JOURNAL OF
MATHEMATICAL
PHYSICS

VOLUME 6, NUMBER 6 JUNE 1965

Unitarity and On-Mass-Shell Analyticity as a Basis for S-Matrix Theories. I*

J. GUNSON
Department of Mathematical Physics, University of Birmingham, Birmingham, England
(Received 20 January 1964; final manuscript received 30 June 1964)

Many-particle unitarity relations are analyzed and related to maximal analyticity conjectures with particular reference to continuations into unphysical sheets. Poles and cuts associated with unstable particles are treated in detail and generalized unitarity relations derived.

1. INTRODUCTION

IN this series of papers we continue a study of the analytic properties of general S-matrix elements which was initiated for the case of two-particle scattering in a previous paper,[1] hereafter denoted by UP. Many of the results appearing here can be regarded as providing support in favor of the analyticity conjecture appearing in the conclusion to UP. The essential point of this conjecture is, in current terminology, that of "maximal analyticity,"[2] according to which we assume that suitably defined scattering amplitudes may be analytically continued arbitrarily to complex values of the energy–momentum variables wherever it is not inconsistent with the unitarity conditions.

As a starting point, we suppose that the scattering amplitudes are themselves analytic functions of the real energy–momentum variables in the physical regions except at thresholds for new physically possible processes. This postulate, together with a specification for approaching the correct physical limit from the associated complex neighborhood of analyticity, is regarded here as a "causality" condition on the theory, by analogy with the properties of the corresponding field theoretical functions which follow from the more usual causality condition. Next we suppose that the amplitudes are analytically continuable from these neighborhoods of the physical region into a certain domain whose bounding singularities must have some physical interpretation, however indirect, in order to be there at all. We propose to show that, apart from selection rules, etc., consistency with unitarity is all that is in principle required to determine the over-all singularity structure.

As in UP, we restrict the discussion to theories with neutral strongly interacting bosons and now work entirely with on-mass-shell amplitudes. It is our contention that off-mass-shell continuations contain no more physical information than on-mass-shell quantities and are nonessential adjuncts to a complete theory, though they are very useful mathematical tools. Of course, this requires that we should be able to express all measurable physical quantities in terms of on-mass-shell S-matrix elements or as limits of such. At present, we know of no counterexample to this statement. We propose to demonstrate in this paper that a complete theory can be constructed in this manner and thus provide another alternative to the field theory and other off-mass-shell approaches. This contrasts with some

* Based on a thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the University of Cambridge (1962).

[1] J. Gunson and J. G. Taylor, Phys. Rev. 119, 112 (1960).

[2] G. F. Chew, *Dispersion Relations* (Oliver and Boyd, Edinburgh, 1961), p. 168.

formulations of the complete theories based on maximal analyticity that have been proposed,[3,4] which make essential use of off-mass-shell continuations.[5]

We commence with a summary of formalism and notation in the section immediately following, together with a more precise statement of the maximal analyticity principle. In Sec. 3 there are listed equations expressing unitarity for processes involving up to four particles in each state and these are used effect continuations onto unphysical sheets. The complex poles and branch points reached on carrying out these continuations are interpreted as threshold singularities associated with states containing unstable particles, by proving the natural extensions of the unitarity conditions for these cuts. Related results concerning these singularities have been obtained independently by Blankenbecler *et al.*[6] Stapp,[7] and recently by Zwanziger,[8] Landshoff,[9] and Olive.[10] In Sec. 4, we apply the methods developed for unstable particle cuts to the stable particle cuts themselves, leading to important self-consistency relations. Analogous problems for the K matrix are treated briefly in Sec. 5. Two appendices deal with the properties of unitary S matrices in an arbitrary number of space dimensions and the proof of unitarity relations for cuts involving two unstable particles.

In Part II, we consider possible methods of approach to the problem of constructing complete dynamical theories of the S matrix based on analyticity and on-mass-shell unitarity. In Part III, we complement the above study of the effects of unitarity on analyticity in the energy variables by developing a method for continuation in the momentum transfer variables.

2. S-MATRIX FORMALISM AND MAXIMAL ANALYTICITY.

As a reasonably realistic model which still partially avoids the complications of spin and charge variables we take the one used in UP of a theory with two stable spin-zero particles, the "nucleon" and the "meson" of nonzero masses m_1 and m_2, respectively. The selection rules are taken to be those obtained from a conserved multiplicative quantum number taking the values -1 for nucleons and $+1$ for mesons.

For these particles we construct spaces of ingoing and outgoing states. Such a state containing several particles is taken to be the direct product of single particle states in which energy, momentum, etc., are additive. In other words, interaction between the particles composing the state is taken to be absent or negligible.

The probability amplitude for the transition from a normalised in-state $|\alpha_{\text{in}}\rangle$ to a normalised out-state $|\beta_{\text{out}}\rangle$ is given by the S-matrix element

$$S_{\beta\alpha} = \langle\beta_{\text{out}}|\alpha_{\text{in}}\rangle = \langle\beta_{\text{in}}|\, S\,|\alpha_{\text{in}}\rangle \tag{2.1}$$

defining the S-operator S. The assumed completeness of the set of in- and out-states requires S to be unitary if the usual quantum mechanical probability amplitude interpretation is to hold, i.e.,

$$SS^* = S^*S = I. \tag{2.2}$$

A basis for the space of in-states may be formed by repeated applications of the creation operators $a_1^*(\mathbf{k})$ and $a_2^*(\mathbf{k})$ for nucleons and mesons of 3-momentum $\mathbf{k}$ onto the vacuum state $|0\rangle$. These operators satisfy the usual commutation relations

$$[a_\mu(\mathbf{k}), a_\nu^*(\mathbf{k}')] = \delta^{(3)}(\mathbf{k} - \mathbf{k}')\,\delta_{\mu\nu};\ \mu, \nu = 1, 2, \tag{2.3}$$

all other commutators vanishing. Setting $k_\nu = (k_\nu^0, \mathbf{k})$ where $k_\nu^0 = \pm(|\mathbf{k}|^2 + m_\nu^2)^{\frac{1}{2}}$, we may define the in-operators

$$\phi_\nu(k_\nu) = \begin{cases} (2k_\nu^0)^{\frac{1}{2}} a_\nu(\mathbf{k}) & k_\nu^0 > 0, \\ (-2k_\nu^0)^{\frac{1}{2}} a_\nu^*(-\mathbf{k}) & k_\nu^0 < 0, \end{cases} \quad \nu = 1, 2. \tag{2.4}$$

The formal normal product expansion for the S matrix then takes the form

$$\begin{aligned} (S - I)/i = &\sum_{\mu=1}^{\infty}\sum_{\nu=1}^{\infty} \frac{1}{\mu!}\frac{1}{\nu!} \\ &\cdot \int \cdots \int d^4k_1 \cdots d^4k_\mu\, d^4q_1 \cdots d^4q_\nu \\ &\cdot \delta^{(4)}(k_1 + \cdots + k_\mu + q_1 + \cdots + q_\nu) \\ &\cdot \prod_{i=1}^{\mu}\prod_{j=1}^{\nu} \delta(k_i k_i - m_1^2)\,\delta(q_j q_j - m_2^2) \\ &\cdot W_{\mu,\nu}(k_1, \cdots, k_\mu; q_1, \cdots, q_\nu) \\ &\cdot :\phi_1(k_1) \cdots \phi_1(k_\mu)\phi_2(q_1) \cdots \phi_2(q_\nu):. \end{aligned} \tag{2.5}$$

Here $W_{\mu,\nu}$ is an amplitude for a scattering process involving a total of μ external nucleons and ν

[3] J. C. Polkinghorne, Nuovo Cimento **23**, 360 (1962) and **25**, 901, (1962).
[4] H. P. Stapp, Phys. Rev. **125**, 2139 (1963).
[5] However, a recent paper by J. C. Polkinghorne, Phys. Rev. **128**, 2898 (1962), shows that this need not be so.
[6] R. Blankenbecler, M. L. Goldberger, S. W. McDowell, and S. B. Treiman, Phys. Rev. **123**, 692 (1961).
[7] H. P. Stapp, "On the Masses and Lifetimes of Unstable Particles," Preprint UCRL-10261.
[8] D. Zwanziger, Phys. Rev. **131**, 888 (1963).
[9] P. V. Landshoff, Nuovo Cimento **28**, 123 (1963).
[10] D. Olive, "On the Analytic Continuation of a Scattering Amplitude through a Three-Particle Cut," Cambridge Preprint (1962).

external mesons. The sign of the energy component k_α^0 or q_α^0 of the αth nucleon or meson determines whether it is to be regarded as ingoing or outgoing. This amplitude is the same as the usual Feynman amplitude except that those contributions involving particles which do not interact[11] are omitted.

TCP invariance, which has been proved by Stapp[4] from postulates implied by ours, leads to the relation

$$W_{\mu,\nu}(k_1, \cdots, k_\mu; q_1, \cdots, q_\nu) = W_{\mu,\nu}(-k_1, \cdots, -k_\mu; -q_1, \cdots, -q_\nu). \quad (2.6)$$

On combining (2.2), (2.5), and (2.6) and applying Wick's theorem,[12] we obtain the unitarity relations

$$\begin{aligned}
&\text{Im } W_{\mu,\nu}(k_1, \cdots, k_\mu; q_1, \cdots, q_\mu) \\
&= \frac{1}{2}\sum_{\alpha=0}^{\infty}\sum_{\beta=0}^{\infty}\frac{1}{\alpha!}\frac{1}{\beta!}\sum_{s=0}^{\mu}\sum_{t=0}^{\nu}P\left(\lambda; \frac{s}{\mu}\right)P\left(\sigma; \frac{t}{\nu}\right) \\
&\cdot\int\cdots\int d^4p_1\cdots d^4p_\alpha\, d^4r_1\cdots d^4r_\beta \\
&\cdot\prod_{i=1}^{\alpha}[\theta(p_i^0)\,\delta(p_ip_i - m_1^2)] \\
&\cdot\prod_{j=1}^{\beta}[\theta(r_j^0)\,\delta(r_jr_j - m_2^2)]\,\delta^{(4)} \\
&\cdot(k_{\lambda_1} + \cdots k_{\lambda_s} + q_{\sigma_1} + \cdots \\
&+ q_{\sigma_t} + \cdots + p_1 + \cdots + p_\alpha + r_1 + \cdots + r_\beta) \\
&\cdot W^*_{s+\alpha,t+\beta}(k_{\lambda_1}, \cdots, k_{\lambda_s}, p_1, \cdots, p_\nu; \\
&\cdot q_{\sigma_1}, \cdots, q_{\sigma_t}, r_1, \cdots, r_\beta) \\
&\cdot W_{\mu-s+\alpha,\nu-t+\beta}(k_{\lambda_{s+1}}, \cdots, k_{\lambda_\mu}, -p_1, \cdots, -p_\alpha; \\
&\cdot q_{\sigma_{t+1}}, \cdots, q_{\sigma_\nu}, -r_1, \cdots, -r_\nu), \quad (2.7)
\end{aligned}$$

where $P(\lambda; s/\mu)$ denotes the sum over all $\mu!/s!(\mu-s)!$ ways of breaking 1, 2, $\cdots$, μ into sets $\lambda_1, \cdots, \lambda_s$ and $\lambda_{s+1}, \cdots, \lambda_\mu$. This form of the unitarity condition has the advantage of explicitly demonstrating the contributions from the various possible channels for a particular process. In general, it is valid only in the physical regions appropriate to the various possible choices of signs for the k^0 and q^0. A graphical method for displaying the separate equations contained in (2.7) is developed in the later sections.

As some of the W amplitudes still contain vacuum singularities, we must define a new type of amplitude denoted by V in which these singularities are eliminated, before applying the analyticity postulates. Equation (5.5) of Sec. 5 gives the defining equations for the V amplitudes to be

$$\begin{aligned}
&V_{\mu,\nu}(k_1, \cdots, k_\mu; q_1, \cdots, q_\nu) \\
&= W_{\mu,\nu}(k_1, \cdots, k_\mu; q_1, \cdots, q_\nu) \\
&- \frac{i}{2}\sum_{s=0}^{\mu}\sum_{t=0}^{\nu}P\left(\lambda; \frac{s}{\mu}\right)P\left(\sigma; \frac{t}{\nu}\right)\delta^{(4)}(k_{\lambda_1} + \cdots + k_{\lambda_s} \\
&+ q_{\sigma_1} + \cdots + q_{\sigma_t})V_{s,t}(k_{\lambda_1}, \cdots, k_{\lambda_s}; q_{\sigma_1}, \cdots, q_{\sigma_t}) \\
&\cdot W_{\mu-s,\nu-t}(k_{\lambda_{s+1}}, \cdots, k_{\lambda_\mu}; q_{\sigma_{t+1}}, \cdots, q_{\sigma_\nu}). \quad (2.8)
\end{aligned}$$

Our basic postulate can now be stated in the form: the amplitudes $\{V_{\mu,\nu}\}$ are, in given physical regions, boundary values of functions analytic in all the Lorentz invariant variables on which they depend and the boundary values themselves are analytic except at the physical thresholds determined by unitarity. The way in which the boundary value is to be taken is of importance. This is obtained from the Feynman prescription of adding small negative imaginary parts to the masses of the intermediate particles and may be regarded as our "causality" condition on the S-matrix theory. This procedure displaces the physical threshold singularities to slightly complex values, making it clear on which sides of the corresponding cuts the physical region lies. If we now transform to the usual prescription of taking the limit with fixed real values of the intermediate particle masses and approaching through slightly complex values of the external invariants, we find that it is not possible to take this limit from a uniquely defined "physical sheet" except in special cases. To take an example, we may choose the set of independent invariants (cf. Appendix A)

$$\begin{array}{lll}
s_{123}\quad s_{12}\quad s_{13}\quad s_{45}\quad s_{46} & & \text{"energy invariants",} \\
\qquad\quad s_{14}\quad s_{15}\quad s_{24} & & \text{"momentum transfer invariants"}
\end{array}$$

for the three-particle scattering process with ingoing momenta k_1, k_2, k_3 and outgoing momenta k_4, k_5, k_6. The vanishing of the Gram determinant of k_1, k_2, k_3, k_4, k_5 as required by the four-dimensionality of the energy–momentum space then precludes us from adding small imaginary parts all of the same sign to s_{123}, s_{12}, s_{23}, s_{13}, s_{45}, s_{46} in the physical region near threshold. This that we have to go into an "unphysical" sheet in at least one energy channel in order to take our point of observation off all the physical cuts in the amplitude. This phenomenon occurs only for processes involving six or more external particles.

The situation can be remedied if the dimensionality of momentum space is relatively unimportant for the particular problem being considered. By

[11] Examples of such contributions appear in diagrammatical form in Eqs. (3.13) and (3.16).

[12] G. C. Wick, Phys. Rev. 80, 268 (1950).

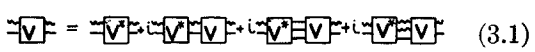
relaxing the dimensionality restrictions, we may consider an extended function defined on the $[n(n-3)/2]$-dimensional complex Euclidean space of the $s_{\alpha\beta}$ (cf. Appendix A). This extension is not uniquely determined by the physical values of the amplitudes, as the physical region is not a real environment[13] in this larger space. Some of the consequences of imposing a generalized unitarity condition for the case of arbitrary dimensionality are discussed in Appendix A.

An alternative procedure is to accept the extra complication as it stands and forget about precisely defined physical sheets, treating the function as a whole.

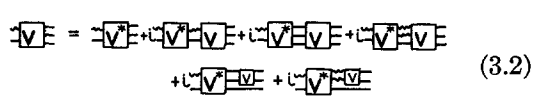
The next stage in applying the analyticity postulate is to continue these analytic functions away from the neighborhoods of the physical regions along all possible paths. The only singularities attained are those required for consistency with the unitarity conditions, the functions being otherwise holomorphic. The following sections contain a discussion of the self-consistency of the postulates for the simplest cases. In order to apply the unitarity conditions, we require also the possibility of being able to continue analytically to the complex-conjugate value V^* by taking the opposite limits on appropriate sheets.[14]

3. UNITARITY CONDITIONS FOR PROCESSES OF LOW ORDER

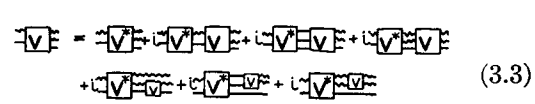
We commence by writing the unitarity relations for the V amplitudes in an obvious graphical notation. The nucleon and meson are denoted by straight and wavy lines respectively. If $3m_1/2 < m_2 < 2m_1$, then we may write the following relations, valid in physical regions up to a total c.m. energy W of $3m_1 + m_2$



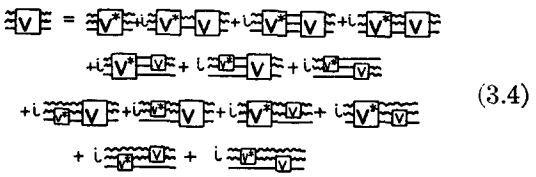
(3.1)



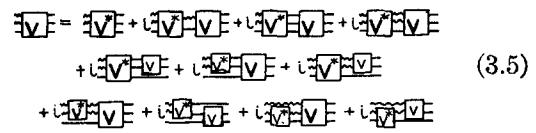
(3.2)



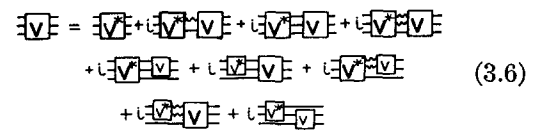
(3.3)



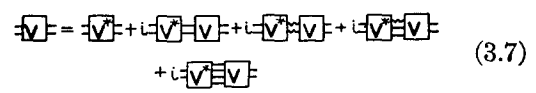
(3.4)



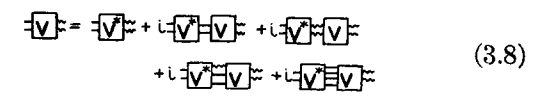
(3.5)



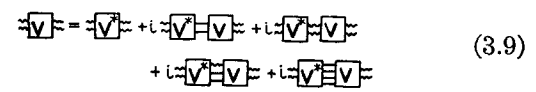
(3.6)



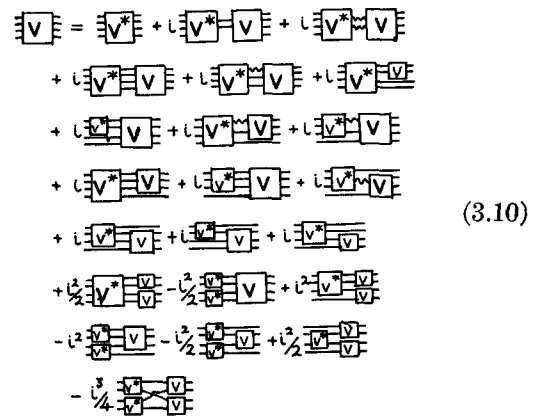
(3.7)



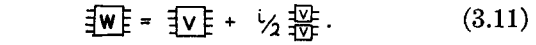
(3.8)



(3.9)



(3.10)



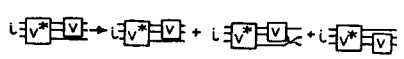
(3.11)

These relations are obtained directly from (2.7) and (2.8) by retaining only the nonvanishing terms allowed by the selection rules and energy–momentum conservation. Most of the physical thresholds can be obtained by direct inspection of the various terms, but some of the anomalous ones appear only in the more detailed analysis of Sec. 3.3 *et seq.* A summation convention has been used: whenever a term on the rhs does not possess the symmetry

[13] S. Bochner and W. T. Martin, *Several Complex Variables* (Princeton University Press, Princeton, New Jersey, 1948), p. 33.

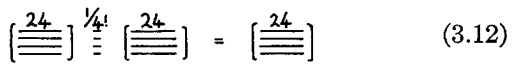
[14] D. Olive, Nuovo Cimento **26**, 73 (1962).

of the lhs amplitude under permutations of ingoing or outgoing particle lines, then a sum over the smallest number of permutations applied to the term actually appearing, required to restore the symmetry, is implied (neglecting any symmetries between component amplitudes). This is most easily illustrated by examples. One of the terms in (3.6) when written in full becomes

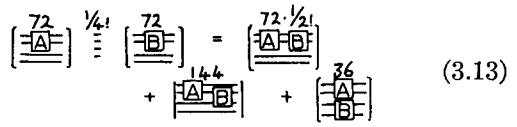


(3.6a)

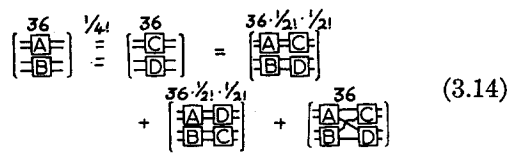
and the final term in (3.11) stands for a sum over 36 different terms, constructed by permutatuions of external lines. Examples of the technique for multiplication of these quantities are as follows:



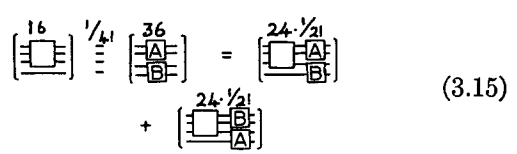
(3.12)



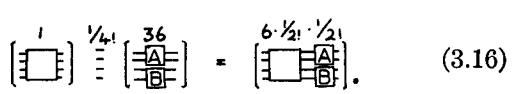
(3.13)



(3.14)



(3.15)



(3.16)

The integers above the brackets denote the number of terms in the symmetrized sum, and the inverse factorials give the $1/n!$ factor appropriate to each integration over an n-particle intermediate state connecting any two individual components in the graph [this is the factor $1/\alpha!$ in (2.7)].

3.1 Introduction of Interpolating Amplitudes

In the physical regions there are many overlapping branch cuts associated with the threshold singularities. The physical scattering amplitude is defined by only one limiting process onto these cuts out of many different possible ways. The alternative limits define new functions in the physical regions which are related to the original scattering amplitudes and in this subsection we will show how to obtain some of them in an operational formalism. These will be called "interpolating" amplitudes by virtue of the fact that they form natural sequences of functions terminating at either end with the physical amplitude and its complex conjugate.

Consider a scattering process involving i ingoing and j outgoing particles with $i \leq j$. If we restrict the theory to that of a single mass m for the present, then the physical region is a subset of the domain $W^2 > (jm)^2$ where $W^2 = (k_1 + k_2 + \cdots + k_i)^2$. We dissect the space of invariants $s_{\alpha\beta} = (k_\alpha + k_\beta)^2$ after the manner of Appendix A into subsets $J(\beta)$. The reason for this subdivision is that for all points in a single $J(\beta_0)$ in the physical region, there is a uniquely determined set of terms in our form of the unitarity conditions which give nonzero contributions, being the ones allowed by energy–momentum conservation. For this given subregion $J(\beta_0)$, which we take to lie in the range $(nm)^2 < W^2 < [(n+1)m]^2$, we see that the usual unitarity condition is the same as that obtained from the modified expression

$$T = T^* + iTP_nT^* = T^* + iT^*P_nT, \qquad iT = S - I, \tag{3.17}$$

where P_n is the projection operator onto states which contain no more than n particles.

For this same range in W^2, we define sequences of operators T_ν, $0 \leq \nu \leq n$, which "interpolate" between T and T^* in the following manner

$$T_\nu = T^* + iT^*P_\nu T_\nu = T^* + iT_\nu P_\nu T^*, \tag{3.18}$$

in which P_ν is the projection operator onto a certain subset of the set of all states containing not more than n particles, such that in the subspace of m-particle states, P_ν is either the unit or null operator. The sequences in which we are interested are the maximal sequences linearly ordered by inclusion relations of the form $P_{\nu_1} \subset P_{\nu_2}$, there being $(n+1)!$ such sequences.

When $P_\nu = P_n$, (3.18) reduces to (3.17) and we set $T_\nu = T_n$. When $P_\nu = 0$, then $T_0 = T^*$ is the first element of all sequences. The formal solution of (3.18) is

$$T_\nu = (I - iT^*P_\nu)^{-1}T^* = T^*(I - iP_\nu T^*)^{-1}, \tag{3.19}$$

and the existence of the inverses will be verified later for certain cases from the existence of solutions to the integral equations to be derived from (3.19). The unitarity equations for the interpolating op-

erators can be derived as follows:

$$T_{\nu_1} - T_{\nu_2} = (I - iT^*P_{\nu_1})^{-1}T^* - T^*(I - iP_{\nu_2}T^*)^{-1}$$
$$= i(I - iT^*P_{\nu_1})^{-1}(T^*(P_{\nu_1} - P_{\nu_2})T^*)(I - iP_{\nu_2}T^*)^{-1}$$
$$= iT_{\nu_1}(P_{\nu_1}(I - P_{\nu_2}) - (I - P_{\nu_1})P_{\nu_2})T_{\nu_2}. \quad (3.20)$$

When $P_{\nu_1} \supset P_{\nu_2}$, then $(I - P_{\nu_1})P_{\nu_2} = 0$; hence

$$T_{\nu_1} = T_{\nu_2} + iT_{\nu_1}P_{\nu_1}(I - P_{\nu_2})T_{\nu_2}, \quad (P_{\nu_1} \supset P_{\nu_2}), \quad (3.21)$$

connecting any two terms in a particular sequence. On taking matrix elements we get the required interpolating unitarity equations.

It is apparent that these amplitudes are not associated just with the particular energy range and amplitude with which we started, but are of general application to the whole of the physical regions of the S-matrix elements. In conjunction with the analyticity conjectures, it will appear later that these interpolating amplitudes are indeed obtained by taking appropriate limits between the overlapping cuts on the physical region and so do indeed partly resolve the problem of separating the discontinuities across the various cuts. This separation however is not complete, as we have lumped together all cuts involving a definite number of particles in the intermediate states of the main channel. The refinement of the interpolation process to individual cuts is carried out later in this section for special cases.

The complex conjugates of the interpolating amplitudes possess similar properties and in particular we have the relations

$$T^*_{\nu_1} = T(I + iP_{\nu_1}T)^{-1} = (I + iTP_{\nu_1})^{-1}T, \quad (3.22)$$

$$T^*_{\nu_1} - T^*_{\nu_2} = -iT^*_{\nu_1}(P_{\nu_1} - P_{\nu_2})T^*_{\nu_2}, \quad (3.23)$$

$$T_{\nu_1} - T^*_{\nu_2} = iT_{\nu_1}P_{\nu_1}T^*_{\nu_2} - iT_{\nu_1}(I - P_{\nu_2})T^*_{\nu_2}, \quad (3.24)$$

$$T_{\nu} - T^*_{\nu} = iT_{\nu}P_{\nu}T^*_{\nu} - iT_{\nu}(I - P_{\nu})T^*_{\nu}. \quad (3.25)$$

3.2 Interpolating Amplitudes for Simple Processes

The threshold masses for the various intermediate states in the equations (3.1) to (3.10) can be divided into two classes, with odd and even nucleon number respectively (Fig. 1).

There are two interpolating amplitudes for an odd nucleon number sequence and three for an even

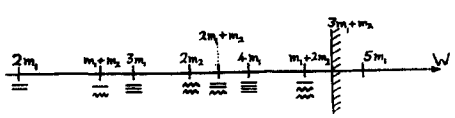


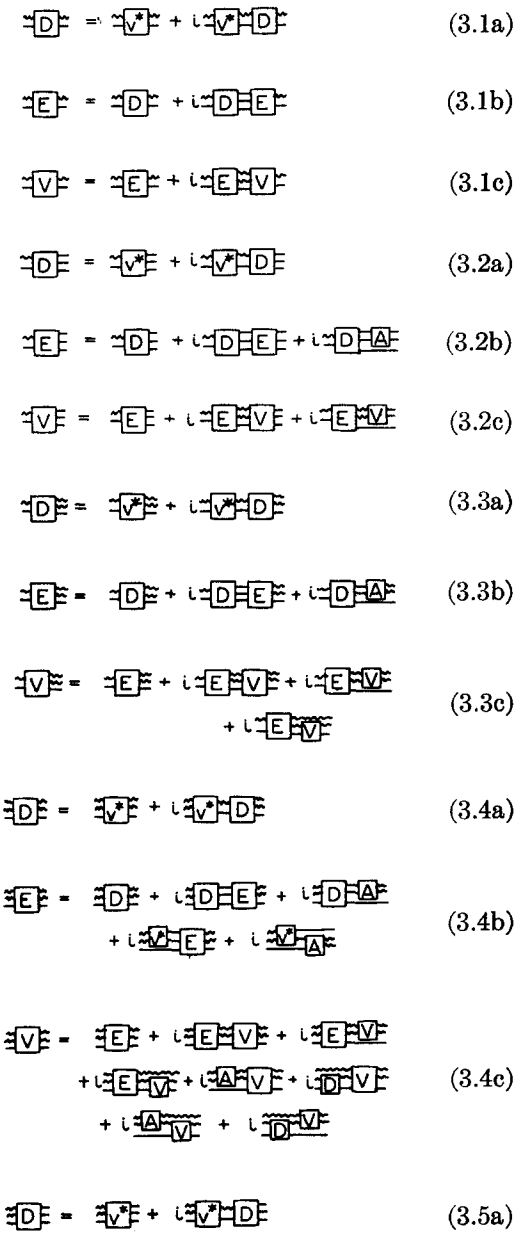
FIG. 1. Mass spectrum of intermediate states for the scalar nucleon–meson model with $m_2 \cong 7m_1/4$.

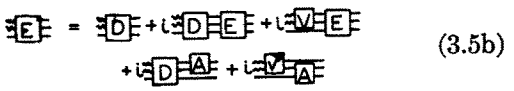
sequence. For example, consider the principal sequences,

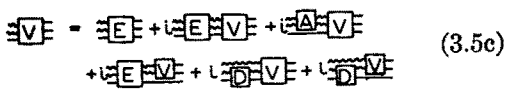
$$V^* \Longrightarrow A \Longrightarrow B \Longrightarrow C \Longrightarrow V \quad (3.26)$$

$$V^* \Longrightarrow D \Longrightarrow E \Longrightarrow V, \quad (3.27)$$

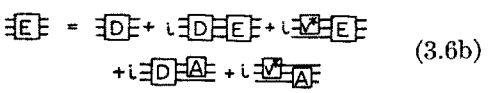
where the order in which we "jump" across sets of cuts is clearly indicated, in this case being in order of increasing mass. Treating the odd-nucleon-number case in detail, we obtain the following interpolating unitarity conditions for (3.1) to (3.9)

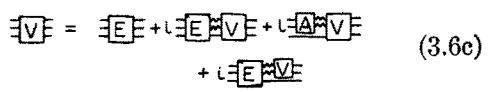


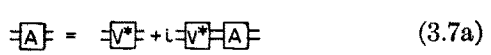

(3.5b)

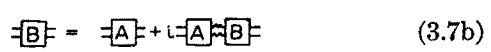

(3.5c)


(3.6a)


(3.6b)

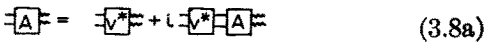

(3.6c)


(3.7a)

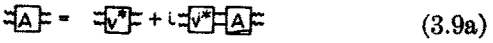

(3.7b)

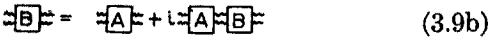

(3.7c)


(3.7d)


(3.8a)

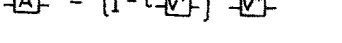

(3.8b)


(3.9a)


(3.9b)


(3.9c)

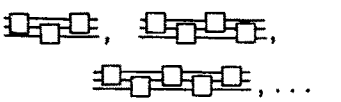
These equations are valid up to $W = 3m_1 + m_2$. Below the various thresholds in this range, the set of equations can be reduced in number as some become simple identities, e.g., $V = E$ below $W = m_1 + 2m_2$. As the equations connect boundary values from different sheets and form a complete set of coupled equations expressing all the other amplitudes in terms of, say, the V^* amplitudes, a study of the analytic properties of their solutions provides a connection between the analytic properties of the overall function on different sheets. This has been carried out in UP for the simplest case (3.7a). This integral equation was first reduced to a set of uncoupled algebraic equations by projecting out partial wave amplitudes and solved to give the analytic properties on the first unphysical sheet. The analytic properties of the total amplitude on this sheet have also been treated in an earlier paper.[15] In this paper, we will use the formal expression


(3.28)

to denote the solution to the integral equation. This notation will be used for more complicated cases when we have shown by other means that the solution exists and that the inverse is uniquely defined. The solutions for processes involving further particles depend essentially on the solutions of the three-particle integral equations (3.4b) and (3.6b).

3.3 Analytic Properties of the Solution to a Three-Particle Integral Equation

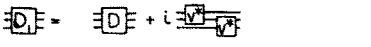
For our basic example we take the equation (3.6b). There are 16 normal thresholds occurring in this equation: one 3-particle, six 2-particle and nine 1-particle thresholds. In addition, there are an infinite number of anomalous thresholds corresponding to the boundaries of the regions in which the terms

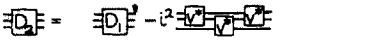

(3.29)

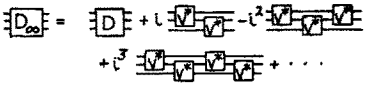
are allowed by energy momentum conservation with physical momenta. These will be discussed in more detail in Parts II and III.

We wish to transform this equation into a set of Fredholm equations so as to be able to apply the standard theory of the analytic properties of the denominator solutions to these latter equations.

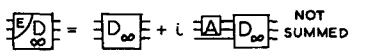
The first step is to introduce a finer subdivision into the sequence of interpolating amplitudes. Consider the set of equations


(3.30a)

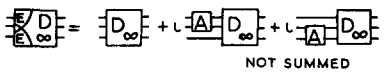

(3.30b)


(3.30c)

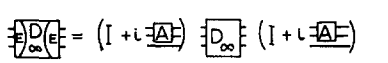
[15] J. Gunson and J. G. Taylor, Phys. Rev. **121**, 343, (1961).



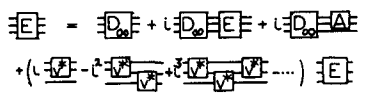
(3.31a)



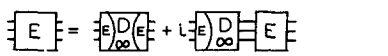
(3.31b)



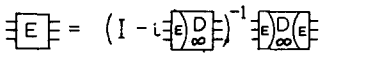
(3.31c)



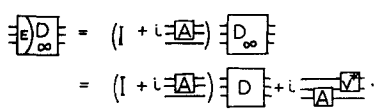
(3.32)



(3.33)



(3.34)



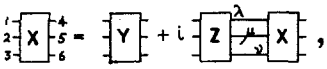
(3.35)

The interpretation of these equations is as follows. The D_1 amplitude on the lhs of (3.30a) is obtained from the D amplitude by "jumping" across the one particle poles, i.e., it is a boundary value taken in the same sense as the D amplitude except that the one-particle poles are approached from the other side. Similarly for the D_2 amplitude on the lhs of (3.30b) we approach the triangle graph anomalous threshold cuts from the "physical" side. The D_∞ amplitude on the lhs of (3.30c) is the limiting case where all of the "three-nucleon" anomalous threshold graphs are included. The statement that the various terms added actually express the discontinuities across the corresponding poles and cuts must at present be taken as part of the analyticity–unitarity connection postulate which appears at the end of Sec. 2. However, it is not clear whether this postulate is independent of the other postulates; certainly it appears consistent with them.

The expressions denoting the corresponding jumps across two-particle cuts appear in (3.31a, b, c). These may be taken in any order. The resulting unitarity equations are (3.32) and (3.33). Equation (3.33) is a pure three-particle equation of a type suitable for conversion into a standard Fredholm form and (3.34) is used to denote the corresponding solution in a formal manner.

We note an important consequence of these equations. In some cases, the discontinuities across the various cuts are independent of the order in which they are taken. For the normal thresholds, the cases in which this occurs are precisely those in which we have a Steinmann relation[16] holding for the corresponding field-theoretical generalised retarded function. A Steinmann relation, or permutability of discontinuities, can occur when the physical cuts from two different channels overlap and the channels possess at least one particle in common, whichever set of particles is used to specify the channel. A complete set of conditions for its validity is given in Araki's paper.[16] We conjecture that in analysis of the problem more detailed than that carried out here, the Steinmann relations would be found to hold in all the cases implied by the field theory results.

The problem has now been reduced to an equation of the generic type



(3.36)

where the region of integration in the final term must be deformed where necessary to slightly complex values of the intermediate variables in the manner indicated by the above discussion, so as to avoid passing through pole terms and branch points. Written in full, the three-particle unitarity integration operator in the c.m. system is (apart from some constant factors)

$$i\int\cdots\int d^4k_\lambda\, d^4k_\mu\, d^4k_\nu\, \delta^{(4)}((W,\mathbf{0})-k_\lambda-k_\mu-k_\nu)\cdot\prod_{\alpha=\lambda,\mu,\nu}\theta(k_\alpha^0)\,\delta(k_\alpha k_\alpha-m^2). \tag{3.37}$$

If we follow the partial wave projection procedure of Appendix A for both sides of (3.36), we obtain the set of simultaneous Fredholm equations

$$X_i^{m,n'}(W^2,s_{12},s_{23},s_{45},s_{46}) = Y_i^{m,m'}(W^2,s_{12},s_{23},s_{45},s_{46}) + \frac{1}{W^2}\sum_{m''=-j}^{+j}\iint ds_{\lambda\mu}\,ds_{\lambda\nu}\,Z_i^{m,m''}(W^2,s_{12},s_{23},s_{\lambda\mu},s_{\lambda\nu})\cdot X_i^{m'',m'}(W^2,s_{\lambda\mu},s_{\lambda\nu},s_{45},s_{46}), \tag{3.38}$$

where the region of integration is given by (cf. Appendix A)

$$\begin{aligned} &s_{\lambda\mu}+s_{\lambda\nu}+s_{\mu\nu}=W^2+3m_1^2,\\ &s_{\lambda\mu}\,s_{\lambda\nu}\,s_{\mu\nu}\geq m_1^2(W^2-m_1^2)^2,\\ &s_{\lambda\mu},\,s_{\lambda\nu},\,s_{\mu\nu}\geq 4m_1^2 \quad\text{(before continuation)}, \end{aligned} \tag{3.39}$$

with the appropriate deformations around poles and branch points. It is sufficient for the initial study

[16] H. Araki, J. Math. Phys. 2, 163 (1961).

of analytic properties to consider the simpler equation

$$f(x, y; t) = g(x, y; t) + \iint_G K(x, y, x', y'; t) f(x', y'; t)\, dx'\, dy', \quad (3.40)$$

where G is given by

$$x + y + z = t + 3,$$
$$xyz \geq (t - 1)^2, \quad (3.41)$$
$$x, y, z \geq 4 \quad \text{(before continuation)}.$$

The correspondence with (3.38) and (3.39) is obvious. It follows from our basic postulates of analyticity for the region of integration that the Fredholm solution must exist in the physical region for $9 < t < t_0$, for some $t_0 > 9$. Furthermore, the usual series expansions[17] of the Fredholm determinant and the first Fredholm minor are absolutely and uniformly convergent on any compact subset in t of this interval. The analyticity of these quantities then follows from the analyticity in t of each term in the expansion.

When we attempt to continue the determinantal solution off the real t axis to complex values, we see that analyticity must be maintained unless

(i) a singularity present in the kernel K or inhomogeneous term g on a surface t = constant is reached,

(ii) singularities depending only on the x and/or y variables appear on the boundary of the region of integration in such a manner that it cannot be deformed away from them,

(iii) a singularity of the equations defining the boundaries of the region of integration is reached,

(iv) the Fredholm determinant $D(t)$ vanishes for the value of t attained,

(v) two singularities in the kernel or inhomogeneous term coincide and "pinch"[18] the contour.

Case (ii) is of most interest in the present context and leads to new singularities. In the real x–y plane, the available phase space for a typical value of t is the shaded region of Fig. 2.

This two-dimensional surface is arbitrarily continuously deformable to complex values of x and y, under the restrictions that the boundaries remain on the manifold defined by taking the equality signs

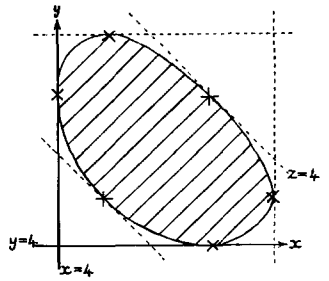


FIG. 2. Three-particle phase space, equal-mass case.

in the first two equations of (3.41). For a given value of t there are six points (indicated by crosses in Fig. 2) permanently fixed and specified by

$$\left.\begin{aligned} x &= 4 \\ y = z &= \tfrac{1}{2}(t-1) \end{aligned}\right\} \quad \left.\begin{aligned} x &= (t^{\frac{1}{2}} - 1)^2 \\ y = z &= t^{\frac{1}{2}} + 1 \end{aligned}\right\}. \quad (3.42)$$

and equations with $x \to y \to z \to x$.

Suppose there is a pole at $x = a^2/m^2$ in $g(x, y; t)$ corresponding to a complex pole on the unphysical sheet of the two-particle amplitude at $s_{\lambda\mu} = a^2$. If t is such that

$$(t^{\frac{1}{2}} - 1)^2 = a^2/m_1^2, \quad \text{i.e.,} \quad t = (a/m_1 + 1)^2 \quad (3.43)$$

on taking the appropriate root, then we can have a singularity of the solution at this point, because of the nondeformability of the integration contour away from this pole. For the equation (3.38) this becomes

$$W^2 = (a + m_1)^2. \quad (3.44)$$

Such a singularity will be shown to possess a simple interpretation as a threshold singularity for the production of a pair of particles, one unstable with complex mass a and one stable with mass m_1. This compares with the unstable particle interpretation of the complex poles as discussed in UP. Case (iv) introduces further possible poles in the same manner as for the two-particle unphysical sheet.

On summing the partial wave expansion we expect these properties to be maintained. A remark about the convergence of this expansion is in order at this point. The presence of the single-particle poles and anomalous threshold branch points in the physical region of the momentum transfer variables precludes any possibility of convergence at complex points. As causality specifies the manner in which we must integrate around these singularities, the individual partial wave amplitudes are well-defined and tend uniformly to zero on any compact set in the energy variables in a region of analyticity. An adaptation of a classical theorem for power series

[17] E. T. Whittaker and G. N. Watson, *Modern Analysis* (Cambridge University Press, New York, 1927), 4th ed., p. 213.

[18] The idea of "pinching" of coincident singularities is described in papers of J. C. Polkinghorne and G. R. Screaton, Nuovo Cimento **15**, 289, 925 (1960). See also lecture notes by R. J. Eden, University of Maryland Technical Report 211 (1961).

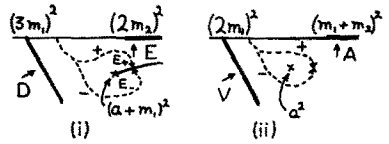


FIG. 3. Paths of continuation to unstable particle cuts, (i) W^2-plane, (ii) two-particle channels.

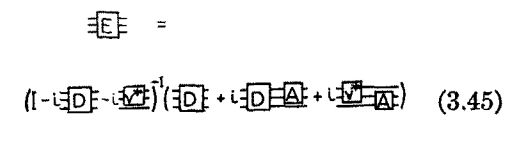
expansions[19] then shows that the analyticity at points other then the singular points mentioned requires that the partial wave expansion converges uniformly on any compact set of the physical region not containing any of the singular points. However, the converse of this does not hold in general, uniform convergence implying only continuity of the sum function. A treatment of analyticity in momentum transfer variables not using partial wave expansions will be given in Part III.

Having in this manner established the existence of solutions to (3.6b), we can now revert to more formal methods for the determination of the nature of the new singularities. We must of course still follow the above order of treatment of the various cuts so as to ensure validity of the procedure.

3.4 UNITARITY RELATIONS ON UNPHYSICAL SHEETS

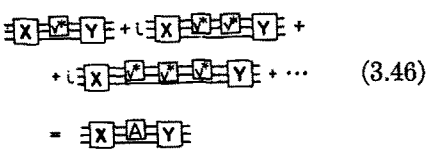
The first case that we treat is the calculation of the discontinuity across the cut attached to the threshold branch point at the location given by (3.44). Starting from the original equation (3.6b), we perform formal continuations to points lying on opposite sides of the unstable particle cut, in the manner of Fig. 3.

The formal solution of (3.6b) is



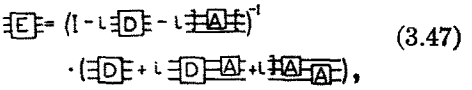
(3.45)

As we must treat the two-particle terms first, we partially sum the formal series expansion of this expression, summing over all sets of terms of the type



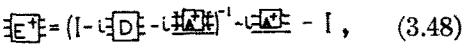
(3.46)

for all fixed arbitrary sequences X and Y of the D and V^* amplitudes. Thus (3.45) may be written

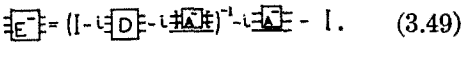


(3.47)

where the sign] [indicates that all terms of the form of the rhs of (3.46), in which the two-particle A amplitude is replaced by a chain of two or more, are omitted from any formal product. The two formal continuations depicted in Fig. 3 can now be written

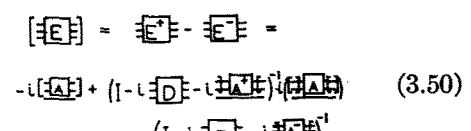


(3.48)



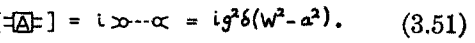
(3.49)

Subtracting, the expression for the discontinuity becomes



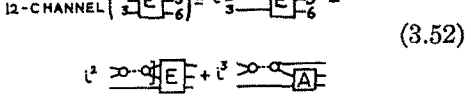
(3.50)

Following our previous assumption that the two-particle amplitude possesses a pole at $s_{12} = a^2$ on the two-particle unphysical sheet, we use the residue to define the single-particle unitarity term (S-wave pole only)



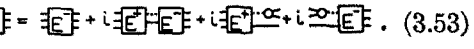
(3.51)

We introduce the following definition of the unstable particle amplitude



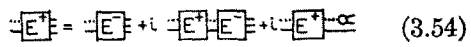
(3.52)

in terms of the residue at the appropriate pole on $s_{12} = a^2$. The right-hand expression is obtained from (3.47). Equation (3.50) then takes its final form, on using (3.48), (3.49), and (3.52),

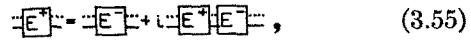


(3.53)

From the basic relation (3.53) we may readily derive further unitarity relations, e.g.,

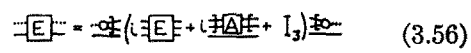


(3.54)



(3.55)

where



(3.56)

is derived in a similar manner to (3.52).

[19] P. Dienes, *The Taylor Series* (Oxford University Press, New York, 1931), p. 467.

The entire process carried out above can be carried out in reverse. If we take as given the three-particle E amplitude and its continuation into the appropriate unphysical sheet together with the unstable particle branch cut and its attendant unitarity relation (3.53), then we may show that the corresponding continuation of the three-particle D amplitude is analytic in the same region. In addition, the unstable particle branch point disappears, being canceled out by a further contribution from the unstable particle pole.

4. CUTS IN HIGHER UNPHYSICAL SHEETS AND APPLICATIONS TO STABLE PARTICLE CUTS

On carrying through the above program for the basic four-particle integral equation relating the four-particle V and C amplitudes, we find that both of the expected cuts at $(a + 2m)^2$ and $(2a)^2$ are produced, with appropriate discontinuity equations. However, the algebra becomes much more lengthy and is relegated to Appendix B. We can see no obstacle (other then complexity) to treating the general case in the same way, but we have not discovered a suitable generating function formalism for treating all the n-particle equations simultaneously.

The important feature which emerges is how the discontinuities across the higher cuts are built up in bits and pieces from the lower-order cuts in minor energy channels so as to give precisely the expected unitarity condition when we adopt the pole-residue definition for scattering amplitudes involving unstable particles. However, considering the over-all collection of threshold singularities and cuts associated with at least one unstable, particle of mass a we may ask whether this interrelation of discontinuities is the most general one consistent with unitarity. It seems that the answer must be negative, for, if we ignore for the moment any violations of causality, we may allow extra contributions to the discontinuities in the form of cuts which are already present *in the physical sheet.* For example, consider the presence of an extra cut in the lower half-plane of the three-particle D amplitude of (3.45) with discontinuity

(4.1)

expressed in terms of given amplitudes involving two new particles denoted by the zigzag and dashed lines. We find as before that unitarity is maintained for the corresponding cut in the continuation of the three-particle E amplitude (assuming that these new amplitudes satisfy the usual unitarity equations with

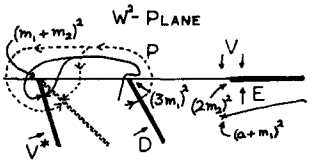


FIG. 4. Illustrating the problem of pion–nucleon cut on three-nucleon unphysical sheet.

stable particle intermediate states and hence that continuations onto other sheets can be defined). In addition we find that on higher sheets the expected threshold singularities from states with one or more particles added occur. With the correct definition of the component amplitudes, unitarity is again satisfied.

This latter type of singularity structure can be ruled out by the causality condition except for the case when the new particles are stable. However, in this case they can only be particles already appearing in the original unitarity equations and the question arises: what are the discontinuities of the stable particle thresholds on unphysical sheets attained by going round a suitable branch point? In particular, we may ask for the discontinuity of the nucleon + meson cut (dotted) of the three-nucleon amplitude on the unphysical sheet reached by following the path P in Fig. 4.

Of course, if we were given the information that there was a meson pole term with appropriate coupling constant on the physical sheet of the two-nucleon scattering amplitude, then there would be no cut on this unphysical sheet. To see this, we may imagine the unstable particle cut being transported up onto the physical sheet so as to coincide with the nucleon + meson cut and in the previous analysis replace unstable particles by mesons. We do not wish to introduce a direct postulate of the existence of meson poles in our theory and so we consider alternative possibilities which are equivalent to such a postulate. In Part II, we show that the absence of the nucleon + meson cut in the unphysical sheet discussed above is actually equivalent to the presence of the meson bound-state pole with correct residue. We prefer this form of the bound state postulate as being more qualitative in nature and avoiding difficulties over what we mean by the "correct residue" at a bound-state pole.

It is an important question to investigate whether there is any physical restriction which limits the possibilities to the above, thus making a separate postulate unnecessary. Until a satisfactory investigation of the existence and uniqueness of solutions of a complete dynamical theory can be carried

through, we cannot say with any confidence that other possibilities are excluded. However there are still some plausibility arguments which may be employed.

The simplest alternative occurs when no meson pole terms are present in nucleon–nucleon scattering, in which case the nucleon + meson cut is entirely self-generating as we pass from sheet to sheet. Consistency arguments then show that a bound-state meson pole in the meson–meson amplitude would induce a nonvanishing pole in the nucleon amplitude, contradicting our hypothesis. This results in much weaker restrictions being placed by unitarity on the coupled set of meson scattering equations in that the amplitudes with even and odd numbers of mesons are no longer related by a bound-state pole decomposition law of the type dicussed in Part II. In fact, it would seem perfectly consistent if all the amplitudes involving an odd meson number now vanished, as though a selection rule of the type introduced for the nucleons were operating. As we require this not to happen, it seems plausible that there is associated at least one extra parameter with the set of odd meson number amplitudes, as compared to the original case. The primitive picture that we have at the moment of a complete dynamical S-matrix calculation suggests that the number of parameters allowed is severely restricted. At this point we may consider an analogy with quantum field theory in the perturbation solution, in which only a very limited number of couplings are allowed if the theory is to be renormalizable. It is usually held that nonrenormalizable interactions involve one in an infinity of undetermined parameters, which is the reason for excluding them from consideration. In S-matrix theories, the corresponding restrictions are suggested by the work of Martin,[20] in the matter of the uniqueness of a two-particle amplitude with a given spectral function under the hypothesis of polynomial boundedness at infinity (and more specifically, the validity of a Mandelstam representation). The very powerful restrictions which seem implicit in results such as these and those as yet unknown extensions to many-particle amplitudes will presumably have the effect on strictly limiting the numbers of free parameters. Thus it may be possible to exclude the above example on such a basis, as probably requiring the specification of an infinite number of parameters in any complete theory based on it. This assumes that uniqueness theorems no longer hold when we drop the restriction of polynomial boundedness. We note that, for the nucleons, the selection rule that we have imposed consistently eliminates all odd-nucleon amplitudes as well as the three-nucleon vertex, and so this case may be expected not to suffer from the same parameter trouble. In fact, the two-nucleon branch point is absent on the unphysical sheet reached round the *four-nucleon* branch point, without the introduction of any extra hypotheses. This can be shown by appropriately modifying the methods of Appendix B for the case of the single nucleon pole in the main channel of a three-particle scattering amplitude.

Other cases in which the nucleon + meson cut is partly self-generating and partly induced by a meson pole are more difficult to discuss though it seems that the remarks made in the previous paragraph still apply.

[20] A. Martin, Phys. Rev. Letters 9, 410 (1962).

5. K-MATRIX AND RELATED METHODS

The usefulness of the K matrix[21] for the study of analytic properties of the elastic branch point of a two-particle scattering amplitude has been indicated by Zimmermann.[22] The general definition is

$$S = \frac{I + iK/2}{I - iK/2}, \qquad K = 2i\frac{I - S}{I + S},$$
$$T = K + iTK/2 = K + iKT/2, \qquad S = I + iT, \tag{5.1}$$

for which the K amplitude $K_{\mu,\nu}$ is defined by the normal product expansion (2.5) with $(S - I) \to iK$ and $W_{\mu,\nu} \to K_{\mu,\nu}$. From (5.1) and these formal expansions, we may derive the equations which define the $K_{\mu,\nu}$ amplitudes in terms of the $W_{\mu,\nu}$ amplitudes in the same manner as the general unitarity equations (2.7).

5.1 Analytic Continuation of the K-Matrix Elements and Interpolating Sequences.

Let us carry through the procedure introduced in Sec. 3.1 for the K-matrix equation (5.1), by defining a set of "interpolating" K matrices. Using the same type of notation, these are defined by equations like

$$T = K_\nu + iK_\nu P_\nu T/2$$
$$= K_\nu + iTP_\nu K_\nu/2 : K_n = K : K_0 = T, \tag{5.2}$$

satisfying

$$K_{\nu_1} - K_{\nu_2} = -iK_{\nu_1}(P_{\nu_1}(I - P_{\nu_2}) - (I - P_{\nu_1})P_{\nu_2})K_{\nu_2}/2 \tag{5.3}$$

[21] B. A. Lippmann and J. Schwinger, Phys. Rev. 79, 499 (1950).

[22] W. Zimmermann, Nuovo Cimento 23, 249 (1961). Also R. Oehme, *ibid.* 20, 334 (1961).

$$K_\nu - K_\nu^* = iK_\nu(I - P_\nu)K_\nu^* = iK_\nu^*(I - P_\nu)K_\nu. \tag{5.4}$$

On taking matrix elements of the latter equation, we obtain an iterative sequence of integral equations for the interpolating amplitudes, for which the problem of analytic continuation can be studied by the methods of Sec. 3.3. However, unlike S-matrix elements, we find that these continued functions are not analytic continuations of each other and we now have a different function for every subregion of analyticity of the physical region. The vacuum and one-particle intermediate states are particularly important, and we separate these first, writing the equations in the symmetrized form which is applicable to all channels (cf. 2.7).

5.2 Treatment of Vacuum and One-Particle Intermediate States.

Let us define the symmetrized interpolating amplitudes $K_n^{(0)}$ (taking a single mass theory for simplicity) by

$$K_n^{(0)}(k_1, \cdots, k_n) = W_n(k_1, \cdots, k_n) - \frac{i}{2}\sum_{\nu=1}^{n-1} P\left(\alpha; \frac{\nu}{n}\right) K_\nu(k_{\alpha_1}, \cdots, k_{\alpha_\nu}) \cdot W_{n-\nu}(k_{\alpha_{\nu+1}}, \cdots, k_{\alpha_n})\, \delta^{(4)}(k_{\alpha_1} + \cdots + k_{\alpha_n}). \tag{5.5}$$

This expression is identical in form to equations for the removal of vacuum singularities as treated by Zimmermann[23] and Watanabe[24] and so we naturally identify the $K_n^{(0)}$ as the V_n amplitudes of Sec. 2. The absence of vacuum singularities for this case in any channel can be demonstrated by direct substitution in the appropriate form of the symmetrized unitarity relation (2.7), when all terms containing disconnected parts cancel out. Note that the $K_n^{(0)}$ as defined are not present in the crude interpolating sequences defined in (5.3), but only in certain refinements.

The one-particle equations now take the form

$$K_n^{(1)}(k_1, \cdots, k_n) = K_n^{(0)}(k_1, \cdots, k_n) - \frac{i}{2}\sum_{\nu=1}^{n-1} P\left(\alpha; \frac{\nu}{n}\right) \cdot \int d^4q\, \theta(q^0)\, \delta(q^2 - m^2)\, \delta^{(4)}(k_{\alpha_1} + \cdots + k_{\alpha_\nu} - q) \cdot K_{\nu+1}^{(0)}(k_{\alpha_1}, \cdots, k_{\alpha_\nu}; -q) K_{n-\nu+1}^{(1)}(k_{\alpha_{\nu+1}}, \cdots, k_{\alpha_n}; q), \tag{5.6}$$

defining the next symmetrized interpolating amplitude $K_n^{(1)}$. After carrying out the integration over the δ functions and combining terms which differ only in the argument of the θ functions, we see that the $K^{(1)}$ amplitudes contain no single-particle singularities of the type $\delta((k_{\alpha_1} + \cdots + k_{\alpha_\nu})^2 - m^2)$ in the physical regions, assuming that the $K_n^{(0)}$ possess the pole terms indicated by the physical region unitarity relations. As before, this is demonstrated by substitution into (2.7) and observing the cancellation of all single-particle terms.

In (5.5) and (5.6) we have the first expressions in what is a symmetrized version of the interpolating sequences (5.3), but which seem less useful for many-particle states. In addition, it does seem possible to express it in terms of operator equations and so we treat only the cases (5.5) and (5.6), in which no trouble over the existence of solutions arises.

APPENDIX A: KINEMATICS OF MANY-PARTICLE PROCESSES.

As we go to processes involving increasingly more external particles from the two-particle elastic case, the number of useful choices of a set of independent Lorentz invariant variables soon becomes embarrassingly large. The most convenient choice depends on the particular problem being considered, but we can make several general remarks about them.

In general, if we have i incoming and j outgoing particles, $N = i + j$, the number of independent invariants is $3N - 10$ for $N \geq 4$ and 0 for $N < 4$ (for three-dimensional momentum space). We are not treating the external masses as variables in this enumeration. On the other hand, the number of associated processes involving the same total number of external particles is $2^{(N-1)} - N - 1$ and each corresponding "total energy" variable $s_{\alpha_1\alpha_2\cdots} = (k_{\alpha_1} + k_{\alpha_2} + \cdots)^2$ has a claim to be considered as "basic."

The physical regions are connected subsets of the space of invariants, being defined by the condition that all external momenta are real.

It is sometimes convenient to regard the space of invariants as being embedded in the $[\frac{1}{2}N(N-1)]$-dimensional space of the $s_{\alpha\beta}$. The $s_{\alpha_1\alpha_2\cdots\alpha_r}$ for $r > 2$ can always be expressed as linear combinations of the $s_{\alpha\beta}$ and external masses. Before imposing energy–momentum conservation and four-dimensionality of Lorentz space, we must regard the k_α as N dimensional. Energy–momentum conservation then imposes N independent linear conditions. On successively reducing the dimensions of momentum space to 4, we are imposing 1, 2, 3, $\cdots$, $(N - 5)$ independent nonlinear conditions successively, which may be written as the vanishing of the Gram determinants of some of the k_α. The resulting space

[23] W. Zimmermann, Nuovo Cimento, **13**, 503 (1959).
[24] I. Watanabe, Progr. Theoret. Phys. (Kyoto) **4**, 371 (1953).

is thus a complex algebraic variety, almost everywhere of dimension $3N - 10$. The nonlinear conditions affect only processes involving six or more external particles.

The physical regions are conveniently dissected into subregions bounded by the surfaces

$$s_{\alpha_1\alpha_2\cdots} = 4m^2, 9m^2, 16m^2, \cdots \qquad \text{(A1)}$$

(for the equal mass case). Here $\alpha_1\alpha_2 \cdots$ runs over all possible channels. These subregions are denoted collectively by $J_{(\beta)}$ where (β) is a label specifying the number of particles in the highest mass states in every channel for which there is sufficient energy to see it as a real state for a process occurring in the region considered. The set of all $J_{(\beta)}$ then possesses a natural ordering ">" defined by: $J_{(\beta)} > J_{(\beta')} \leftrightarrow (\beta) > (\beta')$, where the latter inequality holds when it is satisfied in the usual sense for integers of the sets (β), (β') in every channel.

Consider the process depicted in Fig. 5, in which we have inserted p intermediate particles.

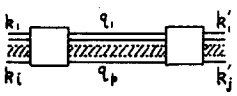


FIG. 5. Unitarity diagram with p intermediate particles.

We work in the c.m. system and consider how many invariants are required to specify completely a "configuration" of incoming, intermediate, or outgoing particles. For example, the i 3-vectors $\mathbf{k}_\mu$ are specified completely as to their relative orientation by $(3i-7)$ invariants in conjunction with the total energy square $W^2 = (k_1 + k_2 + \cdots + k_i)^2$. The total of $3(i + j) - 10$ is then made up from W^2, $(3i - 7)$ "incoming" invariants, $(3j - 7)$ "outgoing" invariants and three other invariants specifying the relative orientation of the incoming and outgoing configurations.

The problem of transforming integral equations of the type (3.36) into Fredholm form is simplified by the introduction of partial-wave amplitudes. These may be regarded as a straightforward generalisation of the procedure followed in UP for two-particle intermediate states. We work with an arbitrary dimensionality of the momentum space for the reasons discussed in Sec. 2. When the dimensionality is sufficiently high, our choice of independent invariants for the $(i \to j)$ particle amplitude becomes independent of the dimension and the previous prescription is changed to

$W^2 = (k_1 + k_2 + \cdots k_i)^2$

$(i+1)(i-2)/2$ other "ingoing invariants," $i > 3$,

$(j+1)(j-2)/2$ other "outgoing invariants," $j > 3$,

$(i - 1)(j - 1)$ invariants specifying the relative orientations of the "ingoing" and "outgoing" configurations.

The total is $\frac{1}{2}N(N - 3)$.

We now consider the amplitude, for fixed and physical values of the ingoing and outgoing invariants as an analytic function defined on the underlying manifold of the rotation group R_τ of the $\tau[=(\sigma - 1)]$-dimensional Euclidean space. Consider a complete set of unitary irreducible representations of this group, denoted by $D_{(\alpha)}(\rho)$, $\rho \in R_\tau$, satisfying the usual orthogonality and completeness relations

$$\int_{\Omega_\tau} d\rho\, D^{m_1,m_1'}_{(\alpha_1)}(\rho)\, D^{*m_2,m_2'}_{(\alpha_2)}(\rho) = \frac{\delta_{(\alpha_1)(\alpha_2)}\,\delta_{m_1m_2}\,\delta_{m_1'm_2'}}{d_{(\alpha_1)}}$$

$$0 < m, m' < d_{(\alpha)} - 1 \qquad \text{(A2)}$$

in which d is the (finite) dimension of the representation and $d\rho$ is the normalized invariant measure on the group manifold. The symbol (α) denotes a set of $[\tau/2]$ numbers specifying the irreducible representation.[25] The completeness relation can be expressed as

$$\sum_{(\alpha)} d_{(\alpha)}\, D^{m,m'}_{(\alpha)}(\rho_1)\, D^{*m,m'}_{(\alpha)}(\rho_2) = \delta(\rho_1, \rho_2), \qquad \text{(A3)}$$

in which $\delta(\rho_1, \rho_2)$ is the invariant δ function of two elements of the group satisfying $\delta(\rho_1, \rho_2) = 0$ if $\rho_1 \neq \rho_2$ and $\int_{\Omega_t} d\rho_1\, \delta(\rho_1, \rho_2) = 1$ (interpreted as Schwartz distribution equations). The group property is expressed in

$$D^{m,m'}_{(\alpha)}(\rho_3) = \sum_{m''=0}^{d(\alpha)-1} D^{m,m''}_{(\alpha)}(\rho_2)\, D^{m'',m'}_{(\alpha)}(\rho_1) \quad \rho_3 = \rho_2\rho_1. \qquad \text{(A4)}$$

A convenient parametrization of the group, the Euler angle parametrization, is defined by the following factorization of a group element into elementary rotations

$$\rho = T_{12}(\alpha_{12})T_{23}(\beta_{23})T_{12}(\beta_{12})T_{34}(\gamma_{34})T_{23}(\gamma_{23}) \cdots T_{(\tau-1)\tau} \cdot(\mu_{(\tau-1)\tau}) \cdots T_{23}(\mu_{23})T_{12}(\mu_{12}), \qquad \text{(A5)}$$

$$0 \leq \alpha_{12}, \beta_{12}, \cdots, \mu_{12} < 2\pi, \quad 0 \leq \beta_{23}, \text{ etc.}, \leq \pi.$$

$T_{pq}(\theta)$ is the rotation $x \to x'$, defined by

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_p \\ x_q \end{bmatrix} = \begin{bmatrix} x_p' \\ x_q' \end{bmatrix} \quad \text{and} \quad x_r = x_r' \quad \text{for} \quad r \neq p, q$$

[25] F. D. Murnaghan, *The Theory of Group Representations* (Johns Hopkins Press, Baltimore, Maryland, 1938).

of the τ-dimensional vector x. The invariant measure in this case can be calculated by standard methods[26] to give

$$d\rho = \left(\prod_{\epsilon=1}^{\tau-1} S_\epsilon\right)^{-1} d\alpha_{12}(\sin\beta_{23}\, d\beta_{23}\, d\beta_{12}) \cdot(\sin^2\gamma_{34}\, d\gamma_{34} \sin\gamma_{23}\, d\gamma_{23}\, d\gamma_{12})(\sin^3\gamma_{45}\cdots d\mu_{12}), \tag{A6}$$

where S_ϵ is the "content" of the ϵ-dimensional sphere of unit radius: $S_1 = 2\pi$, etc. Using fundamental theorems for functions defined on the rotation groups,[27,28] we assert the existence of an expansion of the $(i \to j)$ particle amplitude in terms of element of these representations, which is uniformly convergent on any compact subset of the group manifold on which the amplitude is holomorphic and except for singularities associated with the particular parametrization used. This is our generalized partial wave expansion.

Consider next a typical integration over intermediate states of the form

$$\int\cdots\int \prod_{\epsilon=1}^{p} \{d^\sigma q_\epsilon \theta(q_\epsilon^0)\ \delta(q_\epsilon q_\epsilon - m^2)\} \cdot\delta^{(\sigma)}((W, \mathbf{0}) - q_1 - q_2 - \cdots - q_p), \tag{A7}$$

where we leave understood the two amplitudes normally appearing in the integrand, as the integration over the δ functions has only the effect of picking out mass- and energy-shell values of these amplitudes.

A change of variables of the form

$$d^\sigma q_\epsilon \to |\mathbf{q}_\epsilon|^{\sigma-3}\,(4W)^{-1}\, d(q_\epsilon)^2\, d(\lambda_\epsilon)^2\, d\Omega_\epsilon, \tag{A8}$$

where $d\Omega$ is an element of solid angle in τ-dimensional momentum space, $|\mathbf{q}_\epsilon|$ the τ-momentum and $\lambda_\epsilon = (W, \mathbf{0}) - q_\epsilon = q_1 + q_2 + \cdots + q_{\epsilon-1} + q_{\epsilon+1} + \cdots + q_p$, gives

$$(4W)^{p-1}\int\cdots\int d\Omega_1\cdots d\Omega_{p-1}\, d(\lambda_1)^2\cdots d(\lambda_{p-1})^2 \cdot(|\mathbf{q}_1|\cdots|\mathbf{q}_{p-1}|)^{\sigma-3} \cdot\delta^{(\sigma)}(((W, 0) - q_1 - \cdots - q_{p-1})^2 - m_1^2), \tag{A9}$$

after carrying out some of the δ-function integrations. A further change of variables is then made:

$$d\Omega_1\cdots d\Omega_{p-1} \to d\Omega_1\, d\Omega_{12}\, d\Omega_{123}\cdots d\Omega_{12\ldots(p-1)}, \tag{A10}$$

in which the integration over unit spheres in momentum space, all expressed initially in a coordinate system fixed with respect to, say, the orientation of the ingoing particle configuration, are progressively made dependent on the directions of the intermediate particles in some specified order. For example, $d\Omega_{123}$ is an element of solid angle of the direction of $\mathbf{q}_3$, measured in a coordinate system fixed with respect to the orientations of $\mathbf{q}_1$ and $\mathbf{q}_2$ and a further set of orthogonal vectors independent of the orientations of the remaining $\mathbf{q}_\epsilon$. In a suitable spherical coordinate system, this element of solid angle is

$$d\Omega_{123} = \sin^{\tau-1}\theta_1\, d\theta_1 \sin^{\tau-2}\theta_2\, d\theta_2 \sin^{\tau-3}\lambda_3 \cdots \sin\lambda_{\tau-2}\, d\lambda_{\tau-2}\, d\lambda_{\tau-1}, \tag{A11}$$

where the components of $\hat{\mathbf{q}}_3$ in Cartesian coordinates are

$$\begin{aligned} &\cos\theta_1 \\ &\sin\theta_1\cos\theta_2 \\ &\sin\theta_1\sin\theta_2\cos\lambda_3 \\ &\quad\cdots \\ &\sin\theta_1\sin\theta_2\cdots\sin\lambda_{\tau-2}\sin\lambda_{\tau-1}. \end{aligned} \tag{A12}$$

In (A11) we have now two types of angles: θ_1 and θ_2 are angles defined entirely by the relative orientation of the intermediate particle momenta q_1, q_2, and q_3, whereas the angles $\{\lambda\}$ form part of an Euler angle parametrization of a rotation from a coordinate system fixed with respect to the incoming configuration of c.m. momenta into one fixed with respect to the intermediate configuration. The integrations over directions in the momentum spaces of intermediate particles can thus be replaced by the product of two forms

$$\prod_{\epsilon=1}^{p-1} d\Omega_\epsilon \to d\rho_1\, d\mu. \tag{A13}$$

Here $d\rho_1$ is an invariant measure of the τ-dimensional rotation group, formed from the angles λ and, if necessary, a redundant measure involving fictitious intermediate particle to take into account an excess of τ over $(p - 1)$. The second factor $d\mu$ is part of the usual measure of phase space of the p-particle intermediate state, the remainder being $d(\lambda_1)^2\cdots d(\lambda_{p-1})^2$, restricted by the δ function in (A9).

Consider first the factor $d\rho_1$, which refers to rotations between coordinate systems fixed in the ingoing and intermediate particle configurations. When we project out partial-wave amplitudes from (A7) we find that each projection can be expressed

[26] E. P. Wigner, *Group Theory and its Application to the Quantum Theory of Atomic Spectra* (Academic Press Inc., New York, 1959).

[27] C. Chevalley, *Theory of Lie Groups I* (Princeton University Press, Princeton, New Jersey, 1948).

[28] L. Pontrjagin, *Topological Groups* (Princeton University Press, Princeton, New Jersey, 1939).

in a form involving only integration over intermediate phase space, as demonstrated below.

Partial-wave projection is carried out by multiplication with a $D^{m,m'}_{(\alpha)}(\rho_3)$ and integration over the group manifold of measure $d\rho_3$. We can now change the variables of integration in the manner $d\rho_1\, d\rho_3 \to d\rho_1\, d\rho_2$, because ρ_3, a rotation between ingoing and outgoing configurations, can be factorized into $\rho_1\rho_2$, a product of two rotations between ingoing and intermediate and between intermediate and outgoing configurations, respectively. On substituting the representation equation (A4), both integrations may be performed immediately to project out partial wave amplitudes from the suppressed amplitudes in the integrand. Apart from a numerical factor, the partial-wave projection of the integral then becomes, on restoring the suppressed amplitudes

$$\sum_{m''=0}^{d(\alpha)-1} \int \cdots \int d\chi \{\text{amplitude for } p \to j \text{ process}\}^{m,m''}_{(\alpha)} \cdot \{\text{amplitude for } i \to p \text{ process}\}^{m'',m'}_{(\alpha)}. \qquad \text{(A14)}$$

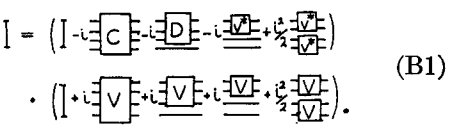
We now require two properties of the intermediate phase space; (i) its element of volume $d\chi$ as defined above and (ii) the limits of the region of integration. The latter is determined by the requirements that momenta be positive real in magnitude, angles be real and that the condition imposed by the remaining δ-function in (A9) holds. It is possible to give simple formulas only in the lowest particle number cases, e.g., for $p = 3$, dimensionality ≥ 3, equal mass case, phase space is given by

$$\begin{aligned} &s_{12}s_{13}s_{23} \geq m^2(W^2 - m^2)^2, \\ &s_{12} + s_{13} + s_{23} = W^2 + 3m^2, \\ &s_{12} = (\lambda_3)^2 \geq 4m^2, \quad \text{etc.}, \end{aligned} \qquad \text{(A15)}$$

and for dimensionality 3 only: $d\chi = ds_{12}\, ds_{23}$.

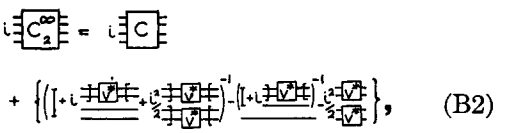
APPENDIX B: UNITARITY CONDITIONS FOR SOME UNSTABLE PARTICLE CUTS.

Using the notation of Sec. 3, the interpolating unitarity condition for the four-nucleon scattering amplitude involving only four-nucleon intermediate states in the main channel, is given by



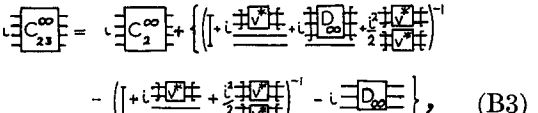
(B1)

The generalisation of the procedure used in Sec. 3 to extract an equation representable in Fredholm form follows the sequence:



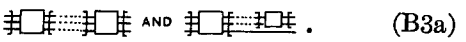
(B2)

in which all connected diagrams build up entirely from two-nucleon V^* amplitudes are eliminated,



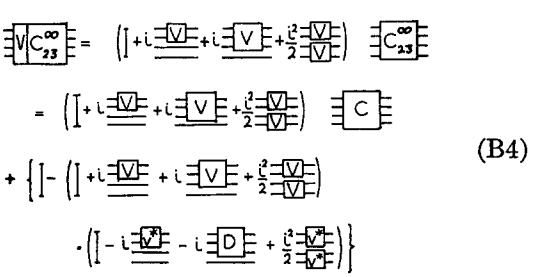
(B3)

in which all connected diagrams involving the two-nucleon V^* amplitudes and at least one three-particle D_∞ amplitude are eliminated. A "]" sign crossing *three* particle lines forbids all products of the type



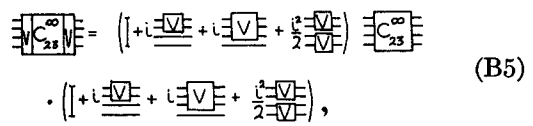
(B3a)

The next steps are



(B4)

and



(B5)

in which the two- and three-particle terms of the minor energy channels are eliminated.

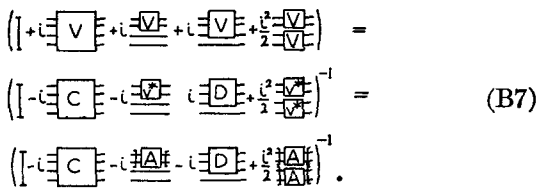
From these definitions and (B1) we deduce the following pure four-particle equation

(B6)

in analogy to (3.33). This equation can now be treated by the Fredholm methods of Sec. 3.3 so as to demonstrate the existence of the continuations to unstable particle thresholds at $W^2 = (a + 2m_1)^2$ and $(2a)^2$.

In order to derive the unitarity conditions, it is simpler to return to the original expression (B1) and use the formal techniques of Sec. 3.4.

By partial summation, it can be shown that



(B7)

Consider first the $(a + 2m_1)^2$ threshold branch point. We follow the paths of continuation A and B depicted in Fig. 6. These paths are consistent with the linear relations between the invariants in the way the various singularities are bypassed, but the choice is not unique for given paths in the W^2 plane. However, the other possible choices all give appropriately different results and so we consider only this particular one.

We must now consider what happens to the four-particle intermediate phase space contour. This is a complicated surface in a five-dimensional space (cf. Appendix A), but the integrations in three of the variables can be carried out without introducing the singularities of interest as we need only consider a term containing poles in only say the s_{12} and s_{34} planes at once, as in Fig. 7.

The remaining part of the phase space takes the form (shaded) in Fig. 8, which has some freedom of deformation, the three edges being tied down to

$$\begin{array}{lll} \text{EF:} & \text{EG:} & \text{FG:} \\ s_{12} = 4m^2 & s_{34} = 4m^2 & s_{12}^{\frac{1}{2}} + s_{34}^{\frac{1}{2}} = W \end{array} \qquad \text{(B8)}$$

and the three corners to

$$\begin{array}{lll} \text{E:} & \text{F:} & \text{G:} \\ s_{12} = 4m^2 & s_{12} = 4m^2 & s_{12} = (W - 2m)^2 \\ s_{34} = 4m^2 & s_{34} = (W - 2m)^2 & s_{34} = 4m^2. \end{array}$$

As in Eq. (6.5), the integrand possesses square-root branch points on the edges of the region of integration, which arise from the integrations over the three other independent variables, these having been

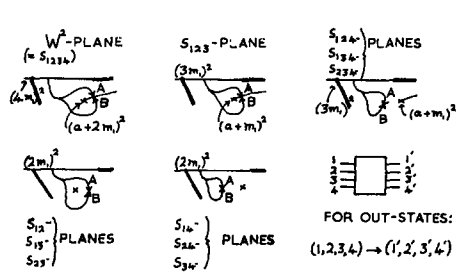


FIG. 6. Paths of continuation.

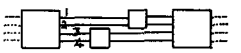
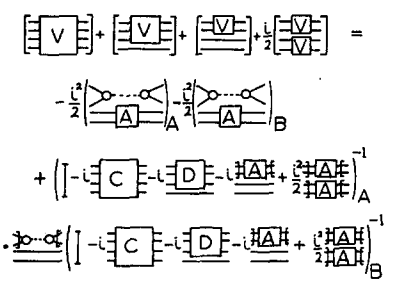


FIG. 7. Unitarity diagram giving rise to unstable particle cuts.

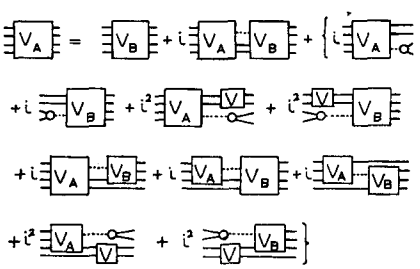
carried out first. As before, going once round one of these branch points just changes the sign of the integrand. We see that the $(a + 2m_1)^2$ branch point occurs when the corners F and G hit the poles at $s_{12} = a^2$ and $s_{34} = a^2$. In each term of the formal expansion of the final expression in (B7) the discontinuity arising from following the two paths A and B, which forces the contour opposite ways round the poles, can be evaluated and summed to give



(B10)

in which square brackets $[\cdots]$ denote a discontinuity and the suffices A and B denote the path on which the inverses are evaluated.

Using (B7) again on (B10) and noticing that the disconnected parts cancel, we get



(B11)

being interpreted as follows: the terms in curly brackets are present only if consistent with the original paths of continuation chosen (Fig. 6), the amplitudes involving unstable particles are obtained

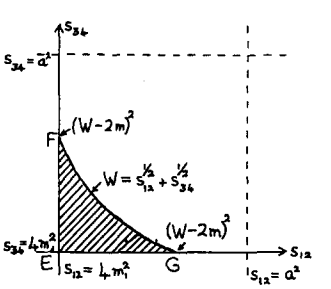
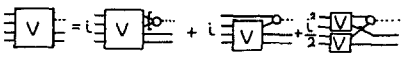


FIG. 8. Four-particle phase space.

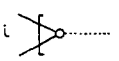
from the pole-residue definition [cf. (3.52)], which takes the form



(B12)

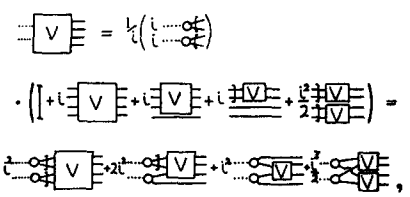
on applying to (B7). The logarithmic nature of this three-particle branch point can be obtained directly from the Fredholm solutions by a more detailed analysis of the coinciding poles and integration contour.

On using the pole-residue definition for amplitudes involving unstable particles, we find the following prescription for its expansion following from the unitarity formulas: write down the S-matrix element as a sum of V amplitudes, contract with a sufficient number of the vertex constants

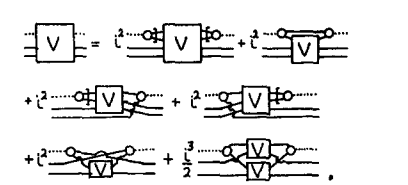


(B12a)

on either side and pick out the connected terms so produced. For example, we have



(B13)



(B14)

We next treat the $W^2 = (2a)^2$ singularities. It is not difficult to visualise the possible motions and deformations of the contour of integration in Fig. 8. This can be used to show that the $W^2 = (2a)^2$ singularity arises from the coincidence of the leading edge of the contour on $W = s_{12}^{\frac{1}{2}} + s_{34}^{\frac{1}{2}}$ with both poles simultaneously. This occurs whichever way we go round the $W^2 = (a + 2m_1)^2$ branch point initially, and so we expect branch points on *both* sheets.

Taking case (a) first (see Fig. 9) and the path A^+,

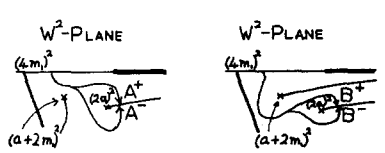
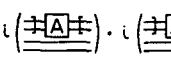


FIG. 9. Paths of continuation for the singularities at $W^2 = (2a)^2$.

the contour of integration lies "above" all the poles right up to the endpoint. For the path A^-, however, though the corners F and G pass above the poles as we increase W, we require the contour to pass below the point $s_{12} = s_{34} = a^2$. It is necessary to slit the contour slightly in order to do this, and then the essential difference between the two contours is a bicylinder enclosing both poles. For the case (b) we have a similar construction. The form of the residues themselves is obtained as for (B10). The coefficient attached to a

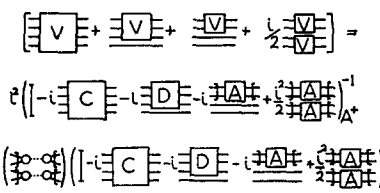
(B14a)

factor in any term of the expansion of (B10) is $\frac{1}{2}i^2$, of which $-\frac{1}{2}i^2$ arises from that term in (B10) itself and i^2 from the appropriate term in the product [cf. (3.13)]



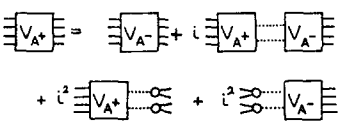
(B14b)

The discontinuity between the values reached along the paths A^+ and A^- is



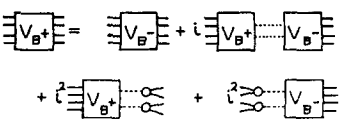
(B15)

Treating in the same way as (B10), we get finally



(B16)

Similarly for the B^+ and B^- paths, we get



(B17)

Notice that in evaluating the double residue to obtain (B15), a factor 2 appears which cancels with the factor $\frac{1}{2}$ in the coefficient discussed above. A check on our result is obtained by carrying out our continuation procedure on the two-nucleon + unstable particle cuts in the various amplitudes using the unitarity relations as derived from (B11).

Unitarity and On-Mass-Shell Analyticity as a Basis for S-Matrix Theories. II

J. GUNSON
Department of Mathematical Physics, University of Birmingham, Birmingham, England
(Received 20 January 1964; final manuscript received 30 June 1964)

Sets of postulates for S-matrix theories are given and used to construct potentially complete dynamical theories for scattering of strongly interacting particles, based essentially on the unitarity relations and on-mass-shell analyticity of scattering amplitudes. In one formulation, the unitarity relations for stable particles are used to derive the corresponding relations for cuts associated with unstable particle states. In a second formulation, an attempt is made to treat stable and unstable particles on a more equal footing. Crossing symmetries and decomposition laws for single-particle poles appear as consequences of the postulates.

ON the basis of results obtained in the first paper of this series (hereafter denoted by I), we attempt below to construct sets of axioms for complete dynamical theories of strongly interacting particles. For the reasons stated in the introduction to I, we wish to do this entirely in terms of on-mass-shell analyticity. This immediately plunges us into all the usual difficulties associated with unphysical regions, but which we regard as not insuperable by on-mass-shell methods.

The first step is to obtain a general prescription for obtaining all the singularities of scattering amplitudes which are allowed by the maximal analyticity postulate. The singularities occurring in the physical regions we have already obtained directly from the unitarity relations. The most important unphysical region singularities are the one-particle poles of two-particle scattering amplitudes, whose existence can be proved in many cases from field theory, making essential use of off-mass-shell continuations.[1,2] In order to obtain these singularities in an on-mass-shell theory, we are forced to start from physical region unitarity and invert the methods of Sec. 3.3 of I in order to deduce the existence of single-particle poles from the discontinuities given by physical region unitarity across certain cuts. We need one additional postulate of the type mentioned at the end of Sec. 3 of I, i.e., we require the nucleon + meson branch point to be absent on the particular unphysical sheet discussed there and on corresponding sheets of the other amplitudes with three-nucleon branch points (cf. Fig. 4 of I). We now claim that the existence of the nucleon + meson cut on the physical sheet with discontinuity given by the usual unitarity condition (3.6a) of I, implies uniquely the existence of the pole terms to be expected if the meson is to be regarded as a bound state of two nucleons.

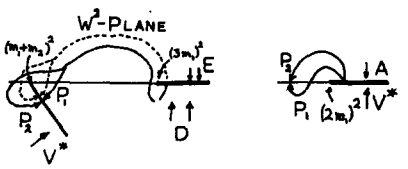
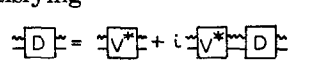
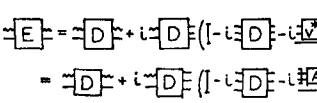


FIG. 1. Paths of continuation for the meson–nucleon cut on physical and unphysical sheets.

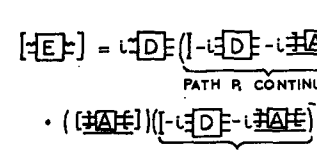
1. PROOF OF THE EXISTENCE OF BOUND STATE POLES.

We start with the meson–nucleon scattering amplitudes satisfying


(1)

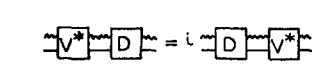

(2)

as obtained from (3.1a, b) and (3.2b) of I. Following the paths P_1 and P_2 of Fig. 1 (cf. Fig. 4 of I) and using the postulated nonexistence of the two-particle cut on the three-particle unphysical sheet, we deduce the expression



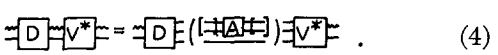
(3)

for the difference between the values attained by following paths P_1 and P_2. But the lhs expression in (3) is just the discontinuity


(3a)

[1] H. J. Bremmermann, R. Oehme, and J. G. Taylor, Phys. Rev. 109, 2178 (1958).
[2] J. Gunson and J. G. Taylor, Nuovo Cimento 15, 806 (1961).

across the two-particle cut of the meson–nucleon scattering amplitude, and so we perform the identification



(4)

Let us consider this expression in each partial wave separately. In order for the rhs to be nonzero, there must be some singularity present in the two-nucleon scattering amplitude so as to give rise to the discontinuity term in (4). We also know certain further properties of the lhs of (4), (i) it is the discontinuity across a square-root type branch point (elastic threshold), (ii) near the threshold $q^2 = W^2 - (m_1+m_2)^2 = 0$, the λth partial wave of the discontinuity behaves like $cq^{4\lambda+1}$, $c \neq 0$ (one factor $q^{2\lambda}$ of this behavior arises from the external particle states and is irrelevant for our discussion).

We now claim that the only singularity in the two-nucleon scattering amplitude which can give this behavior is an S-wave pole at $W^2 = m_2^2$, with an appropriate residue. To see this, we first consider in more detail the integration over three-particle intermediate phase space (cf. Sec. 3.3 of I). This is of the form

$$\int_{4m_1^2}^{(W-m_1)^2} ds_{12} \int_{\alpha+(s_{12},W^2)}^{\alpha-(s_{12},W^2)} ds_{13}\, f(s_{12}, s_{13}, W^2), \qquad (5)$$

where

$$\alpha_\pm(s_{12}, W^2) = \tfrac{1}{2}(W^2 + 3m_1^2 - s_{12}) \pm \tfrac{1}{2}\{[s_{12} - (W + m_1)^2] \cdot [s_{12} - (W - m_1)^2](s_{12} - 4m_1^2)s_{12}^{-1}\}^{\frac{1}{2}},$$

in which we suppose that s_{12} is the energy variable for the two-nucleon amplitude which contains our conjectured singularity. The integration over the s_{13} variable clearly produces a function which possesses square-root branch points on $s_{12} = (W + m)^2$ and $s_{12} = (W - m)^2$. Moreover, the function is simply changed in sign if we encircle one of these branch points, as the only effect of this encirclement is to reverse the direction of integration in the s_{13} variable between the same endpoints. The s_{12} integration can thus be replaced by a contour integral which encircles the $s_{12} = (W - m_1)^2$ branch point, as in Fig. 2. The method of generation of the square-root

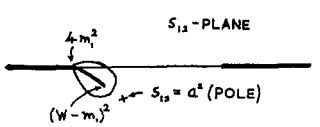


FIG. 2. This figure shows how the phase-space integration contour is pinched, producing a two-particle threshold.

two-particle threshold branch point from a pole at, say, $s_{12} = a^2$, is now clear, using the methods of continuation used extensively for perturbation theory amplitudes. Conversely, given the square-root branch point at $W^2 = (m_1 + m_2)^2$, we see that this must arise from an isolated singularity at $s_{12} = m_2^2$. This singularity cannot have a cut attached to it, as this would not give a square-root branch point. Furthermore, it cannot contain any singular component other than a simple pole, as this would give an unbounded value to the rhs of (4) at the threshold $W^2 = (m_1 + m_2)^2$. We thus deduce the following form for the singularity

$$g^2(\cos\theta)/2\pi(s_{12} - m^2), \qquad (6)$$

where the residue $g^2(\cos\theta)$ may yet depend on the cm scattering angle cosine invariant. Expressing this as a Legendre series (we assume that the analytic properties allow this)

$$\frac{1}{2\pi}\left\{\frac{g_0^2}{s_{12} - m_2^2} + \frac{g_1^2 P_1(\cos\theta)}{s_{12} - m_2^2} + \cdots\right\}, \qquad (7)$$

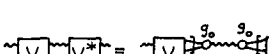
we see that the rhs of (4) reduces to a sum of expressions, one for each integer s, like

(8)

where the patterned line represents an intermediate particle of spin s and mass m_2, associated with the term $g_s^2 P_s(\cos\theta)/2\pi(s_{12} - m_2^2)$. However, such a term as (8) can be shown to possess a threshold momentum dependence expressed as a function of the total angular momentum j, of the form $q^{2\alpha(j)+1}$ (neglecting the q dependence arising from the external states), with $\alpha(j)$ of the form $|j - s|$. The rhs of (4) has the corresponding momentum dependence factor q^{2j+1}. Hence inconsistency rules out all terms in (6) with $s \geq 1$ and we are left with the relation

$$ig_0^2\, \delta(s_{12} - m_2^2) \qquad (9)$$

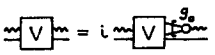
for the discontinuity across the pole in the two-nucleon amplitude. Clearly, this uniqueness property still holds if we start with a meson of nonzero spin. Comparison of (9) and (4) now gives us the value of g_0^2, from



(10)

setting $D \equiv V$ in the two-particle region. As each partial-wave component of these expressions can be expressed as the product of two partial-wave amplitudes, the most general factorization of the equation

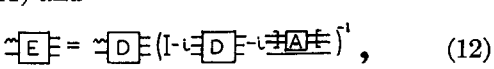
gives



(11)

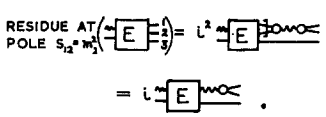
up to an ambiguity of sign, which can be absorbed into the arbitrary sign factor of g_0 (we make essential use of the property that V and V^* are continuations of each other in eliminating constant complex phase factors). Note that g_0 is real in view of the reality of the two-nucleon amplitude.

From (11) and



(12)

we obtain the factorization property



(13)

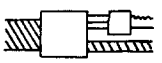
This demonstrates that the two-particle scattering amplitude for meson on nucleon is itself in its entirety contained in the production amplitude (meson + nucleon) → (three nucleons) as a factor of the residue at an appropriate meson pole. The maximal analyticity principle ensures this and hence that the complete singularity structure of the two-particle amplitude is determined by that of the higher amplitude.

2. THE UNPHYSICAL REGION DECOMPOSITION LAW

The above factorization property forms the natural extension of the physical region decomposition law given by the factorization of residues at single particle poles in the physical regions [cf. Eq. (5.6) of I]. We show in this subsection that this unphysical region decomposition law is almost certainly true in general, on the basis of our postulates, and thus forms a powerful tool for determining singularity structure. The physical necessity of the single-particle decomposition law is quite clear, as direct scattering experiments can be made arbitrarily close to the single-particle poles, in which case we would expect that the major contributions to a many-particle scattering process would arise from two successive lower-order scattering process connected by a real stable intermediate particle state, which may exist for macroscopic distances and times. It is physically reasonable to suppose that such a factorization property is approximately true for poles which are very close to the physical region, and indeed the maximal analyticity principle now shows that the factorization property is still exact at the pole itself.

