperfections may be expressed in terms of the functions D_β. We consider the basic equation (44) which expresses the change in the density of states, N, as the imaginary part of the logarithmic derivative of the determinant of the matrix $I - [1/(E - H_0)]V$. The determinant of a matrix is unchanged by a unitary transformation, so we may write formally

$$\Delta N = -\frac{1}{\pi} \operatorname{Im} \frac{d}{dE} \times \ln \det \left[I - U \frac{1}{E^+ - H_0} U^+ U V U^+ \right], \quad (75)$$

where U is the unitary transformation introduced in Eq. (52). This transformation causes the determinant to factor into portions corresponding to the various irreducible representations. It will be seen from (62) that these factored parts are just the functions D_β. Hence

$$\det \left[I - U \frac{1}{E^+ - H_0} U^+ U V U^+ \right] = \prod_\beta D_\beta(E), \quad (76)$$

where the index β includes the rows of the representation. The change in the density of states may therefore be expressed as a sum of contributions from the irreducible representations,

$$\Delta N = \sum_\beta \Delta N_\beta, \quad (77a)$$

where

$$\Delta N_\beta = -\frac{1}{\pi} \operatorname{Im} \left(\frac{1}{D_\beta} \frac{dD_\beta}{dE} \right). \quad (77b)$$

In (77a), one term appears for each row of a degenerate irreducible representation.

We can determine the locations and widths of resonances directly from D_β. We will suppose that, for some energy $E_0^{(\beta)}$,

$$\operatorname{Re} D_\beta(E_0^{(\beta)}) = 0. \quad (78)$$

In addition let $D_{\beta,r}$ and $D_{\beta,i}$ be the real and imaginary parts of D_β, and let a prime indicate differentiation with respect to energy. We expand D_β near E_0,

$$D_\beta(E) = (E - E_0)D'_{\beta,r} + i[D_{\beta,i}(E_0) + (E - E_0)D'_{\beta,i}(E_0)]. \quad (79)$$

As before, we write $E = E_{\beta,r} - i\Gamma_\beta/2$, set $D_\beta = 0$, separate real and imaginary parts; and solve the resulting pair of equations.

$$E_{\beta,r} = E_0 - \frac{D_{\beta,i}D'_{\beta,i}}{(D'_{\beta,r})^2 + (D'_{\beta,i})^2} \approx E_0 - \frac{D_{\beta,i}D'_{\beta,i}}{(D'_{\beta,r})^2},$$
$$\Gamma_\beta = \frac{2D_{\beta,i}D'_{\beta,r}}{(D'_{\beta,r})^2 + (D'_{\beta,i})^2} \approx \frac{2D_{\beta,i}}{D'_{\beta,r}}. \quad (80)$$

In the second step of each of Eq. (80), we have assumed that $|D'_{\beta,r}| \gg |D'_{\beta,i}|$. These results are, of course, quite similar to Eqs. (39 a,b), and in fact reduce to them through the substitution, valid for the model discussed there, that $D_{\beta,r} = 1 - V_0 I(E)$; $D_{\beta,i} = \pi V_0 G(E)$. We also have,

$$\Delta N_\beta \approx \frac{\Gamma_\beta}{2\pi} \frac{1}{(E - E_0)^2 + \frac{1}{4}\Gamma_\beta^2}, \quad (81a)$$

and, for $E = E_0$

$$\Delta N_\beta(E_0) = 2/\pi\Gamma_\beta. \quad (81b)$$

These expressions agree exactly with Eqs. (46) and (47).

Now let us return to Eq. (72) which expresses the contribution to the scattering amplitude from the βth irreducible representation. One general result may be deduced. The plane-wave combinations $C^{(0)}_{\beta,\nu}(k, R)$, if expanded in powers of kR, yield symmetrized combinations of spherical harmonics which belong to β_ν (Kubic harmonics). Hence for sufficiently small k, the scattering amplitude is proportional to k^{2l}, where l is the "angular momentum" corresponding to the lowest-order Kubic harmonic appearing in the expansion of $C^{(0)}_{\beta,\nu}$. Exceptions to this result can arise only when cancellations occur in the sum, in which case the scattering amplitude will be proportional to a higher power of k.

To see this result in more detail, and in order to investigate the angular dependence of the scattering amplitude, let us consider the case of p-wave scattering in a simple cubic lattice with nearest neighbor interactions only. By p wave, we mean that the representation Γ_{15} of the cubic point group is considered. In this case (72) becomes

$$f_\nu = -2\pi^2 g(\mathbf{k}') D_\nu^{-1} V_{\nu,11} \times \sum_\nu C^{(0)*}_{\nu,\nu}(k', R_1) C^{(0)}_{\nu,\nu}(k_0, R_1). \quad (82)$$

Basis functions for the rows of the Γ_{15} representation are functions transforming as x, y, z under the cubic

point group. We must make a linear combination of plane waves $e^{ik_s a}$, etc., which transform as required.[25] This is easily done using group theoretic methods, as described in Ref. 25. The normalized functions are

$$C_{p_x}^{(0)}(k, R_i) = [\Omega^{\frac{1}{2}}/(2\pi)^{\frac{3}{2}}]\sqrt{2}\sin k_x a, \tag{83}$$

etc. Hence we find for f_p

$$f_p = [-\Omega g(\mathbf{k}')/2\pi D_p]V_{p,11}[\sin k_x' a \sin k_{0x} a + \sin k_y' a \sin k_{0y} a + \sin k_z' a \sin k_{0z} a]. \tag{84}$$

If we expand the sines, and retain only the first term, we have an approximation valid for long wavelengths,

$$f_p = [-\Omega a^2 g(\mathbf{k}')/2\pi D_p](k_x' k_{0x} + k_y' k_{0y} + k_z' k_{0z})V_{p,11} = -\Omega a^2 g(\mathbf{k}') k_0^2 a^2 V_{p,11}\cos\theta/(2\pi D_p), \tag{85}$$

in which θ is the angle between k' and k_0, i.e., between the incoming and outgoing waves. This angular dependence of f_p is just what is found for p-wave scattering in elementary quantum mechanics. In addition, one naturally finds that the s-wave (Γ_1) amplitude is independent of angle (but only in the long wavelength limit).

Results for other representations are not particularly simple because the spherical harmonics of the same "l" may be divided between several representations, and one cannot make use of the addition theorem. The greater complexity of the present "partial-wave" formula [Eq. (72)] as compared with the formulas of elementary quantum mechanics results from the lower symmetry of the present problem. In particular, it does not seem to be possible to express the scattering amplitudes solely in terms of $|\mathbf{k}' - \mathbf{k}_0|$, that is, in terms of $|\mathbf{k}'|$ and the angle between $\mathbf{k}'$ and $\mathbf{k}_0$.

IV. PERTURBATION CONNECTING BANDS

Additional complexities arise when the defect potential is allowed to have nonvanishing matrix elements between different bands. In this situation, most significant physical quantities become matrices in the band index. We will consider this problem first within the framework of a generalized Koster–Slater model, in which we write, instead of Eq. (12),

$$(s, l\,|V|\,t, n) = V_{st}'\delta_{l,0}\delta_{n,0}. \tag{86}$$

This potential is localized on a single site, but connects different bands. It is necessary here and in the following to assume that the matrix of the potential is of finite dimension, or at least that it may be truncated. The equation for the scattering functions is

$$B_s(\mathbf{R}_m) = B_s^{(0)}(\mathbf{R}_m) + \mathcal{G}_s(\mathbf{R}_m)\sum_t V_{st}'B_t(0). \tag{87}$$

As usual, we consider (87) for $\mathbf{R}_m = 0$. Let the incident wave belong to band u and have wave vector k_0. Then

$$B_s(0) = [\Omega^{\frac{1}{2}}/(2\pi)^{\frac{3}{2}}]\delta_{s,u} + \mathcal{G}_s(0)\sum_t V_{st}'B_t(0). \tag{88}$$

We define the matrices Q' and P' similarly to (58) and (61),

$$Q_{st}' = \delta_{st} - \mathcal{G}_s(0)V_{st}', \tag{89}$$

and

$$[Q'^{-1}]_{st} = D^{-1}P_{st}', \tag{90}$$

in which $D' = \det Q'$. Then we find (91)

$$B_s(0) = [\Omega^{\frac{1}{2}}/(2\pi)^{\frac{3}{2}}](P_{su}/D'), \tag{92}$$

from which we obtain

$$B_s(\mathbf{R}_m) = \frac{\Omega^{\frac{1}{2}}}{(2\pi)^{\frac{3}{2}}}\left[e^{i\mathbf{k}_0\cdot\mathbf{R}_m}\delta_{su} + \frac{\mathcal{G}_s(\mathbf{R}_m)}{D'}\sum_t V_{st}'P_{tu}'\right] = \frac{\Omega^{\frac{1}{2}}}{(2\pi)^{\frac{3}{2}}}\left[e^{i\mathbf{k}_0\cdot\mathbf{R}_m} - \frac{\Omega}{4\pi}g_s(\mathbf{q}_0)\frac{e^{iq_0R_m}}{D'R_m}\sum_t V_{st}'P_{tu}'\right]. \tag{93}$$

In the second line of (93), we have passed to the limit of large $\mathbf{R}_m$ with the use of Eq. (29). The outgoing wave has wave vector $\mathbf{q}_0$.

The scattering amplitude which is obtained from Eq. (93) is evidently

$$f_{su} = \frac{-\Omega}{4\pi}\frac{g_s(q_0)}{D'}\sum_t V_{st}'P_{tu}', \tag{94}$$

which is a component of a matrix. This result may be interpreted as follows: A wave of propagation vector $\mathbf{k}_0$ in band u is incident upon the defect. Scattered waves are obtained not only in band u but in all other bands in which it is possible to satisfy the requisite conditions, Eq. (19). We note that in this case, the magnitude of the wave vector $\mathbf{q}_0$, of the outgoing wave in band s will not be the same as $\mathbf{k}_0$, since the energy-wave vector relation will be different in different bands.

The conditions for the existence of a bound state or resonance is, as usual,

$$D'(E) = 0. \tag{95}$$

Since D' is the determinant of a matrix containing the interband couplings, several bound states and (or) resonances associated with different bands may appear.

[25] J. Callaway, *Energy Band Theory* (Academic Press Inc., New York, 1963), Chap. II.

These results may be, unfortunately, purely formal. Unless the potential energy matrix V'_{st} is finite, an exact calculation of Q' is not possible. If the potential matrix is infinite, approximation techniques must be employed.

The general case in which the potential-energy operator is not diagonal in either the band index or the site index is similar in principle. It will be assumed that only a finite number of bands need be considered. The basic equation is:

$$\mathbf{B}(\mathbf{R}_m) = \mathbf{B}^{(0)}(\mathbf{R}_m) + \sum \mathbf{G}(\mathbf{R}_m - \mathbf{R}_n)\mathbf{V}(\mathbf{R}_n, \mathbf{R}_p)\mathbf{B}(\mathbf{R}_p), \quad (96)$$

in which $\mathbf{B}$ is a vector with respect to the band index, while $\mathbf{V}$ and $\mathbf{G}$ are matrices in this quantity (however, $\mathbf{G}$ is diagonal). We can still make the unitary transformation which separates the irreducible representations since the coefficients depend only on the representation and not the band. In this way, we obtain an equation analogous to (56), except that all the quantities are matrices or vectors in the band index,

$$\mathbf{C}_\alpha(R_m) = \mathbf{C}_\alpha^{(0)}(R_m) + \sum_{pn}{}'' \boldsymbol{\Gamma}_{\alpha,mp}\mathbf{V}_{\alpha,pn}\mathbf{C}_\alpha(R_n). \quad (97)$$

The technique for solution of this equation is similar to that of Eq. (56) namely, one first solves the portion involving sites for which the potential is not zero. The dimension of the matrix Q, etc. is now larger, since for each site there are as many elements as there are bands involved. The resonance bound-state equation

$$D_\beta(E) = 0$$

may have an appropriately larger number of roots. We may therefore proceed directly to the answer, which is a generalized form of Eq. (71) or (72), since nothing fundamental is altered in the analysis.

Let $C^{(0)}_{\alpha,s}$ denote a linear combination of plane waves belonging to band s. The dependence on s comes from the fact that if a specific energy is given, the wave vectors corresponding to that energy will be different in the different bands. The scattering amplitude is now a matrix, characterized by band indices, as is already evident from Eq. (94). We have

$$f_{st} = \sum_\beta f_{\beta,st},$$

where the index β still describes the irreducible representation and

$$f_{\beta,st} = -2\pi^2 g_s(\mathbf{q}_0)D_\beta^{-1} \sum_{pnl}{}'' \sum_u V_{\beta,pn,su}P_{\beta,nl,ut} \times \sum_\nu C^{(0)*}_{\beta\nu,s}(q_0, R_p)C_{\beta\nu,t}(k_0, R_l). \quad (98)$$

In this equation, the indices su, ut, etc., on V and P designate the bands. In particular,

$$V_{\beta,pn,su}\delta_{\alpha\beta} = \sum_{pn}{}' U(\alpha, R_p) \times (s, p\,|V|\,u, n)U^+(R_n, \beta), \quad (99)$$

and

$$P_{\beta,nl,ut} = D_\beta[Q_\beta^{-1}]_{nl,ut} \quad (100)$$

in which

$$Q_{\beta,nl,ut} = \delta_{nl}\delta_{ut} - \sum_j \Gamma_{\beta,nj,u}V_{\beta,jl,ut} \quad (101)$$

with

$$\Gamma_{\beta,nj,u}\delta_{\alpha\beta} = \sum_{nj}{}' U(\alpha, \mathbf{R}_n)\mathcal{G}_u(\mathbf{R}_n - \mathbf{R}_j)U^+(\mathbf{R}_j, \beta). \quad (102)$$

It will be observed that $f_{\beta,st}$ is proportional to the sum of matrix elements of the operator VQ^{-1}, evaluated on the basis of symmetrized linear combinations of plane waves belonging to the rows of the βth representation and belonging to bands s and t.

The physical interpretation of the scattering amplitude is the same as for Eq. (94).

ACKNOWLEDGMENTS

I wish to thank Professor P. E. Kaus for stimulating my interest in scattering theory and for many valuable discussions on these problems. I also wish to thank the staff and management of Bell Telephone Laboratories for their hospitality during a portion of the summer of 1963, when the writing of this paper was begun.

APPENDIX A

We will consider the application of the preceding general theory to the scattering of phonons by mass defects and to the scattering of spin waves by magnetic defects. An equation analogous to Eq. (8) is derived in each case. In the lattice dynamics problem, we consider a crystal with one atom in each unit cell. In the presence of defects, the time-independent equation of motion may be written as

$$\sum_{n,\beta} [M\omega^2\delta_{ln}\delta_{\alpha\beta} - \Phi_{\alpha\beta}(\mathbf{R}_l, \mathbf{R}_n)]U_\beta(\mathbf{R}_n) = \sum_{n,\beta} \Delta^{ln}_{\alpha\beta}U_\beta(\mathbf{R}_n), \quad (A1)$$

in which $U_\alpha(\mathbf{R}_n)$ is the αth rectangular component of the displacement of the atom in the cell centered at $\mathbf{R}_n$, M is the atomic mass, $\Phi_{\alpha\beta}$ is the potential energy matrix, and $\Delta^{ln}_{\alpha\beta}$ characterizes the defects.

The notation is that of Ref. 11 to which the reader is referred for additional information. It will be observed that this equation already has the form of Eq. (5), in which the component index α is analogous to the band index j and the displacements u replace the B's of that equation. Notice also that the energy is replaced by $M\omega^2$.

We can invert (A1) immediately to obtain

$$U_\alpha(\mathbf{R}_l) = U_\alpha^{(0)}(\mathbf{R}_l) + \sum_{\substack{\beta,\nu \\ m,n}} \mathcal{G}_{\alpha\beta}(\mathbf{R}_l - \mathbf{R}_n)\Delta_{\beta\nu}^{mn} U_\nu(\mathbf{R}_m), \tag{A2}$$

in which the Green's function $\mathcal{G}$ is a solution of

$$\sum_{m,\beta} [M\omega^2\delta_{\alpha\beta}\delta_{ln} - \Phi_{\alpha\beta}(\mathbf{R}_l, \mathbf{R}_m)] \times \mathcal{G}_{\beta\nu}(\mathbf{R}_m - \mathbf{R}_n) = \delta_{ln}\delta_{\alpha\nu}, \tag{A3}$$

and is given by

$$\mathcal{G}_{\alpha\beta}(\mathbf{R}_l - \mathbf{R}_m) = \frac{\Omega}{(2\pi)^3} \sum_j \int d^3k \frac{e_\alpha\left(\begin{matrix}\mathbf{k}\\ j\end{matrix}\right) e_\beta^*\left(\begin{matrix}\mathbf{k}\\ j\end{matrix}\right)}{M[\omega^2 - \omega_j^2(\mathbf{k})]} e^{i\mathbf{k}\cdot(\mathbf{R}_l - \mathbf{R}_m)}, \tag{A4}$$

in which the e_α are components of the eigenvectors of the dynamic matrix satisfying the equations

$$\sum_\alpha e_\alpha^*\left(\begin{matrix}\mathbf{k}\\ j\end{matrix}\right) e_\alpha\left(\begin{matrix}\mathbf{k}\\ j'\end{matrix}\right) = \delta_{jj'}, \qquad \sum_j e_\alpha^*\left(\begin{matrix}\mathbf{k}\\ j\end{matrix}\right) e_\beta\left(\begin{matrix}\mathbf{k}\\ j\end{matrix}\right) = \delta_{\alpha\beta}, \tag{A5}$$

$$\sum_\beta D_{\alpha\beta}(\mathbf{k}) e_\beta\left(\begin{matrix}\mathbf{k}\\ j\end{matrix}\right) = \omega_j^2(\mathbf{k}) e_\alpha\left(\begin{matrix}\mathbf{k}\\ j\end{matrix}\right), \tag{A6}$$

and

$$MD_{\alpha\beta}(\mathbf{k}) = \sum_{(\mathbf{R}_l - \mathbf{R}_m)} e^{i\mathbf{k}\cdot(\mathbf{R}_l - \mathbf{R}_m)} \Phi_{\alpha\beta}(\mathbf{R}_l, \mathbf{R}_m). \tag{A7}$$

Finally we note that $U^{(0)}$ is a solution of the homogeneous equation obtained from A(1) by replacing the right side by zero.

$$U_\alpha^{(0)}(\mathbf{R}_l) = [\Omega^{\frac{1}{2}}/(2\pi)^{\frac{3}{2}}] e^{i\mathbf{k}\cdot\mathbf{R}_l} e_\alpha\left(\begin{matrix}\mathbf{k}\\ j\end{matrix}\right). \tag{A8}$$

The proofs of these statements are straightforward. First one verifies that the expression (A8) for $U^{(0)}$ is correct:

$$\begin{aligned}\sum_{\beta,m} [M\omega^2\delta_{\alpha\beta}\delta_{lm} - \Phi_{\alpha\beta}(\mathbf{R}_l, \mathbf{R}_m)] U_\beta^{(0)}(\mathbf{R}_m) &= M\omega^2 U_\alpha^{(0)}(\mathbf{R}_l) - \sum_{\beta,m} e^{i\mathbf{k}\cdot\mathbf{R}_l} e^{i\mathbf{k}\cdot(\mathbf{R}_m - \mathbf{R}_l)} \Phi_{\alpha\beta}(\mathbf{R}_m, \mathbf{R}_l) e_\beta\left(\begin{matrix}\mathbf{k}\\ j\end{matrix}\right) \\ &= M\omega^2 U_\alpha^{(0)}(R_l) - M \sum_\beta e^{i\mathbf{k}\cdot\mathbf{R}_l} D_{\alpha\beta}(\mathbf{k}) e_\beta\left(\begin{matrix}\mathbf{k}\\ j\end{matrix}\right) \\ &= M[\omega^2 - \omega_j^2(\mathbf{k})] U_\alpha^{(0)}(R_l). \end{aligned} \tag{A9}$$

In the third line of (A9) we have converted the sum over m to one over $m - l$. Hence, we see that if $\omega = \omega_j(k)$, $U_\alpha^{(0)}(\mathbf{R}_l)$ is a solution of the homogeneous equation. It corresponds to an incident wave of definite frequency.

Next, one verifies that (A2) is a solution of (A1). This is accomplished easily by substitution of (A2) into (A1), with the use of (A3). The process is completed by showing that the Green's function given by (A4) satisfies (A3) as required. We have, on substitution,

$$\begin{aligned}&\frac{\Omega}{(2\pi)^3} \sum_{j\beta m} \int d^3k \frac{[M\omega^2\delta_{lm}\delta_{\alpha\beta} - \Phi_{\alpha\beta}(\mathbf{R}_l, \mathbf{R}_m)]}{M[\omega^2 - \omega_j^2(\mathbf{k})]} \times e_\beta\left(\begin{matrix}\mathbf{k}\\ j\end{matrix}\right) e_\nu^*\left(\begin{matrix}\mathbf{k}\\ j\end{matrix}\right) e^{i\mathbf{k}\cdot(\mathbf{R}_m - \mathbf{R}_n)} \\ &= \frac{\Omega}{(2\pi)^3} \sum_{j\beta} \int d^3k \frac{[\omega^2\delta_{\alpha\beta} - D_{\alpha\beta}(\mathbf{k})]}{[\omega^2 - \omega_j^2(\mathbf{k})]} \times e_\beta\left(\begin{matrix}\mathbf{k}\\ j\end{matrix}\right) e_\nu^*\left(\begin{matrix}\mathbf{k}\\ j\end{matrix}\right) e^{i\mathbf{k}\cdot(\mathbf{R}_l - \mathbf{R}_m)} \\ &= \frac{\Omega}{(2\pi)^3} \sum_j \int d^3k e_\alpha\left(\begin{matrix}\mathbf{k}\\ j\end{matrix}\right) e_\nu^*\left(\begin{matrix}\mathbf{k}\\ j\end{matrix}\right) e^{i\mathbf{k}\cdot(\mathbf{R}_l - \mathbf{R}_m)} \\ &= \frac{\Omega}{(2\pi)^3} \delta_{\alpha\nu} \int d^3k e^{i\mathbf{k}\cdot(\mathbf{R}_l - \mathbf{R}_m)} = \delta_{\alpha\nu}\delta_{lm}. \end{aligned} \tag{A10}$$

Hence, (A2) is analogous to (8) as required. In the case of a defect differing from the host lattice only through its mass ($M + \delta M$) and located at the origin, we have

$$\Delta_{\beta\nu}^{mn} = \delta_{\beta,\nu}\delta_{m,n}\delta_{n,0}\delta M\omega^2, \tag{A11}$$

so that (A2) takes the simpler form

$$U_\alpha(R_l) = U_\alpha^{(0)}(\mathbf{R}_l) + \delta M\omega^2 \sum_\beta \mathcal{G}_{\alpha\beta}(\mathbf{R}_l) U_\beta(0). \tag{A12}$$

This equation is similar to that obtained in the Koster–Slater model except that the "potential" here is frequency-dependent. In particular, the condition for resonance and bound states is just

$$\det [\delta_{\alpha\beta} - \delta M\omega^2 \mathcal{G}_{\alpha\beta}(0)] = 0. \tag{A13}$$

We next consider the spin-wave problem.[12] It is assumed that the spin waves are obtained from a simple isotropic exchange Hamiltonian,

$$\mathcal{H} = -\sum_{m,\Delta} J(\mathbf{R}_m, \mathbf{R}_m + \boldsymbol{\Delta})\mathbf{S}(\mathbf{R}_m)\cdot\mathbf{S}(\mathbf{R}_m + \boldsymbol{\Delta}). \tag{A14}$$

Here $\boldsymbol{\Delta}$ is a vector connecting an atom to its nearest

neighbors. We neglect any external magnetic field as largely irrelevant to the present discussion. Next we introduce the usual spin raising and lowering operators $S_\pm(\mathbf{R}) = S_x(\mathbf{R}) \pm iS_y(\mathbf{R})$. Then

$$\mathcal{H} = -\sum_{m,\Delta} J(\mathbf{R}_m, \mathbf{R}_m + \boldsymbol{\Delta})[S_z(\mathbf{R}_m)S_z(\mathbf{R}_m + \boldsymbol{\Delta}) + S_-(\mathbf{R}_m + \boldsymbol{\Delta})S_+(\mathbf{R}_m)]. \quad \text{(A15)}$$

The ground state of the ferromagnet is described by the ket $|0\rangle$ and characterizes a condition of complete spin alignment,

$$S_+(\mathbf{R}_m)\,|0\rangle = 0. \quad \text{(A16)}$$

We neglect the interaction between spin waves and define a complete set of orthonormal states corresponding to a single spin deviation at site R_l:

$$|l\rangle = [2S(\mathbf{R}_l)]^{-\frac{1}{2}}S_-(\mathbf{R}_l)\,|0\rangle. \quad \text{(A17)}$$

The operator $S_-(\mathbf{R}_l)$ lowers the z component of the spin at the site $\mathbf{R}_l$, and therefore creates a single spin deviation. $S(\mathbf{R}_l)$ is the "magnitude" of the spin at $\mathbf{R}_l$ in the sense that the eigenvalue of $\mathbf{S}^2(\mathbf{R}_l)$ in the ground state is $S(\mathbf{R}_l)[S(\mathbf{R}_l) + 1]$.

An arbitrary one spin-wave state $|\Psi\rangle$ may be expanded in the kets $|l\rangle$,

$$|\Psi\rangle = \sum_l \phi(\mathbf{R}_l)\,|l\rangle. \quad \text{(A18)}$$

Evidently, $\phi(\mathbf{R}_l)$ is the projection of $|\Psi\rangle$ on $|l\rangle$, and $|\phi(\mathbf{R}_l)|^2$ may be interpreted as the probability of finding a spin deviation at R_l. These functions are analogous to the quantities $B(R_l)$ introduced in the previous discussion. The equation satisfied by ϕ can be determined by writing the Schrödinger equation for $|\Psi\rangle$, and forming the scalar product with some bra $\langle m|$, so that we have

$$\sum_l \langle m|\,\mathcal{H}\,|l\rangle\phi(\mathbf{R}_l) = E\phi(\mathbf{R}_m). \quad \text{(A19)}$$

When the matrix elements are calculated, (A19) takes the form

$$2\sum_\Delta J(\mathbf{R}_l, \mathbf{R}_l + \boldsymbol{\Delta})[S(\mathbf{R}_l + \boldsymbol{\Delta})\phi(\mathbf{R}_l) - \{S(\mathbf{R}_l)S(\mathbf{R}_l + \boldsymbol{\Delta})\}^{\frac{1}{2}}\phi(\mathbf{R}_l + \boldsymbol{\Delta})] = E\phi(\mathbf{R}_l), \quad \text{(A20)}$$

in which the zero of energy has been chosen to correspond to the ground-state energy of the system. We now suppose that a defect which has spin S' and exchange integral J' coupling it to its neighbors is located at the origin. (The host lattice has these quantities equal to S_0 and J_0, respectively).

We wish to write (A20) in a form equivalent to Eq. (5). To this end, we define

$$\langle m|\,E - H_0\,|l\rangle = (E - 2J_0S_0z)\delta_{l,m} + 2J_0S_0\delta_{R_l-R_m,\Delta}, \quad \text{(A21)}$$

in which Z is the number of nearest neighbors. Also,

$$\langle m|\,V\,|l\rangle = \delta_{R_l,R_m} \times [2Sz(J' - J_0)\delta_{R_l,0} + 2(J'S' - J_0S_0)\delta_{R_l,\Delta}] - 2[J'(S'S_0)^{\frac{1}{2}} - J_0S_0]\delta_{R_l-R_m,\Delta}(\delta_{R_l,0} + \delta_{R_m,0}). \quad \text{(A22)}$$

We may now pass directly to Eq. (8). This equation now has the form

$$\phi(\mathbf{R}_m) = \phi^{(0)}(\mathbf{R}_m) + \sum_{pn} \mathcal{G}(\mathbf{R}_m - \mathbf{R}_p)\langle p|\,V\,|n\rangle\phi(\mathbf{R}_n), \quad \text{(A23)}$$

in which $\phi^{(0)}$ is a solution of

$$\sum_l \langle m|\,H_0\,|l\rangle\phi_0(\mathbf{R}_l) = E(\mathbf{k})\phi^{(0)}(\mathbf{R}_m), \quad \text{(A24)}$$

and $\mathcal{G}$ satisfies

$$\sum_l \langle m|\,E - H_0\,|l\rangle\mathcal{G}(\mathbf{R}_l - \mathbf{R}_n) = \delta_{m,n}. \quad \text{(A25)}$$

We easily find that

$$\phi^{(0)}(\mathbf{R}_l) = [\Omega^{\frac{1}{2}}/(2\pi)^{\frac{3}{2}}]e^{i\mathbf{k}\cdot\mathbf{R}_l}, \qquad E(\mathbf{k}) = 2J_0S_0(z - \sum_\Delta \cos\mathbf{k}\cdot\boldsymbol{\Delta}). \quad \text{(A26)}$$

The Green's function $\mathcal{G}(\mathbf{R}_m - \mathbf{R}_p)$ is given by

$$\mathcal{G}(\mathbf{R}_m - \mathbf{R}_p) = \frac{\Omega}{(2\pi)^3}\int d^3k\,\frac{e^{i\mathbf{k}\cdot(\mathbf{R}_m-\mathbf{R}_p)}}{E - E(\mathbf{k})}, \quad \text{(A27)}$$

just as in Eq. (17).

One point of difference arising in the consideration of defect scattering of phonons and spin waves should be noted. In the case of phonon scattering, we have in the case of the isotopic impurity, a physical situation approximating the Slater–Koster model of an impurity potential which is different from zero only in a single cell. In the case of spin waves, however, a magnetic defect influences the coupling of spins, and therefore the effective potential extends over nearest neighbors of the defect.

APPENDIX B

In this appendix, we will consider the justification for the use of the term "scattering amplitude" to describe the coefficient of e^{ikR}/R in the expression for $B(\mathbf{R})$. To do this, it is desirable to consider the probability current density vector. We express the current density as

$$\mathbf{j} = \tfrac{1}{2}[\psi^*\mathbf{V}\psi + \psi\mathbf{V}^*\psi^*], \quad \text{(B1)}$$

in which $\mathbf{V}$ is the velocity operator. In the present

case, we introduce a lattice vector operator conjugate to the wave vector,[26] and write

$$V = (i/\hbar)[HR - RH]. \tag{B2}$$

We wish to evaluate (B1) using (B2) in the scattering problem. Therefore we work in a representation in which the wavefunctions ψ appearing in (B1) are the scattering functions $B_s(\mathbf{R}_n)$. This interpretation is justified physically since $|B_s(\mathbf{R}_n)|^2$ gives the probability of finding the excitation on the site located at $\mathbf{R}_n$. It is necessary to determine the velocity. To do this, we consider a time-dependent form of Eq. (5) for values of $\mathbf{R}_n$ and $\mathbf{R}_m$ sufficiently large so that the matrix elements of the potential vanish. This is

$$\sum (s, m|\ \delta_{mn} i\hbar(\partial/\partial t) - H_0\ |s, n) B_s(\mathbf{R}_n) = 0. \tag{B3}$$

The matrix elements of H_0 can be obtained from Ref. 25, Sec. 2.11,

$$(s, m|\ H_0\ |S, n) = \mathcal{E}_s(\mathbf{R}_m - \mathbf{R}_n), \tag{B4}$$

in which

$$\mathcal{E}_s(\mathbf{R}_m - \mathbf{R}_n) = \frac{\Omega}{(2\pi)^3} \int e^{i\mathbf{k}\cdot(\mathbf{R}_m - \mathbf{R}_n)} E_s(\mathbf{k})\, d^3k, \tag{B5}$$

and the integration includes the Brillouin zone.

We now express (B5) in terms of a differential operator by the following device. The functions B_s in which we are interested are characterized by some wave vector, $\mathbf{k}_0$. We expand the energy as a power series in k around $\mathbf{k}_0$,

$$\mathcal{E}_s(\mathbf{k}) = \sum_{l=0}^{\infty} \left[\frac{(\mathbf{k} - \mathbf{k}_0)\cdot\nabla_k}{l!}\right]^l E_s = \sum_{l=0}^{\infty} \frac{(\boldsymbol{\kappa}\cdot\nabla)^l}{l!} E_s$$
$$= \sum_{pqr} A_{\alpha\beta\gamma}^{(p,q,r)} \kappa_\alpha^p \kappa_\beta^q \kappa_\gamma^r. \tag{B6}$$

In this equation, we have defined

$$\boldsymbol{\kappa} = \mathbf{k} - \mathbf{k}_0,$$

and κ_α, κ_β, κ_γ are rectangular components of $\boldsymbol{\kappa}$. The quantities $A_{\alpha\beta\gamma}^{(p,q,r)}$ are coefficients which depend on the energy and its derivatives at the point $\mathbf{k}_0$. Eq. (B6) is inserted into (B5), which becomes

$$\mathcal{E}_s(\mathbf{R}_m - \mathbf{R}_n) = \frac{\Omega}{(2\pi)^3} e^{i\mathbf{k}_0\cdot(\mathbf{R}_m - \mathbf{R}_n)}$$
$$\times \sum_{pqr} A_{\alpha\beta\gamma}^{(pqr)} \int e^{i\boldsymbol{\kappa}\cdot(\mathbf{R}_m - \mathbf{R}_n)} \kappa_\alpha^p \kappa_\beta^q \kappa_\gamma^r\, d^3\kappa.$$

At this point, we treat the lattice vectors as continuous variables so that we may write

$$\mathcal{E}_s(\mathbf{R}_m - \mathbf{R}_n) = \frac{\Omega}{(2\pi)^3} e^{i\mathbf{k}_0\cdot(\mathbf{R}_m - \mathbf{R}_n)}$$
$$\times \sum_{pqr} A_{\alpha\beta\gamma}^{(pqr)} \left(-i\frac{\partial}{\partial R_{m\alpha}}\right)^p \left(-i\frac{\partial}{\partial R_{m\beta}}\right)^q \left(-i\frac{\partial}{\partial R_{m\gamma}}\right)^r$$
$$\times \int e^{i\boldsymbol{\kappa}\cdot(\mathbf{R}_m - \mathbf{R}_n)}\, d^3\kappa = \delta_{mn} E_s(-i\nabla_m). \tag{B7}$$

Then Eq. (B3) becomes

$$i\hbar(\partial/\partial t) B(\mathbf{R}) = E_s(-i\nabla_{\mathbf{R}}) B(\mathbf{R}). \tag{B8}$$

Evidently the Hamiltonian in this representation is just $E_s(-i\nabla_{\mathbf{R}})$. We can now proceed to evaluate the commutator in (B2). This is done by expanding the operator $E_s(-i\nabla_{\mathbf{R}})$. Evidently only the linear term in the expansion contributes, and we find

$$\mathbf{V} = (1/\hbar)(\nabla_k E_s)_{k_0}. \tag{B9}$$

This gives

$$\mathbf{j} = \hbar^{-1}(\nabla_k E_s)|B(\mathbf{k}_0, \mathbf{R})|^2, \tag{B10}$$

which is just the result one would expect.