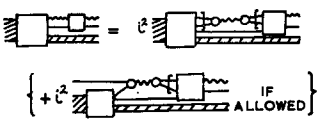
The general decomposition law is deduced from a more complete form of the bound-state postulate. This may be regarded as in some sense the equivalent of a postulate of the existence of the corresponding three-particle interaction in a Lagrangian field theory. In effect, we postulate the absence of all threshold singularities involving mesons on the unphysical sheet of any amplitude attained by encircling the branch points for associated thresholds for which one or more of the mesons are replaced by one or more nucleon pairs. We cannot at present give a completely satisfactory proof owing to the many complications of overlapping cuts in the general case.

Working with a quite general amplitude, we need to consider only meson–nucleon pairs and meson–meson pairs of either both incoming or both outgoing particles. Taking the first case, among the various unitarity conditions, we always find an expression for the discontinuity across a three-particle cut of the form



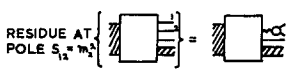
(14)

Using the analysis of Sec. 3 of I as adapted to this case, we obtain the analog of (10):



(15)

From (11) and the unitarity conditions, we obtain the analog of (13):



(16)

The meson–meson pairs can be treated in a similar way via the meson + two-nucleon intermediate states. This is sufficient to obtain the decomposition law for every case in which one of the factors is the three-particle vertex.

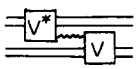
We again stress that this is only an outline proof. We are accepting its conclusions only because it works in the simpler cases and we can see no means by which it can fail for the more complicated cases. Indeed, we might take the decomposition law as a fundamental postulate, if it were possible to show the existence of the unphysical region poles more directly.

3. CROSSING SYMMETRY AND THE DETERMINATION OF SINGULARITY STRUCTURE.

Up to this point in the paper we have made no use of the crossing symmetry properties of scattering

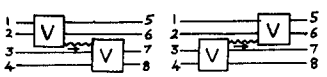
amplitudes. In field theory, these properties appear as direct consequences of the basic postulates in, for example, the proof of dispersion relations.[3] Correspondingly, we would like to deduce these properties from our postulates. We start with examples showing the possibility of such a deduction.

The set of terms



(17)

contribute to the physical region unitarity condition for the four-nucleon scattering amplitude if $W^2 > (4m_1 + m_2)^2$. If we label the particles and add arrows to denote the direction of positive energy, then the diagrams



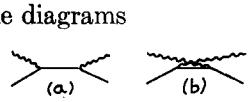
(18)

represent the discontinuities across physical region poles on the physical side of the cuts. However, they are essentially the same pole as they both lie on the surface

$$(p_1 + p_2 - p_5 - p_6)^2 = m_2^2, \tag{19}$$

and the principle of maximal analyticity thus requires these terms to be analytic continuations of each other if the continuation is at all possible. This would imply that, apart from an irrelevant phase factor, the component amplitudes are indeed continuations of each other, as required by the crossing symmetry principle. It is clear that by this mechanism any two physical amplitudes which differ by interchange of some of the ingoing and outpoing particles can be connected by a sequence of terms like (17) in which they appear and from which the crossing property can be deduced by an application of the maximal analyticity principle to physical region poles.

We assume that all these continuations of residues are indeed possible, so that crossing symmetry follows from our postulates. Of course, it is not necessary to assume that all these continuations are valid, as they are not independent. Another approach is to establish that the residues at the poles corresponding to the diagrams



(20)

are the same. The pole term (20b) (crossed nucleon pole) is obtained at the end of this subsection, whereas the first one follows from an application of the bound state postulate (cf. Sec. 2) to a nucleon–meson pair coupled with a single nucleon.

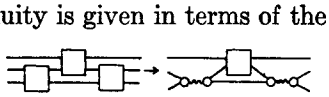
This latter postulate must be made simultaneously with the nucleon-pair–meson bound-state postulate if we are not to contradict elementary crossing properties. As we indicate below, the complete singularity structure of the scattering amplitudes as generated by these poles is that obtained by considering the singularities of a complete set of perturbation theory graphs (extended to include unstable particle lines). This spectrum of discontinuities across the complete set of perturbation theory singularities is manifestly crossing symmetric, and so the amplitudes themselves must be, insofar as they are determined by their discontinuities.

Our prescription for obtaining the singularity structure is to take the complete set of diagrams of the perturbation theory type which are allowed by selection rules (note that our diagrams always represent a unitarity discontinuity if they involve two or more amplitudes, and are not Feynman diagrams). For our purposes, we include all diagrams for which unstable particle lines are allowed and so the total amplitude cannot be regarded as a formal sum of all these diagrams. In addition, at all four and higher particle number vertices, we always write in the complete scattering amplitude appropriate to the number of lines meeting. To take an example, the discontinuity of the triangle graph,[4,5] which appears as

(21)

in our notation occurs in the meson–nucleon elastic scattering amplitude. We now claim that the singularity structure associated with this graph and all other possible ones is present in the complete meson–nucleon scattering amplitude. The main difficulty is in determining the sheet structure of these discontinuities.

Using the decomposition law, we can state that this discontinuity is given in terms of the residue



(22)

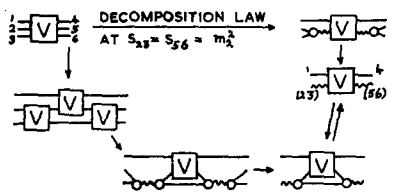
of suitable poles in the three-particles scattering amplitude. But we know that the latter unitarity graph can be continued to the physical region, where its discontinuity appears as a direct consequence of

[3] This is clearly demonstrated in the review by J. D. Jackson in *Dispersion Relations*, edited by G. R. Screaton (Oliver and Boyd, Edinburgh, 1961).

[4] R. Karplus, C. M. Sommerfeld, and F. H. Wichmann, Phys. Rev. **111**, 1187 (1958).

[5] R. Cutkosky, Rev. Mod. Phys. **33**, 448, (1961). Also J. Math. Phys. **1**, 429, (1960).

unitarity (cf. 3.30 of I). Thus the principle of maximal analyticity verifies the presence of the singularity associated with (21) and moreover gives the sheets on which it occurs or is absent. We may follow the steps of the process in the diagram



(23)

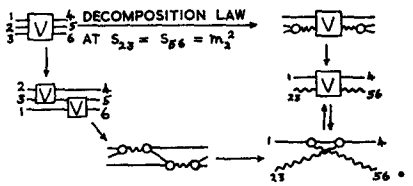
For the case $s_{23} = s_{45}$, perturbation theory calculations give the location of the anomalous threshold term branch point at

$$s_{14} = 4m_1^2[1 - (1 - s_{23}/2m_1^2)^2], \tag{24}$$

and on continuing to the meson pole at $s_{23} = s_{45} = m_2^2$, we see that the branch point stays on the physical sheet or moves onto the unphysical sheet depending on whether $m_2^2 > 2m_1^2$ or $< 2m_1^2$. This gives the usual condition for the presence or absence of the appropriate anomalous threshold on the physical sheet of the meson–nucleon scattering amplitude. This process is in some ways similar to off-mass-shell continuation methods of obtaining the same result.[6]

To generalize the argument, we claim that the discontinuity represented by any graph can be embedded in some amplitude involving a larger number of external particles, such that the corresponding singularities enter the physical region of that higher-order process. Physical region unitarity and the decomposition law then confirm the existence of the singularity in the lower-order amplitude even if it lies entirely in unphysical regions, and in principle completely determine the sheet or sheets on which it lies. While it is clear that any graph can be built up by repeated applications of the singularity multiplication process of Part III on certain elementary diagrams, there still remains the possibility of the presence of other singularities which cannot be attributed to discontinuities given by any of the perturbation-theory graphs. As we have not yet discovered any means by which these extra singularities may be generated and still be in accord with our analyticity postulates, we at present omit them and regard the above specification as complete Similar considerations have been given by Polkinghorne and Zwanziger (see Introduction to I).

To conclude this subsection, we now derive the crossed nucleon pole of (20a) by the steps [cf. (23)],



(25)

4. ALTERNATIVE FORMS OF THE BASIC POSTULATES.

The program for constructing a complete on-mass-shell theory of the S matrix can be based on the following postulates:

(i) the postulate of the existence of the usual spaces of Lorentz covariant in- and out-states involving the stable particles of the theory and the the existence of a unitary S matrix connecting them (see Sec. 2 of I);

(ii) the maximal analyticity postulate (see Secs. 1 and 2 of I) for the on-mass-shell continuations of the scattering amplitudes and the analyticity unitarity connection (see the end of Sec. 2.2 and Sec. 3.3 of I);

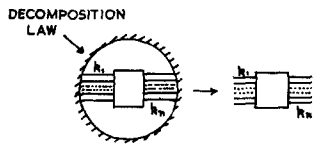
(iii) the bound-state postulate, which in our presentation takes the form of requiring the absence of certain threshold singularities on unphysical sheets whenever the bound-state condition is satisfied (see the second paragraph of the introduction and Sec. 2). In accordance with present trends, we may regard any particle as a bound state of any set of particles with the same over-all quantum numbers. This condition makes the bound-state postulate more precise, but can be readily modified if it is required to take certain particles as "elementary."

We have demonstrated possible ways for deducing the complete singularity structure and crossing symmetry properties, at least in principle, for all scattering amplitudes and in terms of a few given coupling constants and masses. This is the basis on which we claim that, if we knew how to solve the resulting equations, the physical scattering amplitudes are given uniquely by the equations representing the analyticity and unitarity conditions. However, at present we cannot even write down a complete set of equations.

Alternative formulations of the above three postulates can no doubt be devised. In particular, it should be possible to construct them in such a way as to cause stable and unstable particles to enter on a more equal footing, in contrast with the above formulation. Presumably in such a formulation one must take the S matrix as a given entity independent of the existence of any in- and out-state spaces,

[6] S. Mandelstam, Phys. Rev. Letters 4, 84, (1960).

with the one-particle decomposition law taken as fundamental. Then all amplitudes appear as derived quantities in some over-all "universal" functional amplitude, being obtained by extracting a component which is a simultaneous singularity at several single-particle poles



(26)

This is depicted in (26), in which we define the n-point amplitude in terms of the coefficient of $\delta^{(4)}(k_1 + k_1 + \cdots + k_n)$ in the simultaneous residue at n single-particle poles. The required amplitude is extracted from the residue by removing a constant factor representing the rest of the "universal" amplitude. Reasons for taking this factor as constant are given later in the section. We have not yet found a more satisfactory way of formulating these ideas, in particular the precise functional form of the "universal" amplitude, but considerations such as these seem inevitable when one abandons asymptotic states. For unstable particle and unphysical bound state poles, the energy–momentum vectors are complex and so we must regard the residue as being observed from a neighboring region which lies in a physical region. A sufficiently high degree of accuracy in the knowledge of the function in the physical region enables one, on the basis of the analyticity principle, to separate out in principle just the pole contribution. This is precisely what is done in "polological" methods for measuring coupling constants.[7]

The unitarity condition can now be formulated. There is now no expansion like (2.5) of I, but the equations obtained from it can be reinterpreted as the unique expression of the conservation of probability for a complete set of particle states. We are regarding the value of any point on the entire Riemann surface as in principle measureable, either "on the spot" in the case of a physical region or by extrapolation at other points. The class of analytic functions from which an extrapolation should be chosen must take due regard of the general analytic properties of scattering amplitudes, in order to avoid pathological cases of the type found by Dyson,[8] in which infinite accuracy of measurement is required in order to extrapolate with finite error to any unphysical point whatsoever.

For a general complex value of the total energy invariant for incoming particles, there appears to be no natural dichotomy of particle states into those that are "allowed" because there is sufficient c.m. energy available and those that are "not allowed." This contrasts with the real energy case. We therefore take *any* set of particle states (including unstable ones) as a "complete" set for a particular total energy and with this complete set apply the probability amplitude interpretation and require conservation of probability. We may allow the "probability" to become complex away from the physical regions as long as it attains a physically reasonable value in the physical region itself. From a suitable formulation of this idea, it appears reasonable that we should be able to derive the interpolating type of unitarity condition discussed in Sec. 3.2 of I. We have already shown that this latter set of equations is self-consistent, even when we include the unstable particle singularities. We simply interpret the various equations as connecting the values of the amplitude on different sheets, all of which are measurable quantities in the above sense. The physical region is then just a real domain on one of these sheets where "on the spot" measurements can be made and this is the sole reason for singling it out.

It is apparent that the reason we obtained only stable particle states in the unitarity conditions was our insistence on using the spaces of asymptotic in- and out-states taken over from field theory. We consider that formulations of the type discussed above are more satisfactory than the earlier ones just because the need for these dubious asymptotic states is eliminated. Another possible reason for preferring it is that it forms a more natural setting for certain modifications of present relativistic quantum theory to take into account a "fundamental length" or some other way of expressing the finiteness of the universe. In many ways, both directly from the commutation rules[9] and indirectly through the unitarity conditions,[10] it has been shown that there is no natural upper limit to the number of particles in a given state in conventional quantum theory (neither in the form of a sharp cutoff nor as a smoother one which inhibits states of sufficiently high energy, particle number, etc., without actually

[7] M. J. Moravcsik, review in *Dispersion Relations*, edited by G. R. Screaton (Oliver and Boyd, Edinburgh, 1961), p. 117.

[8] K. Symanzik, in *Lectures on Field Theory and the Many Body Problem*, edited by E. R. Caianello (Academic Press Inc., New York), p. 90.

[9] The infinite dimensionality of any representation of the boson commutation relations has been proved many times, e.g., see L. Gårding and A. Wightman, Proc. Nat. Acad. Sci. U.S., **40**, 622 (1954).

[10] O. W. Greenberg, J. Math. Phys. **3**, 31 (1962). Also R. Acharya, Nuovo Cimento **27**, 1151 (1963).

excluding them). We attribute this property in our formulation to the convention that the "rest of the universe" contribution in (26) is a constant factor. This expresses the idea that the rest of the universe "produces" and "accepts" any particles we wish to use in our experiments with an entirely uniform probability distribution. In other words, the preparation of a particle in a given state of charge, momentum, etc., for use in a scattering experiment in no way affects our chances of being able, in principle, to prepare any other particle state for use in that experiment. Similar remarks apply to the outgoing particle states. This situation is found to be true to a very high degree of approximation in almost all of present-day elementary particle physics (except for gravitational physics), but we feel that it should not be applied at arbitrarily high values of the energy, momenta, angular momenta, etc., except as a simplifying approximation. Instead we propose that the unitarity conditions should be modified by requiring only conservation of probability jointly for both the components in the factorization (26), one of which represents our scattering experiment and the other one the rest of the universe. However, we know of no suitable formulation of this idea.

In both conventional field theory and our unmodified version of S-matrix theory it is in effect postulated that we can separate out a particular part of the universe for detailed study to an arbitrary degree of disconnection from events occurring in the rest of the universe. However, many people accept it as a "principle of impotence" that one cannot push this degree of disconnection arbitrarily far. This principle has been accepted from the start in Einsteinian gravitational theory, and this partly accounts for the widely held view that gravitational theory seems somewhat set apart from the rest of elementary particle physics. If the above proposed modification of quantum theory could be satisfactorily formulated, then the gravitational interaction would find a more natural place along with the other, stronger types of interaction, differing from the rest in that the interaction is so weak that the "few-particle approximation," as we may term ordinary S-matrix theory, has only a small domain of applicability.

5. RELATIONS WITH OFF-MASS-SHELL THEORIES.

As all physical measurements are made on the mass shell, no off-mass-shell quantity is measurable in the sense used in Sec. 4. However, it may still happen that there exist such continuations, which satisfy a generalized form of unitarity condition and the various symmetry properties, etc., but which are not now expected to be unique. This is in accordance with the non-uniqueness which occurs in the definition of a local interpolating field associated with a given S matrix.[11] Indeed, if we could show that a particular off-mass-shell "continuation" satisfied enough of the properties of the generalized r functions[12] of field theory, then we could demonstrate the existence of a local field by using the equivalence theorems of Wightman.[13] However, this problem is still a long way from final solution and is not treated in this paper.

[11] H. J. Borchers, Nuovo Cimento **15**, 784 (1960).
[12] H. Araki, J. Math, Phys. **2**, 163 (1961).
[13] A. Wightman, Phys. Rev. **101**, 860, (1956).

Unitarity and On-Mass-Shell Analyticity as a Basis for S-Matrix Theories. III

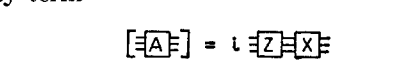
J. GUNSON
Department of Mathematical Physics, University of Birmingham, Birmingham, England
(Received 20 January 1964; final manuscript 30 June 1964)

The singularity structure of many-particle scattering amplitudes in momentum transfer variables is investigated in terms of the many-particle unitarity conditions. Cauchy-type kernels for analytic functions defined on complex rotation group in three dimensions are constructed and related to a theory of local representations of the rotation group, corresponding to complex angular momenta.

IN Parts I and II of this series, we have studied singularities of scattering amplitudes whose location depends only on the energy invariants for one particular channel, in particular the total energy W. By generalizing the treatment given in a previous paper,[1] we investigate below the implications that the many-particle unitarity relations hold for the structure of singularities whose locations depend on one or more of the momentum transfer variables for the chosen channel. For this purpose, we have found it convenient to construct a special Cauchy-type reproducing kernel for holomorphic functions defined on the manifold of the group $O(3, C)$, regarded as a complex space of three (complex) dimensions. In the original derivation of this kernel, there appeared incidentally certain formulas which had, at least locally, some of the structural properties of representations of the rotation group in three dimensions. These formulas have subsequently proved to be of importance in the complex angular momentum analysis of many-particle amplitudes and so a detailed mathematical treatment of certain basic properties is given in Sec. 2. Applications to scattering theory will appear in later papers.

1. CONTINUATION IN MOMENTUM TRANSFER VARIABLES

As an example we consider a three-particle unitarity term



(1)

occurring in (3.36) of I, whose partial wave projection appears in (3.38). The total amplitude

$$A(W^2, \cdots, \rho) = \sum_{j=0}^{\infty} \sum_{m=-j}^{+j} \sum_{m'=-j}^{+j} A_j^{m,m'}(W^2, \cdots)\, \mathfrak{D}_j^{m,m'}(\rho) \quad (2)$$

is, according to our postulates, analytic in ρ on the group manifold of the complex orthogonal group $O(3, C)$ (or rather the component of this group which is connected with the identity element) except for certain sets of singularities. These must include at least the ones which intersect the physical region itself, i.e., the submanifold which forms the real rotation group. These comprise the single-particle pole and anomalous threshold terms introduced in Sec. (3.3) of I and further branch points at higher energies.

In Sec. 2, there is constructed an invariant kernel function associated with a special type of closed hypercontour M in the manifold of $O(3, C)$ and possessing the reproducing property

$$f(R) = \oint_M dS\, K(R, S) f(S) \quad (3)$$

for any function f holomorphic in a neighborhood of the physical region. In order to apply this theorem, we must reintroduce the $(-i\epsilon)$ addenda to the masses of the intermediate particles as in Sec. 2 of I, in order to produce a neighborhood of the physical region in which the functions are holomorphic. Let us now write the amplitudes occurring in the unitarity integral (1) in this form, i.e.,

$$A(R_3) = \int_G dR_1 \int_G dR_2\, \delta(R_3, R_2R_1)\, Z(R_2)\, X(R_1) \quad (4)$$

after omitting irrelevant variables and integrations over the intermediate phase space. G is the manifold of $SO(3)$. Using (3) on Z alone, we get

$$A(R_3) = \int_R dR_1 \oint_M dS_2\, K(R_3R_1^{-1}, S_2)\, Z(S_2)\, X(R_1). \quad (5)$$

Applying to both Z and X, we get

$$A(R_3) = \oint_M dS_1 \oint_M dS_2 H(R_3, S_1, S_2) Z(S_2) X(S_1), \quad (6)$$

[1] J. Gunson and J. G. Taylor, Phys. Rev. **121**, 343 (1961).

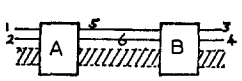
where

$$H(R_3, S_2, S_1) = \int_G dR_1 \int_G dR_2 \cdot \delta(R_3, R_2R_1)K(R_2, S_2)K(R_1, S_1). \quad (7)$$

These formulas show how we may generalize the formula for multiplication of singularities[1] to the many-particle case. The most immediate result is that if $Z(S_2)$ and $X(S_1)$ possess singularities on the sets $\{\beta_i\}$ and $\{\gamma_j\}$, respectively, where β_i and γ_j are elements of the complex orthogonal group, then we find that the singularities of $A(R_3)$ are among the points $\{\beta_i\gamma_j\}$. This result has also been obtained by Lardner.[2] This formula reduces to Eq. (A7) of Ref. (1) in the two-particle scattering case. In contrast with the latter case, however, two singularities not intersecting the physical region can give rise to a multiplied singularity which intersects the physical region; in other words, the product of two non-real rotations may be real.

The nature of the multiplied singularities is determined partly by those of the H-kernel function (7). An integral representation of this function is obtained by substituting the explicit expression (46) for the K kernels and integrating over the δ function. We have not been able to integrate the resulting expression and so we use the technique of deformed contours of integration to analytically continue the function and to determine the singularities. For fixed R_3, we obtain three surfaces of poles which intersect on the three (complex)-dimensional manifold $R_3 = R_2R_1$, square-root branching surfaces on $z_3^2 + z_2^2 + z_1^2 - 1 - 2z_1z_2z_3 = 0$ and logarithmic branch points on $z_1 = \pm 1$, $z_2 = \pm 1$. The latter two types of singularity are present in the simpler kernel[3] for the two-particle scattering case.

For many purposes it is more convenient to use the invariants $z_{\alpha\beta} = \cos\theta_{\alpha\beta}$, where $\theta_{\alpha\beta}$ is the angle between the c.m. momenta of particles α and β in the initial and final states respectively, rather than the Euler angle invariants. In terms of the new invariants, the formulas for multiplication of singularities again takes a simple form. Consider any selection of six particles appearing in a unitarity integral



(8)

The imposition of the three-dimensionality on momentum space requires that we restrict the variables to lie on the variety described by the vanishing of the determinant

$$D(1234) = \begin{vmatrix} 1 & z_{12} & z_{13} & z_{14} \\ z_{12} & 1 & z_{23} & z_{24} \\ z_{13} & z_{23} & 1 & z_{34} \\ z_{14} & z_{24} & z_{34} & 1 \end{vmatrix} \quad (9)$$

and likewise for $D(1256)$ and $D(3456)$. A three-dimensional rotation R_a is determined uniquely by the requirement that the pair of c.m. momenta $\mathbf{k}_1$, $\mathbf{k}_2$ be rotated until $\mathbf{k}_1$ is parallel to $\mathbf{k}_5$ and $\mathbf{k}_1 \times \mathbf{k}_2$ is parallel to $\mathbf{k}_5 \times \mathbf{k}_6$. In terms of this choice of coordinates let the amplitude A possess a singularity at R_a', for fixed z_{12}, z_{56}, etc. In combination with a singularity of B at R_b', the previous results show that we may expect a singularity of the integral at

$$R_c' = R_b'R_a' \quad (10)$$

in the coordinate system for which $(\mathbf{k}_1, \mathbf{k}_2)$ is rotated into $(\mathbf{k}_3, \mathbf{k}_4)$. Taking the simplest case in which there are poles at R_a' and R_b' (note that these must lie on surfaces of poles, because of the continuity properties of singularities for functions of two or more complex variables), we then find the multiplied singularity in the form of a pole at the location given by

$$D(1653) = D(1564) = D(2563) = D(2564) = 0, \quad (11)$$

in which z_{15}, z_{16}, z_{25}, z_{26}, z_{35}, z_{45}, z_{36}, z_{46} are set equal to the values given by R_a', R_b'. The solution of each of these equations in general gives two sets of values for z_{13}, z_{23}, z_{14}, z_{24}, but only one set of solutions is consistent with the requirement $D(1234) = 0$, except in certain special cases. In a similar manner the square-root branching surfaces appear on the surfaces described by the solutions of

$$D(153) = 0 \quad \text{and} \quad D(163) = 0, \quad (12)$$

etc., and logarithmic singularities on $z_{15}' = 0$, etc. These singularities will be modified somewhat when integration over intermediate particle states is carried out.

This method clearly produces the same results as the more usual methods when applied to perturbation theory graphs such as the five-point loop graph.[4] The location of the singularities is determined in both methods essentially be the same conditions of internal lines being upon the mass shell and the momentum space being three-dimensional. The particular set of diagrams (3.29) in I which give phys-

[2] R. Lardner, Nuovo Cimento **23**, 323 (1962); **24**, 763 (1962).

[3] S. Mandelstam, Phys. Rev. **115**, 1741 (1959).

[4] L. F. Cook and J. Tarski, J. Math. Phys. **3**, 1 (1962).

ical-region thresholds form a closed set under multiplication of singularities as described above if we include the physical-region pole terms. This again shows that unitarity gives the physical region singularities consistently.

2. LOCAL REPRESENTATIONS AND THE CONSTRUCTION OF CAUCHY KERNEL FUNCTIONS FOR THE ROTATION GROUP IN THREE DIMENSIONS.

In this section, we associate with the rotation groups a set of mathematical constructions which we have termed "local representations" and which give a natural extension of the well-known global representations of the three-dimensional rotation group[5] for $j = 0, \frac{1}{2}, 1, \frac{3}{2}, \cdots$ to arbitrary complex values of j. Some of these representations are then used in the construction of a Cauchy kernel for functions holomorphic in domains on the manifold of the complex rotation group in three dimensions.

We first give a precise definition of a local representation for an arbitrary Lie group, though this is not essential for the rest of the section and can be omitted by readers unfamiliar with the terminology. It is known that with any topological group there are associated many local groups which are locally isomorphic with it. The abstract definition of a local group is given by Pontrjagin[6] and Cohn,[7] but as we are interested only in Lie groups, we may replace the usual concept by one which is more restrictive. This we do by making the local group also a branched covering space[8] of the underlying manifold of the group.

Definition. A *branched covering local Lie group* (BCLLG) over a connected Lie group G is a quadruple (X, τ, G, H) in which

(i) X is a connected analytic manifold;

(ii) H is an analytic submanifold of G not containing the identity and for which $H^{-1} = H$, and $G - H$ is dense in G;

(iii) $\tau : X \to G$ is a projection of X into G which is locally analytic;

(iv) there is in X a binary relation $(x, y) \to xy$, taking values in X and a unitary relation $x \to x^{-1}$ also taking values in X which are defined for every $x, y \in X$ for which all of $\tau(x)$, $\tau(y)$, $\tau(x)\tau(y) \notin H$ and moreover the relations: $\tau(x)\tau(y) = \tau(xy)$, $\{\tau(x)\}^{-1} = \tau(x^{-1})$ are satisfied in the group operations of G;

(v) there exists an element e in X such that $e^{-1} = e$ and $(x, e) \to x$ for all $x \in X$.

The specification of conditions on X, τ, G, and H for which such a construction is possible would seem to be a very difficult mathematical problem and we have not met any nontrivial constructions of this nature in the mathematical literature. The concept of a representation of a BCLLG can be formulated in an obvious way as a global analytic homomorphism into the ring of bounded operators on a suitable Hilbert space. The irreducibility of such a representation is also well defined, this being a property of any suitable set of operators.[9] We can now define a "local representation" of a connected Lie group G as a true representation of a BCLLG over G.

By actual construction below we demonstrate the existence of local representations which are not true representations for the group $SU(2)$.

2.1. Local Representations of $SU(2)$.

The local representations in which we are interested are analytic continuations of some of the infinite-dimensional true representations of the noncompact locally compact group $SL(2, R)$, first constructed by Bargmann.[10] The connection arises from the occurrence of both $SU(3)$ and $SL(2, R)$ as analytic subgroups of the same complex group $SL(2, C)$. These representations are then also local representations of $SO(3)$ as this group is locally isomorphic with $SU(2)$.

A convenient parametrization of $SU(2)$ is given by the three Euler angles α, β, γ,[5] in terms of which the normalized invariant measure on the group manifold is

$$d\rho = (\sin\beta/16\pi^2)\, d\alpha\, d\beta\, d\gamma : 0 \leq \beta \leq \pi$$
$$-2\pi \leq (\alpha + \gamma), (\alpha - \gamma) < 2\pi. \qquad (13)$$

The basic spinor representation can be written[5]

$$\begin{bmatrix} \omega_3 \\ \tilde{\omega}_3 \end{bmatrix} = \begin{bmatrix} \cos(\frac{1}{2}\beta)\, e^{\frac{1}{2}i(\alpha+\gamma)} & \sin(\frac{1}{2}\beta)\, e^{\frac{1}{2}i(\alpha-\gamma)} \\ -\sin(\frac{1}{2}\beta)\, e^{\frac{1}{2}i(\gamma-\alpha)} & \cos(\frac{1}{2}\beta)\, e^{-\frac{1}{2}i(\alpha+\gamma)} \end{bmatrix} \begin{bmatrix} \omega_2 \\ \tilde{\omega}_2 \end{bmatrix}. \qquad (14)$$

Let us define

[5] A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton University Press, Princeton, New Jersey, 1957).

[6] L. Pontrjagin, *Topological Groups* (Princeton University Press, Princeton, New Jersey, 1939).

[7] P. M. Cohn, *Lie Groups* (Cambridge University Press, New York, 1957).

[8] L. V. Ahlfors and I. Sario, *Riemann Surfaces* (Princeton University Press, Princeton, New Jersey, 1957), p. 39.

[9] M. A. Naimark, *Normed Rings* (P. Noordhoff Ltd., Groningen, The Netherlands, 1960), p. 491.

[10] V. Bargmann, Ann. Math. 48, 568 (1947).

$$\chi^+ = e^{\frac{1}{2}i(\alpha+\gamma)}; \quad \chi^- = e^{\frac{1}{2}i(\alpha-\gamma)}; \quad z = \cos\beta;$$
$$\cos\tfrac{1}{2}\beta = [\tfrac{1}{2}(1+z)]^{\frac{1}{2}}; \quad \sin\tfrac{1}{2}\beta = [\tfrac{1}{2}(1-z)]^{\frac{1}{2}}$$

and the space $H(j, \lambda)$ as the linear vector space spanned by the products

$$\frac{\omega^{j+m}\tilde{\omega}^{j-m}}{\{\Gamma(j+m+1)\Gamma(j-m+1)\}^{\frac{1}{2}}} = e^r_{j,\lambda}$$
$$\mathrm{Re}\, m \equiv \mathrm{Re}\,\lambda (\mathrm{mod}\ 1) \qquad 0 \leq \mathrm{Re}\,\lambda \leq 1$$
$$s = \tilde{\omega}/\omega \qquad (16)$$
$$\mathrm{Im}\, m = \mathrm{Im}\,\lambda$$
$$r = \mathrm{Re}\, m - \mathrm{Re}\,\lambda = 0, \pm 1, \pm 2, \cdots,$$

in which j is any complex number in C_1, the complex plane.

We determine the transformation properties of the $e^r_{j,\lambda}$ by substituting (14), to give

$$\omega_3^{j+m}\tilde{\omega}_3^{j-m} = \omega_2^{j+m}\tilde{\omega}_2^{j-m}[\cos(\tfrac{1}{2}\beta)\chi^+ + \sin(\tfrac{1}{2}\beta)\chi^- s_2]^{j+m} \cdot\left(-\sin(\tfrac{1}{2}\beta)\frac{1}{\chi^- s_2} + \frac{\cos\frac{1}{2}\beta}{\chi^+}\right). \qquad (17)$$

We may expand this as a Laurent series in s_2 for certain restricted values of the group parameters. The most important case is depicted in Fig. 1

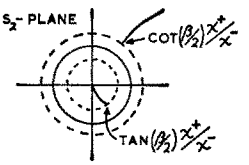


FIG. 1. Integration contour in the s_2-plane, for the case $|\cot(\frac{1}{2}\beta)\chi^+/\chi^-| > |S_2| > |\tan(\frac{1}{2}\beta)\chi^+/\chi^-|$.

thus defining the quantities $D_j^{m,m'}(R)$, $R \in SU(2)$

$$e^r_{j,\lambda} = \sum_{r'=-\infty}^{+\infty} D_j^{m,m'}(R)\, e^{r'}_{j,\lambda} \qquad \begin{matrix} r' = \mathrm{Re}\, m' - \mathrm{Re}\,\lambda \\ m' = m + (r' - r), \end{matrix} \qquad (18)$$

from which we deduce the group properties

$$D_j^{m,m'}(R_3) = \sum_{m''} D_j^{m,m''}(R_2)\, D_j^{m'',m'}(R_1), \quad R_3 = R_2 R_1, \qquad (19)$$

valid at least when all the defining series converge.

An explicit expression for the $D_j^{m,m'}$ is obtained by picking out the coefficients of the Laurent series by contour integration

$$D_j^{m,m'}(R) = \left\{\frac{\Gamma(j+m'+1)\Gamma(j-m'+1)}{\Gamma(j+m+1)\Gamma(j-m+1)}\right\}^{\frac{1}{2}} \cdot\oint \frac{(\cos(\frac{1}{2}\beta)\chi^+ + \sin(\frac{1}{2}\beta)\chi^- S_2)^{j+m}\left(-\sin(\frac{1}{2}\beta)\frac{1}{\chi^-} + \cos(\frac{1}{2}\beta)\frac{S_2}{\chi^+}\right)^{j-m}}{S_2^{j-m'+1}} \frac{dS_2}{2\pi i}, \qquad (20)$$

which can be converted into a hypergeometric form of integral,[11] to give

$$D_j^{m,m'}(\chi^+, z, \chi^-) = (\chi^+)^{m+m'}(\chi^-)^{m-m'}\, d_j^{m,m'}(z), \qquad (21)$$

where the $d_j^{m,m'}(z)$ can be written in the form

$$d_j^{m,m'}(z) = \left\{\frac{\Gamma(j+m+1)\Gamma(j-m'+1)}{\Gamma(j+m'+1)\Gamma(j-m+1)}\right\}^{\frac{1}{2}} \cdot\left(\frac{1+z}{2}\right)^{\frac{1}{2}(m+m')}\left(\frac{1-z}{2}\right)^{\frac{1}{2}(m-m')} \cdot\frac{{}_2F_1[-j+m, j+m+1; (1+m-m') : \frac{1}{2}(1-z)]}{\Gamma(1+m-m')} \qquad (22)$$

and three other forms obtained from the relations

$$d_j^{m,m'}(z) = (-1)^{m-m'}\, d_j^{-m,-m'}(z) = (-1)^{m-m'}\, d_j^{m',m}(z) = d_j^{-m',-m}(z), \qquad (23)$$

which are identical to the usual expressions for global representations.[5] These functions are analytic in the cut z plane of Fig. 2. This shows that the branching surface H for these representations is just the manifold $z = -1$.

Further important properties of the $d_j^{m,m'}(z)$ are the asymptotic relations

$$|d_j^{m,m'}(z)| \to \text{const}\ |m'|^{m-\frac{1}{2}} \left|\frac{1-z}{1+z}\right|^{\pm m'/2},$$
$$|\arg(1+z)| < \pi - \epsilon \qquad \mathrm{Re}\, m' \to \pm\infty \qquad (24)$$

for fixed j, m, z, and large $|m'|$, obtained by using the appropriate asymptotic expansions for the hypergeometric functions.[12] This can be used to calculate

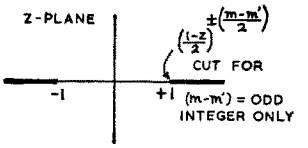


FIG. 2. Cut for $(n - m')$ = odd integer only.

[11] J. Gunson, Ph.D. thesis, Cambridge, 1962 (unpublished).

[12] A. Erdelyi *et al.*, *Higher Transcendental Functions* (McGraw-Hill Book Company, Inc., New York, 1956), Vol. I. p. 75.

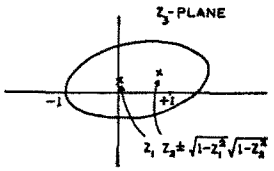


FIG. 3. Elliptical domain of convergence for the series in Eq. (19).

the domain of convergence of (19) in an elementary manner. The result is that the series converges in

$$\left|\left(\frac{1-z_2}{1+z_2}\cdot\frac{1-z_1}{1+z_1}\right)^{\frac{1}{2}}\right| < |e^{i(\gamma_2+\alpha_1)}| < \left|\left(\frac{1+z_2}{1-z_2}\cdot\frac{1+z_1}{1-z_1}\right)^{\frac{1}{2}}\right|, \tag{25}$$

which finds an alternative expression in

$$\cos(\gamma_2+\alpha_1) \in E\left(\frac{1+z_1z_2}{(1-z_1^2)^{\frac{1}{2}}(1-z_2^2)^{\frac{1}{2}}}\right) \cdot\left|\frac{1-z_2}{1+z_2}\right|\cdot\left|\frac{1-z_1}{1+z_1}\right| < 1, \tag{26}$$

where $E(a)$ is the ellipse with foci ± 1 and passing through the point a. In terms of $z_3 = z_1z_2 + (1-z_1^2)^{\frac{1}{2}} \times (1-z_2^2)^{\frac{1}{2}} \cos(\gamma_2+\alpha_1)$, the ellipse becomes that shown in Fig. 3. These are precisely the conditions given by Henrici[13] for the classical special case of (19), viz.,

$$P_j(z_1z_2 + (1-z_1^2)^{\frac{1}{2}}(1-z_2^2)^{\frac{1}{2}}\cos\omega) = P_j(z_1)P_j(z_2) + 2\sum_{\mu=1}^{\infty}(-1)^{\mu}\frac{\Gamma(j-\mu+1)}{\Gamma(j+\mu+1)}P_j^{\mu}(z_1)P_j^{\mu}(z_2)\cos\omega\mu. \tag{27}$$

2.2. Functions of the Second Kind.

Another region in which a Laurent series expansion of (17) can be made is for

$$\left|\cot\left(\frac{\beta}{2}\right)\frac{\chi^+}{\chi^-}\right| < |s_2| < \left|\tan\left(\frac{\beta}{2}\right)\frac{\chi^+}{\chi^-}\right| \Rightarrow \mathrm{Re}\cos\beta < 0 \tag{28}$$

this defines the functions $C_j^{m,m'}(R)$ in

$$\frac{\omega_3^{j+m}\bar{\omega}_3^{j-m}}{\{\Gamma(j+m+1)\Gamma(j-m+1)\}^{\frac{1}{2}}} = \sum_{m'} C_j^{m,m'}(\chi^+, \cos\beta, \chi^-) \cdot\frac{\omega_2^{j-m'}\bar{\omega}_2^{j+m'}}{\{\Gamma(j+m'+1)\Gamma(j-m'+1)\}^{\frac{1}{2}}} \tag{29}$$

or

$$e_{j,\lambda}^{r} = \sum_{r'=-\infty}^{+\infty} C_j^{m,m'}(R)\, e_{j,1-\lambda}^{r'},$$
$$r' = -\mathrm{Re}\, m' - \mathrm{Re}(1-\lambda) = 0, \pm 1, \pm 2, \cdots$$
$$m' = m - r - r' - 1. \tag{30}$$

Evaluating $C_j^{m,m'}(R)$ as for $D_j^{m,m'}(R)$ we get

$$C_j^{m,m'}(R) = (\chi^+)^{m-m'}(\chi^-)^{m+m'}c_j^{m,m'}(z), \tag{31}$$

with

$$c_j^{m,m'}(z) = \left\{\frac{\Gamma(j+m+1)\Gamma(j-m'+1)}{\Gamma(j+m'+1)\Gamma(j-m+1)}\right\}^{\frac{1}{2}} \cdot\left(\frac{1+z}{2}\right)^{\frac{1}{2}(m-m')}\left(\frac{1-z}{2}\right)^{\frac{1}{2}(m+m')} \cdot\frac{{}_2F_1[-j+m, j+m+1; 1+m-m' : \frac{1}{2}(1+z)]}{e^{\mp i(j-m)\pi}\Gamma(1+m-m')}$$

with $\pm$ for $\mathrm{Im}\, z \gtrless 0$. (32)

The symmetry relations (21) are still satisfied with $d \to c$. From (29) and (18) we derive the group property

$$C_j^{m,m'}(R_3) = \sum_{\substack{m''=m,\\ m\pm1,\cdots}} D_j^{m,m''}(R_2)C_j^{m'',m'}(R_1), \qquad R_3 = R_2R_1. \tag{33}$$

In (28), it is seen that the $C_j^{m,m'}(R)$ connect the spaces $H(j,\lambda)$ and $H(j, 1-\lambda)$. In the special cases $\lambda = 0, \frac{1}{2}$ these spaces are the same, and so for these cases we may define a function of the second kind $e_j^{m,m'}(z)$ by

$$d_j^{m,m'}(z) = c_j^{m,-m'}(z) + 2/\pi\, e^{\pm i\pi(j-m)} \cdot\sin\pi(j-m)\, e_j^{m,m'}(z), \qquad \lambda = 0, \tfrac{1}{2}, \tag{34}$$

from which the expression

$$e_j^{m,m'}(z) = \tfrac{1}{2}\{\Gamma(j+m+1)\Gamma(j-m+1) \cdot\Gamma(j+m'+1)\Gamma(j-m'+1)\}^{\frac{1}{2}} \cdot\left(\frac{1+z}{2}\right)^{\frac{1}{2}(m+m')}\left(\frac{1-z}{2}\right)^{-\frac{1}{2}(m-m')}\left(\frac{z-1}{2}\right)^{-j-m'-1} \cdot\frac{{}_2F_1[j+m'+1, j+m+1; 2j+2 : 2/(1-z)]}{\Gamma(2j+2)}, \qquad \lambda = 0, \tfrac{1}{2} \tag{35}$$

may be deduced by using the linear relations between hypergeometric functions.[14]

On defining

$$E_j^{m,m'}(\chi^+, z, \chi^-) = (\chi^+)^{m+m'}\, e_j^{m,m'}(z)(\chi^-)^{m-m'}, \tag{36}$$

[13] P. Henrici, J. Ratl. Mech. Anal. 4, 983 (1955).

[14] Reference 12, p. 108.

we get from (19) and (33) the grouplike property

$$E_j^{m,m'}(R_3) = \sum_{m''=-m,m\pm1,\cdots} D_j^{m,m''}(R_2)E_j^{m'',m}(R_1),$$
$$R_3 = R_2R_1, \qquad (37)$$

which is a generalization of the addition theorem for Legendre functions of the second kind. Indeed, for $\lambda = 0, \frac{1}{2}$, and integral $(j - m)$, we have the identity

$$e_j^{m,m'}(z) = (-1)^{m-m'}\left\{\frac{(j-m)!\,(j+m)!}{(j-m')!\,(j+m')!}\right\}^{\frac{1}{2}} \cdot\left(\frac{1+z}{2}\right)^{\frac{1}{2}(m+m')}\left(\frac{1-z}{2}\right)^{\frac{1}{2}(m-m')} Q_{j-m}^{m-m',m+m'}(z) \qquad (38)$$

on using Szëgo's definition of the Jacobi functions of the second kind.[15] The asymptotic properties at large m' follow directly from (34) to give

$$|e_j^{m,m'}(z)| \sim \max\left(|m'|^{m-\frac{1}{2}}\left|\frac{1-z}{1+z}\right|^{m'/2}, |m'|^{-m}\left|\frac{1+z}{1-z}\right|^{m'/2}\right)$$
for $m' \to \pm\infty$, fixed j, m, z, $|\arg(z-1)| < \pi - \epsilon$. (39)

The expressions for the convergence domain corresponding to (25) and (26) are

$$\mathrm{Re}\, z_1 > 0 : \left|\frac{1-z_2}{1+z_2}\frac{1+z_1}{1-z_1}\right| < |e^{2i(\gamma_2+\alpha_1)}| < \left|\frac{1+z_2}{1-z_2}\frac{1-z_1}{1+z_1}\right|$$
$$\mathrm{Re}\, z_1 < 0 : \left|\frac{1-z_1}{1+z_1}\frac{1-z_2}{1+z_2}\right| < |e^{2i(\gamma_2+\alpha_1)}| < \left|\frac{1+z_2}{1-z_2}\frac{1+z_1}{1-z_1}\right| \qquad (40)$$

$$\mathrm{Re}\, z_1 > 0 : \cos(\gamma_2+\alpha_2) \in E\left(\frac{z_1z_2-1}{(1-z_1^2)^{\frac{1}{2}}(1-z_2^2)^{\frac{1}{2}}}\right) : \left|\frac{1-z_2}{1+z_2}\cdot\frac{1+z_1}{1-z_1}\right| < 1$$
$$\mathrm{Re}\, z_1 < 0 : \cos(\gamma_2+\alpha_1) \in E\left(\frac{z_1z_2+1}{(1-z_1^2)^{\frac{1}{2}}(1-z_2^2)^{\frac{1}{2}}}\right) : \left|\frac{1-z_1}{1+z_1}\cdot\frac{1-z_2}{1+z_2}\right| < 1. \qquad (41)$$

As these functions are cut from $+1$ to $-\infty$ in the z plane, care must be taken as to which sheets the functions are evaluated. This is treated in more detail by Henrici,[13] who obtains the above regions of convergence by different methods for the case $\mathrm{Re}\, z_1 > 0$, but does not mention the case $\mathrm{Re}\, z_1 < 0$.

[15] G. Szëgo, *Orthogonal Polynomials* (American Mathematical Society, New York, 1939).

3. CAUCHY KERNELS FOR THE COMPLEX ROTATION GROUP

The manifold of $SU(2)$ is a three-dimensional real analytic manifold which is topologically a three-dimensional sphere S_3. The associated complex manifold is thus $S_3 \times R_3$, and so we cannot use the usual Cauchy kernels of C_3 (which topologically is $R_3 \times R_3$) for all types of contour.

The type of representation which we are seeking is

$$f(\rho_1) = \int_M d\rho_2 K(\rho_1, \rho_2) f(\rho_2) \qquad (42)$$

holding for any function f which is holomorphic on the manifold of $SU(2)$ (which is the "physical region" in applications) and hence holomorphic in a complex neighborhood. We choose M as the special contour which is the product of three single variable contours as depicted in Fig. 4.

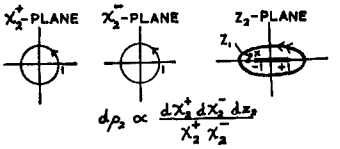


FIG. 4. Projections of the surface of integration M in three different planes.

The contour in the z_2 plane takes the form of a pair of ellipses with foci ±1, one on each sheet of the Riemann surface of $(z_2^2 - 1)^{\frac{1}{2}}$.

The choice of M is motivated by its use in projecting our partial waves by

$$f_j^{m,m'} = \int_G d\rho f(\rho)\, D_j^{*m,m'}(\rho) = \frac{1}{2\pi i}\int_M d\rho f(\rho) E_j^{*m,m'}(\rho). \qquad (43)$$

This can be proved from the expressions of the $E_j^{m,m'}(\rho)$ in terms of Jacobi functions of the second kind [cf. (38)]. The kernel $K(\rho_1, \rho_2)$ thus possesses the following expansion

$$K(\rho_1,\rho_2) = \sum_{m+m'=-\infty}^{+\infty}\sum_{m-m'=-\infty}^{+\infty}\sum_{j-\max(|m|,|m'|)=0}^{\infty}(2j+1) \cdot D_j^{m,m'}(\rho_1)E_j^{*m,m'}(\rho_2)$$
for $|z_1 + (z_1^2-1)^{\frac{1}{2}}| < |z_2 + (z_2^2-1)^{\frac{1}{2}}|$;
$$\left|\frac{1+z_1}{1+z_2}\right| < \left|\frac{\chi_1^+}{\chi_2^+}\right|^2 < \left|\frac{1+z_2}{1+z_1}\right|;$$
$$\left|\frac{1-z_1}{1-z_2}\right| < \left|\frac{\chi_1^-}{\chi_2^-}\right|^2 < \left|\frac{1-z_2}{1-z_1}\right|. \qquad (44)$$

The summations can be carried out explicitly on using the classical expansion[15]

$$\frac{1}{z_2 - z_1} = \sum_{j-m=0}^{\infty} (2j+1) \cdot \left\{\frac{(j-m)!\,(j+m)!}{(j-m')!\,(j+m')!}\right\}\left(\frac{z_2-1}{2}\right)^{m-m'}\left(\frac{z_2+1}{2}\right)^{m+m'} \cdot P_j^{(m-m',\,m+m')}(z_1)Q_j^{(m-m',\,m+m')}(z_1)$$

$$\text{for} \quad |z_1 + (z_1^2-1)^{\frac{1}{2}}| < (z_2 + (z_2^2-1)^{\frac{1}{2}}|, \qquad (45)$$

giving

$$K(\rho_1, \rho_2) = (z_2 - z_1)\left\{\left[\frac{(1+z_2)^{\frac{1}{2}}}{\chi_2^+} - \frac{(1+z_1)^{\frac{1}{2}}}{\chi_1^+}\right] \cdot [(1+z_2)^{\frac{1}{2}}\chi_2^+ - (1+z_1)^{\frac{1}{2}}\chi_1^+]\left[\frac{(1-z_2)^{\frac{1}{2}}}{\chi_2^-} - \frac{(1-z_1)^{\frac{1}{2}}}{\chi_1^-}\right][(1-z_2)^{\frac{1}{2}}\chi_2^- - (1-z_1)^{\frac{1}{2}}\chi_1^-]\right\}^{-1}. \qquad (46)$$

The particularly simple form of this kernel is a direct consequence of our particular choice of variables and contours which take into account the essential topological structure of the complex space.

The relation of (46) and (44) to the completeness condition [cf. Eq. (A3) of I] is immediately obtained by evaluating the appropriate discontinuity to give the invariant δ function on the group manifold.

We list below some basic properties of the kernel function

$$f(R) = \oint_{\rho^{-1}(M)} dSK(\rho R, \rho S)f(S), \quad \text{(invariance)} \qquad (47)$$

where ρ is a real rotation and $\rho^{-1}(M)$ is the integration contour consisting of the elements $\{\rho^{-1}\tau; \tau \in M\}$,

$$f(R) = \oint_{M^{-1}} dSK(R^{-1}, S^{-1})f(S) \qquad (M^{-1} = \{S : S^{-1} \in M\}) \qquad (48)$$

$$\oint_M dSK(R, S)K(S, T) = K(R, T). \qquad (49)$$

ACKNOWLEDGMENTS

This work was done during the tenure of a D.S.I.R. Research Studentship at the University of Cambridge and a D.S.I.R. Research Fellowship at the University of Birmingham.

I should like to thank Professor Peierls for the hospitality of the Department of Mathematical Physics during the past year and Professor Mandelstam for reading through the manuscripts and for helpful discussions.

I am also indebted to Dr. J. G. Taylor and Dr. R. J. Eden for their advice and encouragement during the early stages of this investigation.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 6 JUNE 1965

Long-Time Behavior of the Electric Potential and Stability in the Linearized Vlasov Theory

ALBERT W. SÁENZ
United States Naval Research Laboratory, Washington, D. C.
(Received 7 July 1964)

In this paper we study in a mathematically rigorous manner how the electric potential, produced by small electronic charge density oscillations of definite wavenumber vector **k** in a plasma, behaves in the long-time limit and the connection between this behavior and the stability of a given steady, spatially uniform, distribution of the plasma electrons. Our work is based on the linearized Vlasov equation and on the associated Poisson equation. We formulate a very general initial-value problem concerning this system of equations, writing the above electric potential at a given position vector **r** and time t as $\varphi(t)e^{i\mathbf{k}\cdot\mathbf{r}}$ multiplied by a suitable constant, where $\varphi(t)$ is independent of **r**. We establish the existence and uniqueness of solution of this problem by exploiting the fact that, in the linear theory, $\varphi(t)$ obeys an inhomogeneous Volterra integral equation of convolution type, which is rigorously derived here. A detailed study of the asymptotic properties of the solutions of this equation for $t \to \infty$ is made, including the establishment of necessary and sufficient conditions on the initial perturbations (perturbations of the steady electron distribution function at $t = 0$) for $\varphi(t)$ to be of negative exponential order as $t \to \infty$. As a byproduct of this asymptotic investigation, we give a precise discussion of the Landau damping of long wavelength plasma oscillations in an initially Maxwellian plasma, concluding that in this case $\varphi(t)$ exhibits such damping for a broad range of initial perturbations and that the damping decrement is essentially that first computed by Landau. We introduce criteria of stability and instability based on the boundedness and unboundedness, respectively, in the limit $t \to \infty$ of certain nonnegative quantities $W_p(t)$, which are defined as suitable norms of the perturbed electron distribution function. New sufficient conditions for stability and instability are proved for extensive classes of initial distributions and initial perturbations. These results are compared with conclusions on stability and instability reached by Backus.

I. INTRODUCTION

THIS paper is devoted to a mathematically precise treatment of the way in which the electric potential, originated by electronic charge density waves of small amplitude in a plasma, behaves in the long-time limit and of the relationship between this behavior and the stability of plasma electronic distributions which are independent of spatial coordinates and time. Our investigation is confined to plasma waves of definite wavenumber vector **k**.

Recent results of Hayes[1,2] and Backus[3] on various problems of the linearized Vlasov theory involving plasma disturbances of this last type are of particular relevance to the present work. We proceed to outline these results.

In Hayes' two papers just cited, a basic role is played by an inhomogeneous Volterra integral equation of convolution type. Within the framework of the linearized Vlasov theory, this equation, derived in Ref. 1, governs a certain function $\varphi(t)$ [Eq. (2.3)[4]] which contains the entire dependence on the time t of the above electric potential. A special form of this integral equation had been found earlier by Rosenbluth[5] for the case of Maxwellian initial electron distributions. A subject of prime concern in Refs. 1 and 2 is the derivation of solutions $\varphi(t)$ of the integral equation in question which damp more rapidly than any exponential when $t \to \infty$. An example of a $\varphi(t)$ exhibiting this "non-Landau" damping is given in Ref. 1 for initially Maxwellian plasmas. Necessary and sufficient conditions on the initial perturbations (perturbations of the steady electron distribution function at $t = 0$, due to the plasma oscillation of interest) for this last type of damping to occur are derived in Ref. 2 for a more general type of initial distributions. These results show that Landau's well-known conclusions[6] on the damping of plasma waves in initially Maxwellian plasmas do not have the domain of validity originally ascribed to them.

Backus'[3] beautiful investigations on stability in the cited linear theory are more general and rigorous than any of the earlier work on this subject. This author defines an initial electron distribution to be stable (unstable) if a certain nonnegative quantity $A(t)$ [Eq. (4.2)], involving certain integrations over the perturbed electron distribution, is bounded (unbounded) as $t \to \infty$. $A(t)$ has several physical and

[1] J. N. Hayes, Phys. Fluids **11**, 1387 (1961).
[2] J. N. Hayes, Nuovo Cimento **30**, 1048 (1963).
[3] G. Backus, J. Math. Phys. **1**, 178 (1960).
[4] References inside of square brackets in this Introduction refer to the present paper.
[5] M. Rosenbluth, *Plasma in a Magnetic Field* (Stanford University Press, Stanford, California, 1958), p. 23.
[6] L. Landau, J. Phys. (USSR) **10**, (1946).

mathematical advantages which he points out very lucidly. In contrast to the approach of Refs. 1 and 2, the discussions in Ref. 3 are centered on an integral equation governing the appropriate perturbed electronic distribution functions. Important new sufficient conditions for stability and instability are derived in Ref. 3.

We now summarize the results of the present investigation.

In Sec. II we formulate a very general initial-value problem for the coupled system composed of the linearized Vlasov equation and of the appropriate Poisson equation, and we establish an existence theorem and a uniqueness theorem [Theorems 2.1 and 2.2] for this problem. These theorems are proved by reducing the solution of the initial-value problem of interest to that of solving the previously mentioned integral equation for $\varphi(t)$, which is rigorously derived in this section. Previous derivations[1,5] have been of a purely formal character. Of course, the cited integral equation employed in Ref. 3, which was obtained formally there, can also be given a precise derivation within the context of the linear theory by the methods of Sec. II.

Section III is devoted primarily to the investigation of the solutions of the integral equation for $\varphi(t)$ in the limit $t \to \infty$ for an extensive class of initial distributions and initial perturbations. It should be kept in mind that the major portion of the results in this section, although of obvious relevance to plasma physics, are of much more general applicability, referring as they do to asymptotic properties of inhomogeneous Volterra integral equations of convolution type under certain well defined conditions. Laplace-transform language is used extensively in Sec. III. A basic role in the further work of this section is played by a lemma thereof [Lemma 3.1] which concerns the zeros of the so-called Landau denominator $\Delta(s)$ (defined in Lemma 3.1) in the plane of the Laplace transform variable s and which applies to the class of initial distributions alluded to earlier in this paragraph. In Sec. III we derive a very general formula for $\varphi(t)$ in the limit $t \to \infty$. In this section we also establish a theorem [Theorem 3.2] giving necessary and sufficient conditions on initial perturbations for $\varphi(t)$ to be of negative exponential order in this limit as well as for $\varphi(t)$ to belong to a more general class of functions. We conclude this section by giving a precise discussion of the Landau damping (asymptotically exponential damping) of plasma oscillations in the case when the initial distribution is Maxwellian and when a certain dimensionless number κ, proportional to $|\mathbf{k}|$ [see the paragraph containing (3.14)], is a sufficiently small positive quantity. We find that such damping occurs in this case for a large class of initial perturbations, the asymptotic damping rate being essentially (but not exactly) the same as that computed by Landau.[6] Our discussion of Landau damping is intimately related to the proof of the last cited theorem of this section and rests on rigorous results of the present author [Eqs. (3.15)] concerning the asymptotic location of the zeros of $\Delta(s)$ in the limit $\kappa \to +0$.

In Sec. IV we introduce a family of nonnegative quantities $W_p(t)$ [Eq. (4.1)], which are defined as certain norms of the perturbed electron distribution function. For a given p, we define stability and instability of the $f_0(\mathbf{v})$ of interest in the $W_p(t)$ sense in terms of the behavior of $W_p(t)$ as $t \to \infty$, in a way parallel to that of Ref. 3. We are of the opinion that, within the context of the present paper, $W_1(t)$ is the most natural member of the above family for use in stability studies in the linear theory and that $W_1(t)$ has all the advantages possessed by $A(t)$ in this regard. However, the behavior of $W_p(t)$ in the limit $t \to \infty$ is also of interest when $p \neq 1$. In Sec. IV, the behavior in question is studied for $p \geq 1$ by exploiting the asymptotic form of $\varphi(t)$ in this limit. New sufficient conditions for stability [Theorems 4.1 and 4.3] and instability [Theorems 4.2 and 4.4] are established for a wide variety of initial distributions and initial perturbations. On the basis of these results we are led to the important result that, generally speaking, stability in either our sense or in that of Backus has very little to do with the smoothness properties, as specified in Sec. IV, of the initial perturbations. We conclude Sec. IV with a comparison of our theorems on stability and instability with those of Backus.

Examples of unperturbed distributions illustrating theorems of Sec. IV are given in the Appendix.

II. A BASIC INITIAL-VALUE PROBLEM FOR PLASMA OSCILLATIONS IN THE LINEAR THEORY AND RIGOROUS DERIVATION OF AN INTEGRAL EQUATION FOR THE ELECTRIC POTENTIAL IN THAT THEORY

Let $\mathbf{r} = (x, y, z)$ and $\mathbf{v} = (u, v, w)$ stand for the position and velocity vectors, respectively, of an unbound electron pertaining to a completely ionized gas, the cited components of $\mathbf{r}$ and $\mathbf{v}$ being referred to a common Cartesian frame which will be conveniently chosen in a manner specified below. Let $f_0(\mathbf{v})$ represent a one-electron distribution function with respect to the variables x, y, z, u, v, w, this function being independent of $\mathbf{r}$ and of the time t.

The change induced in this distribution function, which we regard as the initial electron distribution, by externally produced disturbances will be denoted by $f_1(\mathbf{r}, \mathbf{v}, t)$. We shall work within the usual approximation in which $f_1(\mathbf{r}, \mathbf{v}, t)$ and the electric potential $\Phi(\mathbf{r}, t)$ produced by the charged particles of the plasma satisfy the linearized Vlasov equation

$$\partial f_1(\mathbf{r}, \mathbf{v}, t)/\partial t + \mathbf{v}\cdot\nabla_r f_1(\mathbf{r}, \mathbf{v}, t) + (e/m)[\nabla_v f_0(\mathbf{v})\cdot\nabla_r\Phi(\mathbf{r}, t)] = 0, \tag{2.1a}$$

and the Poisson equation

$$\nabla_r^2\Phi(\mathbf{r}, t) = 4\pi en\int_V f_1(\mathbf{r}, \mathbf{v}, t)\, d\mathbf{v} \tag{2.1b}$$

appropriate to the normalization

$$\int_V f_0(\mathbf{v})\, d\mathbf{v} = 1. \tag{2.2}$$

Here e and m stand for the absolute value of the electronic charge and for the electronic mass respectively, and n represents the average electronic number density. The notation $\int_V R(\mathbf{v})\, d\mathbf{v}$ serves to denote the integral of a given function $R(\mathbf{v})$ over the entire velocity space $V = \{(u, v, w) \mid |\mathbf{v}| < \infty\}$.

We are interested in solutions of (2.1a) and (2.1b) such that

$$f_1(\mathbf{r}, \mathbf{v}, t) = G(\mathbf{v}, t)e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{v}t)}, \tag{2.3a}$$

$$\Phi(\mathbf{r}, t) = -(4\pi en/k^2)\varphi(t)e^{i\mathbf{k}\cdot\mathbf{r}}, \tag{2.3b}$$

where $G(\mathbf{v}, t)$ and $\varphi(t)$ are independent of $\mathbf{r}$ and, of course, $k \equiv |\mathbf{k}|$ and n are always taken to be positive.

It is expedient to begin by a purely formal derivation of the pertinent equations obeyed by $G(\mathbf{v}, t)$ and $\varphi(t)$.

In this derivation and throughout this paper, we shall choose the Cartesian frame mentioned earlier in a familiar and convenient way, namely, in such a way that $\mathbf{k} = (k, 0, 0)$, i.e., so that u is the component of $\mathbf{v}$ parallel to $\mathbf{k}$.

Hence Eqs. (2.1a) and (2.3a) yield

$$\partial G(\mathbf{v}, t)/\partial t - H(\mathbf{v})e^{iku t}\varphi(t) = 0, \tag{2.4}$$

where

$$H(\mathbf{v}) \equiv (i\omega_p^2/k)\partial f_0(\mathbf{v})/\partial u, \tag{2.5}$$

and where $\omega_p \equiv (4\pi ne^2/m)^{\frac{1}{2}}$. Therefore

$$G(\mathbf{v}, t) = G(\mathbf{v}) + H(\mathbf{v})\int_0^t e^{iku t'}\varphi(t')\, dt', \tag{2.6}$$

if $G(\mathbf{v}, t)$ is subjected to the initial condition

$$G(\mathbf{v}) = G(\mathbf{v}, 0). \tag{2.7}$$

From (2.1b), (2.3a), and (2.3b), one obtains:

$$\varphi(t) = \int_V e^{-iku t}G(\mathbf{v}, t)\, d\mathbf{v}. \tag{2.8}$$

Substituting (2.6) into (2.8), it is seen that $\varphi(t)$ satisfies the integral equation

$$\varphi(t) = g(t) + \int_0^t h(t - t')\varphi(t')\, dt', \tag{2.9}$$

where

$$g(t) \equiv \int_V e^{-iku t}G(\mathbf{v})\, d\mathbf{v}, \tag{2.10}$$

$$h(t) \equiv \int_V e^{-iku t}H(\mathbf{v})\, d\mathbf{v}. \tag{2.11}$$

Within the context of these heuristic considerations, we conclude that the problem of determining a function $G(\mathbf{v}, t)$ with a prescribed initial value $G(\mathbf{v})$ can be reduced to the much simpler problem of solving (2.9), since the desired $G(\mathbf{v}, t)$ is then given by (2.6). The remainder of this section is devoted to a rigorous investigation of solutions of (2.1a) and (2.1b) obeying (2.3a) and (2.3b). This investigation will provide a firm mathematical basis for (2.9) and for the way of computing $G(\mathbf{v}, t)$ just mentioned.

Henceforth, all references to measurability in the text should be understood in the Lebesgue sense. In this paper, integrability will mean, in general, only Lebesgue integrability[7] and the integrals of this paper are intended to be interpreted solely as Lebesgue integrals in the general case. The fact that in particular instances (for example, for the one-dimensional integrations over intervals of the t axis which occur here) our remarks on integrability in the text and the pertinent integrals can also be interpreted in the sense of Riemann should be perfectly obvious to the reader and will not be mentioned any further.

Besides always requiring that $f_0(\mathbf{v})$ be integrable over V, as expressed by (2.2), we shall invariably suppose that this integrability property is possessed by $G(\mathbf{v})$ and $H(\mathbf{v})$, i.e., that

$$\int_V |G(\mathbf{v})|\, d\mathbf{v} < \infty, \quad \int_V |H(\mathbf{v})|\, d\mathbf{v} < \infty. \tag{2.12}$$

We pause to derive some useful properties of $g(t)$ and $h(t)$. Notice that (2.12) and Fubini's theorem

[7] In this paper, Lebesgue integrability will always be taken to mean Lebesgue summability, i.e., a function will be regarded here as being integrable in the Lebesgue sense over a given set if its integral in this sense over the set in question exists and is finite.

imply that (2.10) and (2.11) can be rewritten as follows:

$$g(t) = \int_{-\infty}^{\infty} e^{-iku t} G(u)\, du, \tag{2.13}$$

$$h(t) = \int_{-\infty}^{\infty} e^{-iku t} H(u)\, du, \tag{2.14}$$

where

$$G(u) \equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(\mathbf{v})\, dv\, dw, \tag{2.15}$$

$$H(u) \equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H(\mathbf{v})\, dv\, dw. \tag{2.16}$$

From (2.12), (2.15), (2.16), and Fubini's theorem, we conclude that $G(u)$ and $H(u)$ are in $L_1(-\infty, \infty)$. Therefore, $g(t) \to 0$ and $h(t) \to 0$ as $|t| \to \infty$ by the Riemann–Lebesgue lemma, and $g(t)$ and $h(t)$ are continuous functions of t for all (finite) t by an elementary property of Fourier transforms of functions in $L_1(-\infty, \infty)$. Thus, $g(t)$ and $h(t)$ are uniformly bounded over the real line.

We now turn to the formulation of two basic problems of the linear theory.

The next paragraph is devoted to the statement of the first of these problems, which is one of the most general physically interesting initial-value problems relative to solutions of the linearized Vlasov equation and of the Poisson equation obeying (2.3a) and (2.3b). In that paragraph and in the remainder of this Section, V' represents the subset of points of V on which both of the functions $G(\mathbf{v})$ and $H(\mathbf{v})$ exist and are finite. It is plain from the integrability conditions (2.12) that V' differs from V by at most a set of (three-dimensional Lebesgue) measure zero.[8]

Problem A. *For prescribed $G(\mathbf{v})$ and $H(\mathbf{v})$, find functions $f_1(\mathbf{r}, \mathbf{v}, t)$ and $\Phi(\mathbf{r}, t)$ with the following properties. The latter two functions have the respective forms (2.3a) and (2.3b) for all $\mathbf{r}$, every $\mathbf{v} \in V'$, and each $t \geq 0$. The corresponding functions $G(\mathbf{v}, t)$ and $\varphi(t)$ are required to be finite at each such $\mathbf{v}$ and t. Moreover, such a $G(\mathbf{v}, t)$ is required to satisfy the initial condition (2.7) and to possess these further properties: (1) $\partial G(\mathbf{v}, t)/\partial t$ exists and is finite for each $\mathbf{v}$ in the cited range and for each $t > 0$, and for every such $\mathbf{v}$ and for $t = 0$ the right-hand derivative of $G(\mathbf{v}, t)$ with respect to t exists and is finite; (2) for every finite closed interval $[0, T]$,[9] $G(\mathbf{v}, t)$ is integrable over the product space $V \times [0, T]$, i.e.,*

$$\iint_{V\times[0,T]} |G(\mathbf{v}, t)|\, d\mathbf{v}\, dt < \infty. \tag{2.17}$$

Finally, the said two functions $f_1(\mathbf{r}, \mathbf{v}, t)$ and $\Phi(\mathbf{r}, t)$ are required to satisfy (2.1a) for each $\mathbf{r}$, all $\mathbf{v} \in V'$, and every $t > 0$; to satisfy for each such $\mathbf{r}$ and $\mathbf{v}$ and for $t = 0$ the modification of (2.1a) in which $\partial f_1(\mathbf{r}, \mathbf{v}, t)/\partial t$ is replaced by the corresponding right-hand derivative with respect to t; and to satisfy (2.1b) over these last ranges of $\mathbf{r}$ and $\mathbf{v}$ for all $t \geq 0$.

The second problem is

Problem B. *For prescribed $g(t)$ and $h(t)$, i.e., for given $G(u)$ and $H(u)$, find a function $\varphi(t)$ which exists and is finite at all $t \geq 0$, which is integrable over each $[0, T]$, and which obeys (2.9) for every $t \geq 0$.*

Since $g(t)$ and $h(t)$ are continuous in t over this last range, one knows from the theory of Volterra integral equations of the second kind that Problem B has a unique solution $\varphi(t)$ which is continuous on every closed interval $[0, T]$, i.e., right-continuous at $t = 0$ and continuous at each $t > 0$.

The existence and uniqueness of solution of Problem B will be exploited below to establish corresponding properties for Problem A.

Our first theorem maps each solution of Problem A on the unique solution of Problem B.

Theorem 2.1. *Let $f_1(\mathbf{r}, \mathbf{v}, t)$ and $\Phi(\mathbf{r}, t)$ constitute a solution of Problem A for prescribed $G(\mathbf{v})$ and $H(\mathbf{v})$. Then the function $\varphi(t)$ appearing in (2.3b) is the solution of Problem B for the given $G(\mathbf{v})$ and $H(\mathbf{v})$, and the function $G(\mathbf{v}, t)$ in (2.3a) is determined in terms of this $\varphi(t)$ by (2.6) for every $\mathbf{v} \in V'$ and each $t \geq 0$.*

Proof: First we prove that (2.6) is true in the sense of the theorem.

If the functions $f_1(\mathbf{r}, \mathbf{v}, t)$ and $\Phi(\mathbf{r}, t)$ fulfill the requirements of Problem A, (2.4) is evidently valid for the corresponding functions $G(\mathbf{v}, t)$ and $\varphi(t)$ for each $\mathbf{v} \in V'$ and each $t > 0$. Now, property (1) implies that this $G(\mathbf{v}, t)$ is an absolutely continuous function of t over each $[0, T]$ for every fixed $\mathbf{v} \in V'$. Hence, since (2.7) holds over V' in this case, we conclude that the $G(\mathbf{v}, t)$ in question obeys the equation

$$G(\mathbf{v}, t) = G(\mathbf{v}) + \int_0^t \frac{\partial G(\mathbf{v}, t')}{\partial t'}\, dt' \tag{2.18}$$

[8] The dimensionality of the Lebesgue measure corresponding to the terms "a.e." and "almost all" in the text should be obvious always from the context in which these terms occur. It should be kept in mind that henceforth in this paper t is always finite, so that, for example, "each $t \geq 0$" should be understood to mean "each finite $t \geq 0$." The only restriction imposed on $\mathbf{r}$ in this paper is that $|\mathbf{r}| < \infty$.

[9] As usual, the symbol $[a, b]$ denotes a closed interval. The symbol $[0, T]$ should always be understood as a finite closed interval.

over this range of $\mathbf{v}$ for each $t \geq 0$. Combining (2.18) with (2.4) we deduce that (2.6) obtains in the manner prescribed in the theorem.

Next we establish that the function $\varphi(t)$ in the previous paragraph, which is finite at each $t \geq 0$ because of a requirement of Problem A, is the solution of Problem B for the $G(\mathbf{v})$ and $H(\mathbf{v})$ of interest.

It should be obvious that, for $t \geq 0$, (2.8) is rigorously correct for the present $G(\mathbf{v}, t)$ and $\varphi(t)$. From (2.8), property (2), and Fubini's theorem, we see that this $\varphi(t)$ is integrable over any $[0, T]$. Hence, because of (2.12), the said $\varphi(t)$ is such that $H(\mathbf{v})\varphi(t)$ is integrable over each $V \times [0, T]$. This last integrability property and Fubini's theorem allow us to interchange the order of the integrations over V and $[0, t]$ occurring in

$$\int_V \left(\int_0^t e^{ikut'} \varphi(t')\, dt' \right) e^{-ikut} H(\mathbf{v})\, d\mathbf{v}.$$

Our heuristic derivation of (2.9) is easily justified for the $\varphi(t)$ in question for $t \geq 0$ by exploiting, in particular, the validity of this interchange of integrals and the noted validity of (2.6) and (2.8) under the conditions of interest. Hence this $\varphi(t)$ is indeed the solution of the pertinent problem B, so that the proof of the theorem is complete.

Our second theorem shows that the solution of Problem B yields a solution of Problem A and that this last solution is unique to within suitable null sets.

Theorem 2.2. *Let $G(\mathbf{v})$ and $H(\mathbf{v})$ be prescribed. Let $\varphi(t)$ be the solution of the corresponding Problem B and define $G(\mathbf{v}, t)$ by (2.6) everywhere on V' for $t \geq 0$ in terms of the functions $G(\mathbf{v})$ and $\varphi(t)$ just mentioned. Define $f_1(\mathbf{r}, \mathbf{v}, t)$ in terms of this $G(\mathbf{v}, t)$ by (2.3a) for every $\mathbf{r}$ and over the range of $\mathbf{v}$ and t last cited, and define $\Phi(\mathbf{r}, t)$ in terms of this $\varphi(t)$ by (2.3b) for all such $\mathbf{r}$ and t. Then $f_1(\mathbf{r}, \mathbf{v}, t)$ and $\Phi(\mathbf{r}, t)$ constitute a solution of the Problem A pertaining to this $G(\mathbf{v})$ and $H(\mathbf{v})$. Moreover, any functions $f_1'(\mathbf{r}, \mathbf{v}, t)$ and $\Phi'(\mathbf{r}, t)$ satisfying the latter Problem A coincide with the respective functions $f_1(\mathbf{r}, \mathbf{v}, t)$ and $\Phi(\mathbf{r}, t)$ of the last sentence for all $\mathbf{r}$, all $\mathbf{v} \in V'$, and each $t \geq 0$.*

Proof: Define $\varphi(t)$ as stated in the theorem. Then the right-hand side of (2.6) is defined and finite for $\mathbf{v} \in V'$ and $t \geq 0$, since then $\varphi(t)$, being the solution of the pertinent Problem B, is integrable over all $[0, T]$. Hence the specification of $G(\mathbf{v}, t)$ by means of (2.6) mentioned in the theorem actually yields a function $G(\mathbf{v}, t)$ which assumes a finite value at each $\mathbf{v} \in V'$ and $t \geq 0$. From the fact that (2.6) holds for every such $\mathbf{v}$ and t for the functions $G(\mathbf{v}, t)$ and $\varphi(t)$ of the theorem, and from the continuity of this $\varphi(t)$ on every $[0, T]$, it follows first that these two functions satisfy (2.4) over the last cited range of $\mathbf{v}$ and for all $t > 0$, and second that the two functions in question obey for all such $\mathbf{v}$ and for $t = 0$ a modified version of (2.4) in which $\partial G(\mathbf{v}, t)/\partial t$ is replaced by the corresponding right-hand derivative with respect to t. Hence it should be clear that the $G(\mathbf{v}, t)$ under discussion possesses property (1).

To prove that this $G(\mathbf{v}, t)$ also has property (2), we observe that the circumstance that (2.6) is satisfied over V' for $t \geq 0$ in the case of interest implies that the inequality

$$|G(\mathbf{v}, t)| \leq |G(\mathbf{v})| + |H(\mathbf{v})| \int_0^t |\varphi(t')|\, dt' \qquad (2.19)$$

is true on the same ranges of $\mathbf{v}$ and t. Because the integral over $[0, t]$ in (2.19) is itself integrable over any $[0, T]$, due to the fact that the present $\varphi(t)$ enjoys this property, and since (2.12) obtains, the right-hand side of (2.19) is integrable over each $V \times [0, T]$. Therefore, the $G(\mathbf{v}, t)$ under discussion fulfills (2).

Define the functions $f_1(\mathbf{r}, \mathbf{v}, t)$ and $\Phi(\mathbf{r}, t)$ as prescribed in the theorem. Then the fact that (2.4) and its cited modification are valid over the above ranges and an elementary computation show that (2.1a) and its modification mentioned in the statement of Problem A are satisfied by these two functions in the sense specified by the latter statement.

A trivial calculation shows that the last two functions also satisfy (2.1b) as specified in Problem A, provided that the corresponding $\varphi(t)$ and $G(\mathbf{v}, t)$ obey (2.8) for $t \geq 0$. That (2.8) holds in this sense follows with the aid of (2.6), (2.9), to (2.11), and Fubini's theorem. The latter is employed in a way similar to that in the proof of Theorem 2.1, its present use being justified for reasons identical to those mentioned in that earlier application.

We have thus established that the functions $f_1(\mathbf{r}, \mathbf{v}, t)$ and $\Phi(\mathbf{r}, t)$ defined in the present theorem constitute a solution of Problem A. We proceed to show that the solution of this problem is unique in the specified sense, thereby completing the proof of the theorem.

Consider two pairs of functions, say $f_1(\mathbf{r}, \mathbf{v}, t)$, $\Phi(\mathbf{r}, t)$ and $f_1'(\mathbf{r}, \mathbf{v}, t)$, $\Phi'(\mathbf{r}, t)$, both of these pairs satisfying the conditions of Problem A for the same $G(\mathbf{v})$ and $H(\mathbf{v})$. If $G(\mathbf{v}, t)$ and $\varphi(t)$ correspond to the first pair and $G'(\mathbf{v}, t)$ and $\varphi'(t)$ to the second one, in the sense of (2.3a) and (2.3b), the uniqueness assertion of the theorem is equivalent to the state-

ment that $G'(\mathbf{v}, t) = G(\mathbf{v}, t)$ and $\varphi'(t) = \varphi(t)$ on V' and for $t \geq 0$. Now, Theorem 2.1 implies that $G(\mathbf{v}, t)$ obeys (2.6) over such ranges of $\mathbf{v}$ and t and that $G'(\mathbf{v}, t)$ is equal to the right-hand side of (2.6) over the ranges in question, provided that $\varphi(t)$ is replaced by $\varphi'(t)$ in this right-hand side. Using Theorem 2.1, it is also seen that $\varphi(t)$ and $\varphi'(t)$ are solutions of Problem B for the same $G(\mathbf{v})$ and $H(\mathbf{v})$, and hence that $\varphi'(t) = \varphi(t)$ in the stated sense by virtue of the uniqueness of solution of this problem. One therefore concludes from the structure of (2.6) that $G'(\mathbf{v}, t) = G(\mathbf{v}, t)$ in the desired sense, so that the above uniqueness assertion concerning Problem A has been proved.

III. BEHAVIOR OF THE ELECTRIC POTENTIAL FOR LONG TIMES

Before discussing the behavior of $\varphi(t)$ in the long-time limit, it is expedient to present some definitions and auxiliary results, which include Lemma 3.1 on the zeros of the Landau denominator. Naturally, the functions $\varphi(t)$ occurring henceforth should always be understood to be solutions of the appropriate Problem B.

We first introduce the class of functions $E(a)$, constituted by those functions $f(t)$ for which $f(t)e^{at} \in L_2(0, \infty)$ for some real a. Classes of this type will play an essential part in the remainder of this paper. We believe that the attention paid to such classes in this investigation is justified *a posteriori* by the broad scope of new results on the temporal evolution of plasma oscillations obtained here by invoking the properties of functions pertaining to the classes in question.

Let us denote the Laplace transform of $f(t)$ by $\bar{f}(s)$, where $s = \sigma + i\tau$ (σ, τ real), i.e.,

$$\bar{f}(s) \equiv \int_0^\infty e^{-st} f(t)\, dt. \tag{3.1}$$

Theorems of Paley and Wiener[10] imply that a necessary and sufficient condition for $f(t)$ to belong to $E(a)$ for a given real a is that $\bar{f}(s)$ have the following properties:

(a) $\bar{f}(s)$ is regular for $\sigma > -a$;

(b) $\sup_{\sigma > -a} \int_{-\infty}^{\infty} |\bar{f}(\sigma + i\tau)|^2\, d\tau < \infty$.

[10] R. E. A. C. Paley and N. Wiener, *Fourier Transforms in the Complex Domain* (American Mathematical Society, Providence, Rhode Island, 1934), Vol. XIX. See especially Theorem I in Sec. 2 and Theorems IV and V in Sec. 3.

These properties will be particularly useful in the present section.

Although the functions $g(t)$ and $h(t)$ of this paper are automatically in every class $E(-\epsilon)$, $\epsilon > 0$, because of the integrability of $G(u)$ and $H(u)$ over $[-\infty < u < \infty]$ guaranteed by (2.12), this integrability is obviously insufficient to ensure that $g(t)$ or $h(t)$ belong to a class $E(c)$ with $c > 0$. In the case of such physically significant distributions $f_0(\mathbf{v})$ as the Maxwellian and Lorentzian ones, $h(t)$ fulfills this last condition. In general, the additional requirements beyond mere integrability which are imposed on $H(u)$ by demanding that $h(t)$ belong to such a class are much stronger than the corresponding additional requirements on $G(u)$ entailed by the condition that $g(t)$ belong to a class of this type. Consider $H(u)$ first. Employing (2.5) and the reality of $f_0(\mathbf{v})$, we see from (2.14) that $H(u)$ is purely imaginary, so that $|h(-t)| = |h(t)|$. Thus, if $h(t)e^{bt} \in L_2(0, \infty)$, then also $h(t)e^{-bt} \in L_2(-\infty, 0)$. From this remark and theorems of Paley and Wiener,[11] we infer that $h(t) \in E(b)$ for some $b > 0$ if and only if $H(u)$ has the properties: (a′) it can be continued analytically into the strip $-b/k < u' < b/k$ of the $u + iu'$ plane [modifying $H(u)$, if necessary, on a set of measure zero]; (b′) the analytic continuation $H(u + iu')$ of $H(u)$ possesses the property that

$$\sup_{|u'| < b/k} \int_{-\infty}^{\infty} |H(u + iu')|^2\, du < \infty.$$

Since $G(u)$ need not be real nor purely imaginary, nor even or odd, the specification of $g(t)$ for $t \geq 0$ tells one nothing about $g(t)$ when $t < 0$ in the general case. This fact is reflected in the circumstance that, even if $g(t) \in E(b)$ for some $b > 0$, $G(u)$ may, for example, be unbounded[12] and not belong to $L_2(-\infty, \infty)$, or it may have unbounded derivatives.

Lemma 3.1. *The function*

$$\Delta(s) \equiv 1 - \bar{h}(s) \tag{3.2}$$

has no zeros for $\sigma > M \equiv \max_{t \geq 0} h(t)$. *Moreover, if* b *is any real number such that* $h(t) \in E(b)$, *then* $\Delta(s)$ *has at most a finite number of zeros in any vertical strip*

[11] See the cited Theorems I and IV of Paley and Wiener's book in the previous footnote.

[12] Of course, one is naturally inclined to distrust results obtained from the present linearization of the Vlasov theory for the case of unbounded initial perturbations $G(u)$. Whether this intuitive feeling is always in accordance with the facts is quite another matter, which plainly lies outside of the scope of this investigation. Hence it has seemed more prudent not to exclude explicitly such perturbations from this paper.

$$-b < \sigma < \beta, \tag{3.3}$$

where β is any real number greater than $-b$.[13,14]

Proof: The assertion that $\Delta(s) \neq 0$ when $\sigma > M$ is implied by the inequality

$$|\bar{h}(s)| \leq \int_0^\infty |h(t)|\, e^{-\sigma t}\, dt \leq \frac{M}{\sigma}. \tag{3.4}$$

To prove the second assertion of the lemma, we notice that the fact that $h(t) \in E(b)$ for some real b implies that $h(t)e^{-\sigma t} \in L_1(0, \infty)$ for each $\sigma > -b$, by virtue of the Schwarz inequality. Invoking the Riemann–Lebesgue lemma, we therefore conclude that

$$\lim_{|\tau|\to\infty} \bar{h}(\sigma + i\tau) = 0, \tag{3.5}$$

for fixed $\sigma > -b$. Therefore, for each such σ, $\Delta(s) \to 1$ as $|\tau| \to \infty$ in the present case. Combining this result with the regularity of $\Delta(s)$ for $\sigma > -b$, entailed by the hypothesis that $h(t) \in E(b)$, and with the isolated nature of the zeros of $\Delta(s)$ lying in regions of regularity of this function, we easily conclude that $\Delta(s)$ has at most a finite number of zeros in vertical strips (3.3).

In the next theorem, we give a formula for $\varphi(t)$ which is particularly pertinent in the study of the asymptotic properties of this function for long times. In this theorem, we introduce the symbol $\rho \equiv \min\{a, b\}$, where a and b are real numbers such that $g(t) \in E(a)$ and $h(t) \in E(b)$. We have seen earlier in this section that such real numbers always exist for the functions $g(t)$ and $h(t)$ of interest.

Theorem 3.1. *As $t \to \infty$, we have*

$$\varphi(t) = \sum_{\mathrm{Re}\, s_i > -\rho} P_i(t)e^{s_i t} + g(t) + O(e^{-(\rho-\epsilon)t}). \tag{3.6}$$

Here, $P_i(t)e^{s_i t}$ is the residue of $\bar{g}(s)e^{st}/\Delta(s)$ at the zero s_i of $\Delta(s)$, so that

$$P_i(t) = \sum_{m=0}^{n_i-1} A_{i,m}t^m, \tag{3.7}$$

where n_i is the multiplicity of s_i and the $A_{i,m}$ are constants; the sum in (3.6) ranges over the finite set of zeros, if any, in the half-plane $\sigma > -\rho$, and is defined as zero if no such zeros occur; and $\epsilon > 0$ is arbitrary.

Proof: We first show that

$$\varphi(t) = g(t) + \frac{1}{2\pi i}\int_{\alpha-i\infty}^{\alpha+i\infty} \frac{\bar{g}(s)\bar{h}(s)}{\Delta(s)}\, e^{st}\, ds \tag{3.8}$$

for each $t \geq 0$, where α is any real number greater than M. Thus $\Delta(s) \neq 0$ for $\sigma \geq \alpha$.

To derive (3.8), we begin by observing that if $\beta > -\rho$ is a fixed real number such that $\Delta(\beta + i\tau) \neq 0$ for $|\tau| < \infty$ then

$$\left|\int_{\beta-i\infty}^{\beta+i\infty} \frac{\bar{g}(s)\bar{h}(s)}{\Delta(s)}\, e^{st}\, ds\right| \leq e^{\beta t}\int_{-\infty}^{\infty} \left|\frac{\bar{g}(\beta + i\tau)\bar{h}(\beta + i\tau)}{\Delta(\beta + i\tau)}\right| d\tau < N(\beta)e^{\beta t} \tag{3.9}$$

for each real t, where $N(\beta)$ is a finite number independent of t. In fact, $\bar{g}(s)$ and $\bar{h}(s)$ are square-integrable along any vertical line in the half-plane $\sigma > -\rho$, so that $\bar{g}(s)\bar{h}(s)$ is absolutely integrable along any such line. Moreover, for any such β, the regularity of $\Delta(s)$ for $\sigma > -\rho$ and the established property that, for any such σ, $\Delta(s) \to 1$ as $|\tau| \to \infty$ imply that $\inf_{|\tau|<\infty} |\Delta(\beta + i\tau)| > 0$. The remarks in the preceding two sentences yield (3.9).

The basic element in our proof of (3.8) is a powerful theorem of Titchmarsh,[15] whose hypotheses are satisfied by $g(t)$ and $h(t)$ because of the continuity and boundedness of these functions. This theorem implies that (3.8) holds for $t \geq 0$ in the mean-square sense, i.e., with

$$\int_{\alpha-i\infty}^{\alpha+i\infty}$$

replaced by

$$\lim_{\tau\to\infty} \int_{\alpha-i\tau}^{\alpha+i\tau}.$$

But (3.9) implies that the integral in (3.8) converges absolutely for every real t, and therefore (3.8) holds a.e. on $t \geq 0$. Now, the right-hand side of (3.8) is

[13] Notice that Lemma 3.1 holds under the sole assumption that $h(t)$ is an *arbitrary* element of some class $E(b)$, provided that we appropriately modify the definition of M in the said lemma.

[14] Results on the zeros of $\Delta(s)$ are obtained also in Refs. 2 and 3. (a) The italicized result on the zeros of $\Delta(s)$ on p. 1060 of Ref. 2 is a special case of Lemma 3.1. (b) Lemma 3, pp. 181–182 of Ref. 3 concerns the positions of the zeros of $\Delta(s)$ under the sole conditions that $df_0(u)/du$ and, presumably, $f_0(u)$ [see Eq. (A1) of the Appendix] are in $L_1(-\infty, \infty)$. However, these conditions are insufficient to imply all the conclusions of this last lemma. In fact, the proof in Ref. 3 of the lemma in question involves the use of the formula for $\mathcal{L}(s)$ occurring in Eq. (21) of this reference and lying closest to the center of p. 181 thereof, a formula which is not generally valid under the conditions on $f_0(u)$ just alluded to. When s is not real the formula in question, which can be obtained by an obvious partial integration, holds, for example, if $f_0(u)$ is absolutely continuous over all closed finite intervals, $|f_0(u)/u| \to 0$ as $|u| \to \infty$, and $df_0(u)/du \in L_1(-\infty, \infty)$. A similar criticism can be made concerning the validity of the partial integration with respect to u employed by Backus to prove the inequality (14) of Ref. 3.

[15] E. C. Titchmarsch, *Introduction to the Theory of Fourier Integrals* (Oxford University Press, New York, 1948), 2nd ed., p. 132, Theorem 147. See, in particular, Eq. (11.5.3) in this last page.

continuous at each real t because $g(t)$ has this continuity property and because the integral in (3.8) also possesses it by virtue of its absolute convergence and of the Riemann–Lebesgue lemma. Since the left-hand side of (3.8) is continuous over each $[0, T]$ because it is the solution of Problem B, it follows that (3.8) is not only satisfied a.e. on $t \geq 0$, but everywhere on this range.

Second, we derive (3.6) from (3.8). By Lemma 3.1, the number of zeros s_i in the half-plane $\sigma > -\rho$ is at most finite. Therefore, there exists a constant $\delta > 0$ such that $\Delta(s) \neq 0$ on the straight line $[-(\rho - \delta) - i\infty, -(\rho - \delta) + i\infty]$ and such that the zeros of $\Delta(s)$ at the right of this line are the same as those in the half-plane $\sigma > -\rho$. Now, the integrand of the integral in (3.8) tends to zero as $|\tau| \to \infty$ for fixed $\sigma > -\rho$, by virtue of the fact that (3.5) and the analogous equation for $\bar{g}(s)$ hold for all such σ. Hence, the usual argument of bounded convergence and the analyticity of $\bar{g}(s)$ and $\bar{h}(s)$ when $\sigma > -\rho$ allow us to evaluate the integral in (3.8) by replacing the Bromwich path $[\alpha - i\infty, \alpha + i\infty]$ by $[-(\rho - \delta) - i\infty, -(\rho - \delta) + i\infty]$, taking into account the poles of the corresponding integrand crossed during this path displacement. Since the residues $P_i(t)e^{s_i t}$ of $\bar{g}(s)e^{st}/\Delta(s)$ are equal to those of $\bar{g}(s)\bar{h}(s)e^{st}/\Delta(s)$ in the common region of analyticity of $\bar{g}(s)$ and $\bar{h}(s)$, we thus find

$$\varphi(t) = g(t) + \sum_{\mathrm{Re}\, s_i > -\rho} P_i(t)e^{s_i t} + \frac{1}{2\pi i}\int_{-(\rho-\delta)-i\infty}^{-(\rho-\delta)+i\infty} \frac{\bar{g}(s)\bar{h}(s)}{\Delta(s)} e^{st}\, ds \qquad (3.10)$$

for $t \geq 0$. Since $\delta > 0$ can be chosen smaller than any $\epsilon > 0$ and since the integral in (3.10) is $O(e^{-(\rho-\delta)t})$ as $t \to \infty$ by virtue of (3.9), the theorem follows.

Necessary and sufficient conditions for the occurrence of an important type of damped functions $\varphi(t)$ are given in

Theorem 3.2. *Let $h(t)$ satisfy the double requirement that $h(t) \in E(b)$ for some $b > 0$ and that $\Delta(s) \neq 0$ for $\sigma \geq 0$. Then a necessary and sufficient condition for $\varphi(t)$ to be of negative exponential order when $t \to \infty$ is that $g(t)$ have this same property. Moreover, if $h(t)$ fulfills this double requirement, a necessary and sufficient condition for $\varphi(t) \in E(c)$ for some $c > 0$ is that $g(t) \in E(a)$ for some $a > 0$.*

Proof: We shall only prove the portion of this theorem concerning necessary and sufficient conditions for $\varphi(t)$ to be of negative exponential order in the limit $t \to \infty$. This is the most interesting portion from the standpoint of the subsequent discussions of this section. The remainder of the theorem can be established by arguments practically identical to those given below.

(a) *Necessity.* Let positive constants b and c exist such that $h(t) \in E(b)$ and $\varphi(t) = O(e^{-ct})$ for $t \to \infty$. Denote by μ an arbitrary positive constant such that $\mu < \min\{b, c\}$. The assumptions on $h(t)$ and $\varphi(t)$ just made evidently imply that $h(t)$, $\varphi(t) \in E(\mu)$. Exploiting (2.9) and the Schwarz inequality, we therefore find under the present circumstances:

$$\begin{aligned} |g(t)| &\leq |\varphi(t)| + \int_0^t |h(t-t')|\, |\varphi(t')|\, dt' \\ &= |\varphi(t)| + e^{-\mu t}\int_0^t |e^{\mu(t-t')}h(t-t')|\, |e^{\mu t'}\varphi(t')|\, dt' \\ &\leq |\varphi(t)| + e^{-\mu t}\left\{\int_0^\infty |e^{\mu t'}h(t')|^2\, dt'\right\}^{\frac{1}{2}} \\ &\times \left\{\int_0^\infty |e^{\mu t'}\varphi(t')|^2\, dt'\right\}^{\frac{1}{2}} = O(e^{-ct}) + O(e^{-\mu t}) \qquad (3.11) \end{aligned}$$

for $t \to \infty$, which concludes our proof of necessity.