The coefficients $B(\mathbf{R})$ have the general form

$$B(\mathbf{R}) = \frac{\Omega^{\frac{1}{2}}}{(2\pi)^{\frac{3}{2}}} \left[e^{i\mathbf{k}_0\cdot\mathbf{R}} + f\frac{e^{ik'R}}{R}\right] \tag{B11}$$

for a scattering problem. Then the probability per unit time that the scattered particle will pass through an element of area $ds = R^2 d\omega$ at a distance R from the scattering center is $|\nabla_k E_s|_{k'}^2 \cdot |f|^2 d\omega$. The ratio of this to the probability current in the incident wave is

$$d\sigma = [|\nabla_k E_s|_{k'}^2 / |\nabla_k E_s|_{k_0}]\ |f|^2\, d\omega. \tag{B12}$$

There is no interference between the two terms of (B10) if we assume that the incident beam of particles is actually of limited extent (a wave packet), in which case the current in the scattered wave is evaluated at a point at which there is actually no incident wave. This restriction leads to the conventional definition of the differential cross section.

$$\sigma_d = d\sigma/d\omega = [|\nabla_k E_s|_{k'}^2 / |\nabla_k E_s|_{k_0}]\ |f|^2. \tag{B13a}$$

If we consider energies low enough so that the band is spherical, $\mathbf{k}' = \mathbf{k}_0$, and

$$\sigma_d = |f|^2. \tag{B13b}$$

One further point is worth discussing. The resemblance of $B_s(\mathbf{R})$ to a wavefunction of ordinary scattering theory is obvious. In addition, it can be shown that in the limit of (1) a vanishing crystal (periodic) potential, and (2) vanishing lattice constant, B becomes just such a wavefunction.

[26] G. Wannier, *Elements of Solid State Theory* (Cambridge University Press, London, 1959), p. 174.

If the crystal potential vanishes, the Bloch functions are just plane waves,

$$\psi(\mathbf{k}, \mathbf{r}) = [1/(2\pi)^{\frac{3}{2}}]e^{i\mathbf{k}\cdot\mathbf{r}}. \tag{B14}$$

The Wannier functions are

$$a(\mathbf{r} - \mathbf{R}_n) = \frac{\Omega^{\frac{1}{2}}}{(2\pi)^{\frac{3}{2}}} \int d^3k e^{-i\mathbf{k}\cdot\mathbf{R}_n} \psi(\mathbf{k}, \mathbf{r}).$$

In the limit of vanishing lattice constant, the Brillouin zone expands to include all k space. Also there is only one band,

$$a(\mathbf{r} - \mathbf{R}) = \frac{\Omega^{\frac{1}{2}}}{(2\pi)^3} \int e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{R})}\, d^3r = \Omega^{\frac{1}{2}}\delta(\mathbf{r} - \mathbf{R}). \tag{B15}$$

Then, referring to Eq. (4) of the main text, we see that we must replace the summation over lattice vectors by an integral

$$\sum_n \rightarrow \frac{1}{\Omega} \int d^3R.$$

Hence, Eq. (4) becomes

$$\Psi(\mathbf{r}) = \frac{1}{\Omega} \int d^3R B(\mathbf{R}) a(\mathbf{r} - \mathbf{R}) = \Omega^{-\frac{1}{2}} B(\mathbf{r}). \tag{B16}$$

The factor of $\Omega^{-\frac{1}{2}}$ in (B16) may seem peculiar, but it is required to cancel a factor of $\Omega^{\frac{1}{2}}$ in (B11) and restore the proper normalization to the scattering wavefunction. Also, we evidently have $E(k) = \hbar^2k^2/2m$, so that from Eq. (5) and (B8), it follows that $B(\mathbf{R})$ satisfies the ordinary Schrödinger equation. We have therefore established that the present theory reduces properly to the ordinary scattering theory of elementary quantum mechanics.

Singular Potentials and Peratization. II

A. PAIS
The Rockefeller Institute, New York, New York
and
TAI TSUN WU*
Gordon McKay Laboratory, Harvard University, Cambridge, Massachusetts
(Received 10 February 1964)

In a field theory with more than one unrenormalizable interaction (like the W theory with weak and electromagnetic coupling), the important problem arises of resumming multiple series of individually divergent terms. In order to get a first insight in the new questions which arise for multiple as compared to single series, we study the analogous question for a superposition of two singular potentials, using a family of exactly soluble cases. We ask whether one can expand resummed series for the zero-energy scattering amplitude in powers of one coupling constant with coefficients depending on the other. The answer depends both on the relative magnitude of the coupling constants and on the relative degree of singularity of the interactions. Depending on these two conditions one finds three regimes, one where a convergent power expansion holds, another where an asymptotic expansions obtains, and a third where it is impossible to expand in powers of either single constant separately. It is conjectured that a similar situation will be true in a field theory with leading power singularities only (and no logarithmic ones), if such a theory has meaning.

I. INTRODUCTION

IN a previous paper,[1] certain properties of repulsive singular power potentials were studied by means of the method of summing infinite series of individually divergent terms. In particular such properties were considered for which a closed-form answer is available by other means. The whole point of the work was in fact to get a clearer understanding of the peratization methods by applying it to questions where the answer is guaranteed to exist, something which so far has always to be an assumption in the case of real interest, namely unrenormalizable field theory. In the present paper we study more complicated singular potentials. Their general character is the superposition of two singular potentials, each with its own coupling constant. It should at once be stressed that we do not do this because we wish just to generalize the results found in I for a single power potential. While the motivation for the present investigation stems again from field theory, it is different from the one which led to the previous work, as is now to be explained.

Suppose we have a field theory with an unrenormalizable interaction characterized by a coupling constant ρ_1 which is small. We like to get a meaningful answer for some quantity like a scattering amplitude by first summing the leading singular parts of the terms in the perturbation series. As shown in concrete cases,[2,3] this leads one to ask whether meaning can be given to expressions like

$$\lim_{\Lambda\to\infty} \Lambda^{\alpha} \sum_n a_n \rho_1^n \Lambda^{\beta n}, \tag{1.1}$$

where Λ is a (dimensionless) cutoff, the series is an infinite one, and the numbers a_n, α, and β are characteristic for the problem at hand. (For simplicity we take an example where all the leading singular terms are power singularities.) We then say: if the theory exists and the method has meaning, Eq. (1.1) gives the result $\phi(\infty)\rho_1^{-\alpha/\beta}$ where $\phi(\infty)$ is a number defined by

$$\begin{aligned} \phi(\infty) &= \lim_{z\to\infty} \phi(z), \\ \phi(z) &= z^{\alpha/\beta} \sum_n a_n z^n \qquad (z = \rho_1 \Lambda^{\beta}). \end{aligned} \tag{1.2}$$

If we can define a limiting process which leads to a finite $\phi(\infty)$, we then proceed to sum the next leading singular terms, etc.[4]

Let us now consider a theory with two unrenormalizable interactions, coupling constants ρ_1 and ρ_2. Let the leading singular term $\sim\rho_1^m\rho_2^n$ again be a power singularity (for the sake of argument). Then instead of Eq. (1.1) we must try to give meaning to

---

* Alfred P. Sloan Foundation Fellow.

[1] N. Khuri and A. Pais, Rev. Mod. Phys. (to be published). This paper is hereafter referred to as I. See, also, G. Tiktopoulos and S. B. Treiman, "Weak Coupling Limit for Scattering by Strongly Singular Potentials" (to be published).

[2] G. Feinberg and A. Pais, Phys. Rev. **131**, 2724 (1963).

[3] G. Feinberg and A. Pais, Phys. Rev. **133**, B477 (1964).

[4] For general features of the method see, e.g., A. Pais, "Methods and Problems in the Dynamics of Weak Interactions," in Proceedings of the Siena Conference on Elementary Particles, 1963 (to be published).

$$\lim_{\Lambda\to\infty} \Lambda^{\alpha} \sum_{m,n} a_{mn}\rho_1^m\rho_2^n\Lambda^{\beta m+\gamma n}, \tag{1.3}$$

where the numbers a_{mn} are independent of ρ_1, ρ_2, and Λ. The question now arises: how do we treat this infinite *double* sum? What is the analog of introducing the variable z as we did above?

It is evident that we need a *further* prescription when we deal with infinite double sums. Let us make an example. Suppose we try to represent Eq. (1.3) as follows by a single infinite series,

$$\sum_n c_n(\rho_1)\rho_2^n, \tag{1.4}$$

where

$$c_n(\rho_1) = \rho_1^{-(\alpha+\gamma n)/\beta}\phi_n(\infty),$$

$$\phi_n(\infty) = \lim_{z\to\infty} \phi_n(z),$$

$$\phi_n(z) = z^{(\alpha+\gamma n)/\beta} \sum_m a_{mn}z^m, \qquad (z = \rho_1\Lambda^{\beta}). \tag{1.5}$$

For this to be sensible the following need be true. First, for all n, $\phi_n(\infty)$ should exist. Secondly, the power series expansion in ρ_2 given by Eq. (1.4) should be possible, if not as a convergent expansion, then at least as an asymptotic one. When is it possible to have such an expansion in ρ_1 with coefficients in ρ_2? Or one with the roles of ρ_1 and ρ_2 interchanged? Or one in a variable made up out of both ρ_1 and ρ_2?

Evidently this problem is far more complex than the one for a single interaction. The problem is also extremely acute. To give the perhaps most important example: the W-meson theory of weak interactions is unrenormalizable. So is W electrodynamics. Consider for simplicity (if for no better reasons[5]) the case where the bare gyromagnetic ratio of W mesons = 1, (the case known[6] as $\kappa = 0$.) Then we have a two-coupling problem and we must ask such physical questions as the influence of the electromagnetic interaction on the "purely weak" phenomena; and vice versa. As long as we have no established technique for two-coupling problems, we certainly cannot be content to say that certain effects can be ignored just because of the smallness of some coupling constant. Another example of two such couplings is W electrodynamics with $\kappa \neq 0$; and there are more complicated situations.

For the "weak-plus electromagnetic" problem it has been conjectured[5] that the main damping of the higher-order perturbation contributions comes from electromagnetic effects, for both electromagnetic and weak phenomena. That is to say, one should first peratize in the electromagnetic effects and then obtain a series of the structure of equation (1.4) in powers of the weak coupling constant with coefficients in the electromagnetic coupling constants. However, it seems difficult to reconcile this conjuncture with the results of a recent study[7] of the combined weak and electromagnetic effects for $\kappa \neq 0$. We do not enter at this time in any further detail about these questions, nor do we discuss here the problem "weak vs electromagnetic" for the case $\kappa = 0$ (which is more delicate). All we wish to bring out for the moment is that there are new problems to be faced when we have to do with multiple series of divergent terms. So far there exists no related situation in physics that we know of where such questions are treated. This is the reason why we thought it worthwhile to examine the case of two singular potentials. As in I, we need not worry about existence of answers. Also, as in I, we shall deal with a number of cases where the answer is known in closed form by other means than the summation of series of divergent terms.

In Sec. II we discuss potentials[8] of the form $g^2x^{-2\tau-2} + fx^{-\tau-2}$ with $\tau > 1$. It was explained in I why we can confine ourselves to the zero-energy case and the same is true here. The zero-energy wave function can be given in closed form; hence the same is true for the zero-energy scattering amplitude. We consider the latter quantity as a function of f and g^2 and ask (Sec. III) when we can expand in f (for fixed g) or in g^2 (for fixed f). The results are summarized in Fig. 1. There is a region (1) where the expansion in f converges. There is no region where the expansion in g converges. However, there is a domain near the positive f axis (2) where one has an asymptotic expansion in g. The extent of this domain has the vagueness inavoidably connected with asymptotic series. There is a remaining domain (3) where one can expand neither in f nor in g; only double series can be used here. Finally, for $g = 0$, $f < 0$ (4), there does not exist a scattering amplitude, as follows from the well-known work of Case.[9] In terms of the present methods, what happens in this instance is the

[5] J. Bernstein and T. D. Lee, Phys. Rev. Letters **11**, 512, (1963).

[6] T. D. Lee and C. N. Yang, Phys. Rev. **128**, 885 (1962); T. D. Lee, *ibid.* **128**, 899 (1962).

[7] Y. Pwu and T. T. Wu, Phys. Rev. **133B**, 1299 (1964). T. T. Wu, in Weak Interaction Conference, Brookhaven National Laboratory, September 1963 (to be published).

[8] The special case $\tau = 1$ occurs in the work of E. Predazzi and T. Regge, Nuovo Cimento **24**, 518 (1962). For $\tau = 1$ there are some minor complications at large distances which we wish to avoid as they have no bearing on the problem which interests us here.

[9] K. M. Case, Phys. Rev. **80**, 797 (1950).

following. As long as the attractive singular potential is cut off, the scattering amplitude contains a well defined quotient of functions of the cutoff. However, the limit of the quotient does not exist as the cutoff tends to zero [Eq. (3.4) below].

What we have therefore learned can be summarized as follows. We have two couplings, the more singular one characterized by g, the less singular one by f. If the less singular one is also weak compared to the more singular one ($f \ll g$, both quantities have the same dimension), then a convergent power series in the "weak" one is possible. On the other hand, when $g \ll f$ one still cannot develop in g in a convergent way, though there does exist an asymptotic expansion in g, *as long as the theory makes sense for* $g = 0$.

We have shown this for a specific relative degree of singularity. In terms of the general form [Eq. (1.3)] of a double expansion, we have $\rho_1 = f$, $\rho_2 = g^2$, and $\gamma = 2\beta$. We do not believe that the specific relation $\gamma = 2\beta$ is material for the validity of the above result and we would like to state the following conjecture for the potential problem:

Whenever we are led to an expansion like Eq. (1.3) and $\gamma > \beta$, then for sufficiently small $\rho_1^{1/\beta}\rho_2^{-1/\gamma}$ there will exist a convergent expansion in ρ_1. And for sufficiently small $\rho_2^{1/\gamma}\rho_1^{-1/\beta}$ there will exist an asymptotic expansion in ρ_2, as long as the theory makes sense for $\rho_2 = 0$.

If (and only if) this is indeed the case, then we would be encouraged to think that the same situation regarding expansions in ρ_1 and ρ_2 may hold in unrenormalizable field theory. It should be stressed that we only wish to entertain such conjectures in the absence of log Λ factors in the double sum.

We now turn to the mathematical details.

II. A SPECIAL CLASS OF POTENTIALS

We discuss here a special class of potentials for which solutions can be given in explicit form, namely,

$$V(r) = \frac{g^2}{r^{2+2\tau}} + \frac{f}{r^{2+\tau}}, \qquad \tau > 1, \tag{2.1}$$

where we shall take $g > 0$ by convention; f may be of either sign. The s-wave Schrödinger equation at zero energy is

$$d^2\psi/dr^2 - V(r)\psi = 0, \tag{2.2}$$

which we need to solve under the following boundary conditions:

$$\psi(0) = 0, \qquad \psi(r) \sim r \quad \text{as} \quad r \to \infty. \tag{2.3}$$

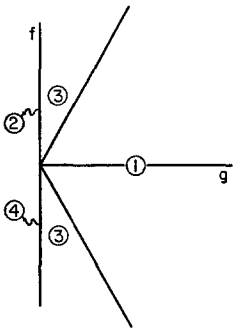

FIG. 1. The different expansion regimes in the f-g half-plane. 1: domain of convergent power series in f. 2 ($g = 0, f > 0$): asymptotic series in g. 3: no single power series in f or g is possible. 4 ($g = 0$, $f < 0$): the theory has no meaning. 1 is separated from 3 by $|f| = g(\tau + 1)$.

Put

$$\psi(r) = r^{\frac{1}{2}(1+\tau)}\phi(y), \qquad y = 2g/(\tau r^\tau). \tag{2.4}$$

Then ϕ satisfies

$$\frac{d^2\phi}{dy^2} + \left[-\frac{1}{4} - \frac{f}{2g^\tau}\frac{1}{y} + \left(\frac{1}{4} - \frac{1}{4\tau^2}\right)\frac{1}{y^2}\right]\phi = 0, \tag{2.5}$$

which is the Whittaker equation.[10] Put

$$\frac{1}{2}\left(1 + \frac{f}{g\tau} + \frac{1}{\tau}\right) = a. \tag{2.6}$$

$$1 + 1/\tau = c. \tag{2.7}$$

Then the solution for $\psi(r)$ is

$$\psi(r) = e^{-\frac{1}{2}y}\Psi(a, c; y). \tag{2.8}$$

Here Ψ is that confluent hypergeometric function which allows us to satisfy the boundary conditions. We have[11]

$$\Psi(a, c; y) = \frac{\Gamma(1-c)}{\Gamma(a-c+1)}\Phi(a, c; y) + \frac{\Gamma(c-1)}{\Gamma(a)}y^{1-c}\Phi(a-c+1, 2-c; y). \tag{2.9}$$

Φ is the confluent hypergeometric function with the small argument expansion [B, p. 248, Eq. (1)]

$$\Phi(a, c; y) = 1 + \frac{a}{c}y + \frac{a(a+1)}{c(c+1)}\frac{y^2}{2!} + \cdots. \tag{2.10}$$

To find the behavior of $\psi(r)$ near $r = 0$ we need the asymptotic behavior of $\Psi(a, c; y)$ as $y \to \infty$ [see Eq. (2.4)]. For all a, c this is a power behavior [B, p. 278, Eq. (1)], which is sufficient to satisfy the first of the conditions (2.3). It follows further from Eqs. (2.4, 8, 9) that, for large r, $\psi(r) = \text{const.}\ (r + A)$. Here A is the zero-energy scattering amplitude and is found to be given by

[10] See Bateman Manuscript Project, *Higher Transcendental Functions*, edited by A. Erdelyi (McGraw-Hill Book Company, Inc., New York, 1953), Vol. I, p. 248. This volume of the Bateman Project is hereafter referred to as B.

[11] Notations are as in B; see p. 257, Eq. (7).

$$A = \left(\frac{2g}{\tau}\right)^{1/\tau} \frac{\Gamma\left(-\frac{1}{\tau}\right)\Gamma\left[\frac{1}{2}\left(1 + \frac{1}{\tau} + \frac{f}{g\tau}\right)\right]}{\Gamma\left(\frac{1}{\tau}\right)\Gamma\left[\frac{1}{2}\left(1 - \frac{1}{\tau} + \frac{f}{g\tau}\right)\right]}. \tag{2.11}$$

In I we obtained the following closed form for the zero-energy scattering amplitude for a single power potential.[12]

$$V = \frac{g^2}{x^m}, \quad m > 3 \rightarrow A = -(\nu g)^{2\nu} \frac{\Gamma(1-\nu)}{\Gamma(1+\nu)},$$

$$\nu = \frac{1}{m-2}. \tag{2.12}$$

It is easily seen that Eq. (2.11) reduces to Eq. (2.12) for $f = 0$, by means of [B, p. 5, Eq. (15)]

$$\Gamma(2\nu) = 2^{2\nu-1}\pi^{-\frac{1}{2}}\Gamma(\nu)\Gamma(\nu + \tfrac{1}{2}).$$

Also, for $g = 0$, we get Eq. (2.12) (with $g^2 \rightarrow f$, $m \rightarrow \tau + 2$) from Eq. (2.11) by means of the Stirling expansion.

For the sake of illustration, let us see how for this potential problem one gets double power series of the type of Eq. (1.3) by considering a closed form for the scattering amplitude in the presence of a regulated version of the potential (2.1). As in I we choose the case $V(r) \rightarrow V(r + \alpha)$, $\alpha > 0$. Instead of Eq. (2.8) we now have

$$\psi_\alpha(r) = e^{-\frac{1}{2}y'}[\Psi(a, c; y') + \rho\Phi(a, c; y')],$$
$$y' = 2g/[\tau(r+\alpha)^\tau],$$

where ρ is fixed by $\psi_\alpha(0) = 0$. It is then easy to write down the zero-energy amplitude $A(\alpha)$ for this case. With the help of Eq. (2.9) one finds

$$A(\alpha) = -\alpha\left\{\frac{\Phi\left[\frac{1}{2}\left(1 + \frac{f}{g\tau} - \frac{1}{\tau}\right), 1 - \frac{1}{\tau}; \frac{2g}{\tau\alpha^\tau}\right]}{\Phi\left[\frac{1}{2}\left(1 + \frac{f}{g\tau} + \frac{1}{\tau}\right), 1 + \frac{1}{\tau}; \frac{2g}{\tau\alpha^\tau}\right]} - 1\right\}. \tag{2.13}$$

This is to be compared with the corresponding expression for the single power potential given in Eq. (2.12) for which it was found that[13]

$$A_g(\alpha) = -\alpha\left[\frac{(\tfrac{1}{2}Z)^\nu I_{-\nu}(Z)\Gamma(1-\nu)}{(\tfrac{1}{2}Z)^{-\nu} I_\nu(Z)\Gamma(1+\nu)} - 1\right], \tag{2.14}$$
$$Z = 2\nu g\alpha^{-1/(2\nu)},$$

with ν given by Eq. (2.12). For $f = 0$, Eq. (2.13) reduces to Eq. (2.14) as [B, p. 265, Eq. (1)]

$$\Gamma(\nu+1)I_\nu(x) = (\tfrac{1}{2}x)^\nu e^{-x}\Phi(\tfrac{1}{2} + \nu, 1 + 2\nu; 2x).$$

Also, for $g = 0$ and *positive* f we obtain Eq. (2.14) (with $g^2 \rightarrow f$) from Eq. (2.13), as[14]

$$\lim_{a\rightarrow\infty} \Phi(a, c; x/a) = \Gamma(c)x^{\frac{1}{2}-\frac{1}{2}c} I_{c-1}(2x^{\frac{1}{2}}).$$

As a final limiting case we consider

$$g = 0, \qquad f < 0, \tag{2.15}$$

which, for $\alpha \neq 0$, corresponds to an *attractive nonsingular* potential. With the help of[15]

$$\lim_{\alpha\rightarrow\infty} \Phi(a, c; -x/a) = \Gamma(c)x^{\frac{1}{2}-\frac{1}{2}c} J_{c-1}(2x^{\frac{1}{2}}),$$

we get for the zero scattering amplitude $A_f^-(\alpha)$

$$A_f^-(\alpha) = -\alpha\left[\frac{(\tfrac{1}{2}\zeta)^\nu J_{-\nu}(\zeta)\Gamma(1-\nu)}{(\tfrac{1}{2}\zeta)^{-\nu} J_\nu(\zeta)\Gamma(1+\nu)} - 1\right], \tag{2.16}$$
$$\zeta = 2\nu\,|f|^{\frac{1}{2}}\,\alpha^{-1/(2\nu)}, \qquad \nu = 1/\tau.$$

We conclude this section by noting the relation between Eqs. (2.11) and (2.13),

$$\lim_{\alpha\rightarrow 0} A(\alpha) = A, \tag{2.17}$$

which follows from the asymptotic behavior of the Φ's, [B, p. 278, Eq. (3)]

$$\Phi(a, c; x) \simeq [\Gamma(c)/\Gamma(a)]e^x x^{a-c}. \tag{2.18}$$

Here we find a similar analytic behavior as a function of α as was noted in I for a single power potential. Both numerator and denominator in Eq. (2.13) have an essential singularity in α at $\alpha = 0$, while the quotient is well-behaved in this limit.

III. PERATIZATION PROBLEMS FOR THE SPECIAL POTENTIALS

Consider the formal double power series expansion of $A(\alpha)$ [Eq. (2.13)], in f and g^2. From Eq. (2.10),

$$A(\alpha) = -\alpha\left[\frac{2}{\tau^2-1}\frac{f}{\tau\alpha^\tau} + \frac{1}{\tau^2(4\tau^2-1)} \times \left(\frac{2-5\tau}{\tau^2-1}f^2 + \tau g^2\right)\frac{1}{\alpha^{2\tau}} + \cdots\right], \tag{3.1}$$

or generally,

$$A(\alpha) = \alpha \sum_{m+n\geq 1} c_{mn}\left(\frac{f}{\alpha^\tau}\right)^m\left(\frac{g}{\alpha^\tau}\right)^{2n}. \tag{3.2}$$

[12] See I Eq. (2.19). For easier comparison we have replaced the symbols a and g of I by A_g and g^2, respectively.

[13] See I, Eqs. (2.25, 26) and also footnote 12 of the present paper.

[14] B, p. 266, Eq. (18). I is a Bessel function for imaginary argument.

[15] B, p. 266, Eq. (16). J is the Bessel function of the first kind.

The coefficients c_{mn} do not depend on α, nor on f and g. (They depend of course on τ but that plays no role in the general considerations.) It should be stressed that the formal expansion (3.2) holds regardless of the specific way in which one regularizes, as follows on dimensional grounds. Of course the c_{mn} depend on the details of the regularization.

Let us first discuss the special cases in which one of the two coupling constants vanish. For $f = 0$, the peratization argument goes as follows. The limit value A of $A(\alpha)$ for $\alpha = 0$ is

$$A = g^{1/\tau} \lim_{z\to\infty} F(z),$$
$$F(z) = z^{-1/\tau} \sum_{n=1} c_{0n} z^n; \qquad z = g/\alpha^\tau. \tag{3.3}$$

If we can give meaning to $F(\infty)$ and if $F(\infty) \neq 0$ then $A = \text{const } g^{1/\tau}$. In the present case we know of course that all is well and that we end up with the answer given in Eq. (2.12). The same reasoning applies to the case $g = 0$, f *positive*.

The case $g = 0$, f *negative* is especially interesting. As follows from Eq. (2.16), the formal expansion (3.2) applies here as well, with only $c_{m0} \neq 0$. Thus if an appropriate $F(\infty)$ were to exist we would have $A = \text{const. } |f|^{1/\tau}$. However, as we are now dealing with an *attractive singular* potential, it follows[9] that there cannot exist an unambiguously defined limit value $F(\infty)$. This is at once evident from Eq. (2.16). From this equation we obtain the formal limit expression

$$\frac{\Gamma\left(1-\frac{1}{\tau}\right)}{\Gamma\left(1+\frac{1}{\tau}\right)} \left(\frac{|f|}{\tau^2}\right)^{1/\tau} \lim_{z\to\infty} \frac{J_{-\nu}(z)}{J_\nu(z)}, \tag{3.4}$$

which clearly has no meaning. This region of breakdown of peratization, due to the breakdown of the theory, is marked 4 in Fig. 1.

We now turn to the general case in which both f and g are nonzero and where f may be of either sign. In accordance with the general considerations given in the Introduction we shall examine whether it is possible to expand the amplitude A in a power series in one of the coupling constants, while the other one has a fixed value.

A. g Fixed, Series in f

We ask under which conditions the expansion

$$A = \sum_{m=0} c_m(g) f^m \tag{3.5}$$

is meaningful. In the framework of the peratization program we must phrase this problem as follows. Start from Eq. (3.2) and consider the order of limits in which one first arrives at each individual $c_m(g)$ by the limit process

$$c_m(g) = \frac{1}{g^m} \lim_{z\to\infty} G_m(z),$$
$$G_m(z) = \sum_{n=1} c_{mn} z^{2n+m}. \tag{3.6}$$

Thus we need to know whether meaning can be given to $G_m(\infty)$, $m = 0, 1, 2, \cdots$. If this is possible, one then discusses the convergence of the expansion of Eq. (3.5) with the $c_m(g)$ so obtained.

Of course, the virtue of dealing with a problem where the explicit answer for A is known is that we can use this answer itself to find out where Eq. (3.5) is meaningful. All we have to do in fact is to find the radius of convergence of the f power series expansion of Eq. (2.11) for fixed g. It should therefore be emphasized that the answer to our question is independent of any particular mode of regularization. We shall therefore have no need for the explicit regularization given in Eq. (2.13) and its expansion Eq. (3.1). It was nevertheless believed to be instructive to exhibit this example.

As $1/\Gamma(z)$ is an entire function of z, it follows that the radius of convergence of Eq. (3.5) is determined by the nearest pole of $\Gamma[\frac{1}{2}(1+f/g\tau+1/\tau)]$. Thus the order of limits defined by Eqs. (3.5) and (3.6) can only work for

$$-g(\tau+1) < f < g(\tau+1). \tag{3.7}$$

In particular, when $|f/g| < 1$, one can first peratize the stronger of the two interactions and then expand in terms of the weaker one. The domain in which Eq. (3.7) is valid is the region 1 in Fig. 1.

B. f Fixed, Series in g

We can next ask for a representation

$$A = \sum_{n=0} c_n'(f) g^{2n}. \tag{3.8}$$

By the same reasoning as before we again inspect Eq. (2.11). It is at once evident that, *independently of the ratio* f/g, there is no value of g for which (3.8) converges. For A has a branch point at $g = 0$ due to the factor $g^{1/\tau}$, and this essential singularity cannot be affected by the Γ functions which have poles only as singular points.

However, we can ask if Eq. (3.8) could make sense as an asymptotic series rather than as a convergent one. If so, we would, nevertheless, have a decent approximation at least for small g. In connection with the discussion of Eq. (3.4), we know that this is at best only possible for *positive* f.

Thus we start with Eq. (2.11) and use the asymptotic expansion [B, p. 47, Eq. (2)]

$$\Gamma(z) = e^{-z+(z-\frac{1}{2})\ln z}(2\pi)^{\frac{1}{2}} \times \left[1 + \frac{1}{12z} + \frac{1}{288z^2} + \cdots\right], \qquad (3.9)$$

from which we get

$$A = -\left(\frac{f}{\tau^2}\right)^{1/\tau} \frac{\Gamma\left(1 - \frac{1}{\tau}\right)}{\Gamma\left(1 + \frac{1}{\tau}\right)} \times \left[1 + \frac{g^2}{6\tau f^2}(\tau^2 - 1) + O\left(\frac{g^4}{f^4}\right)\right]. \qquad (3.10)$$

From the known results for a single repulsive singular potential [see Eq. (2.12)] we know of course that the leading term of Eq. (3.10) is exactly correct at $g = 0$. The "region" where asymptotic series expansions in a coupling constant apply is marked 2 in Fig. 1; it is the positive f axis.

There remains the region 3 of Fig. 1, where no power series expansion in a single coupling constant is valid.

Scattering of Surface Waves on an Infinitely Deep Fluid*

J. E. BURKE
Sylvania Electronic Defense Laboratories, Mountain View, California

The reflection and transmission of small amplitude waves incident on a plane barrier submerged in an infinitely deep fluid are investigated; the barrier has a finite width and is parallel to the undisturbed free surface of the fluid. Green's function techniques are used to represent the velocity potential in terms of its discontinuity (pressure difference) across the barrier. The application of the boundary condition at the barrier leads to a one-dimensional integral equation for the pressure difference. This equation is extended by introducing the fluid velocity normal to the plane of the barrier, and then it is analyzed by the complex Fourier transform methods of Wiener and Hopf. The velocity transforms are found to be characterized by a pair of dual inhomogeneous integral equations which allow systematic approximation by iteration. A convolution provides representations for the velocity potential in the different regions of the fluid, and these are the basis for investigating the reflection and transmission of waves. Results for some related problems (e.g., scattering by a semi-infinite barrier, by a finite dock, etc.) are obtained as limiting cases.

1. INTRODUCTION

VARIOUS techniques have been used to solve problems in which time-harmonic surface waves of small amplitude[1-3] are incident on a fixed body located in a fluid. In particular, an integral-equation method has proved successful when the obstacle has the form of a semi-infinite plane located at, or below, the surface of a fluid of finite or infinite depth.[4-6] Thus, Greene and Heins[6] investigate the reflection and transmission of waves incident on a semi-infinite plane barrier placed parallel to the free surface of an infinitely deep fluid. The propagation normal of the wave motion is assumed oblique to the coordinate axis (z axis) parallel to the edge of the barrier, allowing the factor $e^{ikz}(0 < k < 1)$ to be separated from the potential function. The two-dimensional residual function satisfies a modified potential equation.

Green's function techniques are used to represent the residual function in terms of its derivative normal to the plane of the barrier. The regions above and below this plane are considered separately, and, by matching the representations at the common fluid boundary of their domains, a one-dimensional integral equation for the derivative is obtained. The discontinuity of the potential across the barrier is introduced to extend the equation preliminary to a Fourier transform analysis by the methods of Wiener and Hopf.[5,7-9] A complete separation of the transforms pertaining to the normal velocity and the discontinuity is obtained allowing them to be determined explicitly.

The present paper considers the analogous scattering problem when the barrier has a finite width. A Green's function representation of the velocity potential is obtained in terms of its discontinuity (pressure difference) across the barrier. The requirement that the normal velocity vanish at the barrier leads to a one-dimensional integral equation for the pressure difference. This equation is extended by introducing the normal fluid velocity in the plane of, but off, the barrier, and then analyzed via complex Fourier transform methods.

Here the transform equation differs markedly from that for the semi-infinite barrier in that two normal velocity transforms appear which refer to the regions left and right of the barrier. A complicating feature is that it is not possible to obtain a complete separation of the transforms. However, the velocity transforms are found to be characterized by a pair of dual inhomogeneous integral equations which allow systematic approximation by iteration. In a study of the acoustical diffraction by an infinite slit, Levine[10] obtained a similar characterization of

* This work was partially supported by the Office of Naval Research Contract Nonr-225 (11) (NR-041-086).

[1] H. Lamb, *Hydrodynamics* (Dover Publications, Inc., New York, 1945), Chap. 9.

[2] L. M. Milne-Thompson, *Theoretical Hydrodynamics* (MacMillan and Company, Ltd., London, 1938), Chap. 14.

[3] J. J. Stoker, *Water Waves* (Interscience Publishers, Inc., New York, 1957), pp. 19–22.

[4] A. E. Heins, Can. J. Math. **2**, 210 (1950).

[5] A. E. Heins, Commun. Pure Appl. Math. **9**, 447 (1956).

[6] T. R. Greene and A. E. Heins, Quart. Appl. Math. **11**, 201 (1953).

[7] N. Wiener and E. Hopf, Sitz. Berlin Akad. Wiss., **30/32**, 696 (1931).

[8] R. E. A. C. Paley and N. Wiener, Am. Math. Soc. Coll. Publ. N. Y., **19**, 1934, Chap. 4.

[9] E. C. Titchmarsh, *Introduction to the Theory of Fourier Integrals* (Oxford University Press, Oxford, 1937).