(b) *Sufficiency.* Let $g(t) = O(e^{-\alpha t})$ for some $\alpha > 0$ as $t \to \infty$ and let $h(t)$ obey the double requirement of the theorem. If a is any positive number smaller than α, it then follows that $g(t) \in E(a)$. Therefore, the ρ corresponding to this a and to the b in the first sentence of the theorem is positive and smaller than α, so that $g(t) = O(e^{-\rho t})$ *a fortiori*. Hence, applying Theorem 3.1 to the $g(t)$ and $h(t)$ just specified, keeping in mind that $\Delta(s)$ has no zeros for $\sigma \geq 0$ in the case under discussion, we deduce that

$$\varphi(t) = \sum_{-\rho < \mathrm{Re}\, s_i < 0} P_i(t)e^{s_i t} + O(e^{-(\rho-\epsilon)t}) \qquad (3.12)$$

as $t \to \infty$ for each $\epsilon > 0$ and for the ρ in question, where the sum in (3.12) is taken to range over the finite set of points s_i, if any, in the strip $-\rho < \sigma < 0$ and as zero otherwise. Invoking (3.7), the right-hand side of (3.12) is seen to be a finite sum of functions of negative exponential order. Thus, $\varphi(t)$ is a function of the same kind under the present hypotheses, so that our sufficiency proof is complete.[16]

[16] The sufficiency aspect of Theorem 3.2 concerning solutions $\varphi(t)$ of negative exponential order can be proved without appealing to Theorem 3.1. Let the hypotheses on $g(t)$ and $h(t)$ and the choice of ρ be the same as in part (b) of the proof in the text. Let ν be any positive number smaller than the minimum perpendicular distance of the zeros, if any, in the strip $-\rho < \sigma < 0$ or smaller than ρ if no such zeros occur. Clearly, an equation of the same structure as (2.9) connects the functions $g(t)e^{\nu t}$, $h(t)e^{\nu t}$, and $\varphi(t)e^{\nu t}$. It is easily shown that Theorem 145, p. 304, of Titchmarsh's book in the previous footnote is applicable to this equation satisfied by these three functions. That theorem implies that $\varphi(t) \in E(\nu)$ in

Let us examine (3.12) more closely. When for each s_i in this strip either $\Delta(s_i) = 0$ or $P_i(t) \equiv 0$, (3.12) invites no special comment, for then it merely states that $\varphi(t) = O(e^{-(\rho-\epsilon)t})$ as $t \to \infty$ [$P_i(t) \equiv 0$ for every s_i in the case of the non-Landau damped behavior alluded to in the Introduction and dealt with in Refs. 1 and 2]. Consider next the situation when there exists at least one s_i in the strip $-\rho < \sigma < 0$ such that the corresponding $P_i(t) \not\equiv 0$. In this case, the most significant contribution to $\varphi(t)$ in the limit $t \to \infty$ will obviously be furnished by the group of terms in (3.12) which pertain to the set of s_i in this strip having $P_i(t) \not\equiv 0$ and having the maximum value of $|\mathrm{Re}\, s_i|$. Calling this maximum value γ, we shall say in this case that $\varphi(t)$ damps asymptotically with the rate γ.

It is interesting to discuss the damping of $\varphi(t)$ in the important Maxwellian case:

$$f_0(\mathbf{v}) = \left(\frac{2\pi\Theta}{m}\right)^{\frac{3}{2}} e^{-m|\mathbf{v}|^2/2\Theta} \qquad (\Theta > 0). \tag{3.13}$$

If (3.13) holds, it is easily proved that $h(t) \in E(b)$ for each $b > 0$ and that

$$\Delta(s) = \kappa^{-2}\{1 + \kappa^2 - \pi^{\frac{1}{2}} e^{z^2} \operatorname{erfc}(z)\}, \tag{3.14}$$

where $z \equiv s/2^{\frac{1}{2}}\omega_p\kappa$ and $\kappa \equiv k(m\Theta/4\pi ne^2)^{\frac{1}{2}}$.[17] Evidently, the function $\Delta(s)$ in (3.14) is an entire function of s. It is known that $\Delta(s) \neq 0$ in the half-plane $\sigma \geq 0$ when (3.14) holds.[18] It can also be demonstrated for the Maxwellian distribution that $\Delta(s)$ has an infinite number of zeros in the half-plane $\sigma < 0$, all of these zeros being simple.[19]

Applying Lemma 3.2 to the example under discussion, we conclude that there is a finite number of points s_i which lie closer to the imaginary s axis than all the other s_i when (3.14) holds. In addition, one can show that (3.14) implies that there are exactly two zeros of $\Delta(s)$, s_+ and s_-, closest to this imaginary axis in the limit $\kappa \to +0$. These zeros are given by

$$\begin{aligned} s_\pm &= \tilde{\sigma} \pm i\tilde{\tau}, \\ \tilde{\sigma} &= -\left(\frac{\pi}{8}\right)^{\frac{1}{2}} e^{-\frac{3}{2}} \frac{1}{\kappa^3} e^{-1/2\kappa^2}[1 + O(\kappa^2)] \\ &\qquad + O(e^{-c/2\kappa^2}), \\ \tilde{\tau} &= \omega_p[1 + \tfrac{3}{2}\kappa^2 + O(\kappa^4)] \end{aligned} \tag{3.15}$$

in this limit, where c is a number greater than unity and independent of κ. Equations (3.15) can be proved simply and rigorously by employing the asymptotic formula alluded to in footnote 18 and the classical Lagrange formula for reverting power series. This proof is given elsewhere.[20] Except for the factor $e^{-\frac{3}{2}}$ in the formula for $\tilde{\sigma}$ in Eqs. (3.15), these equations confirm the correctness of the approximate results obtained by Landau[6] for the zeros closest to the imaginary axis in the situation when κ is a small enough positive number.[21]

Let $f_0(\mathbf{v})$ be Maxwellian and suppose that $\kappa > 0$ is so small that s_+ and s_- are closer to the imaginary axis than all the other zeros of $\Delta(s)$. Furthermore, let $G(\mathbf{v})$ be such that $g(t) = O(e^{-at})$ as $t \to \infty$, a being a positive constant. Then $\varphi(t)$ damps asymptotically with the rate $\gamma = |\tilde{\sigma}|$ if and only if $\bar{g}(s)$ does not vanish at both s_+ and s_-.

This assertion is easily proved by using the fact that both $\bar{g}(s)$ and $\Delta(s)$ are regular for $\sigma > -a$ under the assumptions of the last paragraph, by exploiting the simple nature of the zeros s_i in the Maxwellian case, and by applying the above definition of asymptotic damping rates.

IV. STABILITY IN THE LINEAR THEORY

In this section, we shall be interested in the nonnegative quantities

$$W_p(t) \equiv \left[\int_V |G(\mathbf{v}, t)|^p \, d\mathbf{v}\right]^{1/p}, \tag{4.1}$$

---

the present case. Combining the last result with reasoning analogous to that used in obtaining (3.11), we deduce that $\varphi(t) = O(e^{-\rho' t})$ as $t \to \infty$, which completes the desired sufficiency proof. The sufficiency assertion of the final sentence of Theorem 3.2 also follows by the methods of this footnote.

[17] Of the many papers in which formulas equivalent to (3.14) occur, one may cite that of J. D. Jackson, J. Nucl. Energy: Pt C, 1, 171 (1960) [Eq. (A2.8)].

[18] For a proof that (3.14) implies that $\Delta(s) \neq 0$ for $\sigma \geq 0$ see Ref. 3, Sec. VI. One can also prove this fact as follows. Consider a closed path composed of a segment of the imaginary axis and of a semicircle of radius R in the half-plane $\sigma \geq 0$. Using a familiar asymptotic formula for $\operatorname{erfc}(z)$, (3.14) entails that $\Delta(s) \neq 0$ on this path for large enough R and that the change in $\arg \Delta(s)$ around this path approaches zero as $R \to \infty$.

[19] To show that $\Delta(s)$ has an infinite number of zeros in the half-plane $\sigma < 0$ when (3.14) holds, we observe that the right-hand side of (3.14) is an entire function of order 2. Such a function has an infinite number of zeros unless it is of the form $P(z) \exp[Q(z)]$, where $P(z)$ is a polynomial and $Q(z)$ is quadratic in z, by Hadamard's factorization theorem [see, for example, E. C. Titchmarsh, *Theory of Functions* (Oxford University Press, London, 1939), 2nd ed., p. 250]. Since this is not the case here, the present $\Delta(s)$ has an infinite number of zeros, which lie in the half-plane $\sigma < 0$ because of the cited fact that this $\Delta(s)$ has no zeros when $\sigma \geq 0$. Dr. J. N. Hayes (private communication) has observed that the simple nature of all these zeros follows in an elementary way by computing $d\Delta(s_i)/ds$.

[20] See A. W. Sáenz, "Rigorous Treatment of the Propagation and Damping of Small-Amplitude Plasma Waves in an Initially Maxwellian Plasma," NRL Report 6125, August, 1964.

[21] That the factor $e^{-\frac{3}{2}}$ needs to be inserted in the pertinent damping formula of Landau is, of course, well known heuristically. See, for example, A. G. Sitenko and K. N. Stepanov, Soviet Phys.—JETP 4, 512 (1957) [p. 520] and the reference in footnote 17 [Eq. (6.3)]. We are not aware of any rigorous derivation of (3.15) outside of the one cited in the preceding footnote.

where $G(\mathbf{v}, t)$ fulfills the conditions of Problem A and where p is an appropriate positive number.

An attractive feature of the $W_p(t)$ as criteria of stability and instability is that each $W_p(t)$ is small or large, speaking crudely, when $G(\mathbf{v}, t)$ is small or large respectively. In particular, it is plain that $W_p(t) = 0$ for some t if and only if $G(\mathbf{v}, t) = 0$ a.e. on V for that t.

From the standpoint of the present work, $W_1(t)$ has a significant advantage over the $W_p(t)$ with $p \neq 1$ for use in such criteria. In fact, one sees from (2.19) and (4.1) that $W_1(t)$ exists and is finite for each $t \geq 0$ under the basic integrability conditions (2.12) underlying our investigation. However, the $W_p(t)$ with $p \neq 1$ do not exist in general under these conditions alone.

For a given p, we shall say that an undisturbed distribution $f_0(\mathbf{v})$ is stable (unstable) in the $W_p(t)$ sense with respect to initial perturbations $G(\mathbf{v})$ of a given class if, for that p, $W_p(t)$ is bounded (unbounded) as $t \to \infty$ for all (some) $G(\mathbf{v})$ of that class.

Backus[3] has employed the nonnegative quantity

$$A(t) \equiv \int_{-\infty}^{\infty} \left| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(\mathbf{v}, t)\, dv\, dw \right| du \tag{4.2}$$

in his studies of stability in the linear theory. We shall speak of stability (instability) in the $A(t)$ sense if the definition of stability in the previous paragraph holds in terms of $A(t)$, rather than in terms of $W_p(t)$. Notice that stability in the sense of $W_1(t)$ implies stability in that of $A(t)$, while the converse statement obtains for instability in these two senses. The last two assertions follow from the fact that, in particular for $t \geq 0$,

$$A(t) \leq W_1(t), \tag{4.3}$$

which in turn results from (4.1), (4.2), and Fubini's theorem.

In a number of cases, the determination of the boundedness or unboundedness of $W_p(t)$ as $t \to \infty$ for $p > 1$ is no more difficult to study than the corresponding problem for $p = 1$. On the other hand, the methods of this section are inapplicable in the range $p < 1$.[22]

Before embarking on the study of the long-time properties of the $W_p(t)$ with $p \geq 1$ on the basis of the corresponding properties of $\varphi(t)$, we need a generalization of (2.12). This generalization is that there exists a constant $r \geq 1$ such that[23]

$$\left.\begin{aligned} \left[\int_V |G(\mathbf{v})|^p\, d\mathbf{v}\right]^{1/p} &= G_p < \infty, \\ \left[\int_V |H(\mathbf{v})|^p\, d\mathbf{v}\right]^{1/p} &= H_p < \infty, \\ 1 \leq p \leq r. & \end{aligned}\right\} \tag{4.4}$$

Employing (2.3a), (4.1), (4.4), and Minkowski's inequality, one finds that

$$\begin{aligned} W_p(t) &\leq G_p + \left[\int_V \left| H(\mathbf{v}) \int_0^t \varphi(t') e^{ik\mathbf{v}t'}\, dt' \right|^p d\mathbf{v}\right]^{1/p} \\ &\leq G_p + H_p \int_0^t |\varphi(t')|\, dt' \end{aligned} \tag{4.5}$$

over the range of p in (4.4), whenever $t \geq 0$. Using the last inequality (4.5), one finds that $W_p(t) < \infty$ over this range of p for each $t \geq 0$, a straightforward generalization of the parallel result for $W_1(t)$ found earlier.

The conditions (4.4), not stated explicitly in Theorems 4.1, 4.3, and 4.4 below, should be understood, nevertheless, to form part of the respective hypotheses of these three theorems. *Henceforth, we shall usually not state the range of the subscripts p, but, when this omission occurs, it should be understood that* $1 \leq p \leq r$.

Sufficient conditions for stability are furnished by

Theorem 4.1. *Let* $g(t) \in E(a)$ *and let* $h(t) \in E(b)$ *for some* $a > 0$ *and some* $b > 0$. *Let* $\bar{g}(s)/\Delta(s)$ *be analytic for* $\sigma \geq 0$; *for example, let* $\Delta(s) \neq 0$ *for* $\sigma \geq 0$. *Then* $W_p(t)$ *is bounded as* $t \to \infty$.

Proof: By arguments parallel to ones used in the sufficiency proof of Theorem 3.2, we conclude that the hypotheses of the present theorem imply that, for $t \to \infty$, $\varphi(t) - g(t)$ is given by the right-hand side of (3.12), in terms of an appropriate $\rho > 0$. Hence $\varphi(t) - g(t)$ is a continuous function of t ($t \geq 0$) which is of negative exponential order for $t \to \infty$ and hence is in $L_1(0, \infty)$. Since $g(t) \in L_1(0, \infty)$ whenever $g(t) \in E(a)$ for an $a > 0$, we thus see that $\varphi(t) \in L_1(0, \infty)$ if the conditions of the theorem are obeyed. This integrability property of $\varphi(t)$ and (4.5) imply that $W_p(t)$ is bounded as $t \to \infty$ if these conditions are satisfied.

Of the many examples of initial distributions such that $\Delta(s) \neq 0$ for $\sigma \geq 0$ and such that $h(t) \in E(b)$

[22] The restriction to functions $W_p(t)$ with $p \geq 1$ in this section arises because the Minkowski inequality, which is of basic importance in the arguments of this section, is not applicable to the case $p < 1$.

[23] Notice that the existence of an $r \geq 1$ such that $G(\mathbf{v})$ and $H(\mathbf{v})$ obey (4.4) is equivalent to the integrability of $|G(\mathbf{v})|^p$ and $|H(\mathbf{v})|^p$ over V for $p = 1$ and $p = r$, by virtue of an elementary theorem of integration theory.

for some $b > 0$, the most important is the Maxwellian one, in which case $H(\mathbf{v})$ obeys the pertinent condition (4.4) in the range $0 < p < \infty$. We may thus apply Theorem 4.1 to initially Maxwellian plasmas, drawing the conclusion that such plasmas are stable in the $W_p(t)$ sense ($1 \leq p < \infty$) with respect to perturbations for which $g(t) \in E(a)$ for an $a > 0$.

The reader is referred to the Appendix for examples of functions $f_0(\mathbf{v})$ illustrating Theorems 4.2, 4.3, and 4.4 below. Examples of functions $G(\mathbf{v})$ illustrating the theorems of this section are given in the paragraphs containing Eqs. (4.26) and (4.27). The construction of a more comprehensive list of such examples is left to the reader.

Our second theorem provides sufficient conditions for instability.

Theorem 4.2. *Let $f_0(\mathbf{v})$ obey the second inequality (2.12) and let it possess one of the properties (i) or (ii), defined as follows: (i) $\Delta(s)$ has at least one zero in the half-plane $\sigma > 0$; (ii) $h(t) \in E(b)$ for some $b > 0$, $\Delta(s) \neq 0$ if $\sigma > 0$, and $\Delta(s)$ has at least one zero with $n_i \geq 2$ on the line $\sigma = 0$. Let $G(\mathbf{v})$ obey the first inequality (2.12) and, moreover, let it be such that $P_i(t) \not\equiv 0$ for some s_i in the half-plane $\sigma > 0$ when property (i) holds, or such that $g(t) \in E(a)$ for some $a > 0$ and $\bar{g}(s_i) \neq 0$ for some s_i with $n_i \geq 2$ on the line $\sigma = 0$ when (ii) obtains. Then there exist nonnegative constants λ and μ such that $t^{-\lambda}e^{-\mu t}A(t)$, and therefore also $t^{-\lambda}e^{-\mu t}W_1(t)$, is unbounded as $t \to \infty$. If property (i) [(ii)] holds, then $\lambda \geq 0$ (≥ 1) and $\mu > 0$ ($= 0$).*

For each given pair of functions $f_0(\mathbf{v})$ and $G(\mathbf{v})$ which, outside of satisfying all of the respective requirements of the last paragraph, obey (4.4) for an $r > 1$, there are two constants λ and μ with the following properties. These two constants have all of the respective properties mentioned in that paragraph and in addition they are such that, for all p fulfilling the inequality $1 < p \leq r$, $t^{-\lambda}e^{-\mu t}W_p(t)$ is unbounded as $t \to \infty$.

Proof: We begin by considering the portion of the theorem pertaining to the unboundedness of $A(t)$ as $t \to \infty$.

Employing Fubini's theorem, (2.8), and (4.3), we find for $t \geq 0$

$$A(t) \geq |\varphi(t)|. \tag{4.6}$$

This inequality plays a major role in proving the unboundedness property just alluded to.

Let $f_0(\mathbf{v})$ and $G(\mathbf{v})$ satisfy the conditions in the first two sentences of the theorem. Then Theorem 3.1, the elementary fact that $A_{i,n_i-1} \neq 0$ if and only if $\bar{g}(s_i) \neq 0$ [we employ this fact only when property (ii) obtains], and the boundedness of $g(t)$ are easily seen to imply that there are constants σ_0, ν, and $A_{j,\nu}$ having the properties stated below and such that

$$\varphi(t) = t^\nu e^{\sigma_0 t}\left\{\sum_{j=1}^{l} A_{j,\nu}e^{i\tau_j t} + o(1)\right\} \tag{4.7}$$

as $t \to \infty$. When property (i) holds, σ_0 is a positive number equal to the maximum perpendicular distance from the imaginary axis of the zeros s_i in the half-plane $\sigma > 0$ which have $P_i(t) \not\equiv 0$, and σ_0 is zero when (ii) obtains; ν is a nonnegative integer such that $\nu \geq 0$ (≥ 1) when property (i) [(ii)] is present; and τ_j is the imaginary part of the jth zero ($j = 1, \cdots, l \geq 1$) lying on the line $\sigma = \sigma_0$ when either (i) or (ii) occurs. Finally, the constants $A_{j,\nu}$ in (4.7) are defined to be zero if ν is greater than or equal to the multiplicity of the corresponding zero $s_j = \sigma_0 + i\tau_j$ of $\Delta(s)$, this definition being necessary for the validity of (4.7) because of the range of m in (3.7). One finds that not all the $A_{j,\nu}$ ($j = 1, \cdots, l$) are zero in either case (i) or case (ii) [It is elementary to show that (4.7) is no longer true in general in the stated sense for case (ii) if the condition on $\bar{g}(s_i)$ mentioned in the theorem is replaced by the weaker condition that $P_i(t) \not\equiv 0$ for some s_i on the line $\sigma = 0$ with $n_i \geq 2$, even if the remaining requirements on $f_0(\mathbf{v})$ and $G(\mathbf{v})$ invoked above to derive (4.7) are satisfied. This breakdown of (4.7) is symptomatic of the fact that the assertions of the theorem are generally false under these weakened requirements on $G(\mathbf{v})$ when property (ii) holds.]

The sum inside of the curly brackets in (4.7) is an almost periodic function of t which does not vanish identically when either property (i) or property (ii) occurs, by virtue of the last property of the $A_{j,\nu}$. Hence there exists a sequence $\{t_n\}$ ($n = 1, 2, \cdots$) of real numbers, with $t_n \to \infty$ as $n \to \infty$, and a positive number δ independent of n such that

$$\left|\sum_{j=1}^{l} A_{j,\nu}e^{i\tau_j t_n}\right| > \delta > 0 \tag{4.8}$$

for each positive integer n.[24]

[24] That the sum inside of the curly brackets in (4.7) is an almost periodic function of t follows, for example, from the corollary on p. 38 of H. Bohr's *Almost Periodic Functions* (Chelsea Publishing Company, New York, 1947). The result (4.8) is entailed by an elementary property of almost periodic functions $\Lambda(t) \not\equiv 0$, which we shall prove here for completeness. For any given real number $\hat{t}$ such that $\Lambda(\hat{t}) \neq 0$, we select an $\epsilon > 0$ such that $\delta = |\Lambda(\hat{t})| - \epsilon > 0$. By Bohr's

For the functions $f_0(\mathbf{v})$ and $G(\mathbf{v})$ under discussion, we conclude from (4.6), (4.7), and (4.8), and from the stated properties of σ_0 and ν, that there exist nonnegative constants λ and μ which possess the properties specified in the final sentence of the first paragraph of the theorem and which are such that $t_n^{-\lambda}e^{-\mu t_n}A(t_n) \to \infty$ as $n \to \infty$. Hence the proof of the portion of the theorem referring to $A(t)$ is complete.

We now investigate the long-time behavior of $W_p(t)$ in the as yet unexplored range $1 < p \leq r$ [All of the arguments and equations in this investigation hold equally well, of course, for $1 \leq p \leq r$]. In the remainder of the proof of the present theorem, $f_0(\mathbf{v})$ and $G(\mathbf{v})$ will be supposed to obey the requirements laid down in the first sentence of the second paragraph of this theorem.

It is convenient to introduce the notations

$$\|F(u)\|_p \equiv \left[\int_{-\infty}^{\infty} |F(u)|^p\, du\right]^{1/p}, \qquad \|F(u)\|_p^{[a,b]} \equiv \left[\int_a^b |F(u)|^p\, du\right]^{1/p}. \tag{4.9}$$

If $f_0(\mathbf{v})$ and $G(\mathbf{v})$ satisfy the last-mentioned conditions, we obtain by reasoning parallel to that used in arriving at (4.5) and by employing Fubini's theorem:

$$W_p(t) \geq -G_p + \left\|\eta_p(u)\int_0^t \varphi(t')e^{ikut'}\, dt'\right\|_p, \tag{4.10}$$

for $t \geq 0$, where the function

$$\eta_p(u) \equiv \left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |H(\mathbf{v})|^p\, dv\, dw\right]^{1/p} \tag{4.11}$$

is in $L_p(-\infty, \infty)$ as a consequence of this last theorem and of the condition (4.4) on $H(\mathbf{v})$.

Notice that the trivial case $\eta_p(u) = 0$ a.e. for $|u| < \infty$ cannot occur under the present hypotheses on $f_0(\mathbf{v})$. For if $\eta_p(u)$ had this property, then this requirement (4.4) on $H(\mathbf{v})$, (4.11), Fubini's theorem, and an elementary property of Lebesgue integrals would imply that $H(\mathbf{v}) = 0$ a.e. on V. But this would mean that $\Delta(s)$ would have no zeros, in evident contradiction with the occurrence of either property (i) or property (ii). Hence $\|\eta_p(u)\|_p$ is always finite and positive for the $f_0(\mathbf{v})$ obeying such hypotheses.

We can thus choose a finite closed interval I, not containing any of the points $-\tau_j/k$ $(j = 1, \cdots, l)$ and such that the integral of $\eta_p(u)$ over I is finite and positive. Therefore, a simple theorem[25] on Lebesgue integration allows us to infer the existence of a number α which lies in the interior of I and is such that $\|\eta_p(u)\|_p^{[\alpha,\alpha+\omega]}$ is finite and positive for each choice of $\omega > 0$ for which $[\alpha, \alpha + \omega] \subseteq I$. [Obviously, this finiteness and positiveness hold, *a fortiori*, for all $\omega > 0$, since $\eta_p(u) \in L_p(-\infty, \infty)$, but we shall not employ this fact here.] Henceforth, this last condition on $[\alpha, \alpha + \omega]$ will be assumed to obtain always, so that none of the points $-\tau_j/k$ will lie in $[\alpha, \alpha + \omega]$.

Let us now consider the integral over $[0, t]$ occurring in (4.10). From (4.7) and the detailed structure of the $o(1)$ terms therein, we obtain the following formula, which holds for each finite $t > 0$ and each finite u in case (i), and which holds for each such t and each finite u different from all the points $-\tau_j/k$ in case (ii):

$$\int_0^t \varphi(t')e^{ikut'}\, dt' = e^{\sigma_0 t}t' \times \left\{\sum_{j=1}^{l} \frac{A_{j,\nu}e^{i(ku+\tau_j)t}}{[\sigma_0 + i(ku + \tau_j)]} + Q(u, t)\right\}. \tag{4.12}$$

If property (i) holds, the function $Q(u, t)$ [whose explicit form is immaterial here] is bounded and integrable for each $t \neq 0$ over any finite interval on the u axis and

original definition of almost periodicity, there exists a real sequence $\{\tau_n(\epsilon)\}(n = 1, 2, \cdots)$ of (translation) numbers tending to infinity as $n \to \infty$ and such that $|\Delta(\hat{t} + \tau_n(\epsilon)) - \Delta(\hat{t})| \leq \epsilon$ for each positive integer n. Setting $t_n \equiv \hat{t} + \tau_n(\epsilon)$, one then finds that $|\Delta(t_n)| \geq \delta > 0$ for each such n, so that (4.8) obtains as stated.

[25] The theorem in question is as follows. Let there exist an interval J (finite or infinite) such that the function $f(x)$ of the real variable x is integrable over J, the corresponding integral over J is positive, and $f(x) \geq 0$ on J. Then there exists an interior point $c \in J$ such that

$$\int_c^{c+h} f(x)\, dx > 0$$

for every $h > 0$ such that $[c, c + h] \in J$. To prove this theorem, suppose that no such c exists. Then, for each interior point $x \in J$, there exists a $\delta(x) > 0$ such that

$$\int_x^{x+\delta(x)} f(\xi)d\xi = 0.$$

This result and the nonnegativity of $f(x)$ on J imply that

$$\int_x^{x+\tau} f(\xi)d\xi = 0$$

for every such x and each $\tau \in [0, \delta(x)]$. Setting

$$G(x) \equiv \int_a^x f(\xi)\, d\xi,$$

where $a \leq -\infty$ is the inf of the points of J, we thus see that $G(x + \tau) = G(x)$ for all x and τ of this type. Hence the right-hand derivative of $G(x)$ vanishes at each interior point of J, that is, it vanishes a.e. in J. Now, this right-hand derivative is equal to $dG(x)/dx$ whenever the latter exists. But, a.e. in J, $dG(x)/dx$ exists and equals $f(x)$, since $f(x)$ is integrable over J. Hence $f(x) = 0$ a.e., which contradicts our second hypothesis on $f(x)$ and proves the theorem.

$$\lim_{t\to\infty} Q(u, t) = 0 \tag{4.13}$$

for any fixed finite u. If property (ii) occurs, $Q(u, t)$ has poles at the same values $-\tau_j/k$ $(j = 1, 2, \cdots, l)$ of u as has the sum inside of the curly brackets of (4.12). However, when case (ii) obtains, $Q(u, t)$ has the above boundedness and integrability properties for each $t \neq 0$ over each finite closed interval on the u axis not containing any of the points $-\tau_j/k$, and (4.13) holds in this case for each fixed finite u not coinciding with any of the latter points. The presence of poles at these points in case (ii) motivated our choice of intervals $[\alpha, \alpha + \omega]$ not containing any of these points in this proof, intended to apply to cases (i) and (ii).

Since (4.13) is valid for (i) and (ii) whenever $u \in [\alpha, \alpha + \omega]$, we find

$$\lim_{t\to\infty} \|\eta_\nu(u)Q(u, t)\|_p^{[\alpha,\alpha+\omega]} = \|\eta_\nu(u) \lim_{t\to\infty} Q(u, t)\|_p^{[\alpha,\alpha+\omega]} = 0 \tag{4.14}$$

in these two cases by dominated convergence.

To finish the proof of the present theorem, we require a result analogous to (4.8). This result is that, for each sufficiently small ω, there exists a sequence $\{t_n'(\alpha)\}$ $(n = 1, 2, \cdots)$ of real numbers $t_n'(\alpha) \to \infty$ as $n \to \infty$ for fixed α which are such that

$$\left|\sum_{j=1}^{l} \frac{A_{j,\nu}e^{i\tau_j t_n'(\alpha)}}{[\sigma_0 + i(ku + \tau_j)]}\right| > \delta'(\alpha, \omega) > 0 \tag{4.15}$$

for each $u \in [\alpha, \alpha + \omega]$ and each positive integer n. Here $\delta'(\alpha, \omega)$ is independent of u and n.

It is easy to show that (4.15) holds on these ranges of n and u by exploiting the facts about the function $\sum_{j=1}^{l} A_{j,\nu}e^{i\tau_j t}/[\sigma_0 + i(ku + \tau_j)]$ which we proceed to state. First, in particular for any given $u \in [\alpha, \alpha + \omega]$, this function is an almost periodic function of t, which is not identically zero because of the circumstance that not all the $A_{j,\nu}$ vanish under the present conditions. Second, as can be verified immediately, this function is differentiable with respect to u over $[\alpha, \alpha + \omega]$ for each real finite t, the absolute value of the corresponding partial derivative being bounded in this range of u and t by a finite number independent of u and t.[26]

Selecting an $\omega > 0$ so small that (4.15) obtains in the prescribed sense, we find with the aid of (4.12), (4.14), (4.15), and Minkowski's inequality:

$$\begin{aligned}
&\left\|\eta_\nu(u) \int_0^{t_n'(\alpha)} \varphi(t')e^{iku t'}\, dt'\right\|_p \\
&\geq \left\|\eta_\nu(u) \int_0^{t_n'(\alpha)} \varphi(t')e^{iku t'}\, dt'\right\|_p^{[\alpha,\alpha+\omega]} \\
&\geq e^{\sigma_0 t_n'(\alpha)}(t_n'(\alpha))^\nu \\
&\times \left\{\left\|\eta_\nu(u) \sum_{j=1}^{l} \frac{A_{j,\nu}e^{i\tau_j t_n'(\alpha)}}{[\sigma_0 + i(ku + \tau_j)]}\right\|_p^{[\alpha,\alpha+\omega]} - \|\eta_\nu(u)Q(u, t)\|_p^{[\alpha,\alpha+\omega]}\right\} \\
&\geq e^{\sigma_0 t_n'(\alpha)}(t_n'(\alpha))^\nu \\
&\times \{\delta'(\alpha, \omega) - o(1)\} \|\eta_\nu(u)\|_p^{[\alpha,\alpha+\omega]} \to \infty
\end{aligned} \tag{4.16}$$

as $n \to \infty$, since $\delta'(\alpha, \omega)\ \|\eta_\nu(u)\|_p^{[\alpha,\alpha+\omega]} > 0$ for such an ω.

Let $f_0(\mathbf{v})$ and $G(\mathbf{v})$ be given functions fulfilling the conditions of the second paragraph of the theorem. Then we conclude from (4.10) and (4.16), from the above properties of σ_0 and ν, and from our previous discussion of the unboundedness of $A(t)$ in the sense of the theorem, that we can choose two constants λ and μ which are such that $t^{-\lambda}e^{-\mu t}W_p(t)$ is unbounded as $t \to \infty$ when $1 < p \leq r$ and such that they possess all the remaining properties enumerated in the last paragraph of the theorem. Hence our proof of the theorem is complete.

We now consider the case when all the zeros of $\Delta(s)$ in the half-plane $\sigma \geq 0$ are both simple and purely imaginary. The next two theorems give sufficient conditions for stability and instability when $\Delta(s)$ is of this type.

Theorem 4.3. *Let $h(t) \in E(b)$ for some $b > 0$. Let $\Delta(s) \neq 0$ for $\sigma > 0$ and let $\Delta(s)$ have zeros $i\tau_j$ $(j = 1, \cdots, q)$ on the line $\sigma = 0$, all of these zeros being simple. Moreover, let there exist positive constants Γ_p and ξ_p such that*

$$\eta_p(u) \leq \Gamma_p\, |u + \tau_j/k| \tag{4.17}$$

for almost all $u \in [\alpha_{jp}, \beta_{jp}]$ $(j = 1, \cdots, q)$, where $\alpha_{jp} \equiv -\tau_j/k - \xi_p$ and $\beta_{jp} \equiv -\tau_j/k + \xi_p$. Then $W_p(t)$ is bounded as $t \to \infty$ for all $G(\mathbf{v})$ such that $g(t) \in E(a)$ for some $a > 0$.

[26] In fact, let $\Lambda(u, t)$ be a function with the properties enumerated in this paragraph of the text. Consider an arbitrary fixed interval $[\alpha, \alpha + \omega]$, say the interval $[\alpha, \alpha + \omega_0]$. Then $|\Lambda(u, t) - \Lambda(\alpha, t)| \leq |u - \alpha|\ K(\alpha, \omega_0)$ for every $u \in [\alpha, \alpha + \omega_0]$ and each real finite t, $K(\alpha, \omega_0)$ being the finite number independent of u and t cited in the third sentence of that paragraph. Since $\Lambda(\alpha, t) \not\equiv 0$ and since $\Lambda(\alpha, t)$ is almost periodic in t, we infer from the pertinent result of footnote 24 that there exists a sequence $\{t_n'(\alpha)\}$, of the type specified in the paragraph of the text containing (4.15), such that $|\Lambda(\alpha, t_n'(\alpha))| > \delta(\alpha) > 0$ for each positive integer n. Here $\delta(\alpha)$ is a constant independent of n. Now choose a positive $\omega < \omega_0$ so small that $\delta'(\alpha, \omega) \equiv \delta(\alpha) - \omega K(\alpha, \omega_0) > 0$. For such an ω, the inequalities of this footnote imply that $|\Lambda(u, t_n'(\alpha))| > \delta'(\alpha, \omega) > 0$ over the last range of n for each $u \in [\alpha, \alpha + \omega]$. Hence (4.15) obtains in the desired sense.

Proof: Let the hypotheses in the first two sentences of the theorem be fulfilled, and let $g(t)$ have the property specified in the theorem, so that $g(t) \in L_1(0, \infty)$. Then Theorem 3.1 and an argument similar to one used in the proof of Theorem 4.1 lead to the result:

$$\varphi(t) = \sum_{j=1}^{q} A_{j,0} e^{i\tau_j t} + \psi(t), \tag{4.18}$$

where $\psi(t) \in L_1(0, \infty)$.

Let us rewrite (4.5) by means of (4.9), (4.11), and Fubini's theorem. If we then substitute (4.18) in the inequality obtained in this manner and invoke Minkowski's inequality, we find:

$$W_p(t) \leq K_p + \sum_{j=1}^{q} |A_{j,0}| \, \|\eta_p(u)\lambda_j(u, t)\|_p, \tag{4.19}$$

where

$$\lambda_j(u, t) \equiv \frac{1 - e^{i(\tau_j + ku)t}}{\tau_j + ku} \tag{4.20}$$

and where K_p is a positive constant.

If (4.17) holds a.e. on $[\alpha_{ip}, \beta_{ip}]$ $(j = 1, \cdots, q)$, then one has over this range of j:

$$\begin{aligned} &\int_{-\infty}^{\infty} |\eta_p(u)\lambda_j(u, t)|^p \, du \\ &= \left[\int_{-\infty}^{\alpha_{jp}} + \int_{\alpha_{jp}}^{\beta_{jp}} + \int_{\beta_{jp}}^{\infty}\right] |\eta_p(u)\lambda_j(u, t)|^p \, du \\ &\leq \left(\frac{2}{k\xi_p}\right)^p \left[\int_{-\infty}^{\alpha_{jp}} + \int_{\beta_{jp}}^{\infty}\right] \eta_p^p(u) \, du + 2\left(\frac{2\Gamma_p}{k}\right)^p \xi_p < \infty, \end{aligned} \tag{4.21}$$

because of (4.20) and elementary inequalities, and because $\eta_p(u) \in L_p(-\infty, \infty)$.

Inequalities (4.19) and (4.21) imply the boundedness of $W_p(t)$ as $t \to \infty$.

Theorem 4.4. *Let $f_0(\mathbf{v})$ fulfill the conditions in the first two sentences of Theorem 4.3. Let $G(\mathbf{v})$ be such that $g(t) \in E(a)$ for some $a > 0$ and that $\bar{g}(i\tau_\ell) \neq 0$ for some $\ell = 1, \cdots, q$. Moreover, for this same ℓ, let the inequality*

$$\eta_p(u) \geq \eta_{\ell p} > 0 \tag{4.22}$$

hold a.e. on $[\alpha_{\ell p}, \beta_{\ell p}]$, $\eta_{\ell p}$ being a number independent of u. Then $W_p(t) \to \infty$ as $t \to \infty$.

Proof: Under the present hypotheses, (4.10) and reasoning parallel to that used to arrive at (4.19) yield:

$$W_p(t) \geq \left\| \eta_p(u) \sum_{j=1}^{q} A_{j,0}\lambda_j(u, t) \right\|_p - K_p. \tag{4.23}$$

We choose ξ_p so small that the intervals $[\alpha_{ip}, \beta_{ip}]$ $(j = 1, \cdots, q)$ are pairwise disjoint. No loss of generality is involved in this choice, since if (4.22) is satisfied a.e. on a given interval it is also obeyed a.e. on any subinterval of the interval in question.

If (4.22) obtains in the sense of the theorem, we find with the aid of Minkowski's inequality, (4.20), and elementary estimation procedures:

$$\begin{aligned} &\left\| \eta_p(u) \sum_{j=1}^{q} A_{j,0}\lambda_j(u, t) \right\|_p \\ &\geq \left\| \eta_p(u) \sum_{j=1}^{q} A_{j,0}\lambda_j(u, t) \right\|_p^{[\alpha_{\ell p}, \beta_{\ell p}]} \\ &\geq |A_{\ell,0}| \, \eta_{\ell p} \, \|\lambda_\ell(u, t)\|_p^{[\alpha_{\ell p}, \beta_{\ell p}]} \\ &\quad - 2 \sum_{\substack{j=1 \\ (j \neq \ell)}}^{q}{}' \frac{|A_{j,0}|}{||\tau_\ell - \tau_j| - k\xi_p|} \, \|\eta_p(u)\|_p^{[\alpha_{jp}, \beta_{jp}]}. \end{aligned} \tag{4.24}$$

Since $||\tau_\ell - \tau_j| - k\xi_p| > 0$ for $j \neq \ell$, by virtue of the pairwise disjointness of the intervals $[\alpha_{ip}, \beta_{ip}]$ $(j = 1, \cdots, q)$, and since $\|\eta_p(u)\|_p^{[\alpha_{jp}, \beta_{jp}]} < \infty$ under the present hypotheses, the t-independent sum over $j \neq \ell$ in (4.24) is finite. On the other hand, the use of (4.20) yields:

$$\begin{aligned} &\|\lambda_\ell(u, t)\|_p^{[\alpha_{\ell p}, \beta_{\ell p}]} \\ &\quad = \left[\frac{t^{p-1}}{k} \int_{-k\xi_p t}^{k\xi_p t} \left|\frac{e^{i\zeta} - 1}{\zeta}\right|^p d\zeta\right]^{1/p} \to \infty \end{aligned} \tag{4.25}$$

as $t \to \infty$ for any $p \geq 1$.

The hypothesis $\bar{g}(i\tau_\ell) \neq 0$ of the theorem implies that $A_{\ell,0} \neq 0$, because of the assumption that all the zeros $i\tau_j$ are simple. But if $A_{\ell,0} \neq 0$ then (4.24) and (4.25) imply that the right-hand side of (4.23), and therefore $W_p(t)$, tends to infinity as $t \to \infty$ for $1 \leq p \leq r$.

Recalling a remark in the paragraph preceeding Lemma 3.1, it is seen that the functions $G(u)$ pertaining to the initial perturbations $G(\mathbf{v})$ to which Theorems 4.1 to 4.4 apply need not be particularly "smooth," understanding smoothness as referring, for instance, to boundedness of function or derivative for $|u| < \infty$, and that $G(u)$ need not be in $L_2(-\infty, \infty)$. This fact underlies our contention that, generally speaking, stability or instability in the $W_1(t)$ or $A(t)$ sense does not imply the existence of any of the properties just stated and that, in general, stability and instability in the $W_p(t)$ $(p > 1)$ sense is independent of certain of the cited smoothness properties. As far as $A(t)$ is concerned, this last assertion contradicts the point of view espoused by Backus in Sec. VI of Ref. 3 and especially in the last paragraph of that reference.

Our remarks about smoothness, square integrability, and stability will now be restated precisely. Let $f_0(\mathbf{v})$ be such that (2.12) holds, that $h(t) \in E(b)$ for some $b > 0$, and that either one of the cases (i) or (ii) of Theorem 4.2 occurs. Then there exists an initial perturbation $G_1(\mathbf{v})$ with respect to which $f_0(\mathbf{v})$ is unstable in the sense of $A(t)$, and hence in that of $W_1(t)$, and whose corresponding function $G_1(u)$ has analyticity and square integrability properties in a horizontal strip $|u'| < d > 0$ which are of the same type as (a') and (b') in Sec. III of this paper.[27] However, and this is an essential point of our argument, there also exist initial perturbations $G_2(\mathbf{v})$ with respect to which $f_0(\mathbf{v})$ is stable in the sense of $W_1(t)$ and therefore in that of $A(t)$, but whose corresponding functions $G_2(u)$, while they are in $L_1(-\infty, \infty)$, have the "pathological" properties which we now mention. Namely, there is a $G_2(\mathbf{v})$ whose corresponding $G_2(u)$ has the first ((I)) and a $G_2(\mathbf{v})$ whose $G_2(u)$ has the second ((II)) of the following two properties: (I) $G_2(u)$ is bounded but $dG_2(u)/du$ is unbounded; (II) $G_2(u)$ is unbounded and is not in $L_2(-\infty, \infty)$. Moreover, let there exist a constant $r_0 \geq 1$ such that $f_0(\mathbf{v})$ fulfills the second inequality (4.4) when $1 \leq p \leq r_0$, outside of fulfilling the earlier requirements of this paragraph concerning $h(t)$ and the occurrence of cases (i) or (ii). Then one can find a function $G_1(\mathbf{v})$ with the stated properties and the additional one that $f_0(\mathbf{v})$ is unstable in the $W_p(t)$ sense over this range of p. One can also find a function $G_2(\mathbf{v})$ having a $G_2(u)$ with property (I) and such that any $f_0(\mathbf{v})$ of the latter type is stable with regard to this $G_2(\mathbf{v})$ in the $W_p(t)$ sense over the last range.

It is easy to discover many functions $G_1(\mathbf{v})$ obeying the above conditions. For instance, consider the function defined by

$$g_1(t) \equiv Ae^{-\alpha|t|} \tag{4.26}$$

for $|t| < \infty$, where A and α are constants such that $A \neq 0$ and $\alpha > 0$. One finds trivially that $\bar{g}_1(s) \neq 0$, in particular when $\sigma \geq 0$. This $g_1(t)$ is the Fourier transform, in the sense of (2.10), of a function $G_1(u)$ which has the properties mentioned in the previous paragraph and which is in $L_p(-\infty, \infty)$, in particular, when $1 \leq p < \infty$. These integrability properties of $G_1(u)$ imply that this function can be obtained from a $G_1(\mathbf{v})$ which satisfies the first inequality (4.4) over this range of p. Such a $G_1(\mathbf{v})$ possesses all the instability properties and other attributes of the function $G_1(\mathbf{v})$ whose existence was asserted in the last paragraph, as follows from the discussions in the present one and from Theorem 4.2.

To construct examples of functions $G_2(\mathbf{v})$ with the cited properties, we start from the fact that the condition $h(t) \in E(b)$ for some $b > 0$ entails the existence of a positive constant β such that $\Delta(s) \neq 0$ on the line $\sigma = -\beta$. We define

$$\bar{g}_2(s) \equiv B\{\prod_i (s - s_i)^{n_i}\}/(s + \beta)^N, \tag{4.27}$$

where $B \neq 0$, where the product ranges over the finite set of zeros s_i of $\Delta(s)$ in the half-plane $\sigma > -\beta$, n_i being the multiplicity of s_i ,and where N is a positive integer not smaller than $\sum_i n_i + 1$, the range of j in this sum being the same as in the said product. Hence $\bar{g}_2(s)/\Delta(s)$ is certainly regular for $\sigma \geq 0$. Furthermore, $\bar{g}_2(s)$ is the Laplace transform of a function $g_2(t) \not\equiv 0$ which, for $0 \leq t < \infty$, is of the form $e^{-\beta t} \times$ polynomial in t. For $-\infty < t < 0$ we define

$$g_2(t) \equiv g_2(|t|) + g'(t), \tag{4.28}$$

where we choose $g'(t)$ as either $C \sin \pi\gamma t/[1 - \gamma^2 t^2]$ or $D \sin \delta |t|/|t|^{\frac{1}{2}}$, the constants C, D, γ, and δ being nonvanishing, and γ and δ being real. Direct computation shows that $g_2(t)$ is the Fourier transform, in the sense last specified, of a function $G_2(u)$ which is in $L_p(-\infty, \infty)$ for $1 \leq p < \infty$ and $p = 1$, respectively, depending on whether the first or second of these choices of $g'(t)$ is made, and that such a $G_2(u)$ has the respective properties (I) and (II) for these two choices. The relative "roughness" of such functions $G_2(u)$ with respect to $G_1(u)$ is due, as the reader will easily realize, to our selections of $g_2(t)$ for negative t, many other selections leading to similar results. The stated integrability properties of the two functions $G_2(u)$ under discussion entail that each of these two functions can be obtained from a $G_2(\mathbf{v})$ which satisfies the first inequality (4.4) on the same range of p for which the said integrability property of the corresponding $G_2(u)$ holds. Hence, invoking Theorem 4.1 and collecting the relevant results of this paragraph, we conclude that such a function $G_2(\mathbf{v})$, whose $G_2(u)$ is of type (I) or (II), possesses all the stability properties and other attributes ascribed in the penultimate paragraph to the functions $G_2(\mathbf{v})$ of that paragraph having a $G_2(u)$ of the same type. Our proof of the existence of the desired functions $G_1(\mathbf{v})$ and $G_2(\mathbf{v})$ is thus complete.

It is of interest to contrast the theorems of the present section with theorems proved by Backus in

[27] Compare the properties of $G_1(\mathbf{v})$ mentioned so far in the text with analogous results stated in p. 184 and p. 186 of Ref. 3 in connection with Theorems 1 and 6 of that reference.

Sec. VI of his paper. Our Theorem 4.1 should be compared to his Theorems 2 and 4, the first paragraph of Theorem 4.2 to his Theorems 1 and 5, and Theorems 4.3 and 4.4 to his respective Theorems 6 and 7. In a careful mathematical comparison of the theorems of this section with their counterparts of Ref. 3, one should bear in mind that it is implicitly assumed in that reference that $f_0(\mathbf{v})$ is regular enough so that the condition cited immediately after Eq. (A2) of the Appendix holds, a mild requirement which, however, is irrelevant to our studies in the present section. The portion of the first paragraph of Theorem 4.2 referring to case (i), when suitably augmented by examples of the above functions $G_1(\mathbf{v})$, includes Backus' Theorem 1 as a special case if $f_0(\mathbf{v})$ is not restricted by this condition. Otherwise, this last theorem and these augmented results on case (i) are equivalent. The remaining results in Theorems 4.1 to 4.4 are neither implied nor do they imply their counterparts of Ref. 3, independently of the fulfillment of the cited condition.

The functions $G(u)$ belonging to the $G(\mathbf{v})$ considered in the theorems of this section are subjected to weaker requirements of smoothness, in the previously stated sense, than the $G(\mathbf{v})$ embraced by the greater part of the corresponding theorems of Backus. However, exception made of the above portion of Theorem 4.1, no inclusion relation exists between the classes of functions $G(u)$ obeying the hypotheses of Theorems 4.1 to 4.4 and the corresponding classes of such functions in Ref. 3. In contrast to this situation, and with the exception just noted, the functions $f_0(\mathbf{v})$ considered in Theorems 4.1 to 4.4, although very general, are less general than their counterparts in this reference when the above regularity condition on $f_0(\mathbf{v})$ holds, no inclusion relation existing otherwise. Notice, in particular, that when this condition obtains and when $h(t) \in E(b)$ for some $b > 0$, then $df_0(u)/du$ possesses analyticity and square-integrability properties parallel to (a') and (b'), such properties not being assumed in Ref. 3.

Because of the discussions of the last two paragraphs, the present stability investigation and that of Backus are best regarded as complementary.

ACKNOWLEDGMENT

We wish to thank Dr. J. N. Hayes for helpful conversations about his work on plasma oscillations.

APPENDIX

Examples of unperturbed distributions $f_0(\mathbf{v})$ satisfying the hypotheses of Theorems 4.2 to 4.4 are given in this appendix. The reader will find no difficulty in supplying the computational steps left out here for the sake of brevity.

In what follows,

$$f_0(u) \equiv \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_0(\mathbf{v})\, dv\, dw, \tag{A1}$$

and the integral of $f_0(u)$ over the entire u axis is unity because of (2.2).

1. We begin by giving examples of functions $f_0(\mathbf{v})$ satisfying the conditions of Theorem 4.2. Consider functions $f_0(\mathbf{v})$ which obey (4.4) for some $r > 1$ and are regular enough so that

$$H(u) = (i\omega_p^2/k) df_0(u)/du \tag{A2}$$

is valid a.e., that is, regular enough so that $df_0(u)/du = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\partial f_0(\mathbf{v})/\partial u) dv\, dw$, a.e. Moreover, the $f_0(\mathbf{v})$ of interest are such that $f_0(u)$ has one of the forms:

$$f_0(u) = (1/u_0)^2 \theta(|u| - u_0)\, |u|, \tag{A3}$$

$$f_0(u) = (4\epsilon^{5/2}/3\pi^{\frac{1}{2}}) u^4 e^{-\epsilon u^2}, \tag{A4}$$

where $\theta(x) \equiv 1(0)$ if $x \geq 0$ (< 0), and u_0 and ϵ are positive constants. Many such $f_0(\mathbf{v})$ can be readily constructed.

For functions $f_0(\mathbf{v})$ of this type obeying (A4), one finds that $h(t) \in E(b)$ for each $b > 0$, as in the Maxwellian case. Furthermore, the $f_0(\mathbf{v})$ of the last paragraph have the properties: (a) if (A3) holds, $\Delta(s)$ has only the two zeros

$$s = \pm k u_0 \{\exp [(k u_0/\omega_p)^2] - 1\}^{-\frac{1}{2}},$$

i.e., property (i) is present; (b) if (A4) holds and if $k = (2\epsilon\omega_p^2/3)^{\frac{1}{2}}$, the only zero of $\Delta(s)$ for $\sigma \geq 0$ is a double zero at $s = 0$, i.e., property (ii) obtains. Assertion (a) follows by direct calculation, while (b) can be proved by a combination of such a procedure with reasoning parallel to that in the second proof mentioned in Footnote 18.

Hence it is plain that the functions $f_0(\mathbf{v})$ of the class under consideration satisfy the hypotheses in both paragraphs of Theorem 4.2, provided only that $k = (2\epsilon\omega_p^2/3)^{\frac{1}{2}}$ when the corresponding functions $f_0(u)$ have the form (A4).

2. We now exhibit a function $f_0(\mathbf{v})$ fulfilling the hypotheses of Theorem 4.3. Let

$$f_0(\mathbf{v}) = f_0(u) M(v, w). \tag{A5}$$

Here, $M(v, w)$ denotes a nonnegative function which is independent of u, $M^p(v, w)$ is integrable over the entire vw plane when $1 \leq p \leq r_1$ for some constant

$r_1 \geq 1$, and the integral of $M(v, w)$ over this plane is unity; and

$$f_0(u) = \frac{u_1}{2\pi}\left\{\frac{1}{u_1^2 + (u - u_2)^2} + \frac{1}{u_1^2 + (u + u_2)^2}\right\}, \quad \text{(A6)}$$

u_1 and u_2 being positive constants. Hence this $f_0(\mathbf{v})$ obeys the second inequality (4.4) over this last range of p.

For the present example, it is found that $h(t) \in E(b)$ for each $b < ku_1$. We also conclude by direct computation in this case that the only zero of $\Delta(s)$ for $\sigma \geq 0$ is a simple zero at $s = 0$ when

$$x_1 = 2^{-\frac{1}{2}}\{[8x_2^2 + 1]^{\frac{1}{2}} - [2x_2^2 + 1]\}^{\frac{1}{2}}, \qquad 0 < x_2 < 1, \quad \text{(A7)}$$

where $x_i \equiv ku_i/\omega_p$ $(i = 1, 2)$.

With the aid of (2.5), (4.11), and (A5) we find for $|u| < \infty$ and $1 \leq p \leq r_1$:

$$\eta_p(u) = A_p\,|df_0(u)/du|, \quad \text{(A8)}$$

where A_p is a positive constant. But (A6) entails that $df_0(u)/du$ is continuous at each finite u and vanishes at $u = 0$. By (A8), this implies the existence of positive constants Γ_p and ξ_p for each p in the last range, such that (4.17) (with $\tau_j = 0$) is true for every $|u| \leq \xi_p$.

Hence the present example satisfies all the hypotheses of Theorem 4.3 if (A7) holds.

3. Finally, we exhibit an $f_0(\mathbf{v})$ obeying the conditions of Theorem 4.4. Let

$$f_0(\mathbf{v}) \equiv f_0(u)M(v, w) + r(u)N(v, w), \quad \text{(A9)}$$

where $f_0(u)$ is given by (A6); $r(u)$ is absolutely continuous over some finite interval $[-D, D]$ $(D > 0)$; there exists a positive constant r' such that $|dr(u)/du| \geq r'$ a.e. on $[-D, D]$, and $r(u) \in L_p(-\infty, \infty)$ for $1 \leq p \leq r_1$ (r_1 is the same as in Part 2 of this Appendix); $N(v, w) \not\equiv 0$ over a set of positive plane measure in the vw plane, $N^p(v, w)$ is integrable over this whole plane for p in the last range, and the integral of $N(v, w)$ over the plane in question vanishes. Furthermore, $r(u)$ and $M(v, w)$ are independent of v, w and of u, respectively, and are chosen so that $f_0(\mathbf{v}) \geq 0$. Since in the present example $f_0(u) > 0$ for $|u| < \infty$ by (A6), it is easy to see that there are many functions $r(u)$ and $N(v, w)$ with these properties. It should be clear that (A9) is consistent with (A1) and that any $f_0(\mathbf{v})$ of the form (A9) satisfies the second inequality (4.4) over the range of p just cited.

Equation (A2) is easily proved to hold a.e. for examples (A5) and (A9). From this fact, it is seen that, for given u_1 and u_2 in (A6), these two examples have the same $h(t)$ and hence the same $\Delta(s)$.

Equations (2.5), (4.11), and (A9), and Minkowski's inequality yield a.e. on $[-D, D]$ when $1 \leq p \leq r_1$:

$$\eta_p(u) \geq B_p r' - A_p\,|df_0(u)/du|, \quad \text{(A10)}$$

where B_p is a positive constant. Since $df_0(u)/du$ is continuous at each finite u and $df_0(0)/du = 0$, (A10) implies the existence of positive constants η_{1p} and $\xi_p' \leq D$, such that (4.22) (with $\eta_{tp} = \eta_{1p}$) holds a.e. on $[-\xi_p', \xi_p']$ over this range of p.

We thus deduce that the example (A9) satisfies the hypotheses of Theorem 4.4 when (A7) obtains.

Projected Angular Momentum States

N. KARAYIANIS AND C. A. MORRISON
Harry Diamond Laboratories, Washington, D. C.
(22 September 1964; final manuscript received 17 December 1964)

Known results regarding the projection of total spin states from product wavefunctions involving spin-$\frac{1}{2}$ particles are generalized. It is shown that total spin states projected from product wavefunctions involving particles with arbitrary spin, but otherwise restricted to having either maximum or minimum z component of spin, are no more complicated in form than for the spin-$\frac{1}{2}$ case.

I. INTRODUCTION

In recent articles[1-3] several investigators have reported their findings regarding the projection of a total spin state from a product wavefunction involving N spin-$\frac{1}{2}$ particles. Some of the first results along these lines were reported in an earlier article by Löwdin.[4] The use of a projection operator to obtain a total spin state is an alternative to the procedure of using Clebsch–Gordan (CG) coefficients to construct such a state, but in addition it provides a flexibility characteristic of operator techniques that is not shared by the latter method.

In this paper, the methods of group theory are shown to allow a more general, yet simpler treatment of the problem of obtaining a total spin state through the use of a projection operator. It is shown that total spin states projected from product wavefunctions involving particles with arbitrary spin, but otherwise restricted to having either maximum or minimum z component of spin, are no more complicated in form than for the spin-$\frac{1}{2}$ case. The general expression for total spin states projected from such initial product wavefunctions is obtained in Sec. II. In Sec. III the results are specialized to N spin-$\frac{1}{2}$ particles, and the agreement with the results obtained by Sasaki and Ohno,[2] and corroborated by Smith[3] is indicated.

II. THEORY

Let the symbol $[s_1 \cdots s_\mu][-s_{\mu+1} \cdots -s_{\mu+\nu}]$ represent the product wavefunction defined by

$$[s_1 \cdots s_\mu] = |s_1 s_1\rangle \cdots |s_\mu s_\mu\rangle, \tag{1}$$

$$[-s_{\mu+1} \cdots -s_{\mu+\nu}] = |s_{\mu+1} - s_{\mu+1}\rangle \cdots |s_N - s_N\rangle, \tag{2}$$

where

$$\mu + \nu = N, \tag{3}$$

and where each ket $|s\ m\rangle$ represents a normalized single-particle wavefunction, with spin s, and z-component m. Further, let $\mathbf{S}_a$ and $\mathbf{S}_b$ be the total spin operators acting on the first μ particles and the second ν particles respectively, i.e.,

$$\mathbf{S}_a = \sum_1^\mu \mathbf{s}_i \tag{4}$$

and

$$\mathbf{S}_b = \sum_{\mu+1}^N \mathbf{s}_i. \tag{5}$$

Clearly,

$$\mathbf{S}_{az}[s_1 \cdots s_\mu] = S_a[s_1 \cdots s_\mu], \tag{6}$$

and

$$\mathbf{S}_{bz}[-s_{\mu+1} \cdots -s_N] = -S_b[-s_{\mu+1} \cdots -s_N], \tag{7}$$

where

$$S_a = \sum_1^\mu s_i \tag{8}$$

and

$$S_b = \sum_{\mu+1}^N s_i. \tag{9}$$

Since S_a is the maximum obtainable z component of spin for the first μ particles (by construction), it follows that the state $[s_1 \cdots s_\mu]$ must also be an eigenstate of $\mathbf{S}_a^2$ with maximum possible angular momentum, $S_a(S_a + 1)$. A similar argument holds for the state $[-s_{\mu+1} \cdots -s_N]$ with respect to the operator $\mathbf{S}_b^2$. Therefore, it is useful to represent the two wavefunctions by

$$[s_1 \cdots s_\mu] = |S_a S_a\rangle \tag{10}$$

and

$$[-s_{\mu+1} \cdots -s_N] = |S_b - S_b\rangle, \tag{11}$$

in which case,

$$\mathbf{S}_a^2 |S_a S_a\rangle = S_a(S_a + 1) |S_a S_a\rangle \tag{12}$$

[1] J. K. Percus and A. Rotenberg, J. Math. Phys. **3**, 928 (1962).
[2] F. Sasaki and K. Ohno, J. Math. Phys. **4**, 1140 (1963).
[3] V. H. Smith, Jr., J. Chem. Phys. **41**, 277 (1964).
[4] P.-O. Löwdin, Phys. Rev. **97**, 1509 (1955).

and

$$\mathbf{S}_b^2 \, |S_b - S_b\rangle = S_b(S_b + 1) \, |S_b - S_b\rangle. \tag{13}$$

The group-theoretic operator, acting in the spin space of N particles, that projects out a state belonging to the Mth row of the Sth irreducible representation of the three-dimensional rotation group is given by[5]

$$\mathbf{O}_{SM} = \frac{2S+1}{8\pi^2} \oint d\omega D^S_{MM}(\alpha\beta\gamma)^* R, \tag{14}$$

where

$$R = e^{-i\mathbf{S}_z\alpha} e^{-i\mathbf{S}_y\beta} e^{-i\mathbf{S}_z\gamma}, \tag{15}$$

$$\mathbf{S} = \mathbf{S}_a + \mathbf{S}_b = \sum_1^N \mathbf{s}_i, \tag{16}$$

and $D^S_{MM}(\alpha\beta\gamma)$ is the M, Mth component of the Sth irreducible representation of the operator R. The integration in (14) extends over the three Euler angles specifying the rotation, $d\omega = d\alpha \, d(\cos\beta) \, d\gamma$. According to the familiar transformation properties of angular momentum states under rotation, the operation of R on the product wavefunction $|S_aS_a\rangle \, |S_b - S_b\rangle$ gives

$$R \, |S_aS_a\rangle \, |S_b - S_b\rangle = \sum_{\lambda_a\lambda_b} D^{S_a}_{\lambda_aS_a}(\alpha\beta\gamma) D^{S_b}_{\lambda_b - S_b}(\alpha\beta\gamma) \, |S_a\lambda_a\rangle \, |S_b\lambda_b\rangle. \tag{17}$$

The extraction of the Euler angle variables as arguments of the D matrices obtained in this manner enables the integration indicated in (14) to be performed, giving the result

$$\begin{aligned} |SMS_aS_b) &= \mathbf{O}_{SM} \, |S_aS_a\rangle \, |S_b - S_b\rangle \\ &= \sum_{\lambda_a\lambda_b} \langle S_a(S_a)S_b(-S_b) \, |S(M)\rangle \\ &\quad \times \langle S_a(\lambda_a)S_b(\lambda_b) \, |S(M)\rangle \, |S_a\lambda_a\rangle \, |S_b\lambda_b\rangle, \end{aligned} \tag{18}$$

where the symbol $\langle j_1(m_1)j_2(m_2)|j(m)\rangle$ is a CG coefficient. Owing to a well-known property of CG coefficients, the values of λ_a and λ_b are restricted by $\lambda_a + \lambda_b = M$. Similarly, expression (18) vanishes unless the condition $M = S_a - S_b$ is fulfilled, i.e., the initial product wavefunction is composed only of those total spin states $|S \; M \; S_aS_b)$ that admit a z component of spin M equal to the total z component of spin of the initial wavefunction, $S_a - S_b$.

Invoking the orthonormality of the states $|S_a\lambda_a\rangle$ and $|S_b\lambda_b\rangle$ and the following property of CG coefficients,

$$\sum_{m_1m_2} \langle j_1(m_1)j_2(m_2) \, |j'(m)\rangle \times \langle j_1(m_1)j_2(m_2) \, |j(m)\rangle = \delta_{jj'}, \tag{19}$$

the projected total spin states are found to have the following normalization:

$$\begin{aligned} &(S'M'S_aS_b \mid SMS_aS_b) \\ &\quad = \delta_{SS'}\delta_{MM'} \, |\langle S_a(S_a)S_b(-S_b) \mid S(M)\rangle|^2, \end{aligned} \tag{20}$$

where the value of the CG coefficient is given explicitly by,[6]

$$\begin{aligned} |\langle S_a(S_a)S_b(-S_b) \mid S(M)\rangle|^2 &= \delta(M, \, S_a - S_b)(2S+1) \\ &\times \frac{(2S_a)! \, (2S_b)!}{(S + S_a + S_b + 1)! \, (S_a + S_b - S)!}. \end{aligned} \tag{21}$$

The individual wavefunctions $|S_a\lambda_a\rangle$ and $|S_b\lambda_b\rangle$ appearing in (18) are referred to as stretched-case wavefunctions owing to the maximal nature of S_a and S_b. According to the results of the appendix to this paper, wavefunctions of this type can be expressed explicitly in terms of the single-particle wavefunctions as follows:

$$|S_a\lambda_a\rangle = \binom{2S_a}{S_a + \lambda_a}^{-\frac{1}{2}} |s_1 \cdots s_\mu; \, \lambda_a), \tag{22}$$

where

$$\begin{aligned} |s_1 \cdots s_\mu; \, \lambda_a) &= \sum_{m_1 \cdots m_\mu} \delta(\textstyle\sum m_i, \, \lambda_a) \\ &\quad \times \prod_{p=1}^{\mu} \binom{2s_p}{s_p + m_p}^{\frac{1}{2}} |s_pm_p\rangle, \end{aligned} \tag{23}$$

and similarly for $|S_b\lambda_b\rangle$. Thus, the projected total spin states $|S \; M \; S_aS_b)$ are given by the following simple expression in terms of product spin states:

$$\begin{aligned} |SMS_aS_b) &= \sum_\lambda C_\lambda(SM; \, S_aS_b) \\ &\times |s_1 \cdots s_\mu; \, S_a - \lambda) \, |s_{\mu+1} \cdots s_N; \, \lambda - S_b), \end{aligned} \tag{24}$$

where

$$\begin{aligned} &C_\lambda(SM; \, S_aS_b) \\ &\quad = \binom{2S_a}{\lambda}^{-\frac{1}{2}} \binom{2S_b}{\lambda}^{-\frac{1}{2}} \langle S_a(S_a)S_b(-S_b) \mid S(M)\rangle \\ &\quad \times \langle S_a(S_a - \lambda)S_b(\lambda - S_b) \mid S(M)\rangle. \end{aligned} \tag{25}$$

The properties of the coefficients C_λ are contained implicitly in the CG coefficients of expression (25). Consequently, many of the properties of the coefficients derived by the previous authors specifically for the spin-$\frac{1}{2}$ case are both more generally valid, and more simply obtained by exploiting the

[5] E. P. Wigner, *Group Theory* (Academic Press Inc., New York, 1959), p. 114.

[6] M. E. Rose, *Elementary Theory of Angular Momentum* (John Wiley & Sons, Inc., New York, 1961), p. 46.

known properties of the CG coefficients. For example, owing to the property of CG coefficients,

$$\sum_j \langle j_1(m_1')j_2(m_2') \mid j(m)\rangle\langle j_1(m_1)j_2(m_2) \mid j(m)\rangle = \delta_{m_1m_1'}\delta_{m_2m_2'}, \quad (26)$$

one obtains

$$\sum_S C_\lambda(SM;\, S_aS_b) = \delta_{\lambda 0}. \quad (27)$$

Similarly, from (19) the following orthogonality is obtained:

$$\sum_\lambda \binom{2S_a}{\lambda}\binom{2S_b}{\lambda} C_\lambda(S'M;\, S_aS_b)C_\lambda(SM;\, S_aS_b) = \delta_{SS'}\,|\langle S_a(S_a)S_b(-S_b) \mid S(M)\rangle|^2. \quad (28)$$

Finally, from two recursion relationships for CG coefficients given by Rose,[7] the following expressions can be obtained:

$$(2S_a - \lambda)(2S_b - \lambda)C_{\lambda+1}(SM;\, S_aS_b) + [S_a(S_a+1) + S_b(S_b+1) - S(S+1) - 2(2S_a - \lambda)(2S_b - \lambda)]C_\lambda(SM;\, S_aS_b) + \lambda^2 C_{\lambda-1}(SM;\, S_aS_b) = 0, \quad (29)$$

and

$$(2S-1)S(S_a + S_b + S + 2)[(S+1)^2 - M^2] \times C_\lambda(S+1M;\, S_aS_b) + (2S-1)(2S+3) \times \{S(S+1)(2\lambda+1) - (S_a + S_b + 1) \times [S(S+1) - M^2]\}C_\lambda(SM;\, S_aS_b) + (S+1)(2S+3)(S_a + S_b - S + 1) \times (S^2 - M^2)C_\lambda(S-1M;\, S_aS_b) = 0, \quad (30)$$

where

$$M = S_a - S_b. \quad (31)$$

III. SPIN-½ PARTICLES

The relationships between our coefficients C_λ and the coefficients obtained in previous articles[2-4] for spin-½ particles are

$$C_\lambda(S\,M;\, S_aS_b) = c_{S,\lambda} \quad \text{(Sasaki–Ohno)} \quad (32)$$

$$= C_\lambda(S,\, M,\, n) \quad \text{(Smith)}, \quad (33)$$

$$C_\lambda(S\,0;\, S_aS_b) = c_\lambda \quad \text{(Löwdin)}, \quad (34)$$

where

$$n = \tfrac{1}{2}N, \quad (35)$$

$$S_a = \tfrac{1}{2}(n + M), \quad (36)$$

and

$$S_b = \tfrac{1}{2}(n - M). \quad (37)$$

Sasaki and Ohno[2] and, later, Smith[3] have obtained an expression for $C_\lambda[SM;\, \frac{1}{2}(n+M)\ \frac{1}{2}(n-M)]$ which involves a single summation over binomial coefficients. Their expression is proportional to Wigner's explicit form[8] for the second CG coefficient in (25) with the values for S_a and S_b given by (36) and (37). The first CG coefficient in (25) is given by (21) so that even in the more general case the coefficients $C_\lambda(SM;\, S_aS_b)$ involve only one summation over binomial coefficients.

The further specialization of the results obtained in the preceding section to the spin-½ case is, of course, straightforward, however we give two more examples for clarity. For spin-½ particles the state defined by (23) reduces to

$$|\tfrac{1}{2}\cdots\tfrac{1}{2};\, \tfrac{1}{2}\mu - \lambda) = \sum_{m_1\cdots m_\mu} \delta(\textstyle\sum m_i,\, \tfrac{1}{2}\mu - \lambda)\ |\tfrac{1}{2}m_1\rangle\cdots|\tfrac{1}{2}m_\mu\rangle, \quad (38)$$

which Smith represents by the symbol

$$[\alpha^{\mu-\lambda}\beta^\lambda] = |\tfrac{1}{2}\cdots\tfrac{1}{2};\, \tfrac{1}{2}\mu - \lambda). \quad (39)$$

Finally we note that the recursion relation obtained by P.-O. Löwdin for his coefficients c_λ is a special case of (29) with $S_a = S_b = \frac{1}{2}n$.

These examples provide a useful contact with the previous papers, and, in addition, they serve to emphasize the power of the group-theoretic formalism which has been shown in this paper to give more general results at no added expense.

APPENDIX: STRETCHED-CASE WAVEFUNCTIONS

The CG coefficients are useful for constructing a total angular momentum wavefunction from single-particle angular momentum states. A state with total angular momentum j and z-component m is given by

$$|jm\rangle = \sum_{m_1\cdots m_\mu} \langle s_1(m_1)s_2(m_2) \mid j_1(m_1+m_2)\rangle \times \langle j_1(m_1+m_2)s_3(m_3) \mid j_2(m_1+m_2+m_3)\rangle\cdots \times \left\langle j_{\mu-2}\left(\sum_1^{\mu-1} m_i\right)s_\mu(m_\mu) \mid j(m)\right\rangle |s_1m_1\rangle\cdots|s_\mu m_\mu\rangle, \quad (A1)$$

[7] See Ref. 6, p. 222 and p. 224.

[8] See Ref. 6, p. 39.

where the $|s_i m_i\rangle$ are μ single-particle states with spin s_i and z-component m_i. The intermediate angular momenta j_i are free parameters within the restrictive properties of the CG coefficients. In general, no simplification of (A1) is possible. However, when

$$j = \sum_{1}^{\mu} s_i, \tag{A2}$$

the so-called stretched case, each CG coefficient is of the form

$$\langle k_1(\mu_1)k_2(\mu_2) \mid k_1 + k_2(\mu_3)\rangle = \delta_{\mu_3,\mu_1+\mu_2}\binom{2k_1 + 2k_2}{2k_1}^{-\frac{1}{2}} \times \binom{k_1 + k_2 + \mu_1 + \mu_2}{k_1 + \mu_1}^{\frac{1}{2}}\binom{k_1 + k_2 - \mu_1 - \mu_2}{k_1 - \mu_1}^{\frac{1}{2}}, \tag{A3}$$

and (A1) becomes,

$$|\sum s_i m\rangle = \binom{2\sum s_i}{\sum s_i + m}^{-\frac{1}{2}} \sum_{m_1 \cdots m_\mu} \delta(\sum m_i, m) \times \prod_{p=1}^{\mu}\binom{2s_p}{s_p + m_p}^{\frac{1}{2}} |s_p m_p\rangle. \tag{A4}$$

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 6 JUNE 1965

Conserved Quantities Associated with Symmetry Transformations of Relativistic Free-Particle Equations of Motion

D. M. FRADKIN*
Istituto di Fisica dell'Università, Rome
(Received 13 November 1964)

A general technique is presented for associating conservation laws with the symmetry transformations that leave invariant the relativistic equations of motion for a free particle. These transformations may be either continuous with the identity (such as infinitesimal transformations) or discontinuous (such as reflections). It is found that for each transformation there exist two classes of conservation laws. The number of separate laws within a class depends on the spin of the particle. The particular cases of the Dirac equation and Maxwell's equations are investigated in some detail. For the Dirac equation, conserved quantities involving discontinuous transformations and also matrix elements between particle and antiparticle states are obtained, in addition to the usual conservation laws. Application of the general method to Maxwell's equations yields not only the usual conserved quantities and Lipkin's "zilch," but also twenty new gauge-independent conserved quantities and other additional integrals associated with discontinuous transformations.

1. INTRODUCTION

IT is well known that the invariance of a physical system under a given set of symmetry operations implies a corresponding set of properties that are constants for that system. Inversely, the existence of a set of conserved quantities implies associated invariance properties of the equations providing the underlying mathematical description of the system. For a given system, the detailed connection between symmetry transformations and conserved quantities is usually exhibited[1] within the Lagrangian formalism. One shows that if the Lagrangian is invariant under a symmetry transformation continuous with the identity, then an associated quantity is conserved.

It is possible however, to develop the connection between symmetry transformations and conserved quantities directly from the equations of motion themselves. In this paper, we follow such a program for relativistic free particles. In essence, the equations of motion are manipulated in such a way that a vanishing 4-divergence is produced, $\partial_\mu D_\mu \ldots = 0$. Then, if the space components vanish at infinity fast enough, it follows that there exists a conserved quantity $\int D_4 \ldots d\mathbf{x}$.

In order to treat all relativistic particles uniformly, we consider their equations of motion in the Pauli–Fierz–Dirac form.[2] This spinor formulation allows us to establish the pertinent conservation theorems very directly. A short but sufficient algebra of spinors[3] and two-by-two matrices is included in the

* North Atlantic Treaty Organization Postdoctoral Fellow in Science, 1964–65.

[1] See, for example, E. L. Hill, Rev. Mod. Phys. **23**, 253 (1951).

[2] P. A. M. Dirac, Proc. Roy. Soc. (London) **A115**, 447 (1936); M. Fierz, Helv. Phys. Acta **12**, 3 (1939); M. Fierz and W. Pauli, Proc. Roy. Soc. (London) **A173**, 211 (1939).

[3] A general review of spinors, in a notation different from ours, is given by W. L. Bade and H. Jehle, Rev. Mod. Phys. **25**, 714 (1953). See also, E. M. Corson, *Introduction to Tensors, Spinors and Relativistic Wave-Functions* (Blackie and Son, Ltd., Glasgow, 1953); H. Umezawa, *Quantum Field Theory* (North-Holland Publishing Co., Amsterdam, 1956).

where the $|s_i m_i\rangle$ are μ single-particle states with spin s_i and z-component m_i. The intermediate angular momenta j_i are free parameters within the restrictive properties of the CG coefficients. In general, no simplification of (A1) is possible. However, when

$$j = \sum_{1}^{\mu} s_i, \tag{A2}$$

the so-called stretched case, each CG coefficient is of the form

$$\langle k_1(\mu_1)k_2(\mu_2) \mid k_1 + k_2(\mu_3)\rangle = \delta_{\mu_3,\mu_1+\mu_2}\binom{2k_1 + 2k_2}{2k_1}^{-\frac{1}{2}} \times \binom{k_1 + k_2 + \mu_1 + \mu_2}{k_1 + \mu_1}^{\frac{1}{2}}\binom{k_1 + k_2 - \mu_1 - \mu_2}{k_1 - \mu_1}^{\frac{1}{2}}, \tag{A3}$$

and (A1) becomes,

$$|\sum s_i m\rangle = \binom{2\sum s_i}{\sum s_i + m}^{-\frac{1}{2}} \sum_{m_1 \cdots m_\mu} \delta(\sum m_i, m) \times \prod_{p=1}^{\mu}\binom{2s_p}{s_p + m_p}^{\frac{1}{2}} |s_p m_p\rangle. \tag{A4}$$

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 6 JUNE 1965

Conserved Quantities Associated with Symmetry Transformations of Relativistic Free-Particle Equations of Motion

D. M. FRADKIN*
Istituto di Fisica dell'Università, Rome
(Received 13 November 1964)

A general technique is presented for associating conservation laws with the symmetry transformations that leave invariant the relativistic equations of motion for a free particle. These transformations may be either continuous with the identity (such as infinitesimal transformations) or discontinuous (such as reflections). It is found that for each transformation there exist two classes of conservation laws. The number of separate laws within a class depends on the spin of the particle. The particular cases of the Dirac equation and Maxwell's equations are investigated in some detail. For the Dirac equation, conserved quantities involving discontinuous transformations and also matrix elements between particle and antiparticle states are obtained, in addition to the usual conservation laws. Application of the general method to Maxwell's equations yields not only the usual conserved quantities and Lipkin's "zilch," but also twenty new gauge-independent conserved quantities and other additional integrals associated with discontinuous transformations.

1. INTRODUCTION

IT is well known that the invariance of a physical system under a given set of symmetry operations implies a corresponding set of properties that are constants for that system. Inversely, the existence of a set of conserved quantities implies associated invariance properties of the equations providing the underlying mathematical description of the system. For a given system, the detailed connection between symmetry transformations and conserved quantities is usually exhibited[1] within the Lagrangian formalism. One shows that if the Lagrangian is invariant under a symmetry transformation continuous with the identity, then an associated quantity is conserved.

It is possible however, to develop the connection between symmetry transformations and conserved quantities directly from the equations of motion themselves. In this paper, we follow such a program for relativistic free particles. In essence, the equations of motion are manipulated in such a way that a vanishing 4-divergence is produced, $\partial_\mu D_\mu \ldots = 0$. Then, if the space components vanish at infinity fast enough, it follows that there exists a conserved quantity $\int D_4 \ldots d\mathbf{x}$.

In order to treat all relativistic particles uniformly, we consider their equations of motion in the Pauli–Fierz–Dirac form.[2] This spinor formulation allows us to establish the pertinent conservation theorems very directly. A short but sufficient algebra of spinors[3] and two-by-two matrices is included in the

* North Atlantic Treaty Organization Postdoctoral Fellow in Science, 1964–65.

[1] See, for example, E. L. Hill, Rev. Mod. Phys. **23**, 253 (1951).

[2] P. A. M. Dirac, Proc. Roy. Soc. (London) **A115**, 447 (1936); M. Fierz, Helv. Phys. Acta **12**, 3 (1939); M. Fierz and W. Pauli, Proc. Roy. Soc. (London) **A173**, 211 (1939).

[3] A general review of spinors, in a notation different from ours, is given by W. L. Bade and H. Jehle, Rev. Mod. Phys. **25**, 714 (1953). See also, E. M. Corson, *Introduction to Tensors, Spinors and Relativistic Wave-Functions* (Blackie and Son, Ltd., Glasgow, 1953); H. Umezawa, *Quantum Field Theory* (North-Holland Publishing Co., Amsterdam, 1956).

notation section in order to make the treatment here reasonably self-complete. Of course, for a particle of specified spin and mass, the spinor formulation may be transcribed into a more familiar form, and for the cases of the electron (Dirac equation) and the photon (Maxwell's equations), we have performed such a transcription. If the reader does not care to bother with the general spinor formulation, he may go directly to the transcribed formulas.

It is shown that for each symmetry operation, there exist two classes of conserved quantities, the number of elements in a class depending on the spin of the particle under consideration. Moreover, since these classes of conserved quantities are developed directly from the equations of motion and not through a variational technique via a Lagrangian, the present approach provides conserved quantities associated with discontinuous symmetry transformations—such as space reflection.

For the Dirac equation, there is only one element in each of the two classes of conservation laws. One class provides the usual conserved field quantities, i.e., expectation values of momentum, energy, etc. The other class of conserved quantities consists of the matrix elements of these generators of infinitesimal symmetry transformations between the particle state and its charge-conjugate (antiparticle) state. In addition, conserved quantities for discontinuous symmetry operations are obtained which do not correspond to usual expectation values or matrix elements.

For the case of Maxwell's equations, we find for each class eight conserved elements associated with a single symmetry transformation. For one class, we obtain vanishing 4-divergences for two different 4-vectors, and also for a third-rank tensor, antisymmetric in the two indices left free after the divergence contraction. In the other class, we obtain vanishing 4-divergences for two different symmetric second-rank tensors. This is indeed a large number of conservation laws arising from a single symmetry transformation and naturally necessitates detailed investigation. Accordingly, each type of conservation law is examined in detail for a variety of symmetry operations. It is demonstrated how one recovers the usual conservation laws, the quantities "zilch" defined by Lipkin,[4] twenty new conserved quantities for the free electromagnetic field (apparently hitherto undetected) that might be called "angular zilch," and also conserved quantities associated with discontinuous symmetry transformations.

[4] D. M. Lipkin, J. Math. Phys. 5, 696 (1964).

2. NOTATION AND SPINOR ALGEBRA

Latin (italic type) indices with the exception of i through n, are used as spinor indices and have the range 1 to 2; Greek indices range from 1 to 4; Latin (italic type) indices i through n range from 1 to 3. Summation convention on repeated indices is employed unless specifically revoked. In any dimension, δ with two subscripts represents the usual Kronecker delta (1 if its two subscripts are the same, zero otherwise); the general case for a delta with two subscripts and two superscripts is defined in Eq. (2.3).

The equations are numbered in a conventional way except that sometimes the added designation I or II is appended in order to emphasize different classes of equations, as explained in the text. The symbols A^*, $A^\dagger$ denote the complex conjugate and Hermitian conjugate of any quantity A. Also, the 3-vector cross product is given typically by the notation $\mathbf{E} \wedge \mathbf{B}$.

For a discussion of 4-space, we circumvent the use of an explicit metric and the distinctions between covariant and contravariant indices by taking $x_4 = ict$. A contravariant–covariant distinction will only be made explicitly for spinors associated with the symbols χ and Φ, and for simplicity, will be dealt with only in terms of the mechanics of raising spinor lower indices.

Algebra of 2 × 2 Matrices and the Four-Dimensional Antisymmetric Tensor

The four matrices σ_μ are defined as

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \sigma_4 = \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}. \tag{2.1}$$

These matrices satisfy the relations

$$(\sigma_\mu^\dagger)_{ab}(\sigma_\mu)_{cd} = 2\delta_{ad}\delta_{bc}, \tag{2.2}$$

$$(\sigma_2)_{ab}(\sigma_2)_{cd} = -\delta_{cd}^{ab} = -(\delta_{ac}\delta_{bd} - \delta_{ad}\delta_{bc}), \tag{2.3}$$

$$\sigma_2\sigma_\mu = -\sigma_\mu^*\sigma_2, \tag{2.4}$$

$$\sigma_\lambda\sigma_\lambda^\dagger = \sigma_\lambda^\dagger\sigma_\lambda = 1, \qquad \text{no sum on } \lambda, \tag{2.5}$$

$$\sigma_\mu^\dagger\sigma_\nu = \delta_{\mu\nu} + \tfrac{1}{2}\epsilon_{\mu\nu\rho\pi}\sigma_\rho^\dagger\sigma_\pi, \tag{2.6}$$

$$\sigma_\mu\sigma_\nu^\dagger = \delta_{\mu\nu} - \tfrac{1}{2}\epsilon_{\mu\nu\rho\pi}\sigma_\rho\sigma_\pi^\dagger. \tag{2.7}$$

Also[5]

[5] The quantities $(\delta_{\gamma\eta}^{\alpha\beta} \pm \epsilon_{\alpha\beta\gamma\eta})$ are self and antidualing operators for the indices α and β in the sense that

$$(\delta_{\gamma\eta}^{\alpha\beta} \pm \epsilon_{\alpha\beta\gamma\eta}) = \pm\tfrac{1}{2}\epsilon_{\alpha\beta\mu\nu}(\delta_{\gamma\eta}^{\mu\nu} \pm \epsilon_{\mu\nu\gamma\eta}).$$

$$\sigma_\alpha\sigma^\dagger_\beta\sigma_\gamma = \delta_{\alpha\beta}\sigma_\gamma - (\delta^{\alpha\beta}_{\gamma\eta} - \epsilon_{\alpha\beta\gamma\eta})\sigma_\eta$$
$$= \delta_{\beta\gamma}\sigma_\alpha + (\delta^{\beta\gamma}_{\alpha\eta} + \epsilon_{\alpha\beta\gamma\eta})\sigma_\eta, \quad (2.8)$$
$$\mathrm{Tr}\,(\sigma^\dagger_\mu\sigma_\nu) = \mathrm{Tr}\,(\sigma_\nu\sigma^\dagger_\mu) = 2\delta_{\mu\nu}. \quad (2.9)$$

The symbol $\epsilon_{\mu\nu\rho\pi}$ represents that tensor which is antisymmetric to the interchange of any two indices and which has the value $\epsilon_{1234} = 1$. The contraction of two of these tensors is given by

$$\epsilon_{\mu\nu\rho\pi}\epsilon_{\mu\alpha\beta\gamma} = \delta_{\alpha\nu}\delta^{\beta\gamma}_{\rho\pi} + \delta_{\alpha\pi}\delta^{\beta\gamma}_{\nu\rho} + \delta_{\alpha\rho}\delta^{\beta\gamma}_{\pi\nu}, \quad (2.10)$$
$$\epsilon_{\mu\nu\rho\pi}\epsilon_{\mu\nu\beta\gamma} = 2\delta^{\rho\pi}_{\beta\gamma}. \quad (2.11)$$

Also, $\epsilon_{ijk} \equiv \epsilon_{ijk4}$.

Spinor Algebra

The quantities χ and Φ (with appropriate indices) represent spinors. The dotted spinor index is related to the undotted spinor index through conjugation, as in the following example:

$$\chi_{\dot a b} = (\chi_{a\dot b})^*. \quad (2.12)$$

Upper spinor indices are defined in terms of lower spinor indices by the operation

$$\chi^a = i(\sigma_2)_{ab}\chi_b \quad (2.13)$$

for each index. The spinor gradient is defined as

$$\partial_{\dot s r} \equiv (\sigma^\dagger_\mu)_{sr}\partial_\mu = \partial_{r\dot s},$$
$$\partial^{r\dot s} = \partial^{\dot s r} = -(\sigma_\mu)_{rs}\partial_\mu, \quad (2.14)$$

where $\partial_\mu = \partial/\partial x_\mu$. When the coordinates are subjected to a Lorentz transformation continuous with the identity, $x'_\mu = a_{\mu\nu}x_\nu$, each lower dotted index of a spinor transforms according to

$$\Phi'_{\dot a}(x') = \Lambda_{ab}\Phi_{\dot b}(x), \quad (2.15)$$

where

$$\Lambda_{ab} = (e^{i\mathbf{s}\cdot\mathbf{k}})_{ab} = (\cos\tfrac{1}{2}k)\delta_{ab}$$
$$+ ik^{-1}(\sin\tfrac{1}{2}k)(\mathbf{k}\cdot\mathbf{s})_{ab}, \quad (2.16)$$
$$\Lambda^\dagger\sigma_\mu\Lambda = a_{\mu\nu}\sigma_\nu. \quad (2.17)$$

From these formulas it follows that for space rotations, the 3-vector $\mathbf{k}$ is in the direction of the rotation axis with magnitude equal to the angle of rotation, and that for a pure Lorentz transformation with relative velocity $\mathbf{v}$, $\mathbf{k}$ is equal to $i(\mathbf{v}/v)$ arctanh (v/c).

3. GENERAL DERIVATION OF CONSERVATION LAWS

The relativistic equations of motion for a free particle with intrinsic spin and mass m can be studied uniformly in the Pauli–Fierz–Dirac representation. In this formalism, the equations of motion for the spinor field quantities are given by the coupled equations

$$\partial^{r\dot s}\chi_{\dot s(v)} = i\kappa\Phi^r_{(v)},$$
$$\partial_{\dot r s}\Phi^s_{(v)} = i\kappa\chi_{\dot r(v)}. \quad (3.1)$$

Here $\kappa = mc/\hbar$, and the symbol (v) denotes an arbitrary number of dotted and undotted upper and lower indices. Of course, the number of indices determines the spin of the particle under discussion. Substitution of one of the Pauli–Fierz–Dirac equations into the other gives the relativistic equation of motion for each uncoupled field quantity,

$$(-\partial_\mu\partial_\mu + \kappa^2)\Phi^r_{(v)} = (-\partial_\mu\partial_\mu + \kappa^2)\chi_{\dot r(v)} = 0. \quad (3.2)$$

By direct algebraic manipulation, the following theorem relating the symmetry properties of the equations of motion to a vanishing 4-divergence is easily established: *For transformations*

$$\chi'(x) = \chi'(\chi(x), \Phi(x)),$$
$$\Phi'(x) = \Phi'(\chi(x), \Phi(x)), \quad (3.3)$$

which leave invariant the equations of motion [i.e., primes may be put on the spinors of Eq. (3.1) while leaving their arguments unchanged], *one may construct two classes of vanishing 4-divergences,*

$$\partial_{\dot r s}[\Phi'^s_{(v_1)}\chi^{\dot r}_{(v_2)} + \chi'^{\dot r}_{(v_1)}\Phi^s_{(v_2)}] = 0, \quad (3.4\mathrm{I})$$

$$\partial_{\dot r s}[\chi'^s_{(v_1)}\chi^{\dot r}_{(v_2)} + \Phi'^{\dot r}_{(v_1)}\Phi^s_{(v_2)}] = 0. \quad (3.5\mathrm{II})$$

That these are indeed 4-divergences is seen immediately from the relation $\partial_{\dot r s} = (\sigma^\dagger_\mu)_{rs}\partial_\mu$.

For massless particles, the equations of motion are obtained from Eqs. (3.1) by replacing Φ by $\kappa\Phi$, then letting $\kappa^2 \to 0$. In this formulation, Φ represents a potential for the field quantities χ. The massless equations of motion become

$$\partial^{r\dot s}\chi_{\dot s(v)} = 0,$$
$$\partial_{\dot r s}\Phi^s_{(v)} = i\chi_{\dot r(v)}. \quad (3.6)$$

As before, the following theorem connecting symmetry properties and vanishing 4-divergences is established: *For transformations given by Eq. (3.3) which leave invariant the equations of motion* [Eqs. (3.6)], *one may construct two classes of vanishing 4-divergences,*

$$\partial_{\dot r s}[\Phi'^s_{(v_1)}\chi^{\dot r}_{(v_2)} + \chi'^{\dot r}_{(v_1)}\Phi^s_{(v_2)}] = 0, \quad (3.7\mathrm{I})$$

$$\partial_{\dot r s}[\chi'^s_{(v_1)}\chi^{\dot r}_{(v_2)}] = 0. \quad (3.8\mathrm{II})$$

The transformations satisfying the premises of the preceding theorems are called symmetry transforma-

tions. If one assumes that the spinor components vanish sufficiently rapidly at spatial infinity, Eqs. (3.4), (3.5) and Eqs. (3.7), (3.8) yield conservation laws for massive and massless particles, respectively, that are associated with a particular symmetry transformation. It is seen that there are two classes of conservation laws (here designated by the equation labeling scheme of I and II) for each symmetry transformation and that the number of laws within a class is given by the square of the number of free indices designated by (v). Also, since the preceding derivation of the conservation laws does not depend on the commutation properties of the spinor components, these laws are still valid when the theories are second-quantized, as long as the symmetry operations are further restricted to leave the commutation assignments unchanged.

We shall now examine these conservation laws for the two cases of the electron (Dirac equation) and the photon (Maxwell's equations). For these cases, the spinor equations will be transcribed into equations containing more familiar quantities.

4. DIRAC EQUATION

The spinor representation of the free-particle Dirac equation is given by the coupled equations

$$\begin{aligned} \partial^{\dot{r}i}\chi_{\dot{i}} &= i\kappa\Phi^{\dot{r}}, \\ \partial_{\dot{r}s}\Phi^{\dot{s}} &= i\kappa\chi_{\dot{r}}. \end{aligned} \tag{4.1}$$

The correspondence between this representation and the usual one in terms of wavefunctions and four-by-four matrices is given by the relations

$$\psi = \begin{bmatrix} \chi_{\dot{1}} \\ \chi_{\dot{2}} \\ \Phi^{\dot{1}} \\ \Phi^{\dot{2}} \end{bmatrix},$$

$$\psi^{\mathrm{ch}} = \gamma_2\psi^* = (C\psi)^*,$$

$$\begin{aligned} C^{-1} &= C^* = C^\dagger, \\ \alpha &= \begin{pmatrix} \mathfrak{d} & 0 \\ 0 & -\mathfrak{d} \end{pmatrix}, \qquad \beta = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ \gamma_5 &= \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \sigma_\mu = \begin{pmatrix} \sigma_\mu & 0 \\ 0 & \sigma_\mu \end{pmatrix}, \\ \boldsymbol{\gamma} &= -i\beta\boldsymbol{\alpha}, \qquad \gamma_4 = \beta, \\ \gamma_5 &= \gamma_1\gamma_2\gamma_3\gamma_4, \qquad C^*\gamma_\mu^*C = -\gamma_4\gamma_\mu\gamma_4. \end{aligned} \tag{4.2}$$

Here, C is the charge-conjugate matrix and ψ^{ch} is the charge-conjugate wavefunction. A convenient identity that exists in the spinor representation is

$$[(\sigma_\mu - \sigma_\mu^\dagger) - \gamma_5(\sigma_\mu + \sigma_\mu^\dagger)] = 2i\gamma_4\gamma_\mu. \tag{4.3}$$

With the aid of the relations given in Eqs. (4.2) and (4.3), the spinor equations may be transcribed into more familiar ones. The equation of motion, given by Eq. (4.1) becomes

$$(\gamma_\mu\partial_\mu + \kappa)\psi = 0 \tag{4.4}$$

and the two classes of vanishing 4-divergences given by Eqs. (3.4I), (3.5II) become

$$\partial_\mu[(\psi'^{\mathrm{ch}})^\dagger\gamma_4\gamma_\mu\psi] = 0, \tag{4.5I}$$

$$\partial_\mu[\psi'^\dagger\gamma_4\gamma_\mu\psi] = 0. \tag{4.6II}$$

There is only one equation for each class since the Dirac equation involves spinors with only one index. Also, one can now forget the derivation of Eqs. (4.5I), (4.6II) from the Pauli–Fierz–Dirac viewpoint and consider directly symmetry transformations of the wavefunction, $\psi'(x) = \psi'(\psi(x))$ which leave the Dirac equation [Eq. (4.4)] invariant.