[10] H. Levine, "Diffraction by an Infinite Slit," TN 61, Applied Mathematics and Statistics Laboratory, Stanford University (1957).

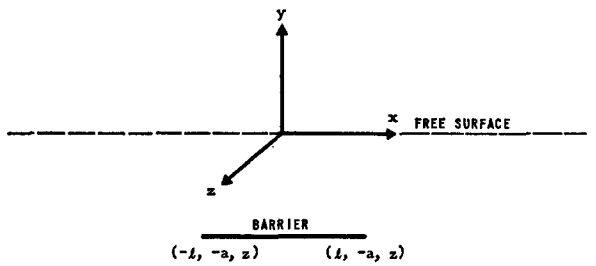

FIG. 1. Geometry for the scattering of waves by a submerged-plane barrier. The width of the barrier is $\delta = 2l$.

the screen distributions, where the inhomogeneous terms feature the distributions corresponding to isolated screens. The present inhomogeneous terms have an analogous interpretation for a semi-infinite barrier, but are complicated by their dependence on unknown wave amplitudes. Thus, in the acoustical problem the transforms may be approximated directly from the integral equations, whereas here consistent approximations for the wave amplitudes are needed.

Representations of the velocity potential are obtained with emphasis on the resulting relations between the wave amplitudes. The amplitudes of the waves above the barrier are proportional to the left velocity transform when its argument has the positive or negative (numerical) value of the propagation constant of the corresponding motion. This result, together with an iterative approximation for the dual integral equations, provides two linear relations between these amplitudes and the reflected wave amplitude. The requirement that no wave be incident from the right results in the third equation needed to specify these three amplitudes. The transmitted amplitude is determined from a general expression which exhibits its dependence on the wave motion above the barrier and the motion below.

The first approximation to the wave amplitudes is found to be appropriate when motion below the barrier is neglected. That is, only the multiple reflections of the waves above the barrier are considered with each edge having the characteristics of an edge of a semi-infinite barrier. The higher approximations, which take into account the motion below, involve integrals that are not readily evaluated. However, approximate procedures are illustrated which are appropriate when the barrier width is large compared with the wavelength $2\pi/k$.

2. FUNDAMENTAL EQUATIONS OF MOTION AND STATEMENT OF THE PROBLEM

This section contains a brief sketch of the procedures which lead to the small amplitude wave theory, and a statement of the problem (based on this theory) considered in subsequent sections.

A nonviscous fluid in irrotational motion is considered and consequently the motion can be specified by a velocity potential, $\Phi(x, y, z; t)$, which is a solution of the differential equation

$$\Phi_{xx} + \Phi_{yy} + \Phi_{zz} = 0 \tag{1}$$

satisfying prescribed conditions on boundary surfaces. The condition at a fixed (time independent) surface is the vanishing of the normal fluid velocity ($\partial\Phi/\partial n = 0$). At a free (constant pressure) surface two boundary conditions must be satisfied.[3] One of them, which expresses the constancy of pressure, arises from Bernoulli's law and has the form

$$gf + \Phi_t + \tfrac{1}{2}(\Phi_x^2 + \Phi_y^2 + \Phi_z^2) = 0, \tag{2}$$

where g denotes acceleration due to gravity and where $y = f(x, z; t)$ defines the free surface at which the pressure has been assumed to be zero. The other, a consequence of the fact that once a fluid particle is on the surface it remains there, is expressed by the equation

$$\Phi_x f_x + \Phi_z f_z + f_t = \Phi_y. \tag{3}$$

The theory of small amplitude waves[1-3] is obtained by dropping the nonlinear terms in (2) and (3), i.e., the velocity and elevation of the free surface, and their derivatives, are assumed small enough to allow the indicated linearization. Elimination of $f(x, z; t)$ between the resulting equations gives the single condition

$$\Phi_{tt} + g\Phi_y = 0, \tag{4}$$

and, subject to an error of the order already neglected, this condition may be imposed at the undisturbed position ($y = 0$) of the free surface. With this approximation the problem is reduced to a classical boundary value problem of potential theory for a fixed geometry.

Consider now a fluid of semi-infinite extent, with x and z axes in the plane of the undisturbed surface, and the y axis directed vertically upward, in which there is located a rigid barrier described by $y = -a$, $-l \leq x \leq l$, for any z (see Fig. 1). A steady state situation is contemplated wherein a train of surface waves is incident on the barrier from one side. The fluid is taken to be incompressible, nonviscuous, and in irrotational, time-harmonic motion, and thus there exists a velocity potential of the form

$$\Phi(x, y, z; t) = \Phi_0(x, y, z)e^{-i\omega t}, \qquad (\partial_x^2 + \partial_y^2 + \partial_z^2)\Phi_0 = 0. \tag{5}$$

The additional asumption that the propagation normal of the wave motion is oblique to the edge of the barrier is made, and consequently

$$\Phi_0(x, y, z) = \varphi(x, y)e^{ikz}, \tag{6}$$

where the residual function, $\varphi(x, y)$, satisfies

$$\varphi_{xx} + \varphi_{yy} - k^2\varphi = 0. \tag{7}$$

Restricting the motion to waves of infinitesimal amplitude, the free surface condition (4), stated in terms of $\varphi(x, y)$, becomes

$$\varphi_y - \beta\varphi = 0, \qquad y = 0, \qquad \beta = \omega^2/g. \tag{8}$$

At the rigid barrier, we have

$$\partial\varphi/\partial y = 0, \qquad y = -a, \qquad -l < x < l. \tag{9}$$

In the absence of the barrier, Eqs. (7) and (8) admit the surface wave solution

$$\varphi(x, y) = e^{\beta y \pm i\kappa x}, \qquad y < 0, \quad \kappa^2 = \beta^2 - k^2, \tag{10}$$

and when the fluid has a finite depth, bounded below by the rigid surface $y = -a$, the analogous wave solution for Eqs. (7), (8), and (9) is

$$\varphi(x, y) = e^{\pm i\kappa_0 x}\cosh\left[\frac{\rho_0}{a}(y + a)\right], \qquad -a < y < 0,$$
$$\kappa_0^2 = (\rho_0/a)^2 - k^2, \tag{11}$$

where

$$\rho \sinh \rho = \beta a \cosh \rho \tag{12}$$

has the smallest real roots $\pm\rho_0$. Equation (12) also has an infinite sequence of imaginary roots which correspond to exponentially damped motion. These solutions are needed to characterize motions when discontinuities are present in the fluid. The two surface-wave solutions provide the basic wave motions relevant for describing the fluid motion in the regions left and right, and above the barrier. Thus we seek a solution of (7), (8), and (9) satisfying

$$\varphi(x, y) \sim e^{\beta y}[\alpha_1 e^{i\kappa(x+l)} + \alpha_2 e^{-i\kappa(x+l)}], \quad x \to -\infty,$$
$$\varphi(x, y) \sim e^{\beta y}[\alpha_3 e^{i\kappa(x-l)} + \alpha_4 e^{-i\kappa(x-l)}], \quad x \to +\infty, \tag{13}$$

and $\varphi \to 0$ at infinite depths ($y \to -\infty$).

3. DERIVATION OF THE INTEGRAL EQUATION

In order to obtain a representation of $\varphi(x, y)$ we introduce a Green's function $G(x, y; x', y')$, associated with Eq. (7), satisfying

$$(\partial_x^2 + \partial_y^2 - k^2)G = -\delta(x - x')\delta(y - y'), \; y < 0, \qquad Gy - \beta G\,|_{y=0} = 0, \tag{14}$$

where $\delta(x - x')$ denotes the Dirac delta function. The Green's function can be constructed with the aid of the solution, G^*, of the auxiliary problem specified by

$$(\partial_x^2 + \partial_y^2 - k^2)G^* = -\delta(x - x')\delta(y - y'), \quad y, y' < 0, \qquad G_y^* + bG^* = 0, \quad b > 0, \quad y = 0. \tag{15}$$

Using the method of images, i.e., assuming sources in the image space ($y > 0$) of the fluid region, G^* may be written in the form

$$G^*(x, y; x', y') = \frac{i}{4}[H_0^{(1)}(ikR) + H_0^{(1)}(ikR')] - \frac{ib}{2}\int_{-y'}^{\infty} e^{-b(y''+y')}H_0^{(1)}(ikR'')\,dy'',$$
$$R^2 = (x - x')^2 + (y - y')^2, \qquad R'^2 = (x - x')^2 + (y + y')^2, \qquad R''^2 = (x - x')^2 + (y - y'')^2, \tag{16}$$

where $H_0^{(1)}$ is the Hankel function of the first kind. Substituting the integral representation

$$H_0^{(1)}(ikR'') = \frac{1}{\pi i}\int_{-\infty}^{\infty} \exp\,[-k(y'' - y)\cosh t + ik(x - x')\sinh t]\,dt, \qquad y'' > y, \tag{17}$$

into (16) and carrying out the y'' integration results in

$$G^*(x, y; x', y') = \frac{i}{4}[H_0^{(1)}(ikR) - H_0^{(1)}(ikR')] + \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{k\cosh t \exp\,[k(y + y')\cosh t + ik(x - x')\sinh t]}{b + k\cosh t}\,dt. \tag{18}$$

Finally, making the change of variable $\zeta = k \sinh t$ in (18), we obtain

$$G^* = \frac{i}{4}[H_0^{(1)}(ikR) - H_0^{(1)}(ikR')] + \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp[i\zeta(x - x') + (\zeta^2 + k^2)^{\frac{1}{2}}(y + y')] \frac{d\zeta}{(\zeta^2 + k^2)^{\frac{1}{2}} + b}. \quad (19)$$

The integrand in (19) has no singularity when $b > 0$, which is the condition underlying the original expression (16) for G^*. On the other hand, for $b = -\beta < 0$, the integrand has simple poles at $\zeta = \pm\kappa = \pm(\beta^2 - k^2)^{\frac{1}{2}}$. Consequently, the desired Green's function,

$$G(x, y; x', y') = \frac{i}{4}[H_0^{(1)}(ikR) - H_0^{(1)}(ikR')] + \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp[i\zeta(x - x') + (\zeta^2 + k^2)^{\frac{1}{2}}(y + y')] \frac{d\zeta}{(\zeta^2 + k^2)^{\frac{1}{2}} - \beta}, \quad (20)$$

is dependent on the path of integration chosen with respect to these poles, the various choices affecting only the asymptotic form of the function for $|x - x'| \to \infty$. We could, for example, require that the Green's function describe outward-going waves for $|x| \to \infty$. Such a Green's function is asymptotic to $(i\beta/2\kappa) \exp[\beta(y + y') + i\kappa|x - x'|]$, and would be obtained by deforming the path below $+\kappa$ and above $-\kappa$. This function, however, is not suitable for taking the bilateral Fourier transform with respect to the variable x'. To insure a favorable behavior we deform the path below both singularities giving a Green's function with asymptotic forms

$$G \sim O(e^{-k|x-x'|}), \quad x' - x \to \infty, \qquad G \sim \frac{-\beta}{\kappa} e^{\beta(y+y')} \sin\kappa(x - x'), \quad x' - x \to -\infty. \quad (21)$$

Applying Green's theorem to φ and G leads to the representation

$$\varphi(x, y) = \int_{C_1+C_2} (G\,\partial n\varphi - \varphi\,\partial nG)\,ds \equiv \{G, \varphi\}_{C_1+C_2}; \quad (22)$$

here C_1 is the boundary of the rectangle formed by the four lines $y = 0$, $y = -M$, $x = L_2$, and $x = -L_1$, and the path C_2 is composed of the upper and lower faces of the barrier. In view of the boundary conditions satisfied by G and φ at the free surface ($y = 0$), this relation reduces to

$$\varphi(x, y) = \int_{-l}^{l} G_{y'}(x, y; x', -a)[\varphi(x', -a + 0) - \varphi(x', -a - 0)]\,dx' + \{G, \varphi\}_{x'=-L_1} + \{G, \varphi\}_{y'=-M} + \{G, \varphi\}_{x'=L_2}. \quad (23)$$

The order conditions on G and φ at infinity imply that the last two integrals tend to zero as the contours recede to infinity ($L_2, M \to \infty$), while the second integral gives

$$\{G, \varphi\}_{x=-L_1} = e^{\beta y}[\alpha_1 e^{i\kappa(x+l)} + \alpha_2 e^{-i\kappa(x+l)}] + O(e^{-\beta M}). \quad (24)$$

Hence, when $M \to \infty$, we obtain the representation

$$\varphi(x, y) = e^{\beta y}[\alpha_1 e^{i\kappa(x+l)} + \alpha_2 e^{-i\kappa(x+l)}] + \int_{-l}^{l} I(x')G_{y'}(x, y; x', -a)\,dx', \qquad I(x') = \varphi(x', -a + 0) - \varphi(x', -a - 0), \quad (25)$$

at any point in the fluid. Here $I(x')$ characterizes the pressure difference across the barrier. The requirement that the normal velocity vanishes at the barrier results in

$$\int_{-l}^{l} I(x')G_{yy'}(x, -a; x', -a)\,dx' + \beta e^{-\beta a}[\alpha_1 e^{i\kappa(x+l)} + \alpha_2 e^{-i\kappa(x+l)}] = 0, \qquad -l < x < l, \quad (26)$$

a one-dimensional integral equation for the discontinuity.

Considerable information can be obtained from the integral equation (26) without actually solving it. In particular, the edge behavior of $I(x)$ can be determined, and this result in conjunction with (25) suffices to specify the singularities of $\partial\varphi/\partial y\,|_{y=-a}$ at the edges of the barrier.

Equation (26) is basically of the form

$$\int_{-l}^{l} I(x')G_{yy'}(x, -a; x', -a)\,dx' = A\cos\kappa x + B\sin\kappa x, \qquad -l < x < l. \quad (27)$$

When x and x' are nearly coincident, the logarithmic singularity characteristic of the Green's function is dominant, i.e.,

$$G_{yy'}(x, -a; x', -a) \propto \frac{\partial^2}{\partial x^2} \ln|x - x'|, \quad x \approx x'. \quad (28)$$

Thus, the significant contribution to the integral in (27) arises from the neighborhood of x, and, consequently, removing the x derivatives from under

the integral sign, we approximate (27) by

$$\int_{-l}^{l} I(t) \ln |t - x| \, dt = A' \cos \kappa x + B' \sin \kappa x + Cx + D, \quad (29)$$

where C and D are arbitrary constants of integration. Making the change of variables $t' = l \cos \theta'$, $x = l \cos \theta$, in (10) leads to

$$\begin{aligned} F(l \cos \theta) &= A' \cos (\kappa l \cos \theta) + B' \sin (\kappa l \cos \theta) \\ &\quad + Cl \cos \theta + D \\ &= l \int_0^{\pi} I(l \cos \theta') \ln |\cos \theta' - \cos \theta| \\ &\quad \times \sin \theta' \, d\theta' + E \\ E &= l \ln l \int_0^{\pi} I(l \cos \theta') \sin \theta' \, d\theta'. \end{aligned} \quad (30)$$

Upon differentiating (30) with respect to θ we observe that

$$-F'(l \cos \theta) = \int_0^{\pi} \frac{\sin \theta' I(l \cos \theta) \, d\theta}{\cos \theta' - \cos \theta}. \quad (31)$$

Imposing the requirement $\int_0^{\pi} F'(l \cos \theta) \, d\theta = 0$, the arbitrary constant C is fixed and is found to be

$$C = -\frac{B'\kappa}{\pi} \int_0^{\pi} \cos (\kappa l \cos \theta') \, d\theta'. \quad (32)$$

Now (31) and (32) together imply[11]

$$I(l \cos \theta') = -\int_0^{\pi} \frac{\sin \theta' F'(l \cos \theta) \, d\theta}{\cos \theta' - \cos \theta}, \quad (33)$$

which, when the integrand is expanded in a Fourier series, becomes

$$I(l \cos \theta') = 2 \int_0^{\pi} F'(l \cos \theta) \sum_{n=1}^{\infty} \sin n\theta' \cos n\theta \, d\theta, \quad (34)$$

or more compactly,

$$I(l \cos \theta') = l \sum_{n=1}^{\infty} C_n \frac{\sin n\theta'}{n}. \quad (35)$$

Thus, we see that $I(l \cos \theta') = O(l \sin \theta')$ as $\theta' \to \pi, 0$, i.e.,

$$I(x') = O(l^2 - x'^2)^{\frac{1}{2}}, \qquad x' \to \pm l, \quad (36)$$

showing the discontinuity vanishes at the edges in accordance with the square root of the distance therefrom.

To determine any singular behavior of $\partial\varphi/\partial y \,|_{y=-a}$ we need only consider the contribution from the integral in (25), i.e.,

$$F = \int_{-l}^{l} I(x') G_{yy'}(x, -a; x', -a) \, dx'. \quad (37)$$

We now suppose that $x \approx -l$, $(x < -l)$, so that the significant range of integration is in the vicinity of the lower limit. Making the small argument approximation for the Green's function in (37), and using (36), we obtain

$$F = O\left(\int_{-l}^{l} \frac{(x' + l)^{\frac{1}{2}}}{(x - x')^2} \, dx'\right) = O\left(\frac{1}{(x + l)^{\frac{1}{2}}}\right), \quad x \to -l. \quad (38)$$

Thus, in the neighborhood of an edge, the derivative $\partial\varphi/\partial y \,|_{y=-a}$ varies inversely as the square root of the distance measured from the edge.

4. FOURIER TRANSFORM ANALYSIS

We first extend the integral equation (26) to the domain $-\infty < x < \infty$, and write

$$\psi(x) = \psi_0(x) + \int_{-l}^{l} I(x') G_{yy'}(x, -a; x', -a) \, dx', \quad (39)$$

where

$$\begin{aligned} \psi &= 0, && |x| < l, \\ &= \partial\varphi/\partial y|_{y=-a}, && |x| > l, \\ I(x) &= 0, && |x| > l, \\ \psi_0(x) &= \beta e^{-\beta a}[\alpha_1 e^{i\kappa(x+l)} + \alpha_2 e^{-i\kappa(x+l)}], \\ & && -\infty < x < \infty. \end{aligned} \quad (40)$$

Introducing $\Psi(\omega)$, $\Psi_0(\omega)$, $J(\omega)$, and $G(\omega)$ as the Fourier transforms of $\psi(x)$, $\psi_0(x)$, $I(x)$, and the Green's function, respectively, the transform version of (39) is

$$\Psi(\omega) = \Psi_0(\omega) + J(\omega)G(\omega), \quad (41)$$

where

$$\begin{aligned} \Psi(\omega) &= \int_{-\infty}^{-l} \psi(x) e^{-i\omega x} \, dx + \int_{l}^{\infty} \psi(x) e^{-i\omega x} \, dx \\ &\equiv \psi_1(\omega) + \psi_2(\omega), \\ \Psi_0(\omega) &= -i\beta e^{-\beta a}\left[\alpha_1 e^{i\kappa l}\left(\frac{1}{\omega - \sigma_+} - \frac{1}{\omega - \sigma_-}\right)\right. \\ &\quad \left. + \alpha_2 e^{-i\kappa l}\left(\frac{1}{\omega + \sigma_-} - \frac{1}{\omega + \sigma_+}\right)\right], \quad \sigma_\pm = \kappa \pm i\epsilon, \\ J(\omega) &= \int_{-l}^{l} I(x) e^{-i\omega x} \, dx, \end{aligned} \quad (42)$$

[11] W. Schmeidler, *Integralgleichungen mit Anwendung in Physik und Technik* (Geest and Portig, Leipzig, 1950), example 26.

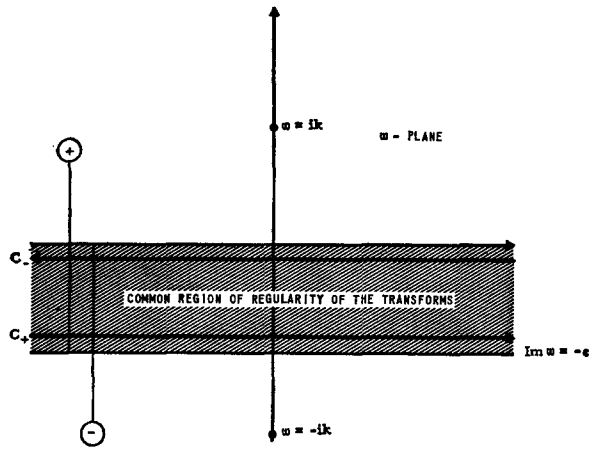

FIG. 2. Complex w-plane showing the region of regularity common to all the transforms in Eq. (41).

$$G(\omega) = \frac{pe^{-pa}}{p-\beta}(\beta \cosh pa - p \sinh pa), \qquad p = (\omega^2 + k^2)^{\frac{1}{2}}.$$

A convergence factor $(e^{-\epsilon|x|})$ has been introduced in $\psi_0(x)$ to facilitate its transformation; the parameter ϵ will subsequently be put equal to zero. The transform $G(\omega)$ has been obtained by noting that, with the aid of the integral representation (17), Eq. (20) reduces to the form

$$G(x, y; x', y') = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega(x-x')} g(\omega, y, y')\, d\omega \qquad (43)$$

so that $G(\omega) = g_{yy'}(\omega, y, y')\,|_{y=y'=-a}$.

A meaning can be attributed to the transform relation (41) if there exists a region of regularity common to all the transforms. Since $J(\omega)$ is the transform of a function defined in a finite range, it exists everywhere in the finite ω-plane, and must become infinite as $|\omega| \to \infty$, otherwise it would be constant. To obtain the asymptotic forms of $J(\omega)$ we note that as $|\omega| \to \infty$ in the lower half plane the major contribution to the integral for $J(\omega)$ comes from the neighborhood of the lower limit. Thus, in view of the behavior of $I(x)$ for $x \sim -l$ [see Eq. (36)] we have

$$J(\omega) \propto \int_{-l}^{l} (x+l)^{\frac{1}{2}} e^{-i\omega x}\, dx = e^{i\omega l}\int_0^{2l} t^{\frac{1}{2}} e^{-i\omega t}\, dt = \frac{e^{i\omega l}}{2i\omega}\int_0^{2l} t^{-\frac{1}{2}} e^{-i\omega t}\, dt + O\left(\frac{e^{-i\omega l}}{\omega}\right). \qquad (44)$$

When $|\omega| \gg 1$ and $\operatorname{Im}\omega < 0$ we may approximate the final integral by replacing the upper limit by $+\infty$ and obtain

$$J(\omega) \sim O\left(\frac{e^{i\omega l}}{\omega^{\frac{1}{2}}}\right), \quad |\omega| \to \infty, \quad \operatorname{Im}\omega < 0, \qquad (45)$$

since the integrated term vanishes in the limit. A similar analysis shows that

$$J(\omega) \sim O\left(\frac{e^{-i\omega l}}{\omega^{\frac{1}{2}}}\right), \quad |\omega| \to \infty, \quad \operatorname{Im}\omega > 0. \qquad (46)$$

When $|x| \to \infty$ the function $\psi(x)$ is $O(e^{\pm ikx - \epsilon|x|})$ in view of the attenuation factor, and consequently $\psi_1(\omega)$ exists for $\operatorname{Im}\omega > -\epsilon$ and $\psi_2(\omega)$ exists for $\operatorname{Im}\omega < \epsilon$. Since the singularities of $\psi(x)$ at $x = \pm l$ are of the form $(x \mp l)^{-\frac{1}{2}}$ considerations analogous to those used for $J(\omega)$ show that

$$\psi_1(\omega) \sim O\left(\frac{e^{i\omega l}}{\omega^{\frac{1}{2}}}\right), \quad |\omega| \to \infty, \quad \operatorname{Im}\omega > -\epsilon, \qquad \psi_2(\omega) \sim O\left(\frac{e^{-i\omega l}}{\omega^{\frac{1}{2}}}\right), \quad |\omega| \to \infty, \quad \operatorname{Im}\omega < \epsilon. \qquad (47)$$

Finally $\Psi_0(\omega)$ is regular in $|\operatorname{Im}\omega| < \epsilon$ while $G(\omega)$ is regular for $-k < \operatorname{Im}\omega < 0$. Thus, there exists a common strip of regularity for the transforms of width ϵ as shown in Fig. 2.

The exponential factors of the transforms $\psi_{1,2}(\omega)$ for large $|\omega|$ are made explicit by writing

$$\psi_1(\omega) = P_1(\omega)e^{i\omega l}, \qquad \psi_2(\omega) = P_2(\omega)e^{-i\omega l}, \qquad (48)$$

where $P_1(\omega)$ and $P_2(\omega)$ have algebraic behavior at infinity in the appropriate half planes. Eq. (41) may thus be written as

$$J(\omega)G(\omega) = P_1(\omega)e^{i\omega l} + P_2(\omega)e^{-i\omega l} - \Psi_0(\omega). \qquad (49)$$

The task now is to find a rearrangement of (49) so that the left and right sides are regular in half planes which overlap, thus defining an entire function in the ω-plane. If $G(\omega)$ can be represented by a ratio of functions $G_+(\omega)$, $G_-(\omega)$, viz.

$$G(\omega) = G_+(\omega)/G_-(\omega), \qquad (50)$$

where $G_+(\omega)$ is regular and nonzero in $\operatorname{Im}\omega > -k$ and $G_-(\omega)$ is regular and nonzero in $\operatorname{Im}\omega < 0$, then multiplication of (49) by the appropriate exponential factor results in the two equations

$$\begin{aligned} J(\omega)e^{-i\omega l}/G_-(\omega) &= \underset{(-)}{P_1(\omega)/G_+(\omega)} \\ &\quad + P_2(\omega)e^{-i\omega 2l}/G_+(\omega) - \underset{(-)}{\Psi_0(\omega)e^{-i\omega l}/G_+(\omega)} \\ J(\omega)e^{i\omega l}G_+(\omega) &= \underset{(+)}{P_1(\omega)e^{i\omega 2l}G_-(\omega)} \\ &\quad + P_2(\omega)G_-(\omega) - \underset{(-)}{\Psi_0(\omega)e^{i\omega l}G_-(\omega)}. \end{aligned} \qquad (51)$$

The underscoring denotes functions regular in the indicated half plane while the remaining are not. However, with a knowledge of $G_\pm(\omega)$ and their

behavior for $|\omega| \to \infty$, the latter can be represented by a sum of terms regular in half planes.

From (42) and (50) we have $G_+(\omega)$ and $G_-(\omega)$ defined in $-k < \operatorname{Im}\omega < 0$ by the relation

$$G_+(\omega)/G_-(\omega) = \frac{pe^{-pa}}{p-\beta}(\beta\cosh pa - p\sinh pa). \tag{52}$$

The calculation of $G_+(\omega)$ is conveniently accomplished by considering each factor on the right-hand side separately, i.e., the factors

$$e^{-pa} = e^{M_+(\omega)}/e^{M_-(\omega)}, \tag{53}$$

$$1 - \beta/p = L_-(\omega)/L_+(\omega), \tag{54}$$

and

$$\beta\cosh pa - p\sinh pa = N_+(\omega)/N_-(\omega), \tag{55}$$

where the subscripts $(\pm)$ denote the half-plane of regularity of the functions. These decompositions have been noted elsewhere[6] and the final results are merely stated in the following.

The required form (53) is obtained by using

$$\begin{gathered} M_\pm(\omega) = \frac{a}{\pi i}(\omega^2 + k^2) f(\mp\omega) \\ f(\omega) = (\omega^2+k^2)^{-\frac{1}{2}} \ln\left(\frac{1+\tau}{1-\tau}\right), \quad \tau = \left(\frac{\omega+ik}{\omega-ik}\right)^{\frac{1}{2}} \\ -\frac{3\pi}{2} < \arg(\omega - ik) < \frac{\pi}{2}, \\ -\frac{\pi}{2} < \arg(\omega + ik) < \frac{3\pi}{2}. \end{gathered} \tag{56}$$

Letting $|\omega| \to \infty$ in the appropriate half plane, these expressions lead to the asymptotic forms

$$\begin{gathered} \exp M_\pm \sim \exp\left(\mp\frac{a\omega}{2} \mp \frac{a\omega}{\pi i}\ln\frac{2}{k} \mp \frac{a\omega}{\pi i}\ln\omega\right), \\ |\omega| \to \infty, \qquad \operatorname{Im}\omega \gtrless \mp k. \end{gathered} \tag{57}$$

The quantities $L_\pm(\omega)$ are specified, except for a constant of integration, by the two differential equations

$$\begin{aligned} \frac{L_-'(\omega)}{L_-(\omega)} &= -\frac{1}{2(\omega - ik)} + \frac{\omega}{\omega^2 - \kappa^2} \\ &\quad - \frac{1}{\beta\pi i}\frac{d}{d\omega}[(\omega^2+k^2)f(\omega)] + \frac{(\omega^2+k^2)f(\omega)}{2\beta\pi i(\omega+\kappa)} \\ &\quad + \frac{(\omega^2+k^2)f(\omega)}{2\pi i(\omega-\kappa)} + \frac{\beta f(+\kappa)}{2\pi i(\omega+\kappa)} + \frac{\beta f(-\kappa)}{2\pi i(\omega-\kappa)} \\ \frac{L_+'(\omega)}{L_+(\omega)} &= \frac{1}{2(\omega+ik)} + \frac{1}{\beta\pi i}\frac{d}{d\omega}[(\omega^2+k^2)f(-\omega)] \\ &\quad - \frac{1}{2\beta\pi i(\omega+\kappa)}[(\omega^2+k^2)f(-\omega) - \beta^2 f(+\kappa)] \\ &\quad - \frac{1}{2\beta\pi i(\omega-\kappa)}[(\omega^2+k^2)f(-\omega) - \beta^2 f(-\kappa)]. \end{aligned} \tag{58}$$

When $|\omega| \to \infty$ we have

$$\begin{aligned} L_-(\omega) &\sim c_-\omega\left[1 - \frac{2\beta}{\pi i\omega}\ln i\omega\right]e^{-i\omega/\pi\beta}, \\ &\qquad \operatorname{Im}\omega < 0, \\ L_+(\omega) &\sim c_+\omega\left[1 + \frac{2\pi}{\pi i\omega}\ln(-i\omega)\right]e^{-i\omega/\pi\beta}, \\ &\qquad \operatorname{Im}\omega > -k. \end{aligned} \tag{59}$$

The explicit determination of the constants $(c_\pm)$ of integration is not needed for our purposes.

Finally, the decomposition (55) is obtained by using the Weierstrass product representation for the entire function on the left side. Thus

$$\begin{aligned} N_+(\omega) &= \prod_{n=1}^{\infty}\left\{\left[1 + \left(\frac{ka}{\rho_n}\right)^2\right]^{\frac{1}{2}} - \frac{i\omega a}{\rho_n}\right\}e^{i\omega a/n\pi}, \\ \frac{1}{N_-(\omega)} &= \frac{a^2\beta}{\rho_0^2}(\omega^2 - \kappa_0^2) \\ &\quad \times \prod_{n=1}^{\infty}\left\{\left[1 + \left(\frac{ka}{\rho_n}\right)^2\right]^{\frac{1}{2}} + \frac{i\omega a}{\rho_n}\right\}e^{-i\omega a/n\pi}, \end{aligned} \tag{60}$$

where $\pm i\rho_n$, $(n = 1, 2, \cdots)$ are the imaginary roots of the transcendental equation (12). When $|\omega| \to \infty$ in the appropriate half plane, the corresponding asymptotic forms of these expressions are

$$\begin{aligned} N_+(\omega) &\sim \frac{1}{\omega}\exp\left[\frac{i\gamma\omega a}{\pi} - \left(\frac{i\omega a}{\pi} - \frac{1}{2}\right)\right. \\ &\qquad \left.\times \ln\frac{i\omega a}{\pi} + \frac{i\omega a}{\pi} - \frac{1}{2}\ln 2\pi\right], \\ &\qquad \operatorname{Im}\omega > -\rho_1\left[1 + \frac{k^2a^2}{\rho_1^2}\right]^{\frac{1}{2}}, \\ \frac{1}{N_-(\omega)} &\sim \omega\exp\left[-\frac{i\gamma\omega a}{\pi} + \left(\frac{i\omega a}{\pi} + \frac{1}{2}\right)\right. \\ &\quad \left.\times \ln\left(\frac{-i\omega a}{\pi}\right) - \frac{i\omega a}{\pi} - \tfrac{1}{2}\ln 2\pi\right], \quad \operatorname{Im}\omega < 0, \end{aligned} \tag{61}$$

where γ denotes Euler's constant.

The asymptotic forms of the factors (53) to (55) indicate that $G_\pm(\omega)$ will have an exponential behavior for $|\omega| \to \infty$, and, consequently, we write

$$\begin{aligned} &G_+(\omega)/G_-(\omega) \\ &\quad = (e^{M_+}N_+L_+e^{-Q(\omega)})(e^{M_-}L_-N_-e^{-Q(\omega)})^{-1}, \end{aligned} \tag{62}$$

where the factor

$$e^{-Q(\omega)} = \exp\left[-\frac{i\omega a}{\pi}\left(\gamma - 1 + \ln\frac{ka}{2\pi}\right)\right] \tag{63}$$

has been introduced to make $G_\pm(\omega)$ of algebraic order at infinity; thus

$$G_+(\omega) \sim O(\omega^{\frac{1}{2}}), \qquad |\omega| \to \infty, \qquad \operatorname{Im}\omega > -k,$$
$$G_-(\omega) \sim O\left(\frac{1}{\omega^{\frac{1}{2}}}\right), \qquad |\omega| \to \infty, \qquad \operatorname{Im}\omega < 0.$$

With $G_\pm(\omega)$ determined, the decomposition of the remaining terms of (51) can be carried out. Consider, for example, the function $P_2(\omega)e^{-i\omega\delta}/G_+(\omega)$ where $\delta = 2l$ is the width of the barrier. For a point ω within a closed rectangular contour (C) in the common strip of regularity, Cauchy's theorem permits the representation

$$P_2(\omega)e^{-i\omega\delta}/G_+(\omega) = \frac{1}{2\pi i}\int_C \frac{P_2(t)e^{-it\delta}}{G_+(t)(t-\omega)}\,dt. \tag{64}$$

Since the integrand is $O(1/t^2)$ for $|t| \to \infty$ in this strip, the ends of the contour may be allowed to recede to infinity with no contribution to the integral. The path C is thus deformed into two lines $C_\pm$ parallel to the real axis of the ω-plane as shown in Fig. 2. Thus, we obtain

$$\begin{aligned} P_2(\omega)e^{-i\omega\delta}/G_+(\omega) &= \eta_+(\omega) + \eta_-(\omega), \\ \eta_\pm(\omega) &= \frac{1}{2\pi i}\int_{C_\pm} \frac{P_2(t)e^{-it\delta}}{G_+(t)(t-\omega)}\,dt, \end{aligned} \tag{65}$$

where the $\eta_\pm$ are regular in half planes. Similarly,

$$\begin{aligned} P_1(\omega)G_-(\omega)e^{i\omega\delta} &= \xi_+(\omega) + \xi_-(\omega), \\ \Psi_0(\omega)e^{-i\omega l}/G_+(\omega) &= \theta_+(\omega) + \theta_-(\omega), \\ \Psi_0(\omega)e^{i\omega l}G_-(\omega) &= \Omega_+(\omega) + \Omega_-(\omega), \end{aligned} \tag{66}$$

where the subscripts $(\pm)$ denote the half planes of regularity of these functions. In each case, for $|\omega| \to \infty$ in the appropriate half plane the functions on the right vanish.