Transformations which may be applied in Eqs. (4.5I), (4.6II) are given in Table I.

It is seen that for class II, given by Eq. (4.6II), we obtain the usual expectation values of the conserved quantities: probability density (or number of particles in the second quantized case), momen-

TABLE I. Wavefunction transformations corresponding to Dirac equation symmetry operations. Both λ_α and $\xi_{\alpha\beta} = -\xi_{\beta\alpha}$ denote infinitesimals.

Wavefunction transformation $\psi'(x) =$	Dirac equation symmetry operation
$\psi(x)$	identity, $x'_\mu = x_\mu$
$(1 + \lambda_\alpha\partial_\alpha)\psi(x)$	infinitesimal displacements, $x'_\mu = x_\mu + \lambda_\mu$
$\{1 + \frac{1}{2}\xi_{\alpha\beta}[(x_\alpha\partial_\beta - x_\beta\partial_\alpha) + \frac{1}{4}(\gamma_\alpha\gamma_\beta - \gamma_\beta\gamma_\alpha)]\}\psi(x)$	infinitesimal Lorentz transf., $x'_\mu = (\delta_{\mu\nu} + \xi_{\mu\nu})x_\nu$
$i\gamma_4\psi(-x^*)$	usual space reflection $x'_\mu = -x^*_\mu$
$i\gamma_4C^*\psi^*(-x^*)$	Wigner–Landau combined inversion, $x'_\mu = -x^*_\mu$
$\gamma_4\gamma_5C^*\psi^*(x^*)$	time inversion, $x'_\mu = x^*_\mu$
$(1 - i\lambda_\mu\epsilon_{\mu\nu\rho\pi}\gamma_\nu\gamma_\rho\partial_\pi)\psi(x)$	little group transf.[a]

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

tum–energy, angular momentum–center of energy, and 4-vector polarization[6] from the transformations generated by the identity. infinitesimal displacements, infinitesimal Lorentz transformations, and little group transformations, respectively. Also, we obtain conserved quantities from the discontinuous transformations—space reflection, Wigner–Landau combined inversion, and time inversion—which are not directly interpretable as expectation values.

The other class of conservation laws given by Eq. (4.5I) yields conserved matrix elements of the generators of the continuous transformations between the particle state and the antiparticle (charge-conjugate) state. A set of conserved quantities corresponding to the discontinuous transformations is also obtained.

5. MAXWELL'S EQUATIONS: TRANSCRIBED RELATIONS

The spinor representation of Maxwell's equations is given by the coupled equations

$$\partial^{r\dot{s}}\chi_{\dot{s}}^{\dot{u}} = 0, \tag{5.1}$$

$$\partial_{\dot{r}s}\Phi^{s\dot{u}} = i\chi_{\dot{r}}^{\dot{u}}, \tag{5.2}$$

in which the spinor field quantities possess the symmetry

$$\chi_{\dot{s}\dot{u}} = \chi_{\dot{u}\dot{s}} \quad (\text{so } \chi_{\dot{1}}^{\dot{1}} = -\chi_{\dot{2}}^{\dot{2}}) \quad \text{and} \quad \Phi^{s\dot{u}} = \Phi^{\dot{u}s}.$$

The correspondence between this representation and the usual one in terms of the antisymmetric electromagnetic field tensor $F_{\mu\nu}$ and the 4-potential A_μ is given by the relations

$$\chi_{\dot{s}}^{\dot{u}} = \tfrac{1}{2}i(\sigma_\mu^\dagger\sigma_\nu)_{\dot{s}u}F_{\mu\nu} = \tfrac{1}{2}i(\sigma_\mu\sigma_\nu^\dagger)_{\dot{s}u}F_{\mu\nu}^*, \tag{5.3}$$

$$\Phi^{s\dot{u}} = -(\sigma_\mu)_{s\dot{u}}A_\mu.$$

One could also write $\chi_{\dot{s}}^{\dot{u}}$ in terms of the space rotation 3-vector $\psi = \mathbf{E} + i\mathbf{B}$, by using the relation $\sigma_\mu^\dagger\sigma_\nu F_{\mu\nu} = 2\sigma_k\psi_k$. However, we shall not use the ψ formalism here since this does not lead to manifestly covariant four-dimensional forms. Here, the electric and magnetic fields, $\mathbf{E}$ and $\mathbf{B}$, are given by

$$F_{\mu\nu} = -F_{\nu\mu}, \quad E_k = iF_{k4}, \quad B_i = \tfrac{1}{2}\epsilon_{ijk}F_{jk}. \tag{5.4}$$

It is also convenient to introduce the antisymmetric electromagnetic dual field tensor $F_{\mu\nu}^{\mathrm{D}}$ defined by

$$F_{\mu\nu}^{\mathrm{D}} = \tfrac{1}{2}i\epsilon_{\mu\nu\rho\tau}F_{\rho\tau}. \tag{5.5}$$

[6] D. M. Fradkin and R. H. Good, Jr., Rev. Mod. Phys. **33**, 343 (1961).

The dual tensor $F_{\mu\nu}^{\mathrm{D}}$ has the same reality properties as $F_{\mu\nu}$ and corresponds to replacing $\mathbf{B} \to \mathbf{E}$, $\mathbf{E} \to -\mathbf{B}$ in the latter tensor.

The spinor equations of motion may be transcribed into more familiar ones by the following procedure. Upon multiplying Eq. (5.1) by $\sigma_\pi^\dagger$ and taking the trace with the aid of the relations given by Eqs. (2.8) and (2.9), one obtains

$$i(\partial_\mu F_{\pi\mu}) + (\partial_\mu F_{\pi\mu}^{\mathrm{D}}) = 0. \tag{5.6}$$

Also, upon multiplying Eq. (5.2) by $\sigma_\pi^\dagger\sigma_\rho$ and subsequently taking the trace, one obtains

$$\{\delta_{\pi\rho}(\partial_\mu A_\mu) + [(\partial_\rho A_\pi - \partial_\pi A_\rho) - F_{\rho\pi}] + \tfrac{1}{2}\epsilon_{\pi\rho\mu\lambda}[(\partial_\lambda A_\mu - \partial_\mu A_\lambda) - F_{\lambda\mu}]\} = 0. \tag{5.7}$$

From a consideration of the reality properties and the independence of the symmetric and antisymmetric parts, one recovers from the preceding two equations the usual relations:

$$\partial_\mu F_{\pi\mu} = 0, \tag{5.8}$$

$$\partial_\mu F_{\pi\mu}^{\mathrm{D}} = 0, \tag{5.9}$$

$$F_{\rho\pi} = (\partial_\rho A_\pi - \partial_\pi A_\rho), \tag{5.10}$$

$$\partial_\mu A_\mu = 0. \tag{5.11}$$

These are the familiar Maxwell's equations in the Lorentz gauge [Eq. (5.11)]. The field equations (5.8) and (5.9) may be written in the alternative forms

$$i\epsilon_{\mu\nu\rho\pi}(\partial_\lambda F_{\lambda\pi}) = 0 = (\partial_\mu F_{\nu\rho}^{\mathrm{D}} + \partial_\rho F_{\mu\nu}^{\mathrm{D}} + \partial_\nu F_{\rho\mu}^{\mathrm{D}}), \tag{5.12}$$

$$-i\epsilon_{\mu\nu\rho\pi}(\partial_\lambda F_{\lambda\pi}^{\mathrm{D}}) = 0 = (\partial_\mu F_{\nu\rho} + \partial_\rho F_{\mu\nu} + \partial_\nu F_{\rho\mu}). \tag{5.13}$$

Now that the spinor equations of motion have been transcribed into the familiar ones, it remains to transcribe the equations [Eqs. (3.7I) and (3.8II)] for the two classes of vanishing 4-divergences for each symmetry operation. For the case of Maxwell's equations, it is easily shown that these equations reduce to

$$\partial_\mu[(\sigma_\rho^\dagger\sigma_\pi\sigma_\mu^\dagger\sigma_\lambda)_{ab}A_\lambda' F_{\rho\pi} - (\sigma_\lambda^\dagger\sigma_\mu\sigma_\pi^\dagger\sigma_\rho)_{ab}F_{\rho\pi}'A_\lambda] = 0, \tag{5.14I}$$

$$\partial_\mu[(\sigma_\alpha^\dagger\sigma_\beta\sigma_\mu^\dagger\sigma_\pi\sigma_\rho^\dagger)_{ab}F_{\rho\pi}'F_{\alpha\beta}] = 0. \tag{5.15II}$$

Here a and b are free indices. In order to write these equations in 4-dimensional form, a trace technique similar to that used to transcribe the equations of motion is employed. Multiplying Eq. (5.14I) by

$\sigma_\alpha^\dagger\sigma_\beta$ and taking the trace with the aid of Eqs. (2.8), its conjugate, and (2.9), one obtains

$$\{\delta_{\alpha\beta}[\partial_\mu(A'_\lambda F_{\lambda\mu} - F'_{\lambda\mu}A_\lambda)] - i\delta_{\alpha\beta}[\partial_\mu(A'_\lambda F^{\mathrm{D}}_{\lambda\mu} - F'^{\mathrm{D}}_{\lambda\mu}A_\lambda)] + \partial_\mu[G_{\alpha\beta,\mu}(A', F', A, F)] + \partial_\mu[\tfrac{1}{2}\epsilon_{\alpha\beta\rho\pi}G_{\rho\pi,\mu}(A', F', A, F)]\} = 0, \quad (5.16\mathrm{I})$$

where

$$G_{\alpha\beta,\mu}(A', F', A, F) = [\delta^{\alpha\beta}_{\lambda\eta}(A'_\lambda F_{\eta\mu} + F'_{\eta\mu}A_\lambda) + i\epsilon_{\alpha\beta\lambda\eta}(A'_\lambda F^{\mathrm{D}}_{\mu\eta} + F'^{\mathrm{D}}_{\mu\eta}A_\lambda)]. \quad (5.17)$$

We note that the tensor G is antisymmetric in the indices α and β, and that an element of it is real or imaginary depending on whether the index 4 occurs an even or odd number of times in α, β, and μ. Consequently, if $G_{\alpha\beta,\mu}$ is real (imaginary) it follows that $\epsilon_{\alpha\beta\rho\pi}G_{\rho\pi,\mu}$ is imaginary (real). Moreover, since the symmetric and antisymmetric (with respect to the indices α and β) parts of Eqs. (5.16I) are completely independent, it follows that

$$\partial_\mu(A'_\lambda F_{\lambda\mu} - F'_{\lambda\mu}A_\lambda) = 0, \quad \partial_\mu(A'_\lambda F^{\mathrm{D}}_{\lambda\mu} - F'^{\mathrm{D}}_{\lambda\mu}A_\lambda) = 0, \quad \partial_\mu G_{\alpha\beta,\mu}(A', F', A, F) = 0. \quad (5.18\mathrm{I})$$

The remaining equation for class II [Eq. (5.15II)] is reduced to four-dimensional form by multiplying it by σ_ν and taking the trace with the aid of Eqs. (2.8), its conjugate, and (2.9). The symmetry with respect to μ and ν can be explicitly exhibited by noting that since Tr $[(\sigma_\alpha^\dagger\sigma_\beta\sigma_\mu^\dagger)(\sigma_\pi\sigma_\rho^\dagger\sigma_\nu)]$ equals Tr $[(\sigma_\nu\sigma_\alpha^\dagger\sigma_\beta)(\sigma_\mu^\dagger\sigma_\pi\sigma_\rho^\dagger)]$, the trace reductions can be made with the average of these two expressions. The result is

$$[\partial_\mu W_{\mu\nu}(F', F)] + i[\partial_\mu Z_{\mu\nu}(F', F)] = 0, \quad (5.19\mathrm{II})$$

where

$$W_{\mu\nu}(F', F) \equiv \tfrac{1}{2}[(F'_{\mu\rho}F_{\rho\nu} + F'_{\nu\rho}F_{\rho\mu}) + (F'^{\mathrm{D}}_{\mu\rho}F^{\mathrm{D}}_{\rho\nu} + F'^{\mathrm{D}}_{\nu\rho}F^{\mathrm{D}}_{\rho\mu})] = [(F'_{\mu\rho}F_{\rho\nu} + F'_{\nu\rho}F_{\rho\mu}) - \tfrac{1}{2}\delta_{\mu\nu}F'_{\rho\pi}F_{\pi\rho}], \quad (5.20)$$

$$Z_{\mu\nu}(F', F) \equiv \tfrac{1}{2}[(F'^{\mathrm{D}}_{\mu\rho}F_{\rho\nu} + F'^{\mathrm{D}}_{\nu\rho}F_{\rho\mu}) - (F'_{\mu\rho}F^{\mathrm{D}}_{\rho\nu} + F'_{\nu\rho}F^{\mathrm{D}}_{\rho\mu})]. \quad (5.21)$$

Again, from investigation of reality properties, it follows that the two bracketed ([ ]) expressions in Eq. (5.19II) are independently zero. Consequently,

$$\partial_\mu W_{\mu\nu}(F', F) = 0, \quad \partial_\mu Z_{\mu\nu}(F', F) = 0. \quad (5.22\mathrm{II})$$

It is apparent that the second-rank tensors $W_{\mu\nu}$ and $Z_{\mu\nu}$ are both symmetric in their indices. We also find that their tensorial traces are both zero.

Now that the spinor equations have been completely transcribed, one may ignore the spinor formulation and consider directly the symmetry transformations of the field quantities $F'_{\mu\nu}(x) = F'_{\mu\nu}(F(x), A(x))$, $A'_\mu(x) = A'_\mu(F(x), A(x))$ which leave invariant the equations of motion given by Eqs. (5.8) to (5.11). Each symmetry transformation gives rise to the vanishing 4-divergences displayed in Eqs. (5.18I) and (5.22II).

In order to see which of the equations of motion are critical for the establishment of the vanishing 4-divergences, we have actually performed the differentiation, using only the antisymmetry of $F_{\mu\nu}$ and the definition of $F^{\mathrm{D}}_{\mu\nu}$ in terms of $F_{\mu\nu}$, but not employing the equations of motion. The result is

$$\partial_\mu(A'_\lambda F_{\lambda\mu} - F'_{\lambda\mu}A_\lambda) = \{A'_\lambda[\partial_\mu F_{\lambda\mu}] - A_\lambda[\partial_\mu F'_{\lambda\mu}] + \tfrac{1}{2}[(\partial_\mu A'_\lambda - \partial_\lambda A'_\mu)F_{\lambda\mu} - F'_{\mu\lambda}(\partial_\lambda A_\mu - \partial_\mu A_\lambda)]\}, \quad (5.23\mathrm{I})$$

$$\partial_\mu(A'_\lambda F^{\mathrm{D}}_{\lambda\mu} - F'_{\lambda\mu}A_\lambda) = \{A'_\lambda[\partial_\mu F^{\mathrm{D}}_{\lambda\mu}] - A_\lambda[\partial_\mu F'^{\mathrm{D}}_{\lambda\mu}] + (i/4)\epsilon_{\lambda\mu\rho\pi}[(\partial_\mu A'_\lambda - \partial_\lambda A'_\mu)F_{\rho\pi} - F'_{\mu\lambda}(\partial_\rho A_\pi - \partial_\pi A_\rho)]\}, \quad (5.24\mathrm{I})$$

$$\partial_\mu G_{\alpha\beta,\mu}(A', F', A, F) = (\delta^{\alpha\beta}_{\gamma\eta}\{A'_\lambda[\partial_\mu F_{\eta\mu}] + A_\lambda[\partial_\mu F'_{\eta\mu}]\} + i\epsilon_{\alpha\beta\lambda\eta}\{A'_\lambda[\partial_\mu F'^{\mathrm{D}}_{\mu\eta}] + A_\lambda[\partial_\mu F'^{D}_{\mu\eta}]\} + \delta^{\alpha\beta}_{\lambda\eta}[(\partial_\mu A'_\lambda - \partial_\lambda A'_\mu)F_{\eta\mu} - F'_{\mu\lambda}(\partial_\eta A_\mu - \partial_\mu A_\eta)] - \{F_{\alpha\beta}[\partial_\mu A'_\mu] + F'_{\alpha\beta}[\partial_\mu A_\mu]\}), \quad (5.25\mathrm{I})$$

$$\partial_\mu W_{\mu\nu}(F', F) = \{F_{\rho\nu}[\partial_\mu F'_{\mu\rho}] + F'_{\rho\nu}[\partial_\mu F_{\mu\rho}] + F^{\mathrm{D}}_{\rho\nu}[\partial_\mu F'^{\mathrm{D}}_{\mu\rho}] + F'^{\mathrm{D}}_{\rho\nu}[\partial_\mu F^{\mathrm{D}}_{\mu\rho}]\}, \quad (5.26\mathrm{II})$$

$$\partial_\mu Z_{\mu\nu}(F', F) = \{F_{\nu\rho}[\partial_\mu F'^{\mathrm{D}}_{\rho\mu}] - F'_{\nu\rho}[\partial_\mu F^{\mathrm{D}}_{\rho\mu}] + F'^{\mathrm{D}}_{\nu\rho}[\partial_\mu F_{\rho\mu}] - F^{\mathrm{D}}_{\nu\rho}[\partial_\mu F'_{\rho\mu}]\}. \quad (5.27\mathrm{II})$$

The right-hand sides of the preceding five equations are all zero (as we already knew) when the equations of motion are employed since then each term in square brackets [ ] is zero. However, $G_{\alpha\beta,\mu}$ is the only quantity that requires the Lorentz condition[7] $\partial_\lambda A_\lambda = 0$ in order that its 4-divergence be zero. The other quantities—the two 4-vectors and the two symmetric second-rank tensors—require only two

[7] Hence, we use the symbol G to imply a gauge-dependent quantity.

TABLE II. Electromagnetic field transformations corresponding to symmetry operations for Maxwell's equations. Both λ_α and $\xi_{\alpha\beta} = -\xi_{\beta\alpha}$ denote infinitesimals.

Field tensor transf., $F'_{\mu\nu}(x)$	Potential transf., $A'_\mu(x)$ =	Maxwell's Eqs. symm. operation
$F_{\mu\nu}(x)$	$A_\mu(x) + \partial_\mu g(x)$	gauge transf.
$F_{\mu\nu}(x)$	$A_\mu(x)$	identity, $x'_\mu = x_\mu$
$(1 + \lambda_\alpha\partial_\alpha)F_{\mu\nu}(x)$	$(1 + \lambda_\alpha\partial_\alpha)A_\mu(x)$	inf. displ., $x'_\mu = x_\mu + \lambda_\mu$
$\{\delta_{\mu\alpha}\delta_{\nu\beta} + \frac{1}{2}\xi_{\lambda\rho}[\delta_{\nu\beta}\delta^{\lambda\rho}_{\mu\alpha} + \delta_{\mu\alpha}\delta^{\lambda\rho}_{\nu\beta} + \delta_{\mu\alpha}\delta_{\beta\nu}(x_\lambda\partial_\rho - x_\rho\partial_\lambda)]\}$ operating on $F_{\alpha\beta}(x)$	$\{\delta_{\mu\nu} + \frac{1}{2}\xi_{\alpha\beta}[\delta^{\alpha\beta}_{\mu\nu} + \delta_{\mu\nu}(x_\alpha\partial_\beta - x_\beta\partial_\alpha)]\}$ operating on $A_\nu(x)$	inf. Lorentz transf., $x'_\mu = (\delta_{\mu\nu} + \xi_{\mu\nu})x_\nu$
$-F^*_{\mu\nu}(-x^*)$	$A^*_\mu(-x^*)$	space refl.[a] $x'_\mu = -x^*_\mu$
$F^*_{\mu\nu}(x^*)$	$A^*_\mu(x^*)$	time invers.[a] $x'_\mu = x^*_\mu$

[a] Under the usual assumption that charge is a scalar by these transformations.

of the remaining equations of motion [Eqs. (5.8) to (5.10)] in order that their 4-divergences vanish.

Symmetry transformations which meet the conditions of our theorems and which may be applied in Eqs. (5.18I) and (5.22II) are given in Table II. A detailed discussion of the associated conserved quantities is given in the following sections.

6. MAXWELL'S EQUATIONS: CLASS I CONSERVED QUANTITIES

To facilitate discussion, we shall define the 4-vectors

$$V_\mu(A', F', A, F) \equiv A'_\lambda F_{\lambda\mu} - F'_{\lambda\mu}A_\lambda, \tag{6.1}$$

$$\bar{V}_\mu(A', F', A, F) \equiv A'_\lambda F^{\mathrm{D}}_{\lambda\mu} - F'^{\mathrm{D}}_{\lambda\mu}A_\lambda. \tag{6.2}$$

Consequently, under adequate assumptions concerning asymptotic behavior, the conserved (time-independent) quantities derivable from the class I equations [Eqs. (5.18I)] may be simply written

$$\int V_4(A', F', A, F)\, d\mathbf{x}, \quad \int \bar{V}_4(A', F', A, F)\, d\mathbf{x},$$

$$\int G_{\alpha\beta,4}(A', F', A, F)\, d\mathbf{x}. \tag{6.3}$$

We shall refer to the integrands of the preceding expressions as densities which correspond to a given symmetry transformation. Also, in discussion of a particular symmetry transformation, we shall suppress the detailed description of the argument and simply write $G_{\alpha\beta,4}(\lambda_\gamma)$, for example, to indicate the resulting expression obtained from the symmetry transformation associated with an arbitrary infinitesimal displacement in the γth direction. Moreover, in the case of an infinitesimal transformation, this designation will apply only to that part of the transformation which is proportional to the infinitesimal (subtracting out the identity transformation).

Conserved Quantities from the 4-Vector $V_\mu(A', F', A, F)$

From the gauge transformation, we obtain the associated density

$$V_4(\text{gauge}) = (\partial_\lambda g)F_{\lambda 4} = \partial_k(gF_{k4}).$$

Consequently, its integral over all space is zero (because of the asymptotic behavior assumed for the relevant quantities) and no nontrivial conserved expression is obtained. Also, the identity transformation gives a density that is identically zero.

The infinitesimal displacements λ_α provide the usual conservation laws of momentum and energy. Specifically, using Maxwell's equations, it is found that

$$V_4(\lambda_i) = -2i(\mathbf{E} \wedge \mathbf{B})_i + \partial_k[2A_iF_{k4} - \delta_{ki}F_{\lambda 4}A_\lambda],$$
$$V_4(\lambda_4) = (E^2 + B^2) + \partial_k[2A_4F_{k4} + F_{\lambda k}A_\lambda]. \tag{6.4}$$

Again, by our hypothesis concerning asymptotic behavior, the integrals over the 3-divergences vanish and we are left with conserved quantities proportional to the usual integrals giving the momentum and energy of the electromagnetic field.

The infinitesimal Lorentz transformations (ξ_{ab} premultipliers) provide the usual conservation laws of angular momentum and center of energy. Specifically, the densities are

$$V_4(\xi_{ij}) = \{-2i[x_i(\mathbf{E} \wedge \mathbf{B})_j - x_j(\mathbf{E} \wedge \mathbf{B})_i]$$
$$+ \partial_k[2(x_iA_j - x_jA_i)F_{k4} + (\delta_{ik}x_j - \delta_{jk}x_i)F_{\lambda 4}A_\lambda]\},$$
$$V_4(\xi_{4j}) = \{[-2ix_4(\mathbf{E} \wedge \mathbf{B})_j - x_j(E^2 + B^2)] \tag{6.5}$$
$$+ \partial_k[2(x_4A_j - x_jA_4)F_{k4} - (\delta_{jk}x_4F_{\lambda 4} + x_jF_{\lambda k})A_\lambda]\}.$$

Once more, the 3-space divergences vanish when integrated over all space and we are left with the usual conservation laws of angular momentum–center of energy. Here, we note that the usual connection between infinitesimal symmetry generators and conservation laws arises from a class I conservation law, in contrast to the Dirac case for which the usual connection came from a class II conservation law,

For the space reflection transformation, use of the reality properties of the field quantities gives the associated conserved quantity

$$\int [\mathbf{A}(-\mathbf{x}, t)\cdot\mathbf{E}(\mathbf{x}, t) - \mathbf{E}(-\mathbf{x}, t)\cdot\mathbf{A}(\mathbf{x}, t)]\, d\mathbf{x}. \quad (6.6)$$

However, $\mathbf{x}$ may be changed to $-\mathbf{x}$ in the last term of this expression, so the above integral becomes identically zero and no nontrivial conserved quantity is obtained. For time reflection, the associated conserved quantity may be written

$$\int [\mathbf{A}(\mathbf{x}, -t)\cdot\mathbf{E}(\mathbf{x}, t) + \mathbf{E}(\mathbf{x}, -t)\cdot\mathbf{A}(\mathbf{x}, t)]\, d\mathbf{x}, \quad (6.7)$$

which is a nontrivial integral of the motion.

Conserved Quantities from the 4-Vector $\bar{V}_\mu(A', F', A, F)$

For the gauge transformation, the identity transformation, the infinitesimal displacements, and the infinitesimal Lorentz transformations—for all these transformations continuous with the identity—we may use Maxwell's equations to show that the associated densities $\bar{V}_4$ are either zero or equal to a 3-space divergence. Consequently, the integrals of these densities over all space yield zero and no new (nontrivial) conserved quantities are obtained.

The space-reflection transformation provides the conserved quantity

$$2\int \mathbf{A}(-\mathbf{x}, t)\cdot\mathbf{B}(\mathbf{x}, t)\, d\mathbf{x},$$

and the time-reflection transformation provides the conserved quantity

$$\int [\mathbf{A}(\mathbf{x}, -t)\cdot\mathbf{B}(\mathbf{x}, t) - \mathbf{B}(\mathbf{x}, -t)\cdot\mathbf{A}(\mathbf{x}, t)]\, d\mathbf{x}.$$

However, using the fact that $\mathbf{B} = \text{curl } \mathbf{A}$, and that surface integrals may be ignored, it is easily shown that the space-reflection conserved quantity has the value zero. Also, the time-reflection integral may be evaluated at $t = 0$, so we find that this conserved quantity is also merely zero. Hence, $\bar{V}_4$ does not provide any nontrivial conserved quantities.

Conserved Quantities from the Tensor $G_{\alpha\beta,\mu}(A', F', A, F)$

As noted before, the tensor $G_{\alpha\beta,\mu}$ (defined by Eq. (5.17) is antisymmetric in the indices α and β, and the tensor has a zero 4-divergence $\partial_\mu G_{\alpha\beta,\mu}$ only if the 4-potential satisfies the Lorentz condition $\partial_\lambda A_\lambda = 0$. In the sense that the Lorentz condition is not a necessary or essential feature of electromagnetism, also the conservation laws arising from the densities $G_{\alpha\beta,4}$ are not expected to be a part of something actually conserved in nature, but only reflect an arbitrary supplementary mathematical condition.

Using Maxwell's equations and ignoring integrals over space divergences, we find that the gauge transformation provides the associated conserved quantity

$$\int G_{\alpha\beta,4}(\text{gauge})\, d\mathbf{x} = \int [g(\partial_4 F_{\alpha\beta}) - (\partial_4 g)F_{\alpha\beta}]\, d\mathbf{x}. \quad (6.8)$$

That this is indeed a conserved quantity could be seen immediately from the fact that if two quantities satisfy the d'Alembertian equation $\partial_\mu\partial_\mu g = \partial_\mu\partial_\mu F_{\alpha\beta} = 0$, then it follows that $\partial_\mu[(\partial_\mu g)F_{\alpha\beta} - g(\partial_\mu F_{\alpha\beta})] = 0$. Here, g satisfies the d'Alembertian equation because of the Lorentz gauge condition. We note that this conservation argument based on solutions of the d'Alembertian equation applies equally well if for g we substitute $F_{\lambda\rho}$ or A_λ.

The identity transformation, here signified by the notation $G_{\alpha\beta,4}(I)$, leads to the associated conserved quantities

$$\int G_{4j,4}(I)\, d\mathbf{x} = \int [-iA_4E_j - (\mathbf{A}\wedge\mathbf{B})_j]\, d\mathbf{x}, \quad (6.9)$$

$$\frac{\epsilon_{ijk}}{2}\int G_{ij,4}(I)\, d\mathbf{x} = \int [-i(\mathbf{A}\wedge\mathbf{E})_k - A_4B_k]\, d\mathbf{x}. \quad (6.9)$$

The infinitesimal displacements lead to no new conserved quantities, since $G_{\alpha\beta,4}(\lambda_\tau) = \frac{1}{2}\partial_\tau G_{\alpha\beta,4}(I)$. Consequently, for space displacements, the associated density is a 3-space divergence, and for time displacements, the associated conserved quantities are zero since these are merely the derivatives with respect to time of those conserved quantities given in Eqs. (6.9).

Similarly, the infinitesimal Lorentz transformations provide no new conserved quantities. The density corresponding to $\xi_{\eta\rho}$ premultipliers is

$$G_{\alpha\beta,4}(\xi_{\eta\rho}) = \{\tfrac{1}{2}\partial_k[(\delta_{\rho k}x_\eta - \delta_{\eta k}x_\rho)G_{\alpha\beta,4}(I) + (\delta_{\eta 4}x_\rho - \delta_{\rho 4}x_\eta)G_{\alpha\beta,k}(I)] + \tfrac{1}{2}(\delta_{\rho 4}x_\eta - \delta_{\eta 4}x_\rho)[\partial_\lambda G_{\alpha\beta,\lambda}(I)] + \tfrac{1}{2}[\delta_{\alpha\eta}G_{\rho\beta,4}(I) - \delta_{\beta\eta}G_{\rho\alpha,4}(I) - \delta_{\alpha\rho}G_{\eta\beta,4}(I) + \delta_{\beta\rho}G_{\eta\alpha,4}(I)]\}. \quad (6.10)$$

However, $\partial_\lambda G_{\alpha\beta,\lambda}(I)$ is zero, the 3-space divergence may be ignored, so consequently we are left only with densities obtained previously from the identity transformation.

For the space-reflection transformation, we obtain the conserved quantities

$$\epsilon_{ijk}\int G_{ij,4}\begin{pmatrix}\text{space}\\ \text{refl}\end{pmatrix} d\mathbf{x} = 4\int \{-i[\mathbf{A}(-\mathbf{x}, t) \wedge \mathbf{E}(\mathbf{x}, t)]_k + A_4(-\mathbf{x}, t)B_k(\mathbf{x}, t)\}\, d\mathbf{x}, \quad \int G_{4i,4}\begin{pmatrix}\text{space}\\ \text{refl}\end{pmatrix} d\mathbf{x} = 0. \quad (6.11)$$

Also, from the time-reflection transformation, we obtain the conserved quantities[8]

$$\frac{\epsilon_{ijk}}{2}\int G_{ij,4}\begin{pmatrix}\text{time}\\ \text{refl}\end{pmatrix} d\mathbf{x} = \int \{-i[\mathbf{A}(\mathbf{x}, -t) \wedge \mathbf{E}(\mathbf{x}, t) + \mathbf{E}(\mathbf{x}, -t) \wedge \mathbf{A}(\mathbf{x}, t)]_k + [A_4(\mathbf{x}, -t)B_k(\mathbf{x}, t) - B_k(\mathbf{x}, -t)A_4(\mathbf{x}, t)]\}\, d\mathbf{x} = 0, \quad (6.12)$$

$$\int G_{4i,4}\begin{pmatrix}\text{time}\\ \text{refl}\end{pmatrix} d\mathbf{x} = \int \{i[A_4(\mathbf{x}, -t)E_i(\mathbf{x}, t) + E_i(\mathbf{x}, -t)A_4(\mathbf{x}, t)] - [\mathbf{A}(\mathbf{x}, -t) \wedge \mathbf{B}(\mathbf{x}, t) - \mathbf{B}(\mathbf{x}, -t) \wedge \mathbf{A}(\mathbf{x}, t)]_i\}\, d\mathbf{x}.$$

Forms Independent of the 4-Potential

We have seen that the conserved quantities

$$\int V_4\, d\mathbf{x} = -i\int [\mathbf{A}'(x)\cdot\mathbf{E}(x) - \mathbf{E}'(x)\cdot\mathbf{A}(x)]\, d\mathbf{x}, \quad (6.13)$$

$$\int \bar{V}_4\, d\mathbf{x} = i\int [\mathbf{A}'(x)\cdot\mathbf{B}(x) - \mathbf{B}'(x)\cdot\mathbf{A}(x)]\, d\mathbf{x} \quad (6.14)$$

do not really depend on the Lorentz gauge, and are in fact gauge-independent. Can these expressions be written in a form such that the potential **A** does not explicitly appear, but is replaced by expressions containing only the field quantities **E** and **B**?

[8] These expressions may be evaluated at $t = 0$, since they are independent of time.

The answer to the preceding question may be obtained by considering an integral expression for **A** in the Coulomb gauge for which div **A** = 0. Since div **B** = 0 and curl **A** = **B**, we may write

$$\mathbf{A}(\mathbf{x}, t) = \frac{1}{4\pi}\int \frac{1}{|\mathbf{x} - \mathbf{y}|}\,\text{curl}\,\mathbf{B}(\mathbf{y}, t)\, d\mathbf{y}. \quad (6.15)$$

Here, the differentiation implied by the curl is with respect to the argument **y**. Substituting this expression into Eqs. (6.13) and (6.14), we obtain the potential-free expressions

$$\int V_4\, d\mathbf{x} = \frac{-i}{4\pi}\int\int \frac{d\mathbf{x}\, d\mathbf{y}}{|\mathbf{x} - \mathbf{y}|}[(\text{curl}\,\mathbf{B}'(y))\cdot\mathbf{E}(x) - (\text{curl}\,\mathbf{E}'(y))\cdot\mathbf{B}(x)], \quad (6.16)$$

$$\int \bar{V}_4\, d\mathbf{x} = \frac{i}{4\pi}\int\int \frac{d\mathbf{x}\, d\mathbf{y}}{|\mathbf{x} - \mathbf{y}|}[(\text{curl}\,\mathbf{B}'(y))\cdot\mathbf{B}(x) - (\text{curl}\,\mathbf{B}'(y))\cdot\mathbf{B}(x)] = 0. \quad (6.17)$$

In obtaining these expressions, we have used the fact that surface terms obtained through integration by parts may be ignored. Also, since the time variables in x and y are the same, we have used the symmetry in the integrals to interchange **x** and **y**. It is completely obvious now that $\int \bar{V}_4\, d\mathbf{x}$ (which is actually zero) is incapable of providing any nontrivial conserved quantities—a fact we have noted earlier in this section.

All the usual connections between infinitesimal transformations and conserved quantities (e.g., space displacement–momentum) arise, as has been demonstrated, from the symmetry transformations involving the 4-vector $V_\mu = A'_\lambda F_{\lambda\mu} - F'_{\lambda\mu}A_\lambda$. The form of the associated conserved quantity displayed in Eq. (6.16) is similar to, but not the same as, that discussed by Good.[9]

7. MAXWELL'S EQUATIONS: CLASS II CONSERVED QUANTITIES

What we have called the class II type conservation law involves the symmetric tensors $W_{\mu\nu}(F', F)$ and $Z_{\mu\nu}(F', F)$ [see Eqs. (5.20), (5.21)]. The corresponding conserved quantities are

$$\int W_{4\nu}(F', F)\, d\mathbf{x} \quad \text{and} \quad \int Z_{4\nu}(F', F)\, d\mathbf{x}. \quad (7.1)$$

It is apparent that these quantities are completely independent of any gauge condition, since they are only functions of the **E** and **B** fields. In this section. we shall use the same conventions (described at the beginning of Sec. 6) to designate a particular symmetry transformation as those employed earlier.

[9] R. H. Good, Jr., Nuovo Cimento **24**, 713 (1962).

Conserved Quantities from the Tensor $W_{\mu\nu}(F', F)$

From the identity transformation (signified by the argument I) we obtain the density

$$W_{4\nu}(I) = -2i\delta_{\nu j}(\mathbf{E} \wedge \mathbf{B})_j + \delta_{\nu 4}(E^2 + B^2). \quad (7.2)$$

Thus, the associated conserved quantities here are just the expressions for momentum and energy that we had obtained earlier by considering the class I quantity V_4. In the earlier case, however, the relevant symmetry transformation was not the identity but the infinitesimal displacements.

Here, the infinitesimal displacements λ_α provide the density

$$W_{4\nu}(\lambda_\alpha) = \tfrac{1}{2}\partial_\alpha W_{4\nu}(I). \quad (7.3)$$

Consequently, the conserved quantites associated with infinitesimal displacements are trivially zero since they are either the integral over a 3-divergence or the time derivative of a conserved quantity.

Also, the infinitesimal Lorentz transformation ($\xi_{\alpha\beta}$ premultiplier) gives no new conserved quantities. In order to demonstrate this, we write the associated density in the form

$$\begin{aligned} W_{4\nu}(\xi_{\alpha\beta}) = \{&\tfrac{1}{2}(\delta_{\beta 4}x_\alpha - \delta_{\alpha 4}x_\beta)[\partial_\lambda W_{\lambda\nu}(I)] \\ &+ \tfrac{1}{2}\partial_k[(\delta_{\beta k}x_\alpha - \delta_{\alpha k}x_\beta)W_{4\nu}(I) \\ &+ (\delta_{\alpha 4}x_\beta - \delta_{\beta 4}x_\alpha)W_{k\nu}(I)] \\ &+ \tfrac{1}{2}[\delta_{\alpha\nu}W_{\beta 4}(I) - \delta_{\beta\nu}W_{\alpha 4}(I)]\}. \end{aligned} \quad (7.4)$$

However, $\partial_\lambda W_{\lambda\nu}(I)$ is zero and we may neglect the 3-space divergence. Thus, we are left with densities already obtained for the identity transformation.

The space-reflection transformation provides the associated conserved quantities

$$\begin{aligned} \int W_{44}\begin{pmatrix}\text{space}\\ \text{refl}\end{pmatrix} d\mathbf{x} &= \int [\mathbf{E}(-\mathbf{x}, t)\cdot\mathbf{E}(\mathbf{x}, t) \\ &\quad - \mathbf{B}(-\mathbf{x}, t)\cdot\mathbf{B}(\mathbf{x}, t)]\, d\mathbf{x}, \\ \int W_{4k}\begin{pmatrix}\text{space}\\ \text{refl}\end{pmatrix} d\mathbf{x} &= i\int [\mathbf{B}(\mathbf{x}, t) \wedge \mathbf{E}(-\mathbf{x}, t) \\ &\quad - \mathbf{B}(-\mathbf{x}, t) \wedge \mathbf{E}(\mathbf{x}, t)]_k\, d\mathbf{x} = 0, \end{aligned} \quad (7.5)$$

and the time-reflection transformation provides the associated conserved quantities[8]

$$\begin{aligned} \int W_{44}\begin{pmatrix}\text{time}\\ \text{refl}\end{pmatrix} d\mathbf{x} &= \int [\mathbf{B}(\mathbf{x}, -t)\cdot\mathbf{B}(\mathbf{x}, t) \\ &\quad - \mathbf{E}(\mathbf{x}, -t)\cdot\mathbf{E}(\mathbf{x}, t)]\, d\mathbf{x}, \\ \int W_{4k}\begin{pmatrix}\text{time}\\ \text{refl}\end{pmatrix} d\mathbf{x} &= i\int [\mathbf{E}(\mathbf{x}, -t) \wedge \mathbf{B}(\mathbf{x}, t) \\ &\quad + \mathbf{B}(\mathbf{x}, -t) \wedge \mathbf{E}(\mathbf{x}, t)]_k\, d\mathbf{x} = 0. \end{aligned} \quad (7.6)$$

In summary, the only new or nontrivial conservation laws provided by consideration of the densities $W_{4\nu}$ are the two expressions obtained from W_{44} for the discontinuous space-and time-reflection transformations.

Conserved Quantities from the Tensor $Z_{\mu\nu}(F', F)$

Using the fact that

$$F_{\mu\rho}F^{\mathrm{D}}_{\rho\nu} = -\delta_{\mu\nu}(\mathbf{E}\cdot\mathbf{B}), \quad (7.7)$$

we may easily show that the tensor $Z_{\mu\nu}(I)$ associated with the identity transformation is identically zero. Thus, for this symmetry transformation, no nontrivial conservation law is provided.

For infinitesimal displacements, the associated tensor $Z_{\mu\nu}(\lambda_\alpha)$ may be written

$$\begin{aligned} Z_{\mu\nu}(\lambda_\alpha) = &\tfrac{1}{2}(\delta_{\mu\xi}\delta_{\nu\eta} + \delta_{\mu\eta}\delta_{\nu\xi}) \\ &\times [(\partial_\alpha F^{\mathrm{D}}_{\xi\rho})F_{\rho\eta} - (\partial_\alpha F_{\xi\rho})F^{\mathrm{D}}_{\rho\eta}]. \end{aligned} \quad (7.8)$$

Now for F and F^{D} antisymmetric, the following expression is an identity:

$$\begin{aligned} &[(\partial_\alpha F^{\mathrm{D}}_{\xi\rho})F_{\rho\eta} - (\partial_\alpha F_{\xi\rho})F^{\mathrm{D}}_{\rho\eta}] \\ &\quad = (\{[\partial_\alpha F^{\mathrm{D}}_{\xi\rho} + \partial_\rho F^{\mathrm{D}}_{\alpha\xi} + \partial_\xi F^{\mathrm{D}}_{\rho\alpha}]F_{\rho\eta} \\ &\qquad - [\partial_\alpha F_{\xi\rho} + \partial_\rho F_{\alpha\xi} + \partial_\xi F_{\rho\alpha}]F^{\mathrm{D}}_{\rho\eta} \\ &\qquad + F^{\mathrm{D}}_{\alpha\xi}[\partial_\rho F_{\rho\eta}] - F_{\alpha\xi}[\partial_\rho F^{\mathrm{D}}_{\rho\eta}] \\ &\qquad - \partial_\rho[F^{\mathrm{D}}_{\alpha\eta}F_{\rho\xi} - F_{\alpha\xi}F^{\mathrm{D}}_{\rho\eta} - F_{\xi\eta}F^{\mathrm{D}}_{\rho\alpha} \\ &\qquad + i\epsilon_{\rho\alpha\eta\beta}F_{\xi\delta}F_{\delta\beta}] + \partial_\xi[F_{\rho\alpha}F^{\mathrm{D}}_{\rho\eta} - F^{\mathrm{D}}_{\rho\alpha}F_{\rho\eta}]\} \\ &\qquad - \partial_\rho\{(F^{\mathrm{D}}_{\alpha\xi}F_{\rho\eta} - F^{\mathrm{D}}_{\alpha\eta}F_{\rho\xi}) + F_{\xi\eta}F^{\mathrm{D}}_{\rho\alpha}\} \\ &\qquad + \{F^{\mathrm{D}}_{\rho\alpha}(\partial_\xi F_{\rho\eta}) - F_{\rho\alpha}(\partial_\xi F^{\mathrm{D}}_{\rho\eta}) \\ &\qquad + i\epsilon_{\rho\alpha\eta\beta}\partial_\rho(F_{\xi\delta}F_{\delta\beta})\}). \end{aligned} \quad (7.9)$$

We note that in the first curly bracket { } of the preceding expression, each term in square brackets is zero because of Maxwell's equation or the defining relation of the dual tensor. Also, the expression in the second curly bracket is antisymmetric in ξ and η. Consequently, we may write $Z_{\mu\nu}(\lambda_\alpha)$ in the alternative form

$$\begin{aligned} Z_{\mu\nu}(\lambda_\alpha) = &\tfrac{1}{2}(\delta_{\mu\xi}\delta_{\nu\eta} + \delta_{\mu\eta}\delta_{\nu\xi})[F^{\mathrm{D}}_{\rho\alpha}(\partial_\xi F_{\rho\eta}) \\ &- F_{\rho\alpha}(\partial_\xi F^{\mathrm{D}}_{\rho\eta}) + i\epsilon_{\rho\alpha\eta\beta}\partial_\rho(F_{\xi\delta}F_{\delta\beta})]. \end{aligned} \quad (7.10)$$

The motivation for writing $Z_{\mu\nu}(\lambda_\alpha)$ in the preceding form is that it may easily be compared with the third-rank tensor "zilch" discussed by Lipkin.[4] We find that

$$Z_{\mu\nu}(\lambda_\alpha) = -iZ_{\mu\nu\alpha} \quad \text{(Lipkin)}. \quad (7.11)$$

Although the relevant tensor is the same, our for-

mulation has demonstrated that $\partial_\mu Z_{\mu\nu}(\lambda_\alpha) = 0$, while Lipkin has shown in detail that $\partial_\alpha Z_{\mu\nu}(\lambda_\alpha) = 0$. That Lipkin's equation is indeed satisfied is obvious from the form of $Z_{\mu\nu}(\lambda_\alpha)$ given by Eq. (7.8), since both F and F^D satisfy the d'Alembertian equation $\partial_\alpha\partial_\alpha F = \partial_\alpha\partial_\alpha F^D = 0$.

Both 4-divergence equations involving $Z_{\mu\nu}(\lambda_\alpha)$ (Lipkin's and ours) lead to the same conserved quantities. This can be seen in the following way. Using Eqs. (7.8), (7.10), and Maxwell's equations, we can prove that

$$\{Z_{\mu\nu}(\lambda_\alpha) - Z_{\alpha\nu}(\lambda_\mu)\} = \tfrac{1}{2}\partial_\rho[F^D_{\rho\mu}F_{\alpha\nu} - F_{\rho\mu}F^D_{\alpha\nu} - i\epsilon_{\rho\mu\nu\beta}F_{\alpha\delta}F_{\delta\beta} - i\epsilon_{\rho\mu\alpha\beta}F_{\nu\delta}F_{\delta\beta}]. \quad (7.12)$$

Consequently, the different densities for the two types of 4-divergence equations, namely $Z_{4\nu}(\lambda_\alpha)$ and $Z_{\alpha\nu}(\lambda_4)$, differ only by a 3-space divergence, so by the usual assumptions concerning asymptotic behavior, the corresponding conserved quantities are the same.

Explicitly, the densities for the conserved quantities are:

$$Z_{44}(\lambda_4) = [(\partial_4\mathbf{E})\cdot\mathbf{B} - \mathbf{E}\cdot(\partial_4\mathbf{B})] = -i[\mathbf{B}\cdot(\text{curl }\mathbf{B}) + \mathbf{E}\cdot(\text{curl }\mathbf{E})], \quad (7.13)$$

$$Z_{44}(\lambda_j) = [(\partial_j\mathbf{E})\cdot\mathbf{B} - \mathbf{E}\cdot(\partial_j\mathbf{B})] = -i[\mathbf{B}\wedge(\partial_4\mathbf{B}) + \mathbf{E}\wedge(\partial_4\mathbf{E})]_j + \partial_k(E_jB_k - B_jE_k), \quad (7.14)$$

$$Z_{4j}(\lambda_4) = -i[\mathbf{B}\wedge(\partial_4\mathbf{B}) + \mathbf{E}\wedge(\partial_4\mathbf{E})]_j, \quad (7.15)$$

$$Z_{4j}(\lambda_i) = i[(\partial_i\mathbf{B})\wedge\mathbf{B} + (\partial_i\mathbf{E})\wedge\mathbf{E}]_j = \{\delta_{ij}[(\partial_4\mathbf{B})\cdot\mathbf{E} - (\partial_4\mathbf{E})\cdot\mathbf{B}] + 2[(\partial_4E_j)B_i - E_i(\partial_4B_j)] - i\partial_k[\epsilon_{kjn}(B_iB_n + E_iE_n)]\}. \quad (7.16)$$

To obtain the conserved quantities, we may neglect the 3-space divergences (which appear in the densities expressed in terms of derivatives with respect to time). Consequently, the conserved quantity obtained from $Z_{44}(\lambda_j)$ is the same as that obtained from $Z_{4j}(\lambda_4)$. Also, since

$$2[(\partial_4E_j)B_i - E_i(\partial_4B_j)] = \{[(\partial_4E_j)B_i - E_j(\partial_4B_i)] + [(\partial_4E_i)B_j - E_i(\partial_4B_j)] - \epsilon_{ijk}\partial_4(\mathbf{E}\wedge\mathbf{B})_k\}, \quad (7.17)$$

and the integral over $\partial_4(\mathbf{E}\wedge\mathbf{B})$ is zero (because momentum is conserved), we see from Eq. (7.16) that $Z_{4j}(\lambda_i)$ gives the same conserved quantity that $Z_{4i}(\lambda_j)$ does. Moreover, $Z_{44}(\lambda_4)$ equals minus $Z_{4k}(\lambda_k)$ (to within a 3-space divergence). Hence, there are a total of nine distinct[10] conserved quantities obtained from the infinitesimal displacements. Recently, these conservation laws for the zilch of the electromagnetic field have also been discussed by Candlin,[11] Kibble,[12] and Morgan.[12a] Here, we see how this type of conservation law fits in with the general description of relativistic free particles—specifically, that it is a consequence of an invariance property (infinitesimal displacements) within a certain class of conservation laws.

For the infinitesimal Lorentz transformations ($\xi_{\alpha\beta}$ premultipliers), the associated tensor $Z_{\mu\nu}(\xi_{\alpha\beta})$ may be written,

$$Z_{\mu\nu}(\xi_{\alpha\beta}) = \{x_\alpha Z_{\mu\nu}(\lambda_\beta) - x_\beta Z_{\mu\nu}(\lambda_\alpha) + i\epsilon_{\alpha\beta\mu\pi}F_{\pi\rho}F_{\rho\nu} + i\epsilon_{\alpha\beta\nu\pi}F_{\pi\rho}F_{\rho\mu}\}. \quad (7.18)$$

Consequently, the densities of the associated conserved quantities are

$$Z_{44}(\xi_{4k}) = \{-ix_4[\mathbf{B}\wedge(\partial_4\mathbf{B}) + \mathbf{E}\wedge(\partial_4\mathbf{E})]_k - x_k[(\partial_4\mathbf{E})\cdot\mathbf{B} - \mathbf{E}\cdot(\partial_4\mathbf{B})] + \partial_n[x_4(E_kB_n - B_kE_n)]\}, \quad (7.19)$$

$$\tfrac{1}{2}\epsilon_{ijk}Z_{44}(\xi_{jk}) = (\{-i\mathbf{x}\wedge[\mathbf{B}\wedge(\partial_4\mathbf{B}) + \mathbf{E}\wedge(\partial_4\mathbf{E})]\} + \partial_k[(\mathbf{x}\wedge\mathbf{E})_iB_k - (\mathbf{x}\wedge\mathbf{B})_iE_k]), \quad (7.20)$$

$$Z_{4i}(\xi_{4k}) = (ix_k[\mathbf{B}\wedge(\partial_4\mathbf{B}) + \mathbf{E}\wedge(\partial_4\mathbf{E})]_i + x_4\{\delta_{ik}[\partial_4\mathbf{B})\cdot\mathbf{E} - (\partial_4\mathbf{E})\cdot\mathbf{B}] + 2[(\partial_4E_i)B_k - E_k(\partial_4B_i)]\} - i\partial_n[\epsilon_{nij}x_4(B_jB_k + E_jE_k)]), \quad (7.21)$$

$$\tfrac{1}{2}\epsilon_{njk}Z_{4i}(\xi_{jk}) = (i\partial_k\epsilon_{kji}[(\mathbf{x}\wedge\mathbf{B})_nB_j + (\mathbf{x}\wedge\mathbf{E})_nE_j] + \epsilon_{njk}x_j\{\delta_{ik}[(\partial_4\mathbf{B})\cdot\mathbf{E} - (\partial_4\mathbf{E})\cdot\mathbf{B}] + 2[(\partial_4E_i)B_k - E_k(\partial_4B_i)]\}). \quad (7.22)$$

Neglecting the 3-space divergences, the 24 densities given by Eqs. (7.19) to (7.22) have the general form

$$Z_{4\nu}(\xi_{\alpha\beta}) = x_\alpha Z_{4\nu}(\lambda_\beta) - x_\beta Z_{4\nu}(\lambda_\alpha),$$

if we use those expressions for $Z_{4\nu}(\lambda_\beta)$, $Z_{4\nu}(\lambda_\alpha)$ that involve the time derivatives [see Eqs. (7.13) to (7.16)] and omit completely the associated 3-space diver-

[10] Lipkin refers to ten quantities but, as he also points out, only nine of them are independent.

[11] D. J. Candlin, "Analysis of the New Conservation Law in Electromagnetic Theory," Nuovo Cimento (to be published).

[12] T. W. B. Kibble, "Conservation Laws for Free Fields," J. Math. Phys. (to be published).

[12a] *Footnote added in proof*: Another approach to the conservation of "zilch" has recently been given by T. A. Morgan, J. Math. Phys. 5, 1659 (1964).

gences. Thus, the densities for the 24 conserved quantities (due to Lorentz transformation invariance) are related to the densities of zilch (due to displacement invariance) in *exactly the same way* that the densities of angular momentum–center of energy (due to Lorentz invariance) are related to the densities of momentum–energy (due to displacement invariance). Without unduly proliferating the nomenclature, one could justifiably call these new conserved quantities $\int Z_{4\nu}(\xi_{\alpha\beta})\, d\mathbf{x}$, by the name "angular zilch."

Actually, not all of the 24 conserved quantities obtained from the densities displayed in Eqs. (7.19) to (7.22) are distinct. There exist the relations;

$$[\epsilon_{ijk}Z_{4k}(\xi_{4j}) + \tfrac{1}{2}\epsilon_{ijk}Z_{44}(\xi_{jk})] = \partial_n[ix_4(B_iB_n + E_iE_n) + (\mathbf{x}\wedge\mathbf{E})_iB_n - (\mathbf{x}\wedge\mathbf{B})_iE_n], \tag{7.23}$$

$$\epsilon_{ijk}Z_{4i}(\xi_{jk}) = i\partial_n[(\mathbf{B}\cdot\mathbf{x})B_n + (\mathbf{E}\cdot\mathbf{x})E_n] + i(E^2 + B^2). \tag{7.24}$$

Consequently, 20 new distinct conserved quantities for the electromagnetic field are obtained from the infinitesimal Lorentz transformations. It is believed that these conservation laws for "angular zilch" are presented here for the first time.

The space reflection transformation provides the associated conserved quantities

$$\begin{aligned}\int Z_{44}\begin{pmatrix}\text{space}\\ \text{refl}\end{pmatrix} d\mathbf{x} &= \int [\mathbf{B}(-\mathbf{x}, t)\cdot\mathbf{E}(\mathbf{x}, t) + \mathbf{E}(-\mathbf{x}, t)\cdot\mathbf{B}(\mathbf{x}, t)]\, d\mathbf{x},\\ \int Z_{4k}\begin{pmatrix}\text{space}\\ \text{refl}\end{pmatrix} d\mathbf{x} &= i\int [\mathbf{E}(-\mathbf{x}, t)\wedge\mathbf{E}(\mathbf{x}, t) - \mathbf{B}(-\mathbf{x}, t)\wedge\mathbf{B}(\mathbf{x}, t)]_k\, d\mathbf{x} = 0,\end{aligned} \tag{7.25}$$

and the time-reflection transformation provides the associated conserved quantities[8]

$$\begin{aligned}\int Z_{44}\begin{pmatrix}\text{time}\\ \text{refl}\end{pmatrix} d\mathbf{x} &= -\int [\mathbf{B}(\mathbf{x}, -t)\cdot\mathbf{E}(\mathbf{x}, t) + \mathbf{E}(\mathbf{x}, -t)\cdot\mathbf{B}(\mathbf{x}, t)]\, d\mathbf{x},\\ \int Z_{4k}\begin{pmatrix}\text{time}\\ \text{refl}\end{pmatrix} d\mathbf{x} &= i\int [\mathbf{B}(\mathbf{x}, -t)\wedge\mathbf{B}(\mathbf{x}, t) - \mathbf{E}(\mathbf{x}, -t)\wedge\mathbf{E}(\mathbf{x}, t)]_k\, d\mathbf{x} = 0.\end{aligned} \tag{7.26}$$

Thus, two new conservation laws (from Z_{44}) are also obtained by considering the discontinuous transformations.

ACKNOWLEDGMENTS

The author is grateful to Professor R. H. Good, Jr., for valuable discussions, and to Professor E. Amaldi and Professor M. Conversi for the hospitality of the Istituto di Fisica "G. Marconi" dell'Università, Roma.

Spatially Inhomogeneous States of Many-Body Systems*

M. EGER AND E. P. GROSS†
Brandeis University, Waltham, Massachusetts
(Received 29 October 1964)

To treat many-body systems in the presence of a static potential or problems of highly collective spatially inhomogeneous motions such as vortex lines, wavefunctions of a type

$$\Psi = \prod_{i=1}^{N} g(\mathbf{x}_i)\Phi(\mathbf{x}_1, \cdots, \mathbf{x}_N)$$

have been proposed. Here Φ is the exact ground state of the homogeneous system and $g(\mathbf{x})$ is a one-particle state introduced to describe the effect of the spatial inhomogeneity. However, to determine $g(\mathbf{x})$ by the variation principle, one needs to know the spatial correlation functions of all orders for the homogeneous many-body system. It is shown that the method of point transformations allows one to work with qualitatively similar but different states. The description of the system in the presence of a static impurity or of a state representing a vortex line in liquid helium requires a knowledge of only the average kinetic energy and x-ray scattering factor for homogeneous liquid helium. Both of these are available from experiments. The treatment of a recoiling impurity atom, strictly speaking, requires a knowledge of the current correlation tensor for the ground state of the homogeneous many-body system. This term vanishes in the Hartree limit of the theory for bosons.

1. INTRODUCTION

ONE important feature of current many-body theory is apparent on looking at the problem of the effect of a static potential on the properties of the system. A typical Hamiltonian is

$$H_U = H + \sum_{i=1}^{N} U(\mathbf{x}_i)$$
$$H = \sum_i \frac{p_i^2}{2M} + \sum_{i<j} V(|\mathbf{x}_i - \mathbf{x}_j|). \tag{1.1}$$

If U is a finite disturbance, even a knowledge of all the stationary states of the homogeneous system governed by H is of little value on the analysis of H_U unless a perturbation treatment of U is valid. If a disturbance can be regarded as small, the response can be expressed in terms of simple correlation functions that characterize the homogeneous system. We are then in the happy position of having nondisturbing probes, as in the theory of x-ray and neutron scattering. However, disturbing probes are frequently important, as for example, in the behavior of foreign bodies in a many-body system. Here it seems to be necessary to develop the theory afresh, without reference to the properties of the system in absence of the foreign body. In certain cases this can actually be done. For example, one can make a theory[1] of foreign ions in a weakly interacting boson system based on an extension to spatially inhomogeneous states of the Hartree–Bogolyubov method. This theory parallels the treatment of the isolated weakly interacting boson system in spirit, and is systematic in the same sense. However, in a more realistic situation (liquid helium) there is no adequate theory that starts from the actual Hamiltonian, and the situation is even worse for the foreign-ion problem.

While this difficulty is perhaps intrinsic insofar as one strives for exact theories, one can take another point of view. This attitude is to make a reasonable but approximate theory of the more complex problem (e.g., a foreign particle plus a many-body system), in which the quantities of interest depend on properties of the isolated many-body system. If one is lucky, these properties may be available directly from experiments. A specific example is Feynman's[2] approach to the excitation spectrum of liquid helium. In the simplest theory the wavefunction of an excitation is assumed to be the Fourier component of the density multiplied by the (unknown) exact ground-state wavefunction. The energy estimate turns out to involve only the two-body spatial correlation function (or x-ray structure function) which can be measured. The limitations as well as the possibilities of this type of approach are brought out in the improved theory of Feynman and Cohen.[3] When the wavefunction of an excitation is approx-

* This work was supported by a grant from the U. S. Air Force Office of Scientific Research (AF 176–63).
† Fulbright Research Scholar, University of Rome, Italy (1963–64).
[1] E. P. Gross, Ann. Phys. (N. Y.) **19**, 234 (1962).
[2] R. P. Feynman, Phys. Rev. **94**, 262 (1954).
[3] R. P. Feynman and M. Cohen, Phys. Rev. **102**, 1189 (1956).

imated by the ground state multiplied by a polynomial in the density Fourier components, the energy depends on the three- and four-body spatial correlation functions. This information is not available experimentally, and one is forced to make estimates—which while not unreasonable, introduce uncertainties. And yet this is perhaps the best theory that has been possible to date.

Let us examine what happens if one adopts this attitude in other problems. Consider, for example, the case of an interacting boson system in the presence of a localized static potential $U(r)$. A reasonable approximation to the ground state is

$$\Psi = \prod_{i=1}^{N} g(r_i)\Phi(\mathbf{x}_1, \cdots, \mathbf{x}_N), \qquad r_i = |\mathbf{x}_i|, \tag{1.2}$$

where Φ is the exact ground state of the system in the absence of the potential. The function $g(r)$ tends to unity at a sufficient distance from the origin. If $U(r)$ has the behavior of a hard core at small distances, $g(r)$ can satisfy the boundary conditions that it vanish at the edges of the core. One would like to calculate the energy with this Ψ and to choose $g(r)$ by making the energy a minimum. However, one immediately encounters the difficulty that the energy depends on the spatial correlation functions of all orders. Without further, rather drastic approximations, the way is blocked. The situation is that when states are of the form

$$\Psi = \sum_{i=1}^{N} g(\mathbf{x}_i)\Phi,$$

the energy depends on only two-body correlation functions, when

$$\Psi = \sum_{i<j} g(\mathbf{x}_i, \mathbf{x}_j)\Phi,$$

a small number of higher correlation functions enter, but that

$$\Psi = \prod_{i=1}^{N} g(\mathbf{x}_i)\Phi$$

brings in all of them. The last form is typical of highly collective states, e.g., vortex lines.

In the present paper, the problem of dealing with this type of collective state is attacked by the method of point transformations.[4,5] We introduce in place of the particle coordinates $\mathbf{x}_i$ new coordinates which contain functions $f(\mathbf{x}_i)$. The transformation deforms the space in the vicinity of the spatial inhomogeneity, but leaves the coordinates unaltered at large distances. This enables us to take account of the spatial inhomogeneity induced, for example, by a static potential. After transformation, we evaluate the expectation value of the energy with a wavefunction in the new frame that is the (unknown) exact ground state of the homogeneous system. One then finds that the energy functional containing $f(\mathbf{x})$ depends only on the two-body correlation functions and on the expectation value of the kinetic energy of the homogeneous system. Both of these quantities can be taken from experiment. This procedure is equivalent to using the variational principle with a trial function in the original frame

$$\Psi = \prod_{i=1}^{N} g(\mathbf{x}_i)\Phi(\mathbf{h}(\mathbf{x}_1), \cdots, \mathbf{h}(\mathbf{x}_N)). \tag{1.3}$$

Here the function $\mathbf{h}(\mathbf{x})$, as well as $g(\mathbf{x})$, tends to unity at large distances. They are both connected with $f(\mathbf{x})$ in a precise and specific way. The choice amounts to a deformation of the Φ in the vicinity of the static potential. While one can state the procedure in this purely variational way, the feature that makes the calculation go through simply is the specific link between $g(\mathbf{x})$ and $\mathbf{h}(\mathbf{x})$. This could hardly be guessed without the method of coordinate transformation. Conversely, a rather drastic approximation to the ground state is thereby implied. One indication that the procedure makes some sense lies in the fact (of Sec. 3) that it reduces to the Hartree theory for weakly interacting bosons.

The point transformation approach to spatially inhomogeneous states is very general, and may be used to treat systems in the presence of walls, the structure of vortex lines, the case of recoiling ions, etc. Since we only make use of transformations that are symmetrical in the *dynamical variables* of the many-body system, our method is applicable, essentially as it stands, to fermions as well as bosons. However, we have had in mind the case of liquid helium in developing the approach, and this bias is reflected in the text of the article. In some of these problems we encounter expectation values taken with the homogeneous ground state that do not seem to be obtainable experimentally. For example, in the recoiling foreign-ion problem one needs to know the single time (two-body) current correlation tensor. The situation is then similar to that of the Feynman–Cohen theory in that some further estimates of these quantities are needed. This brings in the attendant uncertainty as to how much of the error lies in the special form of the trial function. These questions are further elaborated in the text.

In Sec. 2 we find an approximate ground state

[4] M. Eger and E. P. Gross, Ann. Phys. (N. Y.) **24**, 63 (1963).

[5] M. Eger and E. P. Gross, Nuovo Cimento **34**, 1225 (1964).

for a many-body system in the presence of a static potential $U(\mathbf{x})$. It is shown that the determination of the optimum point transformation requires a knowledge of the average kinetic energy and two-body spatial correlation function for the ground state of the many-body system in the absence of the impurity. In Sec. 3 it is shown that the theory reduces to the Hartree treatment of the static potential problem for weakly interacting bosons. This treatment corresponds to complete macroscopic occupancy of a single one-particle spatially inhomogeneous state, and to putting the spatial correlation function equal to a constant. In Sec. 4 the generality of the procedure for the treatment of collective states is emphasized and is illustrated by a sketch of the calculation of the structure of a vortex line in liquid helium. In Sec. 5 we study the ground state of an impurity atom coupled dynamically to a system of interacting particles. The energy functional determining the point transformation differs from that of the static potential problem in two ways. First the mass M of the particle must be replaced by the effective mass $M^* = \mu M/(M + \mu)$, where μ is the mass of the impurity atom. Second, a new term containing the current correlation tensor for the ground state of the homogeneous many-body system makes its appearance. This quantity does not seem to be available from existing experiments. However, it is shown that this term vanishes in the Hartree limit and it may be permissible to neglect it in fixing the point transformation. In Sec. 5, we take up the theory of the effective mass of a foreign body. If one adopts the general form of the wavefunction proposed by Feynman for liquid helium, the point transform allows one to fix the undetermined features of the theory, i.e., the specification of the spatial correlation functions of the ion–boson system. The treatment of the problem strictly within the philosophy of the point-transform approach leads to some generalizations, but these are mainly merely alternative representations of the Feynman theory and its obvious extensions.

2. MANY-BODY SYSTEM IN A STATIC POTENTIAL

Let us consider a many-body system in the presence of a static potential $U(r)$ localized at the origin. The Hamiltonian is

$$H_U = \sum_{i=1}^{N} \frac{p_i^2}{2M} + \sum_{i<j} V(\mathbf{x}_i - \mathbf{x}_j) + \sum_{i=1}^{N} U(\mathbf{x}_i), \quad (2.1)$$

where $\mathbf{p}_i = -i\hbar\, \partial/\partial \mathbf{x}_i$. We introduce the extended point transformation

$$y_i^\mu = x_i^\mu(1 + f(r_i)),$$
$$\mu = 1, 2, 3 \qquad i = 1, 2, \cdots, N. \quad (2.2)$$

Each particle is transformed separately. This makes possible a complete discussion of the inverse transformation, Jacobian, etc. Under the coordinate transformation, a function $Q(\mathbf{x}_1, \cdots, \mathbf{x}_N)$ that depends only on the coordinates goes over to

$$Q'(\mathbf{x}_1, \cdots, \mathbf{x}_N) \equiv Q(\mathbf{y}_1(\mathbf{x}_1), \cdots, \mathbf{y}_N(\mathbf{x}_N)). \quad (2.3)$$

For example, the transformed static potential is

$$\left[\sum_{i=1}^{N} U(\mathbf{x}_i)\right]' = \sum_{i=1}^{N} U(r_i + r_i f(r_i)). \quad (2.4)$$

If one deals with potentials that are Fourier analyzable, one need only find the transform of

$$\tilde{\rho}(k) = \sum_{i=1}^{N} e^{i\mathbf{k}\cdot\mathbf{x}_i}$$
$$\tilde{\rho}(k)' = \sum_{i=1}^{N} e^{i\mathbf{k}\cdot\mathbf{y}_i} = \sum_{i=1} e^{i\mathbf{k}\cdot\mathbf{x}_i} e^{i\mathbf{k}\cdot\mathbf{x}_i f(r_i)}. \quad (2.5)$$

If this is reanalyzed in x space, we find (Ω is the quantization volume)

$$\tilde{\rho}(\mathbf{k})' = (1/\Omega) \sum_{\mathbf{l}} \beta(\mathbf{k}, \mathbf{l})\tilde{\rho}(\mathbf{l}), \quad (2.6)$$

where

$$\beta(\mathbf{k}, \mathbf{l}) \equiv \int e^{-i(\mathbf{k}-\mathbf{l})\cdot\mathbf{x}} e^{-i\mathbf{k}\cdot\mathbf{x}f(r)}\, d^3x. \quad (2.7)$$

The point transformation is a linear transformation on the $\tilde{\rho}(\mathbf{k})$. In this form

$$\sum_{i=1}^{N} U(\mathbf{x}_i)' = \frac{1}{\Omega^2} \sum_{\mathbf{k},\mathbf{l}} \tilde{U}(-\mathbf{k})\beta(\mathbf{k}, \mathbf{l})\tilde{\rho}(\mathbf{l})$$
$$\tilde{U}(\mathbf{k}) \equiv \int e^{-i\mathbf{k}\cdot\mathbf{x}} U(\mathbf{x})\, d^3x \quad (2.8)$$
$$U(\mathbf{x}) \equiv \frac{1}{\Omega} \sum_{k} e^{i\mathbf{k}\cdot\mathbf{x}} \tilde{U}(\mathbf{k}).$$

The transform of the particle–particle interaction energy is

$$\sum_{i<j} V'(\mathbf{x}_i - \mathbf{x}_j) = \sum_{i<j} V(|\mathbf{x}_i(1 + f(r_i)) - \mathbf{x}_j(1 + f(r_j)|) \quad (2.9)$$

For particle–particle potentials that are Fourier analyzable,

$$\sum_{i<j} V'(\mathbf{x}_i - \mathbf{x}_j) = \frac{1}{2\Omega^3} \sum_{\mathbf{k},\mathbf{l},\mathbf{l}'} \tilde{V}(-k)\beta(\mathbf{k}, \mathbf{l}) \times \beta(-\mathbf{k}, \mathbf{l}') \sum_{i \neq j} e^{-i(\mathbf{l}\cdot\mathbf{x}_i + \mathbf{l}'\cdot\mathbf{x}_j)}. \quad (2.10)$$

Equations (2.4) and (2.9), or alternatively (2.8) and (2.10), give the potential-energy terms of the transformed Hamiltonian. After the kinetic energy is transformed (to be accomplished shortly), we are of course free to treat the new Hamiltonian as we please. However, in the present work we will simply evaluate the expectation value of the *transformed* Hamiltonian with the exact ground state of the homogeneous many-body system. Then the effect of the static potential is carried entirely by the point transformation. The expectation value of (2.4) plus (2.9) taken with $\Phi(\mathbf{x}_1, \cdots, \mathbf{x}_N)$ is

$$\frac{N}{\Omega}\int U[r(1+f)]\,d^3x + \frac{N(N-1)}{2}\iint V(|\mathbf{x}_1(1+f_1) - \mathbf{x}_2(1+f_2)|)p_2(1,2)\,d^3x_1\,d^3x_2, \tag{2.11}$$

where

$$p_2(1,2) \equiv p_2(\mathbf{x}_1, \mathbf{x}_2) = \int \Phi^2(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N)\,d^3x_3 \cdots d^3x_N \tag{2.12}$$

is the two-body correlation function for the ground state of the homogeneous boson system. The energy functional thus depends on $f(r)$ and on the correlation function. With the Fourier transform description the energy is

$$(N/\Omega^2)\sum \tilde{U}(k)\beta(\mathbf{k},0) + \frac{1}{2\Omega^3}\sum_{\mathbf{k},\mathbf{l}} \tilde{V}(\mathbf{k})\beta(\mathbf{k},\mathbf{l})\beta(-\mathbf{k},-\mathbf{l})\tilde{G}(\mathbf{l}) \tag{2.13}$$

$$\tilde{G}(\mathbf{l}) = \int \Phi^2 \sum_{i\neq j} e^{i\mathbf{l}\cdot(\mathbf{x}_i-\mathbf{x}_j)}\,d^Nx = \int G(r)e^{-i\mathbf{l}\cdot\mathbf{x}}\,d^3x$$

$$G(\mathbf{x}) = \int \Phi^2 \sum_{i\neq j}\delta(\mathbf{x}_i - \mathbf{x}_j - \mathbf{x})\,d^Nx. \tag{2.14}$$

One notes in the above, and in subsequent calculations, that Φ is translationally invariant and is an eigenstate of the total momentum with eigenvalue zero. To transform the kinetic energy, we must return to the basic point transformation, Eq. (2.2). The momenta conjugate to y_i^μ are defined by

$$P_i^\mu = \tfrac{1}{2}\sum_{\nu=1}^{3}(p_i^\nu\,\partial x_i^\nu/\partial y_i^\mu + (\partial x_i^\nu/\partial y_i^\mu)p_i^\nu), \tag{2.15}$$

where each particle is transformed separately. One finds

$$\frac{\partial x_i^\mu}{\partial y_i^\nu} = \int \delta(\mathbf{x}_i - \mathbf{x})\frac{1}{1+f(r)} \times\left\{\delta_{\mu,\nu} - \frac{f'}{r}\frac{x_\mu x_\nu}{1+(rf)'}\right\}d^3x. \tag{2.16}$$

The Jacobian $B = |\partial x_i^\mu/\partial y_i^\nu|$ is

$$B = \prod_{i=1}^{N} B_1(r_i) = \prod_{i=1}^{N}\frac{1}{(1+f_i)^2}\frac{1}{1+(r_if_i)'}. \tag{2.17}$$

The transform of the kinetic energy $(1/2M)\sum_{i=1}^N p_i^2$ is given by replacing p_i^μ by $P_i^\mu(\mathbf{p}_i, \mathbf{x}_i)$. The significant point is that it is the sum of one-body terms so that when we take the expectation value with $\Phi(\mathbf{x}_1 \cdots \mathbf{x}_N)$ the fact that Φ is an eigenfunction of the momentum introduces drastic simplifications. The calculation is, however, a bit tedious, and we do not include it. The results can be written in the form

$$\frac{1}{2M}\sum_{i,\mu} P_i^\mu P_i^\mu = \sum \frac{p_i^\mu}{2M}\frac{\partial x_i^\mu}{\partial y_i^\lambda}\frac{\partial x_i^\nu}{\partial y_i^\lambda}p_i^\nu + \sum_i W(\mathbf{x}_i), \tag{2.18}$$

where $W(\mathbf{x})$ is a metric potential energy

$$W(\mathbf{x}) = \frac{\hbar^2}{2M}\frac{B_1^{\frac{1}{2}}}{r^2}\frac{d}{dr}\frac{r^2}{[1+(rf)']^2}\frac{d}{dr}\frac{1}{B_1^{\frac{1}{2}}}. \tag{2.19}$$

In the expectation value taken with Φ, only the zero Fourier component survives and we find

$$\int \Phi\sum_{i=1}^{N} W(\mathbf{x}_i)\Phi\,d^Nx = \frac{N}{\Omega}\int W(\mathbf{x})\,d^3x. \tag{2.20}$$

To evaluate the expectation value of the first term of Eq. (2.18), quadratic in the momenta, we note that $\partial\Phi/\partial x_i^\nu$ is an eigenfunction of the total momentum. Thus only the zero wave vector Fourier component of $\partial x_i^\mu/\partial y_i^\lambda(\partial x_i^\nu/\partial y_i^\lambda)$, i.e., its spatial average, contributes. Now

$$\int \frac{\partial x^\mu}{\partial y^\lambda}\frac{\partial x^\nu}{\partial y^\lambda}\,d^3x \equiv \bar{L}(f)\,\delta_{\mu,\nu}$$
$$\bar{L}(f) = \frac{4\pi}{3}\int\left[\frac{2}{(1+f)^2} + \frac{1}{[1+(rf)']^2}\right]r^2\,dr. \tag{2.21}$$

Hence the expectation value of the first term of Eq. (2.18) is

$$\frac{1}{2M}\int \Phi\sum_{i=1}^{N} p_i^\mu\frac{\partial x_i^\mu}{\partial y_i^\lambda}\frac{\partial x_i^\nu}{\partial y_i^\lambda}p_i^\mu\Phi\,d^Nx = \frac{\bar{L}(f)}{\Omega}\int \Phi\sum\frac{p_i^2}{2M}\Phi\,d^Nx. \tag{2.22}$$

Here a new property of the homogeneous boson system enters. It is the expectation value of the kinetic energy in the ground state. Collecting terms, we have for the ground-state energy in the presence of the static impurity,

$$\mathcal{E} = \frac{\tilde{L}(f)}{\Omega}\int \Phi \sum \frac{p_i^2}{2M}\Phi\, d^N x + \frac{N}{\Omega}\int W(\mathbf{x})\, d^3x + \frac{N}{\Omega}\int U(r + rf)4\pi r^2\, dr + \frac{N(N-1)}{2}\iint V(|\mathbf{x}_1(1+f_1) - \mathbf{x}_2(1+f_2)|)p_2(1,2)\, d^3x_1\, d^3x_2. \quad (2.23)$$

This is admittedly a very complicated functional of $f(\mathbf{x})$. But the expectation value of the kinetic energy can be obtained for the many-body system from the cohesive energy, the interparticle potential, and correlation energy. With a suitably flexible choice of the functional form of $f(\mathbf{x})$ containing parameters and using computing machines, it should be possible to obtain satisfactory extimates of $f(r)$ and $\mathcal{E}$.

In the next section we examine the Hartree limit for bosons to show that the preceding theory is reasonable. However, one serious shortcoming should be pointed out immediately. The last term in the expression for $\mathcal{E}$ becomes $\int V(r_{12})p_2(1,2)\, d^3x_1\, d^3x_2$ in the absence of the static potential, i.e., it is the potential contribution to the cohesive energy. Here p_2 drops to zero for numerous many-body systems at interparticle distances where a strong short-range repulsion in $V(|\mathbf{x}_1 - \mathbf{x}_2|)$ sets in. In the presence of the static potential, V is modified to $V[\mathbf{y}_1(\mathbf{x}_1) - \mathbf{y}_2(\mathbf{x}_2)]$, and the distortion of the space can lead to large contributions of repulsive potential energy from regions where $p_2(1,2)$ is not zero. Particles lying on the same radial line from the origin are particularly affected. There is, strictly speaking, no way to cope with this. It is traceable to the oversimplified point transformation used, which in turn is responsible for the neat form of the result. It would certainly be possible to remove the trouble by using a more-general point transformation that couples the particles. But the theory then becomes unwieldy, and it seems preferable to simply cut out the dangerous regions in the spatial integral. The regions in question seem to be unimportant, but a definite uncertainty does creep in. It should be noted, however, that the difficulty disappears entirely in two limiting cases. One case is when $f(\mathbf{x})$ is slowly varying, i.e., the spatial inhomogeneity has a characteristic range large compared to interatomic distances. The second case is when $f(r) = c/r$, which is a form that removes the hard-core interaction between the impurity and many-body system. The domain of $|\mathbf{y}_1 - \mathbf{y}_2|$ then coincides with that of $|\mathbf{x}_1 - \mathbf{x}_1|$.

3. ANALYSIS OF THE APPROXIMATION—THE HARTREE LIMIT

We have discussed how the coordinate and position operators and the Hamiltonian change under the simple transformation. The connection between the wavefunctions in the original and transformed frames is obtained by the requirement that

$$H(p, x)\Phi_E(x) = E\Psi_E(x) \quad (3.1)$$

implies

$$H'(p, x)\Phi_E(x) = H(P(p, x), y(x))\Phi_E(x) = E\Phi_E(x). \quad (3.2)$$

[We frequently write $\Psi(x)$ in place of $\Psi(\mathbf{x}_1 \cdots \mathbf{x}_N)$.] In words, an eigenfunction $\Psi_E(x)$ of energy E of the original Hamiltonian corresponds to an eigenfunction $\Phi_E(x)$ of the transformed Hamiltonian $H'(p, x)$. The connection between the wavefunctions is

$$\Psi_E(y(x)) = [B(x)]^{\frac{1}{2}}\Phi_E(x). \quad (3.3)$$

This guarantees that the normalization $\int |\Psi_E(y)|^2 \times d^N y = 1$ implies the correct normalization for $\Phi_E(x)$. In fact,

$$1 = \int |\Psi_E(y)|^2\, d^N y = \int |\Psi_E(y(x))|^2 \frac{d^N x}{B(x)} = \int |\Phi_E(x)|^2\, d^N x. \quad (3.4)$$

The form of Eq. (3.3) is, however, not the most convenient for our purposes since we usually make some assumption as to $\Phi_E(x)$. Then we obtain only the original wavefunction Ψ_E in terms of $y(x)$ as x ranges over its domain. In examining wavefunctions it is more convenient to think of the inverse transform $x(y)$ and to write

$$\Psi_E(y) = [B(x(y))]^{\frac{1}{2}}\Phi_E(x(y)) = [A(y)]^{-\frac{1}{2}}\Phi_E(x(y)). \quad (3.5)$$

Here y is merely a set of numbers so that the original wavefunction is given directly. $A(y) = |\partial y/\partial x|$ and satisfies $A(y) = 1/B(x(y))$. Thus it can be computed from the Jacobian B and the inverse transform $x(y)$ once a form has been obtained for

$\Phi_B(x)$. Alternatively $A(y)$ can be calculated directly from $x(y)$ in the same way that $B(x)$ was computed from $y(x)$. Thus, if

$$x_i^\mu = y_i^\mu(1 + F(R_i)), \qquad R_i = |\mathbf{y}_i|,$$

$$A(y) = \prod_{i=1}^{N} A_1(y_i), \tag{3.6}$$

$$A_1(y) = \frac{1}{(1+F)^2}\frac{1}{1+(RF)'} \equiv \frac{1}{B_1(x(y))}.$$

This representation is most useful in investigating the Hartree limit. The energy functional is simply

$$\begin{aligned}\mathcal{E} = &-\frac{\hbar^2}{2M}\int [A(y)]^{-\frac{1}{2}}\Phi(x(y))\sum_{i=1}^{N}\frac{\partial}{\partial y_i^\mu}\frac{\partial}{\partial y_i^\mu}\\ &\times [A(y)]^{\frac{1}{2}}\Phi(x(y))\,d^Ny + \int \frac{1}{A(y)}\Phi^2(x(y))\\ &\times\left\{\sum_{i=1}^{N} U(\mathbf{y}_i) + \sum_{i<j} V(\mathbf{y}_i - \mathbf{y}_j)\right\} d^Ny. \end{aligned}\tag{3.7}$$

The Hartree approximation corresponds to taking for $\Phi(x)$ the noninteracting boson wavefunction $\Phi(x) = \Omega^{-N/2}$. The original wavefunction is then

$$\Psi(y_1 \cdots y_N) = \prod_{i=1}^{N}[A_1(y_i)]^{-1}, \tag{3.8}$$

i.e., is given simply by the Jacobian. The energy functional then becomes

$$\begin{aligned}\mathcal{E} = &N\frac{\hbar^2}{2M}\int\left(\nabla_y \frac{1}{A_1^{\frac{1}{2}}}\right)^2 + \frac{N}{\Omega}\int\frac{1}{A_1(y)}U(y)\,d^3y\\ &+\frac{N(N-1)}{2}\int V(\mathbf{y}_1 - \mathbf{y}_2)\frac{1}{A_1(y_1)}\frac{1}{A_1(y_2)}\,d^3y_1\,d^3y_2.\end{aligned}\tag{3.9}$$

This is the energy functional obtained in the treatment[1] of the behavior of foreign bodies in a weakly interacting gas in Hartree approximation. The function $f(y)$ used there is to be identified with the present $[A(y)]^{-\frac{1}{2}}$. While the boson–boson ground-state and correlation functions are those of the noninteracting boson system, the interboson potential does play an essential role, along with the external potential and the kinetic energy, in determining $A^{\frac{1}{2}}$. In the Hartree approach $1/A$ is directly the expectation value of the density in the presence of the potential. In the more-general theory, the expectation value of the density is

$$\begin{aligned}n(\xi) &= \int \Psi^2(y)\sum_{i=1}^{N}\delta(\mathbf{y}_i - \xi)\,d^Ny\\ &= \int \Phi^2(x)\sum_{i=1}^{N}\delta(\mathbf{x}_i(1+f_i) - \xi)\,d^Nx\\ &= \frac{N}{\Omega}\int \delta(\mathbf{x}_1[1+f(r_1)] - \xi)\,d^3x_1\\ &= \frac{N}{\Omega}\int \delta(\mathbf{y}_1 - \xi)\frac{1}{A_1(y_1)}\,d^3y_1 = \frac{N}{\Omega}\frac{1}{A_1(\xi)}.\end{aligned}\tag{3.10}$$

These results are simple because of the product form of the Jacobian. The two-body spatial correlation function in the presence of the static potential is

$$\begin{aligned}n_2(\xi, \mathbf{n}) &= \int \Psi^2(y)\sum_{i<j}\delta(\mathbf{y}_i - \xi)\delta(\mathbf{y}_j - \mathbf{n})\,d^Ny\\ &= \frac{N(N-1)}{2}\int \Phi^2(x)\delta(\mathbf{y}_1(\mathbf{x}_1) - \xi)\delta(\mathbf{y}_2(\mathbf{x}_2) - \mathbf{n})\,d^Nx\\ &= \frac{N(N-1)}{2}\int \delta(\mathbf{y}_1(\mathbf{x}_1) - \xi)\delta(\mathbf{y}_2(\mathbf{x}_2) - \mathbf{n})\\ &\qquad\times p_2(\mathbf{x}_1, \mathbf{x}_2)\,d^3\mathbf{x}_1\,d^3\mathbf{x}_2\\ &= \frac{N(N-1)}{2}\frac{1}{A_1(\xi)}\frac{1}{A_1(\mathbf{n})}p_2(\mathbf{x}_1(\xi), \mathbf{x}_2(\mathbf{n})).\end{aligned}\tag{3.11}$$

Thus we have the Jacobian factors (densities relative to the origin of the potential) and the replacement of $\mathbf{x}_1$ and $\mathbf{x}_2$ by the inverse transforms $\mathbf{x}_1(\mathbf{y}_1 = \xi)$, $\mathbf{x}_2(\mathbf{y}_2 = \mathbf{n})$ in the interacting boson correlation function.

4. THE STRUCTURE OF A VORTEX LINE IN LIQUID ^{4}He

It is shown in the following sections that the application of the point transformation method to the problem of a recoiling foreign ion interacting with bosons cannot be made without some approximations and modifications of the general theory as presented thus far. The theory as it stands can, however, be applied directly to the study of the vortex line in liquid ^{4}He and, for example, the calculation of the radius of such a vortex.

We proceed now to derive the relevant transformation formulas for this problem, and to write down the expression for the ground-state energy in terms of the transformation function. To obtain explicit and numerical results, it will remain only to carry out the indicated variational calculation. This last step is postponed, for the time being, until computational facilities become available.

Let Ψ be the ground state of a system of interacting bosons in which a line-vortex excitation exists. Our experience with the Hartree–Bogolyubov theory in treating this problem for dilute systems[6] suggests that we assume the form $\Psi = e^{+i\Sigma\varphi_i}\Psi_1$, with Ψ_1 real, to bring in the fact that each particle

[6] E. P. Gross, Nuovo Cimento 20, 454 (1961).

is given a unit of angular momentum. Then Ψ_1 is an eigenfunction of

$$H_1 = \sum_{i=1}^{N} \frac{p_i^2}{2M} + \sum_{i<j} V(\mathbf{x}_i - \mathbf{x}_j) + \frac{\hbar^2}{2M} \sum_i \frac{1}{\rho_i^2}, \quad (4.1)$$

where ρ_i is the radial distance from the line vortex. To take account of the density variation in the core of the vortex, we introduce the point transformation,

$$\begin{aligned} y_i^\mu &= x_i^\mu(1 + f(\rho_i)), \qquad \mu = 1, 2, \\ y_i^\mu &= x_i^\mu, \qquad \mu = 3. \end{aligned} \quad (4.2)$$

The relevant transformation formulas for such problems, with cylindrical symmetry, that correspond to Eqs. (2.15)–(2.19) for spherical symmetry, are now given. For the momenta, we have

$$\begin{aligned} P_i^\mu &= \tfrac{1}{2} \sum_{\nu=1}^{2} \left(p_i^\nu \frac{\partial x_i^\nu}{\partial y_\mu^\mu} + \frac{\partial x_i^\nu}{\partial y_i^\mu} p_i^\nu \right), \qquad \mu = 1, 2 \\ P_i^\mu &= p_i^\mu, \qquad \mu = 3. \end{aligned} \quad (4.3)$$

Similarly,

$$\begin{aligned} \frac{\partial x_i^\mu}{\partial y_i^\nu} &= \int \delta(\mathbf{x}_i - \mathbf{x}) \frac{1}{(1 + f)} \\ &\times \left\{ \delta_{\mu,\nu} - \frac{f(\rho)}{\rho} \frac{x_\mu x_\nu}{1 + (\rho f)'} \right\} d^3x, \quad \mu, \nu = 1, 2, \\ \frac{\partial x_i^3}{\partial y_i^\nu} &= \frac{\partial x_i^\nu}{\partial y_i^3} = \delta_{3,\nu}, \end{aligned} \quad (4.4)$$

and the Jacobian takes the new form

$$B = \prod_i^N B_i(\rho_i) = \prod_{i=1}^N \frac{1}{1 + f(\rho_i)} \frac{1}{[1 + (\rho_i f)']}. \quad (4.5)$$

Thus, the transformation of the kinetic energy is

$$\begin{aligned} \sum \frac{P_i^2}{2M} &= \frac{1}{2M} \sum_i (p_i^{(3)})^2 \\ &+ \sum_i^N \sum_{\mu,\nu=1,2} p_\mu \left\{ \frac{\delta_{\mu,\nu}}{(1 + f)^2} + \frac{x_\mu x_\nu}{\rho^2} \mathfrak{K} \right\} p_\nu \\ &+ \frac{\hbar^2}{2M} \sum_{i=1}^N W(\rho_i), \end{aligned} \quad (4.6)$$

where $\mathfrak{K}$ retains the same form,

$$\mathfrak{K}(\rho) \equiv \left[\frac{1}{[1 + (\rho f)']^2} - \frac{1}{(1 + f)^2} \right],$$

while W now becomes

$$W(\rho) = \frac{B^{\frac{1}{2}}}{\rho} \frac{d}{d\rho} \frac{\rho}{[1 + (\rho f)']^2} \frac{d}{d\rho} B^{-\frac{1}{2}}. \quad (4.7)$$

Our problem here is a special case of the Hamiltonian (2.1), and except for the fact that we are now dealing with only two dimensions, the transformation of the Hamiltonian and the evaluation of its average value in the new frame proceeds as in Sec. 2. Following the philosophy discussed in the introduction, the average value of the transformed Hamiltonian is taken with the vortex-free ground state of the bosons. The calculation again makes crucial use of the translation invariance of Φ_1, and we obtain

$$\begin{aligned} \langle \Phi | H_1' | \Phi \rangle &= \bar{\mathfrak{K}}(f) \langle \Phi | \sum \frac{p_i^2}{2M} | \Phi \rangle \\ &+ \frac{N}{\Omega} \frac{\hbar^2}{2M} \bar{W} + \frac{2\pi N \ell}{\Omega} \int U(\rho + \rho f) \rho \, d\rho \\ &+ \frac{N(N-1)}{2} \int V[\mathbf{x}_1(1 + f_1) \\ &- \mathbf{x}_2(1 + f_2)] p_2(1, 2) \, d^3x_1 \, d^3x_2, \end{aligned} \quad (4.8)$$

where ℓ is the length of the quantization cylinder, and

$$\bar{\mathfrak{K}}(f) = \frac{2}{3} \frac{\pi \ell}{\Omega} \int \left\{ \frac{1}{[1 + (\rho f)']^2} + \frac{1}{(1 + f)^2} \right\} \rho \, d\rho + \frac{1}{3}. \quad (4.9)$$

The third term in (4.8) has the characteristic logarithmic dependence on the outer limit of the radial integration. Again, only the kinetic energy and structure factor of liquid helium are needed, and these can be taken from experiments. It should be noted that in its present form the energy is written in terms of B as well as f. There is, of course, only one variational function, since B is a given function of f; however for computational purposes it might be advantageous to leave some terms expressed in terms of B, to vary both B and f, and to relate them by means of a Lagrange multiplier. The best form for Eq. (4.8) would depend on the method of computation, and therefore we make no attempts to achieve further simplifications at this point.

In Ref. 6 an estimate of the size of the vortex was obtained for a dilute system by assuming that the density is constant for ρ smaller than some number, a, and has the vortex-free solution outside this cylinder. The continuity of Ψ then determined the core radius a. It was found to be of the order of the de Broglie wavelength associated with the mean energy of interaction per particle. In our treatment here, the core size is to be found by an examination of the form of the Jacobian after the variational function has been evaluated. Since

$$\Psi_1 = \left[\prod_i B(\rho_i) \right]^{\frac{1}{2}} \Phi$$

and Φ is the vortex-free solution, the Jacobian, which rises to one with increasing ρ, is, in effect, the "core" itself.

5. GROUND STATE OF RECOILING FOREIGN BODY

We turn now to the applications of the point transform method to the treatment of the Hamiltonian

$$H = \frac{p^2}{2\mu} + \sum_{i=1}^{N} U(\mathbf{x}_1 - \mathbf{q}) + \sum_{i=1}^{N} \frac{p_i^2}{2M} + \sum_{i<j} V(\mathbf{x}_i - \mathbf{x}_j). \quad (5.1)$$

Here we have an impurity of mass μ, coordinates q and momenta p interacting dynamically with the many body system. The most interesting fact that will emerge is that we also need the current correlation functions for the homogeneous boson system to fix the point transform. The problem may be put in a form resembling that of the previous sections by introducing the unitary transformation

$$U = e^{i\mathbf{q}\cdot(\sum \mathbf{p}_i)/\hbar} \quad (5.2)$$

Then

$$\mathbf{q}' = U\mathbf{q}U^{-1} = \mathbf{q}, \qquad \mathbf{p}_i' = \mathbf{p}_i \quad (5.3)$$

$$\mathbf{x}_i' = \mathbf{x}_i + \mathbf{q}, \qquad \mathbf{p}' = \mathbf{p} - \sum \mathbf{p}_i$$

$$H' = UHU^{-1} = H(\mathbf{p}'(\mathbf{p}, \mathbf{p}_i), \mathbf{q}'(\mathbf{q}), \mathbf{p}_i'(\mathbf{p}_i), \mathbf{x}_i'(\mathbf{x}_i, \mathbf{q}))$$

$$= \frac{1}{2\mu}\left(\mathbf{p} - \sum_{i=1}^{N} \mathbf{p}_i\right)^2 + \sum_{i=1}^{N} \frac{p_i^2}{2M} + \sum_{i=1}^{N} U(\mathbf{x}_i) + \sum_{i<j} V(\mathbf{x}_i - \mathbf{x}_j). \quad (5.4)$$

An eigenfunction $\Psi(\mathbf{x}_i, \mathbf{q})$ of H is related to an eigenfunction $\Psi'(\mathbf{x}_i, \mathbf{q})$ of H' by

$$\Psi(\mathbf{x}_i', \mathbf{q}) = U^{-1}\Psi'(\mathbf{x}_i, \mathbf{q}) = \Psi'(\mathbf{x}_i - \mathbf{q}, \mathbf{q}). \quad (5.5)$$

However, in H', p is a constant of the motion so that $\Psi'(\mathbf{x}_i, \mathbf{q})$ has the form

$$\Omega^{-\frac{1}{2}} e^{i\mathbf{k}\cdot\mathbf{q}}\Psi'(\mathbf{x}_1 - \mathbf{q}).$$

Thus

$$\Psi(\mathbf{x}_i, \mathbf{q}) = \Omega^{-\frac{1}{2}} e^{i\mathbf{k}\cdot\mathbf{q}}\Psi'(\mathbf{x}_i - \mathbf{q}). \quad (5.6)$$

The wavefunctions are eigenstates of the total momentum (original) $\mathbf{p} + \sum \mathbf{p}_i$, with eigenvalue $\hbar\mathbf{K}$. In H', $\mathbf{p}$ has the significance of the total momentum, since $U(\mathbf{p} + \sum \mathbf{p}_i)U^{-1} = \mathbf{p}$. Thus we may find the states of H' with $\mathbf{p}$ put equal to the constant $\hbar\mathbf{K}$.

In this representation, where the spatial inhomogeneity is relative to the instantaneous value of the particle coordinate $\mathbf{q}$, the problem is almost the same as that treated previously. Consider first the case of the ground state $\mathbf{K} = 0$, i.e., total momentum is zero. We have

$$H' = \frac{1}{2\mu}\sum_{j\neq i}^{N} \mathbf{p}_i, \mathbf{p}_j + \frac{1}{2}\left(\frac{1}{\mu} + \frac{1}{M}\right)\sum_{i=1}^{N} p_i^2 + \sum_{i=1}^{N} U(\mathbf{x}_i) + \sum_{i<j} V(\mathbf{x}_i - \mathbf{x}_j). \quad (5.7)$$

If we calculate the ground state by the point transformation method, the last three terms yield the energy functional of Eq. (2.23), except for replacement of $1/M$ by the effective mass $1/M^* = 1/M + 1/\mu$. We need, therefore, consider only the first term. It is, however, more convenient to consider the term $i = \mathrm{j}$ on the same footing as $i \neq j$. Using the transformation (2.15) and taking the expectation value with the exact homogeneous ground state, one finds for the new term

$$\Delta E = \frac{1}{2\mu}\cdot\frac{1}{4}\int \Phi \sum_{i,j}\left(p_i^\nu \frac{\partial x_i^\nu}{\partial y_i^\mu} + \frac{\partial x_i^\nu}{\partial y_i^\mu} p_i^\nu\right) \times \left(p_i^\lambda \frac{\partial x_i^\lambda}{\partial y_i^\mu} + \frac{\partial x_i^\lambda}{\partial y_i^\mu} p_i^\lambda\right)\Phi\, d^N x. \quad (5.8)$$

We now introduce the local current operator

$$j^\nu(\xi, \mathbf{x}_1, \cdots, \mathbf{x}_N) = \frac{1}{2M}\sum_{i=1}^{N} \{p_i^\nu \delta(\mathbf{x}_i - \xi) + \delta(\mathbf{x}_i - \xi)p_i^\nu\} \quad (5.9)$$

and write

$$\frac{\partial x_i^\nu}{\partial y_i^\mu} = \int \frac{\partial x^\nu}{\partial y^\mu}(\xi)\delta(\mathbf{x}_i - \xi)\, d^3\xi. \quad (5.10)$$

Then

$$\Delta E = \frac{M^2}{2\mu}\int \frac{\partial x^\nu}{\partial y^\mu}(\xi)\frac{\partial x^\lambda}{\partial y^\mu}(\mathbf{n}) \times \int \Phi(x) j^\nu(\xi) j^\lambda(\mathbf{n})\Phi(x)\, d^N x\, d^3\xi\, d^3\eta. \quad (5.11)$$

Since $(\partial x^\nu/\partial y^\mu)(\xi)$ is a functional only of $f(\mathbf{x})$ defining the point transformation, this shows that the additional information needed is the static current correlation tensor at different points in space.

In the Hartree approximation where $\Phi(x) = \Omega^{-N/2}$, the contributions from $i \neq j$ are zero and we find only the effective mass correction referred to above. The irreducible current correlation function consists only of the $i \neq j$ part of $j^\nu(\xi)j^\lambda(\mathbf{n})$. We know of no practical way of finding the current correlation func-

tion from experiment. In the application to liquid helium, we would therefore have to ignore this term in the determination of the point transformation, or alternatively to estimate it from some theory of homogeneous liquid helium.

Let us now examine the density correlation functions of the coupled foreign-body–boson system. We examine first the two-body correlation function

$$n_2(\mathbf{Q}, \mathbf{x}) = \int \Psi^2(\mathbf{q}, \mathbf{x}_1, \cdots, \mathbf{x}_N) \times \delta(\mathbf{q} - \mathbf{Q}) \sum_{i=1}^{N} \delta(\mathbf{x}_i - \mathbf{x}), \quad (5.12)$$

where Ψ is the ground state of (5.1). Under the transformation (5.2) we find

$$n_2(\mathbf{Q}, \mathbf{x}) = \int \Psi'^2(\mathbf{x}_1, \cdots, \mathbf{x}_N) \times \sum \delta(\mathbf{x}_i + \mathbf{Q} - \mathbf{x})\, d^N x = n(\mathbf{x} - \mathbf{Q}). \quad (5.13)$$

According to Eq. (3.10), this is

$$(N/\Omega)[A_1(\mathbf{x} - \mathbf{Q})]^{-1}.$$

The three-body correlation function is

$$\begin{aligned} n_3(\mathbf{Q}, \mathbf{x}, \mathbf{x}') &= \int \Psi^2(\mathbf{q}, \mathbf{x})\delta(\mathbf{q} - \mathbf{Q}) \times \sum_{i<j} \delta(\mathbf{x}_i - \mathbf{x})\delta(\mathbf{x}_j - \mathbf{x}')\, d^3q\, d^N x \\ &= \int \Psi'^2(x) \sum_{i<j} \delta(\mathbf{x}_i + \mathbf{Q} - \mathbf{x})\delta(\mathbf{x}_j + \mathbf{Q} - \mathbf{x})\, d^N x \\ &= \frac{1}{A_1(\mathbf{x} - \mathbf{Q})} \frac{1}{A_1(\mathbf{x}' - \mathbf{Q})} \pi_2(h(\mathbf{x} - \mathbf{Q}), h(\mathbf{x} - \mathbf{Q})). \end{aligned} \quad (5.14)$$

In the boson correlation function π_2, $\mathbf{x}$ is to be replaced by the inverse transform. Apart from this characteristic modification arising from the point transform the result for n_3 is similar to a superposition approximation

$$n_3(\mathbf{Q}, \mathbf{x}, \mathbf{x}') \sim n_2(\mathbf{x}, \mathbf{Q}) n_2(\mathbf{x}', \mathbf{Q}) \pi_2(\mathbf{x}, \mathbf{x}').$$

In the foreign-body problem $n_2(\mathbf{x}, \mathbf{Q})$ may be very different from $\pi_2(\mathbf{x}, \mathbf{x}')$. In general one cannot obtain $n_2(\mathbf{x}, \mathbf{Q})$ from experiment; the point transformation reduces its evaluation to properties of the homogeneous many-body system.

6. EFFECTIVE MASS OF A FOREIGN BODY

In this section we discuss the problem of estimating the effective mass of foreign bodies in a system of interacting bosons. We approach the problem by first discussing Feynman's attempt[2] to calculate the effective mass of a ^{3}He impurity in liquid ^{4}He. He takes the trial function

$$\Psi_K(\mathbf{q}, \mathbf{x}) = \exp(i\mathbf{K}\cdot\mathbf{q}) \times \exp\left[\frac{i}{\hbar} \sum_{i=1}^{N} S_K(\mathbf{x}_i - \mathbf{q})\right] \Psi_g(\mathbf{x}_1, \cdots, \mathbf{x}_N). \quad (6.1)$$

Here Ψ_g is the exact, translationally invariant ground state of the *coupled system*. Ψ_K is an eigenvalue $\hbar\mathbf{K}$. The function $S_K(\mathbf{x} - \mathbf{q})$ represents the velocity potential of the boson flow relative to the foreign body. The expectation value of the energy with this trial function is

$$\begin{aligned} \mathcal{E}(\mathbf{K}) = \mathcal{E}(0) &+ \frac{\hbar^2}{2\mu}\left\{K^2 - 2\mathbf{K}\int p_2(r)\nabla S(x)\, d^3x\right\} \\ &+ \frac{\hbar^2}{2M^*}\int \nabla S(x)\nabla S(x) p_2(r)\, d^3x \\ &+ \frac{\hbar^2}{2\mu}\iint \nabla S(x)\nabla S(x') p_3(\mathbf{x}, \mathbf{x}')\, d^3x\, d^3x'. \end{aligned} \quad (6.2)$$

Here $\mathcal{E}(0)$ is the energy of the state Ψ_g, $p_2(r)$ is the radial distribution function about $\mathbf{q}$, and $p_3(\mathbf{x}, \mathbf{x}')$ is the probability of finding ^{4}He atoms at $\mathbf{x}$ and $\mathbf{x}'$ if the ^{3}He atom is at the origin. Feynman's definitions of these functions are

$$p_2(r'_{12}) = \frac{\Omega}{N} \sum_{i,j} \int \delta(\mathbf{x}_i - \mathbf{x}'_i)\delta(\mathbf{x}_j - \mathbf{x}'_j)\Psi_g^2\, d^N x$$

$$p_2(|\mathbf{x} - \mathbf{Q}|) = (\Omega/N) n_2(\mathbf{Q}, \mathbf{x}) \quad (6.3)$$

$$p_3(\mathbf{x} - \mathbf{Q}, \mathbf{x}' - \mathbf{Q}) = (\Omega/N)\{n_3(\mathbf{x}, \mathbf{x}', \mathbf{Q}) + \delta(\mathbf{x} - \mathbf{x}') n_2(\mathbf{x}, \mathbf{Q})\}.$$

If one knows the distribution functions, the determination of a satisfactory expression for $S_K(\mathbf{x})$ by functional variation of $\mathcal{E}(\mathbf{k})$ is a technical matter. The simple dipolar form $S = Az/r^3$ used by Feynman is not essential to the argument. One could examine flows with more general angular and radial variation.[7] Of course, the key feature of the trial function is that the velocity potential of each boson is relative only to $\mathbf{q}$ and does not depend on the positions of other bosons.

The essential difficulties, however, do not seem to lie in this domain, for the wavefunction does appear to be reasonable at least for neutral impurities. It is rather that the correlation functions

[7] G. V. Chester and T. Burke, unpublished thesis of Burke at the University of Birmingham (1962).

are not known, and rather crude estimates have to be made. For example, Feynman assumes that

$$n_3(\mathbf{x}, \mathbf{x}') \sim n_2(r)n_2(r'), \qquad \mathbf{Q} = 0. \tag{6.4}$$

This is cruder than Kirkwood superposition approximation which states that

$$n_3 \sim n_2(\mathbf{Q}, \mathbf{x})n_2(\mathbf{Q}, \mathbf{x}')\pi_2(\mathbf{x}, \mathbf{x}'), \tag{6.5}$$

i.e., the first assumption neglects boson–boson correlations. The latter assumption seems to be adopted by Feynman and Cohen in a reconsideration of the problem.[3] But the difficulties are not ended at this point, for while $\pi_2(\mathbf{x}, \mathbf{x}')$ can be obtained from experiment, this is not possible for $n_2(\mathbf{Q}, \mathbf{x})$. Feynman simply assumes that the radial distribution about the foreign body is the same as that about a ^{4}He atom. While this crude assumption is not hopelessly wrong for a ^{3}He impurity, one has no way to estimate $n_2(\mathbf{q}, \mathbf{x})$ for a more-general foreign body.

The point transformation may be of use in this regard. For if one follows Feynman's theory and makes either of the above assumptions on n_3, one can calculate $n_2(\mathbf{Q}, \mathbf{x})$. If the neglect of the current correlation term is legitimate, one obtains $n_2(\mathbf{Q}, \mathbf{x})$ for a wide class of potentials $U(x)$ and masses μ of the foreign body. This is a particular example of the fact, expounded elsewhere,[4] that a given point transformation is connected simply to the multiparticle density correlation functions. We note in passing that the self-consistent field theory of the effective mass of ions in a weakly interacting boson system is equivalent to the crudest assumption $n_3(\mathbf{x}, \mathbf{x}', \mathbf{Q}) \sim n_2(\mathbf{x} - \mathbf{Q})n_2(\mathbf{x}' - \mathbf{Q})$. The virtue of the Hartree approach is that one actually calculates $n_2(\mathbf{x} - \mathbf{Q})$. One describes the enhancement of the boson density about the ion under the influence of the long-range attractive potential counter balanced by the kinetic energy and the boson–boson repulsions. This fact, together with the proper velocity flow pattern which differs markedly from dipolar flow, yields a high effective mass for the ion.[1] Examination of the energy functional $\mathcal{E}$, Eq. (2.23), shows that we now have an improved description of the density about the ion. The empirical spatial correlation function for helium enables us to treat the energy due to boson–boson repulsion more adequately. The term containing the average kinetic energy of liquid ^{4}He takes into account the depletion of the macroscopically occupied single-particle state.

In actual fact, the systematic treatment of the effective mass problem dictates not the three-body correlation functions given by (6.4) and (6.5) but rather that of Eq. (5.14). Apart from this fact, the theory is identical with that of Feynman. Suppose one writes for the wavefunction of a moving particle

$$\Phi_K(\mathbf{q}, \mathbf{x}_1 \cdots \mathbf{x}_N) = (e^{i\mathbf{k}\cdot\mathbf{q}}/\Omega^{\frac{1}{2}}) \times \exp\left[i \sum t_K(\mathbf{x}_i)\right]\Phi(\mathbf{x}_1, \cdots, \mathbf{x}_N) \tag{6.6}$$

in the representation reached after the point transformation. In the original representation this corresponds to

$$\Psi_K(\mathbf{q}, \mathbf{x}_1, \cdots, \mathbf{x}_N) = \exp(i\mathbf{K}\cdot\mathbf{q}) \times \exp\left[i \sum_i t_K(\mathbf{h}(x_i - \mathbf{q}))\right]\Psi_g(\mathbf{q}, \mathbf{x}_1 \cdots \mathbf{x}_N). \tag{6.7}$$

Ψ_g is the approximate ground state of the coupled system given by our theory. Here t_K involves the inverse point transformation, but can simply be identified with Feynman's $S_K(\mathbf{x}_i - \mathbf{q})$. Hence the major feature of the transformation is indeed that it determines the impurity–boson correlation functions in all orders.

It is, however, of some interest to set down the results of the point transformation directly. We take the wavefunction (6.6) and set $S(\mathbf{y}) = t(\mathbf{x}(\mathbf{y}))$ so as to make correspondence with other theories. Evaluating the expectation value of (5.7) after performing a coordinate transformation, we find the $\mathbf{K}$-dependent terms

$$\begin{aligned}\mathcal{E}(\mathbf{K}) - \mathcal{E}(0) &= \frac{\hbar^2}{2M^*}\frac{N}{\Omega^3}\sum_{ll'}(\mathbf{l}\cdot\mathbf{l}')\tilde{S}(\mathbf{l})\tilde{S}(\mathbf{l}')\beta(\mathbf{l}+\mathbf{l}', 0) \\ &+ \frac{\hbar^2}{\mu}\left\{\mathbf{K}^2 - \frac{2N}{\Omega^2}\sum_l \mathbf{K}\cdot\mathbf{l}\,\tilde{S}(l)\beta(\mathbf{l}, 0)\right\} \\ &+ \frac{\hbar^2}{\mu}\frac{N^2}{\Omega^3}\sum_{\substack{k,k'\\ l}}(\mathbf{k}\cdot\mathbf{k}')\tilde{S}(\mathbf{k})\tilde{S}(\mathbf{k}') \\ &\times \beta(\mathbf{k}, \mathbf{l})\beta(\mathbf{k}', -\mathbf{l})\tilde{G}(\mathbf{l}).\end{aligned} \tag{6.8}$$

Here S is proportional to $\mathbf{K}$, and $\beta(\mathbf{k}, \mathbf{l})$ and $\tilde{G}(\mathbf{l})$ have been defined in Eqs. (2.7) and (2.14). The form corresponding to Eq. (6.2) is

$$\begin{aligned}\mathcal{E}(\mathbf{K}) - \mathcal{E}(0) &= \frac{\hbar^2}{\mu}\left\{\frac{\mathbf{K}^2 - 2\mathbf{K}}{\Omega}\sum_l \mathbf{l}\tilde{S}(\mathbf{l})\bar{p}_2(l) \right. \\ &\left. + \sum_{kk'}\tilde{S}(k)\tilde{S}(k')\bar{p}_3(\mathbf{k}, \mathbf{k}')\right\} \\ &+ \frac{\hbar^2}{2M^*}\sum_{ll'}(\mathbf{l}\cdot\mathbf{l}')\tilde{S}(\mathbf{l})\tilde{S}(\mathbf{l}')\bar{p}_2(\mathbf{l}+\mathbf{l}').\end{aligned} \tag{6.9}$$

Apart from the effective mass differences, Feynman takes $\bar{p}_3(\mathbf{k}, \mathbf{k}') = \bar{p}_2(\mathbf{k})\bar{p}_2(\mathbf{k}') = \tilde{G}(\mathbf{k})\tilde{G}(\mathbf{k}')$. The point transformation approximation is

$$\bar{p}_3(\mathbf{k}, \mathbf{k}') \to \frac{1}{\Omega}\sum_i \beta(\mathbf{k}, \mathbf{l})\beta(\mathbf{k}', -\mathbf{l})\tilde{G}(\mathbf{l}) \tag{6.10}$$

All of the work of this section is concerned with situations where wavefunctions like (1.2) are indeed reasonable. These wavefunctions imply free mobility of the bosons throughout all space. This is very likely the case for a vortex line. However, Atkins has suggested[8] that the actual situation when a foreign ion is present in liquid helium is very different. He pictures the ion as surrounded by a large number of helium atoms bound to form a solid-like structure, with little spatial mobility for the bound atoms. Outside of the droplet, there is free mobility; we feel that Atkins' picture is substantially correct and that the work of this paper and of Ref. 1 is directly relevant for the fluid outside the droplet. It is, in fact, possible to adapt the approach outlined here to fit Atkins' idea. However, this requires a more detailed description in microscopic terms of the solid-like droplet, and is thus nontrivial from a practical point of view.

7. SUMMARY

It is clear that the coordinate transformation method makes it possible to treat a large number of spatially inhomogeneous collective states in an interacting boson system. In the Hartree approximation the states have the form

$$\Psi(\mathbf{x}_1, \cdots, \mathbf{x}_N) = \prod_{i=1}^{N} g(\mathbf{x}_i),$$

where $g(\mathbf{x})$ is in general a complex one-particle state. $g(\mathbf{x})$ describes the space variation in the velocity and density fields. It is determined[9] by the counterplay of the kinetic energy and of the boson–boson repulsions. The real part, i.e., the density, tends to unity at large distances, where the state is that of an ideal Bose gas. One is then tempted to describe the collective state in liquid helium as

$$\Psi = \prod_{i=1}^{N} g(\mathbf{x}_i)\Phi(\mathbf{x}_1 \cdots \mathbf{x}_N)$$

in order to correct the asymptotic behavior. The remarks in the introduction are now relevant, i.e., it is very difficult to work with this wavefunction. However, a point transformation in which each particle is transformed separately can be computed exactly. The Jacobian of the transformation brings in the spatially inhomogeneous factor $\prod_{i=1}^{N} g(\mathbf{x}_i)$ and there are accompanying shifts in the coordinates of Φ.

ACKNOWLEDGMENT

This work was supported by a grant from the U. S. Air Force Office of Scientific Research. One of the authors (E.P.G.) expresses appreciation for support from Fulbright grant under which tenure the work was completed. We are grateful to Professor G. Careri and to members of the Physics Department at the University of Rome for their gracious hospitality and interest.

[8] K. R. Atkins, Phys. Rev. **116**, 1339 (1959).

[9] E. P. Gross, J. Math. Phys. **4**, 195 (1963).

Structure of Gravitational Sources*

A. I. JANIS AND E. T. NEWMAN
Department of Physics, University of Pittsburgh, Pittsburgh, Pennsylvania
(Received 19 June 1964)

The purpose of this paper is to propose a definition of multipole structure of gravitational sources in terms of the characteristic initial data for asymptotic solutions of the field equations. This definition is based upon a detailed study of the corresponding data for the linearized equations and upon the close analogy between the Maxwell and the linearized gravitational fields.

IT is well known that in linear field theories it is possible to obtain solutions of the equations in a source-free region of space in terms of a distant source distribution. (Frequently, instead of a source distribution, it is useful to use the equivalent multipole moments of the source.) For example, in Maxwell theory, the electric and magnetic fields can be written as integrals of the retarded (or advanced) charge and current densities, or equivalently they can be written as a series over the moments of the densities. However when we come to a nonlinear theory, as the Einstein theory of gravitation, results of this type have so far been impossible to derive. This is due essentially to the extreme difficulties inherent in the nonlinearity of the equations. The situation is actually worse than this; it also has so far been impossible to give a method for systematically generating exact empty-space solutions, even forgetting about the relation of solution to source, except in a few cases with special properties (symmetries,[1-3] etc.).

Nevertheless a great deal of progress has been made in the past few years in a slightly different direction. Though physically interesting exact solutions have not been forthcoming in any abundance, methods have been developed to yield interesting asymptotically (i.e., for large distances from a source) exact solutions.[4-7] These methods, which are intrinsically tied to null or characteristic surfaces and the initial data given on these surfaces, are of sufficient generality that they very likely yield all solutions which are asymptotically flat. The main purpose of this paper is to show how to give a physical interpretation to a large class of these asymptotic solutions or equivalently how to give the interpretation to the initial null surface data which gives rise to these solutions; i.e., we propose a method for identifying the moments of the sources from a study of the asymptotic fields. We rely heavily on the notation and results of two earlier papers and refer to them as NP[6] and NU.[7]

Though, as we have said, the prime purpose of this paper is a study of the gravitational field, a secondary purpose is a study of a Maxwell field and its analogy to the gravitational field. The examination of the Maxwell field will serve several functions: (a) it is of intrinsic interest in itself to see and study a new formulation of Maxwell's equations, (b) pedagogically it will aid in the presentation of the analogous results for the gravitational field, and (c) the remarkable analogy (to be shown) between the two fields indicates the direction for further investigation in the more complicated gravitational theory.

The paper is divided into two main sections; one on the Maxwell field, the other on the gravitational field. In Sec. IA, we present Maxwell's equations in the tetrad, spin-coefficient formalism. In Sec. IB we analyze the algebra of the field. In Sec. IC we study the null surface data and the solutions of Maxwell's equations. In Sec. ID we give a physical meaning to the data and the associated solutions. In Sec. II exactly the same things are done, but this time for the linearized gravitational field, the results then being extrapolated to the exact theory. In the Appendix, we show the relationship in the linearized theory of gravity between the sources and the fields.

I. MAXWELL THEORY

A. Field Equations

We briefly summarize in this section the spin-coefficient formalism for the Maxwell field [see (NP)

* Supported in part by Aerospace Research Labs., Office of Aerospace Research, U. S. Air Force.
[1] A. Taub. Ann. Math. **53**, 472 (1951).
[2] B. K. Harrison, Phys. Rev. **116**, 1285 (1959).
[3] I. Robinson and A. Trautman, Proc. Roy. Soc. (London) **A265**, 463 (1962).
[4] H. Bondi, M. van der Burg, and A. Metzner, Proc. Roy. Soc. (London) **A269**, 21 (1962).
[5] R. K. Sachs, Proc. Roy. Soc. (London) **A270**, 103 (1962).
[6] E. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962).
[7] E. Newman and T. Unti, J. Math. Phys. **3**, 891 (1962).

for details]. In Minkowski space we introduce four linearly independent vectors; two, l^μ and n^μ, are null with $l^\mu l_\mu = n^\mu n_\mu = 0$ and $l_\mu n^\mu = 1$, the other two, m^μ and $\bar{m}^\mu$, are also null but in addition are complex with $m_\mu m^\mu = \bar{m}_\mu \bar{m}^\mu = 0$, $m_\mu \bar{m}^\mu = -1$, and $m_\mu l^\mu = m_\mu n^\mu = 0$. m^μ can be written as $m^\mu = (a^\mu + ib^\mu)/\sqrt{2}$, a^μ and b^μ being unit spacelike vectors. $\bar{m}^\mu$ is the complex conjugate of m^μ. Instead of the Maxwell field tensor $F_{\mu\nu}$, we introduce as our field variables the tetrad components of $F_{\mu\nu}$ by the following definitions:

$$\begin{aligned} \Phi_0 &= F_{\mu\nu} l^\mu m^\nu, \\ \Phi_1 &= \tfrac{1}{2} F_{\mu\nu}(l^\mu n^\nu + \bar{m}^\mu m^\nu), \\ \Phi_2 &= F_{\mu\nu} \bar{m}^\mu n^\nu. \end{aligned} \tag{1.1}$$

The six real components of $F_{\mu\nu}$ are replaced by the three complex Φ's. In terms of these new variables, the Maxwell equations are

$$D\Phi_1 - \bar{\delta}\Phi_0 = (\pi - 2\alpha)\Phi_0 + 2\rho\Phi_1 - \kappa\Phi_2, \tag{1.2}$$

$$D\Phi_2 - \bar{\delta}\Phi_1 = -\lambda\Phi_0 + 2\pi\Phi_1 + (\rho - 2\epsilon)\Phi_2, \tag{1.3}$$

$$\delta\Phi_1 - \Delta\Phi_0 = (\mu - 2\gamma)\Phi_0 + 2\tau\Phi_1 - \sigma\Phi_2, \tag{1.4}$$

$$\delta\Phi_2 - \Delta\Phi_1 = -\nu\Phi_0 + 2\mu\Phi_1 + (\tau - 2\beta)\Phi_2, \tag{1.5}$$

where

$$\begin{aligned} D &= l^\mu \partial/\partial x^\mu, \\ \delta &= m^\mu \partial/\partial x^\mu, \\ \bar{\delta} &= \bar{m}^\mu \partial/\partial x^\mu, \\ \Delta &= n^\mu \partial/\partial x^\mu, \end{aligned} \tag{1.6a}$$

and

$$\begin{aligned} \pi &= -n_{\mu;\nu}\bar{m}^\mu l^\nu, & \alpha &= \tfrac{1}{2}(l_{\mu;\nu} n^\mu \bar{m}^\nu - m_{\mu;\nu}\bar{m}^\mu \bar{m}^\nu), \\ \rho &= l_{\mu;\nu} m^\mu \bar{m}^\nu, & \kappa &= l_{\mu;\nu} m^\mu l^\nu, \\ \lambda &= -n_{\mu;\nu}\bar{m}^\mu \bar{m}^\nu, & \epsilon &= \tfrac{1}{2}(l_{\mu;\nu} n^\mu l^\nu - m_{\mu;\nu}\bar{m}^\mu l^\nu), \\ \mu &= -n_{\mu;\nu}\bar{m}^\mu m^\nu, & \gamma &= \tfrac{1}{2}(l_{\mu;\nu} n^\mu n^\nu - m_{\mu;\nu}\bar{m}^\mu n^\nu), \\ \tau &= l_{\mu;\nu} m^\mu n^\nu, & \sigma &= l_{\mu;\nu} m^\mu m^\nu, \\ \nu &= -n_{\mu;\nu}\bar{m}^\mu n^\nu, & \beta &= \tfrac{1}{2}(l_{\mu;\nu} n^\mu m^\nu - m_{\mu;\nu}\bar{m}^\mu m^\nu). \end{aligned} \tag{1.6b}$$

Before these equations can be solved, a choice must be made for the tetrad vectors as well as the coordinate system in which these vectors will be represented. However before this is done, we will discuss the freedom in the Φ's that results from the freedom in the tetrad system.

B. Tetrad Rotation and the Algebra of the Φ's

The vectors l^μ, n^μ, m^μ, and $\bar{m}^\mu$ (at each point in space) are defined by their "orthogonality" conditions only up to the six-parameter restricted Lorentz group. These transformations can be written in the following way:

(a)
$$\begin{aligned} \tilde{l}^\mu &= l^\mu, \\ \tilde{m}^\mu &= m^\mu + a l^\mu, \\ \tilde{n}^\mu &= n^\mu + a\bar{m}^\mu + \bar{a} m^\mu + a\bar{a} l^\mu, \end{aligned} \tag{1.7}$$

which are the so-called two-parameter null rotations around l^μ (a is an arbitrary complex function);

(b)
$$\begin{aligned} \tilde{l}^\mu &= \lambda l^\mu, \\ \tilde{n}^\mu &= \lambda^{-1} n^\mu, \\ \tilde{m}^\mu &= e^{i\phi} m^\mu, \end{aligned} \tag{1.8}$$

which are the ordinary Lorentz transformations in the l^μ, n^μ plane and a spatial rotation in the m^μ, $\bar{m}^\mu$ plane (λ and ϕ are real functions);

(c)
$$\begin{aligned} \tilde{l}^\mu &= l^\mu + b\bar{m}^\mu + \bar{b} m^\mu + b\bar{b} n^\mu, \\ \tilde{m}^\mu &= m^\mu + b n^\mu, \\ \tilde{n}^\mu &= n^\mu, \end{aligned} \tag{1.9}$$

which are similar to (a).

It is easily calculated that under these transformations the Φ's transform in the following fashion:

(a)
$$\begin{aligned} \tilde{\Phi}_0 &= \Phi_0, \\ \tilde{\Phi}_1 &= \Phi_1 + \bar{a}\Phi_0, \\ \tilde{\Phi}_2 &= \Phi_2 + 2\bar{a}\Phi_1 + \bar{a}^2\Phi_0; \end{aligned} \tag{1.10}$$

(b)
$$\begin{aligned} \tilde{\Phi}_0 &= \lambda e^{i\phi}\Phi_0, \\ \tilde{\Phi}_1 &= \Phi_1, \\ \tilde{\Phi}_2 &= \lambda^{-1} e^{-i\phi}\Phi_2; \end{aligned} \tag{1.11}$$

(c)
$$\begin{aligned} \tilde{\Phi}_0 &= \Phi_0 + 2b\Phi_1 + b^2\Phi_2, \\ \tilde{\Phi}_1 &= \Phi_1 + b\Phi_2, \\ \tilde{\Phi}_2 &= \Phi_2. \end{aligned} \tag{1.12}$$

For the moment, we shall confine the discussion to the transformation (c). From Eqs. (1.9), it can be shown that by an appropriate choice of b, $\tilde{l}^\mu$ can be made to point in any null direction except that of n^μ. If we wish to choose an $\tilde{l}^\mu$ such that $\tilde{\Phi}_0 = 0$, it is easily seen from Eqs. (1.12) that the appropriate b must be a root of the equation

$$0 = \Phi_0 + 2b\Phi_1 + b^2\Phi_2.$$

In the case where the two roots are distinct (non-

null fields) there are two possible $\tilde{l}^\mu$ which are singled out. These two vectors (called the principal null vectors of the Maxwell tensor) are the real null eigenvectors of $F_{\mu\nu}$. If $\tilde{l}^\mu$ is chosen as one of them and $\tilde{n}^\mu$ as the other [using transformation (a) now] the $\tilde{\Phi}$'s become $\tilde{\Phi}_0 = \tilde{\Phi}_2 = 0$ and $\tilde{\Phi}_1 \neq 0$. Note that now when $\tilde{l}^\mu$ and $\tilde{n}^\mu$ are chosen as the principal null vectors and the only freedom we have is the transformation (b), $\tilde{\Phi}_1$ is invariantly defined. Actually it can be simply expressed as a function of $E^2 - B^2$ and $\mathbf{E}\cdot\mathbf{B}$, the only invariants of the Maxwell field.

In the case of the two equal roots (called the null case) there is only one principal null vector. It is easily seen from (1.12) that $\tilde{\Phi}_0 = 0$ implies that $\tilde{\Phi}_1 = 0$, so that $\tilde{\Phi}_2$ is the only nonvanishing component.

Though these results are not new[8] and are well known, their presentation in this form is of use in that it shows the analogy with similar results in gravitational theory (Sec. IIB).

C. Null-Surface Data and Solutions of Maxwell's Equations

As mentioned previously, before we can solve the Maxwell equations (1.2)–(1.5), we must make a choice of coordinate and tetrad system. The simplest choice (for specialized problems other choices may be more useful) consists of adopting null polar coordinates and an associated tetrad. Starting with polar coordinates in the Minkowski line element, $ds^2 = dt^2 - dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2)$, the coordinate transformation, $u = t - r$ (r, θ, and ϕ remaining the same), yields the null-coordinate line element

$$ds^2 = du^2 + 2dudr - r^2(d\theta^2 + \sin^2\theta d\phi^2).$$

The surfaces $u =$ constant are just the light cones emanating from the origin $r = 0$.

At any point in space the tetrad can be chosen in the following manner: l^μ is the outward null vector tangent to the cone, n^μ is the inward null vector pointing toward the origin, and m^μ and $\bar{m}^\mu$ are vectors tangent to the two-dimensional sphere defined by constant r and u. These vectors in the null coordinate system ($x^0 = u$, $x^1 = r$, $x^2 = \theta$, $x^3 = \phi$) have the form

$$\begin{aligned} l^\mu &= \delta^\mu_1, \qquad n^\mu = \delta^\mu_0 - \tfrac{1}{2}\delta^\mu_1, \\ m^\mu &= \frac{1}{\sqrt{2}}\frac{1}{r}\left(\delta^\mu_2 + \frac{i}{\sin\theta}\delta^\mu_3\right), \\ \bar{m}^\mu &= \frac{1}{\sqrt{2}}\frac{1}{r}\left(\delta^\mu_2 - \frac{i}{\sin\theta}\delta^\mu_3\right). \end{aligned} \tag{1.13}$$

[8] J. L. Synge, *Relativity, The General Theory* (North-Holland Publishing Company, Amsterdam, 1960).

By direct calculation from their definitions, Eqs. (1.6b), the spin coefficients are found to be

$$\begin{aligned} &\pi = \kappa = \epsilon = \lambda = \gamma = \nu = \tau = \sigma = 0, \\ &\rho = -\frac{1}{r}, \qquad \alpha = -\frac{1}{2\sqrt{2}\,r}\cot\theta, \\ &\beta = \frac{1}{2\sqrt{2}\,r}\cot\theta, \qquad \mu = -\frac{1}{2r}. \end{aligned} \tag{1.14}$$

The Maxwell equations now become

$$\frac{\partial}{\partial r}\Phi_1 + \frac{2}{r}\Phi_1 = \frac{1}{\sqrt{2}\,r}(\bar{\eth} + \cot\theta)\Phi_0, \tag{1.15a}$$

$$\frac{\partial}{\partial r}\Phi_2 + \frac{1}{r}\Phi_2 = \frac{1}{\sqrt{2}\,r}\bar{\eth}\Phi_1, \tag{1.15b}$$

$$\frac{\partial\Phi_0}{\partial u} - \frac{1}{2}\frac{\partial\Phi_0}{\partial r} - \frac{\Phi_0}{2r} = \frac{1}{\sqrt{2}\,r}\eth\Phi_1, \tag{1.15c}$$

$$\frac{\partial\Phi_1}{\partial u} - \frac{1}{2}\frac{\partial\Phi_1}{\partial r} - \frac{\Phi_1}{r} = \frac{1}{\sqrt{2}\,r}(\eth + \cot\theta)\Phi_2, \tag{1.15d}$$

where $\eth \equiv \partial/\partial\theta + (i/\sin\theta)\,\partial/\partial\phi$ and $\bar{\eth}$ is its complex conjugate.

If Φ_0 is given on one null surface $u = u_0$ as an arbitrary function of r, θ, and ϕ, the first two equations can be integrated immediately, with r as the independent variable (θ and ϕ are to be considered as parameters), yielding

$$\begin{aligned} \Phi_0 &= \Phi_0(r, \theta, \phi), \\ \Phi_1 &= \frac{\Phi_1^0(\theta,\phi)}{r^2} + \frac{1}{r^2}\int \frac{r}{\sqrt{2}}(\bar{\eth} + \cot\theta)\Phi_0\, dr, \\ \Phi_2 &= \frac{\Phi_2^0(\theta,\phi)}{r} + \frac{1}{r}\int \frac{\bar{\eth}\Phi_1}{\sqrt{2}}\, dr, \end{aligned} \tag{1.16}$$

where Φ_1^0 and Φ_2^0 are "constants" of integration.

[Assuming[9] that $\Phi_0 = O(1/r^3)$, and $\bar{\eth}\Phi_0$, $\bar{\eth}^2\Phi_0 = O(1/r^3)$, it is easily seen that $\Phi_1 = O(1/r^2)$ and $\Phi_2 = O(1/r)$. This result is a statement of the well known "peeling off" theorem in electromagnetic theory; namely, in the radiation zone where the fields go as $1/r$, i.e., where we neglect $O(1/r^2)$, the field is null and l^μ coincides with the degenerate or double principal null vector; in the intermediate zone where we neglect $O(1/r^3)$, l^μ is one of the two principal null vectors, and finally where we keep $O(1/r^3)$, l^μ coincides with neither of the principal null vectors.]

If the expressions in Eqs. (1.16) are placed into the two remaining Maxwell equations, the u dependence of Φ_0 and Φ_1^0 is determined. There is no equation which governs the u dependence of Φ_2^0 and

[9] These assumptions hold true in the case of retarded solutions or incoming fields of finite time duration.

hence it can be given as an arbitrary function of u as well as of θ and ϕ.

To summarize, an arbitrary but unique (up to the choice of the original Lorentz frame) solution of Maxwell's equations is obtained by giving $\Phi_0(r, \theta, \phi)$, $\Phi_2^0(u, \theta, \phi)$, and $\Phi_1^0(\theta, \phi)$ as arbitrary functions of their stated arguments. $\Phi_2^0(u, \theta, \phi)$ will be called the news function in that its u dependence governs the u dependence of Φ_0 and Φ_1^0. Physically it is the information sent by a "broadcasting" station. Φ_1^0, for reasons given later, will be called the charge aspect.

At this point we make a special choice of Φ_0, namely

$$\Phi_0 = \sum \frac{\Phi_0^{n-1}}{r^{2+n}}, \qquad n \geq 1, \tag{1.17}$$

which corresponds to retarded multipole solutions.[10] In addition, solely for reasons of simplicity, we will assume axial symmetry; i.e., all functions will be independent of the angular coordinate ϕ. The operator $\mathfrak{D}$ becomes $\partial/\partial\theta$. (If this assumption is dropped, results similar to the ones presented in the remainder of this section and the next can be obtained. However, the added complexity just obscures these results.)

From (1.17) and (1.16) we obtain

$$\begin{aligned} \Phi_1 &= \frac{\Phi_1^0}{r^2} - \frac{1}{\sqrt{2}}\left(\frac{\partial}{\partial\theta} + \cot\theta\right)\sum_{n=1}^{\infty}\frac{\Phi_0^{n-1}}{nr^{2+n}}, \\ \Phi_2 &= \frac{\Phi_2^0}{r} - \frac{1}{\sqrt{2}\,r^2}\frac{\partial\Phi_1^0}{\partial\theta} \\ &\quad + \frac{1}{2}\frac{\partial}{\partial\theta}\left(\frac{\partial}{\partial\theta} + \cot\theta\right)\sum_{1}^{\infty}\frac{\Phi_0^{n-1}}{n(n+1)r^{2+n}}. \end{aligned} \tag{1.18}$$

When the solutions, Eqs. (1.18), are substituted into the last two of the Maxwell equations (1.15c, d), the "time" dependence of Φ_1^0 and Φ_0 is obtained:

$$\dot{\Phi}_1^0 = \frac{1}{\sqrt{2}}\left(\frac{\partial}{\partial\theta} + \cot\theta\right)\Phi_2^0, \tag{1.19a}$$

$$\dot{\Phi}_0^0 = \frac{1}{\sqrt{2}}\frac{\partial\Phi_1^0}{\partial\theta}, \tag{1.19b}$$

$$\dot{\Phi}_0^n = -\frac{(n+1)}{2}\Phi_0^{n-1} - \frac{1}{2n}\left\{\frac{\partial^2\Phi_0^{n-1}}{\partial\theta^2} + \cot\theta\frac{\partial\Phi_0^{n-1}}{\partial\theta} - \frac{1}{\sin^2\theta}\Phi_0^{n-1}\right\}, \qquad (n \geq 1) \tag{1.19c}$$

where the dot signifies the derivative with respect to u. It is easily seen that if Φ_2^0, the "news," is given as an arbitrary function of u and θ, then Φ_1^0, Φ_0^0, and Φ_0^n are determined (up to their initial value, which is specified freely) and hence the entire solution is known.

If we integrate Eq. (1.19a) over the surface of a sphere, i.e., multiply by $2\pi \sin\theta d\theta$ and integrate, we obtain

$$\frac{d}{du}\left\{2\pi\int_0^\pi \Phi_1^0 \sin\theta\, d\theta\right\} = \frac{2\pi}{\sqrt{2}}\int_0^\pi \frac{\partial(\sin\theta\Phi_2^0)}{\partial\theta}\,d\theta = 0. \tag{1.20}$$

This is nothing but the law of conservation of charge with $Q \equiv 2\pi\int_0^\pi \Phi_1^0 \sin\theta\, d\theta$ being proportional to the charge. It is for this reason that Φ_1^0 is called the charge aspect. (Bondi[4] called the analogous quantity in the gravitational theory the mass aspect.) It should be noted that in the derivation of (1.20) it was implicitly assumed that Φ_2^0 was not singular at $\theta = 0$ and π. This assumption is now generalized to state that neither Φ_0, Φ_1, nor Φ_2 shall have an angular singularity; i.e., there shall be no singular functions of θ. Singular functions would correspond to sources which are not isolated but extend to spatial infinity.

D. The Meaning of the Data

It was shown in the previous section that if Φ_0^n and Φ_1^0 are given at one value of u and Φ_2^0 given for all values of u, then the complete u dependence of Φ_0^n and Φ_1^0 is known. It is of interest to study the inverse problem; namely, if the u dependence of Φ_0^n and Φ_1^0 is known, what "news," Φ_0^2, gave rise to that u dependence. In particular, if all the Φ_0^l, $l > N$, are zero, what "news" yielded that result? (These latter solutions will correspond to finite multipole expansions.)

It is easily seen from Eq. (1.19c) that if all Φ_0^l, $l > N$, are initially zero and $\dot{\Phi}_0^{N+1} = 0$, then they remain zero. The latter condition then implies

$$\frac{\partial^2\Phi_0^N}{\partial\theta^2} + \cot\theta\frac{\partial\Phi_0^N}{\partial\theta} + \left[(N+1)(N+2) - \frac{1}{\sin^2\theta}\right]\Phi_0^N = 0. \tag{1.21}$$

The solution is the associated Legendre function

$$\Phi_0^N = a_{N+1}(u)P_{N+1}^1(\cos\theta), \tag{1.22}$$

where $a_{N+1}(u)$ is an arbitrary (complex) function of u. (The subscript $N+1$ is used for notational reasons.) Knowing the angular dependence of Φ_0^N, it is possible to determine, by Eq. (1.19c), Φ_0^{N-1}.

[10] The choice of functions other than a Taylor series in inverse powers of r corresponds to incoming radiation fields.

The result,

$$\Phi_0^{N-1} = A\dot{a}_{N+1}P_{N+1}^1, \qquad A = N/(N+1), \tag{1.23}$$

excludes the solution of the homogeneous equation, which would have introduced a new arbitrary function of u, with angular dependence P_N^1. This new "mode" of the field could be obtained by replacing N by $N-1$ in the previous expressions and hence need not be considered separately.

By repeated applications of (1.19c), one can obtain all the Φ_0^n down to Φ_0^0, namely

$$\Phi_0^{N-m} = A_m(d^m/du^m)a_{N+1}P_{N+1}^1(\cos\theta), \tag{1.24}$$

and in particular

$$\Phi_0^0 = A_N(d^N a_{N+1}/du^N)P_{N+1}^1(\cos\theta), \tag{1.25}$$

A_m being a number depending on N and m. Substituting (1.25) into Eq. (1.19b), we obtain

$$\Phi_1^0 = -\sqrt{2}\,A_N(d^{N+1}a_{N+1}/du^{N+1})P_{N+1}(\cos\theta). \tag{1.26}$$

Using Φ_1^0 in (1.19a), we finally obtain the "news,"

$$\Phi_2^0 = B_N(d^{N+2}a_{N+1}/du^{N+2})P_{N+1}^1(\cos\theta), \tag{1.27}$$

where B_N is also a numeric. Due to the linearity of the theory, it is possible to use for the "news" linear combinations of terms of the type (1.27), i.e.,

$$\Phi_2^0 = \sum_{l=0}^{N} B_l \frac{d^{l+2}a_{l+1}}{du^{l+2}} P_{l+1}^1(\cos\theta). \tag{1.28}$$

One case which remains to be analyzed is that in which the entire Φ_0 is initially zero and remains so. From (1.19a, b) it is easily seen that as a consequence, we must have

$$\begin{aligned} \Phi_1^0 &= a_0(u), \\ \Phi_2^0 &= -\sqrt{2}\cot\theta\dot{a}_0(u). \end{aligned} \tag{1.29}$$

Because of the $\cot\theta$ and the previous assumption of no angular singularities in the solution, we must have $\dot{a}_0 = 0$, which is another statement of the law of conservation of charge.[11] This is also a statement of the fact that there exists no monopole radiation; i.e., that there can be no monopole "news."

Let us now return to the examination of the solution generated by the news, Φ_2^0, from Eq. (1.27). If we compare this solution with that obtained from an arbitrary (axially symmetric) outgoing multipole solution of the Maxwell equations, we find that the two are identical if a multiple of the $a_N(u)$ is identified with the multipole moment. A real a_N corresponds to an electric-type pole and an imaginary one to a magnetic-type pole. a_0 is proportional to the monopole moment (charge), a_1 to the dipole moment, etc.

In analogy with the definition of Φ_1^0 as charge aspect, it is reasonable to refer to Φ_0^0 as the dipole aspect, Φ_0^1 as the quadrupole aspect, etc.

To conclude this section we express three different solutions in terms of the Φ's:

(a) Monopole or Coulomb

$$\begin{aligned} \Phi_0 &= 0, \\ \Phi_1 &= \frac{a_0}{r^2}, \\ \Phi_2 &= 0. \end{aligned} \tag{1.30}$$

(b) Dipole

$$\begin{aligned} \Phi_0 &= a_1(u)\sin\theta/r^3, \\ \Phi_1 &= -\sqrt{2}\,\dot{a}_1\cos\theta/r^2 - \sqrt{2}\,a_1\cos\theta/r^3, \\ \Phi_2 &= -\frac{\ddot{a}_1\sin\theta}{r} - \frac{\dot{a}_1\sin\theta}{r^2} - \frac{a_1\sin\theta}{2r^3}. \end{aligned} \tag{1.31}$$

(c) Quadrupole

$$\begin{aligned} \Phi_0 &= \frac{\dot{a}_2\sin\theta\cos\theta}{2r^3} + \frac{a_2(u)\sin\theta\cos\theta}{r^4}, \\ \Phi_1 &= -\frac{\ddot{a}_2(3\cos^2\theta - 1)}{6\sqrt{2}\,r^2} - \frac{\dot{a}_2(3\cos^2\theta - 1)}{2\sqrt{2}\,r^3} \\ &\quad - \frac{a_2(3\cos^2\theta - 1)}{2\sqrt{2}\,r^4}, \\ \Phi_2 &= -\frac{\overset{(3)}{a}_2\sin\theta\cos\theta}{6r} - \frac{\ddot{a}_2\sin\theta\cos\theta}{2r^2} \\ &\quad - \frac{3\dot{a}_2\sin\theta\cos\theta}{4r^3} - \frac{a_2\sin\theta\cos\theta}{2r^4}. \end{aligned} \tag{1.32}$$

The a_0, a_1, and a_2 are proportional respectively to the charge, dipole moment (electric and magnetic), and quadrupole moment (electric and magnetic).

II. GENERAL RELATIVITY

A. Field Equations

In this section we give a brief summary of the spin-coefficient formalism for general relativity [see (NP) and (NU) for details]. We begin by introducing four linearly independent null vectors, l^μ, n^μ, m^μ, and $\bar{m}^\mu$, in the same way as was done for Maxwell theory in Sec. IA; now, however, space–time is not necessarily flat. The ten independent components of the Weyl tensor are then described by five complex tetrad components as follows:

[11] An imaginary a_0 corresponds to a magnetic charge or monopole. Though there is nothing in the analysis presented here to exclude this possibility, it is excluded by the requirement that div **B** vanish *everywhere*. See also Sec. IID.

$$\begin{aligned}
\Psi_0 &= -C_{\mu\nu\rho\sigma}l^\mu m^\nu l^\rho m^\sigma,\\
\Psi_1 &= -C_{\mu\nu\rho\sigma}l^\mu n^\nu l^\rho m^\sigma,\\
\Psi_2 &= -C_{\mu\nu\rho\sigma}\bar{m}^\mu n^\nu l^\rho m^\sigma,\\
\Psi_3 &= -C_{\mu\nu\rho\sigma}\bar{m}^\mu n^\nu l^\rho n^\sigma,\\
\Psi_4 &= -C_{\mu\nu\rho\sigma}\bar{m}^\mu n^\nu \bar{m}^\rho n^\sigma.
\end{aligned} \tag{2.1}$$

Before writing down the field equations, we shall introduce some restrictions on the choice of coordinates and tetrad vectors, analogous to the choices made for the Maxwell case in Sec. IC. We first choose a family of null hypersurfaces, designated by a parameter u = const. The first tetrad vector is taken to be $l_\mu = u_{,\mu}$. If we then take u to be the coordinate x^0, l_μ takes the simple form $l_\mu = \delta^0_\mu$. The l^μ are tangent to a family of null geodesics lying within the hypersurfaces, and we choose an affine parameter r along these geodesics to be the coordinate x^1. The remaining two coordinates, x^2 and x^3, will be angular coordinates, whose choice singles out a particular one of the null geodesics on each of the hypersurfaces u = const. It should be pointed out that this choice of coordinates is well defined only in some finite coordinate patch; for example, if the coordinate system in an asymptotically flat space is chosen asymptotically (i.e., for large r) to approach the flat-space coordinate system described in Sec. IC, then as the null hypersurfaces are followed inward toward smaller values of r, caustics develop (i.e., the null surfaces cross) and the coordinate values are no longer uniquely defined. Thus the formalism that embodies such a coordinate system can only be used outside a world tube that contains all the caustics.

The second tetrad vector, n^μ, is chosen to be null, and normalized according to the requirement $l_\mu n^\mu = 1$; it thus lies outside the hypersurface u = const. The vectors m^μ and $\bar{m}^\mu$ will then be defined by the conditions $m_\mu m^\mu = \bar{m}_\mu \bar{m}^\mu = m_\mu l^\mu = m_\mu n^\mu = m_\mu \bar{m}^\mu + 1 = 0$, as in Sec. IA. In terms of the coordinate system that has been chosen, the tetrad vectors may be written in terms of three arbitrary real functions, U and X^i, and three arbitrary complex functions, ω and ξ^i, where $i = 2, 3$, as follows:

$$\begin{aligned}
l^\mu &= \delta^\mu_1,\\
n^\mu &= \delta^\mu_0 + U\delta^\mu_1 + X^i\delta^\mu_i,\\
m^\mu &= \omega\delta^\mu_1 + \xi^i\delta^\mu_i,
\end{aligned} \tag{2.2}$$

and $\bar{m}^\mu$ is the complex conjugate of m^μ. We make the further specification that n^μ and m^μ are to be parallelly propagated in the direction of l^μ (i.e., along the null geodesics used in the definition of the coordinate system).

The content of the empty-space Einstein equations can now be expressed by a set of field equations written in terms of the notation of Eqs. (1.6). Although these equations are large in number, they are in many respects easier to work with than the original equations. We note first of all that our choice of coordinate system and tetrad makes the derivatives (1.6a) take the form

$$\begin{aligned}
D &= \partial/\partial r, \qquad \delta = \omega\partial/\partial r + \xi^i\partial/\partial x^i,\\
\Delta &= U\partial/\partial r + \partial/\partial u + X^i\partial/\partial x^i,
\end{aligned} \tag{2.3a}$$

and certain of the spin coefficients (1.6b) satisfy the relations [see (NP), Sec. IV]

$$\kappa = \epsilon = \pi = 0, \qquad \rho = \bar{\rho}, \qquad \tau = \bar{\alpha} + \beta. \tag{2.3b}$$

The field equations may be divided into three groups. The first group consists of equations containing the radial derivative D:

$$D\xi^i = \rho\xi^i + \sigma\bar{\xi}^i, \tag{2.4a}$$

$$D\omega = \rho\omega + \sigma\bar{\omega} - \tau, \tag{2.4b}$$

$$DX^i = \tau\bar{\xi}^i + \bar{\tau}\xi^i, \tag{2.4c}$$

$$DU = \tau\bar{\omega} + \bar{\tau}\omega - (\gamma + \bar{\gamma}), \tag{2.4d}$$

$$D\rho = \rho^2 + \sigma\bar{\sigma}, \tag{2.4e}$$

$$D\sigma = 2\rho\sigma + \Psi_0, \tag{2.4f}$$

$$D\alpha = \alpha\rho + \beta\bar{\sigma}, \tag{2.4g}$$

$$D\beta = \beta\rho + \alpha\sigma + \Psi_1, \tag{2.4h}$$

$$D\gamma = \tau\alpha + \bar{\tau}\beta + \Psi_2, \tag{2.4i}$$

$$D\lambda = \lambda\rho + \mu\bar{\sigma}, \tag{2.4j}$$

$$D\mu = \mu\rho + \lambda\sigma + \Psi_2, \tag{2.4k}$$

$$D\nu = \tau\lambda + \bar{\tau}\mu + \Psi_3, \tag{2.4l}$$

$$D\Psi_1 - \bar{\delta}\Psi_0 = 4\rho\Psi_1 - 4\alpha\Psi_0, \tag{2.4m}$$

$$D\Psi_2 - \bar{\delta}\Psi_1 = 3\rho\Psi_2 - 2\alpha\Psi_1 - \lambda\Psi_0, \tag{2.4n}$$

$$D\Psi_3 - \bar{\delta}\Psi_2 = 2\rho\Psi_3 - 2\lambda\Psi_1, \tag{2.4o}$$

$$D\Psi_4 - \bar{\delta}\Psi_3 = \rho\Psi_4 + 2\alpha\Psi_3 - 3\lambda\Psi_2. \tag{2.4p}$$

The second group of field equations consists of those equations that involve u-derivatives of the Ψ's; that is, those equations that describe the propagation of the Weyl tensor from one null surface to another:

$$\begin{aligned}
&\Delta\Psi_0 - \delta\Psi_1\\
&\quad = (4\gamma - \mu)\Psi_0 - 2(2\tau + \beta)\Psi_1 + 3\sigma\Psi_2,
\end{aligned} \tag{2.5a}$$

$$\Delta\Psi_1 - \delta\Psi_2 = \nu\Psi_0 + 2(\gamma - \mu)\Psi_1 - 3\tau\Psi_2 + 2\sigma\Psi_3, \quad (2.5b)$$

$$\Delta\Psi_2 - \delta\Psi_3 = 2\nu\Psi_1 - 3\mu\Psi_2 - 2\bar{\alpha}\Psi_3 + \sigma\Psi_4, \quad (2.5c)$$

$$\Delta\Psi_3 - \delta\Psi_4 = 3\nu\Psi_2 - 2(\gamma + 2\mu)\Psi_3 - (\tau - 4\beta)\Psi_4. \quad (2.5d)$$

The final group consists of field equations that involve the derivatives Δ and δ, but no derivatives of the Ψ's:

$$\delta X^i - \Delta\xi^i = (\mu + \bar{\gamma} - \gamma)\xi^i + \bar{\lambda}\bar{\xi}^i, \quad (2.6a)$$

$$\delta\bar{\xi}^i - \bar{\delta}\xi^i = (\bar{\beta} - \alpha)\xi^i + (\bar{\alpha} - \beta)\bar{\xi}^i, \quad (2.6b)$$

$$\delta\tilde{\omega} - \bar{\delta}\omega = (\bar{\beta} - \alpha)\omega + (\bar{\alpha} - \beta)\tilde{\omega} + \mu - \bar{\mu}, \quad (2.6c)$$

$$\delta U - \Delta\omega = (\mu + \bar{\gamma} - \gamma)\omega + \bar{\lambda}\tilde{\omega} - \bar{\nu}, \quad (2.6d)$$

$$\Delta\lambda - \bar{\delta}\nu = 2\alpha\nu + (\bar{\gamma} - 3\gamma - \mu - \bar{\mu})\lambda - \Psi_4, \quad (2.6e)$$

$$\delta\rho - \bar{\delta}\sigma = \tau\rho + (\bar{\beta} - 3\alpha)\sigma - \Psi_1, \quad (2.6f)$$

$$\delta\alpha - \bar{\delta}\beta = \mu\rho - \lambda\sigma - 2\alpha\beta + \alpha\bar{\alpha} + \beta\bar{\beta} - \Psi_2, \quad (2.6g)$$

$$\delta\lambda - \bar{\delta}\mu = \bar{\tau}\mu + (\bar{\alpha} - 3\beta)\lambda - \Psi_3, \quad (2.6h)$$

$$\delta\nu - \Delta\mu = \gamma\mu - 2\nu\beta + \bar{\gamma}\mu + \mu^2 + \lambda\bar{\lambda}, \quad (2.6i)$$

$$\delta\gamma - \Delta\beta = \tau\mu - \sigma\nu + (\mu - \gamma + \bar{\gamma})\beta + \bar{\lambda}\alpha, \quad (2.6j)$$

$$\delta\tau - \Delta\sigma = 2\tau\beta + (\bar{\gamma} + \mu - 3\gamma)\sigma + \bar{\lambda}\rho, \quad (2.6k)$$

$$\Delta\rho - \bar{\delta}\tau = (\gamma + \bar{\gamma} - \bar{\mu})\rho - 2\alpha\tau - \lambda\sigma - \Psi_2, \quad (2.6l)$$

$$\Delta\alpha - \bar{\delta}\gamma = \rho\nu - \tau\lambda - \lambda\beta + (\bar{\gamma} - \gamma - \bar{\mu})\alpha - \Psi_3. \quad (2.6m)$$

B. Tetrad Rotation and the Algebra of the Ψ's

The tetrad transformations (1.7), (1.8), and (1.9) express the freedom allowed in the choice of a tetrad subject only to the "orthonormality" restrictions. We have imposed two additional restrictions on the tetrad described in Sec. IIA, each of which reduces the group of allowed transformations. The first restriction is that the direction of l^μ has been geometrically determined, which eliminates the transformations (1.9). The other restriction is that n^μ and m^μ are to be parallelly propagated along l^μ, which makes the parameter a of the transformations (1.7) independent of r. In this subsection, however, we wish to examine the general question of the algebra of the Ψ's before the tetrad is tied down by such restrictions, and hence we shall again consider the full transformations (1.7), (1.8), and (1.9), which we shall again denote by (a), (b), and (c), respectively. A straightforward calculation shows that under these transformations, the Ψ's transform in the following fashion:

(a)
$$\begin{aligned}
\tilde{\Psi}_0 &= \Psi_0, \\
\tilde{\Psi}_1 &= \Psi_1 + \bar{a}\Psi_0, \\
\tilde{\Psi}_2 &= \Psi_2 + 2\bar{a}\Psi_1 + \bar{a}^2\Psi_0, \\
\tilde{\Psi}_3 &= \Psi_3 + 3\bar{a}\Psi_2 + 3\bar{a}^2\Psi_1 + \bar{a}^3\Psi_0, \\
\tilde{\Psi}_4 &= \Psi_4 + 4\bar{a}\Psi_3 + 6\bar{a}^2\Psi_2 + 4\bar{a}^3\Psi_1 + \bar{a}^4\Psi_0.
\end{aligned} \quad (2.7)$$

(b)
$$\begin{aligned}
\tilde{\Psi}_0 &= \lambda^2 e^{2i\phi}\Psi_0, \\
\tilde{\Psi}_1 &= \lambda e^{i\phi}\Psi_1, \\
\tilde{\Psi}_2 &= \Psi_2, \\
\tilde{\Psi}_3 &= \lambda^{-1} e^{-i\phi}\Psi_3, \\
\tilde{\Psi}_4 &= \lambda^{-2} e^{-2i\phi}\Psi_4.
\end{aligned} \quad (2.8)$$

(c)
$$\begin{aligned}
\tilde{\Psi}_0 &= \Psi_0 + 4b\Psi_1 + 6b^2\Psi_2 + 4b^3\Psi_3 + b^4\Psi_4, \\
\tilde{\Psi}_1 &= \Psi_1 + 3b\Psi_2 + 3b^2\Psi_3 + b^3\Psi_4, \\
\tilde{\Psi}_2 &= \Psi_2 + 2b\Psi_3 + b^2\Psi_4, \\
\tilde{\Psi}_3 &= \Psi_3 + b\Psi_4, \\
\tilde{\Psi}_4 &= \Psi_4.
\end{aligned} \quad (2.9)$$

Suppose that we have chosen a tetrad such that neither Ψ_0 nor Ψ_4 vanishes. Such a choice is always possible, since if either or both of Ψ_0 and Ψ_4 did vanish, we could make them both nonzero by means of transformations (a) and (c). (Here and in the following discussion we assume that not all of the Ψ's vanish; i.e., the space-times being considered are not flat.) If we now wish to make $\tilde{\Psi}_0 = 0$, we see that we must use transformation (c), which rotates l^μ into $\bar{l}^\mu$. Each such null rotation is specified by a parameter b, which is a root of the quartic equation

$$\Psi_0 + 4b\Psi_1 + 6b^2\Psi_2 + 4b^3\Psi_3 + b^4\Psi_4 = 0 \quad (2.10)$$

obtained from the first of Eqs. (2.9). Each $\bar{l}^\mu$ corresponding to a root of Eq. (2.10) is said to be a principal null vector of the Weyl tensor. Since $\Psi_4 \neq 0$, there are in general four principal null vectors. If, however, any of the b's satisfying Eq. (2.10) is a multiple root of the equation, then there are fewer than four distinct principal null vectors, and the Weyl tensor is said to be algebraically special. The algebraically general Weyl tensor, corresponding to the case of four distinct roots, is said to be of Type I. If there are two single roots of Eq. (2.10) and one double root, the Weyl tensor is said to be of Type II. If b is chosen to be the double root of Eq. (2.10), then $\tilde{\Psi}_1$ also vanishes.

This may be seen by noting from Eqs. (2.9) that $\tilde{\Psi}_1$ is proportional to the derivative of $\tilde{\Psi}_0$ with respect to b. Thus if b is a double root of the equation $\tilde{\Psi}_0 = 0$, then $d\tilde{\Psi}_0/db$ also vanishes for the same value of b. Consequently, b is also a root of the equation $\tilde{\Psi}_1 = 0$.

By a similar argument it may be seen that if b is a triply degenerate root the first three $\tilde{\Psi}$'s vanish, and if b is quadruply degenerate only $\tilde{\Psi}_4$ is nonzero. In the latter two cases, the Weyl tensor is said to be of Types III and N, respectively. The remaining possibility is that Eq. (2.10) has two double roots, in which case the Weyl tensor is said to be of Type D.

It follows from the definitions (2.1) of the Ψ's that under the interchanges $l^\mu \leftrightarrow n^\mu$ and $m^\mu \leftrightarrow \bar{m}^\mu$, the Ψ's undergo the interchanges $\Psi_0 \leftrightarrow \Psi_4$ and $\Psi_1 \leftrightarrow \Psi_3$, while Ψ_2 is unchanged. This interchange of tetrad vectors also interchanges the roles of transformations (a) and (c). Consequently, if $\tilde{n}^\mu$ is made a principal null vector of the Weyl tensor by use of transformation (a), then $\tilde{\Psi}_4$ will vanish, with analogous results for all of the various cases of algebraically special Weyl tensors. In particular, in the case of a Type-D Weyl tensor, let us choose $\tilde{l}^\mu$ and $\tilde{n}^\mu$ to be the two distinct principal null vectors. Then $\tilde{\Psi}_2$ is the only nonvanishing $\tilde{\Psi}$, and the only remaining tetrad freedom is transformation (b), which leaves $\tilde{\Psi}_2$ unchanged. It is thus possible to specify invariantly defined tetrad components of a Type-D Weyl tensor, in close analogy to the specification of invariantly defined tetrad components of a nonnull Maxwell tensor in Sec. IB.

C. Null-Surface Data for the Einstein Equations

A discussion of the initial-value problem for the gravitational field equations (2.4), (2.5), and (2.6), where an initial surface is one of the null surfaces $u = \text{const.}$, has been given in (NU). That discussion is closely analogous to the one given for the Maxwell field in Sec. IC of the present paper. In this subsection, we shall briefly review the results of (NU); then in the following subsection, we shall discuss the meaning that we propose to give to the initial data in terms of the multipole structure of the sources.

The assumption is made in (NU) that Ψ_0 is of the form

$$\Psi_0 = \Psi_0^0 r^{-5} + O(r^{-6}), \tag{2.11}$$

which seems to correspond to the exclusion of incoming radiation of infinite duration. In addition, the coordinate system is subject to certain restrictions beyond those mentioned in Sec. IIA, the details of which need not concern us here. It then follows that a unique solution to the field equations (2.4), (2.5), and (2.6) is determined by five pieces of initial data. The first of these is Ψ_0 specified on an initial null surface $u = \text{const.}$ The second piece of data is specified on a timelike world-tube which may be thought of as being at spatial infinity; this piece of data is $\sigma^0 \equiv \lim_{r\to\infty} r^2\sigma$, where σ is defined in (1.6b). Whereas these two pieces of data are functions of three variables, the remaining data are functions only of two variables, and may be thought of as being specified on the two-dimensional intersection of the null surface on which Ψ_0 is given and the world-tube on which σ^0 is given. Two of these pieces of data are $\Psi_1^0 \equiv \lim_{r\to\infty} r^4\Psi_1$ and $\Psi_2^0 + \bar{\Psi}_2^0 \equiv \lim_{r\to\infty} r^3(\Psi_2 + \bar{\Psi}_2)$; the remaining piece of data is related to the metric of this two-dimensional surface, and depends on the specific coordinate conditions that have been imposed. It is only the first four pieces of data with which we shall be concerned, and to which we shall give a physical meaning. The quantity σ^0 is called the news function, as it is the u-dependent piece of initial data (Bondi[4] has called the u-derivative of σ^0 the news function), and $\Psi_2^0 + \bar{\Psi}_2^0$ is related to Bondi's mass aspect.[4] The significance of the mass aspect and analogous quantities for multipole moments will be discussed in the following subsection.

D. The Meaning of the Data

We turn now to the linearized theory of gravitation, in which it is possible to make an unambiguous interpretation of solutions in terms of properties of the sources.[12] In the appendix, we discuss the time-independent, axially symmetric, linearized solutions corresponding to multipoles of arbitrary order. Here we shall be concerned with the question of how multipole structure is related to the initial data in the time-dependent case, though for the sake of simplicity we shall keep the restriction of axial symmetry. The general features of the results that we obtain in the linearized theory provide the basis for definitions of multipole moments in the full theory.

Our treatment will be closely analogous to the treatment of Maxwell theory in Sec. I. We begin by assuming Ψ_0 to be of the form

$$\Psi_0 = \sum_{n=2}^{\infty} \frac{\Psi_0^{n-2}}{r^{n+3}}, \tag{2.12}$$

where Ψ_0^{n-2} is a function of u and θ. The form (2.12)

[12] R. K. Sachs and P. G. Bergmann, Phys. Rev. **112**, 674 (1958).

is consistent with the assumption (2.11) as well as with the time-independent multipole expansion discussed in the appendix. We next assume that the sum in (2.12) stops at some finite $n = N$, and see what the initial data must be in order for that assumption to hold for all values of u.

We thus assume that Ψ_0 is given by

$$\Psi_0 = \sum_{n=2}^{N} \frac{\Psi_0^{n-2}}{r^{n+3}}, \tag{2.13}$$

and substitute this expression into Eq. (2.5a), which describes the propagation of Ψ_0 as a function of u. As an illustration of the procedure to be followed, we shall outline the treatment of Eq. (2.5a) in somewhat more detail than we shall give for the succeeding equations. We note first of all that Ψ_1 enters Eq. (2.5a) (Ψ_2 drops out as a result of the linearization). The r-dependence of Ψ_1 may be obtained by integrating Eq. (2.4m), where we can use the flat-space values of ρ, α, ω, and ξ^i since all of the equations are to be linearized. This integration yields a Ψ_1 of the form

$$\Psi_1 = \frac{\Psi_1^0}{r^4} + \frac{1}{r^4}\int r^3\left(\frac{1}{\sqrt{2}}\frac{\partial \Psi_0}{\partial \theta} + \sqrt{2}\,\Psi_0 \cot\theta\right) dr, \tag{2.14}$$

where Ψ_1^0 is a "constant" of integration (i.e., taking account of axial symmetry, Ψ_1^0 is a function of u and θ only). If we use (2.13) for Ψ_0, we find that Ψ_1 is given by

$$\Psi_1 = \frac{\Psi_1^0}{r^4} - \sum_{n=2}^{N} \frac{1}{\sqrt{2}\,(n-1)r^{n+3}} \times \left(\frac{\partial \Psi_0^{n-2}}{\partial \theta} + 2\Psi_0^{n-2}\cot\theta\right). \tag{2.15}$$

If we now substitute (2.13) and (2.15) into Eq. (2.5a), again using the flat-space values of the new spin coefficients that are introduced, we obtain

$$\sum_{n=1}^{N-1} \frac{\dot{\Psi}_0^{n-1}}{r^{n+4}} = \frac{1}{\sqrt{2}\,r^5}\left(\frac{\partial \Psi_1^0}{\partial \theta} - \Psi_1^0 \cot\theta\right) - \sum_{n=2}^{N} \frac{1}{2(n-1)r^{n+4}}\left[\frac{\partial^2 \Psi_0^{n+2}}{\partial \theta^2} + \frac{\partial \Psi_0^{n-2}}{\partial \theta}\cot\theta + n(n+1)\Psi_0^{n-2} - 4\Psi_0^{n-2}\csc^2\theta\right], \tag{2.16}$$

where the dot signifies the derivative with respect to u. Coefficients of like powers of r on the two sides of Eq. (2.16) must be equal. From the coefficients of $r^{-(N+4)}$, we obtain the equation

$$\frac{\partial^2 \Psi_0^{N-2}}{\partial \theta^2} + \frac{\partial \Psi_0^{N-2}}{\partial \theta}\cot\theta + N(N+1)\Psi_0^{N-2} - 4\Psi_0^{N-2}\csc^2\theta = 0. \tag{2.17}$$

The only nonsingular solution of this equation is

$$\Psi_0^{N-2} = a_N P_N^2(\cos\theta), \tag{2.18}$$

where P_N^2 is the Nth second associated Legendre function, and a_N is an arbitrary function of u. Thus, by comparison with the first of Eqs. (A4) of the appendix, this term has the form of a 2^N-pole with a time-dependent moment. *We shall say that by definition this term represents a 2^N-pole whose moment is proportional to a_N*; a real a_N corresponding to an "electric"-type pole, an imaginary a_N to a "magnetic"-type pole (see appendix). From the coefficients of $r^{-(n+4)}$ for $2 \leq n \leq N-1$, we obtain the set of equations

$$\frac{\partial^2 \Psi_0^{n-2}}{\partial \theta^2} + \frac{\partial \Psi_0^{n-2}}{\partial \theta}\cot\theta + n(n+1)\Psi_0^{n-2} - 4\Psi_0^{n-2}\csc^2\theta = -2(n-1)\dot{\Psi}_0^{n-1}, \qquad 2 \leq n \leq N-1. \tag{2.19}$$

Having the solution (2.18) to Eq. (2.17), we can integrate Eq. (2.19) for $n = N-1$; then having the solution to that equation, we can integrate Eq. (2.19) for $n = N-2$; and so on through the whole set. At each step we obtain a new "constant" of integration $a_n(u)$, corresponding to a solution of the homogeneous equation, and thus to a 2^n-pole, just as with Eq. (2.17); the solution to the inhomogeneous equation corresponds to terms that are induced by the time dependence of the higher-order poles. We may write the general (nonsingular) solution to the combined set of equations (2.17) and (2.19) as

$$\Psi_0^{n-2} = \sum_{m=n}^{N} \alpha_m^n \frac{d^{m-n}a_m}{du^{m-n}} P_m^2(\cos\theta), \qquad 2 \leq n \leq N, \tag{2.20a}$$

where

$$\alpha_m^n = \prod_{l=n}^{m-1} \frac{2(l-1)}{m(m+1) - l(l+1)}, \qquad m > n; \quad \alpha_n^n = 1. \tag{2.20b}$$

Finally, from the coefficients of r^{-5} in Eq. (2.16), we obtain

$$\partial \Psi_1^0/\partial\theta - \Psi_1^0 \cot\theta = \sqrt{2}\,\dot{\Psi}_0^0. \tag{2.21}$$

Using the solution (2.20) for Ψ_0^0, we may integrate Eq. (2.21) to obtain (again discarding singular solutions)

$$\Psi_1^0 = -\sqrt{2}\sum_{m=1}^{N} \alpha_m^2 \frac{d^{m-1}a_m}{du^{m-1}} P_m^1(\cos\theta), \qquad \alpha_1^2 = -1/\sqrt{2}, \tag{2.22}$$

where the P_m^1 are first associated Legendre functions. The new "constant" of integration a_1, which comes from the solution of the homogeneous equation, corresponds to a dipole; the other terms are again induced by the time dependence of the higher-order poles.

We now proceed to go through the analogous steps with the remaining u-propagation equations (2.5). The linearized version of Eq. (2.5b) does not involve Ψ_3, and Ψ_2 is obtained by integrating Eq. (2.4n). We find that Ψ_2 is of the form

$$\Psi_2 = \frac{\Psi_2^0}{r^3} + \frac{1}{\sqrt{2}\, r^3} \int r^2 \left(\frac{\partial \Psi_1}{\partial \theta} + \Psi_1 \cot \theta \right) dr, \qquad (2.23)$$

where Ψ_2^0 is a function of u and θ. Eq. (2.5b) then yields the following differential equation for Ψ_2^0 (where we have used the previous results for Ψ_1 and Ψ_0):

$$\frac{\partial \Psi_2^0}{\partial \theta} = -2 \sum_{m=1}^{N} \alpha_m^2 \frac{d^m a_m}{du^m} P_m^1(\cos \theta). \qquad (2.24)$$

The solution to Eq. (2.24) may be written in terms of Legendre polynomials as

$$\Psi_2^0 = 2 \sum_{m=0}^{N} \alpha_m^2 \frac{d^m a_m}{du^m} P_m(\cos \theta), \qquad \alpha_0^2 = \tfrac{1}{2}. \qquad (2.25)$$

The new "constant" of integration a_0 corresponds to a monopole, and the remaining terms are induced by the time dependence of the higher-order poles. Since only the real part of Ψ_2^0 is part of the initial data, the imaginary part of Ψ_2^0 must be expressible in terms of initial data. Using the results of (NU), we may write

$$\Psi_2^0 - \bar{\Psi}_2^0 = \frac{1}{2} \frac{\partial^2 (\sigma^0 - \bar{\sigma}^0)}{\partial \theta^2} + \frac{3}{2} \frac{\partial (\sigma^0 - \bar{\sigma}^0)}{\partial \theta} \cot \theta - (\sigma^0 - \bar{\sigma}^0). \qquad (2.26)$$

By combining Eqs. (2.25) and (2.26), we obtain

$$\frac{\partial^2 (\sigma^0 - \bar{\sigma}^0)}{\partial \theta^2} + 3 \frac{\partial (\sigma^0 - \bar{\sigma}^0)}{\partial \theta} \cot \theta - 2(\sigma^0 - \bar{\sigma}^0) = 4 \sum_{m=0}^{N} \alpha_m^2 \frac{d^m (a_m - \bar{a}_m)}{du^m} P_m(\cos \theta). \qquad (2.27)$$

Before dealing with Eq. (2.27), we shall go on to the further treatment of Eqs. (2.5).

The linearization of Eq. (2.5c) eliminates Ψ_4, and Ψ_3 is obtained by integrating Eq. (2.4o). We find that Ψ_3 is of the form

$$\Psi_3 = \frac{\Psi_3^0}{r^2} + \frac{1}{\sqrt{2}\, r^2} \int r \frac{\partial \Psi_2}{\partial \theta} dr, \qquad (2.28)$$

where Ψ_3^0 is a function of u and θ. Equation (2.5c) then yields the following differential equation for Ψ_3^0:

$$\frac{\partial \Psi_3^0}{\partial \theta} + \Psi_3^0 \cot \theta = 2\sqrt{2} \sum_{m=0}^{N} \alpha_m^2 \frac{d^{m+1} a_m}{du^{m+1}} P_m(\cos \theta). \qquad (2.29)$$

We find that this equation has no nonsingular solution in the presence of a time-dependent monopole moment. We thus take

$$\dot{a}_0 = 0. \qquad (2.30)$$

[The real part of this condition corresponds to the conservation of mass in analogy to the way in which conservation of charge was obtained for the Maxwell field; we shall see later that the imaginary part of (2.30) is empty.] We also find that no new "constant" of integration enters into the nonsingular part of the solution, which is

$$\Psi_3^0 = 2\sqrt{2} \sum_{m=1}^{N} \frac{\alpha_m^2}{m(m+1)} \frac{d^{m+1} a_m}{du^{m+1}} P_m^1(\cos \theta). \qquad (2.31)$$

We may use the results of (NU) to express Ψ_3^0 in terms of initial data as

$$\bar{\Psi}_3^0 = -\frac{1}{\sqrt{2}} \frac{\partial \dot{\sigma}^0}{\partial \theta} - \sqrt{2}\, \dot{\sigma}^0 \cot \theta. \qquad (2.32)$$

(For notational convenience, we have written this equation in terms of $\bar{\Psi}_3^0$ rather than Ψ_3^0.) Combining Eqs. (2.31) and (2.32), we obtain the following condition on σ^0:

$$\frac{\partial \dot{\sigma}^0}{\partial \theta} + 2\dot{\sigma}^0 \cot \theta = -4 \sum_{m=1}^{N} \frac{\alpha_m^2}{m(m+1)} \frac{d^{m+1} \bar{a}_m}{du^{m+1}} P_m^1(\cos \theta). \qquad (2.33)$$

We find that Eq. (2.33) has no nonsingular solution for $\dot{\sigma}^0$ unless we take

$$\ddot{a}_1 = 0, \qquad (2.34)$$

(which corresponds to the absence of gravitational dipole radiation) and again set the new "constant" of integration equal to zero. The solution for $\dot{\sigma}^0$ is then

$$\dot{\sigma}^0 = -4 \sum_{m=2}^{N} \frac{\alpha_m^2}{(m-1)m(m+1)(m+2)} \times \frac{d^{m+1} \bar{a}_m}{du^{m+1}} P_m^2(\cos \theta). \qquad (2.35)$$

For the last of Eqs. (2.5), we obtain Ψ_4 by integrating Eq. (2.4p), with the result

$$\Psi_4 = \frac{\Psi_4^0}{r} + \frac{1}{\sqrt{2}\, r} \int \left(\frac{\partial \Psi_3}{\partial \theta} - \Psi_3 \cot \theta \right) dr. \qquad (2.36)$$

Equation (2.5d) then leads to the differential equation

$$\partial\Psi_4^0/\partial\theta + 2\Psi_4^0 \cot\theta = 4\sum_{m=2}^{N} \frac{\alpha_m^2}{m(m+1)} \frac{d^{m+2}a_m}{du^{m+2}} P_m(\cos\theta), \quad (2.37)$$

where we have made use of the condition (2.34). The nonsingular solution of Eq. (2.37) is

$$\Psi_4^0 = 4\sum_{m=2}^{N} \frac{\alpha_m^2}{(m-1)m(m+1)(m+2)} \times \frac{d^{m+2}a_m}{du^{m+2}} P_m^2(\cos\theta). \quad (2.38)$$

From the results of NU, we may write Ψ_4^0 in terms of initial data as

$$\bar{\Psi}_4^0 = -\partial^2\sigma^0/\partial u^2. \quad (2.39)$$

If we substitute Eq. (2.38) for Ψ_4^0 into Eq. (2.39), we get an expression for $\partial^2\sigma^0/\partial u^2$ that is just the derivative of Eq. (2.35). Thus no new conditions on σ^0 are obtained.

Equation (2.35) is trivially integrated to give

$$\sigma^0 = -4\sum_{m=2}^{N} \frac{\alpha_m^2}{(m-1)m(m+1)(m+2)} \times \frac{d^m\bar{a}_m}{du^m} P_m^2(\cos\theta) + f, \quad (2.40)$$

where f is a function of θ alone. If we substitute the solution (2.40) into the condition (2.27), we obtain the following differential equation for the imaginary part of f, where the prime denotes the derivative with respect to θ:

$$(f'' - \bar{f}'') + 3(f' - \bar{f}')\cot\theta - 2(f - \bar{f}) = -2(a_0 - \bar{a}_0) + 2\sqrt{2}\,[d(a_1 - \bar{a}_1)/du]\cos\theta. \quad (2.41)$$

The nonsingular solution of Eq. (2.41) is

$$f - \bar{f} = (a_0 - \bar{a}_0) - \frac{\sqrt{2}}{3}\frac{d(a_1 - \bar{a}_1)}{du}\cos\theta. \quad (2.42)$$

It will be noted that the right-hand side of Eq. (2.42) is, in accord with Eq. (2.40), not a function of u, due to the previously obtained conditions (2.30) and (2.34).

In order to understand the significance of the right-hand side of Eq. (2.42), let us recall that the monopole moment is proportional to a_0 and the dipole moment to a_1. The real part of a_0 corresponds to an "electric"-type monopole, which is just an ordinary mass monopole, and its constancy (see Eq. [2.30]) is just the mass conservation law. The term $a_0 - \bar{a}_0$ in (2.42) corresponds to a "magnetic"-type monopole, which is known not to exist within the framework of linearized gravitational theory.[12] Before pursuing this point, let us first turn to the case of the dipole. The real part of a_1 corresponds to an ordinary ("electric"-type) mass dipole moment, and the vanishing of $\ddot{a}_1$ (see Eq. [2.34]) implies the conservation of linear momentum. The imaginary part of a_1 is a "magnetic"-type dipole moment, which is proportional to the angular momentum. Thus the term $d(a_1 - \bar{a}_1)/du$ in (2.42) corresponds to a changing angular momentum, so that conservation considerations lead us to expect that this term should also not exist. If we now calculate the metric tensor (in a "Cartesian" coordinate system) associated with either an imaginary a_0 or $\dot{a}_1$, we find that it contains line singularities extending to spatial infinity, corresponding to a space that is not simply connected. (The same type of singularity occurs in the vector potential for a magnetic monopole in the Maxwell field.) We are thus led to conclude that both a_0 and $\dot{a}_1$ are real, and consequently the function f in Eq. (2.40) is real. Since the real f is unrestricted by our considerations on the nature of the sources, it might be expected that f could be made zero by using the remaining freedom[4,7] in the choice of coordinates and tetrad vectors. This expectation is, in fact, found to be true.[13] We are thus led to a news function given by

$$\sigma^0 = -4\sum_{m=2}^{N} \frac{\alpha_m^2}{(m-1)m(m+1)(m+2)} \times \frac{d^m\bar{a}_m}{du^m} P_m^2(\cos\theta). \quad (2.43)$$

Let us now summarize the physical interpretation of the initial data. $\Psi_2^0 + \bar{\Psi}_2^0$ is what Bondi[4] has called the mass aspect. It is clear from the form (2.25) for Ψ_2^0 and the orthogonality properties of the Legendre polynomials that the mass is proportional to $\int_0^\pi (\Psi_2^0 + \bar{\Psi}_2^0)P_0(\cos\theta)\sin\theta\,d\theta = \int_0^\pi (\Psi_2^0 + \bar{\Psi}_2^0)\sin\theta\,d\theta$. Similarly, Ψ_1^0 is the dipole aspect, with the dipole moment proportional to $\int_0^\pi \Psi_1^0 P_1^1(\cos\theta)\sin\theta\,d\theta$; Ψ_0^0 is the quadrupole aspect, with the quadrupole moment proportional to $\int_0^\pi \Psi_0^0 P_2^2(\cos\theta)\sin\theta\,d\theta$; and Ψ_0^{n-2} is the 2^n-pole aspect, with the 2^n-pole moment proportional to $\int_0^\pi \Psi_0^{n-2} P_n^2(\cos\theta)\sin\theta\,d\theta$. The form of the news function, σ^0, whose time dependence is part of the initial data, guarantees that the angular dependence of the rest of the initial data will be independent of which null hypersurface is picked as the initial surface.

We propose the following definition of multipole

[13] T. Unti (private communication).

structure in the full (axially symmetric) Einstein theory of gravitation: When the initial data shows the angular dependence that we have found for a 2^n-pole in the linearized theory, we will say that by definition the source of the field has the structure of a 2^n-pole, with the 2^n-pole moment proportional to

$$\int_0^\pi \Psi_0^{n-2} P_n^2(\cos\theta)\sin\theta\, d\theta$$

for $n \geq 2$, the dipole moment proportional to $\int_0^\pi \Psi_1^0 P_1^1(\cos\theta)\sin\theta\, d\theta$, and the monopole moment (i.e., mass) proportional to $\int_0^\pi (\Psi_2^0 + \bar{\Psi}_2^0)\sin\theta\, d\theta$. An examination of the suitability of these definitions in the full theory is being actively pursued.

We conclude this section by exhibiting the Weyl tensor and news function for four specific examples of linearized multipoles, as well as one example (given without derivation) of a nonaxially symmetric news function:

(a) Monopole

$$\begin{aligned} \Psi_0 &= 0, \\ \Psi_1 &= 0, \\ \Psi_2 &= a_0/r^3, \\ \Psi_3 &= 0, \\ \Psi_4 &= 0, \\ \sigma^0 &= 0, \end{aligned} \tag{2.44}$$

where a_0 is real and constant.

(b) Dipole

$$\begin{aligned} \Psi_0 &= 0, \\ \Psi_1 &= a_1(u)\sin\theta/r^4, \\ \Psi_2 &= -\sqrt{2}\,\dot{a}_1\cos\theta/r^3 - \sqrt{2}\,a_1\cos\theta/r^4, \\ \Psi_3 &= -\dot{a}_1\sin\theta/r^3 - a_1\sin\theta/2r^4, \\ \Psi_4 &= 0, \\ \sigma^0 &= 0, \end{aligned} \tag{2.45}$$

where $\dot{a}_1$ is real and constant.

(c) Quadrupole

$$\Psi_0 = 3a_2\sin^2\theta/r^5,$$

$$\Psi_1 = -\frac{3\sqrt{2}\,\dot{a}_2\sin\theta\cos\theta}{r^4} - \frac{6\sqrt{2}\,a_2\sin\theta\cos\theta}{r^5},$$

$$\Psi_2 = \frac{\ddot{a}_2(3\cos^2\theta - 1)}{r^3} + \frac{3\dot{a}_2(3\cos^2\theta - 1)}{r^4} + \frac{3a_2(3\cos^2\theta - 1)}{r^5}, \tag{2.46}$$

$$\Psi_3 = \frac{\sqrt{2}\overset{(3)}{a}_2\sin\theta\cos\theta}{r^2} + \frac{3\sqrt{2}\,\ddot{a}_2\sin\theta\cos\theta}{r^3} + \frac{9\dot{a}_2\sin\theta\cos\theta}{\sqrt{2}\,r^4} + \frac{3\sqrt{2}\,a_2\sin\theta\cos\theta}{r^5},$$

$$\Psi_4 = \frac{\overset{(4)}{a}_2\sin^2\theta}{2r} + \frac{\overset{(3)}{a}_2\sin^2\theta}{r^2} + \frac{3\ddot{a}_2\sin^2\theta}{2r^3} + \frac{3\dot{a}_2\sin^2\theta}{2r^4} + \frac{3a_2\sin^2\theta}{4r^5},$$

$$\bar{\sigma}^0 = -\tfrac{1}{2}\ddot{a}_2\sin^2\theta.$$

(d) Octupole

$$\Psi_0 = \frac{5\dot{a}_3\sin^2\theta\cos\theta}{r^5} + \frac{15a_3\sin^2\theta\cos\theta}{r^6},$$

$$\Psi_1 = -\frac{\ddot{a}_3\sin\theta(5\cos^2\theta - 1)}{\sqrt{2}\,r^4} - \frac{5\dot{a}_3\sin\theta(5\cos^2\theta - 1)}{\sqrt{2}\,r^5} - \frac{15a_3\sin\theta(5\cos^2\theta - 1)}{2\sqrt{2}\,r^6},$$

$$\Psi_2 = \frac{\overset{(3)}{a}_3\cos\theta(5\cos^2\theta - 3)}{3r^3} + \frac{2\ddot{a}_3\cos\theta(5\cos^2\theta - 3)}{r^4} + \frac{5\dot{a}_3\cos\theta(5\cos^2\theta - 3)}{r^5} + \frac{5a_3\cos\theta(5\cos^2\theta - 3)}{r^6}, \tag{2.47}$$

$$\Psi_3 = \frac{\overset{(4)}{a}_3\sin\theta(5\cos^2\theta - 1)}{6\sqrt{2}\,r^2} + \frac{\overset{(3)}{a}_3\sin\theta(5\cos^2\theta - 1)}{\sqrt{2}\,r^3} + \frac{3\ddot{a}_3\sin\theta(5\cos^2\theta - 1)}{\sqrt{2}\,r^4} + \frac{5\dot{a}_3\sin\theta(5\cos^2\theta - 1)}{\sqrt{2}\,r^5} + \frac{15a_3\sin\theta(5\cos^2\theta - 1)}{4\sqrt{2}\,r^6},$$

$$\Psi_4 = \frac{\overset{(5)}{a}_3\sin^2\theta\cos\theta}{6r} + \frac{5\overset{(4)}{a}_3\sin^2\theta\cos\theta}{6r^2} + \frac{5\overset{(3)}{a}_3\sin^2\theta\cos\theta}{2r^3} + \frac{5\ddot{a}_3\sin^2\theta\cos\theta}{r^4} + \frac{25\dot{a}_3\sin^2\theta\cos\theta}{4r^5} + \frac{15a_3\sin^2\theta\cos\theta}{4r^6},$$

$$\bar{\sigma}^0 = -\tfrac{1}{6}\overset{(3)}{a}_3\sin^2\theta\cos\theta.$$

(e) Arbitrary Quadrupole News Function

$$\begin{aligned} \sigma^0 = {} & g_2(u)e^{2i\phi}[-\sin^2\theta + 2(1 + \cos\theta)] \\ & + g_1(u)e^{i\phi}\sin\theta(1 + \cos\theta) \\ & + g_0(u)\sin^2\theta + g_{-1}(u)e^{-i\phi}\sin\theta(1 - \cos\theta) \\ & + g_{-2}(u)e^{-2i\phi}[-\sin^2\theta + 2(1 - \cos\theta)], \end{aligned} \tag{2.48}$$

where the five g's are the components of an irreducible quadrupole tensor.

ACKNOWLEDGMENT

The authors wish to acknowledge and thank Professor P. G. Bergmann, Professor J. N. Goldberg, Dr. J. J. Stachel, Mr. R. Torrence, and Mr. T. Unti for valuable discussions and suggestions.

APPENDIX

Sachs and Bergmann[12] have given a complete classification of time-independent linearized solutions of the Einstein field equations in terms of multipole structure of the sources. These poles are of two types, which Sachs and Bergmann refer to as "electric" and "magnetic" types of poles. The "electric" poles are static in character, corresponding to masses at rest, whereas the "magnetic"-type solutions are stationary, corresponding to time-independent mass "currents." In this appendix, we shall find the initial data (we shall, in fact, exhibit all of the Ψ's)corresponding to these solutions, under the simplifying assumption of axial symmetry. (It will be noted in some of the following expressions referring to 2^n-poles that n occurs in the denominator which means that we are excluding monopoles. Although a solution corresponding to a "magnetic" monopole doesn't exist,[12] the "electric" monopole solution is just the familiar Schwarzschild solution. By treating the latter case separately, we find that our final results are in fact valid for this case, even though the earlier equations in this appendix are not.)