Incorporating (65) and (66) in Eq. (51) leads to

$$\begin{aligned} &J(\omega)e^{-i\omega l}/G_-(\omega) - \underset{(-)}{\eta_-(\omega)} + \underset{(-)}{\theta_-(\omega)} \\ &\quad = \underset{(+)}{P_1(\omega)/G_+(\omega)} + \underset{(+)}{\eta_+(\omega)} - \underset{(+)}{\theta_+(\omega)} = E_1(\omega) \\ &J(\omega)e^{i\omega l}G_+(\omega) - \underset{(+)}{\xi_+(\omega)} + \underset{(+)}{\Omega_+(\omega)} \\ &\quad = \underset{(-)}{P_2(\omega)G_-(\omega)} + \underset{(-)}{\xi_-(\omega)} - \underset{(-)}{\Omega_-(\omega)} = E_2(\omega), \end{aligned} \tag{67}$$

which, by analytic continuation, define entire functions $E_1(\omega)$ and $E_2(\omega)$ in the ω-plane. The asymptotic behaviors of these expressions for $|\omega| \to \infty$ show that the integral functions are identically zero, and consequently, from the representation of $E_1(\omega)$ in the upper half plane and $E_2(\omega)$ in the lower half plane, we deduce

$$\begin{aligned} P_1(\omega)/G_+(\omega) + \eta_+(\omega) - \theta_+(\omega) &= 0, \\ P_2(\omega)G_-(\omega) + \xi_-(\omega) - \Omega_-(\omega) &= 0, \end{aligned} \tag{68}$$

which, with reference to (65) and (66), constitutes a pair of dual integral equations for the transforms $P_{1,2}(\omega)$. These inhomogeneous equations cannot be solved exactly, but they lend themselves to systematic approximation by iteration.

5. DETERMINATION OF $\varphi(x, y)$

Before discussing the solution of the integral equations (68), it is appropriate to relate the function $\varphi(x, y)$ in the various regions of the fluid to the transforms $P_{1,2}(\omega)$. A complete scheme is thereby presented for determining the fluid motion anywhere.

When the Green's function in Eq. (25) is expressed as a Fourier integral, we obtain, by interchanging orders of integration,

$$\varphi(x, y) = \varphi_0(x, y) + \lim_{\epsilon \to 0}\frac{1}{2\pi}\int_C J(\omega)\mathcal{G}_\pm(\omega)e^{i\omega x}\,d\omega, \tag{69}$$

where

$$\mathcal{G}_\pm(\omega) = \begin{cases} \dfrac{e^{-p\alpha}}{p-\beta}(p\cosh py + \beta\sinh py) & y > -\alpha, \\ \dfrac{e^{py}}{p-\beta}(\beta\cosh p\alpha - p\sinh p\alpha) & y < -\alpha, \end{cases} \tag{70}$$

and where C is a contour in the strip of regularity (parallel to the real axis). In order to evaluate the integral $F(x, y: \epsilon) = 1/2\pi\int_C J(\omega)\mathcal{G}_\pm(\omega)e^{i\omega x}\,d\omega$ we use the representations

$$\begin{aligned} J(\omega)e^{i\omega x} &= [\xi_+(\omega) - \Omega_+(\omega)]/G_+(\omega)e^{i\omega(x-l)}, && x > l, \\ J(\omega)e^{i\omega x} &= \frac{1}{G(\omega)}[P_1(\omega)e^{i\omega(x+l)} \\ &\quad + P_2(\omega)e^{i\omega(x-l)} - \Psi_0(\omega)e^{i\omega x}], && -l < x < l, \\ J(\omega)e^{i\omega x} &= [\eta_-(\omega) - \theta_-(\omega)]G_-(\omega)e^{i\omega(x+l)}, && x < -l, \end{aligned} \tag{71}$$

which are obtained from (67). For $x > l$, and any $y < 0$, the contour may be closed in the upper half plane to give

$$\begin{aligned} F(x, y; \epsilon) &= ie^{i\kappa(x-l)}[\xi_+(\kappa) - \Omega_+(\kappa)]\beta^2 e^{\beta(y-\alpha)}/G_+(\kappa) \\ &\quad - ie^{-i\kappa(x-l)}[\xi_+(-\kappa) - \Omega_+(-\kappa)]\beta^2 e^{\beta(y-\alpha)}/G_+(\kappa) \\ &\quad + \frac{1}{\pi}\int_k^\infty e^{-t(x-l)}\frac{[\xi_+(it) - \Omega_+(it)]}{G_+(it)(\sigma^2+\beta^2)} \end{aligned}$$

$$\times (\beta \cos \sigma a + \sigma \sin \sigma a)(\sigma \cos \sigma y + \beta \sin \sigma y)\, dt, \qquad \sigma^2 = t^2 - k^2; \tag{72}$$

here we have introduced a branch cut from ik to $i\infty$ and noted that $\pm\kappa$ are the only poles of the integrand. To obtain the representation of $\varphi(x, y)$ in this region we require the limiting form of (72) for $\epsilon \to 0$. In view of the definition of $\Psi_0(\omega)$ and the integral relation for the Dirac delta function,

$$\delta(\omega - t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ixt} e^{-i\omega x}\, dx, \tag{73}$$

it follows that

$$\lim_{\epsilon \to 0} \Psi_0(\omega) = \beta e^{-\beta a} 2\pi \times [\alpha_1 e^{+i\kappa l} \delta(\omega - \kappa) + \alpha_2 e^{-i\kappa l} \delta(\omega + \kappa)]. \tag{74}$$

When $\operatorname{Im} \omega > 0$ the integration over the contour C_+ in the definition of $\Omega_+(\omega)$ may be taken along the real axis and we write

$$\Omega_+(\omega) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{\Psi_0(t) G_+(t) e^{itl}}{(t - \omega) G(t)}\, dt. \tag{75}$$

Since $1/G(t)$ has zeros at $t = \pm\kappa$, making use of the limiting behavior of $\Psi_0(\omega)$ when $\epsilon \to 0$, we have

$$\begin{aligned} \lim_{\epsilon \to 0} \Omega_+(\omega) &= 0, && \operatorname{Im} \omega > 0, \\ &= -i\alpha_1 G_+(\kappa) e^{i\kappa\delta} \kappa e^{\beta a}/\beta^2, && \omega = \kappa, \\ &= +i\alpha_2 G_+(-\kappa) e^{-i\kappa\delta} \kappa e^{\beta a}/\beta^2, && \omega = -\kappa. \end{aligned} \tag{76}$$

The same procedures show that

$$\lim_{\epsilon \to 0} \Omega_-(\omega) = 0, \qquad \operatorname{Im} \omega < 0, \tag{77}$$

$$\lim_{\epsilon \to 0} \theta_+(\omega) = i\beta e^{-\beta a} \left\{ \frac{\alpha_1}{G_+(\kappa)(\omega - \kappa)} + \frac{\alpha_2}{G_+(-\kappa)(\omega + \kappa)} \right\}, \qquad \omega \neq \pm\kappa.$$

Turning to $\xi_-(\omega)$ we have from (66)

$$\xi_-(\omega) = \frac{1}{2\pi i} \int_{C_-} \frac{P_1(t) G_-(t) e^{it\delta}}{(t - \omega)}\, dt = \frac{1}{2\pi i} \int_{C_-} \frac{P_1(t) G_+(t) e^{it\delta}}{G(t)(t - \omega)}\, dt. \tag{78}$$

Thus, closing C_- in the upper half plane (again using a branch cut along the positive imaginary axis) we obtain

$$\xi_-(\omega) = -B_0 \left\{ \frac{G_+(\kappa_0) P_1(\kappa_0) e^{i\kappa_0\delta}}{\kappa_0 - \omega} + \frac{G_+(-\kappa_0) P_1(-\kappa_0) e^{-i\kappa_0\delta}}{\kappa_0 + \omega} \right\} - \frac{1}{\pi} \int_k^{\infty} \frac{P_1(it) G_+(it) e^{-t\delta}}{(t + i\omega)(t^2 - k^2)^{\frac{1}{2}}}\, dt, \tag{79}$$

$B_0 = \rho_0 \beta [\kappa_0 \sinh \rho_0 \cosh \rho_0 (\beta^2 a^2 - \beta a - \rho_0^2)]^{-1}$.

Similarly, since

$$\xi_+(\pm\kappa) = \frac{1}{2\pi i} \int_{C_+} \frac{P_1(t) G_-(t) e^{it\delta}}{(t \mp \kappa)}\, dt,$$

comparison with (79) shows that

$$\xi_+(\pm\kappa) = \frac{B_0 P_1(\kappa_0) G_+(\kappa_0) e^{i\kappa_0\delta}}{\kappa_0 \mp \kappa} - \frac{B_0 P_1(-\kappa_0) G_+(-\kappa_0) e^{-i\kappa_0\delta}}{-\kappa_0 \mp \kappa} + \frac{1}{\pi} \int_k^{\infty} \frac{P_1(it) G_+(it) e^{-t\delta}\, dt}{(t \pm i\kappa)(t^2 - k^2)^{\frac{1}{2}}}. \tag{80}$$

Finally, in the definition of $\eta_+(\omega)$ the contour C_+ may be extended in the lower half plane resulting in

$$\eta_+(\omega) = -\frac{1}{\pi} \int_k^{\infty} \frac{P_2(-it) e^{-t\delta}}{(t - i\omega) G_-(-it)(t^2 - k^2)^{\frac{1}{2}}}\, dt. \tag{81}$$

Summarizing, for $x > l$, and all y, we have the representation

$$\varphi(x, y) = ie^{i\kappa(x-l)} \beta^2 e^{\beta(y-a)} \frac{\xi_+(\kappa)}{\kappa G_+(\kappa)} - ie^{-i\kappa(x-l)} \beta^2 e^{\beta(y-a)} \frac{\xi_+(-\kappa)}{\kappa G_+(-\kappa)} + \frac{1}{\pi} \int_k^{\infty} \frac{e^{-t(x-l)} \xi_+(it)}{G_+(it)(\sigma^2 + \beta^2)} (\beta \cos \sigma a + \sigma \sin \sigma a)(\sigma \cos \sigma y + \beta \sin \sigma a)\, dt, \tag{82}$$

where $\xi_+(\omega)$ is given in terms of $P_1(\omega)$ through Eqs. (66) and (80). The transforms $P_{1,2}(\omega)$ are to satisfy (68) when $\epsilon \to 0$, i.e., they are solutions of the pair of inhomogeneous integral equations

$$\begin{aligned} \frac{P_1(\omega)}{G_+(\omega)} &= i\beta e^{-\beta a} \left[\frac{\alpha_1}{G_+(\kappa)(\omega - \kappa)} + \frac{\alpha_2}{G_+(-\kappa)(\omega + \kappa)} \right] + \frac{1}{\pi} \int_k^{\infty} \frac{P_2(-it) e^{-t\delta}}{(t - i\omega) G_-(it)(t^2 - k^2)^{\frac{1}{2}}}\, dt, \\ P_2(\omega) G_-(\omega) &= B_0 \left[\frac{G_+(\kappa_0) P_1(\kappa_0) e^{i\kappa_0\delta}}{\kappa_0 - \omega} + \frac{G_+(-\kappa_0) P_1(-\kappa_0) e^{-i\kappa_0\delta}}{\kappa_0 + \omega} \right] + \frac{1}{\pi} \int_k^{\infty} \frac{P_1(it) G_+(it) e^{-t\delta}}{(t + i\omega)(t^2 - k^2)^{\frac{1}{2}}}\, dt. \end{aligned} \tag{83}$$

Analogous procedures are used to obtain the representation

$$\varphi(x, y) = \varphi_0(x, y) + \frac{1}{\pi}\int_k^\infty e^{t(x+l)}\left[\frac{\eta_-(-it) - \theta_-(-it)}{\sigma^2 + \beta^2}\right]G_-(-it) \times (\beta \cos \sigma a + \sigma \sin \sigma a)(\sigma \cos \sigma y + \beta \sin \sigma y)\, dt, \tag{84}$$

which holds for $x < -l$ and for all y; here

$$\theta_-(\omega) = i\beta e^{-\beta a}\left[\frac{\alpha_1}{G_+(\kappa)(\kappa - \omega)} + \frac{\alpha_2}{G_+(-\kappa)(-\kappa - \omega)}\right],$$

$$\eta_-(\omega) = -\text{P.V.}\,\frac{1}{2\pi i}\int_{-\infty}^\infty \frac{P_2(t)e^{-it\delta}}{G_+(t)(t - \omega)}\, dt. \tag{85}$$

In the region above the barrier ($-l<x<l$, $y>-a$) the representation is found to be

$$\varphi(x, y) = iB_0[P_1(\kappa_0)e^{i\kappa_0(x+l)} - P_1(-\kappa_0)e^{-i\kappa_0(x+l)}] \cosh\left[\frac{\rho_0}{a}(y + a)\right] + \sum_{n=1}^\infty B_n[P_1(i\kappa_n)e^{-\kappa_n(x+l)} + P_2(-i\kappa_n)e^{\kappa_n(x-l)}] \cos\left[\frac{\rho_n}{a}(y + a)\right], \tag{86}$$

where the B_n are defined by (79) ($n = 0$) and by

$$B_n = \rho_n\beta[\kappa_n \cos \kappa_n \sin \kappa_n(\beta^2a^2 - \beta a - \rho_n^2)]^{-1}, \qquad \kappa_n^2 = (\rho_n/a)^2 + k^2, \quad n \geq 1. \tag{87}$$

Finally, below the barrier ($-l < x < l$, $y < -a$) we obtain

$$\varphi(x, y) = \frac{1}{\pi}\int_k^\infty P_1(it) \cos \sigma(y + a)\frac{e^{-t(x+l)}}{\sigma}\, dt + \frac{1}{\pi}\int_k^\infty P_2(-it) \cos \sigma(y + a)\frac{e^{t(x-l)}}{\sigma}\, dt. \tag{88}$$

6. DISCUSSION OF THE REPRESENTATIONS

In the representations for $\varphi(x, y)$ in the various regions of the fluid the integral or infinite sum converges uniformly and absolutely. The favoring exponential when $x \neq \pm l$ allows differentiations to be carried under the integral sign or inside the summation as the case may be. Thus it is possible to verify directly that Eq. (7) and the boundary conditions (8) and (9) are satisfied.

The asymptotic requirements may be verified as follows. Consider, for example, the integral in the representation of $\varphi(x, y)$ for $x < -l$. When $t \to \infty$ we have that $\eta_-(-it) - \theta_-(-it) \sim O(t^{-2})$, and consequently the integrand of (85) is the product of the exponential factor, $e^{t(x+l)}$, and a function which is square integrable in (k, ∞). Applying Schwarz's inequality, we have

$$\left|\int_k^\infty e^{t(x+l)}\left[\frac{\eta_-(-it) - \theta_-(-it)}{\sigma^2 + \beta^2}\right]G_-(-it)(\beta \cos \sigma a + \sigma \sin \sigma a)(\sigma \cos \sigma y + \beta \sin \sigma y)\, dt\right| \leq \left\{\left[\int_k^\infty e^{2t(x+l)}\, dt\right]\left[\int_k^\infty \left|\frac{\eta_- - \theta_-}{\sigma^2 + \beta^2}G_-(-it)(\beta \cos \sigma a + \sigma \sin \sigma a)(\sigma \cos \sigma y + \beta \sin \sigma y)\right|^2 dt\right]\right\}^{\frac{1}{2}} = O\left[\int_k^\infty e^{2t(x+l)}\, dt\right]^{\frac{1}{2}} \to 0 \tag{89}$$

as $x \to -\infty$, independently of y. The analogous result holds for $x \to +\infty$ in the representation for the region $x > l$. The condition that the potential vanish at infinite depths is also fulfilled since the integrals in the representations for the regions below the barrier, $x > l$, and $x < -l$, tend to zero as $y \to -\infty$ by the Riemann–Lebesgue lemma. We now investigate the relations between the various wave amplitude implied by these representations.

In Sec. 2 it was noted that the surface wave solution of (7), (8), and (9), for a fluid of finite depth bounded below by the rigid plane $y = -a$, is of the form

$$\Phi(x, y) = e^{\pm i\kappa_0 x} \cosh\left[\frac{\rho_0}{a}(y + a)\right].$$

If wave motion above the barrier is present it must be of this type, and, indeed, investigation of (86) reveals the wave motion terms

$$W(x, y) = [\gamma_1 e^{i\kappa_0(x-l)} + \gamma_2 e^{-i\kappa_0(x-l)}] \times \cosh\left[\frac{\rho_0}{a}(y + a)\right], \qquad \gamma_1 = iP_1(\kappa_0)B_0e^{i\kappa_0\delta}, \quad \gamma_2 = -iP_1(-\kappa_0)B_0e^{-i\kappa_0\delta}. \tag{90}$$

Omitting the time factor $e^{-i\omega t}$, γ_1 and γ_2 characterize the amplitudes of waves above the barrier traveling to and from the right edge, respectively. When $x \to \infty$ in the representation for the region $x > l$, Eq. (82), $\varphi(x, y)$ becomes (retaining only the nonvanishing terms)

$$\varphi(x, y) = \left[e^{i\kappa(x-l)}\frac{\xi_+(\kappa)}{G_+(\kappa)} - e^{-i\kappa(x-l)}\frac{\xi_+(-\kappa)}{G_+(-\kappa)}\right]\frac{i\beta^2e^{\beta(y-a)}}{\kappa}. \tag{91}$$

Using $\xi_+(\pm\kappa)$ of (80) and γ_i of (90) we rewrite (91) as

$$\varphi(x, y) = [\alpha_3e^{i\kappa(x-l)} + \alpha_4e^{-i\kappa(x-l)}]e^{\beta y},$$

$$\alpha_3 = -\frac{\beta^2 e^{-\beta a}}{\kappa G_+(\kappa)}\left[\frac{G_+(\kappa_0)}{\kappa-\kappa_0}\gamma_1 + \frac{G_+(-\kappa_0)}{\kappa+\kappa_0}\gamma_2 - \frac{i}{\pi}\int_k^\infty \frac{P_1(it)G_+(it)e^{-t\delta}}{(t+i\kappa)(t^2-k^2)^{\frac{1}{2}}}\,dt\right], \quad (92)$$

$$\alpha_4 = -\frac{\beta^2 e^{-\beta a}}{\kappa G_+(-\kappa)}\left[\frac{G_+(\kappa_0)}{\kappa+\kappa_0}\gamma_1 + \frac{G_+(-\kappa_0)}{\kappa-\kappa_0}\gamma_2 + \frac{i}{\pi}\int_k^\infty \frac{P_1(it)G_+(it)e^{-t\delta}}{(t-i\kappa)(t^2-k^2)^{\frac{1}{2}}}\,dt\right].$$

Again omitting the time factor, α_3 is the amplitude of the waves traveling to $x = +\infty$ while α_4 relates to waves coming from $x = +\infty$. We further remark that using γ_i of (90) in the definition of $P_2(\omega)$ leads to

$$P_2(\omega) = \frac{i}{G_-(\omega)}\left[\frac{G_+(\kappa_0)}{\omega-\kappa_0}\gamma_1 + \frac{G_+(-\kappa_0)}{\omega+\kappa_0}\gamma_2\right] + \frac{1}{G_-(\omega)\pi}\int_k^\infty \frac{P_1(it)G_+(it)e^{-t\delta}}{(t+i\omega)(t^2-k^2)^{\frac{1}{2}}}\,dt. \quad (93)$$

The description of the solution for the semi-infinite barrier can be obtained as a special case of our formulas. Consider, for example, a barrier which is semi-infinite to the left. Letting $\delta \to \infty$ in (92) and (93), the limiting forms

$$\alpha_3 = -\frac{\beta^2 e^{-\beta a}}{\kappa G_+(\kappa)}\left[\frac{G_+(\kappa_0)}{\kappa-\kappa_0}\gamma_1 + \frac{G_+(-\kappa_0)}{\kappa+\kappa_0}\gamma_2\right],$$

$$\alpha_4 = -\frac{\beta^2 e^{-\beta a}}{\kappa G_+(-\kappa)}\left[\frac{G_+(\kappa_0)}{\kappa+\kappa_0}\gamma_1 + \frac{G_+(-\kappa_0)}{\kappa-\kappa_0}\gamma_2\right],$$

$$P_2(\omega) = \frac{i}{G_-(\omega)}\left[\frac{G_+(\kappa_0)}{\omega-\kappa_0}\gamma_1 + \frac{G_+(-\kappa_0)}{\omega+\kappa_0}\gamma_2\right] \quad (94)$$

are the same as those derived by Greene and Heins.[6]

The limiting forms in (94) suggest the first approximation for the finite barrier when the width is large compared to the wavelength $2\pi/k$, $(k\delta \gg 1)$. Thus, for this range of the parameters, a first approximation is obtained by neglecting the integrals in (83) and (92). The resulting system of four linear equations, (94) (with $\alpha_4 \equiv 0$) and (90) with the approximation

$$P_1(\omega) = i\beta G_+(\omega) \times \left[\frac{\alpha_1}{G_+(\kappa)(\omega-\kappa)} + \frac{\alpha_2}{G_+(-\kappa)(\omega-\kappa)}\right], \quad (95)$$

can be solved for the amplitudes α_2, α_3, γ_1, γ_2. Thus, we obtain

$$\alpha_2 = \left[R + \frac{Tre^{2i\kappa_0\delta}}{1-e^{2i\kappa_0\delta}r^2}\right]\alpha_1, \quad \alpha_3 = \frac{Tte^{i\kappa_0\delta}}{1-r^2e^{2i\kappa_0\delta}}\alpha_1,$$

$$\gamma_1 = \frac{Te^{i\kappa_0\delta}}{1-r^2e^{2i\kappa_0\delta}}\alpha_1, \quad \gamma_2 = \frac{Tre^{2i\kappa_0\delta}}{1-r^2e^{2i\kappa_0\delta}}\alpha_1, \quad (96)$$

where

$$T = -\frac{\kappa e^{\beta a}(\kappa-\kappa_0)}{\beta^2}\frac{G_+(-\kappa)}{G_+(-\kappa_0)},$$

$$t = -\frac{4\beta^2 e^{-\beta a}G_+(\kappa_0)\kappa_0}{(\kappa+\kappa_0)G_+(\kappa)(\kappa^2-\kappa_0^2)}, \quad (97)$$

$$R = \left(\frac{\kappa-\kappa_0}{\kappa+\kappa_0}\right)\frac{G_+(-\kappa)}{G_+(\kappa)},$$

$$r = -\left(\frac{\kappa-\kappa_0}{\kappa+\kappa_0}\right)\frac{G_+(\kappa_0)}{G_+(-\kappa_0)}.$$

As follows from (94), R and T are the reflection and transmission coefficients for a wave incident, say from the left, on a barrier which is semi-infinite to the right. Similarly, r and t are the reflection and transmission coefficients for a wave incident above the barrier. Thus the approximations (96) take into account multiple reflections of waves above the barrier with each edge having the characteristics of an edge of a semi-infinite barrier. This can be seen directly as follows: A wave with amplitude α_1 incident from the left will give rise to a wave transmitted above the barrier with amplitude $T\alpha_1$. This wave will be reflected at the right edge giving a wave of amplitude $Tre^{i\kappa_0\delta}\alpha_1$ traveling to left. Reflection and transmission take place at the left edge resulting in $Trte^{i2\kappa_0\delta}\alpha_1$ contributed to α_2. This process repeats itself indefinitely, and thus the final amplitude of the reflected wave is

$$\alpha_2 = R\alpha_1 + Trte^{i2\kappa_0\delta}\alpha_1 + Tr^3te^{i4\kappa_0\delta}\alpha_1 + \cdots, \quad (98)$$

which equals α_2 of (96).

The above results suggest an iteration scheme for $P_{1,2}(\omega)$ which leads to the amplitudes (96) as first approximations. Thus, we write

$$P_1^{(n)}(\omega) = F(\omega) + Q_1^{(n)}(\omega)G_+(\omega),$$

$$P_2^{(n)}(\omega) = q^{(n)}(\omega, P_1^{(n)}) + \frac{Q_2^{(n)}(\omega)}{G_-(\omega)}, \quad (99)$$

where

$$F(\omega) = i\beta e^{-\beta a}G_+(\omega)\left[\frac{\alpha_1}{G_+(\kappa)(\omega-\kappa)} + \frac{\alpha_2}{G_+(-\kappa)(\omega+\kappa)}\right],$$

$$Q_1^{(n)}(\omega) = \frac{1}{\pi}\int_k^\infty \frac{P_2^{(n-1)}(-it)e^{-t\delta}\,dt}{(t-i\omega)G_-(-it)(t^2-k^2)^{\frac{1}{2}}},$$

$$Q_1^{(0)}(\omega) = 0, \quad (100)$$

$$q^{(n)}(\omega, P_1^{(n)}) = \frac{B_0}{G_-(\omega)}\left[\frac{G_+(\kappa_0)P_1^{(n)}(\kappa_0)e^{i\kappa_0\delta}}{\kappa_0-\omega} + \frac{G_+(-\kappa_0)P_1^{(n)}(-\kappa_0)e^{-i\kappa_0\delta}}{\kappa_0+\omega}\right],$$

$$Q_2^{(n)}(\omega) = \frac{1}{\pi}\int_k^\infty \frac{P_1^{(n-1)}(it)G_+(it)e^{-t\delta}}{(t+i\omega)(t^2-k^2)^{\frac{1}{2}}}\,dt,$$

$$Q_2^{(0)}(\omega) = 0.$$

With these definitions the corresponding approximations for the amplitudes are specified by the system of equations

$$\begin{aligned}
\gamma_1^{(n)} &= iB_0e^{i\kappa_0\delta}P_1^{(n)}(\kappa_0),\\
\gamma_2^{(n)} &= -iB_0e^{-i\kappa_0\delta}P_1^{(n)}(-\kappa_0),\\
\alpha_3^{(n)} &= -\frac{\beta^2e^{-\beta a}}{\kappa G_+(\kappa)}\left[\frac{G_+(\kappa_0)}{\kappa-\kappa_0}\gamma_1^{(n)}\right.\\
&\qquad \left.+\frac{G_+(-\kappa_0)}{\kappa+\kappa_0}\gamma_2^{(n)} - iQ_2^{(n)}(+\kappa)\right],\\
&\left[\frac{G_+(\kappa_0)}{\kappa+\kappa_0}\gamma_1^{(n)} + \frac{G_+(-\kappa_0)}{\kappa-\kappa_0}\gamma_2^{(n)}\right.\\
&\qquad \left. + iQ_2^{(n)}(-\kappa)\right] = 0.
\end{aligned} \tag{101}$$

We note that when $n = 0$ we obtain the desired results.

The above iteration scheme, although formally simple in its application, is complicated by the fact that the integrals involved are not easily evaluated. Thus, it is of interest to consider methods which will provide suitable approximations to these integrals. In the following we indicate a procedure appropriate when the width δ of the barrier is large compared to the wavelength $2\pi/k$. In particular, we consider the second approximation ($n = 1$), i.e., the system

$$\begin{aligned}
&\frac{G_+(\kappa_0)}{\kappa+\kappa_0}\gamma_1^{(1)} + \frac{G_+(-\kappa_0)}{\kappa-\kappa_0}\gamma_2^{(1)} + iQ_2^{(1)}(-\kappa) = 0,\\
&\gamma_1^{(1)} = iP_1^{(1)}(\kappa_0)B_0e^{i\kappa_0\delta},\\
&\gamma_2^{(1)} = -iP_1^{(1)}(-\kappa_0)B_0e^{-i\kappa_0\delta},\\
&P_1^{(1)}(\omega) = F(\omega) + G_+(\omega)Q_2^{(1)}(\omega).
\end{aligned} \tag{102}$$

Thus we seek an approximation for

$$Q_2^{(1)}(\omega) = \frac{1}{\pi}\int_k^\infty \frac{F(it)G_+(it)e^{-t\delta}}{(t+i\omega)(t^2-k^2)^{\frac{1}{2}}}\,dt. \tag{103}$$

Since the range of parameters is such that $k\delta \gg 1$, the significant contribution to the integral (103) comes from the neighborhood of the lower limit. Omitting a constant multiplicative factor (which can be incorporated into the unknown wave amplitudes) we have $G_+(\omega)\approx(\omega+ik)^{\frac{1}{2}}$ for $\omega\approx ik$, and consequently

$$F(\omega)G_+(\omega) \approx i\beta e^{-\beta a}\left[\frac{\alpha_1}{G_+(\kappa)(\omega-\kappa)} + \frac{\alpha_2}{G_+(-\kappa)(\omega+\kappa)}\right](\omega+ik). \tag{104}$$

Inserting this expression into (103) leads to

$$Q_2^{(1)}(\omega) \approx \frac{i\beta e^{-\beta a}}{\pi}\int_k^\infty \left[\frac{\alpha_1}{G_+(\kappa)(t+i\kappa)} + \frac{\alpha_2}{G_+(-\kappa)(t-i\kappa)}\right]\frac{e^{-t\delta}}{(t+i\omega)}\left(\frac{t+k}{t-k}\right)^{\frac{1}{2}}dt. \tag{105}$$

Consider now the integral

$$I_0 = \int_k^\infty \frac{e^{-t\delta}}{(t+i\kappa)(t+i\omega)}\left(\frac{t+k}{t-k}\right)^{\frac{1}{2}}dt \tag{106}$$

which is fundamental in (105). Again the neighborhood of the lower limit will be the important range of integration, and, to emphasize this, we make the expansions

$$\begin{aligned}
\frac{1}{t+i\kappa} &= \frac{1}{k+i\kappa}\left[1 - \frac{t-k}{k+i\kappa} + \left(\frac{t-k}{k+i\kappa}\right)^2 - \cdots\right],\\
\frac{1}{t+i\omega} &= \frac{1}{k+i\omega}\left[1 - \frac{t-k}{k+i\omega} + \left(\frac{t-k}{k+i\omega}\right)^2 - \cdots\right].
\end{aligned} \tag{107}$$

If only the first term of these expressions be retained (106) reduces to

$$\begin{aligned}
I_0(\omega) &= \frac{1}{(k+i\omega)(k+i\kappa)}\int_k^\infty e^{-t\delta}\left(\frac{t+k}{t-k}\right)^{\frac{1}{2}}dt\\
&= \frac{ke^{-k\delta}}{(k+i\omega)(k+i\kappa)}\left[\left(\frac{2\pi}{k\delta}\right)^{\frac{1}{2}} + O\left(\frac{e^{-k\delta}}{(k\delta)^{\frac{3}{2}}}\right)\right].
\end{aligned} \tag{108}$$

We note that the retention of further terms of (107) results in terms of at least the order $O[e^{-k\delta}/(k\delta)^{\frac{1}{2}}]$; e.g., the next term is proportional to

$$\int_k^\infty e^{-t\delta}(t^2-k^2)^{\frac{1}{2}}\,dt \sim O\left(\frac{e^{-k\delta}}{(k\delta)^{\frac{3}{2}}}\right).$$

Substituting (108) into (105) we obtain to lowest order terms

$$Q_2^{(1)}(\omega) = i\beta e^{-\beta a}\left[\frac{\alpha_1}{G_+(\kappa)(k+i\kappa)(k+i\omega)} + \frac{\alpha_2}{G_+(-\kappa)(k-i\kappa)(k+i\omega)}\right]e^{-k\delta}\left(\frac{2k}{\pi\delta}\right)^{\frac{1}{2}}. \tag{109}$$

Using this expression in (102) gives, for example,

$$\frac{\alpha_1}{G_+(\kappa)}\left\{\frac{G_+^2(\kappa_0)}{\kappa+\kappa_0}B_0\beta e^{i\kappa_0\delta}\left[\frac{1}{(\kappa_0-\kappa)} + \frac{1}{(k+i\kappa_0)(k+i\kappa)}\left(\frac{2k}{\pi\delta}\right)^{\frac{1}{2}}e^{-k\delta}\right]\right.$$

$$+\frac{G_+^2(-\kappa_0)}{\kappa-\kappa_0}B_0\beta e^{-i\kappa_0\delta}\left[\frac{1}{(\kappa_0+\kappa)}-\frac{1}{(k+i\kappa)(k-i\kappa_0)}\left(\frac{2k}{\pi\delta}\right)^{\frac{1}{2}}e^{-k\delta}\right]+\frac{\beta}{(k^2+\kappa^2)}\left(\frac{2k}{\pi\delta}\right)^{\frac{1}{2}}e^{-k\delta}\Bigg\}$$

$$+\frac{\alpha_2}{G_+(-\kappa)}\left\{\frac{G_+^2(\kappa_0)}{\kappa+\kappa_0}B_0\beta e^{i\kappa_0\delta}\times\left[\frac{1}{\kappa_0+\kappa}+\frac{1}{(k+i\kappa_0)(k-i\kappa)}\left(\frac{2k}{\pi\delta}\right)^{\frac{1}{2}}e^{-k\delta}\right]\right.$$

$$+\frac{G_+^2(-\kappa_0)}{\kappa-\kappa_0}B_0\beta e^{-i\kappa_0\delta}\left[\frac{1}{(\kappa_0-\kappa)}-\frac{1}{(k-i\kappa)(k-i\kappa_0)}\left(\frac{2k}{\pi\delta}\right)^{\frac{1}{2}}e^{-k\delta}\right]$$

$$\left.+\frac{\beta}{(k-i\kappa)^2}\left(\frac{2k}{\pi\delta}\right)^{\frac{1}{2}}e^{-k\delta}\right\}=0, \tag{110}$$

a linear relationship between the incident wave amplitude (α_1) and the reflected amplitude (α_2) which takes into account all terms to the order $O(e^{-k\delta}/\delta^{\frac{1}{2}})$.