In De Donder coordinates, with $x^1 = x$, $x^2 = y$, $x^3 = z$, and $x^4 = t$, the metric tensor for 2^n-poles takes the form

$$\begin{aligned}
g_{11} &= g_{22} = g_{33} = -1 + \alpha_n P_n(\cos\theta)/2r^{n+1},\\
g_{44} &= 1 + \alpha_n P_n(\cos\theta)/2r^{n+1},\\
g_{14} &= \beta_n P_n^1(\cos\theta)\sin\phi/2nr^{n+1},\\
g_{24} &= -\beta_n P_n^1(\cos\theta)\cos\phi/2nr^{n+1},\\
g_{12} &= g_{13} = g_{23} = g_{34} = 0,
\end{aligned} \tag{A1}$$

where α_n and β_n are proportional to the "electric" and "magnetic" 2^n-pole moments, respectively; P_n and P_n^1 are the Legendre and first associated Legendre functions, respectively; and (r, θ, ϕ) are related to (x, y, z) by the usual relations between spherical and Cartesian coordinates.

We must now introduce a tetrad and coordinate system satisfying the conditions of Sec. IIA. An appropriate coordinate system is given by

$$\begin{aligned}
u \equiv x^0 &= t - r - \alpha_n P_n(\cos\theta)/2nr^n,\\
r' \equiv x^1 &= r,\\
\theta' \equiv x^2 &= \theta - \alpha_n P_n^1(\cos\theta)/2n(n+1)r^{n+1},\\
\phi' \equiv x^3 &= \phi - \beta_n P_n^1(\cos\theta)\csc\theta/2n(n+1)r^{n+1}.
\end{aligned} \tag{A2}$$

In this coordinate system (dropping the primes) and in the notation of Eq. (2.2), an appropriate tetrad is specified by

$$\begin{aligned}
U &= -\tfrac{1}{2} - \alpha_n P_n(\cos\theta)/4r^{n+1},\\
X^2 &= -\alpha_n P_n^1(\cos\theta)/4(n+1)r^{n+2},\\
X^3 &= -\beta_n P_n^1(\cos\theta)\csc\theta/4(n+1)r^{n+2},\\
\omega &= (n+2)(\alpha_n + i\beta_n)P_n^1(\cos\theta)/4\sqrt{2}\,n(n+1)r^{n+1},\\
\xi^2 &= 1/\sqrt{2}\,r + (\alpha_n + i\beta_n)P_n^2(\cos\theta)/4\sqrt{2}\,n(n+1)r^{n+2},\\
\xi^3 &= \frac{i\csc\theta}{\sqrt{2}\,r} - \frac{i(\alpha_n + i\beta_n)P_n^2(\cos\theta)\csc\theta}{4\sqrt{2}\,n(n+1)r^{n+2}},
\end{aligned} \tag{A3}$$

where P_n^2 is the second associated Legendre function.

The tetrad components of the Weyl tensor, using the tetrad (A3) and the coordinates (A2), are found to be

$$\begin{aligned}
\Psi_0 &= (\alpha_n + i\beta_n)P_n^2(\cos\theta)/4r^{n+3},\\
\Psi_1 &= -(n+2)(\alpha_n + i\beta_n)P_n^1(\cos\theta)/4\sqrt{2}\,r^{n+3},\\
\Psi_2 &= (n+1)(n+2)(\alpha_n + i\beta_n)P_n(\cos\theta)/8r^{n+3},\\
\Psi_3 &= (n+2)(\alpha_n + i\beta_n)P_n^1(\cos\theta)/8\sqrt{2}\,r^{n+3},\\
\Psi_4 &= (\alpha_n + i\beta_n)P_n^2(\cos\theta)/16r^{n+3}.
\end{aligned} \tag{A4}$$

The news function, σ^0, may also be calculated, and is found to be zero.

Note on the Kerr Spinning-Particle Metric*

E. T. NEWMAN AND A. I. JANIS
Physics Department, University of Pittsburgh, Pittsburgh, Pennsylvania
(Received 19 June 1964)

It is shown that by means of a complex coordinate transformation performed on the monopole or Schwarzschild metric one obtains a new metric (first discovered by Kerr). It has been suggested that this metric be interpreted as that arising from a spinning particle. We wish to suggest a more complicated interpretation, namely that the metric has certain characteristics that correspond to a ring of mass that is rotating about its axis of symmetry. The argument for this interpretation comes from three separate places: (1) the metric appears to have the appropriate multipole structure when analyzed in the manner discussed in the previous paper, (2) in a covariantly defined flat space associated with the metric, the Riemann tensor has a circular singularity, (3) there exists a closely analogous solution of Maxwell's equations that has characteristics of a field due to a rotating ring of charge.

INTRODUCTION

RECENTLY, R. Kerr[1] has derived a new solution of the empty-space Einstein field equations which in some sense represents a spinning object with mass; its linearized version is a mass monopole plus the Lens–Thirring spinning-particle metric. The present note has two purposes. In the first section we give a curious "derivation" of the Kerr metric by performing a complex coordinate transformation on the Schwarzschild metric. In the second section we attempt to argue that the Kerr metric has certain characteristics that suggest a metric arising from a ring of mass rotating about its axis of symmetry. There are three points to the argument: (a) In a covariantly defined flat space, the Riemann tensor considered as a field defined on the flat space is singular on a ring, (b) there is a very close analogy between the Kerr metric and a solution of Maxwell's equations having characteristics of a rotating ring of charge, and (c) using the definitions of the gravitational multipoles given in the previous paper[2] it is seen that the Kerr metric is compatible with the structure of a rotating ring of mass.

"DERIVATION" OF KERR METRIC

The Schwarzschild metric, written in standard coordinates, is

$$ds^2 = (1 - r_0/r)\,dx^{0^2} - (1 - r_0/r)^{-1}\,dr^2 - r^2(d\theta^2 + \sin^2\theta\,d\phi^2), \qquad r_0 = 2km/c^2. \tag{1}$$

It can be transformed by the coordinate transformation

$$u = t - r - r_0 \ln(r - r_0), \qquad r' = r, \qquad \theta' = \theta, \qquad \phi' = \phi,$$

into the form (dropping the primes)

$$ds^2 = (1 - r_0/r)\,du^2 + 2\,du\,dr - r^2(d\theta^2 + \sin^2\theta\,d\phi^2). \tag{2}$$

(The surface u = constant is a spherically symmetric null surface.)

The contravariant components of the metric [Eq. (2)], namely

$$g^{00} = 0, \quad g^{11} = -(1 - r_0/r), \quad g^{12} = 1, \quad g^{22} = -1/r^2, \quad g^{33} = -1/r^2\sin^2\theta,$$

can be written in the alternate form

$$g^{\mu\nu} = l^\mu n^\nu + l^\nu n^\mu - m^\mu\bar{m}^\nu - m^\nu\bar{m}^\mu, \tag{3}$$

where

$$l^\mu = \delta^\mu_1, \quad n^\mu = \delta^\mu_0 - \tfrac{1}{2}(1 - r_0/r)\delta^\mu_1, \quad m^\mu = \frac{1}{\sqrt{2}\,r}\left(\delta^\mu_2 + \frac{i}{\sin\theta}\,\delta^\mu_3\right), \quad \bar{m}^\mu = \frac{1}{\sqrt{2}\,r}\left(\delta^\mu_2 - \frac{i}{\sin\theta}\,\delta^\mu_3\right). \tag{4}$$

This complex null tetrad system forms the starting point of the "derivation" of the Kerr metric. "Derivation" is put in quotation marks because there is no simple, clear reason for the series of operations performed on the tetrad to yield a new (different from Schwarzschild) solution of the Einstein equations, or even to yield a solution of the empty-space equations at all. Nevertheless, we do obtain a new solution.

[Kerr has recently shown (in a private communication), from the Einstein field equations, that this

* Supported in part by Aerospace Research Laboratories, Office of Aerospace Research, U. S. Air Force.
[1] R. P. Kerr, Phys. Rev. Letters **11**, 237 (1963).
[2] A. I. Janis and E. T. Newman, J. Math. Phys. **6**, 902 (1965).

type of operation works for the class of solutions, $g_{\mu\nu} = \eta_{\mu\nu} + \lambda^2 l_\mu l_\nu$. This class contains the Schwarzschild metric as a special case.]

The coordinate r is allowed to take complex values and the tetrad is rewritten in the form

$$l^\mu = \delta_1^\mu, \quad n^\mu = \delta_0^\mu - \frac{1}{2}\left[1 - \frac{r_0}{2}\left(\frac{1}{r} + \frac{1}{\bar{r}}\right)\right]\delta_1^\mu,$$
$$m^\mu = \frac{1}{\sqrt{2}\,\bar{r}}\left(\delta_2^\mu + \frac{i}{\sin\theta}\,\delta_3^\mu\right), \tag{5}$$
$$\bar{m}^\mu = \frac{1}{\sqrt{2}\,r}\left(\delta_2^\mu - \frac{i}{\sin\theta}\,\delta_3^\mu\right),$$

$\bar{r}$ being the complex conjugate of r. (Note that part of the algorithm is to keep l^μ and n^μ real and m^μ and $\bar{m}^\mu$ the complex conjugates of each other.) We now formally perform the complex coordinate transformation

$$r' = r + ia\cos\theta, \qquad \theta' = \theta,$$
$$u' = u - ia\cos\theta, \qquad \phi' = \phi, \tag{6}$$

on the vectors l^μ, n^μ, and m^μ. ($\bar{m}'^\mu$ is defined as the complex conjugate of m'^μ.)

If one now allows r' and u' to be real, we obtain the following tetrad:

$$l'^\mu = \delta_1^\mu, \quad n'^\mu = \delta_0^\mu - \tfrac{1}{2}\{1 - r_0[r'/(r'^2 + a^2\cos^2\theta)]\}\,\delta_1^\mu,$$
$$m'^\mu = [\sqrt{2}\,(r' + ia\cos\theta)]^{-1} \times [ia\sin\theta(\delta_0^\mu - \delta_1^\mu) + \delta_2^\mu + (i/\sin\theta)\delta_3^\mu]. \tag{7}$$

The metric $g'^{\mu\nu} = l'^\mu n'^\nu + l'^\nu n'^\mu - m'^\mu\bar{m}'^\nu - m'^\nu\bar{m}'^\mu$ can now be shown by a coordinate transformation to be equivalent to that of Kerr.

INTERPRETATION

The contravariant form of the Kerr metric can be written as (dropping the primes)

$$g^{\mu\nu} = g^{o\mu\nu} + \lambda^2 l^\mu l^\nu,$$
$$\lambda^2 = \frac{r_0 r}{r^2 + a^2\cos^2\theta}, \quad l^\mu = \delta_1^\mu, \tag{8}$$

the $g^{o\mu\nu}$ being easily computed from Eq. (7). The form is dictated by requiring $g^{o\mu\nu}$ to be independent of r_0 (or the mass m). By calculating the Riemann tensor it is seen that if r_0 goes to zero, then the space is flat, which proves that $g^{o\mu\nu}$ can be looked on as a flat-space metric tensor which is covariantly defined by the Kerr metric. Another result from the study of the Riemann tensor is that the space is algebraically special, Petrov type ID, l_μ being one of the double principal null vectors. The vector l_μ is not, as it is in the Schwarzschild case, surface forming or hypersurface orthogonal, the constant a giving a measure of the curl of l_μ. There are two real invariants (or one complex one) which can be computed from the Riemann tensor, namely

$$\Psi_2 \equiv R_{\alpha\beta\gamma\delta}l^\alpha m^\beta n^\gamma \bar{m}^\delta = -r_0/2(r - ia\cos\theta)^3. \tag{9}$$

It should be emphasized that one can not treat the coordinates r, θ, and ϕ as usual polar coordinates, for even in the flat-space limit ($r_0 = 0$), the metric $g^{\mu\nu}$ is not the polar coordinate version of $\eta_{\mu\nu}$, the Minkowski metric. However the following coordinate transformation does lead to polar coordinates, $\tilde{r}$, $\tilde{\theta}$, $\tilde{\phi}$, and $\tilde{u} \equiv t - \tilde{r}$:

$$\tilde{r}^2 = r^2 + a^2\sin^2\theta, \quad \tan\tilde{\phi} = \frac{\tan\phi - a/r}{1 + (a/r)\tan\phi}$$
$$\cos\tilde{\theta} = r\cos\theta/(r^2 + a^2\sin^2\theta)^{\frac{1}{2}}, \tag{10}$$
$$\tilde{u} = u - (r^2 + a^2\sin^2\theta)^{\frac{1}{2}} + r.$$

We have the situation that the Kerr metric has associated with it a flat-space metric $g^{o\mu\nu}$ which allows us to define polar coordinates in the original nonflat space. We can now ask where, as a function of the polar coordinates plotted in the associated flat space, is the Riemann tensor, or its invariants, singular. Clearly Ψ_2 is singular at $r = 0$ and $\theta = \frac{1}{2}\pi$, or in polar coordinates [from Eq. (10)] it is singular on the circle $\tilde{r} = a$, $\tilde{\theta} = \frac{1}{2}\pi$. It is reasonable then to associate with the Kerr metric this ring singularity.

The second point of our interpretation arises from noting the striking analogy between the Kerr metric and a solution of the Maxwell equations. First we will show the analogy between the Schwarzschild metric and the Coulomb field. The single invariant of the Schwarzschild Riemann tensor is $\Psi_2 = -r_0/2r^3$; the analogous invariant[3] of the Coulomb field is

$$\Phi_1 \equiv \tfrac{1}{2}F_{\mu\nu}(l^\mu n^\nu + \bar{m}^\mu m^\nu) = -e/2r^2.$$

If in these two invariants we substitute $r = r' - ia\cos\theta$ [obtained from Eq. (6)], we get the invariant Eq. (9) for the Kerr metric and (dropping the prime again) for the Maxwell field we get the invariant

$$\Phi_1 = -e/2(r - ia\cos\theta)^2.$$

It can be shown that this is a solution of Maxwell's equations expressed in terms of the original coordinate system of the Kerr solution [Eq. (8) with $r_0 = 0$]; i.e., r and θ are not polar coordinates.

[3] For a discussion of the invariants of the Riemann tensor and the Maxwell field tensor and the analogy between them see Ref. 2.

Using the coordinate transformation to polar coordinates, Eq. (10), we can see that this solution is singular only on the circle $\tilde{r} = a$, and $\tilde{\theta} = \frac{1}{2}\pi$. The task of analyzing the multipole structure of this solution was rather laborious and only the first three terms were calculated with the following results: (a) the monopole moment is e; (b) there is no electric dipole and the magnetic dipole moment is proportional to ea; (c) there is no magnetic quadrupole moment and the electric quadrupole moment is proportional to ea^2; (d) there appears to be an alternation back and forth between the electric and magnetic type poles.

This structure plus its singularity leads us to conclude that the field is due to a ring of charge rotating about its axis of symmetry with angular velocity proportional to a.

The analogy between the Kerr solution and this solution of Maxwell's equations suggests that the Kerr metric represents a ring of mass rotating about its symmetry axis. This is substantiated by analyzing the multipole structure of the metric in terms of the definitions given in the previous paper.[2] The method consists of finding null surfaces in the Kerr space and introducing them as coordinate surfaces with an associated null tetrad system. A lengthy but not difficult calculation leads to results similar to that found in the Maxwell case; there exists (a) a monopole moment equal to m; (b) no mass dipole but a spin-pole proportional to ma; (c) no spin quadrupole but a mass quadrupole proportional to ma^2.

From these three points we believe that our interpretation of the Kerr metric is reasonable.[4]

ACKNOWLEDGMENTS

We would like to thank Dr. J. N. Goldberg and Dr. R. P. Kerr for many stimulating discussions.

[4] We wish to thank the referee and R. P. Kerr for pointing out that, although our analysis is correct, the final interpretation of the solutions (both Maxwell and Einstein fields) is probably incorrect due to the unusual (multivalued) behavior of the solutions when a closed loop that threads the singular ring is followed. In order to avoid this multivalued behavior of the solutions, it would be necessary to have a surface distribution of matter (charge in the Maxwell case), the surface being bounded by the singular ring, in such a way that the fields are discontinuous across the surface.

Metric of a Rotating, Charged Mass*

E. T. NEWMAN, E. COUCH, K. CHINNAPARED, A. EXTON, A. PRAKASH, AND R. TORRENCE
Physics Department, University of Pittsburgh, Pittsburgh, Pennsylvania
(19 June 1964)

A new solution of the Einstein–Maxwell equations is presented. This solution has certain characteristics that correspond to a rotating ring of mass and charge.

THE purpose of the present note is to present a new solution of the Einstein–Maxwell equations which in some sense represents a rotating mass and charge. This solution bears the same relation to the charged Schwarzschild metric[1] (Reissner–Nordström) as the Kerr spinning particle metric bears to the Schwarzschild. In fact one can "derive" it by means of a similar trick (complex coordinate transformation) as was used to "derive" the Kerr metric.[2]

The Reissner–Nordström metric in null coordinates[2] has the form

$$ds^2 = (1 - 2m/r + e^2/r^2)\,du^2 + 2\,du\,dr - r^2(d\theta^2 + \sin^2\theta\,d\phi^2), \quad (1)$$

where m and e are the mass and charge respectively and u labels the null surfaces. The contravariant form of the metric can be written

$$g^{\mu\nu} = l^\mu n^\nu + l^\nu n^\mu - m^\mu \bar{m}^\nu - m^\nu \bar{m}^\mu, \quad (2)$$

where

$$l^\mu = \delta_1^\mu, \quad m^\mu = (1/\sqrt{2}\,r)[\delta_2^\mu + (i/\sin\theta)\delta_3^\mu]$$
$$n^\mu = \delta_0^\mu - \left(\frac{1}{2} - \frac{m}{r} + \frac{e^2}{2r^2}\right)\delta_1^\mu, \quad (3)$$

and where $\bar{m}^\mu$ is the complex conjugate of m^μ.

A new metric can now be obtained by the following formal process. The radial coordinate r is allowed to take complex values and the tetrad is rewritten in the form

$$l^\mu = \delta_1^\mu, \quad m^\mu = (1/\sqrt{2}\,\bar{r})[\delta_2^\mu + (i/\sin\theta)\delta_3^\mu]$$
$$n^\mu = \delta_0^\mu - \frac{1}{2}\left(1 - m\left[\frac{1}{r} + \frac{1}{\bar{r}}\right] + \frac{e^2}{r\bar{r}}\right)\delta_1^\mu, \quad (4)$$

$\bar{r}$ being the complex conjugate of r. [It should be noted that if the term $e^2/2r^2$ in n^μ was replaced by $\frac{1}{4}e^2(r^{-2} + \bar{r}^{-2})$ instead of $e^2/2r\bar{r}$, we would not obtain a solution of the Einstein–Maxwell equations.] If we now perform the same complex coordinate transformation as was used in Ref. (2) ($r' = r + ia\cos\theta$, $u' = u - ia\cos\theta$) we obtain the following tetrad,

$$l^\mu = \delta_1^\mu, \quad m^\mu = [\sqrt{2}\,(r' + ia\cos\theta)]^{-1} \times [ia\sin\theta(\delta_0^\mu - \delta_1^\mu) + \delta_2^\mu + (i/\sin\theta)\delta_3^\mu],$$
$$n^\mu = \delta_0^\mu - [\tfrac{1}{2} - (mr' - \tfrac{1}{2}e^2)(r'^2 + a^2\cos^2\theta)^{-1}]\delta_1^\mu \quad (5)$$

and associated metric tensor $g'^{\mu\nu} = l'^\mu n'^\nu + l'^\nu n'^\mu - m'^\mu \bar{m}'^\nu - m'^\nu \bar{m}'^\mu$.

If we take the following Maxwell field (stated in terms of the tetrad components[3] of the field tensor $F_{\mu\nu}$ rather than in terms of the tensor itself)

$$\phi_0 \equiv F_{\mu\nu} l^\mu m^\nu = 0$$
$$\phi_1 \equiv \tfrac{1}{2}F_{\mu\nu}(l^\mu n^\nu + \bar{m}^\mu m^\nu) = e/\sqrt{2}\,(r - ia\cos\theta)^2 \quad (6)$$
$$\phi_2 \equiv F_{\mu\nu}\bar{m}^\mu n^\nu = iea\sin\theta/(r - ia\cos\theta)^3,$$

it can be shown by direct calculation that this field with the metric associated with Eq. (5) constitutes a solution of the Einstein–Maxwell equations. [We wish to point out that there was no simple algorithm which led to Eq. (6). It had to be obtained by integration.]

By arguments similar to those used in (2) we conclude that this solution represents the gravitational and electromagnetic fields of a ring of mass and charge rotating about its axis of symmetry.[4]

The Weyl tensor of this space is type II degenerate, the double null vector being l^μ. l^μ is also a principle null vector of the Maxwell tensor. l^μ is shear free but not hypersurface orthogonal, a measuring its curl.

In conclusion, we give the contra- and covariant

* Supported in part by Aerospace Research Laboratories, Office of Aerospace Research, U. S. Air Force.

[1] R. C. Tolman, *Relativity, Thermodynamics and Cosmology* (Oxford University Press, London, 1934).

[2] E. T. Newman and A. I. Janis, J. Math. Phys. **6**, 915 (1965).

[3] A. I. Janis and E. T. Newman, J. Math. Phys. **6**, 902 (1965).

[4] We wish to thank the referee and R. Kerr for pointing out a difficulty in this interpretation. The remarks in Footnote 4, Ref. 2 apply here as well.

forms of the metric, where $x \equiv (r^2 + a^2 \cos^2\theta)^{-1}$,

$$g^{\mu\nu} = \begin{Vmatrix} x(-a^2 \sin^2\theta) & x(r^2 + a^2) & 0 & -xa \\ \cdot & x[2mr - (r^2 + a^2) - e^2] & 0 & xa \\ \cdot & \cdot & -x & 0 \\ \cdot & \cdot & \cdot & x(-\sin^{-2}\theta) \end{Vmatrix}$$

and

$$g_{\mu\nu} = \begin{Vmatrix} 1 + x(e^2 - 2mr) & 1 & 0 & x(a \sin^2\theta)(2mr - e^2) \\ \cdot & 0 & 0 & -a \sin^2\theta \\ \cdot & \cdot & -x^{-1} & 0 \\ \cdot & \cdot & \cdot & -\sin^2\theta(r^2 + a^2 + ag_{03}) \end{Vmatrix}.$$

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 6 JUNE 1965

Unified Dirac–Von Neumann Formulation of Quantum Mechanics. I. Mathematical Theory*

A. R. MARLOW
Georgetown University, Washington, D. C.
(Received 25 September 1964)

In this paper the results from various areas of mathematical research which are necessary for a consistent unification of the Dirac and von Neumann formulations of quantum mechanics are collected and presented as a single synthesis. For this purpose, direct integral decompositions of Hilbert space must be introduced into Dirac's formulation of spectral theory and representation theory; true unit vectors in the direct integral decomposition spaces replace unnormalizable vectors of infinite length. It then becomes clear that families of modified Dirac projection operators are simply related to the Radon–Nikodym derivative of von Neumann spectral measures. In terms of these mathematical preliminaries a second paper will present the more physical aspects of the resulting unified formulation of quantum mechanics.

THE definitive and beautiful formulation given to quantum mechanics by Dirac[1] has the single disadvantage of requiring the introduction of unnormalizable vectors of infinite length into Hilbert space to represent eigenstates of observables having continuous spectra. This fact not only renders the theory mathematically nonrigorous, but even leads to practical difficulties in physical interpretation whenever powers and products of the functions representing such unnormalizable vectors (δ-functions) appear. The difficulties have been completely solved mathematically by the theories of von Neumann[2] and Schwartz,[3] but the methods used differ substantially from Dirac's approach and have not so far proved practical for physicists. Our purpose here is to present a consistent formulation of quantum mechanics which, while preserving the basic physically useful approach of Dirac, will do away with the need for mathematically objectionable unnormalizable vectors and will allow us at will to pass easily and rigorously from a Dirac-type formulation to the von Neumann formulation.

To do this we first need to present a spectral theory in terms of direct integral decompositions of Hilbert space and then, applying this theory, we can develop a rigorous but practical representation theory. This first paper, then, will be mainly a review and

* Work supported by a National Science Foundation fellowship.

[1] P. A. M. Dirac, *The Principles of Quantum Mechanics* (Clarendon Press, Oxford, England, 1958), 4th ed.

[2] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, New Jersey, 1955).

[3] L. Schwartz, *Théorie des Distributions* (Hermann & Cie., Paris, 1950–1951).

forms of the metric, where $x \equiv (r^2 + a^2 \cos^2\theta)^{-1}$,

$$g^{\mu\nu} = \begin{Vmatrix} x(-a^2 \sin^2\theta) & x(r^2 + a^2) & 0 & -xa \\ \cdot & x[2mr - (r^2 + a^2) - e^2] & 0 & xa \\ \cdot & \cdot & -x & 0 \\ \cdot & \cdot & \cdot & x(-\sin^{-2}\theta) \end{Vmatrix}$$

and

$$g_{\mu\nu} = \begin{Vmatrix} 1 + x(e^2 - 2mr) & 1 & 0 & x(a \sin^2\theta)(2mr - e^2) \\ \cdot & 0 & 0 & -a \sin^2\theta \\ \cdot & \cdot & -x^{-1} & 0 \\ \cdot & \cdot & \cdot & -\sin^2\theta(r^2 + a^2 + ag_{03}) \end{Vmatrix}.$$

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 6 JUNE 1965

Unified Dirac–Von Neumann Formulation of Quantum Mechanics. I. Mathematical Theory*

A. R. MARLOW
Georgetown University, Washington, D. C.
(Received 25 September 1964)

In this paper the results from various areas of mathematical research which are necessary for a consistent unification of the Dirac and von Neumann formulations of quantum mechanics are collected and presented as a single synthesis. For this purpose, direct integral decompositions of Hilbert space must be introduced into Dirac's formulation of spectral theory and representation theory; true unit vectors in the direct integral decomposition spaces replace unnormalizable vectors of infinite length. It then becomes clear that families of modified Dirac projection operators are simply related to the Radon–Nikodym derivative of von Neumann spectral measures. In terms of these mathematical preliminaries a second paper will present the more physical aspects of the resulting unified formulation of quantum mechanics.

THE definitive and beautiful formulation given to quantum mechanics by Dirac[1] has the single disadvantage of requiring the introduction of unnormalizable vectors of infinite length into Hilbert space to represent eigenstates of observables having continuous spectra. This fact not only renders the theory mathematically nonrigorous, but even leads to practical difficulties in physical interpretation whenever powers and products of the functions representing such unnormalizable vectors (δ-functions) appear. The difficulties have been completely solved mathematically by the theories of von Neumann[2] and Schwartz,[3] but the methods used differ substantially from Dirac's approach and have not so far proved practical for physicists. Our purpose here is to present a consistent formulation of quantum mechanics which, while preserving the basic physically useful approach of Dirac, will do away with the need for mathematically objectionable unnormalizable vectors and will allow us at will to pass easily and rigorously from a Dirac-type formulation to the von Neumann formulation.

To do this we first need to present a spectral theory in terms of direct integral decompositions of Hilbert space and then, applying this theory, we can develop a rigorous but practical representation theory. This first paper, then, will be mainly a review and

* Work supported by a National Science Foundation fellowship.

[1] P. A. M. Dirac, *The Principles of Quantum Mechanics* (Clarendon Press, Oxford, England, 1958), 4th ed.

[2] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, New Jersey, 1955).

[3] L. Schwartz, *Théorie des Distributions* (Hermann & Cie., Paris, 1950–1951).

synthesis of the mathematical results we need from various sources. Proofs for the assertions made will not be given here, but they can readily be found in our references. The only novel features not present (to the author's knowledge) in existing mathematical or mathematical–physical literature can be summarized as (1) replacement of unnormalizable vectors in the Dirac representation theory of Hilbert space by families of true vectors in direct integral decomposition spaces, and (2) explicitation of the connection between families of Dirac-type projection operators and the Radon–Nikodym derivatives of von Neumann spectral measures.[4] In our paper II, we will go on to give a physical formulation of quantum mechanics in terms of these mathematical preliminaries in a way that unifies the work of Dirac and von Neumann.

1. SPECTRAL THEORY IN HILBERT SPACE

As the first step in our program we indicate how Hilbert spaces general enough to meet all the needs of quantum mechanics can be built up from simpler spaces. The topological and measure theoretical background needed for this can be found in Halmos.[6]

Let X be any locally compact set, H a mapping defined over X such that $H(x)$ is a Hilbert space for each $x \in X$, and $f, g, \cdots$ various mappings over X such that $f(x)$, $g(x)$ are vectors in $H(x)$. Thus H is a family of Hilbert spaces and $f, g, \cdots$ are families of vectors over X. If now Σ is a suitable σ-ring of subsets of X with the property that $\bigcup \Sigma = X$, the pair (X, Σ) will form a locally compact measurable space. We will say that H is a *measurable family* of Hilbert spaces if there exists a set S of families of vectors over X such that: (1) for all $f \in S$ the function $||f||$, which maps each $x \in X$ into the number $||f(x)||$ [the norm in $H(x)$ of the vector $f(x)$], is a measurable function on (X, Σ); (2) if, for every $f \in S$, whenever the function $\langle g, f\rangle$ is measurable on (X, Σ) then $g \in S$, where $\langle g, f\rangle$ maps each $x \in X$ into the number $\langle g(x), f(x)\rangle$ [the inner product in $H(x)$ of the vectors $g(x)$ and $f(x)$]; (3) there exists a countable set $\{f_n\}$ of elements in S such that for each $x \in X$ the set $\{f_n(x)\}$ is dense in the space $H(x)$. The families of vectors in S are then called *measurable families* of vectors. From (3), it is clear that to form a measurable family of spaces the individual spaces $H(x)$ must be separable.

Now let us assume that we have defined a measure μ on (X, Σ), so that we have a locally compact measure space (X, Σ, μ), and that H is a measurable family of Hilbert spaces on X. A family f of vectors in the spaces $H(x)$ is called square integrable on (X, Σ, μ) if it is measurable and if

$$\int_X ||f(x)||^2 \, d\mu(x) < \infty . \tag{1}$$

The set of square integrable families of vectors on (X, Σ, μ) can be shown to be a Hilbert space $H^{\oplus}$, called the *direct integral* of the spaces $H(x)$, and symbolized as

$$H^{\oplus} = \int_X^{\oplus} H(x) \, d\mu(x). \tag{2}$$

The construction of the space $H^{\oplus}$ is essentially unique in the sense that two measurable families H and H' yielding the same direct integral space can only differ on μ-null sets in X.[5]

This process of synthesis can be reversed and general Hilbert spaces decomposed or analyzed as direct integrals of measurable families of spaces. The study of such decompositions largely constitutes the spectral theory of Hilbert space. We will need only the parts of this theory which deal with decompositions relative to commutative sets of normal operators (i.e., operators that commute with their adjoints); our treatment will be essentially that of Maurin.[7]

Let H be a separable Hilbert space with vectors $f, g, \cdots$ and scalar product $\langle f, g\rangle$; let $\alpha = \{\alpha_i\}$ be a commutative set of normal operators on H, so that $\alpha_i\alpha_j = \alpha_j\alpha_i$ and $\alpha_i\alpha_i^{\dagger} = \alpha_i^{\dagger}\alpha_i$, for $\alpha_i, \alpha_j \in \alpha$ and $\alpha_i^{\dagger}$ the adjoint of α_i. Then there exists a direct integral of Hilbert spaces

$$\hat{H}^{\oplus} = \int_A^{\oplus} \hat{H}(x) \, d\mu(x) \tag{3}$$

over a compact measurable subset A of a locally compact measure space (X, Σ, μ), and a unitary mapping U from H onto $\hat{H}^{\oplus}$ which simultaneously

[4] Since the completion of this paper we noticed in the stimulating and beautiful new book of R. F. Streater and A. S. Wightman, *PCT, Spin and Statistics, and All That* (W. A. Benjamin, Inc., New York, 1964), the following statement (p. 92): "The method (i.e., of Dirac) can be completely justified by using the theory of direct integrals, but we want to avoid that." The reference given there is to a mathematical development of direct integral theory (our reference 5). The present paper, then, can be read as an explanatory footnote to the statement of Streater and Wightman.

[5] For a complete treatment of direct integrals of Hilbert spaces see: J. Dixmier, *Les Algebres d'Operateurs dans l'Espace Hilbertien* (Gauthier-Villars, Paris, 1957).

[6] P. R. Halmos, *Measure Theory* (D. Van Nostrand, Inc., Princeton, New Jersey, 1950), particularly Chaps. V, VI, VII, and X.

[7] K. Maurin, "Spectraldarstellung der Kerne," Bull. Acad. Polonaise Sci. 8, 461 (1959).

diagonalizes the system α; that is, there exists a mapping

$$U: f \to U(f) = \hat{f} \in \hat{H}^{\oplus}, \quad \hat{f}(x) \in \hat{H}(x) \quad \text{for} \quad x \in A \tag{4}$$

such that

$$\langle f, g \rangle = \int_A \langle \hat{f}(x), \hat{g}(x) \rangle \, d\mu(x) = \langle \hat{f}, \hat{g} \rangle \tag{5}$$

and

$$U(\alpha_i f) = (\alpha_i f)\hat{} = \hat{a}_i \hat{f}, \qquad \alpha_i \in \alpha, \tag{6}$$

where the $\hat{a}_i$ are measurable scalar functions $\hat{a}_i: x \to \hat{a}_i(x) \in C$ for $x \in A$. These functions are finite almost everywhere with respect to the measure μ, which is positive over the set A. In case any particular $\alpha_i \in \alpha$ is unitary ($\alpha_i^{\dagger} = \alpha_i^{-1}$) we have $\hat{a}_i(x) = 1$ almost everywhere. If α_i is self-adjoint, then $\hat{a}_i(x)$ is real.

In case the set α contains only finitely many normal operators, say $\alpha_1, \alpha_2, \cdots, \alpha_N$, we have available the so-called *canonical diagonalization*; that is, there exists a unitary mapping $U : H \to \hat{H}^{\oplus}$ such that the space X is C^N (the product of N complex planes), and

$$(\alpha_i f(z))\hat{} = z_i \hat{f}(z), \qquad z_i \in C_i,$$
$$z = (z_1, \cdots, z_i, \cdots, z_N) \in \prod_{i=1}^{N} C_i = C^N. \tag{7}$$

If all the operators $\alpha_i \in \alpha$ are self-adjoint then the set A will be a subset of N-dimensional Euclidean space, R^N. If all the α_i are unitary then A is the product of N unit circles in the complex plane, i.e., $z_i = e^{ix_i}$, $0 \leq x_i < 2\pi$. Also, if β is an operator (bounded operator) on H that commutes with all the α_i then β is decomposable in the following sense:

$$(\beta f(x))\hat{} = \hat{\beta}(x)\hat{f}(x), \qquad x \in A, \tag{8}$$

where $\hat{\beta}(x)$ is an operator (bounded operator) in the space $\hat{H}(x)$. We should note that nothing has been stated concerning either the dimensions of the individual spaces $\hat{H}(x)$ or the particular structure of the measure space (X, Σ, μ). These depend on the particular set of normal operators relative to which the direct integral decomposition is made.

Let us consider now the bilinear form $\langle f, \alpha_i g \rangle$, where f, g, are arbitrary vectors in a Hilbert space H and α_i is a normal operator on H. Using the canonical diagonalization of α_i with respect to a set α of N normal operators such that $\alpha_i \in \alpha$ we have

$$\langle f, \alpha_i g \rangle = \int_A z_i \langle \hat{f}(z), \hat{g}(z) \rangle \, d\mu(z), \qquad z \in C^N, \tag{9}$$

where the measure μ is a product measure $\mu = \prod_{i=1}^{N} \mu_i(z_i)$ over C^N. The integral is taken over a measurable set $A = \prod_{i=1}^{N} A_i$, where $A_i \subset C$ can be any measurable set containing the spectrum of α_i. Now from Fubini's theorem relating multiple and iterated integrals and the converse of the Radon–Nikodym theorem for transformation of measures in integrals,[8] we can rewrite this equation as

$$\langle f, \alpha_i g \rangle = \int_{A_i} z_i \, d\gamma_i(f, g, z_i), \tag{10}$$

where the complex measure

$$\gamma_i(f, g, B) = \int_B \int_{A'} \langle \hat{f}(z), \hat{g}(z) \rangle \, d\mu(z),$$
$$B \subset C, \qquad z \in C^N, \qquad A' = \prod_{j \neq i} A_j, \tag{11}$$

is determined by the vectors $f, g \in H$ and is called the *canonical spectral measure* of the operator α_i with respect to the vectors f and g. We could, of course, choose any or all of the sets A_i to be the whole complex plane C. Equation (10) gives the canonical spectral resolution of the bilinear form $\langle f, \alpha_i g \rangle$, and is unitarily invariant, that is, the measure γ over the complex plane as well as a minimal set $A_i \subset C$ are uniquely defined for every normal operator α, independently of the particular unitary mapping U of H onto some direct integral $\hat{H}^{\oplus}$, and so independently of the choice of a particular set of normal operators with respect to which H is decomposed. In particular, if we choose for the set $\{\alpha_i\}$ a set consisting of a single normal operator α we get

$$\langle f, \alpha g \rangle = \int_C z \, d\gamma(f, g, z), \qquad z \in C, \tag{12}$$

with

$$\gamma(f, g, B) = \int_B \langle \hat{f}(z), \hat{g}(z) \rangle \, d\mu(z), \qquad B \subset C. \tag{13}$$

2. VON NEUMANN FORMULATION.

This is the convenient point in our exposition of spectral theory to introduce the notions von Neumann used in formulating quantum mechanics[2] and to relate them to the formalism we have developed so far. A *resolution of the identity*, E, on a Hilbert space H is a one-parameter family of projection operators taking on values E_t, where t ranges over a finite or infinite interval $[a, b] \subset R$, which satisfies the following conditions: (1) $E_a = 0$, $E_b = I$ (identity operator); (2) $E_{t+0} = E_t$ (con-

[8] Reference 6, pp. 143, ff., and 128, ff.

tinuity from the right); (3) $E_rE_s = E_m$ ($m = \min\{r, s\}$). For our purposes we will consider only those resolutions of the identity defined over the whole real axis. We lose no generality by this apparent restriction since we can extend any E defined initially only on an interval $[a, b]$ to all of R by setting $E_t = 0$ for $t < a$ and $E_t = I$ for $t > b$. The first two theorems we will state show the importance of such resolutions of the identity for spectral theory and give explicit prescriptions for constructing the canonical spectral measures of self-adjoint and unitary operators. Proofs can be found in the standard references on linear operators in Hilbert space quoted with each result.

Let $\mathfrak{A}$ be the set of self-adjoint operators on a Hilbert space H and $\mathfrak{U}$ the set of unitary operators on H. Let $\mathfrak{A}'$ be the set of resolutions of the identity on H defined over the whole real axis and $\mathfrak{U}'$ the subset of $\mathfrak{A}'$ such that $E \in \mathfrak{U}'$ if and only if $E_t = 0$ for $t \leq 0$ and $E_t = I$ for $t \geq 2\pi$.

Theorem 1.[9,10] There exists a one-to-one onto mapping (i.e., a bijection) from $\mathfrak{A}$ onto $\mathfrak{A}'$ such that for $f, g \in H$ and $\alpha \in \mathfrak{A}$ the corresponding resolution of the identity $E(\alpha) \in \mathfrak{A}'$ satisfies

$$\langle f, \alpha g\rangle = \int_{-\infty}^{+\infty} t\, d\langle f, E_t(\alpha)g\rangle. \tag{14}$$

The domain of the operator α is the set of vectors (dense in H) for which the inequality

$$\int_{-\infty}^{+\infty} t^2\, d\langle f, E_t(\alpha)g\rangle < \infty \tag{15}$$

is satisfied.

Theorem 2.[11] There exists a bijection from $\mathfrak{U}$ onto $\mathfrak{U}'$ such that for $f, g \in H$ and $\beta \in \mathfrak{U}$ the corresponding $E(\beta) \in \mathfrak{U}'$ satisfies

$$\langle f, \beta g\rangle = \int_{-\infty}^{+\infty} e^{it}\, d\langle f, E_t(\beta)g\rangle. \tag{16}$$

Now from Eq. (12) and Theorems 1 and 2 we see that for $\alpha \in \mathfrak{A}$, $\beta \in \mathfrak{U}$ and $f, g \in H$ we have

$$\int_{-\infty}^{+\infty} t\, d\langle f, E_t(\alpha)g\rangle = \int_{-\infty}^{+\infty} t\, d\gamma(f, g, \alpha, t) \tag{17}$$

and

$$\int_{-\infty}^{+\infty} e^{it}\, d\langle f, E_t(\beta)g\rangle = \int_{-\infty}^{+\infty} e^{it}\, d\gamma(f, g, \beta, t), \tag{18}$$

where in each case γ is the canonical spectral meaure corresponding to the given operator and the vectors f and g. Thus for an operator δ, which is either self-adjoint or unitary, we see that the function $\langle f, E(\delta)g\rangle$, which maps each $t \in R$ into the complex number $\langle f, E_t(\delta)g\rangle$, acts as a complex distribution over the real axis and generates the complex Lebesgue–Stieltjes canonical spectral measure $\gamma(f, g, \delta)$.

Since $\mathfrak{U}'$ is a subset of $\mathfrak{A}'$, a resolution of the identity $E(\beta)$ corresponding to a unitary operator β also corresponds to a self-adjoint operator, which we will write as $\alpha(\beta)$, such that

$$\langle f, \alpha(\beta)g\rangle = \int_{-\infty}^{+\infty} t\, d\langle f, E_t(\beta)g\rangle. \tag{19}$$

Thus if $a \subset \mathfrak{A}$ is the set of self-adjoint operators such that $E(a) = \mathfrak{U}'$, it is clear that there is a bijective mapping of a onto $\mathfrak{U}$ which we will write as

$$\beta(\alpha) = e^{i\alpha}, \qquad \alpha \in a. \tag{20}$$

From (16) we see that

$$\langle f, e^{i\alpha}g\rangle = \int_{-\infty}^{+\infty} e^{it}\, d\langle f, E_t(\alpha)g\rangle. \tag{21}$$

Corresponding to Eqs. (14), (16), and (21) and under the same conditions, we will write the operator equations

$$\alpha = \int_{-\infty}^{+\infty} t\, dE_t(\alpha), \tag{14'}$$

$$\beta = \int_{-\infty}^{+\infty} e^{it}\, dE_t(\beta), \tag{16'}$$

$$\beta(\alpha) = e^{i\alpha} = \int_{-\infty}^{+\infty} e^{it}\, dE_t(\alpha). \tag{21'}$$

It should be noted that these relationships can be shown to make sense purely as operator identities even without reference to (14), (16), and (21).[12]

From the definition of a resolution of the identity and the fundamental properties of Lebesgue–Stieltjes integration we have the following important relationships:

$$\int_{-\infty}^{+\infty} d\langle f, E_t g\rangle = \langle f, E_\infty g\rangle - \langle f, E_{-\infty} g\rangle = \langle f, g\rangle, \tag{22}$$

that is

$$\int_{-\infty}^{+\infty} dE_t = I \quad \text{for} \quad E \in \mathfrak{A}'. \tag{23}$$

It will also be useful here to recall the relationship between resolutions of the identity and the spectra of self-adjoint and unitary operators. Given an

[9] F. Riesz and B. Sz.-Nagy, *Functional Analysis* (Fredrick Ungar Publishing Company, New York, 1955), p. 320.
[10] N. I. Akhiezer and I. M. Glazman, *Theory of Linear Operators in Hilbert Space* (Frederick Ungar Publishing Company, New York, 1963), Vol. II, p. 36.
[11] Reference 9, p. 281.
[12] Reference 10, Vol. II, p. 21, ff.

$E \in \mathfrak{A}'$ we will say that the point $a \in R$ is a point of growth of E if for all $\delta > 0$ we have $E_{a+\delta} - E_{a-\delta} \neq 0$. Then the spectrum of a self-adjoint operator α can be specified as the set of points of growth of $E(\alpha)$ and the spectrum of a unitary operator β is the set of complex numbers $\{e^{it}\}$, where t ranges over the points of growth of $E(\beta)$.

Before closing this catalog of results from spectral theory, we give one further definition and two alternate versions of a theorem proven by M. H. Stone. A *one-parameter group* of unitary operators on a Hilbert space H will be defined as a family β of unitary operators β_t $(-\infty < t < +\infty)$ possessing the properties: $\beta_0 = I$ and $\beta_r\beta_s = \beta_{r+s}$. We will say that such a family is continuous if the function $\langle f, \beta g\rangle$, which takes each $t \in R$ into the number $\langle f, \beta_t g\rangle$, is a continuous function of t for every pair of vectors $f, g \in H$. If $\mathfrak{C}$ denotes the set of all continuous one-parameter groups of unitary operators on H we can state

Theorem 3. (First version of Stone's theorem).[13] There exists a bijection from the set $\mathfrak{C}$ onto the set $\mathfrak{A}'$ such that for $f, g \in H$ and $\beta \in \mathfrak{C}$ the corresponding $E(\beta) \in \mathfrak{A}'$ satisfies the equation

$$\langle f, \beta_t g\rangle = \int_{-\infty}^{+\infty} e^{its}\, d\langle f, E_s(\beta) g\rangle \tag{24}$$

for all $t \in R$. As before, we will write

$$\beta_t = \int_{-\infty}^{+\infty} e^{its}\, dE_s(\beta) \tag{25}$$

and

$$\beta_t(\alpha) = e^{it\alpha}, \tag{26}$$

where α is the unique self-adjoint operator such that $E(\beta) = E(\alpha)$. Clearly this specifies a bijective correspondence from $\mathfrak{A}$ onto $\mathfrak{C}$. This bijection is further elaborated in

Theorem 4. (Second version of Stone's theorem).[14] Every $\beta \in \mathfrak{C}$ is generated by a unique operator $\alpha \in \mathfrak{A}$ and conversely, as given by the relationships (26) and

$$\alpha(\beta) = \lim_{h \to 0} \frac{i}{h} (I - \beta_h). \tag{27}$$

3. REPRESENTATION THEORY IN HILBERT SPACE

For convenience in our latter application of the formalism to quantum mechanics we now introduce the Dirac notation for the vectors, scalar product and projection operators of a Hilbert space H; that is, for $f \in H$ we write $|f\rangle \in H$ and for a pair of vectors $f, g \in H$ we write $\langle f|g\rangle = \langle f, g\rangle$. This allows us to write the vectors $f^*, g^*, \cdots$, in the dual space H^* (i.e., the bounded linear functionals on H) as $\langle f|$, $\langle g|$, $\cdots$, and to use a symmetric notation for the bilinear functionals on H by defining $\langle f|\ \alpha\ |g\rangle = \langle f, \alpha g\rangle = \langle \alpha^\dagger f, g\rangle$, where $\alpha^\dagger$ is the adjoint of the operator α. For a unit vector $|u\rangle \in H$ we will use the dyadic symbol $|u\rangle\langle u|$ to mean the projection operator that projects onto the one-dimensional subspace spanned by $|u\rangle$. Now let $\mathfrak{B}_H$ be the class of all orthonormal basis sets for the space H, that is the class of all bases for H such that the vectors in each basis are unit vectors and are orthogonal to each other. Then, given a separable Hilbert space H and a set $B = \{|j\rangle\} \in \mathfrak{B}_H$, where the vectors in B are labeled by the index j, we have the three identities

$$\sum_j |j\rangle\langle j| = I \qquad \text{(completeness relation)}, \tag{28}$$

$$|f\rangle = \sum_j |j\rangle\langle j|\ |f\rangle = \sum_j \langle j \mid f\rangle\ |j\rangle \qquad \text{(Fourier expansion)}, \tag{29}$$

$$\langle f \mid g\rangle = \sum_j \langle f|\ |j\rangle\langle j|\ |g\rangle = \sum_j \langle f \mid j\rangle\langle j \mid g\rangle \quad \text{(Parseval's identity)}. \tag{30}$$

Now each decomposition of a separable Hilbert space into a direct integral $\hat{H}^{\oplus}$ produces a corresponding family of separable Hilbert spaces $\hat{H}(x)$, where x ranges over the elements of some locally compact space X and the spaces $\hat{H}(x)$, while not subspaces of H, are related to it in the way that we specified earlier. Unless otherwise stated, we will limit ourselves to canonical decompositions of H with respect to a set of N self-adjoint commuting operators, $\alpha = \{\alpha_i\}$. Thus $X = R^N$. We adopt the Dirac convention regarding scalar variables corresponding to self-adjoint and unitary operators: if α represents a set of N self-adjoint operators, then the corresponding primed symbol α' represents the set of real variables $(\alpha_1', \alpha_2', \cdots, \alpha_N') \in R^N$. This notation will allow us to use such symbolism as $\hat{H}(\alpha')$ to mean a member of the family of spaces resulting from a direct integral decomposition of a space H with respect to the set α. If α is a single self-adjoint operator and $\beta = e^{i\alpha}$, then of course $\beta' = e^{i\alpha'} \in C$, $\alpha' \in R$. We can also write $E_{\alpha'}$ for the projection operator $E_t(\alpha)$ in the family $E(\alpha)$ if we set $t = \alpha'$. We will further agree to use the special rounded bracket symbols $|\alpha', f)$, $|\alpha', g)$, $\cdots$,

[13] Reference 9, p. 383; Ref. 10, Vol. II, p. 29.
[14] Reference 9, p. 385.

without caps, to represent the vectors $\hat{f}(\alpha'), \hat{g}(\alpha'), \cdots$, in one of the spaces $\hat{H}(\alpha')$ resulting from a direct integral decomposition of a space H. Then the scalar product in $\hat{H}(\alpha')$ will be written $(f, \alpha'|\alpha', g)$, and the projection operator corresponding to a unit vector $|\alpha', j)$ in a basis $B \in \mathfrak{B}_{H(\alpha')}$ will be $|\alpha', j)(j, \alpha'|$. It follows of course that the identity operators on the spaces $\hat{H}(\alpha')$ will be given by

$$\hat{I}(\alpha') = \sum_{j=1}^{d(\alpha')} |\alpha', j)(j, \alpha'|, \tag{31}$$

where $d(\alpha')$ is the dimension number of $\hat{H}(\alpha')$. To simplify writing, from now on we adopt the convention that the symbol $\sum_j |\alpha', j)(j, \alpha'|$ will mean a summation over all $d(\alpha')$ projection operators corresponding to the vectors in a set $\{|\alpha', j)\} \in \mathfrak{B}_{\hat{H}(\alpha')}$.

Let N be an integer-valued function defined over some measure space (X, Σ, μ) (where, by convention, we are considering $\aleph_0$, the cardinal number of the integers, to be an integer), and let C_N be the set $\bigcup_{x \in X} C^{N(x)}$. Consider now the set $F_{\mu,N}$ of all mappings $f : X \to C_N$ such that $f(x) \in C^{N(x)}$. If there exists a mapping $r : H \to F_{\mu,N}$, where H is some Hilbert space, such that for any $|f\rangle, |g\rangle \in H$ the mappings $f = r(|f\rangle)$ and $g = r(|g\rangle)$ in $F_{\mu,N}$ have the property:

$$\int_X \sum_{j=1}^{N(x)} f_j^*(x) g_j(x)\, d\mu(x) = \langle f \mid g\rangle, \tag{32}$$

[$f_j^*(x)$ is the complex conjugate of the jth component of $f(x)$], then clearly the set $r(H) \subseteq F_{\mu,N}$ has an induced Hilbert space structure with addition and multiplication by a scalar defined in the obvious way and the scalar product given by (32). In that case we call $r(H)$ a generalized *function space representation* of H. If the mapping r is one-to-one we will say that $r(H)$ is a faithful representation of H. We symbolize the set of all function space representations of a given space H, for arbitrary mappings N and measure spaces (X, Σ, μ), by $\mathfrak{F}_H$.

In terms of these concepts we can now develop the principal results of our formulation of representation theory. Given a Hilbert space H and a set of N self-adjoint operators on H, say $\alpha = \{\alpha_i\}$, let $\hat{H}^{\oplus}_{\alpha,\mu}$ be a direct integral decomposition of H with respect to α, with U the unitary mapping of H onto $\hat{H}^{\oplus}_{\alpha,\mu}$. Let $|f), |g) \in \hat{H}^{\oplus}_{\alpha,\mu}$ be the families of vectors $U(|f\rangle)$ and $U(|g\rangle)$ respectively, and write $|f)(\alpha') = |\alpha', f) \in \hat{H}(\alpha')$ and $|g)(\alpha') = |\alpha', g) \in \hat{H}(\alpha')$, $\alpha' \in R^N$. Now we can choose a family B of orthonormal bases for the spaces $H(\alpha')$ so that for each $\alpha' \in R^N$ we have $B_{\alpha'} = \{|\alpha', j)\} \in \mathfrak{B}_{\hat{H}(\alpha')}$, $1 \leq j \leq d(\alpha')$, and use Eq. (31) to rewrite Eq. (5) as

$$\langle f \mid g\rangle = \int_{R^N} \sum_j (f, \alpha' \mid \alpha', g)\, d\mu(\alpha'), \tag{33}$$

where μ is a real Lebesgue–Stieltjes product measure over R^N. To simplify writing even further we will omit one of the α' symbols in the scalar products $(f, \alpha'|\alpha', g)$ and write simply $(f \mid \alpha', g)$. Clearly the mappings $f : R^N \to C_d$, where the functions $f_j : R^N \to C$ are given by $f_j(\alpha') = (f \mid \alpha', j)$, specify a function space representation of H which we will write as $F_H(\alpha, \mu, B)$. Thus we see that each choice of α, μ, and B satisfying the conditions we have laid down defines a unique representation in $\mathfrak{F}_H$. We will use the following consistent notation: $F_H \subset \mathfrak{F}_H$ is the set of all function space representations generated via direct integral decompositions of a Hilbert space H with respect to sets of self-adjoint operators on H; $F_H(\alpha)$ is the set of all function space representations in F_H corresponding to the particular set of operators α; $F_H(\alpha, \mu)$ is the set of representations specified by the particular set α of N operators and the choice of a suitable measure μ over R^N. From the fact that $U : H \to \hat{H}^{\oplus}_{\alpha,\mu}$ is an isomorphism it is clear that all the representations in F_H are faithful.

Now the dimension function d, giving the dimensions of the spaces $\hat{H}(\alpha')$ in a canonical direct integral decomposition of H relative to a given set α of N self-adjoint operators, is the same (except over μ-null sets in R^N) for every representation in $F_H(\alpha)$.[15] If for some set α we have $d(\alpha') = 1$ almost everywhere with respect to μ, we say that α is a *maximal Abelian* set of self-adjoint operators on H (or in Dirac terminology, a complete commuting set). Obviously, for every representation of H generated by a maximal Abelian set α, Eq. (33) simplifies to

$$\langle f \mid g\rangle = \int_{R^N} (f \mid \alpha')(\alpha' \mid g)\, d\mu(\alpha'), \tag{33'}$$

where $|\alpha')$ is a unit vector spanning the space $\hat{H}(\alpha')$.

In particular, if we choose to decompose a Hilbert space H with respect to a single self-adjoint operator α, from Eqs. (23) and (33) we can write the two equations

$$\langle f \mid g\rangle = \int_{-\infty}^{+\infty} d\langle f|\, E_{\alpha'}\, |g\rangle \tag{34}$$

and

$$\langle f \mid g\rangle = \int_{-\infty}^{+\infty} \sum_j (f \mid \alpha', j)(j, \alpha' \mid g)\, dm(\alpha'), \tag{35}$$

where the function m is a real distribution function for a Lebesgue–Stieltjes measure over R. Then from

[15] Reference 6, page 221, ff.

the properties of Lebesgue–Stieltjes integration and the Radon–Nikodym theorem we see that the Lebesgue–Stieltjes measure

$$\langle f|\, M_\alpha(a, b)\, |g\rangle = \int_a^b d\langle f|\, E_{\alpha'}\, |g\rangle$$

$$= \int_a^b \sum_i (f \mid \alpha', j)(j, \alpha' \mid g)\, dm(\alpha') \qquad (36)$$

is well defined for every $|f\rangle$, $|g\rangle \in H$ and every interval $(a, b] \subset R$, and that the function $(f \mid g) : R^N \to C$ given by

$$(f \mid g)(\alpha') = \sum_j (f \mid \alpha', j)(j, \alpha' \mid g)$$

is the Radon–Nikodym derivative of the function $\langle f|\, E(\alpha)\, |g\rangle$ with respect to the function m, that is

$$[d\langle f|\, E(\alpha)\, |g\rangle/dm](\alpha') = \sum_j (f \mid \alpha', j)(j, \alpha' \mid g) \qquad (37)$$

[in fact, (36) can be taken as the definition of the Radon–Nikodym derivative of $\langle f|\, E(\alpha)\, |g\rangle : R^N \to C : \alpha' \to \langle f|\, E_{\alpha'}\, |g\rangle$].

For quantum mechanical purposes it will be important to consider operators of the type $P_\alpha(a, b)$, defined over a Hilbert space H by

$$P_\alpha(a, b) = \int_a^b dE_{\alpha'} = E_b - E_a. \qquad (38)$$

This is a projection operator and from Eq. (36) we see that the corresponding projection operator over the space $\hat{H}^{\oplus}_{\alpha m}$ is given by

$$\hat{P}_\alpha(a, b) = \int_a^b \sum_i |\alpha', j)(j, \alpha'|\, dm(\alpha')$$

$$= \int_{-\infty}^{+\infty} Q_{a,b}(\alpha') \sum_i |\alpha', j)(j, \alpha'|\, dm(\alpha'), \qquad (39)$$

where the function $Q_{a,b}$ is the characteristic function of the set $(a, b]$, that is ,$Q_{a,b}(\alpha') = 1$ for $\alpha' \in (a, b]$ and $Q_{a,b}(\alpha') = 0$ for $\alpha' \notin (a, b]$. But since the operator $\sum_i |\alpha', j)(j, \alpha'|$ is simply $\hat{I}(\alpha')$, the identity on the space $\hat{H}(\alpha')$, it is clear that the projection operator $P_\alpha(a, b)$ is represented in $\hat{H}(\alpha')$ by the operator $\hat{P}_{\alpha'}(a, b) = Q_{a,b}(\alpha')\hat{I}(\alpha')$, and so the family $\hat{P}_\alpha(a, b)\, |f) \in \hat{H}^{\oplus}_{\alpha,m}$ is given by $\hat{P}_\alpha(a, b)\, |f)(\alpha') = \hat{P}_{\alpha'}(a, b)\, |\alpha', f) = Q_{a,b}(\alpha')\, |\alpha', f)$. Thus, from Eqs. (36) and (37) we see that the family $\hat{I}$ of identities on the spaces $\hat{H}(\alpha')$ can be regarded as a sort of operator derivative of the operator $E(\alpha)$ with respect to the function m.

Now let $S_\alpha \subset R$ be the spectrum of the self-adjoint operator α. From (36) and the fact that the resolution of the identity $E(\alpha)$ is constant over intervals on the real axis not containing points in S_α we see that the only spaces $\hat{H}(\alpha')$ in a direct integral decomposition that will play any part in spectral theory or representation theory are those spaces which correspond to an $\alpha' \in S_\alpha$. The number $d(\alpha')$, the dimension of $\hat{H}(\alpha')$, for any $\alpha' \in S_\alpha$ is called the *multiplicity* of the spectral point α' or the multiplicity of the spectrum at α'. The spectrum of α is said to be *simple* if $d(\alpha') = 1$ for all $\alpha' \in S_\alpha$. In all other cases the spectrum is said to be *multiple*.

The last major question concerning the machinery of function space representation theory must now be asked: How are arbitrary operators γ on a Hilbert space H to be represented in terms of function space representations of H? To study this question we first specify a representation of H, say one in the set $F_H(\alpha, \mu)$, where α is a set of N self-adjoint operators. This enables us to write the equation:

$$\langle f|\, \gamma\, |g\rangle = \int_{R^N} \sum_i (f \mid \alpha', j)(j, \alpha' \mid g_\gamma)\, d\mu(\alpha') \qquad (40)$$

for $|f\rangle$, $|g\rangle \in H$ and $|g_\gamma) = U(\gamma\, |g\rangle) \in \hat{H}^{\oplus}_{\alpha,\mu}$, $|\alpha', g_\gamma) = |g_\gamma)(\alpha') \in \hat{H}(\alpha')$. For each specified value of α' and j, then, the operator γ generates a linear functional $\tilde{\gamma}_{\alpha',j} : H \to C$ by the relation $\tilde{\gamma}_{\alpha',j}[|g)] = (j, \alpha'|g_\gamma)$. If now we define another representation of H, say one in the set $F_H(\beta, \nu)$, where β is a set of M self-adjoint operators [in particular, we could choose this second representation to be the same as the one used in (40)], then it is clear that the operator γ also generates a family of linear functionals over this representation, one functional for each value of α' and j, specified by the relation: $\tilde{\gamma}_{\alpha',j}[(\beta' \mid g)] = (j, \alpha' \mid g_\gamma)$, where $(\beta' \mid g)$ represents the set of $d(\beta')$ functions corresponding to the vector $|g\rangle \in H$ in the new representation. By a slight abuse of notation we use the same symbols $\tilde{\gamma}_{\alpha',j}$ for the functionals defined over H and the ones defined over a representation of H.

In the special case that the operator γ commutes with all the operators in the set α we see from Eq. (8) that the functionals $\tilde{\gamma}_{\alpha',j}$ are given by $\tilde{\gamma}_{\alpha',j}[|g)] = \tilde{\gamma}_{\alpha',j}[(\beta \mid g)] = (j, \alpha'|\, \hat{\gamma}(\alpha')\, |\alpha', g)$, where $\hat{\gamma}$ is a family of operators $\hat{\gamma}(\alpha')$, one acting on each of the spaces $\hat{H}(\alpha')$; if γ is one of the operators in the set α, say α_i, then from Eq. (7) we have $\tilde{\gamma}_{\alpha',j}[|g)] = \tilde{\alpha}_{i,\alpha',j}[|g)] = \alpha_i'(j, \alpha'|g)$. However, in general we must deal with the functionals $\tilde{\gamma}_{\alpha',j}$ generated by an arbitrary operator γ by using the general theory of linear functionals on Hilbert space. In many cases it will be possible to express the functional $\tilde{\gamma}_{\alpha',j}$ as an integral operator on a function

space representation $F_H(\beta, \nu, B)$; then, in analogy with our previous notation, we can write

$$\tilde{\gamma}_{\alpha',j}[(\beta' \mid g)] = \int_{R^M} \sum_{k=1}^{d(\beta')} (\alpha', j| \hat{\gamma} |k, \beta')(\beta', k \mid g)\, d\nu(\beta'). \quad (41)$$

If the set β is maximal Abelian, then of course this reduces to

$$\tilde{\gamma}_{\alpha',j}[(\beta' \mid g)] = \int_{R^M} (\alpha', j| \hat{\gamma} |\beta')(\beta' \mid g)\, d\nu(\beta'). \quad (42)$$

If now we have specified two representations of a Hilbert space H, say $F_H(\alpha, \mu, B)$ and $F_H(\beta, \nu, D)$, where both α and β are maximal Abelian sets, the set of all operators γ such that the functionals $\tilde{\gamma}_{\alpha'}$ can be represented over the space $F_H(\beta, \nu, D)$ as

$$\tilde{\gamma}_{\alpha'}[(\beta' \mid g)] = \int_{R^M} (\alpha'| \gamma |\beta')(\beta' \mid g)\, d\nu(\beta') \quad (43)$$

is called the set of Hilbert–Schmidt operators and has been studied extensively.[16] It should be emphasized, however, that it is not always possible to find integral representations for operators (not even all bounded operators are of Hilbert–Schmidt type). As an elementary but important example we have the identity operator: the functional $\tilde{I}_{\alpha'}$ is very simply given on the function space $F_H(\alpha, \mu, B)$ by the delta functional $\delta_{\alpha'}$ according to the relationship: $\tilde{I}_{\alpha'}[(\alpha''|g)] = \delta_{\alpha'}[(\alpha''|g)] = (\alpha'|g)$, but there exists no kernel for an integral operator representation of I.

4. CONCLUSION

The principal advantage of the present formulation of spectral theory and representation theory lies in the fact that there is no need to introduce unnormalizable "vectors" of infinite length as the eigenvectors of operators with continuous spectra. By using direct integral decompositions we have seen that there exist true unit vectors which serve the same purposes, but these vectors are not in general in the original space H but instead in the decomposition spaces $\hat{H}(\alpha')$. With this modification the Dirac formulation of Hilbert space theory is rigorously brought into accord with the theory of von Neumann.

ACKNOWLEDGMENTS

The author wishes to express his sincere gratitude to Dr. Ralph Henderson for his continual interest and guidance during the work leading to this paper, and to Dr. Witold Bogdanowicz and Dr. Lawrence Conlon for many helpful discussions with regard to the mathematical formulation of the paper.

APPENDIX

To show how the methods presented here may be further applied and to make the paper more useful to quantum field theorists[17] we will develop an extension of von Neumann's resolutions of the identity. Given a Hilbert space H and a set α of N self-adjoint operators on H, we use the canonical diagonalization of Eq. (7) to write the bilinear functional $E(f, g, S)$, defined for every pair $|f\rangle, |g\rangle \in H$ and every measurable set $S \subset R^N$ (measurable, of course, with respect to the σ-ring used for the canonical diagonalization) by the equation

$$E(f, g, S) = \int_S (f, \alpha' \mid \alpha', g)\, d\mu(\alpha'). \quad (44)$$

Here, as previously, for $|h\rangle \in H$, $|\alpha', h) = |h)(\alpha') = U(|h\rangle)(\alpha') \in \hat{H}(\alpha')$, where $U : H \to \hat{H}^{\oplus}_{\alpha,\mu}$ is the unitary mapping of the canonical diagonalization. The mapping E is clearly a bounded bilinear functional and so defines a bounded linear operator E_S [18] by the equation

$$E(f, g, S) = \langle f| E_S |g\rangle. \quad (45)$$

From this definition of E_S we see that, if $\emptyset$ is the null set,

$$E_\emptyset = 0, \qquad E_{R^N} = I \quad \text{and} \quad \sum_{i=1}^{\infty} E_{S_i} = E_{\cup_i S_i}, \quad (46)$$

for any countable family of disjoint sets $\{S_i\}$ (i.e., for $S_j \cap S_k = \emptyset$ when $j \neq k$). We have, moreover, for any measurable sets $S_1, S_2, \subset R^N$,

$$\langle f| E_{S_1 \cap S_2} |g\rangle = \int_{S_1 \cap S_2} (f, \alpha' \mid \alpha', g)\, d\mu(\alpha') = \int_{S_1} Q_{S_2}(\alpha')(f, \alpha' \mid \alpha', g)\, d\mu(\alpha'),$$

where Q_{S_2} is the characteristic function of the set S_2. Now defining the family $|G) \in \hat{H}^{\oplus}_{\alpha,\mu}$ by

$$|G)(\alpha') = |\alpha', G) = Q_{S_2}(\alpha')\, |\alpha', g),$$

we get

$$\langle f| E_{S_1 \cap S_2} |g\rangle = \int_{S_1} (f, \alpha' \mid \alpha', G)\, d\mu(\alpha') = \langle f| E_{S_1} |G\rangle,$$

[16] N. Dunford and J. Schwartz, *Linear Operators, Part II: Spectral Theory* (Interscience Publishing Company, Inc., New York, 1963), p. 1009, ff.

[17] See the remarks of Streater and Wightman on p. 92 of Ref. 4. For another possible application of direct integral theory, cf. M. Guenin and B. Misra, Helvetica Physica Acta 37, 269 (1964).

[18] Reference 10, Vol. I, p. 42.

where $|G\rangle = U^{-1}(|G)$. But also

$$\langle f \mid G\rangle = \int_{R^N} (f, \alpha' \mid \alpha', G)\, d\mu(\alpha')$$

$$= \int_{R^N} Q_{S_s}(\alpha')(f, \alpha' \mid \alpha', g)\, d\mu(\alpha'),$$

and so

$$\langle f \mid G\rangle = \int_{S_s} (f, \alpha' \mid \alpha', g)\, d\mu(\alpha')$$

$$= \langle f|\, E_{S_s}\, |g\rangle,$$

or

$$|G\rangle = E_{S_s}\, |g\rangle.$$

Therefore

$$\langle f|\, E_{S_1 \cap S_s}\, |g\rangle = \langle f|\, E_{S_1} E_{S_s}\, |g\rangle. \tag{47}$$

Now (46) and (47) are the defining conditions for what can be called a *generalized resolution of the identity* (GRI), that is, a family E of operators E_S (they can easily be shown to be projection operators by the same sort of reasoning as we have used to develop their other properties), one for each measurable set $S \subset (R^N, \Sigma, \mu)$, which satisfy (46) and (47). Each GRI is also a projection valued measure over R^N (this statement can be understood in the sense that the bilinear functional $\langle f|\, E_S\, |g\rangle$ is always a well defined Lebesgue–Stieltjes measure over R^N, but it can also be given meaning directly in terms of the operators E_S themselves[19]), and integrals can be defined with respect to these measures by standard measure theoretical techniques. We can conclude, then, that to each set α of N self-adjoint operators on a Hilbert space H there corresponds a unique GRI, symbolized by $E(\alpha)$ and defined over R^N by the canonical diagonalization of α. Since the unitary mapping $U : H \to \hat{H}^{\oplus}_{\alpha,\mu}$ simultaneously diagonalizes all of the self-adjoint operators in α, any product $\prod_i \alpha_i$ of finitely many operators $\alpha_i \in \alpha$ can be written in terms of $E(\alpha)$ as

$$\prod_i \alpha_i = \int_{R^N} \prod_i \alpha'_i\, dE_{\alpha'}, \quad \alpha' \in R^N, \quad \alpha'_i \in R,$$

and, for a product of unitary operators of the form $e^{i\alpha_i}$, $\alpha_i \in \alpha$, we have

$$\prod_i e^{i\alpha_i} = \int_{R^N} \prod_i e^{i\alpha'_i}\, dE_{\alpha'}.$$

Finally, corresponding to Stone's theorem, for a product of one-parameter groups of unitary operators of the form $e^{it_i\alpha_i}$, $\alpha_i \in \alpha$, $t_i \in R$, we have

$$\prod_i e^{it_i\alpha_i} = e^{i\sum_i t_i\alpha_i}$$

$$= \int_{R^N} e^{i\sum_i t_i\alpha'_i}\, dE_{\alpha'}, \quad \alpha'_i \in R, \quad \alpha' \in R^N,$$

or, defining vectors $\mathbf{t} = (t_1, \cdots, t_K) \in R^K$ and $\mathbf{a} = (\alpha'_1, \cdots, \alpha'_K) \in R^K$,

$$e^{i\sum_i t_i\alpha_i} = \int_{R^N} e^{i\mathbf{t}\cdot\mathbf{a}}\, dE_{\alpha'}.$$

[19] See, for example, p. 888, ff., of Ref. 16.

Transformation from a Linear Momentum to an Angular Momentum Basis for Particles of Zero Mass and Finite Spin

H. E. Moses
MIT Lincoln Laboratory, Lexington, Massachusetts*
(Received 16 November 1964)

The infinitesimal generators of the inhomogeneous Lorentz group have been given in a basis in which the components of the linear momentum operators are diagonal and in another basis in which the square of the angular momentum is diagonal for all unitary irreducible ray representations of the group. In the present paper we show how the two bases are related for representations corresponding to zero mass and any (finite) spin. It will be shown how this relation enables one to integrate the infinitesimal generators in the angular momentum basis and thereby permits one to show how the angular momentum of a particle changes under the inhomogeneous Lorentz group. In particular, we study the way that the angular momentum of a massless particle of any spin appears in translated and moving frames of reference.

1. INTRODUCTION AND SUMMARY

THE ten infinitesimal generators of the proper, orthochronous, inhomogeneous Lorentz group will be denoted by H (the energy), P_i ($i = 1, 2, 3$—the components of the linear momentum), J_i (the components of the angular momentum), and $\mathcal{J}_i$ (the space–time infinitesimal generators). They satisfy the well-known commutation rules

$$\begin{aligned}
&[H, P_i] = 0 && [J_2, J_3] = iJ_1\\
&[P_i, P_j] = 0 && [J_3, J_1] = iJ_2\\
&[J_i, P_i] = 0 && [\mathcal{J}_1, \mathcal{J}_2] = -iJ_3\\
&[J_i, H] = 0 && [\mathcal{J}_2, \mathcal{J}_3] = -iJ_1\\
&[J_i, \mathcal{J}_i] = 0 && [\mathcal{J}_3, \mathcal{J}_1] = -iJ_2\\
&[\mathcal{J}_i, H] = iP_i && [J_1, P_2] = [P_1, J_2] = iP_3\\
&[\mathcal{J}_i, P_j] = i\delta_{ij}H && [J_2, P_3] = [P_2, J_3] = iP_1\\
&[J_1, J_2] = iJ_3 && [J_3, P_1] = [P_3, J_1] = iP_2\\
&[J_1, \mathcal{J}_2] = [\mathcal{J}_1, J_2] = i\mathcal{J}_3\\
&[J_2, \mathcal{J}_3] = [J_2, J_3] = i\mathcal{J}_1\\
&[J_3, \mathcal{J}_1] = [\mathcal{J}_3, J_1] = i\mathcal{J}_2. && (1.1)
\end{aligned}$$

The global form of the irreducible unitary ray representations was given by Wigner (Ref. 1) in terms of a basis in which the components of the linear momentum P_i are diagonal. The infinitesimal generators are also usually expressed in such a basis (see, e.g., Ref. 2). For certain physical applications, however, it is useful to work in a basis in which the operators H, $\mathbf{J}^2 = J_1^2 + J_2^2 + J_3^2$, J_3 and $\mathbf{P}\cdot\mathbf{J} = P_1J_1 + P_2J_2 + P_3J_3$ are diagonal. The way that the infinitesimal operators act in such a basis was given in Refs. 3 and 4. The principal objective of the present paper is to show how the two bases are related for the case of particles of zero mass and finite spin. We hope in a later paper to give the relation for the far more complicated case of nonzero mass and arbitrary spin.

Our objective can be reinterpreted in terms of more popular language as "expanding relativistic plane waves into relativistic spherical waves" and vice versa. However, the group-theoretical and Hilbert-space treatment used herein permits us to evade the notion of plane and spherical waves which would occur in a more naïve approach.

One might think that in order to obtain a basis in which the square of the angular momentum is diagonal, one would need to know the expansion before the basis was obtained. Such is the usual procedure in nonrelativistic problems. However, in Refs. 3 and 5 it is shown that the commutation rules (1.1) are sufficient to obtain the angular momentum basis. Knowing how the infinitesimal generators act in terms of the linear and angular momentum bases enables us to find the relation between them.

It will be useful to show how the infinitesimal generators act in both bases for the zero-mass case. In Ref. 1 it is shown that the irreducible representations of the infinitesimal generators are completely specified by the sign of the energy ϵ and by the number α where $\alpha = +s$ or $-s$ where s is the spin

* Operated with support of the U. S. Advanced Research Projects Agency.

[1] E. P. Wigner, Ann. Math. **40**, 149 (1939).

[2] Iu. M. Shirokov, Zh. Eksperim. i Teor. Fiz. **33**, 1196 (1957) [English transl.: Soviet Physics—JETP **6**, 919 (1958)].

[3] W. Pauli, CERN Rept. 56-31, Geneva (1956).

[4] J. S. Lomont and H. E. Moses, J. Math. Phys. **5**, 294 (1964).

[5] J. S. Lomont and H. E. Moses, J. Math. Phys. **5**, 1438 (1964).

of the particle. The number s is thus either a non-negative integer or a positive half-odd integer, while $\epsilon = +1$ or -1.

Let an abstract vector Φ in Hilbert space be represented by $f(\mathbf{p}) = f(p_1, p_2, p_3)$ in the linear momentum basis and by $F(E, j, m)$ in the angular momentum basis. The variables p_i are continuous in the range $-\infty < p_i < +\infty$. The variable E is continuous in the range $0 < E < \infty$ for $\epsilon = +1$ and $-\infty < E < 0$ for $\epsilon = -1$. The function $F(E, j, m)$ is identically zero except for $j = |\alpha|, |\alpha| + 1, |\alpha| + 2, \cdots$ and, for a given j, for $m = -j, -j+1, \cdots, j-1, j$.

The norm of Φ is

$$(\Phi, \Phi) = \int \frac{d\mathbf{p}}{p} |f(\mathbf{p})|^2 = \sum_{j,m} \int \frac{dE}{|E|} |F(E, j, m)|^2,$$

where $p = |\mathbf{p}|$. Accordingly we shall write

$$\Phi \underset{\mathbf{p}}{\leftrightarrow} f(\mathbf{p}) \underset{j}{\leftrightarrow} F(E, j, m).$$

If A is an abstract operator, then we write

$$A\Phi \underset{\mathbf{p}}{\leftrightarrow} (Af)(\mathbf{p}) \underset{j}{\leftrightarrow} (AF)(E, j, m)$$

to indicate how operators act in different bases.

In terms of the linear momentum basis we have (Ref. 6).

$$\begin{aligned}
(P_i f)(\mathbf{p}) &= p_i f(\mathbf{p}), \\
(Hf)(\mathbf{p}) &= \epsilon p f(\mathbf{p}), \\
(J_1 f)(\mathbf{p}) &= \left[-i(\mathbf{p} \times \nabla)_1 + \frac{p_1}{p + p_3} \alpha \right] f(\mathbf{p}), \\
(J_2 f)(\mathbf{p}) &= \left[-i(\mathbf{p} \times \nabla)_2 + \frac{p_2}{p + p_3} \alpha \right] f(\mathbf{p}), \\
(J_3 f)(\mathbf{p}) &= [-i(\mathbf{p} \times \nabla)_3 + \alpha] f(\mathbf{p}), \\
(\mathcal{J}_1 f)(\mathbf{p}) &= \epsilon \left[ip \frac{\partial}{\partial p_1} + \frac{p_2}{p + p_3} \alpha \right] f(\mathbf{p}), \\
(\mathcal{J}_2 f)(\mathbf{p}) &= \epsilon \left[ip \frac{\partial}{\partial p_2} - \frac{p_1}{p + p_3} \alpha \right] f(\mathbf{p}), \\
(\mathcal{J}_3 f)(\mathbf{p}) &= \epsilon ip \frac{\partial}{\partial p_3} f(\mathbf{p}),
\end{aligned} \tag{1.2}$$

where $(\mathbf{p} \times \nabla)_1 = p_2 \partial/\partial p_3 - p_3 \partial/\partial p_2$, etc.

The infinitesimal generators are given in terms of the angular momentum basis by[4]

$$\begin{aligned}
(HF)(E, j, m) &= EF(E, j, m), \\
(J_3 F)(E, j, m) &= mF(E, j, m), \\
((J_2 + iJ_1)F)(E, j, m) &= [(j - m)(j + m + 1)]^{\frac{1}{2}} F(E, j, m + 1), \\
((J_2 - iJ_1)F)(E, j, m) &= [(j + m)(j - m + 1)]^{\frac{1}{2}} F(E, j, m - 1),
\end{aligned}$$

$$\begin{aligned}
(P_3 F)(E, j, m) = |E| \Bigg[& \frac{m\alpha}{j(j+1)} F(E, j, m) \\
& + \frac{1}{j+1} \left[\frac{(j - m + 1)(j + m + 1)(j - \alpha + 1)(j + \alpha + 1)}{(2j + 1)(2j + 3)} \right]^{\frac{1}{2}} F(E, j + 1, m) \\
& + \frac{1}{j} \left[\frac{(j - m)(j + m)(j - \alpha)(j + \alpha)}{(2j - 1)(2j + 1)} \right]^{\frac{1}{2}} F(E, j - 1, m) \Bigg],
\end{aligned}$$

$$\begin{aligned}
((P_2 + iP_1)F)(E, j, m) = |E| \Bigg[& \frac{\alpha}{j(j+1)} [(j - m)(j + m + 1)]^{\frac{1}{2}} F(E, j, m + 1) \\
& - \frac{1}{(j+1)} \left[\frac{(j + m + 1)(j + m + 2)(j - \alpha + 1)(j + \alpha + 1)}{(2j + 1)(2j + 3)} \right]^{\frac{1}{2}} F(E, j + 1, m + 1) \\
& + \frac{1}{j} \left[\frac{(j - m - 1)(j - m)(j - \alpha)(j + \alpha)}{(2j - 1)(2j + 1)} \right]^{\frac{1}{2}} F(E, j - 1, m + 1) \Bigg],
\end{aligned}$$

$$\begin{aligned}
((P_2 - iP_1)F)(E, j, m) = |E| \Bigg[& \frac{\alpha}{j(j+1)} [(j + m)(j - m + 1)]^{\frac{1}{2}} F(E, j, m - 1) \\
& + \frac{1}{j+1} \left[\frac{(j - m + 1)(j - m + 2)(j - \alpha + 1)(j + \alpha + 1)}{(2j + 1)(2j + 3)} \right]^{\frac{1}{2}} F(E, j + 1, m - 1) \\
& - \frac{1}{j} \left[\frac{(j + m - 1)(j + m)(j - \alpha)(j + \alpha)}{(2j - 1)(2j + 1)} \right]^{\frac{1}{2}} F(E, j - 1, m - 1) \Bigg],
\end{aligned}$$

[4] J. S. Lomont and H. E. Moses, J. Math. Phys. 3, 405 (1962).

$$(\mathcal{J}_3F)(E, j, m) = \frac{im\alpha\,|E|}{j(j+1)}\frac{\partial}{\partial E}F(E, j, m)$$
$$+\frac{i}{j+1}\left[\frac{(j-m+1)(j+m+1)(j-\alpha+1)(j+\alpha+1)}{(2j+1)(2j+3)}\right]^{\frac{1}{2}}\left[(j+1)\epsilon + |E|\frac{\partial}{\partial E}\right]F(E, j+1, m)$$
$$-\frac{i}{j}\left[\frac{(j-m)(j+m)(j-\alpha)(j+\alpha)}{(2j-1)(2j+1)}\right]^{\frac{1}{2}}\left[j\epsilon - |E|\frac{\partial}{\partial E}\right]F(E, j-1, m),$$

$$((\mathcal{J}_2 + i\mathcal{J}_1)F)(E, j, m) = \frac{i\alpha\,|E|}{j(j+1)}[(j-m)(j+m+1)]^{\frac{1}{2}}\frac{\partial}{\partial E}F(E, j, m+1)$$
$$-\frac{i}{j+1}\left[\frac{(j+m+1)(j+m+2)(j-\alpha+1)(j+\alpha+1)}{(2j+1)(2j+3)}\right]^{\frac{1}{2}}\left[(j+1)\epsilon + |E|\frac{\partial}{\partial E}\right]F(E, j+1, m+1)$$
$$-\frac{i}{j}\left[\frac{(j-m-1)(j-m)(j-\alpha)(j+\alpha)}{(2j-1)(2j+1)}\right]^{\frac{1}{2}}\left[j\epsilon - |E|\frac{\partial}{\partial E}\right]F(E, j-1, m+1),$$

$$((\mathcal{J}_2 - i\mathcal{J}_1)F)(E, j, m) = \frac{i\alpha\,|E|}{j(j+1)}[(j+m)(j-m+1)]^{\frac{1}{2}}F(E, j, m-1)$$
$$+\frac{i}{(j+1)}\left[\frac{(j-m+1)(j-m+2)(j-\alpha+1)(j+\alpha+1)}{(2j+1)(2j+3)}\right]^{\frac{1}{2}}\left[(j+1)\epsilon + |E|\frac{\partial}{\partial E}\right]F(E, j+1, m-1)$$
$$+\frac{i}{j}\left[\frac{(j+m-1)(j+m)(j-\alpha)(j+\alpha)}{(2j-1)(2j+1)}\right]^{\frac{1}{2}}\left[j\epsilon - |E|\frac{\partial}{\partial E}\right]F(E, j-1, m-1). \tag{1.3}$$

We shall now give the relationship between the representatives $f(\mathbf{p})$ and $F(E, j, m)$ such that $(Af)(\mathbf{p})$ and $(AF)(E, j, m)$ have the same relation where A is any of the infinitesimal generators. Toward this end we introduce the Jacobi polynomials $P_n^{(\alpha,\beta)}(x)$ (for notation see, e.g., Ref. 7). For simplicity of notation we set

$$S(j, m, \alpha, x) = P_{j-m}^{(m-\alpha, m+\alpha)}(x) \tag{1.4}$$
$$p = |\mathbf{p}|, \tag{1.5}$$
$$\mathbf{p} = p(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta). \tag{1.6}$$

Then the relationship between the representatives in the two bases is

$$f(\mathbf{p}) = \pi^{-\frac{1}{2}}p^{-1}\left(1+\frac{p_3}{p}\right)^{\alpha}\sum_{j=|\alpha|}^{\infty}\sum_{m=-j}^{j}(\tfrac{1}{2})^{m+1}\exp\left(\frac{im\pi}{2}\right)$$
$$\times\left[\frac{(j-m)!\,(j+m)!\,(2j+1)}{(j-\alpha)!\,(j+\alpha)!}\right]^{\frac{1}{2}}\left(\frac{p_1+ip_2}{p}\right)^{m-\alpha}S(j, m, \alpha, p_3/p)F(\epsilon p, j, m), \tag{1.7}$$

$$F(E, j, m) = \frac{|E|}{\pi^{\frac{1}{2}}}\left(\frac{1}{2}\right)^{m+1}\exp\left(\frac{-im\pi}{2}\right)\left[\frac{(j-m)!\,(j+m)!\,(2j+1)}{(j-\alpha)!\,(j+\alpha)!}\right]^{\frac{1}{2}}$$
$$\times\int_0^{2\pi}d\varphi\int_0^{\pi}d\theta\,e^{-i(m-\alpha)\varphi}(1+\cos\theta)^{\alpha}S(j, m, \alpha, \cos\theta)(\sin\theta)^{m-\alpha+1}$$
$$\times f(|E|\sin\theta\cos\varphi, |E|\sin\theta\sin\varphi, |E|\cos\theta). \tag{1.8}$$

Furthermore,

$$\int\frac{d\mathbf{p}}{p}|f(\mathbf{p})|^2 = \sum_{j=|\alpha|}^{\infty}\sum_{m=-j}^{j}\int\frac{dE}{|E|}|F(E, j, m)|^2, \tag{1.9}$$

so that the norm and inner product are preserved in the two bases. Equations (1.7)–(1.9) are the principal mathematical results of the present paper. Much of the remainder of the paper deals with the verification of these results.

It is interesting to note that the functions $S(j, m, m', x)$ play an important role in the theory of the representations of finite rotations (see, e.g., Ref. 8.)

A formula which is useful for obtaining the func-

[7] G. Szegö, *Orthogonal Polynomials* (American Mathematical Society, Providence, Rhode Island, 1959), Colloquium Publications, Vol. 23, revised ed.

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

tions $S(j, m, \alpha, x)$ is

$$S(j, m, \alpha, x) = (-1)^{j+\alpha}\,[2^{m-j}/(j+m)!] \times (d^{j+m}/dx^{j+m})\,[(1-x)^{j+\alpha}(1+x)^{j-\alpha}]. \quad (1.10)$$

The formula is proved in Appendix I from a similar formula for the Jacobi polynomials. Another expression for $S(j, m, \alpha, x)$ which is useful is obtained from Rodrigues' formula for the Jacobi polynomials (p. 67, Ref. 7),

$$S(j, m, \alpha, x) = (-1)^{j-m}\frac{2^{m-j}}{(j-m)!}(1-x)^{-(m-\alpha)}(1+x)^{-(m+\alpha)} \times \frac{d^{j-m}}{dx^{j-m}}\,[(1-x)^{j-\alpha}(1+x)^{j+\alpha}]. \quad (1.11)$$

2. VERIFICATION OF THE TRANSFORMATION FORMULAS

We shall now verify the expressions (1.7)–(1.9). The properties of the Jacobi polynomials play an essential part. In particular four identities are useful for the case that α and β are integers. They are

$$dP_n^{(\alpha,\beta)}(x)/dx = \tfrac{1}{2}(n+\alpha+\beta+1)P_{n-1}^{(\alpha+1,\beta+1)}(x), \quad (2.1)$$

$$(1-x^2)\,dP_n^{(\alpha,\beta)}(x)/dx - [(\beta+\alpha)x - (\beta-\alpha)] \times P_n^{(\alpha,\beta)}(x) + 2(n+1)P_{n+1}^{(\alpha-1,\beta-1)}(x) = 0, \quad (2.2)$$

$$\{[2n+\beta+\alpha][2(n+1)+\beta+\alpha]x + \alpha^2 - \beta^2\} \times \{2n+1+\beta+\alpha\}P_n^{(\alpha,\beta)}(x) = 2(2n+\beta+\alpha)(n+1)(n+1+\beta+\alpha) \times P_{n+1}^{(\alpha,\beta)}(x) + 2[2(n+1)+\beta+\alpha] \times [n^2+n(\beta+\alpha)+\beta\alpha]P_{n-1}^{(\alpha,\beta)}(x), \quad (2.3)$$

$$(2n+1+\beta+\alpha)(1-x^2)\,dP_n^{(\alpha,\beta)}(x)/dx - \tfrac{1}{2}[(\beta+\alpha+2)x+\alpha-\beta] \times (2n+1+\beta+\alpha)P_n^{(\alpha,\beta)}(x) = [n^2+n(\beta+\alpha)+\beta\alpha]P_{n-1}^{(\alpha,\beta)}(x) - (n+\beta+\alpha+1)(n+1)P_{n+1}^{(\alpha,\beta)}(x). \quad (2.4)$$

The orthogonality relation for the Jacobi polynomials will also be needed

$$\int_{-1}^{1}(1-x)^{\alpha}(1+x)^{\beta}P_n^{(\alpha,\beta)}(x)P_m^{(\alpha,\beta)}(x)\,dx = \frac{2^{\alpha+\beta+1}}{2n+\alpha+\beta+1}\frac{(n+\alpha)!\,(n+\beta)!}{n!\,(n+\alpha+\beta)!}\,\delta_{mn}. \quad (2.5)$$

Equations (2.1) and (2.5) appear on p. 63 and p. 68, respectively, of Ref. 7. Equation (2.3) appears on p. 169 of Ref. 9. We shall prove (2.2) and (2.4) in Appendix I.

In terms of the functions $S(j, m, \alpha, x)$ the relations (2.1)–(2.5) become

$$dS(j, m, \alpha, x)/dx = \tfrac{1}{2}(j+m+1)S(j, m+1, \alpha, x), \quad (2.6)$$

$$(1-x^2)\,dS(j, m, \alpha, x)/dx - 2(mx-\alpha)S(j, m, \alpha, x) + 2(j-m+1)S(j, m-1, \alpha, x) = 0 \quad (2.7)$$

$$[xj(j+1) - m\alpha](2j+1)S(j, m, \alpha, x) = j[(j+1)^2 - m^2]S(j+1, m, \alpha, x) + (j^2-\alpha^2)(j+1)S(j-1, m, \alpha, x), \quad (2.8)$$

$$(2j+1)(1-x^2)\,dS(j, m, \alpha, x)/dx - [(m+1)x-\alpha](2j+1)S(j, m, \alpha, x) = (j^2-\alpha^2)S(j-1, m, x) - [(j+1)^2-m^2]S(j+1, m, \alpha, x), \quad (2.9)$$

$$\int_{-1}^{1}(1-x)^{m-\alpha}(1+x)^{m+\alpha}S(j, m, \alpha, x)S(j', m, \alpha, x)\,dx = \frac{2^{2m+1}}{2j+1}\frac{(j-\alpha)!\,(j+\alpha)!}{(j-m)!\,(j+m)!}\,\delta_{j,j'}. \quad (2.10)$$

We shall now prove our formulas (1.7)–(1.9). Let us consider all complex functions $F(E, j, m)$ where the ranges of the variables E, j, m are given in Sec. 1 such that

$$\sum_{j=|\alpha|}^{\infty}\sum_{m=-j}^{j}\int |F(E, j, m)|^2\,\frac{dE}{|E|} < \infty.$$

The entire set of such functions form a Hilbert space. Consider now the functions $f(\mathbf{p})$ *defined* by (1.7). From (2.10) such functions have a finite norm given by (1.9). Furthermore, the function $F(E, j, m)$ can be recovered from the function $f(\mathbf{p})$ using (2. 10), the expression for $F(E, j, m)$ being given by (1.8).

We shall now prove that if $F(E, j, m)$ is replaced by $(AF)(E, j, m)$ in (1.7), then $f(\mathbf{p})$ is replaced by $(Af)(\mathbf{p})$ where A is any of the infinitesimal generators.

Let us first take A to be H. Then, since

$$(HF)(E, j, m) = EF(E, j, m)$$

and

$$(Hf)(\mathbf{p}) = \epsilon p f(\mathbf{p}),$$

substituting into (1.7) leads to a trivial identity.

[9] A. Erdélyi *et al.*, *Higher Transcendental Functions* (McGraw-Hill Book Company, Inc., New York, 1953), Vol. 2, Chap. 10.

Next let $A = J_3$. Using (1.2) and (1.3) and substituting into (1.7) one again obtains an identity.

Now let us take $A = J_2 - iJ_1$. Then (1.7) is verified for Af and AF if one uses (2.6). When $A = J_2 + iJ_1$, one uses (2.7) to verify (1.7).

For the case $A = P_3$, the identity (2.8) is used, while for $A = \mathcal{J}_3$, the relations (2.8) and (2.9) are both used.

For the cases that $A = P_2 - iP_1$, $A = P_2 + iP_1$, $A = \mathcal{J}_2 + i\mathcal{J}_1$, and $A = \mathcal{J}_2 - \mathcal{J}_1$, it follows from the commutation rules for $J_2 \pm iJ_1$ with P_3 and $\mathcal{J}_3$ that (1.7) is valid.

We shall now prove that the space of functions $f(\mathbf{p})$ which are defined by (1.7) and which have the norm given by (1.9) is the *entire* Hilbert space of functions with the norm (1.9). For if this space were not the entire Hilbert space, it would be a proper invariant subspace under the entire set of infinitesimal generators. But this set of operators is irreducible. Thus the space of functions $f(\mathbf{p})$ defined by (1.7) is the entire Hilbert space. This completes our proof.

3. INTEGRATION OF THE INFINITESIMAL GENERATORS: APPLICATION TO KINEMATICS OF THE ANGULAR MOMENTUM OF MASSLESS PARTICLES

Let A be any of the infinitesimal generators. We are now in a position to find the finite generator $e^{i\beta A}$ where β is any real number as it appears in the angular momentum representation. That is, we can find $(e^{i\beta A}F)(E, j, m)$. We use the fact that $e^{i\beta A}f(\mathbf{p})$ can be found explicitly in terms of $f(\mathbf{p})$. Indeed, the finite generators were obtained in terms of the linear momentum basis before the infinitesimal generators (Ref. 1). Thus to find $(e^{i\beta A}F)(E, j, m)$, we obtain $f(\mathbf{p})$ from $F(E, j, m)$ by using (1.7) and $(e^{i\beta A}F)(E, j, m)$ from $(e^{i\beta A}f)(\mathbf{p})$ from (1.8).