7. RESULTS FOR RELATED PROBLEMS

In this section we derive the pairs of integral equations, analogous to (83), appropriate to the finite dock and a submerged vibrating plate below the free surface of a semi-infinite fluid. The dock problem is concerned with the resultant fluid motion when a train of infinitesimal time-harmonic surface waves is incident on a plane rigid barrier of finite width located at the free surface of a semi-infinite fluid. The assumptions of Sec. 2 are made regarding the fluid motion, and hence we consider only the two-dimensional residual function $\varphi(x, y)$ of (6) which satisfies the differential equation (7), and the boundary conditions

$$\begin{aligned}\partial\varphi/\partial y|_{y=0} &= 0, && |x|<l,\\ \partial\varphi/\partial y|_{y=0} &= \beta\varphi(x,0), && |x|>l.\end{aligned} \tag{111}$$

the conditions at the dock and free surface, respectively.

Employing the Green's function of Sec. 3 we have the representation

$$\varphi(x,y)=\varphi_0(x,y)-\beta\int_{-l}^{l}\varphi(x',0)G(x,y;x',0)\,dx',$$

$$\varphi_0(x,y)=e^{\beta y}[\alpha_1e^{i\kappa(x+l)}+\alpha_2e^{-i\kappa(x+l)}]. \tag{112}$$

Letting $y = 0$ in (112) gives

$$\varphi(x,0)=\varphi_0(x,0)-\beta\int_{-l}^{l}\varphi(x',0)G(x,0;x',0)\,dx', \tag{113}$$

a one-dimensional integral equation for $\varphi(x, 0)$. In terms of the functions

$$\begin{aligned}\psi_1(x)&=\varphi(x,0), && x<-l, & \psi_2(x)&=\varphi(x,0), && x>l,\\ &=0, && x>-l, & &=0, && x<l,\\ I(x)&=-\varphi(x,0), && |x|<l, & \psi_0(x)&=\varphi_0(x,0), && \text{(all } x)\\ &=0 && |x|>l,\end{aligned} \tag{114}$$

the complex Fourier transform of (113) is

$$\Psi_1(\omega)+\Psi_2(\omega)-J(\omega)=\Psi_0(\omega)+\frac{\beta J(\omega)}{(\omega^2+k^2)^{\frac{1}{2}}-\beta},$$

$$J(\omega)=\int_{-l}^{l}I(x)e^{-i\omega x}\,dx, \quad \text{etc.}$$

where a convergence factor is assumed in $\Psi_0(\omega)$. Collecting terms, this equation may be written as

$$\Psi_1(\omega)+\Psi_2(\omega)-\Psi_0(\omega)=\frac{(\omega^2+k^2)^{\frac{1}{2}}J(\omega)}{(\omega^2+k^2)^{\frac{1}{2}}-\beta}, \tag{115}$$

which we note is the limit of (49) for $a \to 0$. Thus putting $a = 0$ in (51) we have for the dock problem

$$\begin{aligned}J(\omega)e^{-i\omega l}/L_-(\omega)&=P_1(\omega)/L_+(\omega)\\ &\quad+P_2(\omega)e^{-i\omega 2l}/L_+(\omega)-\Psi_0(\omega)e^{-i\omega l}/L_+(\omega),\\ J(\omega)e^{i\omega l}/L_+(\omega)&=P_1(\omega)e^{i\omega 2l}L_-(\omega)\\ &\quad+P_2(\omega)L_-(\omega)-\Psi_0(\omega)e^{i\omega l}L_-(\omega),\end{aligned} \tag{116}$$

which can be decomposed as for the submerged barrier.

Although the transform relations (116) are derived from (51) in a simple way ($a \to 0$) the two problems differ significantly in detail owing to the different role played by an edge at, or below, the surface. As we have seen, the behavior of a potential near an edge affects the asymptotic character of the corresponding transform, and so determines the nature of the entire functions obtained by separating the transform equations. In Sec. 3 it was shown that the first derivative of the potential function in the neighborhood of the edge is proportional to the inverse square root of the distance measured from the edge. Studies of the singularities characteristic to the dock have shown that in the neighborhood of an edge the velocity potential may be expanded in a double power series in r and $\log r$ where r is the distance from the edge (see, for example,

MacCamy[12] and Sparenberg[13]). The logarithmic term corresponds to the "breaking" of the fluid at the barrier, and its coefficient specifies the amount of breaking.[14,15]

In order to obtain a solution through (116) it is sufficient to assume that $\psi_{1,2}(x)$ are integrable in every finite interval of the x-axis. The Riemann–Lebesgue lemma gives $\Psi_{1,2}(\omega) \to 0$ for $|\omega| \to \infty$ in the appropriate half plane. In view of this fact, and since

$$L_-(\omega) \sim 0(\omega), \qquad |\omega| \to \infty, \qquad \operatorname{Im} \omega < 0,$$

we observe that the entire function $E_2(\omega)$ of (67) will no longer be zero since

$$P_2(\omega)L_-(\omega) \sim 0(\omega), \qquad |\omega| \to \infty, \qquad \operatorname{Im} \omega < 0.$$

However, Liouville's theorem states it can be at most a constant. Letting this constant value be A, the integral equations for $P_{1,2}(\omega)$ appropriate to the dock are

$$\begin{aligned} P_1(\omega)/L_+(\omega) + \eta_-(\omega) - \theta_+(\omega) &= 0, \\ P_2(\omega)L_-(\omega) + \xi_-(\omega) - \Omega_-(\omega) &= A. \end{aligned} \tag{117}$$

Employing the methods of Sec. 5 we find, for example, when $x > l$, $\varphi(x, y)$ is given by

$$\begin{aligned} \varphi(x, y) &= ie^{i\kappa(x-l)}\beta^2 e^{\beta y}[A + \xi_+(\kappa)]/\kappa L_+(\kappa) \\ &\quad - ie^{-i\kappa(x-l)}\beta^2 e^{\beta y}[A + \xi_+(-\kappa)]/\kappa L_+(-\kappa) \\ &\quad + \frac{\beta}{\pi}\int_k^\infty \frac{e^{-t(x-l)}}{L_+(it)(\sigma^2 + \beta^2)}[\xi_+(it) + A] \\ &\quad \times [\sigma \cos \sigma y + \beta \sin \sigma y]\, dt, \end{aligned} \tag{118}$$

where

$$\xi_+(it) = \frac{1}{2\pi i}\int_{-\infty}^{\infty} \frac{P_1(\omega)L_+(\omega)[(\omega^2 + k^2)^{\frac{1}{2}} - \beta]e^{i\delta\omega}}{(\omega - it)(\omega^2 + k^2)^{\frac{1}{2}}}\, d\omega, \tag{119}$$

and

$$\xi_+(\pm\kappa) = \frac{1}{\pi}\int_k^\infty \frac{P_1(it)L_+(it)e^{-t\delta}}{(t \pm i\kappa)(t^2 - k^2)^{\frac{1}{2}}}\, dt. \tag{120}$$

When $\delta \to \infty$, we have $\xi_+ = 0$, and (118) reduces to the results of Heins[5] for the semi-infinite dock.

The solution thus obtained has no singularity at the edges, however its x derivative does possess a logarithmic behavior and is a solution of the problem originally posed. Thus, a suitable linear combination of these two (φ and φ_x) may be taken in order to produce a new solution which exhibits a desired amount of "breaking" at the edges.

Consider now the case of a vibrating strip located below the free surface of a semi-infinite fluid. The undisturbed position of the free surface is taken as $y = 0$ and the rest position of the plate is specified by

$$y = -a, \qquad -l < x < l, \qquad -\infty < z < \infty.$$

The amplitude of the plate vibrations is assumed infinitesimal and the resulting fluid motion irrotational and time harmonic. Thus, as for the rigid barrier, the motion of fluid is characterized by a velocity potential $\Phi(x, y, z)e^{-i\omega t}$. We further assume that the z variation is given by the factor e^{ikz}. This amounts to assuming a wave in the plate traveling in the positive z direction. Thus, we seek a residual function $\varphi(x, y)$ which is a solution of (7) satisfying the boundary conditions

$$\begin{aligned} \partial\varphi/\partial y &= \beta\varphi, \qquad y = 0, \qquad -\infty < x < \infty, \\ \partial\varphi/\partial y &= V, \qquad y = -a, \qquad -l < x < l. \end{aligned} \tag{121}$$

In addition, $\varphi(x, y)$ is to describe outgoing waves for $|x| \to \infty$, i.e.,

$$\varphi(x, y) \sim O(e^{\beta y}e^{i\kappa|x|}), \qquad |x| \to \infty. \tag{122}$$

Noting the similarities between the present formulation and that for the rigid barrier we proceed immediately to the representation

$$\begin{aligned} \varphi(x, y) &= \alpha_2 e^{-i\kappa(x+l)+\beta y} \\ &\quad + \int_{-l}^{l} I(x')G_{y'}(x, y; x', -a)\, dx', \end{aligned} \tag{123}$$

where $I(x)$ is used to denote the discontinuity of $\varphi(x, y)$ at the plate. A one-dimensional integral equation for $I(x)$ is obtained by imposing the second condition of (121); thus

$$\begin{aligned} V &= \beta\alpha_2 e^{-i\kappa(x+l)-\beta a} \\ &\quad + \int_{-l}^{l} I(x')G_{yy'}(x, -a; x', -a)\, dx', \qquad -l < x < l. \end{aligned} \tag{124}$$

We extend this equation to all values of x by writing

$$\begin{aligned} \psi(x) &= \psi_0(x) \\ &\quad + \int_{-l}^{l} I(x')G_{yy'}(x, -a; x', -a)\, dx', \end{aligned} \tag{125}$$

[12] R. C. MacCamy, "Linear Boundary Problems Arising in the Diffraction of Water Waves by Surface Obstacles, Part I," Dept. of Mathematics, University of California at Berkeley, Technical Report No. 1 (1955).

[13] J. A. Sparenberg, "The Finite Dock," Technische Hogeschool, University of Delft, Report 17, 1957.

[14] J. J. Stoker, Quart. Appl. Math. 5, 1 (1957).

[15] H. Rubin, Comm. Pure Appl. Math. 7, 317 (1954).

where

$$\begin{aligned}
\psi_0(x) &= \beta\alpha_2 e^{-i\kappa(x+l)-\beta a}, \quad -\infty < x < \infty, \qquad \psi(x) = \psi_1(x) + \psi_2(x) + \psi_3(x),\\
\psi_1(x) &= \partial\varphi/\partial y|_{y=-a}, \quad x < l, \qquad \psi_2(x) = \partial\varphi/\partial y|_{y=-a}, \quad x > l,\\
&= 0, \quad x > l, \qquad = 0, \quad x < l,\\
\psi_3(x) &= V, \quad |x| < l, \qquad I(x) = \varphi(x, -a+0) - \varphi(x, -a-0), \quad |x| < l,\\
&= 0, \quad |x| > l, \qquad = 0, \quad |x| > l.
\end{aligned} \tag{126}$$

The transform version of (125) is

$$\Psi_1(\omega) + \Psi_2(\omega) + \frac{V}{i\omega}(e^{i\omega l} - e^{-i\omega l}) = \Psi_0(\omega) + J(\omega)G(\omega), \tag{127}$$

where the use of the convergence factor $e^{-\epsilon|x|}$ is implicit in the transform $\Psi_0(\omega)$. Employing the methods of Secs. 3 and 4, we obtain

$$\begin{aligned}
\Psi_1(\omega) &\sim O\left(\frac{e^{i\omega l}}{\omega^{\frac{1}{2}}}\right), \quad |\omega| \to \infty, \quad \operatorname{Im}\omega > 0,\\
\Psi_2(\omega) &\sim O\left(\frac{e^{-i\omega l}}{\omega^{\frac{1}{2}}}\right), \quad |\omega| \to \infty, \quad \operatorname{Im}\omega < 0,\\
J(\omega) &\sim O\left(\frac{e^{i\omega l}}{\omega^{\frac{1}{2}}}\right), \quad |\omega| \to \infty, \quad \operatorname{Im}\omega > 0,\\
&\sim O\left(\frac{e^{-i\omega l}}{\omega^{\frac{1}{2}}}\right), \quad |\omega| \to \infty, \quad \operatorname{Im}\omega < 0.
\end{aligned} \tag{128}$$

The transform relation (127) differs from that for the rigid barrier [Eq. (41)] by the presence of the term

$$Q(\omega) = \frac{V}{i\omega}(e^{i\omega l} - e^{-i\omega l}),$$

which stems from the prescribed normal velocity of the plate. Evidently $Q(\omega)$ is regular in every finite region of the ω-plane, and we have

$$\begin{aligned}
Q(\omega)e^{i\omega l} &\sim O\left(\frac{1}{\omega}\right), \quad |\omega| \to \infty, \quad \operatorname{Im}\omega > 0,\\
Q(\omega)e^{-i\omega l} &\sim O\left(\frac{1}{\omega}\right), \quad |\omega| \to \infty, \quad \operatorname{Im}\omega < 0.
\end{aligned} \tag{129}$$

Suppressing the exponential behavior of $\Psi_1(\omega)$ and $\Psi_2(\omega)$ by introducing functions $P_1(\omega)$ and $P_2(\omega)$, we obtain

$$P_1(\omega)e^{i\omega l} + P_2(\omega)e^{-i\omega l} + Q(\omega) = \Psi_0(\omega) + J(\omega)G(\omega), \tag{130}$$

which is analogous to Eq. (49) for the rigid barrier. As in the case of this latter equation, (130) yields the two expressions

$$\begin{aligned}
&G_-(\omega)P_1(\omega)e^{i\omega\delta} + P_2(\omega)G_-(\omega) + Q(\omega)G_-(\omega) - \Psi_0(\omega)G_-(\omega)e^{i\omega l} = J(\omega)G_+(\omega)e^{i\omega l},\\
&\frac{P_1(\omega)}{G_+(\omega)} + \frac{P_2(\omega)}{G_+(\omega)}e^{-i\omega\delta} + \frac{Q(\omega)}{G_+(\omega)}e^{-i\omega l} - \frac{\Psi_0(\omega)}{G_+(\omega)}e^{-i\omega l} = \frac{J(\omega)}{G_-(\omega)}e^{-i\omega l}.
\end{aligned} \tag{131}$$

Comparing these equations with (51) we write by inspection

$$\begin{aligned}
&P_2(\omega)G_-(\omega) + \xi_-(\omega) - \Omega_-(\omega) + \frac{1}{2\pi i}\int_{C_-}\frac{Q(t)G_-(t)e^{itl}}{(t-\omega)}\,dt = 0,\\
&P_1(\omega)/G_+(\omega) + \eta_+(\omega) - \theta_+(\omega) + \frac{1}{2\pi i}\int_{C_+}\frac{Q(t)e^{-itl}}{G_+(t)(t-\omega)}\,dt = 0.
\end{aligned} \tag{132}$$

Closing the contour C_+ in the lower half of the ω-plane, with a branch cut along the imaginary axis from $-ik$ to $-i\infty$, leads to

$$\frac{1}{2\pi i}\int_{C_+}\frac{Q(t)e^{-itl}}{G_+(t)(t-\omega)}\,dt = -\frac{V}{\pi}\int_k^\infty \frac{(1-e^{-t\delta})\,dt}{tG_-(-it)(t-i\omega)(t^2-k^2)^{\frac{1}{2}}}.$$

In the limit $\epsilon \to 0$, we have the integral equation

$$\frac{P_1(\omega)}{G_+(\omega)} = \frac{i\beta\alpha_2 e^{-\beta a}}{G_+(-\kappa)(\omega+\kappa)} + \frac{1}{\pi}\int_k^\infty \frac{[tP_2(-it)e^{-t\delta} + V(t-e^{-t\delta})]}{(t-i\omega)G_-(-it)(t^2-k^2)^{\frac{1}{2}}t}\,dt. \tag{133}$$

The present situation is somewhat simplified in view of the symmetry $\psi_1(x) = \psi_2(-x)$, and consequently $P_1(\omega) = P_2(-\omega)$. Thus, Eq. (133) can be iterated directly for $P_1(\omega)$ (or $P_2(-\omega)$). The methods of Sec. 5 can be used to find representations for the velocity potential.

Diffraction by a Smooth Transparent Object*

YUNG MING CHEN†
Courant Institute of Mathematical Sciences, New York University, New York, New York
(Received 19 November 1963)

In previous investigations the so-called geometrical theory of diffraction was established. Its application has been limited to problems of diffraction of waves by smooth opaque objects.

In the present paper, the geometrical theory of diffraction is extended and applied to problems of diffraction by a smooth transparent object of any shape. For simplicity only scalar fields and two-dimensional problems are considered. However, this method can be modified easily to apply to vector fields and three-dimensional problems.

In Part I, the fields associated with the geometric rays and the diffracted rays are constructed in detail for the case of a circular cylinder. Afterwards a general method for treating diffraction by a smooth transparent cylinder of arbitrary shape is given. In order to determine the coefficients, the exact solution for the case of a circular cylinder is found and evaluated asymptotically for large ka in Part II. Then it is compared with the solution obtained from geometrical theory in Part I.

I. GEOMETRICAL THEORY OF DIFFRACTION

Introduction

IN recent years, J. B. Keller's "Geometrical Theory of Diffraction" has become a useful technique for solving asymptotically the boundary-value problems of wave propagation. Whereas the usual method for solving boundary-value problems of this type depends upon separation of variables, this method does not. Therefore, it can be applied to objects of arbitrary shape. Most uses of this theory have concerned diffraction by opaque objects. However we shall apply it to diffraction by smooth transparent objects.

For simplicity we will consider only scalar fields and two-dimensional problems. However, our method can be modified easily to apply to vector fields and three-dimensional problems. First we shall treat diffraction by a circular cylinder. This has the advantage of illustrating the method without introducing extraneous geometrical details, and it is also a case for which we can obtain the exact solution for comparison. Afterwards, a general recipe will be given for treating diffraction by a smooth transparent cylinder of arbitrary shape.

For the historical background and the general discussion of the geometrical theory of diffraction, the reader is referred to articles by Keller.[1,2] Levy and Keller[3] and Keller.[4] These also contain many additional references.

In the following section we shall construct the fields on the geometric rays and on the diffracted rays which arise in scattering by a circular cylinder. These fields are compared with the leading terms in the short wavelength asymptotic expansion of the exact solution in Part II and the two are found to agree with each other. The coefficients of reflection, transmission, and diffraction are thus determined and the postulates verified. In the last section, we give a procedure for calculating the asymptotic form of the field diffracted by a smooth transparent cylinder of arbitrary shape.

The author wishes to express his appreciation to Professor Joseph B. Keller for suggesting this problem and for his aid and guidance in carrying out the research.

A. Diffraction by a Transparent Circular Cylinder

Let us consider the scalar field μ produced by a time-harmonic line source in the presence of a circular cylinder of a homogeneous material different from that which surrounds it. We shall call the surrounding homogeneous medium "one" and that constituting the cylinder "two." For simplicity we consider the case in which the source is parallel to the cylinder axis. The cylinder is of radius a, the source is at distance r_0 from the axis, and the field point is denoted by (r, Ω), where r is the radial distance from the axis, and Ω is the angle between $\mathbf{r}_0$ and $\mathbf{r}$ measured clockwise (Fig. 1). On the cylinder,

* This research was supported by the Office of Naval Research, under Contract Nonr 285 (48). Reproduction in whole or in part is permitted for any purposes of the U. S. Government.

† Present address: Division of Mathematical Sciences, Purdue University, Lafayette, Indiana.

[1] J. B. Keller, "A Geometrical Theory of Diffraction," Symposium on Microwave Optics, Eaton Electronics Research Laboratory, McGill University, Montreal, Canada, (1953).

[2] J. B. Keller, J. Opt. Soc. Am. 52, 116 (1962).

[3] B. R. Levy and J. B. Keller, Commun. Pure Appl. Math. 12, 159 (1959).

[4] J. B. Keller, IRE Trans. Antennas Propagation 4, 312 (1956).

μ satisfies the boundary conditions

$$\mu(a+, \Omega) = \alpha\mu(a-, \Omega) \text{ and } \mu_r(a+, \Omega) = \beta\mu_r(a-, \Omega),$$

where $a\pm = \lim_{\epsilon\to 0} (a \pm \epsilon)$ with $\epsilon \geq 0$; α and β are given constants depending upon the two media.

1. Geometric Rays

According to the geometrical theory of diffraction, the field at any point is the sum of the fields on all the rays through that point. The field on each ray consists of an amplitude and a phase. The phase is proportional to the optical length of the ray and the proportionality constant is the propagation constant k_1 or k_2. The amplitude varies in accordance with the principle of conservation of energy in a narrow tube of rays, so it is inversely proportional to the square root of the distance between two neighboring rays. The incident rays are straight lines from the source. They reflect and refract at the cylinder surface, traverse the cylinder and reflect or refract again, etc. On the incident ray, the field is specified at the source, and on a reflected or refracted ray, the initial value is obtained by multiplying the field on the incident ray by a reflection or refraction (transmission) coefficient.

Figs. 1 and 2 show some of the tubes of geometric optics rays when $N \equiv k_2/k_1 > 1$ and $N < 1$, respectively. They also show the physical dimensions and notations.

We now let R_{11} be the reflection coefficient in medium 1, R_{22} be that in medium 2, T_{12} be the transmission coefficient from medium 1 to medium 2, T_{21} be that from medium 2 to medium 1, and $p - 1$ be the number of internal reflections a refracted ray undergoes. According to the geometrical theory of diffraction, the coefficients above are Fresnel coefficients and the relation between the angles of incidence φ and refraction θ is given by Snell's law. Therefore we have from the boundary conditions

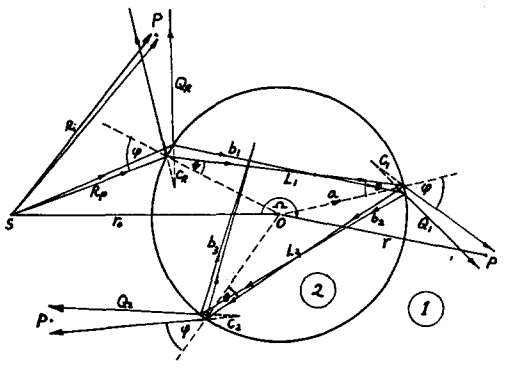

FIG. 1. The behavior of an infinitesimal tube of geometric rays is shown for the case of $N > 1$.

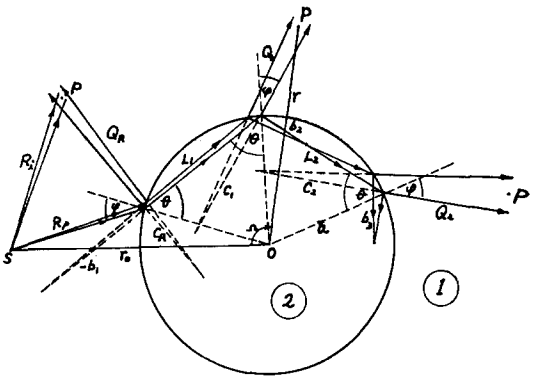

FIG. 2. The behavior of an infinitesimal tube of geometric rays is shown for the case of $N < 1$.

$$R_{11} = \frac{\alpha \cos\varphi - \beta N \cos\theta}{\alpha \cos\varphi + \beta N \cos\theta}$$

where

$$N = k_2/k_1 \neq 1, \tag{1.1}$$

$$R_{22} = -R_{11}, \tag{1.2}$$

$$\alpha T_{12} = 1 + R_{11}, \tag{1.3}$$

$$\alpha^{-1} T_{21} = 1 + R_{22}, \tag{1.4}$$

$$k_1 \sin\varphi = k_2 \sin\theta \tag{1.5}$$

(Snell's law).

In general, at any point of the geometrically illuminated region, the field contains contributions from the incident ray μ_i^G, the reflected ray μ_r^G, and all the multiply reflected rays which we write as

$$\sum_{j=+,-} \sum_{p=1}^{\infty} \sum^{j,p} \mu_{pj}^G,$$

where "+" denotes the contributions from those rays which are incident on the upper half of the cylinder, "−" denotes those which are incident on the lower half of the cylinder, and $\sum^{j,p}$ means a sum of all terms μ_{pj}^G for fixed (j, p).

We now proceed to construct the expressions for μ_i^G, μ_r^G, and μ_{pj}^G. μ_i^G has an amplitude factor $(R_i)^{-\frac{1}{2}}$ which comes from applying the principle of conservation of energy. It also has a phase factor $e^{ik_1R_i}$, where R_i is the distance from the source to the field point. To adjust the incident field to be that of a source of unit strength, we include a factor $\mathfrak{A} \equiv (i/8\pi k_1)^{\frac{1}{2}}$. Thus we have

$$\mu_i^G = \mathfrak{A} R_i^{-\frac{1}{2}} e^{ik_1R_i}, \tag{1.6}$$

μ_r^G has the amplitude factors $(R_r)^{-\frac{1}{2}}$ and $(1+Q_r/C_r)^{-\frac{1}{2}}$, which come from applying the energy principle to the portions of the ray from the source to the external reflection point, and from the external reflection

point to the field point, respectively. Because of the external reflection, the reflection coefficient R_{11} is included. The phase factor is $e^{ik_1(R_r+Q_r)}$. Here R_r is the distance from the source to the point of external reflection, Q_r is the distance from the point of external reflection to the field point, and C_r is the distance from the point of external reflection to the point of convergence of two neighboring rays along Q_r. Thus we have

$$\mu_r^G = \mathfrak{N}\mathfrak{M}_r R_{11} R_r^{-\frac{1}{2}} e^{ik_1(R_r+Q_r)}, \tag{1.7}$$

where

$$\mathfrak{M}_{ij} \equiv (1 + Q_{ij}/C_{ij})^{-\frac{1}{2}}. \tag{1.8}$$

μ_{pj}^G has the amplitude factors

$$1/(R_{pj})^{\frac{1}{2}}, \qquad \mathfrak{L}_{1j}, \mathfrak{L}_{2j}, \cdots, \mathfrak{L}_{pj}, \mathfrak{M}_{pj},$$

where

$$\mathfrak{L}_{ij} = (1 - L_{ij}/b_{ij})^{-\frac{1}{2}}, \tag{1.9}$$

which come from applying the energy principle to the portions of the ray from the source to the external reflection point, from the external reflection point to the first internal reflection point, from the first internal reflection point to the second internal reflection point, $\cdots$, from $(p-1)$st internal reflection point to the pth internal reflection point, and from the pth internal reflection point to the field point, respectively. Since the ray travels through the interface of the two media twice and is internally reflected $(p-1)$ times, μ_{pj}^G has the additional amplitude factor $T_{12}T_{21}R_{22}^{p-1}$ and the phase factor

$$\exp\left[ik_1(R_{pj} + Q_{pj}) + ik_2 \sum_{q=1}^{p} L_{qj}\right].$$

Here L_{qj} is the chord length from the $(q-1)$st internal reflection point to the qth internal reflection point, R_{pj} is the distance from the source to the external reflection point (0th internal reflection point), Q_{pj} is the distance from the pth internal reflection point to the field point, C_{pj} is the distance from the pth internal reflection point to the point of convergence of two neighboring rays along Q_{pj}, and b_q is the distance from the $(q-1)$st internal reflection point to the point of convergence of two neighboring rays along L_q. Hence we have the following expression for μ_{pj}^G:

$$\mu_{pj}^G = \mathfrak{N}\mathfrak{L}_{1j}\mathfrak{L}_{2j}\cdots\mathfrak{L}_{pj}\mathfrak{M}_{pj}(R_{pj})^{-\frac{1}{2}}T_{12}T_{21}R_{22}^{p-1} \times \exp\left[ik_1(R_{pj} + Q_{pj}) + ik_2 \sum_{q=1}^{p} L_{qj}\right]. \tag{1.10}$$

Finally we can write the geometric optics field at (r, Ω) as

$$\mu^G(r, \Omega) = \{[\mu_i^G(r, \Omega) + \mu_r^G(r, \Omega)]H(\Omega) + [\mu_i^G(r, 2\pi - \Omega) + \mu_r^G(r, 2\pi - \Omega)] \times H(-\Omega) \quad \text{in lit region}, \quad 0 \quad \text{in geometric shadow region}\} + \sum_{j=+,-}\sum_{p=1}^{\infty}\sum^{j,p}\mu_{pj}^G, \qquad r > a, \quad 0 \le |\Omega| \le 2\pi, \tag{1.11}$$

with

$$\mu^G(r, 0) = \mu^G(r, 2\pi), \quad \text{and} \quad \mu_i^G(r, 0) = \mu_i^G(r, 2\pi).$$

$H(\Omega - \Omega_0)$ is defined as

$$H(\Omega - \Omega_0) = \begin{cases} 0 & \text{for } \Omega < \Omega_0, \\ \frac{1}{2} & \text{for } \Omega = \Omega_0, \\ 1 & \text{for } \Omega > \Omega_0. \end{cases} \tag{1.12}$$

By geometric considerations we arrive at the following results:

$$R_i = (r^2 + r_0^2 - 2rr_0\cos\Omega)^{\frac{1}{2}}, \quad r > a, \quad 0 \le |\Omega| \le 2\pi, \tag{1.13}$$

$$(C_r)^{-1} = (R_r)^{-1} + (\tfrac{1}{2}a\cos\varphi)^{-1} \tag{1.14}$$

$$\left.\begin{aligned} R_\rho &= \mathfrak{R}_0, \\ Q_\rho &= \mathfrak{R}, \end{aligned}\right\} \quad \rho = r, 1, 2, 3, \cdots, p, \tag{1.15, 1.16}$$

where

$$\mathfrak{R}_0 \equiv (r_0^2 - a^2\sin^2\varphi)^{\frac{1}{2}}, \tag{1.17}$$

and

$$\mathfrak{R} \equiv (r^2 - a^2\sin^2\varphi)^{\frac{1}{2}}, \tag{1.18}$$

$$L_{q_b} = 2a\cos\theta = L, \quad q = 1, 2, 3, \cdots, p, \tag{1.19}$$

$$\frac{1}{b_1} = \frac{1}{a\cos\theta}\left[1 - \left(1 + \frac{a\cos\varphi}{R_1}\right)\frac{\cos\varphi}{N\cos\theta}\right], \tag{1.20}$$

$$\frac{1}{c_1} = \frac{\cos\theta}{\cos\varphi}\left[\frac{1}{(L - b_1)(\cos\varphi/N\cos\theta)} - \frac{(N\cos\theta/\cos\varphi - 1)}{a\cos\theta}\right], \tag{1.21}$$

$$\frac{1}{b_{pj}} = \frac{2}{a\cos\theta} - \frac{1}{L - b_{p-1,j}}, \qquad p = 2, 3, 4, \cdots, \tag{1.22}$$

$$\frac{1}{C_{pj}} = \frac{\cos\theta}{\cos\varphi}\left[\frac{N\cos\theta}{(L - b_{pj})\cos\varphi} - \frac{N}{a\cos\varphi} + \frac{1}{a\cos\theta}\right], \quad p = 1, 2, 3, 4, \cdots, \tag{1.23}$$

for

$$j = \pm, \quad 2\pi M_{p\pm}^{\pm}\Omega = 2\varphi + p(\pi - 2\theta) - \sin^{-1}[(a/r)\sin\varphi] - \sin^{-1}[(a/r_0)\sin\varphi],$$
$$p = r, 1, 2, 3, 4, \cdots, \qquad (1.24)$$

with M_p^+ and M_p^- being the smallest positive integers such that Eq. (1.24) has real solutions.

2. *Diffracted Rays: Ray Tracing*

In the geometrical theory of diffraction, diffracted rays are introduced. Away from the surface diffracted rays behave just like ordinary geometric rays. On the surface a diffracted ray, which is called a surface ray, is a geodesic arc on the fast side of the surface.

The behavior of a surface ray depends on the local curvature of the boundary surface. In general, the local properties of the surface can be classified into two types. A surface, convex towards the fast side, is called a Type A surface, and a surface, convex towards the slow side, is called a Type B surface. For a Type A surface, a diffracted ray can be produced only by an incident ray on the fast side hitting the boundary surface tangentially. When such a surface ray travels along a Type A surface, it continually sheds tangent rays into the fast medium and critically refracted rays into the slow medium. From a Type B surface a diffracted ray can be produced by incident rays in the following two ways: (I) When an incident ray on the slow side hits the boundary surface at the critical angle of incidence. As such a surface ray travels along the surface, it continually sheds critically refracted rays into the slow medium. (II) When an incident ray on the slow side hits the boundary surface tangentially. As such a ray travels along the surface, it continually sheds tangent rays into the slow medium.

For diffraction by a transparent circular cylinder, the diffracted rays in the case $N > 1$ and the case $N < 1$ are quite different, because their boundary surfaces are purely of Type A and Type B, respectively. Thus in the following we give a separate account of all the possible diffracted rays passing an exterior point of the cylinder in each case.

In the case $N > 1$, any exterior point can have infinitely many diffracted rays of the following kinds passing through it (see Fig. 3).

The first kind are rays which start from the source, hit the cylinder tangentially, travel along the surface on the exterior side, and then leave the surface tangentially.

The second kind are rays which start from the source, hit the cylinder tangentially, travel along the surface on the exterior side, are critically refracted into the cylinder, traverse the cylinder, hit the other side of the cylinder, are critically refracted, travel along the surface on the exterior side, and finally leave the surface tangentially.

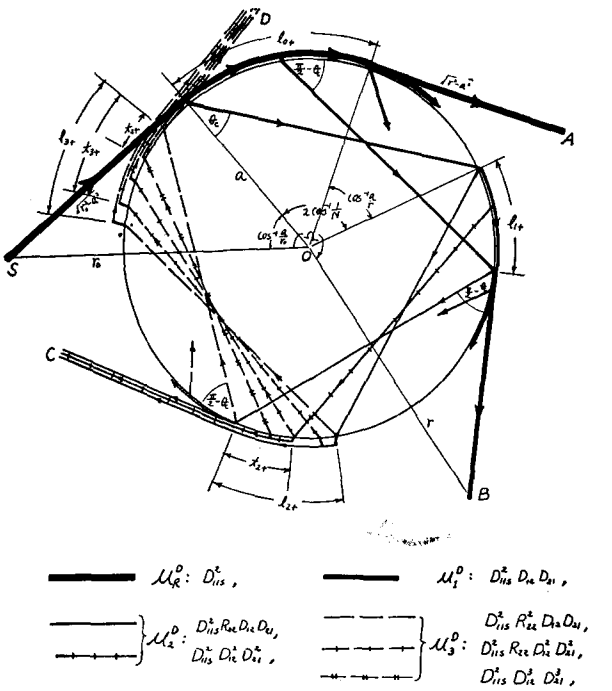

FIG. 3. The ray paths of the diffracted rays in the case of $N > 1$ are shown in detail. The corresponding proportionality constants of the field, associated with each diffracted ray, are given.

The third kind are like rays of the second kind, except when they hit the other side of the cylinder, instead of being critically refracted they are reflected back into the cylinder; then they again traverse the cylinder, hit another side of the cylinder, are critically refracted, travel along the surface on the exterior side, and finally leave the surface tangentially.

The fourth kind are like rays of the second kind except when they travel along the surface, instead of leaving the surface tangentially they are critically refracted again into the cylinder, traverse the cylinder, hit the other side of the cylinder, are critically refracted, travel along the surface on the exterior side, and finally leave the surface tangentially.

The rest of the diffracted rays are just like the rays of third and fourth kinds, except with higher numbers of traversals of the cylinder which are due to either reflection or refraction.

In the case $N < 1$, any exterior point can have

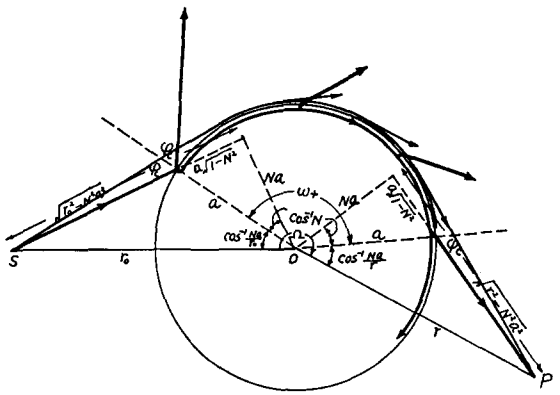

FIG. 4. The ray paths of the diffracted rays in the case of $N > 1$ are shown. The notations for the angles and distances are given.

infinitely many diffracted rays of the following kinds passing through it (see Fig. 4).

The first kind are rays which start from the source, hit the cylinder at a critical angle of incidence, are critically refracted, travel along the surface on the interior side, and are finally critically refracted into the exterior of the cylinder.

The second kind are rays which start from the source, hit the cylinder tangentially, travel along the surface on the exterior side, and then leave the surface tangentially.

Field Calculation. According to the geometrical theory of diffraction, the calculation of the field on a diffracted ray away from the boundary surface is the same as that on a geometric ray. On the surface portion of a diffracted ray, the phase of the field increases in proportion to the optical length of the ray. Along the surface ray, the amplitude varies in accordance with the principle of energy conservation in a narrow strip of surface rays. The amplitude also decays because of radiation from the surface and the rate of decay is governed by a decay exponent which depends upon local properties of the boundary surface and the media. The boundary surface is a caustic of the diffracted fields and, in general, the field within the caustic layer is composed of a number of different modes. Each mode is characterized by its own amplitude and phase, and therefore its own proportionality constant, propagation constant, and decay exponent.

Let the amplitude of the field associated with a diffracted ray which passes through a point of interaction (refraction, reflection, or diffraction) be described by the relation $A_d(c) = D(c)A_i(c)$, where c is the interaction point, $A_d(c)$ is the amplitude on the ray produced at c, $A_i(c)$ is the amplitude on the incident ray at c, and $D(c)$ is the proportionality constant. The proportionality constant is called a transmission, reflection, or diffraction coefficient, according as the interaction is refraction, reflection, or diffraction.

Now we shall construct the diffracted fields of diffraction by a transparent circular cylinder separately for case $N > 1$ and case $N < 1$.

For the case $N > 1$: First we let D_{11} be the diffraction coefficient from medium (1) to medium (1), D_{12} be the refraction coefficient from medium (1) to medium (2), D_{21} be that from medium (2) to medium (1), $\theta_0 = \sin^{-1}(1/N)$ be the angle of total reflection (or angle of critical incidence), and $\nu_s a^{-1}$ be the propagation constant for the waves propagating along the boundary surface only. The subscript s denotes the s mode. Since the wave is propagating in the fast medium [medium 1], we expect that the real part of $\nu_s a^{-1}$ is approximately equal to k_1, and because of the decay exponent we expect that the imaginary part of $\nu_s a^{-1}$ is positive. The $\ell_{0\pm}$, $\ell_{1\pm}$, $\ell_{2\pm}$, $\cdots$ and $t_{2\pm}$, $t_{3\pm}$, $t_{4\pm}$, $\cdots$ are the arc lengths shown in Fig. 3, and defined later.

The fields on the rays, which travel around the cylinder clockwise a distance ℓ_{0+} or around the cylinder counterclockwise a distance l_{0-} are

$$\mu^{D}_{r_{s\pm}}(r, \Omega) = \sum_{s} \mathfrak{N} D^2_{11_s}(\mathfrak{D}\mathfrak{D}_0)^{-\frac{1}{2}} \times \exp\left[ik_1(\mathfrak{D} + \mathfrak{D}_0) + i\frac{\nu_s}{a}\ell_{0\pm}\right], \quad r > a, \quad 0 \leq |\Omega| \leq 2\pi, \tag{1.25}$$

with

$$\ell_{p\pm} = \left(2\pi n \pm \Omega - \cos^{-1}\frac{a}{r} - \cos^{-1}\frac{a}{r_0} - 2p\cos^{-1}\frac{1}{N}\right)a, \quad p = 0, 1, 2, 3, \cdots, \tag{1.26}$$

and n is any positive integer such that $\ell_{p\pm} \geq 0$,

$$\mathfrak{D} = (r^2 - a^2)^{\frac{1}{2}}, \tag{1.27}$$

$$\mathfrak{D}_0 = (r_0^2 - a^2)^{\frac{1}{2}}. \tag{1.28}$$

The total contribution of rays of this type is the sum of the above expressions over all n's.

For the second kind of rays, at the point B of Fig. 3 the field is contributed to by rays diffracted at all points along the whole arc of length ℓ_{1+}. Thus we must sum over these rays by introducing the factor $\int_0^{\ell_{1+}} dt_{1+}/a$ into the expression for $\mu^{D}_{1_{s+}}(r, \Omega)$ and $\int_0^{l_{1-}} dt_{1-}/a$ into the expression for $\mu^{D}_{1_{s-}}(r, \Omega)$, where dt_1 is the infinitesimal arc length measured from the point of tangency of the ray B towards

the point of second refraction of μ_1^D. Then we have

$$\mu_{1n\pm}^D(r, \Omega) = \sum_s \Re D_{11s}^2 D_{12} D_{21} (\mathfrak{D}\mathfrak{D}_0)^{-\frac{1}{2}} \int_0^{t_{1s}} \frac{dt_1}{a} \times \exp\left[ik_1(\mathfrak{D} + \mathfrak{D}_0) + i2k_2 a \cos\theta_0 + i\frac{\nu_s}{a}\ell_{1\pm}\right], \quad r > a, \quad 0 \le |\Omega| \le 2\pi. \tag{1.29}$$

For the third and fourth kinds of rays, we also introduce the factors

$$R_{22}\int_0^{t_2} \frac{dt_2}{a}, \quad \text{and} \quad \int_0^{t_2} \frac{D_{12}D_{21}}{a}\left(\int_0^{t_2}\frac{dt_1}{a}\right) dt_2,$$

respectively where dt_1 is the infinitesimal arc length measured from the point of tangency of the ray C towards the third point of refraction of μ_2^D, t_2 is the arc length measured from the point of tangency of the ray C to the third point of refraction of μ_2^D, and dt_2 is the infinitesimal variation of t_2 measured from t_2 towards the second point of refraction. Then we have

$$\mu_{2n\pm}^D(\Omega, r) = \sum_s \Re D_{11s}^2 D_{12} D_{21} \times \left[\int_0^{t_{2s}} \left(R_{22} + D_{12}D_{21}\int_0^{t_2}\frac{dt_1}{a}\right)\frac{dt_2}{a}\right] \times (\mathfrak{D}\mathfrak{D}_0)^{-\frac{1}{2}} \exp\left[ik_1(\mathfrak{D} + \mathfrak{D}_0) + i4k_2 a\cos\theta_0 + i\frac{\nu_s\ell_{2\pm}}{a} - i\pi\right], \quad r > a, \quad 0 \le |\Omega| \le 2\pi. \tag{1.30}$$

Diffracted rays of type $p = 3$ come from reflection of diffracted rays of type $p = 2$ (See the long dashed lines in Fig. 3). With the same consideration as before, we obtain

$$\mu_{3n\pm}^D(r, \Omega) = \sum_s \Re D_{11s}^2 D_{12} D_{21}\left[\int_0^{t_{3s}}\left\{R_{22} + D_{12}D_{21}\int_0^{t_2}\left(R_{22} + D_{12}D_{21}\int_0^{t_2}\frac{dt_1}{a}\right)\frac{dt_2}{a}\right\}\frac{dt_3}{a}\right] \times (\mathfrak{D}\mathfrak{D}_0)^{-\frac{1}{2}} e^{ik_1(\mathfrak{D}+\mathfrak{D}_0)} \exp\left[+i6k_2 a\cos\theta_0 + i\frac{\nu_s}{a}\ell_{3\pm} - i\frac{3\pi}{2}\right], \quad r > a, \quad 0 \le |\Omega| \le 2\pi. \tag{1.31}$$

If we introduce the following notation for angular distances,

$$\omega_{p\pm} = \ell_{p\pm}/a, \tag{1.32}$$

then the general expressions for $\mu_{pn}^D(r, \Omega)$ are

$$\mu_{pn\pm}^D(r, \Omega) = \sum_s \Re D_{11s}^2 D_{12} D_{21} (\mathfrak{D}\mathfrak{D}_0)^{-\frac{1}{2}} \times \left[R_{22}^{p-1}\omega_{p\pm} + (p-1)R_{22}^{p-2}(D_{12}D_{21})\frac{\omega_{p\pm}^2}{2!} + \frac{(p-1)(p-2)}{2!}R_{22}^{p-3}(D_{12}D_{21})^2\frac{\omega_{p\pm}^3}{3!} + \cdots + \frac{(p-1)!}{(p-1-k)!\,k!}R_{22}^{p-1-k}(D_{12}D_{21})^k\frac{\omega_{p\pm}^{k+1}}{(k+1)!} + \cdots + (D_{12}D_{21})^{p-1}\frac{\omega_{p\pm}^p}{p!}\right] \exp\left[ik_1(\mathfrak{D} + \mathfrak{D}_0) + i2pk_2 a\cos\theta_0 + i\nu_s\omega_{p\pm} - ip(\tfrac{1}{2}\pi)\right]; \quad r > a, \quad 0 \le |\Omega| \le 2\pi, \tag{1.33}$$

where we define the negative powers of R_{22}, D_{12}, and D_{21} to be zero.

As in previous cases, the total diffracted field μ_p^D is the sum of all μ_{pn+}^D's and μ_{pn-}^D's.

For the case $N < 1$: We let $\tilde{\nu}_s a^{-1}$ and $\nu_s a^{-1}$ be the propagation constants for the waves propagating along the boundary surface on the fast side and slow side, respectively. For the waves propagating in the fast medium and slow medium, we expect that the real part of $\tilde{\nu}_s a^{-1}$ and $\nu_s a^{-1}$ are approximately equal to k_2 and k_1, respectively, and because of the decay exponent we expect that both the imaginary parts of $\tilde{\nu}_s a^{-1}$ and $\nu_s a^{-1}$ are positive.

By ray tracing and field calculations we obtain the following results for the first kind of ray:

$$\mu_{pn\pm}^{D1}(r, \Omega) = \sum_s \Re D_{12s} D_{21s} (\mathcal{E}\mathcal{E}_0)^{-\frac{1}{2}} \times \exp\left[ik_1[\mathcal{E} + \mathcal{E}_0 - 2a(1-N^2)^{\frac{1}{2}}] + i\tilde{\nu}_s\tilde{\omega}_{1\pm}\right], \quad r > a, \quad 0 \le |\Omega| \le 2\pi, \tag{1.34}$$

where

$$\mathcal{E} = (r^2 - N^2a^2)^{\frac{1}{2}}, \tag{1.35}$$

$$\mathcal{E}_0 = (r_0^2 - N^2a^2)^{\frac{1}{2}}, \tag{1.36}$$

and

$$\tilde{\omega}_{p\pm} = \left(2\pi n \pm \Omega - \cos^{-1}\frac{Na}{r} - \cos^{-1}\frac{Na}{r_0} + 2p\cos^{-1}N\right), \tag{1.37}$$

with n as any positive integer such that $\tilde{\omega}_p \pm \ge 0$. The physical meaning of $\mathcal{E}$, $\mathcal{E}_0$, $a(1-N^2)^{\frac{1}{2}}$, ω_+, φ_c, etc., are given in Fig. 4. For the second kind of diffracted ray, we have

$$\mu_{n\pm}^{D2}(r, \Omega) = \sum_s \Re D_{11s}^2 (\mathfrak{D}\mathfrak{D}_0)^{-\frac{1}{2}} \times \exp\left[ik_1(\mathfrak{D} + \mathfrak{D}_0) + i\nu_s\omega_{0\pm}\right], \quad r > a, \quad 0 \le |\Omega| \le 2\pi. \tag{1.38}$$

The physical meaning of $\mathfrak{D}$, $\mathfrak{D}_0$, $\cos^{-1}(a/r)$, $\cos^{-1}(a/r_0)$, $\omega_{0\pm}$, etc. are given in Fig. 3. Total

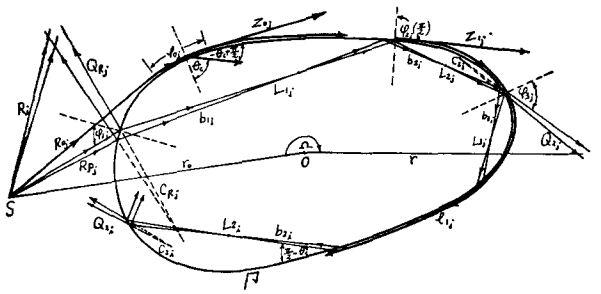

FIG. 5. For the case of $N > 1$, the reactions of a particular tube of rays are shown, and all the notations for the angles and distances are given.

μ^{D1} and μ^{D2} can also be obtained by summing all the possible n's.

From the comparison of the solutions of this part and that of Part II, not only have we verified our postulates but also determined the coefficients as

$$R_{22} = +1, \quad \text{for} \quad N > 1, \tag{1.39}$$

$$D_{12} = (\alpha/\beta)(N^2 - 1)^{-\frac{1}{2}}, \quad \text{for} \quad N > 1, \tag{1.40}$$

$$D_{21} = 2, \quad \text{for} \quad N > 1, \tag{1.41}$$

$$D_{11s}^2 = \left(\frac{2\pi}{k_1}\right)^{\frac{1}{2}} \pi(6)^{-\frac{1}{2}} e^{i(5\pi/12)} (k_1 a)^{-\frac{1}{2}} \frac{\alpha^2}{\beta^2} (N^2 - 1)^{-1} \times [A(q_s)]^{-2}, \quad \text{for} \quad N > 1 \tag{1.42}$$

where $A(q)$ is the Airy function (Appendix I),

$$q_s = (6/k_1 a)^{\frac{1}{3}} e^{-i\frac{1}{3}\pi}(\nu_s - k_1 a), \tag{1.43}$$

and ν_s is a solution of

$$\alpha \frac{H_\nu^{(1)\prime}(k_1 a)}{H_\nu^{(1)}(k_1 a)} = \beta N \frac{H_\nu^{(2)\prime}(k_2 a)}{H_\nu^{(2)}(k_2 a)}. \tag{1.44}$$

The coefficients for the case $N < 1$ will be given in a sequel to this paper.

B. Solution for a Smooth Transparent Cylinder of Arbitrary Shape

In the introduction, we mentioned that the solution of the problem of diffraction by a smooth transparent cylinder of arbitrary shape is highly dependent upon the geometry of the cylinder; therefore it is very difficult to give any general solution of great detail. However, here we shall prescribe a method for obtaining the general solution if one knows the geometry of the body. Due to the fact of the intermingling between geometric rays and diffracted rays in the case $N > 1$ and complete nonintermingling in the case $N < 1$, we will describe these two situations separately.

1. Case $N > 1$

In general, at any point of the configuration space, we may have at most the following fields on the rays passing through this point:

$$\mu_{i\pm}^G, \ \mu_{r\pm}^G, \ \mu_{r+}^D, \ \mu_{r-}^D, \ \mu_{p+}^{DGD\cdots}, \ \cdots,$$

$$\text{and} \quad \mu_{p-}^{DGD\cdots}, \ \cdots, \qquad p = 1, 2, 3, 4, \cdots,$$

where there are 2^{p+1} possible different combinations of DGD $\cdots$ for μ_p. Here we have used previous notation. μ_i^G and μ_r^G are the fields on incident and external reflected rays, respectively, and they appear in the geometrically lit region only. μ_r^D is the field of the diffracted ray just as in the case of the opaque cylinder. $\mu_p^{DGD\cdots}$'s are the fields of the rays with $(p-1)$ internal bounces. They all have $(p+1)$ superscripts and each superscript is either "G" or "D" which denote the behavior of the ray at the interface as either geometric or diffracted type. The criterion of choosing "G" or "D" is whether $\theta < \theta_c$ or $\theta = \theta_c$, respectively.

Since the expressions with subscripts "+" and "−" are the same, we will omit these subscripts from now on. The expressions for μ_i^G, μ_r^G, and μ_r^D are given below and the geometrical notation is defined in Fig. 5. We have

$$\mu_i^G = \mathfrak{N} R_i^{-\frac{1}{2}} e^{ik_1 R_i}, \tag{1.45}$$

$$\mu_r^G = \sum_i \mathfrak{N}\mathfrak{M}_{Ri} R_{ri}^{-\frac{1}{2}} R_{11}(\varphi_{ri}) \times \exp[ik_1(R_{ri} + Q_{ri})], \tag{1.46}$$

$$\mu_r^D = \sum_j \mathfrak{N}(R_{0j} Z_{0j})^{-\frac{1}{2}} \sum_s \mathcal{S} D_{11}^2, \times \exp[ik_1(R_{0j} + Z_{0j}) + i\nu_s \ell_{0j} a^{-1}], \tag{1.47}$$

where the sub-subscript j denotes a particular ray of this type, and

$$\mathcal{S} = (1 - e^{i\nu_s P})^{-1}. \tag{1.48}$$

As for $\mu_p^{DGD\cdots GD}$, we write

$$\mu_p^{DGD\cdots GD} \equiv \sum_{j,s} \mathfrak{N} U_1^{Dj} U_2^{Gj} U_3^{Dj} \cdots U_p^{Gj} U_{p+1}^{Dj}, \tag{1.49}$$

where the subscript on U denotes the position of this particular U in the product.

The expression of U^{Gj}'s are given as

$$U_1^{Gj} = -(R_{pj})^{-\frac{1}{2}} \mathcal{L}_{1j} T_{12}(\varphi_{1j}) \times \exp[ik_1 R_{pj} + ik_2 L_{1j} - iB_{1j}(\tfrac{1}{2}\pi)], \tag{1.50}$$

$$U_x^{Gj} = -\mathcal{L}_{xj} R_{22}(\varphi_{xj}) \exp[ik_2 L_{xj} - iB_{xj}(\tfrac{1}{2}\pi)], \qquad x = 2, 3, 4, \cdots, p, \tag{1.51}$$

$$U_{p+1}^{Gj} = +\mathfrak{M}_{pj} T_{21}(\varphi_{p+1,j}) \exp[ik_1 Q_{pj}], \tag{1.52}$$

where B_T is an operator defined by

$$B_T = 1, \text{ if } L_T \text{ has a caustic, and} \quad B_T = 0, \text{ if } L_T \text{ has no caustic.} \tag{1.53}$$

Before giving any formula for U^{D_i}, we must understand what the subscript of U^{D_i} means. Since $\mu_p^{DGD\cdots D}$ actually represents a chain of $(p+1)$ interactions, $U_x^{D_i}$ of $U_x^{G_i}$ is just the corresponding description for the xth interaction in the chain. Also we give the properties of ℓ_{x+} and ℓ_{x-} as:

> ℓ_{x+} is the signed total distance traveled by the wave, which is the contribution of the incident wave above the line OS, in the $(x+1)$th interaction region. ℓ_{x-} is just the same as ℓ_{x+}, but is due to the contribution of the incident wave below the line OS. The sign of ℓ_{x+} is defined to be positive if the wave is traveling clockwise, and negative if the wave is traveling counterclockwise. The sign of ℓ_{x-} is defined to be the negative of the sign of ℓ_{x+}. ℓ_x may not necessarily be a continuous single segment. Depending on the geometry of the cylinder, it may be a sum of several separate intervals.

Now we have the following expressions:

$$U_1^{D_i} = -(R_{0_i})^{-\frac{1}{2}}\mathfrak{L}_{1_i}D_{11_s}D_{12} \times \exp[ik_1R_{0_i} + ik_2L_{1_i} + iv_s\ell_{0_i}a^{-1} - iB_{1_i}(\tfrac{1}{2}\pi)], \tag{1.54}$$

$$U_x^{D_i} = -\mathfrak{L}_{x_i}[R_{22}(\varphi=\tfrac{1}{2}\pi) + D_{12}D_{21}] \times \exp[ik_2L_{x_i} + iv_s\ell_{(x-1)_i}a^{-1} - iB_{x_i}(\tfrac{1}{2}\pi)], \quad x = 2, 3, 4, \cdots, p, \tag{1.55}$$

$$U_{p+1}^{D_i} = \mathfrak{S}(Z_{p_i})^{-\frac{1}{2}}D_{11_s}D_{21} \times \exp[ik_1Z_{p_i} + iv_s\ell_{p_i}a^{-1}]. \tag{1.56}$$

When the cylinder boundary is noncircular but contains at least two circular segments of common center such that the diffracted rays from one segment hit the other segment, a special treatment similar to the case of circular cylinder is needed. A detailed discussion is given in Ref. 5.

As we can see from above, all the parameters in question are functions of the frequency of oscillation of the source, the properties of the two media, the geometry of the scatterer, and the positions of the source and of the observer. Hence, if we know all about these, we can in principle construct the complete asymptotic solution of this particular problem.

2. Case of $N < 1$

In this case we may have the fields of the rays of types, $\mu_{i\pm}^{G}$, $\mu_{r\pm}^{G}$, μ_{r+}^{D1}, μ_{r-}^{D1}, μ_{r+}^{D2}, μ_{r-}^{D2}, and $\mu_{p\pm}^{GG\cdots G}$ only.

[5] Y. M. Chen, "Diffraction by a Smooth Transparent Object," dissertation for Ph.D. at New York University (June 1963).

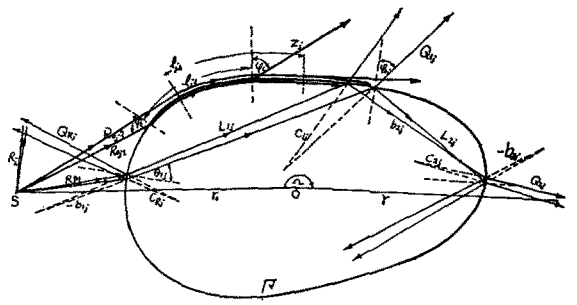

FIG. 6. For the case of $N < 1$, the reactions of a particular tube of rays are shown, and all the notations for the angles and distances are given.

The disappearance of the mixed superscripts in μ_p is due to the fact of nonintermixing between the geometric and the diffracted types of rays. For simplicity, we shall now use all the previous notations, and Fig. 6 serves as an illustration for simple ray tracing.

By omitting the subscripts "+" and "−," we have

$$\mu_i^{G} = \mathfrak{N}(R_i)^{-\frac{1}{2}}\exp[ik_1R_i], \tag{1.57}$$

$$\mu_r^{G} = \sum_j \mathfrak{N}(R_{r_j})^{-\frac{1}{2}}\mathfrak{M}_{r_j}R_{11}(\varphi_{r_j}) \times \exp[ik_1(R_{r_j} + Q_{r_j})], \tag{1.58}$$

$$\mu_r^{D1} = \sum_{j,s} \mathfrak{N}\mathfrak{S}(R_{0_{j_1}}Z_{j_1})^{-\frac{1}{2}}D_{12_s}D_{21_s} \times \exp[ik_1(R_{0_{j_1}} + Z_{j1}) + iv_s\ell_{j1}a^{-1}], \tag{1.59}$$

$$\mu_r^{D2} = \sum_{j,s} \mathfrak{N}\mathfrak{S}D_{11_s}^2(R_{0_{j2}}Z_{j_2})^{-\frac{1}{2}} \times \exp[ik_1(R_{0_{j2}} + Z_{j_2}) + iv_s\ell_{j_2}a^{-1}], \tag{1.60}$$

$$\mu_p^{GG\cdots G} = \sum_j \mathfrak{N}U_1^{G_j}U_2^{G_j}\cdots U_p^{G_j}U_{p+1}^{G_j}, \tag{1.61}$$

where

$$U_1^{G_j} = -\mathfrak{L}_{i_j}(R_{p_j})^{-\frac{1}{2}}T_{12}(\varphi_{1_j}) \times \exp[ik_1R_{p_j} + ik_2L_{1_j} - iB_{1_j}(\tfrac{1}{2}\pi)], \tag{1.62}$$

$$U_x^{G_j} = -\mathfrak{L}_{x_j}R_{22}(\varphi_{x_j})\exp[ik_2L_{x_j} - iB_{x_j}(\tfrac{1}{2}\pi)], \quad x = 2, 3, \cdots, p, \tag{1.63}$$

$$U_{p+1}^{G_j} = \mathfrak{M}_{p_j}T_{21}(\varphi_{p+1_j})\exp[ik_1Q_{p_j}]. \tag{1.64}$$

The necessary and sufficient conditions for constructing the complete asymptotic solution are the same as in the case of $N > 1$.

II. ASYMPTOTIC EXPANSION OF EXACT SOLUTION

A. Introduction

We consider the scattering by a penetrable circular cylinder of the wave from a time-harmonic line

source. As is well known the resulting fields can be expanded explicitly as infinite series of products of functions of the radial and angular variables. The entire difficulty in the study of these fields arises in the attempt to evaluate them when ka is large and to interpret them physically. This was recently done in part by Rubinow[6] for the case of a plane wave incident upon a sphere. For a cylindrical wave incident upon a circular cylinder, evaluation was done in part by Beckmann and Franz.[7] However, they evaluated explicitly only the field which corresponds to transmission with no internal reflection.

We have reconsidered the case of the circular cylinder and have evaluated all the multiply reflected fields. In addition we have evaluated all the diffracted fields, which none of the above authors considered. We have also shown that the asymptotic form of each field—both multiply reflected and diffracted—agrees exactly with the corresponding field constructed by the simpler geometrical method in Part I. We have also found that Rubinow's S-fold substitution is not necessary in the present problem. In addition we have found that the residue evaluation of both Rubinow, and Beckmann and Franz are incorrect.

B. Exact Solution

The field $\mu(r, \Omega)$ produced by a line source at $(r_0, 0)$ in the presence of a cylinder of radius a is defined to be the solution of the following equations:

$$(\nabla^2 + k_1^2)\mu = -\delta(r - r_0)\delta(\Omega)/r, \quad r \geq a, \tag{2.1}$$

$$(\nabla^2 + k_2^2)\mu = 0, \quad r \leq a, \tag{2.2}$$

$$\mu(a+, \Omega) = \alpha\mu(a-, \Omega), \tag{2.3}$$

$$\mu_r(a+, \Omega) = \beta\mu_r(a-, \Omega), \tag{2.4}$$

$$\lim_{r\to\infty} r^{\frac{1}{2}}(\mu_r - ik_1\mu) = 0. \tag{2.5}$$

Here α and β are given constants depending upon the two media.

The unique solution of those equations, readily found by separation of variables, is

$$\mu(r, \Omega) = \tfrac{1}{4}\sum_{\nu=-\infty}^{\infty} e^{i\nu\Omega}(\mathfrak{I}_{10} - P_\nu\mathfrak{I}_{11}), \quad r \geq a, \tag{2.6}$$

where

$$\mathfrak{I}_{10} = H_\nu^{(1)}(k_1 r_>)J_\nu(k_1 r_<), \tag{2.7}$$

$$\mathfrak{I}_{11} = H_\nu^{(1)}(k_1 r_>)H_\nu^{(1)}(k_1 r_<). \tag{2.8}$$

In (2.6) we have used the notation $r_> = \max(r, r_0)$, $r_< = \min(r, r_0)$. The quantity P_ν is defined by

$$P_\nu = \frac{\alpha J_\nu(k_2 a)J_\nu'(k_1 a) - \beta N J_\nu'(k_2 a)J_\nu(k_1 a)}{\alpha J_\nu(k_2 a)H_\nu^{(1)\prime}(k_1 a) - \beta N J_\nu'(k_2 a)H_\nu^{(1)}(k_1 a)}. \tag{2.9}$$

We shall find it useful to rewrite (2.6) by means of the Poisson summation formula as

$$\mu(r, \Omega) = \frac{i}{4}\sum_{n=-\infty}^{\infty}\int_{-\infty}^{\infty} e^{i\nu(\Omega+2\pi n)}(\mathfrak{I}_{10} - P_\nu\mathfrak{I}_{11})\,d\nu, \quad r \geq a, \quad 0 \leq |\Omega| \leq 2\pi. \tag{2.10}$$

To exhibit the physical interpretation of (2.10), we expand P_ν into a geometric series,

$$P_\nu = \frac{1}{2}\left\{1 - \mathfrak{K}_1\left[R_{11} + T_{12}T_{21} \times \sum_{p=1}^{\infty} R_{22}^{p-1}(\mathfrak{K}_2)^p\right]\right\}, \tag{2.11}$$

where

$$R_{11} = -[\alpha \log' H_\nu^{(2)}(k_1 a) - \beta N \log' H_\nu^{(2)}(k_2 a)] \times [\alpha \log' H_\nu^{(1)}(k_1 a) - \beta N \log' H_\nu^{(2)}(k_2 a)]^{-1}, \tag{2.12}$$

$$R_{22} = -[\alpha \log' H_\nu^{(1)}(k_1 a) - \beta N \log' H_\nu^{(1)}(k_2 a)] \times [\alpha \log' H_\nu^{(1)}(k_1 a) - \beta N \log' H_\nu^{(2)}(k_2 a)]^{-1}, \tag{2.13}$$

$$\alpha T_{12} = 1 + R_{11}, \tag{2.14}$$

$$\alpha^{-1}T_{21} = 1 + R_{22}, \tag{2.15}$$

$$\log' H_\nu^{(j)}(z) = H_\nu^{(j)\prime}(z)/H_\nu^{(j)}(z), \quad j = 1, 2, \tag{2.16}$$

$$\mathfrak{K}_1 = H_\nu^{(2)}(k_1 a)/H_\nu^{(1)}(k_1 a), \tag{2.17}$$

$$\mathfrak{K}_2 = H_\nu^{(1)}(k_2 a)/H_\nu^{(2)}(k_2 a). \tag{2.18}$$

By substituting (2.11) into (2.10) we obtain the result

$$\mu(r, \Omega) = \frac{i}{8}\sum_{n=-\infty}^{\infty}\int_{-\infty}^{\infty} d\nu e^{i\nu(\Omega+2\pi n)} \times \left\{2\mathfrak{I}_{10} + [R_{11}\mathfrak{K}_1 - 1]\mathfrak{I}_{11} + T_{12}T_{21}\mathfrak{K}_1 \times \sum_{p=1}^{\infty} R_{22}^{p-1}(\mathfrak{K}_2)^p\mathfrak{I}_{11}\right\}, \quad r \geq a, \quad 0 \leq |\Omega| \leq 2\pi. \tag{2.19a}$$

By changing ν to $-\nu$ in (2.19a) and using the identities $H_{-\nu}^{(1)}(z) = e^{i\nu\pi}H_\nu^{(1)}(z)$ and $H_{-\nu}^{(2)}(z) = e^{-i\nu\pi}H_\nu^{(2)}(z)$, we can write (2.19a) as

$$\mu(r, \Omega) = \frac{i}{8}\sum_{n=-\infty}^{\infty}\int_{-\infty}^{\infty} e^{i\nu[2\pi(1-n)-\Omega]} \times \{\mathfrak{I}_{11} + e^{-i2\nu\pi}\mathfrak{I}_{12} - 2P_{-\nu}\mathfrak{I}_{11}\}\,d\nu, \quad r \geq a, \quad 0 \leq |\Omega| \leq 2\pi, \tag{2.19b}$$

[6] S. I. Rubinow, Ann. Phys. (N. Y.) 14, 305 (1961).
[7] V. P. Beckmann and W. Franz, Z. Naturforsch. 12, 257 (1957).

where

$$\mathfrak{I}_{12} = H_\nu^{(1)}(k_1 r_>)H_\nu^{(2)}(k_1 r_<). \tag{2.20}$$

In order to enterchange the sums of the integrand of (2.19a) and (2.19b) and the integral sign, we have to deform the integration path in such a way that for $\nu > k_1 r_>$, $H_\nu^{(1)}(k_1 r_<)H_\nu^{(1)}(k_1 r_>)$ decays faster than exponential function on the path. This can be accomplished by changing the integration path from the entire real axis to path C. Fig. 7 shows schematically the position of the poles of the integrand of (2.10) as well as the paths of integration C and C_1.

With the restriction of $0 < |\Omega| < 2\pi$, we can easily show that the arc C_0 gives no contribution to the integral by using representation (2.19a) for the terms with $n \geq 0$, and representation (2.19b) for the terms with $n < 0$. Then we can write (2.19a) and (2.19b) as

$$\begin{aligned}\mu(r, \Omega) = \frac{i}{8}\Bigg(&\sum_{n=-\infty}^{-1}\int_C e^{i\nu[2\pi(1-n)-\Omega]} \\ &\times \{\mathfrak{I}_{11} + e^{-i2\pi\nu}\mathfrak{I}_{12} - 2P_{-\nu}\mathfrak{I}_{11}\}\, d\nu \\ &+ \sum_{n=0}^{\infty}\int_C e^{i\nu(\Omega+2\pi n)}\Big\{2\mathfrak{I}_{10} + [R_{11}\mathfrak{C}_1 - 1]\mathfrak{I}_{11} \\ &+ T_{12}T_{21}\mathfrak{C}_1\sum_{p=1}^{\infty}R_{22}^{p-1}\mathfrak{C}_2\mathfrak{I}_{11}\Big\}\, d\nu\Bigg), \\ &r \geq a, \quad 0 < |\Omega| < 2\pi. \end{aligned} \tag{2.21}$$

Since (2.21) is long and complicated, we shall write symbolically $\mu(r, \Omega) = \sum_{n=-\infty}^{\infty}\int_C$ in the rest of the text.