We carry out this procedure for $A = P_3$ and for $A = \mathcal{J}_3$ because, as we now show, these generators give information about the way the probabilities of angular momentum change when one is in a frame of reference translated along the x_3-axis, or moving along the x_3-axis, respectively.

Let us assume that we are in a state $\Phi \leftrightarrow F(E, j, m)$, such that

$$\sum_{j,m} \int \frac{|F(E, j, m)|^2}{|E|}\, dE = 1.$$

Then if we let $\mathcal{I}$ be an interval along the E axis,

$$\int_{\mathcal{I}} \frac{|F(E, j, m)|^2}{|E|}\, dE,$$

is the probability that a simultaneous measurement of H, $\mathbf{J}^2$, and J_3 gives a probability that the value of the energy H is in the interval $\mathcal{I}$, that the value of $\mathbf{J}^2$ is $j(j+1)$ and that the value of J_3 is m.

Let us consider ourselves in a frame of reference whose origin is translated a distance in a direction which can be taken to be the positive z-axis of the original frame.

Then, denoting variables in the original frame of reference without a prime and those in the translated frame with a prime, the space–time coordinates in two frames would be related by the inhomogeneous Lorentz transformation

$$\begin{aligned} x_1' &= x_1 \\ x_2' &= x_2 \\ x_3' &= x_3 - a \\ t' &= t. \end{aligned} \tag{3.1}$$

In the new frame of reference the state is given by $\Phi' = \exp(iaP_3)\Phi$ and thus in the angular momentum representation by $F'(E, j, m) = (\exp(iaP_3)F)(E, j, m)$. As observed in the primed frame of reference, the probability that a simultaneous measurement of H, $\mathbf{J}^2$, and J_3 would yield a value of H in the interval $\mathcal{I}$, a value of $\mathbf{J}^2$ equal to $j(j+1)$ and a value of J_3 equal to m would be given by

$$\int_{\mathcal{I}} \frac{|F'(E, j, m)|^2}{|E|}\, dE.$$

In the next section we shall take the case that in the original frame of reference the state Φ is represented by

$$F(E, j, m) = g(E)\delta_{j,j_0}\delta_{m,m_0}, \tag{3.2}$$

where

$$\int \frac{|g(E)|^2}{|E|}\, dE = 1 \tag{3.2a}$$

and where $g(E)$ has a sharp peak at $E = E_0$ and vanishes for $E < E_0 - b$ and for $E > E_0 + b$, the real number b being taken as small.

Thus the state in the original frame of reference is an eigenstate of $\mathbf{J}^2$ and J_3 with eigenvalues $j_0(j_0+1)$ and m_0 respectively, and an approximate "eigenstate" of H with eigenvalue E_0. We shall then find the probability in the displaced frame of reference that a measurement of H gives a value in the interval $E_0 - b < E < E_0 + b$ and a measurement of $\mathbf{J}^2$ and J_3 a value of $j(j+1)$ and m, respectively.

We shall also consider the case where the second

frame of reference moves with respect to the first in accordance with the Lorentz transformation

$$\begin{aligned} t' &= t\cosh\omega + x_3\sinh\omega \\ x_1' &= x_1 \\ x_2' &= x_2 \\ x_3' &= x_3\cosh\omega + t\sinh\omega. \end{aligned} \tag{3.3}$$

In (3.3) the state in the primed frame of reference is given by $\Phi' = \exp(i\omega\mathcal{J}_3)\Phi$. Hence $F'(E, j, m) = \exp(i\omega\mathcal{J}_3)F)(E, j, m)$. We shall calculate in Sec. 5 probabilities analogous to those which we have discussed for translated frames of reference. We shall thereby have treated aspects of the kinematics of the angular momentum of quantized, relativistic particles of mass zero which hitherto have been neglected, for general spins largely because of the lack of mathematical apparatus to handle such problems, such apparatus having been provided for the first time in the earlier parts of the present paper.

4. THE CHANGE OF ANGULAR MOMENTUM OF A MASSLESS PARTICLE UNDER THE TRANSLATION OF FRAME OF REFERENCE

We first note that

$$(\exp(iaP_3)f)(\mathbf{p}) = \exp(iap_3)f(\mathbf{p}) \tag{4.1}$$

Using the procedure suggested in the previous section, it follows that

$$\begin{aligned} &(\exp(iaP_3)F)(E, j, m) \\ &= \left(\frac{1}{2}\right)^{2m+1}\left[\frac{(j-m)!\,(j+m)!\,(2j+1)}{(j-s)!\,(j+s)!}\right]^{\frac{1}{2}} \\ &\quad\times\sum_{j'=s}^{\infty} H_\alpha(a\,|E|, m: j, j') \\ &\quad\times\left[\frac{(j'-m)!\,(j'+m)!\,(2j+1)}{(j'-s)!\,(j'+s)!}\right]^{\frac{1}{2}} F(E, j', m), \end{aligned} \tag{4.2}$$

where, as before, $s = |\alpha|$ = spin of the particle and $H_\alpha(a, m: j, j')$ is defined by

$$\begin{aligned} H_\alpha(a, m: j, j') &= \int_0^\pi d\theta(\sin\theta)^{2(m-\alpha)+1}e^{ia\cos\theta} \\ &\quad\times(1+\cos\theta)^{2\alpha}S(j, m, \alpha, \cos\theta)S(j', m, \alpha, \cos\theta) \\ &= \int_{-1}^{1} dx(1-x)^{m-\alpha}(1+x)^{m+\alpha}e^{iax} \\ &\quad\times S(j, m, \alpha, x)S(j', m, \alpha, x). \end{aligned} \tag{4.3}$$

Thus the calculation of the function H_α shows us how the state changes under translations of frames of reference. Though we have been unable to obtain a general expression for this function, one can easily calculate this function for small values of α, m, j, j'. This function is just the Fourier transform of a polynomial and thus can be expressed in terms of spherical Bessel functions. We shall give some calculations shortly.

Let us now assume that in the original frame of reference the wavefunction is given by (3.2). It is easily seen that the probability that a measurement of H gives a value outside the interval $E_0 - b < E < E_0 + b$ is zero.

Likewise, the probability that a measurement of J_3 gives a value other than m_0 is zero. Hence the only probability of interest is the probability that a simultaneous measurement of H, $\mathbf{J}^2$, and J_3 gives a value of energy in the region $E_0 - b < E < E_0 + b$, of $\mathbf{J}^2$ equal to $j(j+1)$, and of J_3 equal to m_0. We shall designate this probability by $w_\alpha(a, E_0, m_0: j_0, j)$. Since we are assuming that $g(E)$ is highly peaked at E_0 and that b is very small, we may replace E which appears in $H_\alpha(a\,|E|, m_0: j, j')$ of (4.2) by E_0. Then from (3.2a) and the discussion in the previous section

$$\begin{aligned} w_\alpha(a, E_0, m_0: j_0, j) &= \left(\tfrac{1}{2}\right)^{2(2m_0+1)}|H_\alpha(a\,|E_0|, m_0: j, j_0)|^2 \\ &\quad\times\frac{(j-m_0)!\,(j+m_0)!\,(j_0-m_0)!\,(j_0+m_0)!\,(2j+1)(2j_0+1)}{(j-s)!\,(j+s)!\,(j_0-s)!\,(j_0+s)!}. \end{aligned} \tag{4.4}$$

The probability $w_\alpha(a, E_0, m_0: j_0, j)$ is thus the probability that if in the original frame the particle had an energy E_0, circular polarization α, z-component of angular momentum m_0, and the square of the angular momentum $j_0(j_0+1)$, then in a frame shifted by a distance a the energy is E_0, circular polarization is α, the z-component of angular momentum is m_0, and the square of the angular momentum is $j(j+1)$.

There are relationships between the various probabilities which we shall now state and prove. The relationships are

$$w_\alpha(a, E_0, m_0: j_0, j) = w_\alpha(a, E_0, m_0: j, j_0) \tag{4.5}$$

$$= w_\alpha(-a, E_0, m_0: j_0, j) \tag{4.6}$$

$$= w_{-\alpha}(a, E_0, m_0: j_0, j) \tag{4.7}$$

$$= w_\alpha(a, E_0, -m_0: j_0, j). \tag{4.8}$$

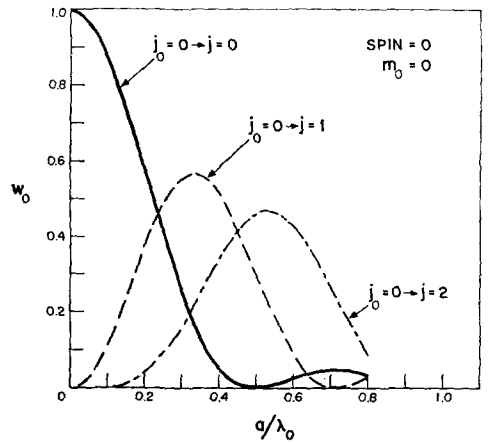


FIG. 1. Probabilities of angular momenta, scalar case.

These relations reduce considerably the number of calculations which one must carry out for a complete description of the kinematics.

Equations (4.5) and (4.6) follow immediately from (4.3) and (4.4). To prove (4.7) we use (1.10) to obtain

$$S(j, m, -\alpha, x) = (-1)^{j-\alpha} \frac{2^{m-j}}{(j+m)!} \frac{d^{j+m}}{dx^{j+m}} (1-x)^{j-\alpha}(1+x)^{j+\alpha}$$
$$= (-1)^{-2\alpha}(-1)^{j+m} S(j, m, \alpha, -x). \quad (4.9)$$

Then (4.7) follows from (4.3) and (4.4).

To prove (4.8) we again use (1.10) and (1.11) to show that

$$S(j, -m, \alpha, x) = 2^{-2m}(-1)^{j+\alpha}(1+x)^{m-\alpha} \times (1-x)^{m+\alpha} S(j, m, \alpha, -x). \quad (4.10)$$

Equation (4.8) then follows from (4.3) and (4.4).

We have derived explicit expressions for some of the probabilities w_α. To make these expressions more useful, we shall introduce units. (Hitherto, we have taken $\hbar = c = 1$.) Accordingly, E_0 is to be replaced by $E_0/\hbar c = 2\pi/\lambda_0$ where λ_0 is the de Broglie wave length of the particle.

For spin-zero cases we have:

$$w_0(a, E_0, 0: 0, 0) = j_0^2(2\pi a/\lambda_0),$$
$$w_0(a, E_0, 0: 0, 1) = 3j_1^2(2\pi a/\lambda_0),$$
$$w_0(a, E_0, 0: 0, 2) = 5j_2^2(2\pi a/\lambda_0), \quad (4.11)$$

where j_n is the nth spherical Bessel function.

For spin-$\frac{1}{2}$ cases (neutrino) we obtain

$$w_{\frac{1}{2}}(a, E_0, \tfrac{1}{2}: \tfrac{1}{2}, \tfrac{1}{2}) = j_0^2(2\pi a/\lambda_0) + j_1^2(2\pi a/\lambda_0)$$
$$w_{\frac{1}{2}}(a, E_0, \tfrac{1}{2}: \tfrac{1}{2}, \tfrac{3}{2}) = 2[j_1^2(2\pi a/\lambda_0) + j_2^2(2\pi a/\lambda_0)]. \quad (4.12)$$

For spin-1 particles (photons) we obtain

$$w_1(a, E_0, 1: 1, 1) = \frac{1}{4}\left\{2j_0\left(2\pi\frac{a}{\lambda_0}\right) - j_2\left(2\pi\frac{a}{\lambda_0}\right)\right]^2 + 9j_1^2\left(2\pi\frac{a}{\lambda_0}\right)\Big\}$$

$$w_1(a, E_0, 1: 1, 2) = \frac{15}{4}\left\{\frac{1}{25}\left[3j_1\left(2\pi\frac{a}{\lambda_0}\right) - 2j_3\left(2\pi\frac{a}{\lambda_0}\right)\right]^2 + j_2^2\left(2\pi\frac{a}{\lambda_0}\right)\right\}$$

$$w_1(a, E_0, 0: 1, 1) = \left[j_0\left(2\pi\frac{a}{\lambda_0}\right) + j_2\left(2\pi\frac{a}{\lambda_0}\right)\right]^2$$

$$w_1(a, E_0, 0: 1, 2) = \frac{9}{5}\left[j_1\left(2\pi\frac{a}{\lambda_0}\right) + j_3\left(2\pi\frac{a}{\lambda_0}\right)\right]^2. \quad (4.13)$$

Of course, in the case of photons the de Broglie wavelength is just wavelength of the light.

The simplicity of these expressions is remarkable. We have plotted these probabilities in Figs. 1–3,

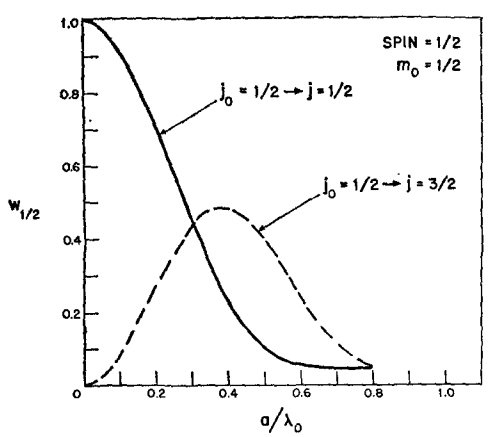


FIG. 2. Probabilities of angular momenta, neutrino.

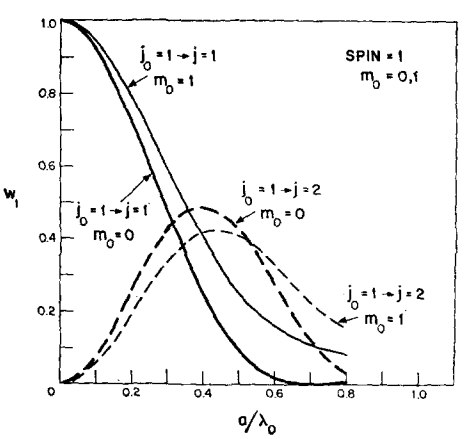


FIG. 3. Probabilities of angular momenta, photon.

using tabulated values for the spherical Bessel functions given in Ref. 10.

5. THE ANGULAR MOMENTUM OF A MASSLESS PARTICLE AS OBSERVED IN A MOVING FRAME OF REFERENCE

We shall now consider how the angular momentum of a massless particle changes when observed in a frame of reference related to the original frame of the homogeneous Lorentz transformation (3.3), give $(\exp(i\omega\mathcal{J}_3)F)(E, j, m)$, and finally obtain probabilities analagous to w_α of the previous section.

First of all, it will be necessary to obtain $(\exp(i\omega\mathcal{J}_3)f)(\mathbf{p})$ from our knowledge of $(\mathcal{J}_3 f)(\mathbf{p})$, i.e., to integrate the infinitesimal generator $\mathcal{J}_3$ in the linear momentum basis. In principle, it would be possible to read off the results from the global treatment of Ref. 1. If we used the results of Ref. 1, however, we should run into the problem of selecting the proper linear momentum basis such that the infinitesimal generators are given by (1.2) of the present paper (i.e., "choosing the phase"). Hence we prefer to integrate the infinitesimal generator directly. The technique which we use can be generalized considerably. It is based upon a theorem about solutions of simple first-order partial differential equations in two variables.

Let us define $g_\omega(\mathbf{p})$ by

$$g_\omega(\mathbf{p}) = (\exp(i\omega\mathcal{J}_3)f)(\mathbf{p}). \tag{5.1}$$

Let us then take a derivative with respect to ω. We obtain

$$\partial g_\omega(\mathbf{p})/\partial\omega = i(\mathcal{J}_3 \exp(i\omega\mathcal{J}_3)f)(\mathbf{p}) = i(\mathcal{J}_3 g_\omega)(\mathbf{p}). \tag{5.2}$$

On using (1.2) we obtain a partial differential equation for g_ω, namely

$$[\partial/\partial\omega + \epsilon p\partial/\partial p_3]g_\omega(\mathbf{p}) = 0. \tag{5.3}$$

Equation (5.3) is a partial differential equation in ω and p_3. The quantities p_1 and p_2 are merely parameters and are to be regarded as constants. Our "initial condition" is

$$g_0(\mathbf{p}) = f(\mathbf{p}), \tag{5.4}$$

which follows from (5.1). To simplify (5.3) we introduce the variable y in place of p_3 by

$$y = \log(p_3 + p). \tag{5.5}$$

In defining $F(p_1, p_2, y)$ by

$$g(\mathbf{p}) = F(p_1, p_2, y), \tag{5.6}$$

we obtain the very simple partial differential equation

$$[\partial/\partial\omega + \epsilon\partial/\partial y]F = 0. \tag{5.7}$$

The general solution of (5.7) is

$$F(p_1, p_2, y) = G(p_1, p_2, \omega - \epsilon y),$$

where $G(p_1, p_2, p_3)$ is a thus far arbitrary function of its arguments. From (5.6)

$$g(\mathbf{p}) = G(p_1, p_2, \omega - \epsilon\log(p_3 + p)). \tag{5.8}$$

This equation can also be written

$$g(\mathbf{p}) = H(p_1, p_2, e^\omega(p_3 + p)^{-\epsilon}), \tag{5.9}$$

where $H(p_1, p_2, p_3)$ is an arbitrary function of its arguments. To find this function H we use (5.4). Then

$$H(p_1, p_2, (p_3 + p)^{-\epsilon}) = f(\mathbf{p}) \tag{5.10}$$

or

$$H(p_1, p_2, p_3) = f(p_1, p_2, \tfrac{1}{2}[p_3^{-\epsilon} - (p_1^2 + p_2^2)p_3^{\epsilon}]). \tag{5.11}$$

On substituting (5.11) into (5.9) we obtain finally $(\exp(i\omega\mathcal{J}_3)f)(\mathbf{p}) = g_\omega(\mathbf{p})$

$$= f(p_1, p_2, p_3\cosh\omega - \epsilon p\sinh\omega). \tag{5.12}$$

Equation (5.12) is "obvious" but still required proof of some kind. As mentioned above, the technique used in this case can be extended to integrate more complicated infinitesimal generators.

We are now ready to give $(\exp(i\omega\mathcal{J}_3)F)(E, j, m)$ by transforming back and forth between the linear momentum and angular momentum representations as described in Sec. 3. Since we are interested in the case that the particle is in an eigenstate of angular momentum in the original frame of reference, we shall let $F(E, j, m)$ be given by (3.2). Then

$$(\exp(i\omega\mathcal{J}_3)F)(E, j, m) = (\tfrac{1}{4})^{m+\frac{1}{2}}e^{-\alpha\epsilon\omega}\left[\frac{(j-m)!\,(j+m)!\,(j_0-m)!\,(j_0+m)!\,(2j+1)(2j_0+1)}{(j-s)!\,(j+s)!\,(j_0-s)!\,(j_0+s)!}\right]^{\frac{1}{2}}\cdot\delta_{m,m_0}$$
$$\times\int_0^\pi d\theta(\cosh\omega - \epsilon\cos\theta\sinh\omega)^{-(m+1)}g(E\cosh\omega - \epsilon E\cos\theta\sinh\omega)$$
$$\times(1+\cos\theta)^{2\alpha}(\sin\theta)^{2(m-\alpha)+1}S(j, m, \alpha, \cos\theta)S\left(j_0, m, \alpha, \frac{\cosh\omega - \epsilon\cos\theta\sinh\omega}{\cos\theta\cosh\omega - \epsilon\sinh\omega}\right). \tag{5.13}$$

[10] National Bureau of Standards, *Tables of Spherical Bessel Functions* (Columbia University Press, New York, 1947).

It will be convenient to choose a new variable of integration. Thus we let

$$F = E(\cosh \omega - \epsilon \cos \theta \sinh \omega). \tag{5.14}$$

It is useful to note that the sign of F is the same as the sign of E. The variable of integration F should not be confused with the function $F(E, j, m)$.

$$(\exp (i\omega J_3)F)(E, j, m) = (\tfrac{1}{2})^{2m+1} \frac{\epsilon^{m+1} e^{-\alpha\epsilon\omega}}{|E|^m \sinh |\omega| \sinh^{2m} \omega} \times \left[\frac{(j-m)!\,(j+m)!\,(j_0-m)!\,(j_0+m)!\,(2j+1)(2j_0+1)}{(j-s)!\,(j+s)!\,(j_0-s)!\,(j_0+s)!}\right]^{\frac{1}{2}} \delta_{m,m_0} I_\alpha(\omega, E, m: j_0, j), \tag{5.15}$$

where

$$I_\alpha(\omega, E, m: j_0, j) = \int_{Ee^{-\epsilon|\omega|}}^{Ee^{\epsilon|\omega|}} \frac{dF}{F^{m+1}} \times (Ee^{\epsilon\omega} - F)^{m+\alpha}(F - Ee^{-\epsilon\omega})^{m-\alpha} g(F)\, S\left(j, m, \alpha, \frac{E\cosh\omega - F}{|E|\sinh\omega}\right) S\left(j_0, m, \alpha, \frac{E - F\cosh\omega}{|F|\sinh\omega}\right). \tag{5.16}$$

Let us define $P_\alpha(\omega, E, m_0: j_0, j)$ by

$$P_\alpha(\omega, E, m_0: j_0, j) = \frac{|(\exp (i\omega \mathcal{J}_3)F)(E, j, m_0)|^2}{|E|}. \tag{5.17}$$

Then

$$\int_{\mathcal{I}} P_\alpha(\omega, E, m_0: j_0, j)\, dE$$

is the probability that if in the original frame of reference the state was given by (3.2), in the moving frame the energy is in the interval $\mathcal{I}$ and a measurement of $\mathbf{J}^2$ and J_3 yields values of $j(j+1)$ and m_0 respectively. (The probability of finding a particle in the moving frame with a measurement of J_3 yielding a value other than m_0 is zero. Hence we need not consider this case.)

We have

$$P_\alpha(\omega, E, m: j_0, j) = (\tfrac{1}{4})^{2m+1} \frac{e^{-2\alpha\epsilon\omega}}{|E|^{2m+1} \sinh^{2(2m+1)} \omega} \times \frac{(j-m)!\,(j+m)!\,(j_0-m)!\,(j_0+m)!\,(2j+1)(2j_0+1)}{(j-s)!\,(j+s)!\,(j_0-s)!\,(j_0+s)!} |I_\alpha(\omega, E, m, : j_0, j)|^2. \tag{5.18}$$

(We have dropped the subscript 0 on m_0 for typographical simplicity.)

Without going into details, one can show by methods used to prove (4.5)–(4.8) that the following "recriprocity relations" are valid.

$$P_\alpha(\omega, E, -m; j_0, j) = P_{-\alpha}(\omega, E, m: j_0, j) = P_\alpha(-\omega, E, m: j_0, j). \tag{5.19}$$

It is not difficult to evaluate I_α and hence P_α for a large class of functions $g(E)$, since the remainder of the integrand is a rational function of the variable F.

The probability density P_α depends rather sensitively on the wavefunction $g(E)$ through I_α [Eq. (5.16)]. We shall see this dependence show up as a dependence on the "line breadth" for special classes of wavefunctions which we shall now consider.

Let us now consider functions $g(E)$ which satisfy (3.2a). We require that

$$g(E) = 0 \quad \text{if} \quad E < E_0 - b \quad \text{and if} \quad E > E_0 + b. \tag{5.20}$$

Subsequently, in order to make explicit calculations, we shall also require that b is very small and therefore it is reasonable to take $g(E)$ to be of one sign, say positive, in the interval $E_0 - b < E < E_0 + b$. Furthermore, we shall require that $|g(E)|^2/|E|$ be a constant in the interval $E_0 - b < E < E_0 + b$. This latter requirement is equivalent to the statement that the probability of finding the energy of the particle in a interval within $E_0 - b < E < E_0 + b$ is proportional to the length of the interval. This assumption is a common one in the theory of probability. These requirements lead to the following expression for $g(E)$.

$$g(E) = (|E|/2b)^{\frac{1}{2}} \text{ for } E_0 - b < E < E_0 + b = 0 \text{ for all other values of } E. \tag{5.21}$$

The wavefunction of a photon which is emitted by an atom would be approximated by $g(E)$. The

value E_0 would be the frequency of the photon and $2b$ gives the line breadth in frequency terms (in units such that $\hbar = c = 1$). Indeed, this interpretation motivates our choice of $g(E)$ as above.

For the moment, however, we shall consider functions $g(E)$ which satisfy (5.20) and not the more severe restrictions (5.21). It will be convenient to introduce nondimensional variables. Accordingly we define ν, $\tilde{\nu}$, σ, $h(\nu)$

$$E/E_0 = \nu,$$
$$b/|E_0| = \sigma,$$
$$F/E_0 = \tilde{\nu},$$
$$h(\nu) = g(E) = g(\nu E_0). \tag{5.22}$$

Equations (3.2a) and (5.20) become

$$\int_{1-\sigma}^{1+\sigma} \frac{|h(\nu)|^2}{\nu}\, d\nu = 1 \tag{5.23}$$

$$h(\nu) = 0 \quad \text{for} \quad \nu < 1 - \sigma \quad \text{or} \quad \nu > 1 + \sigma. \tag{5.24}$$

We also define the probability $p_\alpha(\omega, \nu, m: j_0, j)$ by

$$p_\alpha(\omega, \nu, m: j_0, j) = |E_0|\, P_\alpha(\omega, E, m: j_0, j). \tag{5.25}$$

The quantity

$$\int_{\mathcal{J}} p_\alpha(\omega, \nu, m: j_0: j)\, d\nu$$

is the probability that in the moving frame the measurement of the nondimensional energy E/E_0 will yield a value in the interval $\mathcal{J}$ while measurements of $\mathbf{J}^2$ and J_3 will yield values of $j(j+1)$ and m, respectively.

Then from (5.18) we obtain

$$p_\alpha(\omega, \nu, m: j_0, j) = (\tfrac{1}{4})^{2m+1} \frac{e^{-2\alpha\epsilon\omega}}{\nu^{2m+1} \sinh^{2(2m+1)}\omega} \times \frac{(j-m)!\,(j+m)!\,(j_0-m)!\,(j_0+m)!}{(j-s)!\,(j+s)!\,(j_0-s)!\,(j_0+s)!} \times |\hat{I}_\alpha(\omega, \nu, m: j_0, j)|^2, \tag{5.26}$$

where

$$\hat{I}_\alpha(w, \nu, m: j_0, j) = \int_{\nu e^{-|\omega|}}^{\nu e^{|\omega|}} \frac{h(\tilde{\nu})}{\tilde{\nu}^{m+1}} (\nu e^{\epsilon\omega} - \tilde{\nu})^{m+\alpha} \times (\tilde{\nu} - \nu e^{-\epsilon\omega})^{m-\alpha} S\left(j, m, \alpha, \frac{\nu \cosh\omega - \tilde{\nu}}{\nu \sinh \epsilon\omega}\right) \times S\left(j_0, m, \alpha, \frac{\nu - \tilde{\nu}\cosh\omega}{\tilde{\nu}\sinh\epsilon\omega}\right) d\tilde{\nu}. \tag{5.27}$$

The requirement (5.24) on $h(\nu)$ will, as we shall now see, lead to the Doppler shift in the energy. For some values of ν the integrand of (5.27) is identically zero. To eliminate that portion of the integrand

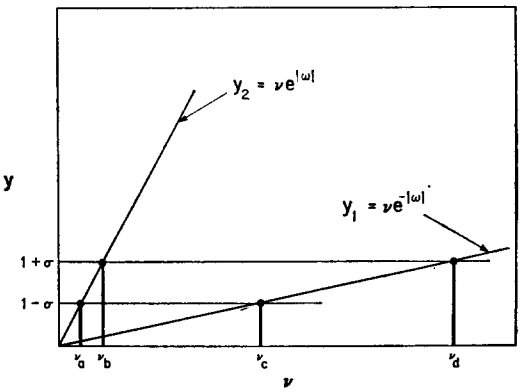


FIG. 4. Definition of ν_i.

which is identically zero, we must consider the limits of the integral as functions of ν. It will be convenient to introduce the following functions of ν:

$$y_1(\nu) = \nu e^{-|\omega|},$$
$$y_2(\nu) = \nu e^{+|\omega|},$$
$$y_3(\nu) = 1 - \sigma, \tag{5.28}$$
$$y_4(\nu) = 1 + \sigma.$$

These functions are plotted as curves in Fig. 4. These curves are all straight lines. The lines corresponding to the functions y_1 and y_2 go through the origin, the former curve having a slope less than 1 and the latter having a slope greater than 1. The curves given by y_3 and y_4, being constants, are straight lines parallel to the ν-axis. For a given value of ν, the integration goes from y_1 to y_2.

Let us define ν_a as the value of ν at which the curves y_2 and y_3 intersect. Likewise ν_b is the value of ν for which the curves y_2 and y_4 intersect, ν_c is the value at which y_1 and y_3 intersect and ν_d is the value at which y_1 and y_4 intersect. It is easily shown that

$$\nu_a = (1 - \sigma)e^{-|\omega|},$$
$$\nu_b = (1 + \sigma)e^{-|\omega|},$$
$$\nu_c = (1 - \sigma)e^{|\omega|}, \tag{5.29}$$
$$\nu_d = (1 + \sigma)e^{|\omega|}.$$

There are two possible orderings of the values ν_i.

$$\nu_a < \nu_b < \nu_c < \nu_d \quad \text{if} \quad \sigma < \tanh|\omega|$$
$$\nu_a < \nu_c < \nu_b < \nu_d \quad \text{if} \quad \sigma > \tanh|\omega|. \tag{5.30}$$

In Fig. 4 we have taken $\sigma < \tanh|\omega|$. In either of the two cases (5.30) it follows that

$$\hat{I}_\alpha(\omega, \nu, m; j_0, j) = 0 \quad \text{if} \quad \nu < \nu_a \quad \text{or if} \quad \nu > \nu_d. \tag{5.31}$$

Hence p_α vanishes for $\nu < \nu_a$ or $\nu > \nu_d$. Thus the Doppler breadth of the "line" is $\nu_d - \nu_a = 2(\sinh |\omega| + \sigma \cosh \omega)$ as contrasted to the "natural breadth" 2σ. The "center of the line" is $\nu = \cosh \omega - \sigma \sinh |\omega|$ as contrasted to $\nu = 1$. If the natural line breadth $\sigma = 0$, then one obtains the familiar results about the Doppler shift in energy and line breadth. It is seen, however, that if σ is of the same order as $|\omega|$, one will obtain quite different results than is given by the usual theory which assumes monochromatic light and hence $\sigma = 0$. In the above discussion we are extending the language which one uses for the photon case.

Because some of the integrand in (5.27) is zero we may replace the lower and upper limits of the integrand in (5.27) by $z_1(\nu)$ and $z_2(\nu)$, respectively

$$\begin{aligned} z_1(\nu) &= 1 - \sigma && \text{for } \nu_a < \nu < \nu_c, \\ z_1(\nu) &= y_1(\nu) = \nu e^{-|\omega|} && \text{for } \nu_c < \nu < \nu_d, \\ z_2(\nu) &= y_2(\nu) = \nu e^{|\omega|} && \text{for } \nu_a < \nu < \nu_b, \\ z_2(\nu) &= 1 + \sigma && \text{for } \nu_b < \nu < \nu_d. \end{aligned} \tag{5.32}$$

Two types of probabilities will prove useful for discussion. We shall denote by $W_\alpha(\omega, m: j_0, j)$ the total probability that in the moving frame of reference a measurement of $\mathbf{J}^2$ and J_3 will yield a value $j(j+1)$ and m, respectively. By $V_\alpha(\omega, m: j_0, j)$ we mean the probability that in the moving frame of reference a measurement of H, will yield a value in the range $E_0 - b < E < E_0 + b$, and measurements of $\mathbf{J}^2$ and J_3 will yield values of $j(j+1)$ and m, respectively. Thus

$$\begin{aligned} W_\alpha(\omega, m: j_0, j) &= \int_{\nu_a}^{\nu_d} p_\alpha(\omega, \nu, m: j_0, j)\, d\nu, \\ V_\alpha(\omega, m: j_0, j) &= \int_{1-\sigma}^{1+\sigma} p_\alpha(\omega, \nu, m: j_0, j)\, d\nu. \end{aligned} \tag{5.33}$$

Let us now consider states which satisfy (5.21). For these states

$$\begin{aligned} h(\nu) &= (\nu/2\sigma)^{\frac{1}{2}} \qquad 1 - \sigma < \nu < 1 + \sigma \\ &= 0, \qquad \nu < 1 - \sigma \quad \text{or} \quad \nu > 1 + \sigma. \end{aligned} \tag{5.34}$$

In order to give an explicit calculation we shall consider a spin-zero particle with $j_0 = j = m_0 = m = 0$. The probabilities are particularly easy to compute. First we take $\sigma > \tanh |\omega|$. Then

$$\begin{aligned} p_0(\omega, \nu, 0: 0, 0) &= 0 \quad \text{for} \quad \nu < \nu_a \\ &= \tfrac{1}{2}(1/\nu\sigma \sinh^2 \omega)[\nu e^{|\omega|} - 2\nu^{\frac{1}{2}} e^{\frac{1}{2}|\omega|} \\ &\quad \times (1 - \sigma)^{\frac{1}{2}} + (1 - \sigma)] \quad \text{for} \quad \nu_a < \nu < \nu_c \\ &= (2/\sigma \sinh^2 \omega) \sinh^2 \tfrac{1}{2}\omega \quad \text{for} \quad \nu_c < \nu < \nu_b \\ &= \tfrac{1}{2}(1/\sigma\nu \sinh^2 \omega)[\nu e^{-|\omega|} - 2\nu^{\frac{1}{2}} e^{-\frac{1}{2}|\omega|} \\ &\quad \times (1 + \sigma)^{\frac{1}{2}} + (1 + \sigma)] \quad \text{for} \quad \nu_b < \nu < \nu_d, \\ &= 0 \quad \text{for} \quad \nu > \nu_d. \end{aligned} \tag{5.35}$$

Also for $\sigma > \tanh |\omega|$,

$$\begin{aligned} W_0(\omega, 0: 0, 0) &= (2/\sigma \sinh^2 \omega) \\ &\quad \times \{[|\omega| - \sinh |\omega|] + \sigma[\cosh \omega - 1]\} \end{aligned} \tag{5.36}$$

$$\begin{aligned} V_0(\omega, 0: 0, 0) &= (1/\sigma \sinh^2 \omega) \\ &\quad \times \{|\omega| - 3 \sinh |\omega| + 4 \sinh \tfrac{1}{2}|\omega| \\ &\quad + \sigma[1 + 3 \cosh \omega - 4 \cosh \tfrac{1}{2}\omega]\}. \end{aligned} \tag{5.37}$$

For $\sigma < \tanh |\omega|$

$$\begin{aligned} p_0(\omega, \nu, 0: 0, 0) &= 0 \quad \text{for} \quad \nu < \nu_a \\ &= (1/2\sigma\nu \sinh^2 \omega)[\nu e^{|\omega|} - 2\nu^{\frac{1}{2}} e^{\frac{1}{2}|\omega|} \\ &\quad \times (1 - \sigma)^{\frac{1}{2}} + (1 - \sigma)] \quad \text{for} \quad \nu_a < \nu < \nu_b, \\ &= (1/\sigma\nu \sinh^2 \omega)[1 - (1 - \sigma^2)^{\frac{1}{2}}] \\ &\qquad\qquad \text{for} \quad \nu_b < \nu < \nu_c \\ &= (1/2\sigma\nu \sinh^2 \omega)[\nu e^{-|\omega|} - 2(1 + \sigma)^{\frac{1}{2}} \nu^{\frac{1}{2}} \\ &\quad \times e^{-\frac{1}{2}|\omega|} + (1 + \sigma)] \quad \text{for} \quad \nu_c < \nu < \nu_d \\ &= 0 \quad \text{for} \quad \nu > \nu_d. \end{aligned} \tag{5.38}$$

Also for $\sigma < \tanh |\omega|$

$$\begin{aligned} W_0(\omega, 0: 0, 0) &= (1/\sigma \sinh^2 \omega)\{2 |\omega| [1 - (1 - \sigma^2)^{\frac{1}{2}}] \\ &\quad - 2\sigma + (1 - \sigma^2)^{\frac{1}{2}} \log [(1 + \sigma)/(1 - \sigma)]\}, \end{aligned} \tag{5.39}$$

while for $\tanh \tfrac{1}{2}|\omega| < \sigma < \tanh |\omega|$

$$\begin{aligned} V_0(\omega, 0: 0, 0) &= (1/\sigma \sinh^2 \omega)\{|\omega| [1 - 2(1 - \sigma^2)^{\frac{1}{2}}] \\ &\quad - \sinh |\omega| + 4 \sinh \tfrac{1}{2}|\omega| \\ &\quad + \sigma(1 + \cosh \omega - 4 \cosh \tfrac{1}{2}\omega) \\ &\quad + (1 - \sigma^2)^{\frac{1}{2}} \log [(1 + \sigma)/(1 - \sigma)]\}, \end{aligned} \tag{5.40}$$

and for $\sigma < \tanh \tfrac{1}{2}|\omega|$,

$$\begin{aligned} V_0(\omega, 0: 0, 0) &= (1/\sigma \sinh^2 \omega) \\ &\quad \times [1 - (1 - \sigma^2)^{\frac{1}{2}}] \log [(1 + \sigma)/(1 - \sigma)]. \end{aligned} \tag{5.40}$$

The expressions for W_0 and V_0 simplify considerably if we take $\sigma \ll 1$ and $\omega \ll 1$.

It will be useful to regard σ as fixed and ω as varying. It can be shown from the exact expressions that

$$\begin{aligned} W_0(\omega, 0: 0, 0) &= 1 - |\omega|/3\sigma \quad \text{for} \quad |\omega| < \sigma \\ &= \sigma/|\omega| \quad \text{for} \quad |\omega| > \sigma. \end{aligned} \tag{5.41}$$

It is interesting to note that W_0 decreases with increasing ω until it reaches a value of $\frac{2}{3}$. Then a jump occurs at $|\omega| = \sigma$ to the value 1, after which the probability decreases. Thus there is a "resonance" for $|\omega| = \sigma$ (i.e., where the Doppler half breadth equals the natural half breadth). The discontinuity of W_0 occurs because we have worked in the limit of small quantities.

V_0 exhibits a similar behavior

$$\begin{aligned} V_0(\omega, 0{:}\,0, 0) &= 1 - 5\,|\omega|/12\sigma \quad \text{for} \quad |\omega| < \sigma \\ &= \sigma/|\omega| \quad \text{for} \quad \sigma < |\omega| < 2\sigma \\ &= \sigma^2/|\omega|^2 \quad \text{for} \quad |\omega| > 2\sigma. \end{aligned} \tag{5.42}$$

It is possible to obtain a general expression for W_0 and V_0 for the asymptotic case $\sigma \ll |\omega| \ll 1$, $(j + |m|)\,|\omega| \ll 1$, $(|\alpha| + |m|)\,|\omega| \ll 1$. In the expressions of $\hat{I}_\alpha$ one first lets $\sigma \to 0$ and uses the facts that $\nu_a \to \nu_b \to e^{-|\omega|}$ and $\nu_c \to \nu_d e^{|\omega|}$. The quantities ν and $\bar{\nu}$ are replaced by 1 and $e^{\epsilon\omega} \to 1 + \epsilon\omega$.

Without going into detail one obtains after some juggling

$$\begin{aligned} W_\alpha(\omega, m{:}\,j_0, j) &= (\tfrac{1}{2})^{4m}[S(j_0, m, \alpha, 0)]^2 \\ &\quad \times [S(j, m, \alpha, 0)]^2\sigma/|\omega| \end{aligned} \tag{5.43}$$

$$\begin{aligned} V_\alpha(\omega, m{:}\,j_0, j) &= (\tfrac{1}{2})^{4m}[S(j_0, m, \alpha, 0)]^2 \\ &\quad \times [S(j, m, \alpha, 0)]^2\sigma^2/|\omega|^2. \end{aligned}$$

By comparing these results with the exact results of the scalar case treated earlier, these results appear to be valid when $\sigma < \frac{1}{2}\,|\omega| \ll 1$.

Since we have already treated the scalar case, we shall consider some results for neutrons and photons.

$$\begin{aligned} W_{\frac{1}{2}}(\omega, \pm\tfrac{1}{2}{:}\,\tfrac{1}{2}, \tfrac{1}{2}) &= \tfrac{1}{4}\sigma/|\omega| \\ V_{\frac{1}{2}}(\omega, \pm\tfrac{1}{2}{:}\,\tfrac{1}{2}, \tfrac{1}{2}) &= \tfrac{1}{4}\sigma^2/|\omega|^2 \\ W_{\frac{1}{2}}(\omega, \pm\tfrac{1}{2}{:}\,\tfrac{1}{2}, \tfrac{3}{2}) &= \tfrac{1}{16}\sigma/|\omega| \\ W_{\frac{1}{2}}(\omega, \pm\tfrac{1}{2}{:}\,\tfrac{1}{2}, \tfrac{3}{2}) &= \tfrac{1}{16}\sigma^2/|\omega|^2 \end{aligned} \tag{5.44}$$

$$W_1(\omega, 0{:}\,1, 1) = W_1(\omega, \pm 1{:}\,1, 1) = \sigma/|\omega|$$

$$V_1(\omega, 0{:}\,1, 1) = V_1(\omega, \pm 1{:}\,1, 1) = \sigma^2/|\omega|^2. \tag{5.45}$$

APPENDIX I: SOME PROPERTIES OF JACOBI POLYNOMIALS

In this Appendix we prove some useful properties of the Jacobi polynomials.

Theorem: Let n be a nonnegative integer and α and β be integers such that $n + \alpha \geq 0$, $n + \beta \geq 0$. $n + \alpha + \beta \geq 0$. Then the Jacobi polynomial $P_n^{(\alpha,\beta)}(x)$ is given by

$$\begin{aligned} P_n^{(\alpha,\beta)}(x) &= \frac{(-1)^{n+\beta}}{2^n(n + \alpha + \beta)!}\,\frac{d^{n+\alpha+\beta}}{dx^{n+\alpha+\beta}} \\ &\quad \times [(1 - x)^{n+\alpha}(1 - x)^{n+\beta}]. \end{aligned} \tag{AI.1}$$

Proof: We multiply the identity

$$\begin{aligned} &[(2n + \alpha + \beta)x + \beta - \alpha] \\ &\quad \times [(2n + \alpha + \beta - 2)x + \beta - \alpha] \\ &\quad - (1 - x^2)(2n + \alpha + \beta) \\ &\quad - [(2n + \alpha + \beta)x + \beta - \alpha] \\ &\quad \times [(2n + \alpha + \beta - 2)x + \beta - \alpha] \\ &\quad + (2n + \alpha + \beta)(1 - x^2) = 0 \end{aligned} \tag{AI.2}$$

by $(1 + x)^{n+\alpha-1}(1 - x)^{n+\beta-1}$. On using

$$\begin{aligned} &(d/dx)[(1 + x)^{n+\alpha}(1 - x)^{n+\beta}] = -(1 + x)^{n+\alpha-1} \\ &\times (1 - x)^{n+\beta-1}[(2n + \alpha + \beta)x + \beta - \alpha] \end{aligned} \tag{AI.3}$$

and

$$\begin{aligned} &(d^2/dx^2)[(1 + x)^{n+\alpha}(1 - x)^{n+\beta}] \\ &\quad = (1 + x)^{n+\alpha-2}(1 - x)^{n+\beta-2} \\ &\quad \times \{[(2n + a + \beta)x + \beta - \alpha] \\ &\quad \times [(2n + \alpha + \beta - 2)x + \beta - \alpha] \\ &\quad - (1 - x^2)(2n + \alpha + \beta)\}, \end{aligned} \tag{AI.4}$$

we obtain

$$\begin{aligned} &(1 - x^2)(d^2/dx^2)[(1 + x)^{n+\alpha}(1 - x)^{n+\beta}] \\ &+ [(2n + \alpha + \beta - 2)x + \beta - \alpha] \\ &\times (d/dx)[(1 + x)^{n+\alpha}(1 - x)^{n+\beta}] \\ &+ (2n + \alpha + \beta)(1 + x)^{n+\alpha}(1 - x)^{n+\beta} = 0. \end{aligned} \tag{AI.5}$$

We differentiate this equation $n + \alpha + \beta$ times. After some rearrangement of terms we find

$$\begin{aligned} &(1 - x^2)[d^2Q(x)/dx^2] \\ &\quad + [\beta - \alpha - (\alpha + \beta - 2)x]\,dQ(x)/dx \\ &\quad + n(n + \alpha + \beta + 1)Q(x) = 0, \end{aligned}$$

where

$$Q(x) = (d^{n+\alpha+\beta}/dx^{n+\alpha+\beta})[(1 + x)^{n+\alpha}(1 - x)^{n+\beta}].$$

Thus $Q(x)$ satisfies the same differential equation as the Jacobi polynomial $P_n^{(\alpha,\beta)}(x)$. It thus follows that

$$P_n^{(\alpha,\beta)}(x) = CQ(x),$$

where C is a constant. To determine C we compare the coefficient of the highest power of x in $Q(x)$

with the highest power of x in $P_n^{(\alpha,\beta)}(x)$. The latter can be obtained from Rodrigues' formula (Ref. 7). The theorem then follows. Expression (AI.1) appears to be a new result.

From the theorem, Eqs. (1.10) and (2.1) can be obtained.

Proof of equation (2.2): We differentiate (AI.5) $n + \alpha + \beta - 1$ times. On using (AI.1), Eq. (2.2) follows.

Proof of equation (2.4): We start with the somewhat formidable identity

$$2[(2n + \alpha + \beta)x + \beta - \alpha][(2n + \alpha + \beta - 2)x + \beta - \alpha](2n + \alpha + \beta + 1) - 2(2n + \alpha + \beta + 1)$$
$$\times (2n + \alpha + \beta)(1 - x^2) - (2n + \alpha + \beta + 1)[(4n + 3\alpha + 3\beta - 6)x + \beta - \alpha][(2n + \alpha + \beta)x + \beta - \alpha]$$
$$- (2n + \alpha + \beta + 1)(n + \alpha + \beta - 1)(2n + \alpha + \beta - 2)(1 - x^2) = -4[n^2 + n(\alpha + \beta) + \alpha\beta](n + \alpha + \beta)$$
$$+ (n + 1)[(2n + \alpha + \beta + 2)x + \beta - \alpha][(2n + \alpha + \beta)x + \beta - \alpha] - (n + 1)(1 - x^2)(2n + \alpha + \beta + 2).$$

We multiply this identity through by $(1+x)^{n+\alpha-1} \times (1 - x)^{n+\beta-1}$ and use Eqs. (AI.3) and (AI.4) to obtain

$$2(2n + \alpha + \beta + 1)(1 - x^2)$$
$$\times (d^2/dx^2)[(1 + x)^{n+\alpha}(1 - x)^{n+\beta}] + (2n + \alpha + \beta + 1)[(4n + 3\alpha + 3\beta - 6)x + \beta - \alpha]$$
$$\times (d/dx)[(1 + x)^{n+\alpha}(1 - x)^{n+\beta}] - (2n + \alpha + \beta + 1)(n + \alpha + \beta - 1)(2n + \alpha + \beta - 2)$$
$$\times [(1 + x)^{n+\alpha}(1 - x)^{n+\beta}] = -4[n^2 + n(\alpha + \beta) + \alpha\beta)](n + \alpha + \beta)$$
$$\times [(1 + x)^{n+\alpha-1}(1 - x)^{n+\beta-1}] + (n + 1)(d^2/dx^2)[(1 + x)^{n+\alpha+1}(1 - x)^{n+\beta+1}].$$

One differentiates the above equation $n+\alpha+\beta-1$ times. On using (AI.1), Eq. (2.4) follows.

Calculation of Some Homology Groups Relevant to Sixth-Order Feynman Diagrams*

PAUL FEDERBUSH†
Department of Physics and the Laboratory of Nuclear Science, Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received 18 August 1964)

A homology group that determines an upper bound for the number of linearly independent analytic functions connected with the sixth-order ladder diagram is here computed. The formalism is that of Fotiadi, Froissart, Lascoux, and Pham. Calculations use the standard methods of homology theory. We find that there are at most 127 such functions in general.

INTRODUCTION

A STUDY of the analytic properties of individual Feynman diagrams, has recently been initiated by Fotiadi, Froissart, Lascoux, and Pham. This study is more detailed than such investigations in the past. The idea is to obtain complete information about the amplitudes on all sheets, the discontinuities around all singularities. The works in Ref. 1 deal almost exclusively with diagrams with only a single internal loop, and for this case evidence quite complete results. When there is more than one internal loop things are apparently much more difficult.

We undertake to solve some portion of the problems associated with a particular Feynman diagram containing two internal loops. We essentially limit ourselves to calculating a relevant homology group. The rank of the homology group puts an upper bound on the number of invariant functions associated with the diagram. (If one considers a contour integral involving a single complex variable, then whenever there is a pole in the integrand the value of the integral may be changed by integrating around the pole by a different path. The rank of the homology group calculated is a generalization of the number of such poles, to an integral involving more than one complex variable.) The calculation of this group is rather technical. The methods used are probably sufficient to calculate the homology groups related to more complicated diagrams.

The present work probably is most important in so far as it indicates the type of mathematical problems one encounters in the homology approach to Feynman diagrams. A vast amount of effort must still be expended to obtain for this diagram the extent of understanding we have of the single-loop cases. Aside from further work on this particular diagram, other directions of investigation are to be suggested. Possibly there is a better compact manifold to perform the integrals in than the one we selected. Maybe there is a better way to calculate the homology groups, using the cohomology of holomorphic differential forms, for example. Many possibilities suggest themselves, and the mathematical tools useful may involve physicists with branches of mathematics heretofore largely unknown to them.

MATHEMATICAL TOOLS AND NOTATION

$H_i(A, B)$, $H^i(A, B)$ the singular homology and cohomology groups of the pair A, B ($B \subset A$) omitting torsion.[2]

T1. Exact sequence of a triple

$$\to H_{i+1}(A, B) \to H_i(B, C) \to H_i(A, C) \to$$
$$\to H_i(A, B) \to H_{i-1}(B, C) \to,$$
$$\to H^{i-1}(B, C) \to H^i(A, B) \to H^i(A, C)$$
$$\to H^i(B, C) \to H^{i+1}(A, B) \to.$$

T2. Mayer–Vietoris exact sequence

$$\to H_{i+1}(A \cup B) \to H_i(A \cap B) \to H_i(A) + H_i(B)$$
$$\to H_i(A \cup B) \to H_{i-1}(A \cap B) \to.$$

As a special application of T2 consider M a complex analytic open manifold and N a submanifold of codimension one, given by the zeros of a single analytic function. Then T2 applies with $A \sim M - N$, $B \sim N$, $A \cap B \sim S^1 \times N$, $A \cup B \sim M$. (In general S^k denotes the k sphere.) The signs $\sim$ imply homotopy equivalences.

T3. Poincaré duality. M a compact manifold with boundary B and real dimension n. Then $H_i(M, B) = H^{n-i}(M - B)$.

* Work supported in part through funds provided by the Atomic Energy Commission under Contract No. AT(30-1)-2098.

† Part of this work was completed when the author was with the Department of Mathematics, Massachusetts Institute of Technology.

[1] D. Fotiadi, M. Froissart, J. Lascoux, and F. Pham (to be published).

[2] A reference for the topology used is P. J. Hilton and S. Wylie, *Homology Theory* (Cambridge University Press, New York, 1960).

T4. If E is a fibre space with fibre F and base B, we denote the situation as

$$\begin{matrix} (F \rightarrow E \\ \downarrow \\ B). \end{matrix}$$

We will mainly use the $E_{p,q}^2$ terms in the spectral sequence. The following two types of Stein manifolds[3] of complex dimension n are used:

Type 1. All the common zeros of a set of polynomials in a complex affine space (if they form a manifold).

Type 2. A Stein manifold of Type 1 minus all the zeros of a polynomial (in the variables of the ambient affine space).

T5. For such a Stein manifold S, $H_i(S) = 0$, $i > n$.

T6. Let S be a Stein manifold in $C^N(x_1, \cdots, x_N)$ of Type 1 or 2 and P a polynomial in $x_1, \cdots, x_N$ and $a_1, \cdots, a_M$, $a_i \in C^M$. It is a reasonable conjecture in algebraic geometry that off an algebraic set in C^M one may move the a_i and the corresponding zeros of P in S around and extend this to an isotopy of the ambient manifold S. We refer to the a_i off this algebraic set as being in general position. If we assume this then

$$H_i(S - N) = H_i(S) + H_{i-1}(N)$$

with $N = \{x \in S \mid P(x) = 0\}$ provided N is a submanifold of S in general position and can be pushed to infinity (outside any given compact set) by varying the a_i. This follows by the argument of Froissart in Ref. 1. In our problem the a_i are the parameters, external momenta and internal masses. A further consequence of the above conjecture is that the singularities of Feynman diagrams fall on algebraic sets in the variables considered.

T7. The decomposition theorem of Ref. 1, that can be generalized using T6 to some manifolds in a Stein manifold. We will want to know the following specific manifolds:

$$\{x_i \in C^N \mid x_1^{\alpha_1} + \cdots + x_N^{\alpha_N} = 1, \alpha_i \text{ integer} > 0\} = M(\alpha_1, \cdots, \alpha_N)$$

$$\{x_i \in C^N - \text{origin} \mid x_1^2 + \cdots + x_N^2 = 0\} = TB_{N-1}.$$

T8. $M(\alpha_1, \cdots, \alpha_N)$ is homotopic to the wedge of $\prod_i (\alpha_i - 1)S^{N-1}$'s (this many spheres joined at a point). TB_{N-1} is homotopic to the tangent bundle of unit vectors on S^{N-1}. See Appendix I.

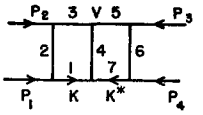


FIG. 1. The sixth-order ladder diagram.

T9. Consider a nonsingular manifold given as the common zeros in C^N of $g_1(x)$, $g_2(x)$, $F(x, a)$, $x \in C^N$, $a \in C^M$. Suppose $\nabla g_i \cdot (\nabla g_2)^* = 0$ (∇ is the gradient in C^N, $*$ indicates complex conjugation, ∇_a is the gradient in C^M). Define

$$\nabla_\perp f = \nabla f - \sum_1^2 \frac{\nabla g[\nabla f \cdot (\nabla g)^*]}{|\nabla g|^2}.$$

When the equation

$$dx = -(\nabla_\perp f)^* \nabla_a f \cdot da / |\nabla f_\perp|^2$$

is continuously integrable between values a_1 and a_2 along some path and over all the manifold, the equation provides an explicit homotopy between the manifolds defined for the different values of a.

THE FEYNMAN DIAGRAM

Refer to Fig. 1. k and k^* are the integration variables. We do not impose the condition $P_1 + P_2 + P_3 + P_4 = 0$. The effect of imposing this condition is discussed in Sec. 8. The numbers will label the denominators. We then transform the integration into the product of two complex projective spaces:

$(k, k^*) \rightarrow (x_i, y_i) \in CP_1^5 \times CP_2^5$ where

$$(x_1, \cdots, x_6) \in CP_1^5 \qquad (y_1, \cdots, y_6) \in CP_2^5$$

by

$$\begin{aligned} x_\alpha &= k_\alpha \qquad \alpha = 1, \cdots, 4 \\ y_\alpha &= k_\alpha^* \\ x_5 &= \tfrac{1}{2}(1 - k^2), \qquad y_5 = \tfrac{1}{2}[1 - (k^*)^2], \\ x_6 &= 1 + k^2, \qquad y_6 = 1 + (k^*)^2. \end{aligned} \tag{1}$$

We "normalize" the k and k^* so that we can set $M_1^2 = M_7^2 = 1$, M_i the mass in denominator i. For simplicity we have assumed M_1 equals M_7, an unnecessary condition. The points realized under this mapping into $CP^5 \times CP^5$ satisfy the equations

$$\left.\begin{aligned} \sum_{i=1}^5 x_i^2 &= \frac{x_6^2}{4} \\ \sum_{i=1}^5 y_i^2 &= \frac{y_6^2}{4} \end{aligned}\right\} = W_1 \times W_2 \subset CP^5 \times CP^5. \tag{2}$$

Denominators are associated with the points in $W_1 \times W_2$ on which they are zero. When the param-

[3] R. C. Gunning, "Local Theory of Several Complex Variables," H. Rossi, "Global Theory of Several Complex Variables." Both are Princeton University notes.

eters are off the singularities, the only case we consider, denominator 4 maps into a singular manifold in this product $W_1 \times W_2$, the other denominators into nonsingular manifolds. We also consider a "desingularized" Feynman diagram with denominator 4 modified by the addition of a term $\epsilon(1 + k^2)(1 + k^{*2})$. Under this modification denominator 4 maps into a manifold nonsingular off the union of the manifolds denominators 1 and 7 map into.

The purpose of this paper is to calculate the rank of the eight-dimensional homology group of the space comprising the complement in $W_1 \times W_2$ of the images of denominators 1 to 7.

In the "desingularized" case we apply T7 to $\{W_1 \times W_2 - 1 \cup 7\} = St_1$ the space $W_1 \times W_2$ off the denominators 1 and 7. In the actual case we apply T7 to $\{W_1 \times W_2 - 1 \cup 7 \cup 4\} = St_2$. By T7 the calculation of $H_i(W_1 \times W_2 - 1 \cup \cdots \cup 7\}$ is reduced to calculating certain homology groups of intersections of the denominators in St_1 or St_2.

In particular:
In the "desingularized" case

$$H_8(W_1 \times W_2 - 1 \cup \cdots \cup 7)$$
$$= \underset{A}{H_8(St_1)} + \underset{B}{4H_7(2 \cap St_1)} + \underset{C}{H_7(4 \cap St_1)} + \underset{D}{2H_6(2 \cap 3 \cap St_1)} + \underset{E}{4H_6(2 \cap 5 \cap St_1)}$$
$$+ \underset{F}{4H_6(2 \cap 4 \cap St_1)} + \underset{G}{4H_5(2 \cap 3 \cap 5 \cap St_1)} + \underset{H}{2H_5(2 \cap 3 \cap 4 \cap St_1)} + \underset{I}{4H_5(2 \cap 4 \cap 5 \cap St_1)}$$
$$+ \underset{J}{H_4(2 \cap 3 \cap 5 \cap 6 \cap St_1)} + \underset{K}{4H_4(2 \cap 3 \cap 4 \cap 5 \cap St_1)} + \underset{L}{H_3(2 \cap 3 \cap 4 \cap 5 \cap 6 \cap St_1)} \quad (3)$$

and in the physical case

$$H_8(W_1 \times W_2 - 1 \cup \cdots \cup 7) = \underset{M}{H_8(St_2)} + \underset{N}{4H_7(2 \cap St_2)} + \underset{O}{2H_6(2 \cap 3 \cap St_2)}$$
$$+ \underset{P}{4H_6(2 \cap 5 \cap St_2)} + \underset{Q}{4H_5(2 \cap 3 \cap 5 \cap St_2)} + \underset{R}{H_4(2 \cap 3 \cap 5 \cap 6 \cap St_2)}. \quad (4)$$

The essential symmetry between 2 and 3, 5 and 6, and 2, 3 and 5, 6 is used to give the above formulas. The capital letters will label the homology groups to be calculated. These will be calculated in the following sections, but not in order. The rank of the first group we calculate to be 142 and that of the second 126. Since the initial integration surface is a relative cycle in $(W_1 \times W_2, 1 \cup \cdots \cup 7)$ there are a possible 127 functions associated with the diagram if $P_1 + P_2 + P_3 + P_4 = 0$ is not imposed.

Section 1 (A, B, D, E, G, J)

1(a) $A = H_8(St_1)$

$$St_1 = \begin{cases} \sum_1^5 x_i^2 = 1 \\ \sum_1^5 y_i^2 = 1 \end{cases} \sim S^4 \times S^4 \text{ by T8}; \quad (5)$$

$A = Z$ (an infinite Abelian group on a single generator).

In giving such equations, trivial changes of variables will be made without comment.

1(b) $B = H_7(2 \cap St_1)$

$$2 \cap St_1 = \begin{cases} \sum_1^4 x_i^2 = 1 \\ \sum_1^5 y_i^2 = 1 \end{cases} \sim S^3 \times S^4; \quad B = Z. \quad (6)$$

1(c) $D = H_6(2 \cap 3 \cap St_1)$

$$2 \cap 3 \cap St_1 = \begin{cases} \sum_1^3 x_i^2 = 1 \\ \sum_1^5 y_i^2 = 1 \end{cases} \sim S^2 \times S^4; \quad D = Z. \quad (7)$$

1(d) $E = H_6(2 \cap 5 \cap St_1)$

$$2 \cap 5 \cap St_1 = \begin{cases} \sum_1^4 x_i^2 = 1 \\ \sum_1^4 y_i^2 = 1 \end{cases} \sim S^3 \times S^3; \quad E = Z. \quad (8)$$

1(e) $G = H_5(2 \cap 3 \cap 5 \cap St_1)$

$$2 \cap 3 \cap 5 \cap St_1 = \begin{cases} \sum_1^3 x_i^2 = 1 \\ \sum_1^4 y_i^2 = 1 \end{cases} \sim S^2 \times S^3; \quad G = Z. \quad (9)$$

1(f) $J = H_4(2 \cap 3 \cap 5 \cap 6 \cap St_1)$

$$2 \cap 3 \cap 5 \cap 6 \cap St_1 = \begin{cases} \sum_1^3 x_i^2 = 1 \\ \sum_1^3 y_i^2 = 1 \end{cases} \sim S^2 \times S^2; \quad J = Z. \quad (10)$$

Section 2 (*C, F, H*)

2(a) $$C = H_7(4 \cap St_1)$$

$$4 \cap St_1 = \begin{cases} \sum_1^5 x_i^2 = \frac{1}{4} \\ \sum_1^5 y_i^2 = \frac{1}{4} \\ \sum_1^4 x_i y_i + \frac{1}{2}(M^2 - 2)x_5 y_5 \\ \quad + \frac{1}{4}M^2(x_5 + y_5) = -\frac{1}{4} - \frac{1}{8}M^2 - \epsilon. \end{cases} \tag{11}$$

By arguing along the lines of T9 we may work with the following set of equations instead:

$$4 \cap St_1 = \begin{cases} \sum_1^5 x_i^2 = 1 \\ \sum_1^5 y_i^2 = 1 \\ \sum_1^4 x_i y_i + \alpha x_5 y_5 = 0, \quad \alpha^2 \neq 1. \end{cases} \tag{12}$$

If one chooses to avoid an analysis along the lines of T9 one may work with the original set of equations. In looking at group M we will work with a set of equations quite analogous to this set. From now on we will denote by π the projection from a point on a surface to its y coordinates. In this case this map gives a fibration of fibre $\sim S^3$ off the non-singular submanifold

$$LR = \{\sum_1^4 y_i^2 + \alpha^2 y_5^2 = 0\}. \tag{13}$$

Over LR there is a fibration of fibre $\sim S^2$. Each of these two fibrations is a manifold, the second a submanifold of $4 \cap St_1$.

$$\left.\begin{aligned} \textstyle\sum_1^5 y_i^2 = 1 \\ \textstyle\sum_1^4 y_i^2 + \alpha^2 y_5^2 = 0 \end{aligned}\right\} \Rightarrow y_5 = \pm(1 - \alpha^2)^{-\frac{1}{2}}. \tag{14}$$

We now apply T2 to $4 \cap St_1$ and the submanifold $\pi^{-1}(LR)$:

$$\rightarrow H_i(BR) \rightarrow H_i\{(S^3 \rightarrow M - R \;\} + H_i(R) \rightarrow H_i(M) \rightarrow \qquad \downarrow \pi \qquad S'^4 - LR) \tag{15}$$

with

$$S'^4 = \{\sum_1^5 y_i^2 = 1\}; \qquad M = 4 \cap St_1; \qquad BR = S^1 \times \pi^{-1}(LR); \qquad R = \pi^{-1}(LR).$$

R is fibred as follows

$$(S^2 \rightarrow R \qquad \downarrow \qquad LR \sim S^3 \cup S^3).$$

The $E^2_{p,q}$ term in the spectral sequence for R is $E^2_{p,q} = Z^{A(p)B(q)}$, where

$$\begin{aligned} A(p) &= 2, \quad p = 0, 3 \qquad & B(q) &= 1, \quad q = 0, 2 \\ &= 0, \quad p \neq 0, 3 & &= 0, \quad q \neq 0, 2 \end{aligned} \tag{16}$$

and for the complement of R, $M - R$, $E^2_{p,q} = Z^{A(p)B(q)}$, where

$$\begin{aligned} A(p) &= 1, \quad p = 0 \qquad & B(q) &= 1, \quad q = 0, 3 \\ &= 2, \quad p = 1 & &= 0, \quad q \neq 0, 3 \\ &= 3, \quad p = 4 \\ &= 0, \quad p \neq 0, 1, 4 \end{aligned} \tag{17}$$

using T7 on $S'^4 - LR$. In all the fibrations we deal with, π_1, (base) acts trivially on the homology of the fibre.

A portion of (15) is now evidenced,

$$0 \rightarrow Z^3 + 0 \rightarrow H_7(M) \rightarrow Z^2 \rightarrow 0. \tag{18}$$

Thus we see that $C = Z^5$.

2(b) $$F = H_6(2 \cap 4 \cap St_1)$$

$$2 \cap 4 \cap St_1 = \begin{cases} \sum_1^5 x_i^2 = 1 \\ \sum_1^4 y_i^2 = 1 \\ \sum_1^3 x_i y_i + \alpha x_4 y_4 = 0, \quad \alpha^2 \neq 1. \end{cases} \tag{19}$$

Proceeding as in 2(a), with a few trivial changes of the form $i \rightarrow i - 1$ we find that $F = Z^5$.

2(c) $$H = H_5(2 \cap 3 \cap 4 \cap St_1)$$

$$2 \cap 3 \cap 4 \cap St_1 = M = \begin{cases} \sum_1^5 x_i^2 = 1 \\ \sum_1^3 y_i^2 = 1 \\ x_1 y_1 + x_2 y_2 + \alpha x_3 y_3 = 0, \quad \alpha^2 \neq 1. \end{cases} \tag{20}$$

We are forced to change notation between sections; analogous objects in different sections may have the same name. Retracing 2(a),

$$\rightarrow H_i(BR) \rightarrow H_i\{(S^3 \rightarrow M - R \;\} + H_i(R) \rightarrow H_i(M) \rightarrow; \qquad \downarrow \qquad S'^2 - LR) \qquad (S^2 \rightarrow R \qquad \downarrow \qquad LR \sim S^1 \cup S^1). \tag{21}$$

The $E^2_{p,q}$ term for R is $E^2_{p,q} = Z^{A(p)B(q)}$ with

$$\begin{aligned} A(p) &= 2, \quad p = 0, 1 \qquad B(q) = 1, \quad q = 0, 2 \\ &= 0, \quad p \neq 0, 1 \qquad\quad = 0, \quad q \neq 0, 2. \end{aligned} \tag{22}$$

The $E^2_{p,q}$ term for $M - R$ is $E^2_{p,q} = Z^{A(p)B(q)}$ with

$$\begin{aligned} A(p) &= 1, \quad p = 0 \qquad\quad B(q) = 1, \quad q = 0, 3 \\ &= 2, \quad p = 1 \qquad\quad\ = 0, \quad q \neq 0, 3 \\ &= 3, \quad p = 2 \\ &= 0, \quad p \neq 0, 1, 2 \end{aligned} \tag{23}$$

and (18) is replaced by

$$0 \to Z^3 + 0 \to H_5(M) \to Z^2 \xrightarrow{f} Z^2. \tag{24}$$

We need to know map f to deduce $H_5(M)$. f is equal to zero. One way to see this is by considering the symmetry transformation $x \to -x$, $y \to y$. Under this transformation, $H_4(BR)$ changes sign and $H_4(M - R)$ does not.

Then, $H = Z^5$.

Section 3 (I, K, L)

3(a) $\qquad I = H_5(2 \cap 4 \cap 5 \cap St_1)$

$$2 \cap 4 \cap 5 \cap St_1 = M = \begin{cases} \sum_1^4 x_i^2 = 1 \\ \sum_1^4 y_i^2 = 1 \\ x_1y_1 + x_2y_2 + \alpha x_3y_3 + \beta x_4y_4 = 0, \\ 1 \neq \alpha^2 \neq \beta^2 \neq 1. \end{cases} \tag{25}$$

Again try to fibre over y. Off

$$LR = \begin{cases} y_1^2 + y_2^2 + \alpha^2 y_3^2 + \beta^2 y_4^2 = 0 \\ \sum_1^4 y_i^2 = 1 \end{cases}$$

there is a fibre space of fibre $\sim S^2$. $\pi^{-1}(LR)$ is a non-singular manifold R fibred over LR with fibre $\sim S^1$.

$$\begin{matrix} (S^1 \to R \\ \quad\downarrow \text{ has } E^2_{p,q} = Z^{A(p)B(q)} \text{ with} \\ LR) \end{matrix}$$

$$\begin{aligned} A(p) &= 1, \quad p = 0, 1 \qquad B(q) = 1, \quad q = 0, 1 \\ &= 4, \quad p = 2 \qquad\quad = 0, \quad q \neq 0, 1. \\ &= 0, \quad p \neq 0, 1, 2 \end{aligned} \tag{26}$$

For the homology of LR see Appendix I.

$$\begin{matrix} (S^2 \to M - R \\ \quad\downarrow \quad \text{has } E^2_{p,q} = Z^{A(p)B(q)} \text{ where} \\ S'^3 - LR) \end{matrix}$$

$$\begin{aligned} A(p) &= 1, \quad p = 0, 1, 2 \qquad B(q) = 1, \quad q = 0, 2 \\ &= 5, \quad p = 3 \qquad\quad = 0, \quad q \neq 0, 2. \\ &= 0, \quad p \neq 0, 1, 2, 3 \end{aligned} \tag{27}$$

A portion of the sequence

$$\to H_i(BR) \to H_i(M - R) + H_i(R) \to H_i(M) \to$$

now is

$$0 \to Z^5 \to H_5(M) \to Z^4 \xrightarrow{f} Z \to. \tag{28}$$

By the same argument as in 2(c), $f = 0$; $I = Z^9$.

3(b) $\qquad K = H_4(2 \cap 3 \cap 4 \cap 5 \cap St_1)$

$$2 \cap 3 \cap 4 \cap 5 \cap St_1 = \begin{cases} \sum_1^4 x_i^2 = 1 \\ \sum_1^3 y_i^2 = 1 \\ x_1y_1 + \alpha x_2y_2 + \beta x_3y_3 = 0 \end{cases} \tag{29}$$

where $1 \neq \alpha^2 \neq \beta^2 \neq 1$;

$$LR = \begin{cases} \sum_1^3 y_i^2 = 1 \\ y_1^2 + \alpha^2 y_2^2 + \beta^2 y_3^2 = 0. \end{cases}$$

$$\begin{matrix} (S^1 \to R \\ \quad\downarrow \text{ has } E^2_{p,q} = Z^{A(p)B(q)} \text{ with} \\ LR) \end{matrix}$$

$$\begin{aligned} A(p) &= 1, \quad p = 0 \qquad B(q) = 1, \quad q = 0, 1 \\ &= 5, \quad p = 1 \qquad\quad = 0, \quad q \neq 0, 1. \\ &= 0, \quad p \neq 0, 1 \end{aligned} \tag{30}$$

For the homology of LR see Appendix I.

$$\begin{matrix} (S^2 \to M - R \\ \quad\downarrow \quad \text{has } E^2_{p,q} = Z^{A(p)B(q)} \text{ with} \\ S'^2 - LR) \end{matrix}$$

$$\begin{aligned} A(p) &= 1, \quad p = 0, 1 \qquad B(q) = 1, \quad q = 0, 2 \\ &= 6, \quad p = 2 \qquad\quad = 0, \quad q \neq 0, 2. \\ &= 0, \quad p \neq 0, 1, 2 \end{aligned} \tag{31}$$

$$\to H_i(S^1 \times R) \to H_i(M - R) + H_i(R) \to H_i(M) \to$$

yields

$$0 \to Z^6 \to H_4(M) \to Z^5 \xrightarrow{f} Z \to. \tag{32}$$

f is equal to zero for the same reason as in 3(a). Thus, $K = Z^{11}$.

3(c) $\qquad L = H_3(2 \cap 3 \cap 4 \cap 5 \cap 6 \cap St_1)$

$$2 \cap 3 \cap 4 \cap 5 \cap 6 \cap St_1 = \begin{cases} \sum_1^3 x_i^2 = 1 \\ \sum_1^3 y_i^2 = 1 \\ x_1y_1 + \alpha x_2y_2 + \beta x_3y_3 = 0 \end{cases} \tag{33}$$

where $1 \neq \alpha^2 \neq \beta^2 \neq 1$;

$$LR = \begin{cases} y_1^2 + y_2^2 + y_3^2 = 1 \\ y_1^2 + \alpha^2 y_2^2 + \beta^2 y_3^2 = 0. \end{cases}$$

$$\begin{pmatrix} S^0 \to R \\ \downarrow \quad \text{has } E_{p,q}^2 = Z^{A(p)B(q)} \text{ with} \\ LR \end{pmatrix}$$

$$\begin{aligned} A(p) &= 1, \quad p = 0 \qquad B(q) = 2, \quad q = 0 \\ &= 5, \quad p = 1 \qquad\quad = 0, \quad q \neq 0. \\ &= 0, \quad p \neq 0, 1 \end{aligned} \tag{34}$$

$$\begin{pmatrix} S' \to M - R \\ \downarrow \quad \text{has } E_{p,q}^2 = Z^{A(p)B(q)} \text{ with} \\ S'^2 - LR \end{pmatrix}$$

$$\begin{aligned} A(p) &= 1, \quad p = 0, 1 \quad B(q) = 1, q = 0, 1 \\ &= 6, \quad p = 2 \qquad\quad = 0, q \neq 0, 1. \\ &= 0, \quad p \neq 0, 1, 2 \end{aligned} \tag{35}$$

$$\to H_i(BR) \to H_i(M - R) + H_i(R) \to H_i(M) \to$$

yields

$$0 \to Z^6 \to H_3(M) \to Z^{10} \xrightarrow{f} (Z^7 \text{ or } Z^6) \to. \tag{36}$$

The doubt indicated by the parentheses arises because we have not calculated the differential d^2 in the relevant spectral sequence. f has rank 5. We do not indicate the proof of this; it is similar to the argument that follows Eq. (53).

Thus we have $L = Z^{11}$.

Section 4 (M)

4(a)

Groups M through R are more difficult to calculate than those we have already calculated. The procedure, however, is quite similar.

$$M = H_8(St_2)$$

$$St_2 = \begin{cases} \sum_1^5 x_i^2 = \frac{1}{4} \\ \sum_1^5 y_i^2 = \frac{1}{4} \end{cases}$$

$$- \begin{cases} \sum_1^4 x_i y_i + \frac{1}{2}(M^2 - 2)x_5 y_5 \\ \qquad + \frac{1}{4}M^2(x_5 + y_5) = -\frac{1}{4} - \frac{1}{8}M^2 \end{cases}$$

$$= W_1 \times W_2 - 4 - \infty. \tag{37}$$

∞ denotes the "hyperplane at infinity" in $W_1 \times W_2$ equal to $1 \cup 7$. By T3, $H_8(W_1 \times W_2 - 4 - \infty) = H^8(W_1 \times W_2, 4 \cup \infty)$. By T1

$$\begin{aligned} &H^7(W_1 \times W_2, \infty) \to H^7(4 \cup \infty, \infty) \\ &\quad \| \text{ by T3} \\ &\quad 0 \\ &\to H^8(W_1 \times W_2, 4 \cup \infty) \xrightarrow{g} H^8(W_1 \times W_2, \infty). \\ &\qquad\qquad\qquad\qquad\qquad\qquad\quad \| \\ &\qquad\qquad\qquad\qquad\qquad\qquad\quad Z \end{aligned} \tag{38}$$

In Appendix II we show $g = 0$ so that

$$H^8(W_1 \times W_2, 4 \cup \infty) = H^7(4 \cup \infty, \infty).$$

We now look at the map π on 4. This yields a fibration of fibre $\sim S^3$ except where y satisfies one of the three conditions:

$$\begin{aligned} &\text{(a)} \quad \sum_1^4 y_i^2 + \left(\frac{M^2 - 2}{2} y_5 + \frac{M^2}{4}\right)^2 = 0, \\ &\qquad \left(-\frac{1}{4} - \frac{1}{8}M^2 - \frac{M^2}{4} y_5\right) = 0. \end{aligned} \tag{39.1}$$

In case (a), $\pi^{-1}(pt.) \sim S^2$ and is a nonsingular manifold.

$$\begin{aligned} &\text{(b)} \quad \sum_1^4 y_i^2 + \left(\frac{M^2 - 2}{2} y_5 + \frac{M^2}{4}\right)^2 = 0, \\ &\qquad \left(-\frac{1}{4} - \frac{M^2}{8} - \frac{M^2}{4} y_5\right) \neq 0. \end{aligned} \tag{39.2}$$

In case (b), $\pi^{-1}(pt.) \sim pt.$ and is a nonsingular manifold.

$$\begin{aligned} &\text{(c)} \quad \sum_1^4 y_i^2 + \left(\frac{M^2 - 2}{2} y_5 + \frac{M^2}{4}\right)^2 \\ &\qquad = 4\left(-\frac{1}{4} - \frac{1}{8}M^2 - \frac{M^2}{4} y_5\right)^2 \\ &\qquad\quad \left(-\frac{1}{4} - \frac{M^2}{8} - \frac{M^2}{4} y_5\right) \neq 0 \end{aligned} \tag{39.3}$$

In case (c), $\pi^{-1}(pt.) \sim pt.$ and is a singular manifold, singular at a single point.

A little algebra shows that (a) does not occur, (c) occurs when $y_5 = -\frac{1}{2}$, and (b) occurs when $y_5 = C_1$ and $y_5 = C_2$ where C_1 and C_2 are distinct values not equal to $\pm\frac{1}{2}$. Define LR_1 equal the manifold in $\sum_1^5 y_i^2 = \frac{1}{4}$ with $y_5 = C_1$ or C_2 and LR_2 equal the set of y's in $\sum_1^5 y_i^2 = \frac{1}{4}$ with $y_5 = -\frac{1}{2}$. Let p equal the point with $y_5 = -\frac{1}{2}$, $y_\alpha = 0$, $\alpha \neq 5$ and SR_2 equal the union of the singular points over LR_2, one point in SR_2 for each point in LR_2. Finally define $SN = SR_2 \cup \pi^{-1}(p)$. Algebraically one verifies that $4 - SN$ is a nonsingular manifold. Look at the exact sequence

$$\begin{aligned} &H^6(4 \cup \infty, \infty) \\ &\to H^6(SN \cup \infty, \infty) \to H^7(4 \cup \infty, SN \cup \infty) \\ &\to H^7(4 \cup \infty, \infty) \to H^7(SN \cup \infty, \infty). \end{aligned} \tag{40}$$

The fourth term indicated is what we wish to know. We will calculate the second and fifth terms to be Z^2 and 0, respectively. We will calculate the third term to be Z^6. The first term is zero, for, looking at the continuation of Eq. (38),

$$\begin{array}{ccc} H^6(W_1 \times W_2, \infty) \to H^6(4 \cup \infty, \infty) \to H^7(W_1 \times W_2, 4 \cup \infty). \\ \| \qquad\qquad\qquad\qquad\quad \| \text{ by T3 and T5} \\ 0 \qquad\qquad\qquad\qquad\quad 0 \end{array} \tag{41}$$

Therefore $H^7(4 \cup \infty, \infty) = Z^4$ and $M = Z^4$.

4(b) *Study of* $H^i(SN \cup \infty, \infty)$

We look at a few terms in the exact sequence

$$\begin{aligned} &\to H^i(SN \cup \infty, \pi^{-1}(p) \cup \infty) \\ &\to H^i(SN \cup \infty, \infty) \to H^i(\pi^{-1}(p) \cup \infty, \infty) \to. \end{aligned} \tag{42}$$

We first study

$$\begin{aligned} &H^i(\pi^{-1}(p) \cup \infty, \infty) \\ &\to H^i(\pi^{-1}(p) \cup \infty, \pi^{-1}(p) \cap SR_2 \cup \infty) \\ &\qquad \to H^i(\pi^{-1}(p) \cup \infty, \infty) \to \\ &\quad \to H^i(\pi^{-1}(p) \cap SR_2 \cup \infty, \infty) \to \\ &H^i(\pi^{-1}(p) \cap SR_2 \cup \infty, \infty) = Z \quad \text{for} \quad i = 0 \\ &\qquad\qquad\qquad\qquad\quad = 0 \quad \text{for} \quad i \neq 0 \end{aligned}$$

$$\begin{aligned} &H^i(\pi^{-1}(p) \cup \infty, \infty) \\ &\quad = H^i(\pi^{-1}(p) \cup \infty, \pi^{-1}(p) \cap SR_2 \cup \infty), \quad i > 1, \\ &H^i(\pi^{-1}(p) \cup \infty, \pi^{-1}(p) \cap SR_2 \cup \infty) \\ &\qquad = H_{6-i}(\pi^{-1}(p) - \pi^{-1}(p) \cap SR_2) \\ &\qquad = H_{6-i}(TB_3) \\ &\qquad = Z, \quad i = 6, 4, 3 \\ &\qquad = 0, \quad i \neq 6, 4, 3. \end{aligned} \tag{43}$$

See Ref. 4 for the last result.

We next consider $H^i(SN \cup \infty, \pi^{-1}(P) \cup \infty)$

$$\begin{aligned} H^i(SN \cup \infty, \pi^{-1}(p) \cup \infty) &= H_{6-i}(SR_2 - \pi^{-1}(p)) \\ &= H_{6-i}(LR_2 - p) \\ &= H_{6-i}(TB_3) \\ &= Z, \quad i = 6, 4, 3, 1 \\ &= 0, \quad i \neq 6, 4, 3, 1. \end{aligned} \tag{44}$$

Referring back to Eq. (42) we see

$$\begin{aligned} H^6(SN \cup \infty, \infty) &= Z^2, \\ H^7(SN \cup \infty, \infty) &= 0. \end{aligned} \tag{45}$$

4(c) *Study of* $H^7(4 \cup \infty, SN \cup \infty)$

$$H^7(4 \cup \infty, SN \cup \infty) = H_7(4 - SN).$$

We now use the discussion following T2 to write

$$\begin{array}{l} \to H_i(BR_1) + H_i(BR_2) \to H_i\{(S^3 \to 4 - SN - R_1 - R_2\} + H_i(R_1) + H_i(R_2) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \downarrow \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad S'^4 - LR_1 - LR_2) \end{array} \tag{46}$$

$\to H_i(4 - SN) \to$ with

$$BR_i = S^1 \times R_i, \qquad R_1 = \pi^{-1}(LR_1), \qquad R_2 = \pi^{-1}(LR_2 - p).$$

We have the fibration for R_1,

$$\begin{array}{l} (pt. \to R_1 \\ \qquad \downarrow \quad \sim LR_1 \sim S^3 \cup S^3 \\ \qquad LR_1), \end{array} \tag{47}$$

and for R_2,

$$\begin{array}{l} (TB_3 \to R_2 \\ \qquad\quad \downarrow \\ \qquad\quad LR_2 - p) \end{array}$$

with $E^2_{p,q} = Z^{A(p)B(q)}$ where

$$\begin{aligned} A(p) &= 1, \quad p = 0, 2, 3, 5 \\ &= 0, \quad p \neq 0, 2, 3, 5 \\ B(q) &= 1, \quad q = 0, 2, 3, 5 \\ &= 0, \quad q \neq 0, 2, 3, 5; \end{aligned} \tag{48}$$

$$\begin{aligned} &H_i(S'^4 - LR_1 - LR_2) \\ &\qquad = H_i(S'^4 - LR_2) + H_{i-1}(LR_1) \text{ by T6,} \\ &H_i(S'^4 - LR_2) = H^{8-i}(S'^4 \cup \infty, LR_2 \cup \infty) \text{ by T3.} \end{aligned}$$

[4] N. Steenrod, *The Topology of Fibre Bundles* (Princeton University Press, Princeton, New Jersey, 1951).

Look at the sequence

$$\to H^i(S'^4 \cup \infty, \infty \cup LR_2) \xrightarrow{i^i} H^i(S'^4 \cup \infty, \infty) \to H^i(LR_2 \cup \infty, \infty) \to. \tag{49}$$

Note that $H^i(LR_2 \cup \infty, \infty \cup p) = Z$ when $i = 6, 4, 3, 1$; $= 0$ when $i \neq 6, 4, 3, 1$.

Equation (49) now looks like

$$G_1 \to 0 \to 0 \to G_2 \to 0 \to 0 \to G_3 \to 0 \to Z \to G_4 \xrightarrow{r^4} Z \to Z \to G_5 \to 0 \to 0 \to G_6 \to 0 \to Z$$
$$\to G_7 \to 0 \to 0 \to G_8 \to Z \to 0; \qquad G_i = H^i(S'^4 \cup \infty, \infty \cup LR_2). \tag{50}$$

Map r^4 is shown to be zero in Appendix II

$$H_i(S'^4 - LR_2) = Z, \quad i = 0, 1, 4$$
$$= 0, \quad i \neq 0, 1, 4.$$

Then

$$(S^3 \to 4 - SN - R_1 - R_2 \downarrow S'^4 - LR_1 - LR_2)$$

has $E^2_{p,q} = Z^{A(p)B(q)}$ with

$$\begin{aligned} A(p) &= 1, \quad p = 0 \qquad & B(q) &= 1, \quad q = 0, 3 \\ &= 3, \quad p = 1, 4 & &= 0, \quad q \neq 0, 3. \\ &= 0, \quad p \neq 1, 4, 0 & & \end{aligned} \tag{51}$$

If we knew d^2 and d^3 in (48) we would know $E^\infty_{p,q}$. We claim $d^3 = 0$ as the bundle has a cross section. The cross section can be chosen as

$$\begin{gathered} (y_\alpha, y_5 = -\tfrac{1}{2}) \to (y_\alpha, y_5, x_\alpha, x_5) \\ (y_\alpha, -\tfrac{1}{2}) \to (y_\alpha, -\tfrac{1}{2}, y_\alpha, -\tfrac{1}{2}) \\ \alpha = 1, \cdots, 4. \end{gathered} \tag{52}$$

Another argument that shows that d^2 and d^3 are both zero is given after (60). It then follows that $H_5(R_2) = Z^4$. We now write a portion of the sequence (46)

$$0 \to (0 \text{ or } Z) \to Z^3 \xrightarrow{i} H_7(4 - SN) \to Z^4 \to 0. \tag{53}$$

We have here made the convention that we set $H_i(R)$ equal zero and in calculating $H_i(S^1 \times R)$ keep only terms from $H_1(S^1)$ and not $H_0(S^1)$ as these terms automatically cancel. We use this convention henceforth. We need to know the rank of i. It is 2. To obtain this we consider the subspace of 4 satisfying $\sum_1^5 |x_i|^2 = M(y)$ where $M(y)$ is a continuous positive function whose numerical value is chosen large enough so that $\pi^{-1}(pt.) \sim TB_3$. We will have to avoid LR_1 where this is impossible. We then map a sequence like (46) but written for this subspace into (46). Using the fact that one of the maps carries $H_3(TB_3)$ isomorphically into $H_3(S^3)$, $H_3(S^3)$ in the fibre of

$$(S^3 \to 4 - SN - R_1 - R_2 \downarrow S'^4 - LR_1 - LR_2),$$

$H_3(TB_3)$ in the fibre of the term corresponding to this, we deduce the rank of i to be 2.

Thus, $M = Z^4$.

Section 5 (N, O)

5(a) $\qquad N = H_7(2 \cap St_2)$

$$2 \cap St_2 = \begin{cases} \sum_1^5 x_i^2 = 1 \\ \sum_1^4 y_i^2 = 1 \end{cases} - \{ \sum_1^3 x_i y_i + \alpha x_4 y_4 = 1 \tag{54}$$
$$\alpha^2 \neq 1.$$

We have here made the assumption that (54) represents $2 \cap St_2$ faithfully enough—using an argument like that in T9 on the singular manifolds involved. Four things make this seem legitimate.

(a) The first two equations are identical to the exact general case.

(b) The third equation has the same quadratic terms as the exact system, missing only terms linear in the variables, and therefore has the same intersection with the "hyperplane at infinity."

(c) The exact system and this system have singular loci of the same type.

(d) They have isomorphic associated "desingularized" manifolds.

We have checked this assumption at least so far as it applies to the homology group of interest.

$$H_i(2 \cap St_2) = H^{14-i}(2 \cup \infty, 4 \cup \infty)$$
$$0 \to H^6(4 \cap 2 \cup \infty, \infty) \to H^7(2 \cup \infty, 4 \cup \infty) \to$$
$$\xrightarrow{g} H^7(2 \cup \infty, \infty) \tag{55}$$
$$\| \qquad Z.$$

By Appendix II, g is onto.

$$N = Z + H^6(4 \cap 2 \cup \infty, \infty).$$

Define similarly to the last section

$$LR_1 = \{\textstyle\sum_1^4 y_i^2 = 1,\ \sum_1^3 y_i^2 + \alpha^2 y_4^2 = 0\}.$$

The fibre over LR_1 is nonsingular $\sim pt.$,

$$LR_2 = \{\textstyle\sum_1^4 y_i^2 = 1,\ \sum_1^3 y_i^2 + \alpha^2 y_4^2 = 1\}.$$

The fibre over LR_2 is $\sim pt.$ and singular at a single point. SN is the union of the singular points over LR_2.

$$R_1 = \pi^{-1}(LR_1), \qquad R_2 = \pi^{-1}(LR_2)$$

$$\begin{array}{ccccccccc} \to H^5(SN \cup \infty, \infty) & \to & H^6(4 \cap 2 \cup \infty, SN \cup \infty) & \to & H^6(4 \cap 2 \cup \infty, \infty) & \to & H^6(SN \cup \infty, \infty), \\ \| & & & & & & \| & & \\ 0 & & & & & & 0 & & \end{array} \tag{56}$$

$$H^6(4 \cap 2 \cup \infty, \infty) = H^6(4 \cap 2 \cup \infty, \infty \cup SN) = H_6(4 \cap 2 - SN).$$

$$\begin{array}{l} (pt. \to R_1 \\ \qquad \downarrow \text{ has } E_{p,q}^2 = Z^{A(p)B(q)} \text{ with} \\ \qquad LR_1) \end{array}$$

$$\begin{array}{llll} A(p) = 2, & p = 0, 2 & B(q) = 1, & q = 0 \\ = 0, & p \neq 0, 2 & = 0, & q \neq 0. \end{array} \tag{57}$$

$$\begin{array}{l} (TB_3 \to R_2 \\ \qquad \downarrow \text{ has } E_{p,q}^2 = Z^{A(p)B(q)} \text{ with} \\ \qquad LR_2) \end{array}$$

$$\begin{array}{llll} A(p) = 1, & p = 0, 2 & B(q) = 1, & q = 0, 2, 3, 5 \\ = 0, & p \neq 0, 2 & = 0, & q \neq 0, 2, 3, 5. \end{array} \tag{58}$$

$$\begin{array}{l} (S^3 \to 4 \cap 2 - R_1 - R_2 - SN \\ \qquad \downarrow \\ \qquad S'^3 - LR_1 - LR_2) \end{array}$$

has $E_{p,q}^2 = Z^{A(p)B(q)}$ with

$$\begin{array}{llll} A(p) = 1, & p = 0 & B(q) = 1, & q = 0, 3 \\ = 3 & p = 1 & = 0, & q \neq 0, 3. \\ = 4, & p = 3 & & \\ = 0, & p \neq 0, 1, 3 & & \end{array} \tag{59}$$

The relevant Mayer–Vietoris sequence is

$$Z^2 \xrightarrow{k} Z^4 \to H_6(4 \cap 2 - SN) \to (Z \text{ or } 0) \to 0. \tag{60}$$

The rank of k can be seen to be one by looking at the two terms preceding it in the series, or by an argument similar to that after (53). The doubt indicated by (Z or 0) is due to our lack of knowledge about d^2 in (58). This is now investigated. The fibration of (58) is a trivial fibration since LR_2 can be shrunk to a point in $S'^3 - LR_1$ in the construction following (53).

Therefore $d^2 = 0$ and $N = Z^5$.