The form (2.11) is introduced because R_{11} and T_{12} are, respectively, the reflection and transmission coefficients for the converging wave $e^{i\nu\Omega}H_\nu^{(2)}(k_1 r)$ incident upon the cylinder surface from outside. Similarly R_{22} and T_{21} are the corresponding coefficients for the diverging wave $e^{i\nu\Omega}H_\nu^{(1)}(k_2 r)$ incident upon the interface from inside. The ratio $H_\nu^{(1)}(k_2 a)/H_\nu^{(2)}(k_2 a)$ gives the phase shift factor $e^{-\frac{1}{2}i\pi}$ when the transmitted wave passes through the axis, which is a caustic. Therefore the various terms in (2.11) have obvious physical meanings which we shall explain further below.

The first term in the integrand in (2.21) represents the incident wave μ_i and the next term represents the wave $\mu_{r,-i}$ externally reflected from the cylinder in lit region and negative of incident wave in the shadow region. The pth term in the sum, which we shall denote by μ_p, represents a wave transmitted into the cylinder, reflected $p = 1$ times internally from the interface, having passed p times through the cylinder, and finally transmitted out into the surrounding medium. This interpretation will be borne out when the various terms are expanded asymptotically for short wavelength and the results compared with the predictions of geometrical optics.

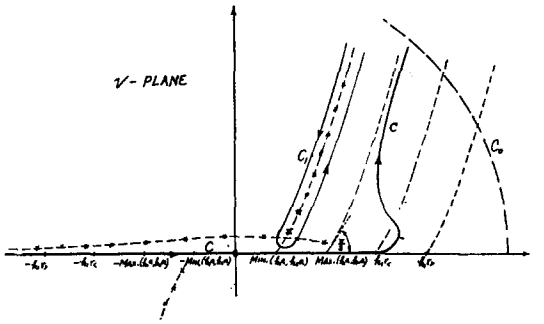

FIG. 7. This figure shows schematically the positions of the poles of the integrand of (2.9) as well as the path of integration C in the ν-plane.

C. Geometric Optics Field

The scattered wave can be decomposed into a geometric optics field, denoted by superscript G, and a diffracted field, denoted by superscript D. The criterion for the proper identification of the geometric optics field from (2.21) is the existence of real saddle points ν_{0_j} ($j = 1, 2, \cdots$) such that $0 < |\nu_{0_j}| < k_1 r_<$.

In the following discussion, we shall discuss the case $N > 1$ in great detail. For simplicity, the k's are restricted to be real. As for the case $N < 1$, when the incident angle $\varphi(r_>, r_<, \Omega)$ is smaller than the angle of total reflection, the solutions have the same forms as in the case $N > 1$, except that the solutions of the saddle-point equations are changed. If φ is larger than or equal to the angle of total reflection, then μ_i^G and μ_r^G have the same form as in the case $N > 1$, but the μ_p^G's vanish.

We will now evaluate each term of (2.21) asymptotically in k_1 and k_2. Since it is found that the main contribution of the integrals comes from the range in which $|\mathrm{Re}\ \nu| < k_1 r_<$, it suffices to use the Debye asymptotic forms, given in Appendix I, for $H_\nu^{(1)}(k_1 r_>)$ and $H_\nu^{(1)}(k_1 r_<)$.

1. Incident and Externally Reflected Fields

The incident field μ_i, as is well known, is given by

$$\mu_i(r, \Omega) \sim \mathfrak{A}(R_i)^{-\frac{1}{2}}e^{ik_1 R_i}. \tag{2.22}$$

The above result can also be obtained by evaluating the integral of the first term in (2.21) asymptotically by the saddle-point method. We proceed by rewriting the first term of (2.21) as

$$\begin{aligned}\mu_i(r, \Omega) \sim \frac{i}{8}\sum_{n=-\infty}^{\infty}\Bigg[&\int_c e^{i\nu(\Omega+2\pi n)}\mathfrak{I}_{11}\, d\nu \\ &+ \int_c e^{i\nu(\Omega+2\pi n)}\mathfrak{I}_{12}\, d\nu\Bigg], \quad r \geq a, \ 0 < |\Omega| < 2\pi. \end{aligned} \tag{2.23}$$

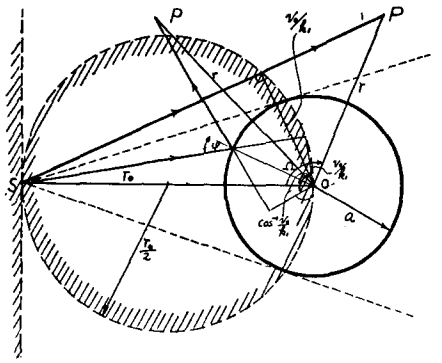

FIG. 8. The regions of validity of Eqs. (2.24) and (2.25) are shown, and the geometrical interpretation of ν_0/k_1 is given here.

After using the proper asymptotic forms in (2.23) and evaluating the resulting integral by the saddle-point method, we obtain Eq. (2.22). We have investigated each term in the infinite series of (2.23) and found that only the term with $n = 0$ has a real saddle point. The corresponding saddle-point equations for the first and second integrals of (2.23), respectively, are

$$\cos^{-1}(\nu_0/k_1 r_>) + \cos^{-1}(\nu_0/k_1 r_<) = \Omega, \qquad (2.24)$$

$$\cos^{-1}(\nu_0/k_1 r_>) - \cos^{-1}(\nu_0/k_1 r_<) = \Omega. \qquad (2.25)$$

It is found that Eq. (2.24) has a real nonzero solution only in the unshaded region, and Eq. (2.25) has a real nonzero solution only in the shaded region of Fig. (2.2). The geometrical interpretation of ν_0/k_1 is given in Fig. 8. The positive solutions of both (2.24) and (2.25) correspond to the field points above the line OS. The negative solutions of those equations correspond to the field points below the line OS.

We now consider

$$\mu_{r,-i}(r, \Omega) \sim \frac{i}{8} \sum_{n=-\infty}^{\infty} \int_c e^{i\nu(\Omega+2\pi n)}[R_{11}\mathfrak{C}_1 - 1]\mathfrak{I}_{11}\, d\nu,$$
$$r \geq a, \quad 0 < |\Omega| < 2\pi. \qquad (2.26)$$

We first notice that the second term in the integrand of (2.26) is the negative of the first term in the integrand of (2.23); hence they cancel each other at all times. After using the correct asymptotic forms in the first term in the integrand of (2.26) and trying to evaluate it by the saddle-point method, we find that there are possible two real saddle points. One falls in the range $0 < |\nu| < k_1 a$, and one falls in the range $k_1 a < |\nu| < k_1 r_<$.

For the case of the saddle point ν_0, such that $k_1 a < |\nu_0| < k_1 r_<$, we have $H_\nu^{(1)}(k_1 a) \sim -H_\nu^{(2)}(k_1 a)$, and then $R_{11}\mathfrak{C}_1 \sim 1$. Therefore this contribution of the first integral of (2.26) gives the incident field only in the intersection of the unshaded area of Fig. 8 and the lit region (excluding the geometric shadow line).

For the case of the saddle point ν_0 which lies in the range of $0 < |\nu_0| < k_1 a$, it gives the contribution to the external reflected field. Here we have used the Debye asymptotic forms for all of the Hankel functions in (2.26). The saddle-point equation of a typical term in the infinite series of (2.26) is

$$\cos^{-1}\frac{\nu_0}{k_1 r_0} + \cos^{-1}\frac{\nu_0}{k_1 r} - 2\cos^{-1}\frac{\nu_0}{k_1 a} = \Omega + 2\pi n. \qquad (2.27)$$

It is found that, only for $n = 0$ and the field point lying in the lit region above the line OS, can (2.27) have a positive real solution, and only for $n = -1$ and the field point lying in the lit region below the line OS can (2.27) have a negative real solution. The geometrical interpretation of (2.27) is given in Fig. 8. The solution of (2.27), if written as a function of the physical angle $\varphi = \varphi(r, r_0, \Omega)$, is

$$\nu_0 = \pm k_1 a \sin\varphi. \qquad (2.28)$$

By substituting (2.28) into (2.27), one obtains the relation between Ω and φ as

$$2\pi - \Omega = 2\varphi - \sin^{-1}[(a/r_0)\sin\varphi] - \sin^{-1}[(a/r)\sin\varphi]. \qquad (2.29)$$

It is remarkable that (2.29) agrees exactly with the relation between Ω and φ obtained by geometric means as equation (1.24) of Sec. A, Part I. Finally, the asymptotic expression for $\mu_R^a(r, \Omega)$ is given by (1.7).

Franz and Beckmann[8] have given an analysis similar to that above.

2. *pth Transmitted Fields*

For the pth transmitted field, we have to evaluate asymptotically the following equation by the saddle-point method:

$$\mu_p(r, \Omega) \cong \frac{i}{8} \sum_{n=-\infty}^{\infty} \int_c e^{i\nu(\Omega+2\pi n)}\mathfrak{I}_{11}T_{12}T_{21}R_{22}^{p-1}\mathfrak{C}_1\mathfrak{C}_2^p\, d\nu,$$
$$0 < |\Omega| < 2\pi, \quad r \geq a, \quad p = 2, 3, 4, \cdots. \qquad (2.30)$$

The saddle-point equation is found to be

$$\cos^{-1}\frac{\nu_0}{k_1 r_0} + \cos^{-1}\frac{\nu_0}{k_1 r} + 2p\cos^{-1}\frac{\nu_0}{Nk_1 a} - 2\cos^{-1}\frac{\nu_0}{k_1 a} = \Omega + 2\pi n. \qquad (2.31)$$

[8] W. Franz and V. P. Beckmann, IRE Trans. Antennas Propagation 4, 203 (1956).

As in the previous cases, the relation between ν_0 and φ is given by (2.28). Also the positive solutions give the contributions of the incident field above the line OS of Fig. 8, and the negative solutions give the contributions of the incident field below the line OS. By substituting (2.28) into (2.31), we have

$$\begin{matrix} 2\pi M_p^+ + \Omega \\ 2\pi M_p^- - \Omega \end{matrix} = 2\varphi + p(\pi - 2\theta) - \sin^{-1}\left(\frac{a}{r}\sin\varphi\right) - \sin^{-1}\left(\frac{a}{r_0}\sin\varphi\right), \quad (2.32)$$

where $\sin\theta = (1/N)\sin\varphi$. Equation (2.32) has solutions only if M_p^+ and M_p^- are properly chosen such that (2.32) is identical with the relation between Ω and φ obtained by geometric construction. Since the "n" in (2.31) is any integer, we find that the s-fold substitution of Rubinow[6] is not necessary in the present problem.

The typical form of contribution from any saddle-point ν_{0s} can be easily derived as (1.10). Beckmann and Franz[7] have given a similar analysis for $p=1$ only.

D. Diffracted Field

Any integral in (2.21) which does not have a real saddle point may still be evaluated by the method of residues. After evaluating them asymptotically, we shall find that all of these integrals give contributions to the diffracted field. We shall discuss the case $N > 1$ in great detail here. As for the case $N < 1$, although the situation is more complicated, the same approach will work. A sequel of this paper will deal with it.

It is easy to see that $\mu_r(r, \Omega)$ has simple poles in the integrand and $\mu_p(r, \Omega)$ has poles of $(p+1)$ order. The positions of the poles of all the integrals are determined by (1.44). The approximate positions of the roots ν_s of (1.44) in the upper complex ν-plane are

$$N > 1, \quad \nu_s = k_1 a + q_s\left(\frac{k_1 a}{6}\right)^{\frac{1}{3}} e^{\frac{1}{3}i\pi} + \cdots, \quad (2.33)$$

where q_s is a number determined by the following equation:

$$N > 1, \quad e^{-\frac{1}{3}i\pi}\left(\frac{6}{k_1 a}\right)^{\frac{1}{3}} \frac{A'(q_s)}{A(q_s)} = i\frac{\beta}{\alpha} N\left[1 - \left(\frac{\nu_s}{k_2 a}\right)^2\right]^{\frac{1}{2}}, \quad (2.34)$$

and $A(z)$ is the Airy function.

For the convenience of physical interpretation, we shall introduce two equivalent representations of the integrals before proceeding to evaluate them:

$$\mu_{p_n}(r, \Omega) = \frac{i}{8}\int_{C_1} e^{i(\Omega + 2\pi n)\nu} \mathfrak{I}_{11} T_{12} T_{21} R_{22}^{p-1} \mathfrak{K}_1 \mathfrak{K}_2^p \, d\nu, \quad 0 < |\Omega| < 2\pi, \quad r \geq a, \quad (2.35)$$

and

$$\mu_{p_n}(r, \Omega) = \frac{i}{8}\int_{C_1} e^{i[(p-n)2\pi - \Omega]\nu} \mathfrak{I}_{11} T_{12} T_{21} R_{22}^{p-1} \mathfrak{K}_1 \mathfrak{K}_2^p \, d\nu, \quad 0 < |\Omega| 2\pi, \quad r \geq a. \quad (2.36)$$

Also,

$$\mu_{r_n}(r, \Omega) = \frac{i}{8}\int_{C_1} e^{+i(\Omega + 2\pi n)\nu} \mathfrak{I}_{11}[R_{11}\mathfrak{K}_1 - 1] \, d\nu, \quad 0 < |\Omega| < 2\pi, \quad r \geq a \quad (2.37)$$

is equivalent to

$$\mu_{r_n}(r, \Omega) = \frac{i}{8}\int_{C_1} e^{-i(\Omega + 2\pi n)\nu} \mathfrak{I}_{11}[R_{11}\mathfrak{K}_1 - e^{i2\pi\nu}] \, d\nu, \quad 0 < |\Omega| < 2\pi, \quad r \geq a. \quad (2.38)$$

The above representations are the same representations of (2.19) which serve as mathematical tricks for closing the integration path in the upper ν-plane (Fig. 7). However, it is more than just a trick, because the criterion of choosing a particular representation has a definite physical meaning which will be seen later.

In evaluating $\mu_{r_n}(r, \Omega)$ we replace $H_\nu^{(1)}(k_1 r_<)$, $H_\nu^{(1)}(k_1 r_>)$, $H_\nu^{(1)}(k_2 a)$, and $H_\nu^{(2)}(k_2 a)$ by their Debye asymptotic forms, and replace $H_\nu^{(1)}(k_1 a)$ and $H_\nu^{(2)}(k_1 a)$ by their Airy function asymptotic forms (Appendix I). After evaluating them by the method of residues and neglecting all the terms of $O[1/k_1 a(N^2-1)^{\frac{1}{2}}]$, we obtain from (2.37) and (2.38), respectively, the following:

$$\mu_{r_n}^{D}(r, \Omega) \cong \sum_s \mathfrak{N} D_{11_s}^2 (\mathfrak{D}\mathfrak{D}_0)^{-\frac{1}{2}} \exp[ik_1(\mathfrak{D} + \mathfrak{D}_0) + i\nu_s\omega_{0+}], \quad r > a, \quad 0 < |\Omega| < 2\pi, \quad N > 1, \quad (2.39)$$

$$^{D}{}_{n}(r, \Omega) \cong \sum_s \mathfrak{N} D_{11_s}^2 (\mathfrak{D}\mathfrak{D}_0)^{-\frac{1}{2}} \exp[ik_1(\mathfrak{D} + \mathfrak{D}_0) + i\nu_s\omega_{0-}], \quad r > a, \quad 0 < |\Omega| < 2\pi, \quad N > 1, \quad (2.40)$$

where $D_{11_s}^2$ is given by (1.42). The criterion of choosing Eq. (2.39) or (2.40) is: For fixed Ω, r, r_0, a, n, (2.39) is used if and only if $\omega_{0+} \geq 0$, and (2.40) is used if and only if $\omega_{0-} \geq 0$. They are mutually exclusive.

In this particular case, the choice is simple. When the field point lies in the geometric shadow region, (2.39) is used for $n = 0, 1, 2, \cdots$, and

(2.40) is used for $n = -1, -2, \cdots$. For the field point situated in the upper lit region, (2.39) is used for $n = 1, 2, \cdots$, and (2.40) is used for $n = -1, -2, -3, \cdots$; for the field point situated in the lower lit region, (2.39) is used for $n = 0, 1, 2, \cdots$, and (2.40) is used for $n = -2, -3, -4, \cdots$. The missing terms in the above give the geometric optics field in different lit regions as had been discussed in Sec. C.

The above criterion is so chosen that (2.39) and (2.40) have direct physical interpretations. (2.39) represents the portion of incident wave which hits the upper surface of the cylinder tangentially. It decays exponentially while traveling clockwise along the interface an angular distance ω_{0+} on the side of medium 1, and then leaves the surface tangentially toward the observing point. Equation (2.40) represents the portion of incident wave which hits the lower surface of the cylinder tangentially. It decays exponentially while traveling counter-clockwise along the interface an angular distance ω_{0-} on the side of medium 1, and then leaves the surface tangentially towards the field point. Therefore, all the criteria do is to make sure that the angular distances are positive in both cases. The physical distances and angles are shown in Fig. 3.

Now we proceed to evaluate (2.35) and (2.36) by the method of residues (see Appendix II). By using the same kind of asymptotic forms for Hankel functions as before and neglecting the terms of $O[1/k_1a(N^2 - 1)^{\frac{1}{2}}]$, we obtain (1.32), (1.33), (1.39), (1.40), (1.41), and (1.42).

APPENDIX I

Debye asymptotic forms for large argument and index, with $\nu < x$, are

$$H_\nu^{(1)}(x) \sim (\tfrac{1}{2}\pi x \sin\tau)^{-\frac{1}{2}} e^{ix(\sin\tau - \tau\cos\tau) - \frac{1}{4}i\pi} \equiv F(x),$$

$$H_\nu^{(2)}(x) \sim (\tfrac{1}{2}\pi x \sin\tau)^{-\frac{1}{2}} e^{-ix(\sin\tau - \tau\cos\tau) + \frac{1}{4}i\pi} \equiv G(x),$$

$$(\partial/\partial x)H_\nu^{(1)}(x) \sim i \sin\tau F(x),$$

$$(\partial/\partial x)H_\nu^{(2)}(x) \sim -i \sin\tau G(x),$$

where $\tau = \cos^{-1}(\nu/x)$. For argument and index both large, and $\nu \simeq x$, we have the Airy function representations for the Hankel functions as

$$H_\nu^{(1)}(x) \sim (2/\pi)yA(q),$$

$$H_\nu^{(2)}(x) \sim (2/\pi)zA(t),$$

$$H_\nu^{(1)\prime}(x) \sim -(2/\pi)y^2A'(q),$$

$$H_\nu^{(2)\prime}(x) \sim -(2/\pi)z^2A'(t),$$

$$(\partial/\partial\nu)H_\nu^{(1)}(x) \sim (2/\pi)y^2A'(q),$$

$$(\partial/\partial\nu)H_\nu^{(2)}(x) \sim (2/\pi)z^2A'(t),$$

$$(\partial/\partial\nu)H_\nu^{(1)\prime}(x) \sim (2/\pi)(y^3/3)qA(q),$$

and

$$(\partial/\partial\nu)H_\nu^{(2)\prime}(x) \sim (2/\pi)(z^3/3)tA(t),$$

where

$$y = (6/x)^{\frac{1}{3}}e^{-\frac{1}{4}i\pi}, \qquad z = (6/x)^{\frac{1}{3}}e^{\frac{1}{4}i\pi},$$

and

$$q = y(\nu - x), \qquad t = z(\nu - x).$$

APPENDIX II

Theorem. Let $g(z)$ be an analytic function of z with simple zeros at z_s and let $f(z)$ be an analytic function of z with no zeros at z_s. Then

$$F(z) = f(z)/[g(z)]^h, \qquad h = 1, 2, 3, \cdots, n$$

has the hth-order poles at z_s, and the residue of $F(z)$ at z_s is

$$R_s = \frac{1}{(h-1)!}\left(\frac{\partial^{h-1}}{\partial z^{h-1}}\left\{\frac{f(z)(z - z_s)^h}{[g(z)]^h}\right\}\right)\Bigg|_{z=z_s}.$$

For convenience, we carry out the differentiation in R_s and obtain the following: for $h = 2$,

$$R_s = \frac{1}{[(\partial/\partial z)g(z)]^2}\left\{\frac{\partial f(z)}{\partial z} - \frac{f(z)[\partial^2 g(z)/\partial z^2]}{\partial g(z)/\partial z}\right\}\Bigg|_{z=z_s},$$

for $h = 3$,

$$R_s = \frac{1}{2\left[\dfrac{\partial g(z)}{\partial z}\right]^3}\left\{\frac{\partial^2 f(z)}{\partial z^2} - \frac{3\,\dfrac{\partial f(z)}{\partial z}\,\dfrac{\partial^2 g(z)}{\partial z^2}}{\dfrac{\partial g(z)}{\partial z}} - \frac{f(z)\,\dfrac{\partial^3 g(z)}{\partial z^3}}{\dfrac{\partial g(z)}{\partial z}} + \frac{3f(z)\left[\dfrac{\partial^2}{\partial z^2}\,g(z)\right]^2}{\left[\dfrac{\partial}{\partial z}\,g(z)\right]^2}\right\}\Bigg|_{z=z_s},$$

for $h = 4$,

$$R_s = \frac{1}{6\left[\dfrac{\partial g(z)}{\partial z}\right]^4}\left\{\frac{\partial^3 f(z)}{\partial z^3} - \frac{6\,\dfrac{\partial^2 f(z)}{\partial z^2}\,\dfrac{\partial^2 g(z)}{\partial z^2}}{\dfrac{\partial g(z)}{\partial z}} - \frac{4\,\dfrac{\partial f(z)}{\partial z}\,\dfrac{\partial^3 g(z)}{\partial^3 z}}{\dfrac{\partial g(z)}{\partial z}} + \frac{15\,\dfrac{\partial f(z)}{\partial z}\left[\dfrac{\partial^2 g(z)}{\partial z^2}\right]^2}{\left[\dfrac{\partial^2}{\partial z^2}\,g(z)\right]^2} - \frac{f(z)\,\dfrac{\partial^4 g(z)}{\partial z^4}}{\dfrac{\partial g(z)}{\partial z}} + 10\,\frac{f(z)\,\dfrac{\partial^2 g(z)}{\partial z^2}\,\dfrac{\partial^3 g(z)}{\partial z^3}}{\left[\dfrac{\partial g(z)}{\partial z}\right]^2} - 15\,\frac{f(z)\left[\dfrac{\partial^2}{\partial z^2}\,g(z)\right]^3}{\left[\dfrac{\partial}{\partial z}\,g(z)\right]^3}\right\}\Bigg|_{z=z_s}.$$

Both Rubinow[6] and Beckmann and Franz[7] have given the residue formula incorrectly for $h > 1$, retaining only the first term in the expanded form of R_s.

Generalizations of the Jost Functions

H. E. Moses
Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts*
(Received 20 January 1964)

The Jost functions have proved valuable in the study of the analytic properties of the scattering phase for the radial Schrödinger equation. In the present paper we shall present an alternative definition of the Jost functions, prove the equivalence of the new definition to the usual one, and generalize the new definition to the one-dimensional Schrödinger equation ($-\infty < x < \infty$), the three-dimensional nonseparated Schrödinger equation, and the three-dimensional nonseparated Dirac equation. It is hoped that these generalizations lead to a better understanding of the analytic properties of the scattering operator for these and related dynamical systems. The generalized Jost functions are shown to be operators in the variables which label the degeneracy of the continuous spectrum of the Hamiltonians which are considered.

1. INTRODUCTION

LET us consider the radial equation for zero angular momentum.

$$(-d^2/dr^2)\psi(r \mid k) + V(r)\psi(r \mid k) = k^2\psi(r \mid k). \quad (1.1)$$

For uses in the quantum mechanical theory of scattering one wishes solutions of (1.1) which satisfy the boundary conditions

$$\psi(0 \mid k) = 0. \quad (1.2)$$

Such solutions constitute the eigenfunctions of the operator $-d^2/dr^2 + V(r)$ corresponding to the eigenvalue k^2.

Usually a second boundary condition is given to specify the solutions of (1.1) uniquely, namely $(d/dr)\ \psi(0 \mid k)$ is prescribed—usually to be 1. We shall find it convenient to make another choice for this boundary condition later.

If $V(r)$ dies down sufficiently rapidly as r becomes large, the solutions of (1.1) which satisfy (1.2) will have the asymptotic form

$$\lim_{r\to\infty} \psi(r \mid k) = A(k) \sin [kr + \delta(k)], \quad (1.3)$$

where $A(k)$ is a generally complex function of k which depends upon our choice of boundary condition for the derivative of $\psi(r \mid k)$ with respect to r at $r = 0$, and $\delta(k)$ is a real function of k called the phase shift.

The scattering operator $S(k)$ is defined in terms of $\delta(k)$ by

$$S(k) = e^{2i\delta(k)}.$$

The Jost functions were introduced in the following way[1]: We look for solutions of (1.1) which we shall denote $f(k, r)$ which satisfy instead of the boundary condition (1.2), the condition

$$\lim_{r\to\infty} e^{ikr} f(k, r) = 1. \quad (1.4)$$

The Jost function $f(k)$ is then defined by

$$f(k) = f(k, 0). \quad (1.5)$$

It is shown in Ref. 1 that

$$S(k) = f(k)/f(-k) \quad (1.6)$$

and that

$$f(-k) = f^*(k). \quad (1.7)$$

The analytic properties of the function $f(k)$ thus determines the analytic properties of the scattering operator $S(k)$.

Jost functions were defined for scattering operators associated with higher angular momenta in Ref. 2.

The Jost functions have been of great importance in obtaining the scattering potential from the scattering phase. Generalizations of the Jost functions have been given in this connection for other radial equations or systems of radial equations.[3] A summary of such generalizations and a bibliography are given in Ref. 4. These generalizations are all characterized by the introduction of functions analogous to $f(k, r)$.

In the present paper we propose another definition of the Jost functions which can be applied to more general dynamical systems in which the interaction vanishes for large distances. We shall make our definition in terms of the time-dependent description of scattering. When applied to the radial equation

* Operated with support from the U. S. Advanced Research Projects Agency.

[1] R. Jost, Helv. Phys. Acta 20, 256 (1947).

[2] R. Jost and W. Kohn, Kgl. Danske Videnskab. Selskab, Mat. Fys. Medd. 27, No. 9 (1953).

[3] R. G. Newton and R. Jost, Nuovo Cimento 1, 590 (1955).

[4] L. D. Faddeyev (transl. by B. Seckler), J. Math. Phys. 4, 72 (1963).

(1.1), we will be led to the same results obtained previously.

The new definition will enable us to give the generalization of the Jost function to the case of the one-dimensional Schrödinger equation, $(-\infty < x < \infty)$ the three-dimensional nonseparated Schrödinger equation and the three-dimensional nonseparated Dirac equation. In these cases our definition obviates the need to introduce functions of the form $f(k, r)$, and hence the need to introduce phases or separation of variables to obtain a radial coordinate.

It should be mentioned that the author, together with Kay, already obtained the generalization of the Jost functions for the one-dimensional and three-dimensional Schrödinger equation in Ref. 5 in connection with the problem of obtaining scattering potentials from the scattering operator. However, the connection with the original Jost definition was not shown, and also the emphasis upon particular inverse scattering problems in Ref. 5 obscures the generalization of the Jost function. The case of the unseparated three-dimensional Dirac equation which is given in the present paper has not been published previously.

We now give the general definition of Jost functions.

Let us consider a Hamiltonian operator

$$H = H_0 + V, \tag{1.8}$$

where H_0 is the unperturbed Hamiltonian which we shall assume has a purely continuous spectrum. We shall assume that the spectrum of H has a continuous spectrum which coincides with that of H_0 and possible point eigenvalues.[6]

Let us denote the eigenfunctions of H_0 by $\psi_0(x \mid E, \alpha)$ where α denotes the degeneracy of the spectrum of H_0. It could, for example, denote the direction of momentum in the three-dimensional cases and includes other variables. The variable x is used to denote collectively all the variables needed to specify coordinate representation. Then

$$H_0\psi_0(x \mid E, \alpha) = E\psi_0(x \mid E, \alpha). \tag{1.9}$$

The eigenfunctions of H which correspond to the continuous spectrum are not unique. There are, however, two sets which are useful for solving the time-dependent Schrödinger equation. We shall denote them by $\psi_-(x \mid E, \alpha)$ and $\psi_+(x \mid E, \alpha)$. These functions are defined by

$$H\psi_\pm(x \mid E, \alpha) = E\psi_\pm(x \mid E, \alpha), \tag{1.9a}$$

$$\lim_{t\to-\infty} e^{iH_0t}e^{-iHt} \int d\alpha \int dE\psi_-(x \mid E, \alpha)f(E, \alpha)$$

$$= \int d\alpha \int dE\psi_0(x \mid E, \alpha)f(E, \alpha), \tag{1.10}$$

$$\lim_{t\to+\infty} e^{iH_0t}e^{-iHt} \int d\alpha \int dE\psi_+(x \mid E, \alpha)f(E, \alpha)$$

$$= \int d\alpha \int dE\psi_0(x \mid E, \alpha)f(E, \alpha), \tag{1.11}$$

where $f(E, \alpha)$ is any quadratically integrable function of its arguments. (The integration over α is to be replaced by summation for discrete values of α.)

Hence $\psi_\pm$ are obtained from the solution of initial- or final-value problems.[7] Still another way of specifying the eigenfunctions of H corresponding to the continuous spectrum is to use boundary conditions on the eigenfunctions.

We shall assume, as is usually done, that the perturbation V is a function of a coordinates such that, when some become sufficiently large, V approaches zero. In the simplest cases these coordinates are the entire set of coordinates as in ordinary potential scattering. But in more complicated cases such as the scattering of an electron by an atom, only the coordinates which enter the electron–atom interaction are considered. We shall call such coordinates the interaction coordinates.

We wish now to define the set of eigenfunctions $\psi(x \mid E, \alpha)$ by the boundary condition

$$\lim_{|x|\to\infty} \int d\alpha \int dE\psi(x \mid E, \alpha)f(E, \alpha)$$

$$= \lim_{|x|\to\infty} \int d\alpha \int dE\psi_0(x \mid E, \alpha)f(E, \alpha), \tag{1.12}$$

and, of course, also satisfy

$$H\psi(x \mid E, \alpha) = E\psi(x \mid E, \alpha). \tag{1.9b}$$

The limit in (1.12) is the limit as the interaction coordinates approach infinity in a range of directions sufficiently large so that (1.12) determines all $\psi(x \mid E, \alpha)$ for all E and α.

One may ask the question whether a solution of the differential equation (1.9b) really can be found

[5] I. Kay and H. E. Moses, Part I, Nuovo Cimento **2**, 917 (1955); Part II, *ibid.* **3**, 66 (1956); Part III, *ibid.* **3**, 276; Part IV, *ibid.* Suppl. **5**, 230 (1957); Part V, *ibid.* **22**, 689 (1961).

[6] These restrictions on the spectra of the Hamiltonians are usual for nonrelativistic problems, but are somewhat severe in the case of relativistic problems where the energy gap may be changed by the perturbation. These restrictions may be relaxed but they lead to a somewhat more complicated formalism which we wish to avoid in this exposition.

[7] In the usual treatments of scattering theory, the minus and plus subscripts on $\psi_\pm$ are reversed in the defining equations. We adhere to the conventions of Ref. 5.

which satisfied the boundary condition (1.12). We shall show that this is possible for "realistic" problems, i.e., those in which we have not separated variables. For the radial equation, condition (1.12) cannot be used to give a solution of (1.9b) because the eigenfunction ψ_0 and ψ both must be zero for $r = 0$. Since (1.3) holds for *all* solutions of the radial equation, we see that this requirement is too severe in general.

Hence, for the radial equation we replace the boundary condition (1.12) by another boundary condition, namely that as $r \to 0$, $\int dE\ \psi(r \mid E)f(E)$ approaches in value $\int dE\ \psi_0(r \mid E)f(E)$ as closely as possible. (α does not appear because the radial equation has no degeneracy. We have replaced the coordinate variable x by r.) This requirement is equivalent to the requirement

$$\begin{aligned} \psi(0 \mid E) &= \psi_0(0 \mid E) = 0, \\ \psi'(0 \mid E) &= \psi_0'(0 \mid E), \end{aligned} \tag{1.12a}$$

where the prime means derivative with respect to r. One might wish to have all derivatives of the perturbed and unperturbed eigenfunctions to equal to each other at $r = 0$. But this requirement can be fulfilled only if $V(r) = 0$, and thus (1.12a) is as severe as we can make our conditions generally.

Thus we have defined the sets of eigenfunctions $\psi_-(x \mid E, \alpha)$, $\psi_+(x \mid E, \alpha)$ and $\psi(x \mid E, \alpha)$. Since each set of eigenfunctions spans the same portion of Hilbert space, each eigenfunction must be a linear combination of the others over the degeneracy variable α. We define the integral operators (which will be matrices if α takes on discrete values)

$$\begin{aligned} \psi(x \mid E, \alpha) &= \int d\alpha' \psi_-(x \mid E, \alpha')\mu_{-E}(\alpha' \mid \alpha) \\ &= \int d\alpha' \psi_+(x \mid E, \alpha')\mu_{+E}(\alpha' \mid \alpha). \end{aligned} \tag{1.13}$$

We maintain that the kernel $\mu_{+E}(\alpha \mid \alpha')$ is the generalization of the Jost function $f(k)$ and $\mu_{-E}(\alpha \mid \alpha')$ is the generalization of $f(-k)$. Indeed, we show in Sec. 3 that, for the radial equation, $\mu_{+E} = f(E^{\frac{1}{2}})$.

2. SOME FORMULAS FROM SCATTERING THEORY

We find some results from the formal theory of scattering useful. We are essentially summarizing the formulas in Part I of Ref. 5.

The scattering operator $S_E(\alpha \mid \alpha')$ is defined by

$$\begin{aligned} \lim_{t \to +\infty} e^{iH_0 t}e^{-iHt} \int d\alpha \int dE \psi_-(x \mid E, \alpha)f(E, \alpha) \\ = \int d\alpha \int d\alpha' \int dE \psi_0(x \mid E, \alpha)S_E(\alpha \mid \alpha')f(E, \alpha'). \end{aligned} \tag{2.1}$$

The inverse of the scattering operator will be denoted by $S_E^{-1}(\alpha \mid \alpha')$. It satisfied the relation

$$\begin{aligned} \int d\alpha' S_E^{-1}(\alpha \mid \alpha')S_E(\alpha' \mid \alpha'') &= \delta(\alpha, \alpha'') \\ &= \int d\alpha' S_E(\alpha \mid \alpha')S_E^{-1}(\alpha' \mid \alpha''), \end{aligned} \tag{2.2}$$

where $\delta(\alpha, \alpha')$ is a Dirac or Kronecker δ function, depending upon whether α is continuous or discrete.

Since the scattering operator is unitary, it satisfied the relation

$$S_E^{-1}(\alpha \mid \alpha') = S^*(\alpha' \mid \alpha), \tag{2.3}$$

where the asterisk means complex conjugate.

The inverse of the scattering operator can also be obtained from the asymptotic time limit

$$\begin{aligned} \lim_{t \to -\infty} e^{iH_0 t}e^{-iHt} \int d\alpha \int dE \psi_+(x \mid E, \alpha)f(E, \alpha) \\ = \int d\alpha \int d\alpha' \int dE \psi_0(x \mid E, \alpha)S_E^{-1}(\alpha \mid \alpha')f(E, \alpha'). \end{aligned} \tag{2.4}$$

The scattering operator and its inverse are related to ψ_- and ψ_+ as follows:

$$\begin{aligned} S_E(\alpha \mid \alpha') &= \delta(\alpha, \alpha') \\ &- 2\pi i \int dx \psi_0^*(x \mid E, \alpha)V(x)\psi_-(x \mid E, \alpha'), \end{aligned} \tag{2.4a}$$

$$\begin{aligned} S_E^{-1}(\alpha \mid \alpha') &= \delta(\alpha, \alpha') \\ &+ 2\pi i \int dx \psi_0^*(x \mid E, \alpha)V(x)\psi_+(x \mid E, \alpha'). \end{aligned} \tag{2.4b}$$

In (2.4a) and (2.4b) the integration over x is to represent integration over all the variables in coordinate space and includes summation over spinor indices in the relativistic case. Also, $V(x)$ in the relativistic case is a matrix in the spinor indices.

Another expression for the scattering operator is

$$S_E(\alpha \mid \alpha') = \int dx \psi_+^*(x \mid E, \alpha)\psi_-(x \mid E, \alpha'). \tag{2.5}$$

From (2.5) and (1.13) we obtain

$$S_E(\alpha \mid \alpha') = \int d\alpha'' \mu_{+E}(\alpha \mid \alpha'')\mu_{-E}^{-1}(\alpha'' \mid \alpha'), \tag{2.6}$$

where μ_{-E}^{-1} is the kernel of the inverse of the operator whose kernel is μ_{-E}.

Equation (2.6) is the generalization of Jost's relation (1.6).

The eigenfunctions ψ_- and ψ_+ satisfy integral equations of the form

$$\psi_-(x \mid E, \alpha) = \psi_0(x \mid E, \alpha) + \int dx' G_{-E}(x \mid x') V(x') \psi_-(x' \mid (e, \alpha), \quad (2.7a)$$

$$\psi_+(x \mid E, \alpha) = \psi_0(x \mid E, \alpha) + \int dx' G_{+E}(x \mid x') V(x') \psi_+(x' \mid E, \alpha), \quad (2.7b)$$

where G_{-E} and G_{+E} are Green's functions which are given by the following:

$$G_{-E}(x \mid x') = \lim_{\epsilon \to 0} \int d\alpha' \int dE' \psi_0(x \mid E', \alpha') \times \frac{1}{E - E' + i\epsilon} \psi_0^*(x' \mid E', \alpha'), \quad (2.8a)$$

$$G_{+E}(x \mid x') = \lim_{\epsilon \to 0} \int d\alpha' \int dE' \psi_0(x \mid E', \alpha') \times \frac{1}{E - E' - i\epsilon} \psi_0^*(x' \mid E', \alpha'). \quad (2.8b)$$

3. THE RADIAL EQUATION FOR ZERO ANGULAR MOMENTUM. IDENTIFICATION OF THE JOST FUNCTION

In the present section we show that in the case of the radial equation for zero angular momentum the function μ_{+E} is identical to the Jost function $f(k)$ where $k = E^{\frac{1}{2}}$. A similar proof can be used to show $\mu_{-E} = f(-k)$.

The eigenstates $\psi_0(r \mid E)$ are given by

$$\psi_0(r \mid E) = (2/\pi)^{\frac{1}{2}} \sin (E^{\frac{1}{2}} r). \quad (3.1)$$

These eigenfunctions span the Hilbert space of function $\varphi(r)$ such that $\varphi(0) = 0$ and they satisfy the Schrödinger equation

$$H_0 \psi_0(r \mid E) = E \psi_0(r \mid E) \qquad (0 < E < \infty), \quad (3.2)$$

where

$$H_0 = -d^2/dr^2. \quad (3.2a)$$

The eigenfunctions $\psi_+(r \mid E)$ of H where

$$H = H_0 + V(r)$$

are given by

$$\psi_+(r \mid E) = \psi_0(r \mid E) + \int_0^\infty dr' G_{+E}(r \mid r') V(r') \psi_+(r' \mid E), \quad (3.3)$$

where

$$G_{+E}(r \mid r') = -(1/k)[\eta(r - r') \sin kr' e^{-ikr} + \eta(r' - r) e^{-ikr'} \sin kr]. \quad (3.4)$$

In (3.4),

$$k = E^{\frac{1}{2}},$$

and $\eta(x)$ is the Heaviside function

$$\eta(x) = 1 \text{ for } x > 0, \quad \eta(x) = 0 \text{ for } x < 0. \quad (3.4a)$$

Now μ_{+E} is defined by

$$\psi(r \mid E) = \mu_{+E} \psi_+(r \mid E), \quad (3.5)$$

where $\psi(r \mid E)$ satisfies the boundary conditions (1.12a). On letting $r = 0$ in (3.5), the first of the boundary conditions (1.12a) leads to the trivial identity $0 = 0$. However, when one takes derivatives with respect to r in (3.5) and uses the second boundary condition of (1.12a) one obtains the following expression for μ_+:

$$\mu_{+E} = \frac{(2/\pi)^{\frac{1}{2}} k}{(2/\pi)^{\frac{1}{2}} k - \int_0^\infty dr' e^{-ikr'} V(r') \psi_+(r' \mid E)}. \quad (3.6)$$

Now let us consider the solution $g(k, r)$ of the following integral equation:

$$g(k, r) = \frac{1}{(2\pi)^{\frac{1}{2}}} e^{-ikr} + \int_0^\infty dr' G_{+E}(r \mid r') V(r') g(k, r'), \quad (k = E^{\frac{1}{2}}). \quad (3.7)$$

It can be shown that $g(k, r)$ satisfies Eq. (1.1) with the boundary conditions

$$\lim_{r \to \infty} e^{ikr} g(k, r) = \frac{1}{(2\pi)^{\frac{1}{2}}} - \frac{1}{k} \int_0^\infty dr' \times \sin kr' V(r') g(k, r'), \quad (3.8a)$$

$$g(k, 0) = 1/(2\pi)^{\frac{1}{2}}. \quad (3.8b)$$

From (3.8a) and (1.4) it is seen that the Jost function $f(k, r)$ is given by

$$f(k, r) = g(k, r) \times \left[\frac{1}{(2\pi)^{\frac{1}{2}}} - \frac{1}{k} \int_0^\infty \sin kr' V(r') g(k, r')\, dr'\right]^{-1}. \quad (3.9)$$

Hence, since $f(k) = f(k, 0)$, we have as an expression for the Jost function on using (3.8b)

$$f(k) = \left(\frac{2}{\pi}\right)^{\frac{1}{2}} k \left[\left(\frac{2}{\pi}\right)^{\frac{1}{2}} k - 2 \int_0^\infty dr' \times \sin kr' V(r') g(k, r')\right]^{-1}. \quad (3.10)$$

We shall now show that

$$2\int_0^\infty dr V(r)\sin krg(k,r) = \int_0^\infty dr V(r)e^{-ikr}\psi_+(r \mid E). \quad (3.11)$$

From this equation it follows from (3.10) and (3.6) that $\mu_{+E} = f(k)$, which we wished to prove.

The proof makes use of a reciprocity theorem of the type used in variational calculations for scattering operators. Consider the equations

$$a = Ky, \qquad a' = Ky', \quad (3.12)$$

where a and a' are known vectors, y and y' are unknown vectors, and K is a symmetric operator. One can easily prove that

$$(a, y') = (a', y), \quad (3.13)$$

where (a, b) denotes the real inner product of two vectors. We shall rewrite Eqs. (3.3) and (3.7) in a form resembling (3.12):

$$\left(\frac{2}{\pi}\right)^{\frac{1}{2}} V(r)\sin kr = \int_0^\infty dr'[V(r)\delta(r-r') + V(r)G_{+E}(r \mid r')V(r')]\psi_+(r' \mid E), \quad (3.3a)$$

$$\frac{1}{(2\pi)^{\frac{1}{2}}} V(r)e^{-ikr} = \int_0^\infty dr'[V(r)\delta(r-r') + V(r)G_{+E}(r \mid r')V(r')]g(k,r'). \quad (3.7a)$$

Let us make the identifications

$$a \to (2/\pi)^{\frac{1}{2}}V(r)\sin kr, \qquad a' \to [1/(2\pi)^{\frac{1}{2}}]V(r)e^{-ikr},$$
$$y \to \psi_+(r \mid E), \qquad y' \to g(k,r),$$
$$K \to [V(r)\delta(r-r') + V(r)G_{+E}(r \mid r')V(r')].$$

Then (3.11) is a consequence of (3.13).

4. THE ONE-DIMENSIONAL SCHRÖDINGER EQUATION

We now give the functions $\mu_{+E}(\alpha \mid \alpha')$ and $\mu_{-E}(\alpha \mid \alpha')$ for the one-dimensional Schrödinger equation. The Hamiltonian has the form

$$H = H_0 + V(x), \quad (4.1a)$$

where

$$H_0 = -d^2/dx^2 \qquad (-\infty < x < \infty). \quad (4.1b)$$

Unlike the radial equation, the Schrödinger equation has a degenerate continuous spectrum. Despite this complication, the treatment of the one-dimensional problem has certain simplicities, because, rather than having to use boundary conditions at the origin, we use boundary conditions for $|x| \to \infty$.

The eigenfunctions of H_0 are given by

$$\psi_0(x \mid E, \alpha) = [(E)^{-\frac{1}{4}}/2\pi^{\frac{1}{2}}]e^{i\alpha E^{\frac{1}{2}}x}, \quad (4.2)$$

where α represents the direction of the momentum and can equal $+1$ or -1. The value of the eigenvalue E is in the range $0 < E < \infty$. These eigenfunctions satisfy the orthonormality and completeness relations

$$\int_{-\infty}^{+\infty} dx\psi_0^*(x \mid E, \alpha)\psi_0(x \mid E', \alpha') = \delta(E-E')\delta(\alpha, \alpha'), \quad (4.3a)$$

$$\sum_\alpha \int_0^\infty dE\psi_0^*(x \mid E, \alpha)\psi_0(x' \mid E, \alpha) = \delta(x-x'). \quad (4.3b)$$

The eigenfunctions of H which we have denoted by ψ_+ and ψ_- satisfy the integral equations (2.7b) and (2.7a), respectively, where the range of integration of x' is from $-\infty$ to $+\infty$. The Green's functions $G_{+E}(x \mid x')$ and $G_{-E}(x \mid x')$ are given by

$$G_{+E}(x \mid x') = (i/2k)e^{-ik|x-x'|}, \quad (4.4a)$$

$$G_{-E}(x \mid x') = -(i/2k)e^{ik|x-x'|}, \quad (4.4b)$$

where $k = E^{\frac{1}{2}}$.

From (2.7a) and (2.7b), the explicit forms of G_{+E} and G_{-E}, and (2.4a) and (2.4b) we can show that

$$\lim_{x\to\pm\infty} \psi_-(x \mid E, \alpha) = \psi_0(x \mid E, \alpha) + \psi_0(x \mid E, \pm 1)[S_E(\pm 1 \mid \alpha) - \delta(\pm 1, \alpha)], \quad (4.5a)$$

$$\lim_{x\to\pm\infty} \psi_+(x \mid E, \alpha) = \psi_0(x \mid E, \alpha) + \psi_0(x \mid E, \mp 1)[S_E^{-1}(\mp 1 \mid \alpha) - \delta(\mp 1, \alpha)]. \quad (4.5b)$$

The eigenfunction of H which we denote by $\psi(x \mid E, \alpha)$ will be required to satisfy the boundary condition

$$\lim_{x\to-\infty} \psi(x \mid E, \alpha) = \psi_0(x \mid E, \alpha). \quad (4.6)$$

If Eq. (4.6) leads to a unique function $\psi(x \mid E, \alpha)$, then the boundary condition (1.12) is also satisfied, which is our general requirement for defining $\psi(x \mid E, \alpha)$.

For convenience we shall find μ_{-E}^{-1} and μ_{+E}^{-1} rather than μ_{-E} and μ_{+E} themselves.

Then the first of Eqs. (1.13) leads to

$$\sum_{\alpha'} \psi(x \mid E, \alpha')\mu_{-E}^{-1}(\alpha' \mid \alpha) = \psi_-(x \mid E, \alpha). \quad (4.7)$$

In (4.7) we let x approach $-\infty$ and use (4.5a) and (4.6),

$$\sum_{\alpha'} \psi_0(x \mid E, \alpha')\mu_{-E}^{-1}(\alpha' \mid \alpha) = \psi_0(x \mid E, \alpha)$$
$$+ \psi_0(x \mid E, -1)[S_E(-1 \mid \alpha) - \delta(-1, \alpha)]. \quad (4.8)$$

Since the eigenfunctions $\psi_0(x \mid E, \alpha)$ for different values of α are linearly independent, we obtain

$$\mu_{-E}^{-1}(1 \mid 1) = 1, \quad \mu_{-E}^{-1}(1 \mid -1) = 0,$$
$$\mu_{-E}^{-1}(-1 \mid 1) = S_E(-1 \mid +1), \quad (4.9)$$
$$\mu_{-E}^{-1}(-1 \mid -1) = S_E(-1 \mid -1).$$

We can obtain μ_{+E}^{-1} similarly:

$$\mu_{+E}^{-1}(1 \mid 1) = S_E^{-1}(1 \mid 1), \quad \mu_{+E}^{-1}(-1 \mid 1) = 0,$$
$$\mu_{+E}^{-1}(1 \mid -1) = S_E^{-1}(1 \mid -1), \quad \mu_{+E}^{-1}(-1 \mid -1) = 1. \quad (4.10)$$

Unlike the situation for the radial equation, the kernels $\mu_{\pm E}^{-1}$ are related directly to the matrix elements of the scattering operator and its inverse.

Also unlike the situation for the radial equation is the fact that these kernels are not altogether unique. We might equally well have chosen $\psi(x \mid E, \alpha)$ using the boundary condition $\lim_{x\to+\infty} \psi(x \mid E, \alpha) = \psi_0(x \mid E, \alpha)$. Then the kernels $\mu_{\pm E}^{-1}$ obtained from such a definition would be related to different elements of the scattering matrix.

5. THE THREE-DIMENSIONAL SCHRÖDINGER EQUATION

We now consider the three-dimensional Hamiltonian

$$H = H_0 + V(\mathbf{x}), \quad (5.1a)$$

where $\mathbf{x}$ denotes collectively the three Cartesian coordinates and

$$H_0 = -\nabla^2. \quad (5.1b)$$

We choose as degeneracy variables for the spectrum of the Hamiltonian H_0 the angle θ which is the angle made by the momentum vector with the z axis and φ which is the angle made by the projection of the momentum operator on the x–y plane with the x axis. We take $0 \leq \theta \leq \pi$ and $0 \leq \varphi \leq 2\pi$.

It will also be useful to introduce polar angles for the vector $\mathbf{x}$. The angles λ and σ which give the direction of $\mathbf{x}$ correspond to θ and φ, respectively.

We shall take as the eigenfunctions of H_0

$$\psi_0(x \mid E, \theta, \varphi) = \frac{(E)^{\frac{1}{4}}}{\sqrt{2}} (\sin\theta)^{\frac{1}{2}} \frac{1}{(2\pi)^{\frac{1}{2}}} e^{i\mathbf{p}\cdot\mathbf{x}}, \quad (5.2)$$

where

$$\mathbf{p} = (E)^{\frac{1}{2}}(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta). \quad (5.3)$$

These eigenfunctions satisfy the completeness and orthonormality relations

$$\int d\mathbf{x}\psi_0^*(\mathbf{x} \mid E, \theta, \varphi)\psi_0(\mathbf{x} \mid E', \theta', \varphi')$$
$$= \delta(E - E')\delta(\theta - \theta')\delta(\varphi - \varphi'), \quad (5.4a)$$

$$\int_0^\infty dE \int_0^\pi d\theta \int_0^{2\pi} d\varphi \psi_0^*(\mathbf{x} \mid E, \theta, \varphi)$$
$$\times \psi_0(\mathbf{x}' \mid E, \theta, \varphi) = \delta(\mathbf{x} - \mathbf{x}'). \quad (5.4b)$$

Before proceeding further it is useful to recollect that the limit of a wave packet of plane waves integrated over all momentum directions is a spherical wave as $|\mathbf{x}| \to \infty$.

That is, in the sense of distributions,

$$\lim_{|\mathbf{x}|\to\infty} e^{i\mathbf{p}\cdot\mathbf{x}} = \frac{2i\pi}{|\mathbf{p}|\,|\mathbf{x}|\sin\theta}[-e^{i|\mathbf{p}||\mathbf{x}|}\delta(\theta - \lambda)$$
$$\times \delta(\varphi - \sigma) + e^{-i|\mathbf{p}||\mathbf{x}|}\delta(\theta - \hat{\lambda})(\varphi - \hat{\sigma})], \quad (5.5)$$

where θ and φ are the polar angles of $\mathbf{p}$, λ and σ are the polar angles of $\mathbf{x}$, and $\hat{\lambda}$ and $\hat{\sigma}$ are the polar angles of $-\mathbf{x}$.

This theorem has frequently been used in discussions of propagation of light but does not seem to have been used too often in quantum mechanical discussions. A formal proof of (5.5) is given in Part V of Ref. 5. Equation (5.5) is the principal reason for the use of boundary condition (1.12) in terms of wave packets. In the case of the radial equation and the one-dimensional equation, it was not necessary to assume as much as (1.12) to obtain $\psi(x \mid E, \alpha)$ from the boundary condition.

The integral equations for the eigenfunctions $\psi_\pm(\mathbf{x} \mid E, \theta, \varphi)$ are

$$\psi_\pm(\mathbf{x} \mid E, \theta, \varphi) = \psi_0(\mathbf{x} \mid E, \theta, \varphi)$$
$$+ \int d\mathbf{x}' G_{\pm E}(\mathbf{x} \mid \mathbf{x}')V(\mathbf{x}')\psi_\pm(\mathbf{x}' \mid E, \theta, \varphi). \quad (5.6)$$

Here

$$G_{\pm E}(\mathbf{x} \mid \mathbf{x}') = -\frac{1}{4\pi\,|\mathbf{x} - \mathbf{x}'|} e^{\mp i|\mathbf{p}||\mathbf{x}-\mathbf{x}'|}, \quad (5.7)$$

where $|\mathbf{p}| = E^{\frac{1}{2}}$.

On using (2.4), (5.2), (5.5), (5.6), and (5.7), we obtain

$$\lim_{|\mathbf{x}|\to\infty} \psi_-(\mathbf{x} \mid E, \theta, \varphi) = \frac{i}{2\pi^{\frac{1}{2}}E^{\frac{1}{4}}\,|\mathbf{x}|}$$
$$\times \left\{\frac{1}{(\sin\theta)^{\frac{1}{2}}}[-e^{i|\mathbf{p}||\mathbf{x}|}\delta(\theta - \lambda)\delta(\varphi - \sigma)\right.$$

$$+ e^{-i|p||x|}\delta(\theta - \hat{\lambda})\delta(\varphi - \hat{\sigma})] - \frac{1}{(\sin\lambda)^{\frac{1}{2}}}$$

$$\times e^{i|p||x|}[S_E(\lambda, \sigma \mid \theta, \varphi) - \delta(\lambda - \theta)\delta(\sigma - \varphi)]\Big\}, \tag{5.8}$$

$$\lim_{|x|\to\infty} \psi_+(\mathbf{x} \mid E, \theta, \varphi) = \frac{i}{2\pi^{\frac{1}{2}}E^{\frac{1}{2}}|\mathbf{x}|}$$

$$\times \Big\{\frac{1}{(\sin\theta)^{\frac{1}{2}}}[-e^{i|p||x|}\delta(\theta - \lambda)\delta(\varphi - \sigma)$$

$$+ e^{-i|p||x|}\delta(\theta - \hat{\lambda})\delta(\varphi - \hat{\sigma})] + \frac{1}{(\sin\hat{\lambda})^{\frac{1}{2}}}$$

$$\times e^{-i|p||x|}[S_E^{-1}(\lambda, \sigma \mid \theta, \varphi) - \delta(\lambda - \sigma)\delta(\sigma - \varphi)]\Big\}. \tag{5.9}$$

Then, on writing

$$\int_0^{\pi} d\theta' \int_0^{2\pi} d\varphi' \psi(\mathbf{x} \mid E, \theta', \varphi')\mu_{-E}^{-1}$$

$$\times (\theta', \varphi' \mid \theta, \varphi) = \psi_-(\mathbf{x} \mid E, \theta, \varphi), \tag{5.10}$$

we let $|\mathbf{x}| \to \infty$. For definiteness we restrict θ to the range $0 \leq \theta \leq \frac{1}{2}\pi$. This range of values will be sufficient to determine μ_{-E}^{-1} uniquely.

From (1.12), (5.2), (5.5), and (5.8), we obtain (after cancelling some common factors)

$$-e^{i|p||x|}\mu_{-E}^{-1}(\lambda, \sigma \mid \theta, \varphi)\frac{1}{(\sin\lambda)^{\frac{1}{2}}}$$

$$+ e^{-i|p||x|}\mu_{-E}^{-1}(\lambda, \sigma \mid \theta, \varphi)\frac{1}{(\sin\hat{\lambda})^{\frac{1}{2}}}$$

$$= -e^{i|p||x|}S_E(\lambda, \sigma \mid \theta, \varphi)\frac{1}{(\sin\lambda)^{\frac{1}{2}}}$$

$$+ e^{-i|p||x|}\delta(\theta - \hat{\lambda})\delta(\varphi - \hat{\sigma})\frac{1}{(\sin\hat{\lambda})^{\frac{1}{2}}}. \tag{5.11}$$

Since $e^{i|p||x|}$ and $e^{-i|p||x|}$ are linearly independent we have

$$\mu_{-E}^{-1}(\lambda, \sigma \mid \theta, \varphi) = S_E(\lambda, \sigma \mid \theta, \varphi),$$

$$\mu_{-E}^{-1}(\hat{\lambda}, \hat{\sigma} \mid \theta, \varphi) = \delta(\hat{\lambda} - \theta)\delta(\hat{\sigma} - \varphi). \tag{5.12}$$

Since λ lies in the range $0 \leq \lambda \leq \frac{1}{2}\pi$, $\hat{\lambda}$ is in the range $\frac{1}{2}\pi < \hat{\lambda} \leq \pi$. We can thus write

$$\mu_{-E}^{-1}(\theta, \varphi \mid \theta', \varphi') = \eta(\tfrac{1}{2}\pi - \theta)S_E(\theta, \varphi \mid \theta', \varphi')$$

$$+ \eta(\theta - \tfrac{1}{2}\pi)\delta(\theta - \theta')\delta(\varphi - \varphi'), \tag{5.13}$$

where $\eta(x)$ is the Heaviside function defined by

$$\eta(x) = 1 \quad \text{for} \quad x > 0,$$

$$= 0 \quad \text{for} \quad x < 0,$$

In a similar way we can obtain μ_{+E}^{-1}:

$$\mu_{+E}^{-1}(\theta, \varphi \mid \theta', \varphi') = \eta(\tfrac{1}{2}\pi - \theta)\delta(\theta - \theta')\delta(\varphi - \varphi')$$

$$+ \eta(\varphi - \tfrac{1}{2}\pi)S_E^{-1}(\theta, \varphi \mid \theta', \varphi'). \tag{5.14}$$

6. DIRAC'S EQUATION IN THREE DIMENSIONS

We denote the spinor index which appears in the wavefunctions of the Dirac Hamiltonian by ω, where $\omega = 1, 2, 3, 4$. Thus $(\mathbf{x}, \omega)$ is the set of variables which describe states in the coordinate representation.

The Dirac Hamiltonian is then written

$$H = H_0 + V_{\mathbf{x}}(\omega \mid \omega'), \tag{6.1}$$

where

$$H_0 = i\sum_{i=1}^{3} \alpha_i(\omega \mid \omega')\frac{\partial}{\partial x_i} - m\beta(\omega \mid \omega'). \tag{6.2}$$

In (6.1), $V_{\mathbf{x}}(\omega \mid \omega')$ are the elements of a Hermitian matrix in the variables ω which are functions of $\mathbf{x}$ in such a way that the elements vanish when $|\mathbf{x}| \to \infty$. The matrices $\alpha_i(\omega \mid \omega')$ and $\beta(\omega \mid \omega')$ are the usual Dirac matrices.

The eigenfunctions of H_0 will be denoted by $\psi_0(x, \omega \mid E, \theta, \varphi, \tau)$, where E is the eigenvalue of $H_0(-\infty < E < -m$ or $m < E < \infty)$, θ and φ are the polar angles which give the direction of the momentum $\mathbf{p} = (E^2 - m^2)^{\frac{1}{2}}$ (sin θ cos φ, sin θ sin φ, cos θ), and τ equals 1 if the eigenvalue of the helicity is positive and -1 if the eigenvalue is negative. The eigenfunctions ψ_0 can be written

$$\psi_0(\mathbf{x}, \omega \mid E, \theta, \varphi, \tau) = (\sin\theta)^{\frac{1}{2}}|E|^{\frac{1}{2}}$$

$$\times (E^2 - m^2)^{\frac{1}{4}}\chi(\omega \mid E, \theta, \varphi, \tau)\frac{e^{i\mathbf{p}\cdot\mathbf{x}}}{(2\pi)^{\frac{3}{2}}}, \tag{6.3}$$

where $\chi(\omega \mid E, \theta, \varphi, \tau)$ are the spinor coefficients which can be chosen to satisfy the following orthonormality and completeness relations:

$$\sum_{\omega} \chi^*(\omega \mid E, \theta, \varphi, \tau)\chi(\omega \mid E, \theta, \varphi, \tau') = \delta_{\tau\tau'}, \tag{6.4a}$$

$$\sum_{\omega} \chi^*(\omega \mid E, \theta, \varphi, \tau)\chi(\omega \mid -E, \theta, \varphi, \tau') = 0, \tag{6.4b}$$

$$\sum_{\tau} [\chi(\omega \mid E, \theta, \varphi, \tau)$$

$$\times \chi^*(\omega', E, \theta, \varphi, \tau) + \chi(\omega \mid -E, \theta, \varphi, \tau)$$

$$\times \chi^*(\omega' \mid -E, \theta, \varphi, \tau)] = \delta_{\omega\omega'}. \tag{6.4c}$$

In Eqs. (6.4) the δ function is a Kronecker δ.

The eigenfunctions ψ_0 can then be shown to satisfy the following orthonormality and completeness relations:

$$\sum_{\omega} \int d\mathbf{x} \psi_0^*(\mathbf{x}, \omega \mid E, \theta, \varphi, \tau) \psi_0(\mathbf{x}, \omega \mid E', \theta', \varphi', \tau') = \delta(E - E') \delta(\theta - \theta') \delta(\varphi - \varphi') \delta_{\tau\tau'}, \quad (6.5a)$$

$$\sum_{\tau} \int dE\, d\theta\, d\varphi \psi_0(\mathbf{x}, \omega \mid E, \theta, \varphi, \tau) \times \psi_0^*(\mathbf{x}', \omega' \mid E, \theta, \varphi, \tau) = \delta(\mathbf{x} - \mathbf{x}') \delta_{\omega\omega'}. \quad (6.5b)$$

An identity which will prove useful later is

$$[E \delta_{\omega\omega'} - \sum_i p_i \alpha_i(\omega \mid \omega') - m\beta(\omega \mid \omega')] = 2E \sum_{\tau} \chi(\omega \mid E, \theta, \varphi, \tau) \chi^*(\omega' \mid E, \theta, \varphi, \tau). \quad (6.6)$$

The eigenfunctions $\psi_\pm$ of H satisfy the integral equations

$$\psi_\pm(\mathbf{x}, \omega \mid E, \theta, \varphi, \tau) = \psi_0(\mathbf{x}, \omega \mid E, \theta, \varphi, \tau) + \sum_{\omega', \omega''} \int d\mathbf{x}' G_{\pm E}(\mathbf{x}, \omega \mid \mathbf{x}', \omega') V_{\mathbf{x}'}(\omega' \mid \omega'') \times \psi_\pm(x', \omega'' \mid E, \theta, \varphi, \tau), \quad (6.7)$$

where the Green's functions $G_{\pm E}$ are given by

$$G_{\pm E}(\mathbf{x}, \omega \mid \mathbf{x}', \omega') = [E + H_0] \left[-\frac{\exp i\epsilon(E) p \mid \mathbf{x} - \mathbf{x}' \mid}{4\pi \mid \mathbf{x} - \mathbf{x}' \mid} \right]. \quad (6.8)$$

In (6.8) the operator H_0 acts on the variables $\mathbf{x}$. Also $p = (E^2 - m^2)^{\frac{1}{2}}$. The quantity $\epsilon(E)$ equals 1 for $E > m$, and $\epsilon(E)$ equals -1 for $E < -m$.

One can also show that, on using (6.7), (6.8), and (6.6) and (2.4a),

$$\lim_{\mid \mathbf{x} \mid \to \infty} \psi_-(\mathbf{x}, \omega \mid E, \theta, \varphi, \tau) = \lim_{\mid \mathbf{x} \mid \to \infty} \psi_0(\mathbf{x}, \omega \mid E, \theta, \varphi, \tau) - \frac{i}{(2\pi)^{\frac{1}{2}}} \frac{\epsilon(E) \mid E \mid^{\frac{1}{2}}}{(\sin \theta')^{\frac{1}{2}} (E^2 - m^2)^{\frac{1}{4}}} \frac{e^{i\epsilon(E) p \mid \mathbf{x} \mid}}{\mid \mathbf{x} \mid} \times \sum_{\tau'} \chi(\omega \mid E, \theta', \varphi', \tau') [S_E(\theta', \varphi', \tau' \mid \theta, \varphi, \tau) - \delta(\theta - \theta') \delta(\varphi - \varphi') \delta_{\tau\tau'}], \quad (6.9)$$

where

$$\begin{matrix} \theta' = \lambda \\ \varphi' = \sigma \end{matrix} \text{ for } E > m, \text{ and } \begin{matrix} \theta' = \hat{\lambda} \\ \varphi' = \hat{\sigma} \end{matrix} \text{ for } E < -m.$$

Also,

$$\lim_{\mid \mathbf{x} \mid \to \infty} \psi_+(\mathbf{x}, \omega \mid E, \theta, \varphi, \tau) = \lim_{\mid \mathbf{x} \mid \to \infty} \psi_0(\mathbf{x}, \omega \mid E, \theta, \varphi, \tau) + \frac{i}{(2\pi)^{\frac{1}{2}}} \frac{\epsilon(E) \mid E \mid^{\frac{1}{2}}}{(\sin \theta')^{\frac{1}{2}} (E^2 - m^2)^{\frac{1}{4}}} \frac{e^{-i\epsilon(E) p \mid \mathbf{x} \mid}}{\mid \mathbf{x} \mid} \times \sum_{\tau'} \chi(\omega \mid E, \theta', \varphi', \tau') [S_E^{-1}(\theta', \varphi', \tau' \mid \theta, \varphi, \tau) - \delta(\theta - \theta') \delta(\varphi - \varphi') \delta_{\tau\tau'}], \quad (6.10)$$

where

$$\begin{matrix} \theta' = \hat{\lambda} \\ \varphi' = \hat{\sigma} \end{matrix} \text{ for } E > m, \text{ and } \begin{matrix} \theta' = \lambda \\ \varphi' = \sigma \end{matrix} \text{ for } E < -m.$$

Also, on using (5.5) and (6.3),

$$\lim_{\mid \mathbf{x} \mid \to \infty} \psi_0(\mathbf{x}, \omega \mid E, \theta, \varphi, \tau) = \frac{i}{(2\pi)^{\frac{1}{2}}} \frac{\mid E \mid^{\frac{1}{2}}}{(\sin \theta)^{\frac{1}{2}} \mid \mathbf{x} \mid (E^2 - m^2)^{\frac{1}{4}}} \chi(\omega \mid E, \theta, \varphi, \tau) \times [-e^{ip \mid \mathbf{x} \mid} \delta(\theta - \lambda) \delta(\varphi - \sigma) + e^{-ip \mid \mathbf{x} \mid} \delta(\theta - \hat{\lambda}) \delta(\varphi - \hat{\sigma})]. \quad (6.11)$$

We can obtain μ_{+E}^{-1} in much the same way as in Sec. 5.

For $E > m$,

$$\mu_{-E}^{-1}(\theta, \varphi, \tau \mid \theta', \varphi', \tau') = \eta(\tfrac{1}{2}\pi - \theta) S_E(\theta, \varphi, \tau \mid \theta', \varphi', \tau') + \eta(\theta - \tfrac{1}{2}\pi) \delta(\theta - \theta') \delta(\varphi - \varphi') \delta_{\tau\tau'}, \quad (6.12)$$

$$\mu_{+E}^{-1}(\theta, \varphi, \tau \mid \theta', \varphi', \tau') = \eta(\tfrac{1}{2}\pi - \theta) \delta(\theta - \theta') \delta(\varphi - \varphi') \delta_{\tau\tau'} + \eta(\theta - \tfrac{1}{2}\pi) S_E^{-1}(\theta, \varphi, \tau \mid \theta', \varphi', \tau'). \quad (6.13)$$

For $E < -m$,

$$\mu_{-E}^{-1}(\theta, \varphi, \tau \mid \theta', \varphi', \tau') = \eta(\tfrac{1}{2}\pi - \theta) \delta(\theta - \theta') \delta(\varphi - \varphi') \delta_{\tau\tau'} + \eta(\theta - \tfrac{1}{2}\pi) S_E(\theta, \varphi, \tau \mid \theta', \varphi', \tau'), \quad (6.14)$$

$$\mu_{+E}^{-1}(\theta, \varphi, \tau \mid \theta', \varphi', \tau') = \eta(\tfrac{1}{2}\pi - \theta) S_E^{-1}(\theta, \varphi, \tau \mid \theta', \varphi', \tau') + \eta(\theta - \tfrac{1}{2}\pi) \delta(\theta - \theta') \delta(\varphi - \varphi') \delta_{\tau\tau'}. \quad (6.15)$$

The techniques for generalizing the Jost functions can be extended even further. One could, for example, consider the scattering of particles by compound particles. However, we refrain in the present paper from carrying out the procedure.