5(b) $\qquad O = H_6(2 \cap 3 \cap St_2)$

$$2 \cap 3 \cap St_2 = \begin{cases} \sum_1^5 x_i^2 = 1 \\ \sum_1^3 y_i^2 = 1 \end{cases}$$

$$- \{x_1 y_1 + x_2 y_2 + \alpha x_3 y_3 = 1 \qquad \alpha^2 \neq 1 \tag{61}$$

$$H_6(2 \cap 3 \cap St_2) = H^6(2 \cap 3 \cup \infty, 4 \cup \infty).$$

As in 5(a) define

$$LR_1 = \{y_3 = \pm(1 - \alpha^2)^{-\frac{1}{2}}\}$$

$$LR_2 = \{y_3 = 0\}.$$

SN is the union of singular points over LR_2

$$R_1 = \pi^{-1}(LR_1), \qquad R_2 = \pi^{-1}(LR_2)$$

$$\begin{array}{l} O \to H^5(2 \cap 3 \cap 4 \cup \infty, \infty) \\ \qquad \to H^6(2 \cap 3 \cup \infty, 2 \cap 3 \cap 4 \cup \infty) \to \\ \qquad \to H^6(2 \cap 3 \cup \infty, \infty). \\ \qquad\qquad\qquad\qquad \| \\ \qquad\qquad\qquad\qquad Z \end{array} \tag{62}$$

Therefore

$$O = Z + H^5(2 \cap 3 \cap 4 \cup \infty, \infty)$$

$$\begin{array}{ccccccccc} \to H^4(SN \cup \infty, \infty) & \to & H^5(2 \cap 3 \cap 4 \cup \infty, SN \cup \infty) & \to & H^5(2 \cap 3 \cap 4 \cup \infty, \infty) & \to & H^5(SN \cup \infty, \infty) \\ \| & & & & & & \| \\ 0 & & & & & & 0 \end{array} \tag{63}$$

$$O = Z + H^5(2 \cap 3 \cap 4 \cup \infty, SN \cup \infty)$$
$$= Z + H_5(2 \cap 3 \cap 4 - SN).$$

$(pt. \to R_1$
$\downarrow$ has $E^2_{p,q} = Z^{A(p)B(q)}$ with
$LR_1)$

$$A(p) = 2, \quad p = 0, 1 \qquad B(q) = 1, \quad q = 0$$
$$= 0, \quad p \neq 0, 1 \qquad = 0, \quad q \neq 0. \tag{64}$$

$(TB_3 \to R_2$
$\downarrow$ has $E^2_{p,q} = Z^{A(p)B(q)}$ with
$LR_2)$

$$A(p) = 1, \quad p = 0, 1 \qquad B(q) = 1, \quad q = 0, 2, 3, 5$$
$$= 0, \quad p \neq 0, 1 \qquad = 0, \quad q \neq 0, 2, 3, 5. \tag{65}$$

$$(S^3 \to 4 \cap 2 \cap 3 - SN - R_1 - R_2$$
$$\downarrow$$
$$S'^2 - LR_1 - LR_2)$$

has $E^2_{p,q} = Z^{A(p)B(q)}$ with

$$A(p) = 1, \quad p = 0 \qquad B(q) = 1, \quad q = 0, 3$$
$$= 3, \quad p = 1 \qquad = 0, \quad q \neq 0, 3.$$
$$= 4, \quad p = 2$$
$$= 0, \quad p \neq 0, 1, 2 \tag{66}$$

The relevant Mayer–Vietoris sequence is

$$Z \xrightarrow{k} Z^4 \to H_5(2 \cap 3 \cap 4 - SN) \to Z^2 \xrightarrow{g} Z^3. \tag{67}$$

k has rank 1 by the same argument as (53) and g has rank 1 similarly.

Thus we have $O = Z^5$.

Section 6 (P, Q)

6(a) $\qquad P = H_6(2 \cap 5 \cap St_2)$

$$2 \cap 5 \cap St_2 = \begin{cases} \sum_1^4 x_i^2 = 1 \\ \sum_1^4 y_i^2 = 1 \end{cases}$$
$$- \{x_1y_1 + x_2y_2 + \alpha x_3y_3 + \beta x_4y_4 = 1, \tag{68}$$

where $1 \neq \alpha^2 \neq \beta^2 \neq 1$;

$$H_i(2 \cap 5 \cap St_2) = H^{12-i}(2 \cap 5 \cup \infty, 4 \cup \infty);$$

4 denotes $4 \cap 2 \cap 5$.

$$\begin{array}{ccccccccc} H^5(2 \cap 5 \cup \infty, \infty) & \to & H^5(4 \cup \infty, \infty) & \to & H^6(2 \cap 5 \cup \infty, 4 \cup \infty) & \xrightarrow{f} & H^6(2 \cap 5 \cup \infty, \infty). \\ \| & & & & & & \| \\ 0 & & & & & & Z \end{array} \tag{69}$$

f is onto by an argument as in Appendix II.

$$P = H^5(4 \cup \infty, \infty) + Z.$$

Now looking at the fibration π and considering the three types of singular places as in (39)

(a) none

(b) $$LR_1 = \begin{cases} \sum_1^4 y_i^2 = 1 \\ y_1^2 + y_2^2 + y_3^2\alpha^2 + y_4^2\beta^2 = 0 \end{cases}$$

with contractible nonsingular fibre.

(c) $$LR_2 = \begin{cases} \sum_1^4 y_i^2 = 1 \\ y_1^2 + y_2^2 + \alpha^2 y_3^2 + \beta^2 y_4^2 = 1 \end{cases}$$

with contractible fibre singular at a single point.

The union of singular points is SR_2.

LR_1 is analyzed in Appendix I.

$$LR_2 = \begin{cases} (1 - \alpha^2)y_3^2 + (1 - \beta^2)y_4^2 = 0 \\ y_3 = \pm[-(1 - \beta^2)/(1 - \alpha^2)]^{\frac{1}{2}} y_4. \end{cases} \tag{70}$$

Thus LR_2 looks like two S^2's joined along $SM = \{y_3 = y_4 = 0\}$ and has

$$H_0(LR_2) = Z$$
$$H_2(LR_2) = Z^3$$
$$H_i(LR_2) = 0, \quad i \neq 1, 2$$

$$SN = SR_2 \cup \pi^{-1}(SM)$$

$$R_1 = \pi^{-1}(LR_1), \qquad R_2 = \pi^{-1}(LR_2 - SM)$$

$$\begin{array}{ccc} H^4(4 \cup \infty, \infty) & \to & H^4(SN \cup \infty, \infty) \\ \| & & \| \\ 0 & & Z^2 \end{array}$$

$$\to H^5(4 \cup \infty, SN \cup \infty) \to H^5(4 \cup \infty, \infty) \to$$
$$\to H^5(SN \cup \infty, \infty) \xrightarrow{k} H^6(4 \cup \infty, SN \cup \infty). \tag{71}$$
$$\|$$
$$Z$$

We now deduce the indicated values of the second and fifth terms in the above sequence and calculate the rank of k to be one. Therefore

$$H^5(4 \cup \infty, \infty) = H^5(4 \cup \infty, SN \cup \infty)/Z^2.$$

Look at

$$\to H^i(SN \cup \infty, \pi^{-1}(SM) \cup \infty) \to H^i(SN \cup \infty, \infty)$$
$$\to H^i(\pi^{-1}(SM) \cup \infty, \infty) \to. \tag{72}$$

Studying the first term

$$H^i(SN \cup \infty, \pi^{-1}(SM) \cup \infty) = H_{4-i}(SN - \pi^{-1}(SM))$$

$$\begin{aligned} H_i(SN - \pi^{-1}(SM)) &= Z^2, \quad i = 0, 1 \\ &= Z^4, \quad i = 2 \\ &= 0, \quad i \neq 0, 1, 2 \end{aligned} \tag{73}$$

by T6. Studying the third term

$$\begin{aligned} &\to H^i(\pi^{-1}(SM) \cup \infty, SR_2 \cap \pi^{-1}(SM) \cup \infty) \\ &\quad \to H^i(\pi^{-1}(SM) \cup \infty, \infty) \\ &\quad \to H^i(SR_2 \cap \pi^{-1}(SM) \cup \infty, \infty) \to \end{aligned} \tag{74}$$

$$H^i(\pi^{-1}(SM) \cup \infty, SR_2 \cap \pi^{-1}(SM) \cup \infty) = H_{6-i}(\pi^{-1}(SM) - SR_2)$$

$$\begin{aligned} H^i(SR_2 \cap \pi^{-1}(SM) \cup \infty, \infty) &= 0, \quad i \neq 1, 2 \\ &= Z, \quad i = 1, 2, \end{aligned}$$

$\pi^{-1}(SM) - SR_2$ is a fibre space

$$\begin{array}{ccc} (TB_2 & \to & \pi^{-1}(SM) - SR_2 \\ & & \downarrow \\ & & S^1 \sim SM) \end{array}$$

with $E_{p,q}^2 = Z^{A(p)B(q)}$ where

$$\begin{aligned} A(p) &= 1, \quad p = 0, 1 & B(q) &= 1, \quad q = 0, 3 \\ &= 0, \quad p \neq 0, 1 & &= 0, \quad q \neq 0, 3. \end{aligned} \tag{75}$$

We deduce

$$\begin{array}{ccccc} H^5(4 \cup \infty, \infty) \to & & H^5(SN \cup \infty, \infty) & \xrightarrow{k} & H^6(4 \cup \infty, SN \cup \infty) \\ \downarrow & & \downarrow \cong & & \downarrow \\ H^5(4' \cup \infty, \infty) \to & H^5((SN)' \cup \infty, \infty) & \xrightarrow{\text{onto}} & H^6(4' \cup \infty, (SN)' \cup \infty). & \\ & \| & & \| & \\ & Z & & Z & \end{array} \tag{76}$$

The isomorphism and onto map are geometrically interpretable. To proceed to our Mayer–Vietoris sequence we need

$$\begin{aligned} H_i(S'^3 - LR_2 - LR_1) \\ H_i(S'^3 - LR_1 - LR_2) &= H_i(S'^3) \\ &\quad + H_{i-1}(LR_2) + H_{i-2}(S') \\ &\quad + 2H_{i-1}(S^2) \text{ by T7} \\ &= Z^8, \quad i = 3 \\ &= Z, \quad i = 0 \\ &= Z^2, \quad i = 2 \\ &= Z^3, \quad i = 1 \\ &= 0, \quad i \neq 0, 1, 2, 3 \end{aligned}$$

$$\begin{aligned} H^i(\pi^{-1}(SM) \cup \infty, \infty) &= Z, \quad i = 5 \\ &= 0, \quad i = 4 \\ H^i(SN \cup \infty, \pi^{-1}(SM) \cup \infty) &= 0, \quad i = 5 \\ &= Z^2, \quad i = 4. \end{aligned}$$

We then find

$H^5(SN \cup \infty, \infty) = Z$ and $H^4(SN \cup \infty, \infty) = Z^2$

provided in

$$H^3(\pi^{-1}(SM) \cup \infty, \infty) \xrightarrow{f} H^4(SN \cup \infty, \pi^{-1}(SM) \cup \infty),$$

f is the zero map. To show it is, map into

$$\begin{array}{ccc} H^3(\pi^{-1}(SM) \cap SR_2 \cup \infty, \infty) & & \\ \downarrow \cong & & \\ \to H^4(SN \cup \infty, \pi^{-1}(SM) \cap SR_2 \cup \infty) & & \end{array}$$

and note

$$H^3(\pi^{-1}(SM) \cap SR_2 \cup \infty, \infty) = 0.$$

Now to study k in (71). We consider the subset of 4:

$$4' = \begin{cases} \sum_1^4 x_i^2 = 1 \\ y_1 = \cosh \lambda, \; y_2 = i \sinh \lambda, \; iy_3 = y_4, \; \lambda \text{ real} \\ x_1 \cosh \lambda + ix_2 \sinh \lambda + \alpha x_3 y_3 + i\beta x_4 y_3 = 1. \end{cases}$$

There is a nonsingular fibration by π over all $4'$ with fibre S^2 except over $y_3 = 0$. Let $\pi^{-1}(\{y_3 = 0\}) = (SN)'$. Now look at the maps

$$\begin{array}{ccc} (pt. & \to & R_1 \\ & & \downarrow \quad \text{has } E_{p,q}^2 = Z^{A(p)B(q)} \text{ where} \\ & & LR_1) \end{array}$$

$$\begin{aligned} A(p) &= 1, \quad p = 0, 1 & B(q) &= 1, \quad q = 0 \\ &= 4, \quad p = 2 & &= 0, \quad q \neq 0. \\ &= 0, \quad p \neq 0, 1, 2 \end{aligned} \tag{77}$$

$$\begin{array}{ccc} (TB_2 & \to & R_2 \\ & & \downarrow \quad \text{has } E_{p,q}^2 = Z^{A(p)B(q)} \text{ where} \\ & & LR_2 - SM) \end{array}$$

$$\begin{aligned} A(p) &= 2, \quad p = 0, 1 & B(q) &= 1, \quad q = 0, 3 \\ &= 4, \quad p = 2 & &= 0, \quad q \neq 0, 3. \\ &= 0, \quad p \neq 0, 1, 2 \end{aligned} \tag{78}$$

$$\begin{array}{ccc} (S^2 & \to & 4 - SN - R_1 - R_2 \\ & & \downarrow \quad \text{has } E_{p,q}^2 = Z^{A(p)B(q)} \\ & & S'^3 - LR_1 - LR_2) \end{array}$$

where

$$A(p) = 1, \quad p = 0 \qquad B(q) = 1, \quad q = 0, 2$$
$$= 3, \quad p = 1 \qquad = 0, \quad q \neq 0, 2.$$
$$= 2, \quad p = 2$$
$$= 8, \quad p = 3$$
$$= 0, \quad p \neq 0, 1, 2, 3 \tag{79}$$

Finally then, a portion of the Mayer–Vietoris sequence is

$$Z^2 \to Z^8 \to H_5(4 - SN) \to Z^2 \to Z^2 \tag{80}$$
$$H_5(4 - SN) = Z^{10}$$

then $P = Z^9$.

6(b) $\qquad Q = H_5(2 \cap 3 \cap 5 \cap St_2)$

$$2 \cap 3 \cap 5 \cap St_2 = \begin{cases} \sum_1^4 x_i^2 = 1 \\ \sum_1^3 y_i^2 = 1 \end{cases}$$
$$- \{x_1y_1 + \alpha x_2y_2 + \beta x_3y_3 = 1 \tag{81}$$
$$1 \neq \alpha^2 \neq \beta^2 \neq 1$$

$$H_i(2 \cap 3 \cap 5 \cap St_2) = H^{10-i}(2 \cap 3 \cap 5 \cup \infty, 4 \cup \infty)$$

$$\begin{array}{c} H^4(2 \cap 3 \cap 5 \cup \infty, \infty) \to H^4(4 \cup \infty, \infty) \to H^5(2 \cap 3 \cap 5 \cup \infty, 4 \cup \infty) \\ \| \\ 0 \\ \to H^5(2 \cap 3 \cap 5 \cup \infty, \infty) \to. \\ \| \\ Z \end{array} \tag{82}$$

Thus $H^5(2 \cap 3 \cap 5 \cup \infty, 4 \cup \infty) = Z + H^4(4 \cup \infty, \infty)$.

Looking at the singular points of the fibration

(a) none

(b) $$LR_1 = \begin{cases} \sum_1^3 y_i^2 = 1 \\ y_1^2 + \alpha^2 y_2^2 + \beta^2 y_3^2 = 0 \end{cases}$$

contractible nonsingular fibre.

(c) $$LR_2 = \begin{cases} \sum_1^3 y_i^2 = 1 \\ y_1^2 + \alpha^2 y_2^2 + \beta^2 y_3^2 = 1 \end{cases}$$

contractible singular fibre.

SR_2 is the union of the singular points over LR_2. For a description of LR_1 see Appendix I.

$$LR_2 = \{y_2 = \pm[-(1 - \beta^2)/(1 - \alpha^2)]^{\frac{1}{2}} y_3\}$$

LR_2 looks like two S^1's joined on an S^0,

$$\{y_2 = y_3 = 0\} = SM.$$

Define $\quad SN = SR_2 \cup \pi^{-1}(SM), \quad R_1 = \pi^{-1}(LR_1), \quad$ and $\quad R_2 = \pi^{-1}(LR_2 - SM)$.

$$\begin{array}{c} H^3(4 \cup \infty, \infty) \to H^3(SN \cup \infty, \infty) \to H^4(4 \cup \infty, SN \cup \infty) \to H^4(4 \cup \infty, \infty) \\ \| \qquad\qquad\qquad \| \\ 0 \qquad\qquad\qquad 0 \\ \to H^4(SN \cup \infty, \infty). \\ \| \\ Z^2 \end{array} \tag{83}$$

$$H^4(4 \cup \infty, \infty) = H^4(4 \cup \infty, SN \cup \infty).$$

$$\to H^i(SN \cup \infty, \pi^{-1}(SM) \cup \infty) \to H^i(SN \cup \infty, \infty) \to H^i(\pi^{-1}(SM) \cup \infty, \infty) \to. \tag{84}$$

Studying the first term

$$H^i(SN \cup \infty, \pi^{-1}(SM) \cup \infty) = H_{2-i}(SN - \pi^{-1}(SM))$$
$$H_i(SN - \pi^{-1}(SM)) = Z^2, \quad i = 0$$
$$= Z^6, \quad i = 1$$
$$= 0, \quad i \neq 0, 1.$$

Studying the third term

$$\to H^i(\pi^{-1}(SM) \cup \infty, SR_2 \cap \pi^{-1}(SM) \cup \infty)$$
$$\to H^i(\pi^{-1}(SM) \cup \infty, \infty)$$
$$\to H^i(SR_2 \cap \pi^{-1}(SM) \cup \infty, \infty) \to \quad (85)$$
$$H^i(\pi^{-1}(SM) \cup \infty, SR_2 \cap \pi^{-1}(SM) \cup \infty)$$
$$= H_{4-i}(\pi^{-1}(SM) - SR_2)$$
$$H^i(SR_2 \cap \pi^{-1}(SM) \cup \infty, \infty) = Z^2, \quad i = 0$$
$$= 0, \quad i \neq 0.$$

$\pi^{-1}(SM) - SR_2$ is a fibre space

$$(TB_2 \to \pi^{-1}(SM) - SR_2$$
$$\downarrow \text{ with } E^2_{p,q} = Z^{A(p)B(q)} \text{ where}$$
$$S^0)$$

$$A(p) = 2, \quad p = 0 \qquad B(q) = 1, \quad q = 0, 3$$
$$= 0, \quad p \neq 0 \qquad = 0, \quad q \neq 0, 3. \quad (86)$$

We deduce

$$H^4(SN \cup \infty, \infty) = Z^2$$
$$H^3(SN \cup \infty, \infty) = 0.$$

$$(pt. \to R_1$$
$$\downarrow \text{ has } E^2_{p,q} = Z^{A(p)B(q)} \text{ where}$$
$$LR_1)$$

$$A(p) = 1, \quad p = 0 \qquad B(q) = 1, \quad q = 0$$
$$= 5, \quad p = 1 \qquad = 0, \quad q \neq 0. \quad (87)$$
$$= 0, \quad p \neq 0, 1$$

$$(TB_2 \to R_2$$
$$\downarrow \qquad \text{has } E^2_{p,q} = Z^{A(p)B(q)} \text{ where}$$
$$LR_2 - SM)$$

$$A(p) = 1, \quad p = 0 \qquad B(q) = 2, \quad q = 0, 3$$
$$= 3, \quad p = 1 \qquad = 0, \quad q \neq 0, 3. \quad (88)$$
$$= 0, \quad p \neq 0, 1$$

$$(S^2 \to 4 - SN - R_1 - R_2$$
$$\downarrow$$
$$S'^2 - LR_1 - LR_2)$$

has $E^2_{p,q} = Z^{A(p)B(q)}$ where

$$A(p) = 1, \quad p = 0 \qquad B(q) = 1, \quad q = 0, 2$$
$$= 3, \quad p = 1 \qquad = 0, \quad q \neq 0, 2.$$
$$= 10, \quad p = 2$$
$$= 0, \quad p \neq 0, 1, 2 \quad (89)$$

A portion of the relevant Mayer–Vietoris sequence is

$$Z^2 \to Z^{10} \to H_4(4 - SN) \to 0. \quad (90)$$

Then, $Q = Z^{11}$.

Section 7 (R)

$$R = H_4(2 \cap 3 \cap 5 \cap 6 \cap St_2)$$

$$2 \cap 3 \cap 5 \cap 6 \cap St_2 = \begin{cases} \sum_1^3 x_i^2 = 1 \\ \sum_1^3 y_i^2 = 1 \end{cases}$$
$$- \{x_1y_1 + \alpha x_2y_2 + \beta x_3y_3 = 0 \quad (91)$$
$$1 \neq \alpha^2 \neq \beta^2 \neq 1.$$

This will be easy because $2 \cap 3 \cap 5 \cap 6$ in its general position intersects 4 nonsingularly.

$$H_4(2 \cap 3 \cap 5 \cap 6 \cap St_2)$$
$$= H_4(2 \cap 3 \cap 5 \cap 6 - 4)$$
$$= H_4(2 \cap 3 \cap 5 \cap 6)$$
$$+ H_3(2 \cap 3 \cap 4 \cap 5 \cap 6) \text{ by T6}$$
$$= J + L$$

Then, $R = Z^{12}$.

DISCUSSION

If the condition $P_1 + P_2 + P_3 + P_4 = 0$ is imposed, no longer neglecting momentum conservation at vertex V, then when manifold 3 is fixed manifold 5 cannot be moved to infinity by varying the free parameters. Of necessity the rank of the homology group under this condition may not increase over its value when the condition is not imposed. In addition we do not know how many of the possible functions predicted by this calculation are relevant to the Feynman integral. Both of these questions are related to the general problem of calculating the discontinuities around the singularities in terms of a geometric action on the homology group [the action of $\pi_1(T - L)$ on

$$H_8(W_1 \times W_2 - \bigcup_1^7 i)].$$ [1]

APPENDIX I

I.1 $$x_1^2 + \cdots + x_N^2 = 0 - \{\text{origin}\}$$

This manifold is homotopic to its section along $\sum_1^N |x_i|^2 = 2$ by moving points in a radial direction. Breaking x_i into real and imaginary parts $x_i = u_i + v_i$ we get the equations

$$\sum_1^N [(u_i^2 - v_i^2) + 2iu_iv_i] = 0$$
$$\sum_1^N (u_i^2 + v_i^2) = 2.$$

By adding and subtracting these two equations and taking real and imaginary parts we find

$$\sum_1^N u_i^2 = 1; \qquad \sum_1^N v_i^2 = 1; \qquad \sum_1^N u_i v_i = 0.$$

This is clearly the bundle of unit tangent vectors to S^{N-1}.

I.2 $x_1^{\alpha_1} + \cdots + x_N^{\alpha_N} = 1, \qquad \alpha_i$ integers > 0

We evidence a sequence of steps:

(a) Keep real part of $x_i^{\alpha_i}$ fixed and let the imaginary parts of $x_i^{\alpha_i}$ go to zero. All x_i are to be changed simultaneously to keep the sum of imaginary parts fixed at zero. This can be done by breaking x_i into two sets, in the first, imaginary part $x_i^{\alpha_i} \geq 0$, in the second, imaginary part $x_i^{\alpha_i} \leq 0$. Move points in the two sets decreasing the absolute value of the imaginary parts monotonically, moving only points with the largest $|\text{Im } x_i^{\alpha_i}|$ in each set.

(b) All the x_i now lie along lines with angles to the real axis $k\pi/\alpha_i$; k an integer. Now push points on such lines toward the origin monotonically, maintaining $\sum x_i^{\alpha_i} = 1$ until all points on odd k lines are at the origin. This may be done continuously by sending

$$x_i \to rx_i \qquad x_i \in \text{even } k \text{ lines}$$
$$x_i \to sx_i \qquad x_i \in \text{odd } k \text{ lines}$$

r and s real, starting at $r = 1$ and $s = 1$ decreasing r and s monotonically, s to zero, at the same time to keep $\sum x_i^{\alpha_i} = 1$.

(c) Some reflection now shows that what we have left looks like $\{\alpha_1\} * \{\alpha_2\} * \cdots * \{\alpha_N\}$ where $\{\alpha_i\}$ indicates a set of α_i points, and $*$ indicates the topological join operation. This is the wedge of $\prod_i (\alpha_i - 1)S^{N-1}$'s.

I.3
$$\left.\begin{aligned} y_1^2 + y_2^2 + y_3^2 + y_4^2 &= 1 \\ y_1^2 + y_2^2 + \alpha^2 y_3^2 + \beta^2 y_4^2 &= 0 \end{aligned}\right\}$$
$$\Leftrightarrow \begin{cases} y_1^2 + y_2^2 + \dfrac{\alpha^2 - \beta^2}{\alpha^2 - 1} y_4^2 = \dfrac{\alpha^2}{\alpha^2 - 1} \\ y_1^2 + y_2^2 + \dfrac{\beta^2 - \alpha^2}{\beta^2 - 1} y_3^2 = \dfrac{\beta^2}{\beta^2 - 1}. \end{cases}$$

Look at

$$LB_N = \begin{cases} \sum_1^N x_i^2 + y^2 = 1 \\ \sum_1^N x_i^2 + z^2 = a^2 \end{cases} \qquad a^2 \text{ real} > 1.$$

By an argument as in I.2 we can push away the imaginary parts of $x_1, \cdots, x_N$ and z and make y either pure real or pure imaginary. Consider the two possible regions

(a) $\qquad 1 > y^2 > 0$

(b) $\qquad 0 > y^2 > 1 - a^2$

with boundary regions ($y = \pm 1$, $z = \pm a$), ($y = 0$, $z = \pm(a^2 - 1)^{\frac{1}{2}}$), ($y = \pm(1 - a^2)^{\frac{1}{2}}$, $z = 0$). In region (a), things look like $S^{N-1} \cup S^{N-1} \cup S^{N-1} \cup S^{N-1}$ In region (b), things look the same. In boundary region one, like four points, in boundary region two, like $S^{N-1} \cup S^{N-1}$, in boundary region three like $S^{N-1} \cup S^{N-1}$. Observing the situation we see

$$\left.\begin{aligned} H_N(LB_N) &= Z^4 \\ H_1(LB_N) &= Z \end{aligned}\right\} = Z^5 \quad \text{if} \quad N = 1$$
$$H_0(LB_N) = Z$$
$$H_i(LB_N) = 0, \quad i \neq 0, 1, N.$$

APPENDIX II

II.1 $H^4(S^4 \cup \infty, LR_2 \cup \infty) \to H^4(S^4 \cup \infty, \infty)$ is zero.

Look at the homology dual $H_4(S^4 \cup \infty, \infty) \to H_4(S^4 \cup \infty, LR_2 \cup \infty)$ The generator of

$$H_4(S^4 \cup \infty, \infty) = \pi_4(S^4 \cup \infty, \infty)$$
$$E = \begin{cases} y_i = u_i + iv_i, & u_i \text{ and } v_i \text{ real} \\ u_i = 0 & i \neq 5 \\ u_5 = (1 + \sum v_i^2)^{\frac{1}{2}}, & v_5 = 0. \end{cases}$$

We show $\pi_4(S^4 \cup \infty, \infty) \to \pi_4(S^4 \cup \infty, LR_2 \cup \infty)$ is zero by distorting E into LR_2.

Pick a continuous unit vector field on S^3, associating S^3 with the direction of v_i, $n(v_i)$. Now map E into LR_2 by

$$\left.\begin{aligned} u_i &= 0 \\ u_5 &= (1 + \sum v_i^2)^{\frac{1}{2}} \\ v_i, v_5 &= 0 \end{aligned}\right\} \to \begin{cases} \lambda \sum v_i^2 n(v_i) \\ [1 + (1 - \lambda^2) \sum v_i^2]^{\frac{1}{2}} \\ v_i, v_5 = 0 \end{cases}$$

λ ranges continuously from zero to one.

II.2 $H_8(S^4 \times S^4 - 4) \xrightarrow{f} H_8(S^4 \times S^4)$ is zero.

Look at a cycle in $H_8(S^4 \times S^4)$. It looks like a product of multiples of two spherical cycles. Can a cycle in $H_8(S^4 \times S^4 - 4)$ cover such a cycle? For a given y_i ($S^4 \times S^4$ with coordinates x_i, y_i as in paper), $\pi^{-1}(y_i)$ has a spherical cycle if y_i is not on the singular surface $y_5 = -\frac{1}{2}$. However there is no spherical cycle on S^4 that does avoid $y_5 = -\frac{1}{2}$ by II.1. Therefore f cannot map into a cycle of the given type. In the further case, $H^8(S^4 \times S^4 \cap 2 \cup \infty, 4 \cap 2 \cup \infty) \to H^8(S^4 \times S^4 \cap 2 \cup \infty, \infty)$ for example, the analagous cycle can be covered as one can avoid the surfaces such as $y_4 = C$ and construct a cycle of S^3, provided $y_4 = C$ is not one of the two singular surfaces tangent to S^3.

Solution of the Källén–Pauli Equation*

ANTONIO PAGNAMENTA
Department of Physics and Astronomy, University of Maryland, College Park, Maryland
(Received 23 November 1964)

We give the relevant solution to the integral equation which Källén and Pauli have derived for elastic $V\Theta$-scattering in the Lee model. As an application of this result the exact Tamm–Dancoff solution for the entire $V\Theta$-sector is obtained. This includes the amplitudes for $V\Theta$- and $N2\Theta$ elastic scattering and $V\Theta$-$N2\Theta$ production, as well as their extensions off the mass shell.

I. INTRODUCTION

THE Lee model, an interesting model field theory, is generally called exactly soluble because one can obtain expressions in closed form for the renormalization constants and the amplitudes of the first sector. It would however be desirable to find the exact solutions also in the higher sectors, where more complicated processes are possible. This aim has been frustrated by the appearance of a singular-integral equation which is not of the classical type, which has been found in both the off-shell methods that have been applied; namely, the Tamm–Dancoff method and the N quantum approximation. Dispersion theory and infinite-order perturbation theory calculations have yielded exact T-matrix elements in the second sector; however these calculations only give the mass-shell values of the amplitudes, and do not determine the Heisenberg fields.

In Sec. II we give an exact statement of the problem and point out its difficulties. In Sec. III we obtain a solution of the integral equation in question. Using the Tamm–Dancoff method we then obtain the statevector relevant to $V\Theta$-scattering in Sec. IV. The $V\Theta$-amplitude is computed. We derive a new integral equation for the second state vector in this sector, solve it and obtain the $N2\Theta$-elastic scattering amplitude. From the two state vectors together the $V\Theta - N\Theta\Theta$ production amplitude follows.

II. STATEMENT OF THE PROBLEM

Källén and Pauli[1] in their famous paper on the Lee model[2] have shown that the Tamm–Dancoff method leads in the $V\Theta$-sector to the singular linear integral equation.[3]

$$H(\omega_0 - \omega)\psi_1(\mathbf{k}, \mathbf{k}_0) = -\frac{f(\omega)f(\omega_0)}{\omega} - f(\omega)\int \frac{f(\omega')\psi_1(\mathbf{k}', \mathbf{k}_0)}{\omega' - \omega_0 + \omega - i\epsilon} d^3k'. \quad \text{(I)}$$

An equivalent to Eq. (I) has also been mentioned by Lee[2] and again by Heisenberg.[4] It further appeared in related problems of more recent authors[5,6] and also in our attempt to solve the Lee model by the N quantum approximation.[7]

In Eq. (I) we have

$$H(\omega) = \omega G(\omega) \quad (1)$$

and $G(\omega)$ satisfies the once-subtracted dispersion relation

$$G(\omega) = 1 + \frac{\omega}{\pi}\int_\mu^\infty \frac{\text{Im } G(\omega')\, d\omega'}{\omega'(\omega' - \omega - i\epsilon)}. \quad (2)$$

$G(\omega)$ is known in principle since Im $G(\omega)$ is known function. In case of the Lee model we have

$$\text{Im } G(\omega) = 4\pi^2 f^2(\omega) k\Theta(\omega) \quad (3)$$

and we call $f(\omega)$ the cutoff function. Its essential property is that it goes to zero fast enough for large values of the argument in order to make our integrals convergent. The cutoff function is assumed to be only a function of $\omega = (\mathbf{k}^2 + \mu^2)^{\frac{1}{2}}$.

Equation (2) shows that $G(\omega)$ can be extended to an analytic function, regular in the complex ω-

* Work supported in part by the U. S. Air Force under contract AFOSR 500-64 and by the National Science Foundation under contract NSF-FP-3221.

[1] G. Källén and W. Pauli, Kgl. Danske Videnshab. Selskab, Mat. Fys. Medd. 30, No. 7 (1955). This paper will be referred to as KP.

[2] T. D. Lee, Phys. Rev. 95, 1929 (1954).

[3] Equation (62) in KP. Our function $H(\omega)$ is identical with $h(\omega)$ of KP. We also put $f(\omega) = g/(2\pi)^{\frac{1}{2}}(f(\omega)/(2\omega)^{\frac{1}{2}}$, where f is the cutoff function of KP and g the renormalized coupling constant.

[4] W. Heisenberg, Nucl. Phys. 4, 532 (1957).

[5] M. S. Maxon and R. B. Curtis, Bull. Am. Phys. Soc. 9, 86 (1964) and M. S. Maxon, thesis, Indiana University (unpublished).

[6] H. Chew, Phys. Rev. 132, 2756 (1963).

[7] O. W. Greenberg, Bull. Am. Phys. Soc. 9, 448 (1964); A. Pagnamenta, *ibid.*, p. 449.

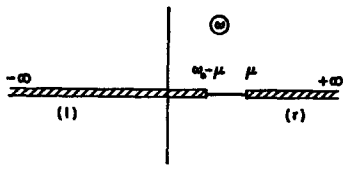


FIG. 1. Showing right-hand cut (r) and left-hand cut (l) of our functions in ω-plane. The integral just runs above (r).

plane, cut along $\mu \leq \omega \leq +\infty$. Again from (2) we derive

$$G(0) = 1$$

$$G(|\infty|) = 1 - \frac{1}{\pi}\int_{\mu}^{\infty} \frac{\operatorname{Im} G(\omega')}{\omega'} d\omega' \equiv Z. \tag{4}$$

If we assume Z, the wavefunction renormalization constant, to be positive (no ghost assumption), then a straightforward discussion[1] shows that $G(\omega)$ has no zeros in the ω-plane with this right-hand cut. Equivalent to (2) is the representation

$$G(\omega) = Z + \frac{1}{\pi}\int_{\mu}^{\infty} \frac{\operatorname{Im} G(\omega')\, d\omega'}{\omega' - \omega - i\epsilon}, \tag{5}$$

where now $\omega = \infty$ is the point of subtraction.

$\psi_1(\mathbf{k}, \mathbf{k}_0)$ is an off-mass-shell extension of the elastic $V\Theta$-scattering amplitude for an infinitely heavy V-particle. $\mathbf{k}_0$ denotes the momentum of the incoming Θ-particle with mass μ, $\mathbf{k}$ the momentum of the outgoing one. Equation (I) has to be solved under the boundary condition that ψ_1 contains only outgoing waves. Therefore here and later a proper $i\epsilon$ has been added in the denominators.

Amado[8] in 1961 using the dispersion theoretic methods of Goldberger and Treiman[9] succeeded in finding the T-matrix for elastic $V\Theta$-scattering which we write in the somewhat different form

$$T(\omega) = \frac{f^2(\omega)}{H(\omega)} \frac{1 + H(\omega)A(\omega)}{1 - H(\omega)A(\omega)}. \tag{6}$$

where

$$A(\omega) = \frac{1}{\pi}\int_{\mu}^{\infty} \left(\operatorname{Im} \frac{1}{H(\omega')}\right) \frac{d\omega'}{H(\omega - \omega')}. \tag{7}$$

This T-matrix is the mass-shell limit of our function ψ_1. The exact relation is

$$T(\omega) = \lim_{\omega_0 \to \omega} (\omega_0 - \omega)\psi_1(\mathbf{k}, \mathbf{k}_0). \tag{8}$$

To get rid of uninteresting factors it is useful for the mathematical part of the discussion to put

$$\psi_1(k, k_0) = f(\omega)f(\omega_0)\psi(\mathbf{k}, \mathbf{k}_0); \tag{9}$$

then Eq. (I) becomes

$$H(\omega_0 - \omega)\psi(\mathbf{k}, \mathbf{k}_0) = \frac{-1}{\omega} - \int \frac{f^2(\omega')\psi(\mathbf{k}', \mathbf{k}_0)}{\omega' - \omega_0 + \omega - i\epsilon} d^3k'. \tag{10}$$

The integral expression in (10) having a Cauchy kernel defines an analytic function in the complex ω-plane, cut on the real axis from $-\infty$ to $\omega_0 - \mu$. Therefore ψ_1, through $\mathbf{k}$ a function of ω, and in the integrand of ω', has the same cut. The integral however runs from $\omega' = \mu$ to $\omega' = +\infty$. So integral and cut do not coincide. (See Fig. 1.) This fact makes it impossible to succeed with the by now classical techniques of Muskhelisvili[10] and Omnes.[11] The particular difficulty is that in contradistinction to the classical problems[12,13] here the related Hilbert problem leads to a functional equation which connects the function at one point to its discontinuity at a different point.

III. THE SOLUTION

One obvious solution of Eq. (I) can be verified immediately:

$$\psi_1^{(1)}(\mathbf{k}, \mathbf{k}_0) = -\delta(\mathbf{k} - \mathbf{k}_0) \tag{11}$$

since

$$(\omega - \omega_0)\delta(\mathbf{k} - \mathbf{k}_0) = 0.$$

However this particular solution is not of great interest. It would lead to a trivial state vector and violate our boundary condition.

To get to the solution of our problem we first observe that Eqs. (I) or (10) are linear integral equations of the third kind. We are only allowed to divide by the factor $H(\omega_0 - \omega)$ if we take proper care of its zeros. Since we have excluded the ghost the only such zero is a simple one at $\omega = \omega_0$ and we may write[14]

$$\psi(\mathbf{k}, \mathbf{k}_0) = (\omega_0 - \omega + i\epsilon)^{-1}\psi_2'(\mathbf{k}, \mathbf{k}_0) + \alpha\delta(\mathbf{k} - \mathbf{k}_0).$$

The coefficient α has to be determined from the boundary conditions. ψ_2' satisfies

$$G(\omega_0 - \omega)\psi_2'(\mathbf{k}, \mathbf{k}_0) = -\frac{1 + \alpha}{\omega} - \int \frac{f^2(\omega')}{\omega_0 - \omega' + i\epsilon} \frac{\psi_2'(\mathbf{k}', \mathbf{k}_0)}{\omega' - \omega_0 + \omega - i\epsilon} d^3k'.$$

[8] R. D. Amado, Phys. Rev. **122**, 697 (1961). Our definition of the T-matrix differs from that of Ref. 8 in that we have not separated out the cutoff function.

[9] M. L. Goldberger and S. B. Treiman, Phys. Rev. **113**, 1663 (1959).

[10] N. J. Muskhelisvili, *Singular Integral Equations* (P. Nordhoff Ltd., Groningen, The Netherlands, 1953).

[11] R. Omnes, Nuovo Cimento **8**, 316 (1958).

[12] R. Blankenbecler and S. Gartenhaus, Phys. Rev. **116**, 1297 (1957).

[13] P. G. Federbusch, M. L. Goldberger, and S. B. Treiman, Phys. Rev. **120**, 1926(1960).

[14] For S-waves, relation (71) may be used.

Put $\psi_2' = (1 + \alpha)\psi_2$, $(1 + \alpha) \neq 0$, then

$$G(\omega_0 - \omega)\psi_2(\mathbf{k}, \mathbf{k}_0) = \frac{-1}{\omega} - \int \frac{f^2(\omega')\psi_2(\mathbf{k}', \mathbf{k}_0)}{(\omega_0 - \omega' + i\epsilon)(\omega' - \omega_0 + \omega - i\epsilon)} d^3k' \tag{12}$$

As a working hypothesis, we may assume that there exists a solution

$$\psi_2 = \phi(\omega, \omega_0)$$

of the above equation that only depends on the ω's. In this case we can do the angular integrals in (12) and get

$$G(\omega_0 - \omega)\phi(\omega, \omega_0) = \frac{-1}{\omega} + \frac{1}{\pi}\int_\mu^\infty \frac{\operatorname{Im} H(\omega')}{\omega' - \omega_0 - i\epsilon} \frac{\phi(\omega', \omega_0)\, d\omega'}{\omega' - \omega_0 + \omega - i\epsilon}. \tag{13}$$

Equation (13) shows that $\phi(\omega, \omega_0)$ is a meromorphic function in the ω-plane with some cuts. Here ω_0 plays the role of a parameter and can be considered as a constant for the process of solution.

We actually solve this equation by constructing an ansatz. To this end we first study the detailed pole-cut structure of $\phi(\omega, \omega_0)$ in the ω-plane, then determine the free constants in the ansatz by substituting it into the equation.

The function $H^{-1}(\omega)$ has the following integral representation[15]

$$\frac{1}{H(\omega)} = \frac{1}{\omega} + \frac{1}{\pi}\int_\mu^\infty \left(\operatorname{Im}\frac{1}{H(\omega')}\right) \frac{d\omega'}{\omega' - \omega - i\epsilon}. \tag{14}$$

This is verified solving the integral by contour integration. The infinite circle gives no contribution and the derivative $H'(0) = 1$, which follows from (1, 2). Multiplying (14) by ω we get

$$\frac{1}{G(\omega)} = 1 + \frac{\omega}{\pi}\int_\mu^\infty \left(\operatorname{Im}\frac{1}{H(\omega')}\right) \frac{d\omega'}{\omega' - \omega - i\epsilon} \tag{15}$$

and substituting $\omega \to \omega_0 - \omega$ in (15)

$$\frac{1}{G(\omega_0 - \omega)} = 1 + \frac{\omega_0 - \omega}{\pi} \times \int_\mu^\infty \left(\operatorname{Im}\frac{1}{H(\omega')}\right) \frac{d\omega'}{\omega' - \omega_0 + \omega - i\epsilon}. \tag{16}$$

This last function has a left-hand cut which only enters the physical range of ω if $\omega_0 > 2\mu$. While $G(\omega)$ is dimensionless $H(\omega)$ has the dimension of an energy. A partial fraction decomposition of the two pole terms in the integrand of (13) gives

$$\frac{1}{(\omega' - \omega_0 - i\epsilon)(\omega' - \omega_0 + \omega - i\epsilon)} = \frac{1}{\omega}\left[\frac{1}{\omega' - \omega_0 - i\epsilon} - \frac{1}{\omega' - \omega_0 + \omega - i\epsilon}\right]. \tag{17}$$

At $\omega = 0$ the square bracket also vanishes, but since we are interested in obtaining an equation for a function with as simple an ω-dependence as possible we now extract an over-all factor ω^{-1}. We put

$$\phi(\omega, \omega_0) = (-1/\omega)\phi_1(\omega, \omega_0)$$

and find for ϕ_1 the integral equation

$$G(\omega_0 - \omega)\phi_1(\omega, \omega_0) = 1 + \frac{1}{\pi}\int_\mu^\infty \frac{\operatorname{Im} G(\omega')}{\omega' - \omega_0 - i\epsilon}\phi_1(\omega', \omega_0)\, d\omega' - \frac{1}{\pi}\int_\mu^\infty \frac{\operatorname{Im} G(\omega')}{\omega' - \omega_0 + \omega - i\epsilon}\phi_1(\omega', \omega_0)\, d\omega'. \tag{18}$$

Multiplying with $G^{-1}(\omega_0 - \omega)$ and using (15) we find

$$\begin{aligned}\phi(\omega, \omega_0) &= 1 + \frac{\omega_0 - \omega}{\pi}\int_\mu^\infty \left(\operatorname{Im}\frac{1}{H(\omega')}\right)\frac{d\omega'}{\omega' - \omega_0 + \omega - i\epsilon} \\ &+ \frac{1}{\pi}\int_\mu^\infty \frac{\operatorname{Im} G(\omega')}{\omega' - \omega_0 - i\epsilon}\phi_1(\omega', \omega_0)\, d\omega' \\ &- \frac{1}{\pi}\int_\mu^\infty \frac{\operatorname{Im} G(\omega')}{\omega' - \omega_0 + \omega - i\epsilon}\phi_1(\omega', \omega_0)\, d\omega' \\ &+ \frac{\omega_0 - \omega}{\pi}\int_\mu^\infty \left(\operatorname{Im}\frac{1}{H(\omega')}\right)\frac{d\omega''}{\omega'' - \omega_0 + \omega - i\epsilon} \\ &\times \frac{1}{\pi}\int_\mu^\infty \frac{\operatorname{Im} G(\omega')}{\omega' - \omega_0 - i\epsilon}\phi_1(\omega', \omega_0)\, d\omega' \\ &- \frac{\omega_0 - \omega}{\pi}\int_\mu^\infty \left(\operatorname{Im}\frac{1}{H(\omega'')}\right)\frac{d\omega''}{\omega'' - \omega_0 + \omega - i\epsilon} \\ &\times \frac{1}{\pi}\int_\mu^\infty \frac{\operatorname{Im} G(\omega')}{\omega' - \omega_0 + \omega - i\epsilon}\phi_1(\omega', \omega_0)\, d\omega'. \end{aligned} \tag{19}$$

Due to the cutoff function all our integrals are uniformly convergent and we may in the following interchange orders of integration. For the last line we may write

$$-\frac{\omega_0 - \omega}{\pi}\int_\mu^\infty \left(\operatorname{Im}\frac{1}{H(\omega'')}\right) d\omega'' \frac{1}{\pi} \times \int_\mu^\infty \frac{\operatorname{Im} G(\omega')\phi_1(\omega', \omega_0)}{(\omega'' - \omega_0 + \omega - i\epsilon)(\omega' - \omega_0 + \omega - i\epsilon)} d\omega'. \tag{20}$$

[15] M. Levy, Nuovo Cimento **13**, 115 (1959).

Next we decompose

$$\frac{1}{(\omega' - \omega_0 + \omega - i\epsilon)(\omega'' - \omega_0 + \omega - i\epsilon)} = \frac{1}{\omega' - \omega'' + i\epsilon}\left[\frac{1}{\omega'' - \omega_0 + \omega - i\epsilon} - \frac{1}{\omega' - \omega_0 + \omega - i\epsilon}\right]. \tag{21}$$

Since for $\omega' = \omega''$ the square bracket also vanishes, there is actually no singularity in the entire expression and we may select this zero as we wish in order to make the single integrals well defined. We did add $+i\epsilon$, since then the following calculation becomes shortest.

To get the left-hand cut in ω explicitly we rearrange (20):

$$-\frac{\omega_0 - \omega}{\pi}\int_\mu^\infty \left(\operatorname{Im}\frac{1}{H(\omega'')}\right)\frac{d\omega''}{\omega'' - \omega_0 + \omega - i\epsilon} \times \frac{1}{\pi}\int_\mu^\infty \frac{\operatorname{Im} G(\omega')}{\omega' - \omega'' + i\epsilon}\phi_1(\omega', \omega_0)\, d\omega'$$

$$-\frac{\omega_0 - \omega}{\pi}\int_\mu^\infty \frac{\operatorname{Im} G(\omega')}{\omega' - \omega_0 + \omega - i\epsilon}\phi_1(\omega', \omega_0)\, d\omega' \times \frac{1}{\pi}\int_\mu^\infty \left(\operatorname{Im}\frac{1}{H(\omega'')}\right)\frac{d\omega''}{\omega'' - \omega' - i\epsilon}.$$

The very last integral can be done, using (14). Since

$$\frac{1}{\omega' - \omega_0 + \omega - i\epsilon} = \frac{1}{\omega'} + \frac{\omega_0 - \omega}{\omega'(\omega' - \omega_0 + \omega - i\epsilon)}$$

the third integral in (19) can be written

$$-\frac{1}{\pi}\int_\mu^\infty \frac{\operatorname{Im} G(\omega')}{\omega' - \omega_0 + \omega - i\epsilon}\phi_1(\omega', \omega_0)\, d\omega' = \frac{-1}{\pi}\int_\mu^\infty \frac{\operatorname{Im} G(\omega')}{\omega'}\phi_1(\omega', \omega_0)\, d\omega' - \frac{\omega_0 - \omega}{\pi}\int_\mu^\infty \frac{\operatorname{Im} G(\omega')\phi_1(\omega', \omega_0)}{\omega'(\omega' - \omega_0 + \omega - i\epsilon)}\, d\omega'.$$

Now we collect all the terms

$$\phi_1(\omega, \omega_0) = 1 + \frac{1}{\pi}\int_\mu^\infty \frac{\operatorname{Im} G(\omega')}{\omega' - \omega_0 - i\epsilon}\phi_1(\omega', \omega_0)\, d\omega' - \frac{1}{\pi}\int_\mu^\infty \frac{\operatorname{Im} G(\omega')}{\omega'}\phi_1(\omega', \omega_0)\, d\omega' + \frac{\omega_0 - \omega}{\pi}\int_\mu^\infty \left(\operatorname{Im}\frac{1}{H(\omega')}\right)\frac{d\omega'}{\omega' - \omega_0 + \omega - i\epsilon} - \frac{\omega_0 - \omega}{\pi}\int_\mu^\infty \frac{\operatorname{Im} G(\omega')\phi_1(\omega', \omega_0)}{\omega'(\omega' - \omega_0 + \omega - i\epsilon)}\, d\omega'$$

$$+ \frac{\omega_0 - \omega}{\pi}\int_\mu^\infty \left(\operatorname{Im}\frac{1}{H(\omega')}\right)\frac{d\omega'}{\omega' - \omega_0 + \omega - i\epsilon} \times \frac{1}{\pi}\int_\mu^\infty \frac{\operatorname{Im} G(\omega'')}{\omega'' - \omega_0 - i\epsilon}\phi_1(\omega'', \omega_0)\, d\omega''$$

$$- \frac{\omega_0 - \omega}{\pi}\int_\mu^\infty \left(\operatorname{Im}\frac{1}{H(\omega')}\right)\frac{d\omega'}{\omega' - \omega_0 + \omega - i\epsilon} \times \frac{1}{\pi}\int_\mu^\infty \frac{\operatorname{Im} G(\omega'')}{\omega'' - \omega' + i\epsilon}\phi_1(\omega'', \omega_0)\, d\omega''$$

$$- \frac{\omega_0 - \omega}{\pi}\int_\mu^\infty \frac{\operatorname{Im} G(\omega')}{H(\omega')}\frac{\phi_1(\omega', \omega_0)}{\omega' - \omega_0 + \omega - i\epsilon}\, d\omega' + \frac{\omega_0 - \omega}{\pi}\int_\mu^\infty \frac{\operatorname{Im} G(\omega')\phi_1(\omega', \omega_0)}{\omega'(\omega' - \omega_0 + \omega - i\epsilon)}\, d\omega'. \tag{22}$$

The first two integrals can be combined, the fourth cancels against the eighth. Using

$$\operatorname{Im}\,[1/H(\omega')] = -\operatorname{Im} H(\omega')/H(\omega')H^*(\omega')$$

in the seventh integrand, where H^* is the complex conjugate of H, and ordering the terms after their ω-dependence, we obtain

$$\phi_1(\omega, \omega_0) = 1 + \frac{\omega_0}{\pi}\int_\mu^\infty \frac{\operatorname{Im} G(\omega')}{\omega'(\omega' - \omega_0 - i\epsilon)}\phi_1(\omega', \omega_0)\, d\omega' + \frac{\omega_0 - \omega}{\pi}\int_\mu^\infty \left(\operatorname{Im}\frac{1}{H(\omega')}\right) \times \frac{d\omega'}{\omega' - \omega_0 + \omega - i\epsilon}X(\omega', \omega_0), \tag{23}$$

where

$$X(\omega', \omega_0) = 1 + G^*(\omega')\phi_1(\omega', \omega_0) + \frac{1}{\pi}\int_\mu^\infty \frac{\operatorname{Im} G(\omega'')}{\omega'' - \omega_0 - i\epsilon}\phi_1(\omega'', \omega_0)\, d\omega'' - \frac{1}{\pi}\int_\mu^\infty \frac{\operatorname{Im} G(\omega'')}{\omega'' - \omega' + i\epsilon}\phi_1(\omega'', \omega_0)\, d\omega''. \tag{24}$$

Put

$$C_1(\omega_0) = 1 + \frac{\omega_0}{\pi}\int_\mu^\infty \frac{\operatorname{Im} G(\omega')\phi_1(\omega', \omega_0)}{\omega'(\omega' - \omega_0 - i\epsilon)}\, d\omega' \tag{25}$$

and observe that ϕ_1 has the form

$$\phi_1(\omega, \omega_0) = C_1(\omega_0) + \frac{\omega_0 - \omega}{\pi}\int_\mu^\infty \left(\operatorname{Im}\frac{1}{H(\omega')}\right)\frac{X(\omega', \omega_0)\, d\omega'}{\omega' - \omega_0 + \omega - i\epsilon}. \tag{26}$$

$C_1(\omega_0)$ is a constant in ω. Substituting (23) and the complex conjugate of (2) for $G^*\phi_1$ in (24) one can see that the apparent right-hand cut of $X(\omega', \omega_0)$ cancels; therefore $X(\omega', \omega_0)$ is a dimensionless an-

alytic function of ω' having only the left-hand cut of $\phi_1(\omega', \omega_0)$. To find $X(\omega', \omega_0)$ we have to guess. There appeared only two dimensionless functions with left-hand cut in our calculation namely $G(\omega_0 - \omega')$ and $G^{-1}(\omega_0 - \omega')$. We try

$$X(\omega', \omega_0) = C_2(\omega_0)G^{-1}(\omega_0 - \omega'). \tag{27}$$

Substituting this in (26) we have

$$\phi_1(\omega, \omega_0) = C_1(\omega_0) + C_2(\omega_0)\frac{\omega_0 - \omega}{\pi}\int_\mu^\infty \left(\operatorname{Im}\frac{1}{H(\omega')}\right) \times \frac{d\omega'}{G(\omega_0 - \omega')(\omega' - \omega_0 + \omega - i\epsilon)} \tag{28}$$

and this ansatz is consistent if we can find C_1 and C_2 independent of ω' and ω such that it fulfills (18).

It is convenient to introduce

$$I_0(\omega) = \frac{-1}{\pi}\int_\mu^\infty \left(\operatorname{Im}\frac{1}{H(\omega')}\right) \times \frac{d\omega'}{G(\omega_0 - \omega')(\omega' - \omega - i\epsilon)} \tag{29}$$

and to observe that $I_0(\omega_0) = A(\omega_0)$ of Eq. (7). Our ansatz (28) becomes

$$\phi_1(\omega, \omega_0) = C_1(\omega_0) - C_2(\omega_0)(\omega_0 - \omega)I_0(\omega_0 - \omega). \tag{30}$$

There are several ways to find C_1 and C_2. We find it most convenient to insert (30) directly into the integral equation (18) for ϕ_1. Using (5) for $G(\omega)$ this gives

$$\begin{aligned} &C_1G(\omega_0 - \omega) - C_2H(\omega_0 - \omega)I_0(\omega_0 - \omega) \\ &= 1 + C_1G(\omega_0) - C_1G(\omega_0 - \omega) \\ &\quad + C_2\omega[G(\omega_0 - \omega)I_0(\omega) + H^{-1}(\omega) - \omega^{-1}]. \end{aligned} \tag{31}$$

The square bracket comes from the fact that, successively,

$$\begin{aligned} J &\equiv \frac{-1}{\pi}\int_\mu^\infty \operatorname{Im} G(\omega')\left[\frac{1}{\omega' - \omega_0 - i\epsilon} - \frac{1}{\omega' - \omega_0 + \omega - i\epsilon}\right](\omega_0 - \omega')I_0(\omega - \omega')\,d\omega' \\ &= \frac{\omega}{\pi}\int_\mu^\infty \frac{\operatorname{Im} G(\omega')}{\omega' - \omega_0 + \omega - i\epsilon} I_0(\omega_0 - \omega')\,d\omega' \\ &= \omega[G(\omega_0 - \omega)I_0(\omega) + H^{-1}(\omega) - \omega^{-1}], \end{aligned} \tag{32}$$

which we show in Appendix I. In (31) we divide by $G(\omega_0 - \omega)$ and get

$$\begin{aligned} &2C_1 - (\omega_0 - \omega)C_2I_0(\omega_0 - \omega) \\ &= [1 + C_1G(\omega_0)]/G(\omega_0 - \omega) + -C_2/G(\omega_0 - \omega) \\ &\quad + \omega C_2[I_0(\omega) + \{1/H(\omega)G(\omega_0 - \omega)\}]. \end{aligned} \tag{33}$$

We further prove in Appendix I

$$\omega\{I_0(\omega) + [H(\omega)G(\omega_0 - \omega)]^{-1}\} = G(\omega_0)^{-1} + \omega_0A(\omega_0) - (\omega_0 - \omega)I_0(\omega_0 - \omega). \tag{34}$$

Using this, the terms in $I_0(\omega_0 - \omega)$ cancel and when we multiply by $G(\omega_0)G(\omega_0 - \omega)$, (33) becomes

$$\begin{aligned} &[2C_1G(\omega_0) - C_2 - C_2H(\omega_0)A(\omega_0)]G(\omega_0 - \omega) \\ &\quad = G(\omega_0)[1 + C_1G(\omega_0) - C_2]. \end{aligned} \tag{35}$$

This equation can only hold identically in ω if each side separately is zero. Therefore we obtain for C_1 and C_2 the system

$$\begin{aligned} 2C_1G(\omega_0) - C_2(1 + H(\omega_0)A(\omega_0)) &= 0 \\ -C_1G(\omega_0) + C_2 &= 1. \end{aligned}$$

We solve it to find

$$\begin{aligned} C_1(\omega_0) &= \frac{1}{G(\omega_0)}\frac{1 + H(\omega_0)A(\omega_0)}{1 - H(\omega_0)A(\omega_0)}, \\ C_2(\omega_0) &= 2/[1 - H(\omega_0)A(\omega_0)], \end{aligned} \tag{36}$$

both independent of ω. This shows that our ansatz is consistent and verifies the assumption that there exists a solution $\phi(\omega, \omega_0)$ which does not depend on the vectors. We get

$$\phi(\omega, \omega_0) = \frac{-1}{\omega}\left[\frac{1}{G_0}\frac{1 + H_0A_0}{1 - H_0A_0} - \frac{2(\omega_0 - \omega)}{1 - H_0A_0}I_0(\omega_0 - \omega)\right], \tag{37}$$

where $H_0 = H(\omega_0)$ and $A_0 = A(\omega_0)$.[16]

We could still add to ϕ as the solution of Eq. (15) any amount of a solution ϕ_0 of the corresponding homogeneous equation.[17]

$$G(\omega_0 - \omega)\phi_0(\omega, \omega_0) = \frac{1}{\pi}\int_\mu^\infty \frac{\operatorname{Im} H(\omega')}{\omega' - \omega_0 - i\epsilon}\frac{\phi_0(\omega', \omega_0)\,d\omega'}{\omega' - \omega_0 + \omega - i\epsilon}. \tag{38}$$

We have not been able to find any nontrivial solution of this equation. If we remark that Im $H(\omega')$ is proportional to g^2, the coupling constant, we realize

[16] After this work was completed[7] we were informed that R. P. Kenschaft and R. D. Amado also have found a particular solution of Eq. (13). R. P. Kenschaft and R. D. Amado, J. Math. Phys. 5, 1340 (1964). Their result agrees with our solution (37). I want to thank them for sending me a preprint prior to publication.

[17] Having found the two clearly linearly independent solutions (11) and, with some factors, (37) of Eq. (I) we can construct a solution of its homogeneous equation. This has taken care of by the factor α. After dividing out $H(\omega_0 - \omega)$ one could still add any amount of a solution of the new homogeneous equation (38).

that this homogeneous equation cannot have any solution which for small g^2 has a power-series expansion in g^2, such as, for instance, our solution of the inhomogeneous equation. Also for ω_0 large and negative, the term with the integral of Eq. (18) can be made arbitrarily small. The equation can then, in principle, be solved by a convergent process of iteration. In this case it has again our unique solution.

This does not exclude the existence of homogeneous solutions for particular values of g^2 and ω_0. Here we assume the T-matrix to be an analytic function in g^2 around $g^2 = 0$. We expect little scattering for small g^2. Then, for small enough g^2, the homogeneous equation has no solution and the most general solution of the KP-equation (I) is given by

$$\psi_1(\mathbf{k}, \mathbf{k}_0) = [(1 + \alpha)f(\omega)f(\omega_0)/(\omega_0 - \omega + i\epsilon)] \times \phi(\omega, \omega_0) + \alpha\delta(\mathbf{k} - \mathbf{k}_0) \quad (39)$$

with the ϕ of Eq. (37).

IV. THE $V\Theta$-SECTOR

Having solved Eq. (I) we can, as the natural application of result and method, give the exact Tamm–Dancoff solution for the entire $V\Theta$-sector in the Lee model. The $V\Theta$-elastic scattering amplitude is obtained by imposing the proper boundary conditions in (39). For the $N2\Theta$ elastic amplitude we derive a new integral equation which can be solved with the knowledge gained in Sec. III. Knowing the two relevant state vectors the production amplitude follows as a byproduct.

Again proceeding from KP, to which we refer for notation and details, we write the Lee model Hamiltonian in renormalized form

$$\mathfrak{H} = \mathfrak{H}_0' + \mathfrak{H}_1'$$

$$\mathfrak{H}_0' = ZmV^\dagger V + mN^\dagger N + \int \omega(\mathbf{k})a^\dagger(\mathbf{k})a(\mathbf{k})\, d^3k \quad (40)$$

$$\mathfrak{H}_1' = \int f(\omega_k)[V^\dagger N a_\mathbf{k} + a_\mathbf{k}^\dagger N^\dagger V]\, d^3k - Z\delta m V^\dagger V.$$

The wavefunction renormalization constant Z has been given in (4). KP further find, written in our notation,

$$Z\delta m = -\int \frac{f^2(\omega')}{\omega'} d^3k' = -\int_\mu^\infty \mathrm{Im}\, G(\omega')\, d\omega'. \quad (41)$$

We use

$$[V, V^\dagger]_+ = Z^{-1} \quad (42)$$

and

$$[N, N^\dagger]_+ = 1, \qquad [a_\mathbf{k}, a_{\mathbf{k}'}^\dagger] = \delta(\mathbf{k} - \mathbf{k}') \quad (43)$$

all other (anti) commutators being trivial. We construct the bare Tamm–Dancoff states by

$$|N\rangle = N^\dagger\, |0\rangle,$$
$$|a_\mathbf{k}\rangle = a_\mathbf{k}^\dagger\, |0\rangle;$$

but

$$|V\rangle = Z^{\frac{1}{2}}V^\dagger\, |0\rangle$$

and

$$|N\Theta_1\Theta_2\rangle = a_{\mathbf{k}_1}^\dagger a_{\mathbf{k}_2}^\dagger\, |0\rangle, \quad (45)$$

so they are all normalized to one, or a δ-function in the continuum. We take $k_1 \neq k_2$.

For the scattering of a V-particle on a Θ-particle we want an eigenstate of the total Hamiltonian of the form

$$|(V\Theta_{\mathbf{k}_0})^{\mathrm{in}}\rangle = a_{\mathbf{k}_0}^\dagger\, |\mathbf{V}\rangle + |\chi\rangle^+ \quad (46)$$

with

$$\mathfrak{H}\, |(V\Theta_{\mathbf{k}_0})^{\mathrm{in}}\rangle = (m + \omega_0)\, |(V\Theta_{\mathbf{k}_0})^{\mathrm{in}}\rangle, \quad (47)$$

$|\mathbf{V}\rangle$ denotes the stationary eigenstate of the total Hamiltonian describing a physical V-particle of mass m. In terms of Tamm–Dancoff states, it is

$$|\mathbf{V}\rangle = Z^{\frac{1}{2}}\, |V\rangle - \int \frac{f(\omega')}{\omega'} |N\Theta_{\mathbf{k}'}\rangle\, d^3k', \quad (48)$$

and $|\chi\rangle^+$—the scattering part—can be written

$$|\chi\rangle^+ = \int \psi_1(\mathbf{k}', \mathbf{k}_0)Z^{\frac{1}{2}}\, |V\Theta_{\mathbf{k}'}\rangle\, d^3k' + \int \psi_2(\mathbf{k}', \mathbf{k}''; \mathbf{k}_0)\, |N\Theta_{\mathbf{k}'}\Theta_{\mathbf{k}''}\rangle\, d^3k'\, d^3k'' \quad (49)$$

with outgoing waves only in ψ_1 and ψ_2. Using the successive relations

$$\begin{aligned}\mathfrak{H}a_{\mathbf{k}_0}^\dagger\, |\mathbf{V}\rangle &= a_{\mathbf{k}_0}^\dagger \mathfrak{H}\, |\mathbf{V}\rangle + [\mathfrak{H}, a_{\mathbf{k}_0}^\dagger]\, |\mathbf{V}\rangle \\ &= ma_{\mathbf{k}_0}^\dagger\, |\mathbf{V}\rangle + (\omega_0 a_\mathbf{k}^\dagger + f(\omega_0)V^\dagger N)\, |\mathbf{V}\rangle \\ &= (m + \omega_0)a_{\mathbf{k}_0}^\dagger\, |\mathbf{V}\rangle \\ &\quad - \frac{f(\omega_0)}{\sqrt{Z}} \int \frac{f(\omega')}{\omega'} |V\Theta_{\mathbf{k}'}\rangle\, d^3k', \end{aligned} \quad (50)$$

the Schrödinger eigenvalue equation gives

$$\begin{aligned} &Z(\omega_0 - \omega + \delta m)\psi_1(\mathbf{k}, \mathbf{k}_0) \\ &\quad = \frac{f(\omega_0)f(\omega)}{\omega} + 2\int f(\omega')\psi_2(\mathbf{k}', \mathbf{k}; \mathbf{k}_0)\, d^3k' \\ &(\omega_0 - \omega_1 - \omega_2)\psi_2(\mathbf{k}_1, \mathbf{k}_2; \mathbf{k}_0) \\ &\quad = (\tfrac{1}{2})[f(\omega_1)\psi_1(\mathbf{k}_2; \mathbf{k}_0) + f(\omega_2)\psi_1(\mathbf{k}_1; \mathbf{k}_0)]. \end{aligned} \quad (51)$$

From the last equation together with our boundary condition (49)

$$\psi_2(\mathbf{k}_1, \mathbf{k}_2; \mathbf{k}_0) = [2(\omega_0 - \omega_1 - \omega_2 + i\epsilon)]^{-1} \times [f(\omega_1)\psi_1(\mathbf{k}_2; \mathbf{k}_0) + f(\omega_2)\psi_1(\mathbf{k}_1; \mathbf{k}_0)], \quad (52)$$

and substituting this into the first equation of (51) we find

$$\left[Z(\omega_0 - \omega) + Z\delta m + \int \frac{f^2(\omega')\, d^3k'}{\omega' - \omega_0 + \omega - i\epsilon}\right]\psi_1(\mathbf{k}, \mathbf{k}_0) = -\frac{f(\omega)f(\omega_0)}{\omega} - f(\omega)\int \frac{f(\omega')\psi_1(\mathbf{k}', \mathbf{k}_0)}{\omega' - \omega_0 + \omega - i\epsilon}\, d^3k'. \quad (53)$$

Using (41) and

$$H(\omega) = Z\omega + Z\delta m + \int \frac{f^2(\omega')\, d^3k'}{\omega' - \omega - i\epsilon}, \quad (54)$$

which is easily checked and is already implicit in KP, one obtains for ψ_1 the integral equation (I).

The most general solution of it (under the restrictions mentioned) has been given in (39). If we insert this and the relevant parts of (48), (49) in (46) we obtain

$$|(V\Theta_{\mathbf{k}_0})^{\text{in}}\rangle = Z^{\frac{1}{2}}(|V\Theta_{\mathbf{k}_0}\rangle + \alpha\,|V\Theta_{\mathbf{k}_0}\rangle) + \cdots.$$

This shows that any $\alpha \neq 0$ would change the normalization of the plane wave contribution to the state. Therefore our boundary condition implies $\alpha = 0$ and the relevant solution of the KP-equation (I) is

$$\psi_1(\omega, \omega_0) = \frac{f(\omega)f(\omega_0)}{\omega(1 - H_0A_0)} \times \left[\frac{1}{\omega - \omega_0 - i\epsilon}\frac{1 + H_0A_0}{G(\omega_0)} + 2I_0(\omega_0 - \omega)\right] \quad (\text{I}')$$

or, somewhat handier,

$$\psi_1(\omega, \omega_0) = \frac{f(\omega)f(\omega_0)}{\omega}\left[\frac{C_1}{\omega - \omega_0 - i\epsilon} + C_2I_0(\omega_0 - \omega)\right]$$

with C_1 and C_2 given in (36).

Using this in (52) we find for ψ_2

$$\psi_2(\omega_1, \omega_2; \omega_0) = \frac{f(\omega_0)f(\omega_1)f(\omega_2)}{\sqrt{2}\,(\omega_0 - \omega_1 - \omega_2 + i\epsilon)} \times \left[C_1(\omega_0)\left(\frac{1}{\omega_1(\omega_1 - \omega_0 - i\epsilon)} + \frac{1}{\omega_2(\omega_2 - \omega_0 - i\epsilon)}\right) + C_2(\omega_0)\left(\frac{1}{\omega_1}I_0(\omega_0 - \omega_1) + \frac{1}{\omega_2}I_0(\omega_0 - \omega_2)\right)\right], \quad (55)$$

where we have indicated that ψ_2 as ψ_1 depends on the energies only.

The state vector corresponding to an incoming V and Θ-particle can now be written down and from it the $V\Theta$-scattering matrix can be computed. We use the definitions

$$S_{V\Theta} = \langle(V\Theta_{\mathbf{k}_1})^{\text{out}} \mid (V\Theta_{\mathbf{k}})^{\text{in}}\rangle \quad (56a)$$

$$= \delta(\mathbf{k} - \mathbf{k}_1) + 2\pi i\,\delta(\omega - \omega_1)T(\omega). \quad (56b)$$

To compute the S-matrix by use of (56a) and explicit substitution of (I') and (55) is possible but extremely lengthy. Therefore we give a different derivation for $T(\omega)$ in Appendix II. Similar to the $N\Theta$-sector (where there is no renormalization) one finds that the T-matrix is given by the residue of ψ_1 in (I') at the pole for $\omega = \omega_0$. This fact we have expressed in relation (8). Applying this to (I') we see that on the energy shell the second term, which carries the left-hand cut, drops out. Due to the δ-function in (56b) the residue of the first term reduces to expression (6) in agreement with Amado's result.

To discuss $N2\Theta$ scattering we have to find the eigenstate of the Hamiltonian of the form

$$|(N\Theta_1\Theta_2)^{\text{in}}\rangle = |N\Theta_1\Theta_2\rangle + \int \phi_1(\mathbf{k}'; \mathbf{k}_1\mathbf{k}_2)Z^{\frac{1}{2}}\,|V\Theta_{\mathbf{k}'}\rangle\, d^3k' + \int \phi_2(\mathbf{k}', \mathbf{k}''; \mathbf{k}_1, \mathbf{k}_2)\,|N\Theta_{\mathbf{k}'}\Theta_{\mathbf{k}''}\rangle\, d^3k'\, d^3k'', \quad (57)$$

with eigenvalue $m + \omega_1 + \omega_2$ and outgoing waves only in ϕ_1 and ϕ_2. The Schrödinger equation leads to the system ($\omega_3 = \omega_1 + \omega_2$):

$$Z(\omega_3 - \omega + \delta m)\phi_1(\mathbf{k}; \mathbf{k}_1, \mathbf{k}_2) = \tfrac{1}{2}f(\omega_1)\delta(\mathbf{k} - \mathbf{k}_1) + \tfrac{1}{2}f(\omega_2)\delta(\mathbf{k} - \mathbf{k}_2) + 2\int \phi_2(\mathbf{k}, \mathbf{k}'; \mathbf{k}_1, \mathbf{k}_2)f(\omega')\, d^3k',$$

$$(\omega_3 - \omega' - \omega'')\phi_2(\mathbf{k}', \mathbf{k}''; \mathbf{k}_1, \mathbf{k}_2) = \tfrac{1}{2}f(\omega'')\phi_1(\mathbf{k}'; \mathbf{k}_1, \mathbf{k}_2) + \tfrac{1}{2}f(\omega')\phi_1(\mathbf{k}''; \mathbf{k}_1, \mathbf{k}_2). \quad (58)$$

Solving the second equation and substituting into the first one, using (54) we find this time for ϕ_1 the singular integral equation

$$H(\omega_3 - \omega)\phi_1(\mathbf{k}; \mathbf{k}_1, \mathbf{k}_2) = \tfrac{1}{2}f(\omega_1)\delta(\mathbf{k} - \mathbf{k}_2) + \tfrac{1}{2}f(\omega_2)\delta(\mathbf{k} - \mathbf{k}_1) - f(\omega)\int \frac{f(\omega')\phi_1(\mathbf{k}'; \mathbf{k}_1, \mathbf{k}_2)}{\omega' - \omega_3 + \omega - i\epsilon}\, d^3k'. \quad (59)$$

It is convenient to make the substitution

$$\phi_1(\mathbf{k}; \mathbf{k}_1, \mathbf{k}_2) = \tfrac{1}{2}H^{-1}(\omega_3 - \omega) \times [f(\omega_1)\delta(\mathbf{k} - \mathbf{k}_2) + f(\omega_2)\delta(\mathbf{k} - \mathbf{k}_1)] - [f(\omega)f(\omega_1)f(\omega_2)/2\,H(\omega_1)H(\omega_2)]\chi(\mathbf{k}; \mathbf{k}_1, \mathbf{k}_2). \quad (60)$$

Then (59) becomes

$$H(\omega_3-\omega)\chi(\mathbf{k};\mathbf{k}_1,\mathbf{k}_2)=\frac{H_1}{\omega-\omega_2-i\epsilon}+\frac{H_2}{\omega-\omega_1-i\epsilon}-\int\frac{f^2(\omega')\chi(\mathbf{k}';\mathbf{k}_1,\mathbf{k}_2)}{\omega'-\omega_3+\omega-i\epsilon}\,d^3k'. \tag{II}$$

This integral equation has a certain similarity to Eq. (I) or better, Eq. (10). Their homogeneous equations are the same. Also we may discard immediately the δ-function solution recalling the above boundary condition. The inhomogeneous term introduces a somewhat more complicated analytic structure than we had in (I). But using the technique of Sec. III and the relations in the Appendix, this is the justification for having been so explicit there, we can still find the solution.[18]

A detailed study of the analytic structure of χ leads us to the ansatz

$$\chi(\omega;\omega_1,\omega_2)=\frac{1}{\omega_3-\omega+i\epsilon}\left[\frac{\omega_1}{\omega-\omega_2-i\epsilon}+\frac{\omega_2}{\omega-\omega_1-i\epsilon}+\frac{B}{\omega}\right]+\frac{C}{\omega}I_3(\omega_3-\omega), \tag{61}$$

where $I_3(\omega)=I_0(\omega)$ for $\omega_0=\omega_3$ in (29). B and C are constant in ω. They have to be determined such that (61) satisfies (60). Substitution in (60) and doing the integrals, where one again will use the relations in the Appendix, we see that the ansatz works if

$$B=2\omega_3H(\omega_3)A(\omega_3)/[1-H(\omega_3)A(\omega_3)] \tag{62}$$

and

$$C=-2H(\omega_3)/[1-H(\omega_3)A(\omega_3)].$$

Therefore [$H_3=H(\omega_3)$, $A_3=A(\omega_3)$]

$$\chi(\omega;\omega_1,\omega_2)=\frac{1}{\omega_3-\omega+i\epsilon}\left[\frac{\omega_1}{\omega-\omega_2-i\epsilon}+\frac{\omega_2}{\omega-\omega_1-i\epsilon}+\frac{\omega_3}{\omega}\frac{2H_3A_3}{1-H_3A_3}\right]-\frac{2H_3}{\omega}\frac{I_3(\omega_3-\omega)}{1-H_3A_3}. \tag{II'}$$

Using this and (60) in the second equation of (58), we find

$$\phi_2(\mathbf{k}',\mathbf{k}'';\mathbf{k}_1,\mathbf{k}_2)=\frac{1}{2}\left[\frac{f(\omega_1)f(\omega'')\delta(\mathbf{k}'-\mathbf{k}_2)}{(\omega_1-\omega''+i\epsilon)H(\omega_1)}+\frac{f(\omega_1)f(\omega')\delta(\mathbf{k}''-\mathbf{k}_2)}{(\omega_1-\omega'+i\epsilon)H(\omega_1)}+\frac{f(\omega_2)f(\omega')\delta(\mathbf{k}''-\mathbf{k}_1)}{(\omega_2-\omega'+i\epsilon)H(\omega_2)}+\frac{f(\omega_2)f(\omega'')\delta(\mathbf{k}'-\mathbf{k}_1)}{(\omega_2-\omega''+i\epsilon)H(\omega_2)}\right]-\frac{f(\omega_1)f(\omega_2)f(\omega')f(\omega'')}{2(\omega_3-\omega'-\omega''+i\epsilon)H_1H_2}\times[\chi(\omega';\omega_1,\omega_2)+\chi(\omega'';\omega_1,\omega_2)]. \tag{63}$$

With this we have obtained the relevant state vector. Call $\tilde{M}_2$ the elastic $N2\Theta$ T-matrix element defined by

$$S_{N2\Theta}=\langle(N\Theta_{\omega'}\Theta_{\omega''})^{\text{out}}\mid(N\Theta_{\omega_1}\Theta_{\omega_2})^{\text{in}}\rangle=\tfrac{1}{2}[\delta(\mathbf{k}'-\mathbf{k}_1)\delta(\mathbf{k}''-\mathbf{k}_2)+\delta(\mathbf{k}'-\mathbf{k}_2)\delta(\mathbf{k}''-\mathbf{k}_1)]+2\pi i\delta(\omega_1+\omega_2-\omega'-\omega'')\tilde{M}_2(\omega_1,\omega_2;\omega'). \tag{64}$$

We read $\tilde{M}_2$ off as the residue of the corresponding pole in $\phi_2(\mathbf{k}',\mathbf{k}'';\mathbf{k}_1,\mathbf{k}_2)$. We have taken care of the identity of the bosons Θ_1 and Θ_2 in (45). This gives

$$\begin{aligned}\delta(\omega_1+\omega_2-\omega'-\omega'')\tilde{M}_2(\omega_1,\omega_2;\omega')&=\delta(\omega_1+\omega_2-\omega'-\omega'')M_2(\omega_1,\omega_2;\omega')\\&\quad-[\delta(\mathbf{k}'-k_2)\delta(\omega''-\omega_2)\\&\quad+\delta(\mathbf{k}''-k_2)\delta(\omega'-\omega_1)]f^2(\omega')/2H(\omega_1)\\&\quad-[\delta(\mathbf{k}''-k_1)\delta(\omega'-\omega_2)\\&\quad+\delta(\mathbf{k}'-\mathbf{k}_1)\delta(\omega''-\omega_2)]f^2(\omega_2)/2H(\omega_2),\end{aligned}$$

and

$$M_2(\omega_1,\omega_2;\omega')=[f(\omega_1)f(\omega_2)f(\omega f')(\omega'')/2H_1H_2]\times[\chi(\omega';\omega_1,\omega_2)+\chi(\omega'';\omega_1,\omega_2)],$$

where $\omega''=\omega_3-\omega'=\omega_1+\omega_2-\omega'$ is understood. Thus

$$\begin{aligned}M_2(\omega_1,\omega_2;\omega)&=\frac{f(\omega_1)f(\omega_2)f(\omega)f(\omega_3-\omega)}{2H(\omega_1)H(\omega_2)}\\&\quad\times\left[\frac{1}{\omega_3-\omega+i\epsilon}\left(\frac{\omega_1}{\omega-\omega_2-i\epsilon}+\frac{\omega_2}{\omega-\omega_1-i\epsilon}+\frac{\omega_3}{\omega}\frac{2H_3\tilde{H}_3}{1-H_3\tilde{H}_3}\right)\right.\\&\quad+\frac{1}{\omega}\left(\frac{\omega_1}{\omega_1-\omega-i\epsilon}+\frac{\omega_2}{\omega_2-\omega-i\epsilon}\right.\\&\quad\left.\left.+\frac{\omega_3}{\omega_3-\omega}\frac{2H_3\tilde{H}_3}{1-H_3\tilde{H}_3}\right)-\frac{2H_3}{1-H_3\tilde{H}_3}\right.\\&\quad\left.\times\left(\frac{1}{\omega}I_3(\omega_3-\omega)+\frac{1}{\omega_3-\omega}I_3(\omega)\right)\right].\end{aligned} \tag{65}$$

The first two lines can be simply combined. The two expressions in the last line combine by use of

[18] The equations in the higher sectors of the Lee model again have the form of (I) and (II). The author believes that they too can be solved along these lines. However the inhomogeneous terms become more and more complicated such that it requires a great amount of work to do all the integrals.

(34). Doing all the cancellations we find[19]

$$M_2(\omega_1, \omega_2; \omega) = i\pi[\delta(\omega - \omega_1) + \delta(\omega - \omega_2)] \frac{f^2(\omega_1)f^2(\omega_2)}{H(\omega_1)H(\omega_2)} + \frac{H(\omega_1 + \omega_2)}{1 - H(\omega_1 + \omega_2)A(\omega_1 + \omega_2)} \times \frac{f(\omega_1)f(\omega_2)f(\omega)f(\omega_1 + \omega_2 - \omega)}{H(\omega_1)H(\omega_2)H(\omega)H(\omega_1 + \omega_2 - \omega)}. \quad (66)$$

The production amplitude P, defined by

$$S_{\text{prod}} = \langle (N\Theta_1\Theta_2)^{\text{out}} \mid (V\Theta_0)^{\text{in}} \rangle = 2\pi i\delta(\omega_1 + \omega_2 - \omega_0)P(\omega_1, \omega_0) \quad (67)$$

is given by the term containing the $\delta(\omega_1 + \omega_2 - \omega_0)$ in $\psi_2(\omega_1, \omega_2, \omega_0)$ of (55). Substituting $\omega_2 = \omega_0 - \omega_1$ we find

$$P(\omega_1, \omega_0) = (f(\omega_1)f(\omega_0)f(\omega_0 - \omega_1)/\sqrt{2}\,\omega_1(\omega_0 - \omega_1 + i\epsilon)) \times [2C_1(\omega_0) - C_2(\omega_0)(\omega_0 - \omega_1) \times I_0(\omega_0 - \omega_1) - C_2(\omega_0)\omega_1 I_0(\omega_1)]. \quad (68)$$

Again using (34) for the last term, the second term in the bracket cancels, and this simplifies enormously to read

$$P(\omega_1, \omega_0) = \frac{f(\omega_1)f(\omega_0)f(\omega_0 - \omega_1)}{H(\omega_1)H(\omega_0 - \omega_1)} \cdot \frac{1}{1 - H(\omega_0)A(\omega_0)}. \quad (69)$$

which again agrees with Amado's work.

For completeness we may mention that the $N\Theta$-elastic amplitude, defined by

$$S_{N\Theta} = \delta(\mathbf{k} - \mathbf{k}') + 2\pi i\delta(\omega - \omega')M_1(\omega)$$

in our notation becomes,

$$M_1(\omega) = -f^2(\omega)H^{-1}(\omega). \quad (70)$$

Since we have only s-wave scattering:

$$\delta(\mathbf{k} - \mathbf{k}') = \delta(\omega - \omega')/4\pi k\omega$$

and we define

$$S_{N\Theta} = (4\pi k\omega)^{-1}\delta(\omega' - \omega)S_0(\omega) \quad (71)$$

where the reduced S-matrix S_0 can now be written

$$S_0(\omega) = \exp[2i\delta(\omega)] = H^*(\omega)/H(\omega) \quad (72)$$

in terms of its phase shift or our function $H(\omega)$.

It is remarkable that all three amplitudes of this sector [(6), (66), (69)] contain the $N\Theta$-amplitude $M(\omega)$ in a simple way. All three amplitudes further contain the denominator $1 - H_0A_0$. It is therefore the study of this function $(1 - H_0A_0)$ that will bring an answer to the question of possible existence of a Peierls pole[20] in the Lee model; a question we think has not yet been answered[19] convincingly.

For $\omega_0 < \mu$ $(\omega_3 < 2\mu)$ a solution of the homogeneous equation (38) would clearly correspond to a stationary eigenstate of $\mathfrak{H}$, which would either be a bound state or a ghost. The denominator $I - H_0A_0$ can vanish between $0 < \omega < \mu$, where it is real, if g^2 is large. This shows that in the Lee model a $V\Theta$-bound state can occur. However such a state, having the $V\Theta$-quantum numbers, will not change the expressions for the amplitudes in this sector, since due to energy conservation (stable V) it would not be accessible from either the $V\Theta$ or $N2\Theta$ channel.

ACKNOWLEDGMENTS

I wish to thank Professor O. W. Greenberg for his continuous interest in this work. I would like especially to thank him and Professor G. Källén for many helpful discussions.

Note added in proof: Using a more deductive method, it has been possible to solve the homogeneous equation and to show uniqueness of the above solution.

APPENDIX I

To prove the relations (32) we write

$$J \equiv \frac{-1}{\pi}\int_\mu^\infty \text{Im}\, G(\omega')\left[\frac{1}{\omega' - \omega_0 - i\epsilon} - \frac{1}{\omega' - \omega_0 + \omega - i\epsilon}\right](\omega_0 - \omega')I_0(\omega_0 - \omega')\,d\omega',$$

make a common denominator and observe that the factor $(\omega_0 - \omega')$ is canceled. So

$$J = \frac{\omega}{\pi}\int_\mu^\infty \frac{\text{Im}\, G(\omega')\,d\omega'}{\omega' - \omega_0 + \omega - i\epsilon} I_0(\omega_0 - \omega').$$

To get further we insert the definition (29) of I_0.

$$J = \frac{-\omega}{\pi}\int_\mu^\infty \frac{\text{Im}\, G(\omega')\,d\omega'}{\omega' - \omega_0 + \omega - i\epsilon} \times \frac{1}{\pi}\int_\mu^\infty \left(\text{Im}\frac{1}{H(\omega'')}\right) \times \frac{d\omega''}{G(\omega_0 - \omega'')(\omega'' - \omega_0 + \omega' - i\epsilon)}.$$

Changing orders of integration,

$$J = -\frac{\omega}{\pi}\int_\mu^\infty \left(\text{Im}\frac{1}{H(\omega'')}\right)\frac{d\omega''}{G(\omega_0 - \omega'')} \times \frac{1}{\pi}\int_\mu^\infty \frac{\text{Im}\, G(\omega')\,d\omega'}{(\omega' - \omega_0 + \omega - i\epsilon)(\omega' - \omega_0 + \omega'' - i\epsilon)},$$

[19] The $N2\Theta$ amplitude has first been derived by P. K. Srirastava, Phys. Rev. 131, 461 (1963). We do not agree with the results in Appendix B of that paper [Eqs. (B10), (B14)].

[20] R. F. Peierls, Phys. Rev. Letters 6, 641 (1961).

the last integral can be done to give

$$\frac{1}{\omega'' - \omega - i\epsilon}\,[G(\omega_0 - \omega) - G(\omega_0 - \omega'')]$$

and then

$$J = -\frac{\omega}{\pi}\,G(\omega_0 - \omega) \times \int_\mu^\infty \left(\operatorname{Im}\frac{1}{H(\omega')}\right)\frac{d\omega'}{G(\omega_0 - \omega')(\omega' - \omega - i\epsilon)} + \frac{\omega}{\pi}\int_\mu^\infty \left(\operatorname{Im}\frac{1}{H(\omega')}\right)\frac{d\omega'}{\omega' - \omega - i\epsilon}.$$

Doing the last integral and using (29) again

$$J = \omega[G(\omega_0 - \omega)I_0(\omega) + [H(\omega)]^{-1} - \omega^{-1}] \qquad \text{QED.}$$

The expression (34):

$$\mathcal{J}_1 \equiv \omega\{I_0(\omega) + [H(\omega)G(\omega_0 - \omega)]^{-1}\}$$

apparently has right- and left-hand cuts combined. The right-hand cut however does cancel out. To show this we use (29) and (14) to write

$$J_1 = G^{-1}(\omega_0 - \omega) + \frac{\omega}{\pi}\int_\mu^\infty \left(\operatorname{Im}\frac{1}{H(\omega')}\right)\frac{d\omega'}{\omega' - \omega - i\epsilon} \times \left[\frac{1}{G(\omega_0 - \omega)} - \frac{1}{G(\omega_0 - \omega')}\right].$$

Now

$$\frac{1}{G(\omega_0 - \omega)} - \frac{1}{G(\omega_0 - \omega')} = \frac{1}{\pi}\int_\mu^\infty \left(\operatorname{Im}\frac{1}{H(\omega'')}\right) \times \left[\frac{\omega_0 - \omega}{\omega'' - \omega_0 + \omega - i\epsilon} - \frac{\omega_0 - \omega'}{\omega'' - \omega_0 + \omega' - i\epsilon}\right] d\omega''$$

and, making one common denominator in the brackets,

$$= \frac{1}{\pi}\int_\mu^\infty \left(\operatorname{Im}\frac{1}{H(\omega'')}\right) \times \frac{\omega''(\omega' - \omega)\,d\omega''}{(\omega'' - \omega_0 + \omega - i\epsilon)(\omega'' - \omega_0 + \omega' - i\epsilon)}.$$

If we insert this into the expression for J_1 we find for the integral alone

$$\frac{\omega}{\pi}\int_\mu^\infty \left(\operatorname{Im}\frac{1}{H(\omega'')}\right)\frac{\omega''\,d\omega''}{\omega'' - \omega_0 + \omega - i\epsilon} \times \frac{1}{\pi}\int_\mu^\infty \left(\operatorname{Im}\frac{1}{H(\omega')}\right)\frac{d\omega'}{\omega' - \omega_0 + \omega'' - i\epsilon}.$$

Doing the last integral this becomes

$$\frac{\omega}{\pi}\int_\mu^\infty \left(\operatorname{Im}\frac{1}{H(\omega')}\right)\frac{\omega'\,d\omega'}{\omega' - \omega_0 + \omega - i\epsilon} \times \left[\frac{1}{H(\omega_0 - \omega')} - \frac{1}{\omega_0 - \omega' + i\epsilon}\right]$$

and by (15) J_1 becomes

$$J_1 = 1 + \frac{\omega}{\pi}\int_\mu^\infty \left(\operatorname{Im}\frac{1}{H(\omega')}\right)\frac{d\omega'}{\omega' - \omega_0 + \omega - i\epsilon} \times \left[\frac{\omega_0 - \omega}{\omega} + \frac{\omega'}{H(\omega_0 - \omega')} - \frac{\omega'}{\omega_0 - \omega' + i\epsilon}\right].$$

This expression obviously only has a left-hand cut in the ω-plane. To get it into the form (34) we combine the first and the last term in the square brackets and use (1) in the middle term to write it

$$\frac{\omega_0(\omega_0 - \omega' - \omega)}{\omega(\omega_0 - \omega' + i\epsilon)} + \frac{\omega'}{(\omega_0 - \omega' + i\epsilon)G(\omega_0 - \omega')}.$$

Then ω times this can be written

$$\omega_0\frac{\omega' - \omega_0 + \omega}{\omega' - \omega_0 - i\epsilon} - \frac{\omega\omega'}{(\omega' - \omega_0 - i\epsilon)G(\omega_0 - \omega')}$$

this leaves for J_1

$$J_1 = 1 + \frac{\omega}{\pi}\int_\mu^\infty \left(\operatorname{Im}\frac{1}{H(\omega')}\right)\frac{d\omega'}{\omega' - \omega_0 - i\epsilon} - \frac{\omega}{\pi}\int_\mu^\infty \left(\operatorname{Im}\frac{1}{H(\omega')}\right) \times \frac{\omega'\,d\omega'}{G(\omega_0 - \omega')(\omega' - \omega_0 - i\epsilon)(\omega' - \omega_0 + \omega - i\epsilon)}.$$

The first integral is $G_0^{-1} - 1$. The last integrand we decompose

$$\frac{\omega\omega'}{(\omega' - \omega_0 - i\epsilon)(\omega' - \omega_0 + \omega - i\epsilon)} = \omega'\left[\frac{1}{\omega' - \omega_0 - i\epsilon} - \frac{1}{\omega' - \omega_0 + \omega - i\epsilon}\right],$$

then

$$J_1 = G_0^{-1} + \frac{-1}{\pi}\int_\mu^\infty \left(\operatorname{Im}\frac{1}{G(\omega')}\right)\frac{d\omega'}{G(\omega_0 - \omega')(\omega' - \omega_0 - i\epsilon)} + \frac{1}{\pi}\int_\mu^\infty \left(\operatorname{Im}\frac{1}{G(\omega')}\right)\frac{d\omega'}{G(\omega_0 - \omega')(\omega' - \omega_0 + \omega - i\epsilon)}.$$

The first integral (including the minus) is

$$\omega_0 A_0 - \int_\mu^\infty \left(\operatorname{Im}\frac{1}{H(\omega')}\right)\frac{d\omega'}{G(\omega_0 - \omega')}.$$

The second integral is

$$-(\omega_0 - \omega)I_0(\omega_0 - \omega) + \int_\mu^\infty \left(\operatorname{Im}\frac{1}{H(\omega')}\right)\frac{d\omega'}{G(\omega_0 - \omega')}.$$

These two relations are easily verified by doing partial fractions in (7) and (29). Adding up we find

$$J_1 = G_0^{-1} + \omega_0 A_0 - (\omega_0 - \omega)I_0(\omega_0 - \omega)$$

or

$$J_1 = G_0^{-1}[1 + H_0 A_0 - (\omega_0 - \omega) G_0 I_0(\omega_0 - \omega)],$$

which is the form used in (34).

APPENDIX II

Here we give a simple derivation of the $T_{V\Theta}$-matrix from the state vector. This derivation is interesting because the $V\Theta$-sector is the first scattering sector in which renormalization enters in a nontrivial way. It further provides an interesting integral relation between the T-matrix and its off-shell extension.

We start from the eigenvalue equation (47), insert (46), use the relations (50) and find formally

$$|\chi\rangle^+ = \frac{-f(\omega) Z^{-\frac{1}{2}}}{m + \omega - \mathfrak{H} + i\epsilon} \int \frac{f(\omega')}{\omega'} |V\Theta_{\mathbf{k}'}\rangle \, d^3k'. \quad \text{(II.1)}$$

Since our theory is invariant under time reversal, the out state is just the Hermitan conjugate of the in state. Therefore from (II.1) one concludes that

$$|(V\Theta_{\mathbf{k}})^{\text{in}}\rangle = |(V\Theta_{\mathbf{k}})^{\text{out}}\rangle + 2\pi i \delta(m + \omega - \mathfrak{H}) \times \frac{f(\omega)}{\sqrt{Z}} \int \frac{f(\omega')}{\omega'} |V\Theta_{\mathbf{k}'}\rangle \, d^3k'$$

and from (56a)

$$\begin{aligned} S_{V\Theta} &= \langle (V\Theta_1)^{\text{out}} \mid (V\Theta)^{\text{out}}\rangle + 2\pi i \frac{f(\omega)}{\sqrt{Z}} \\ &\quad \times \int \frac{f(\omega')}{\omega'} \langle (V\Theta_1)^{\text{out}} | \, \delta(m + \omega - \mathfrak{H}) \, |V\Theta_{\mathbf{k}'}\rangle \, d^3k' \\ &= \delta(\mathbf{k} - \mathbf{k}_1) + 2\pi i \delta(\omega - \omega_1) \frac{f(\omega)}{\sqrt{Z}} \\ &\quad \times \int \frac{f(\omega')}{\omega'} \langle (V\Theta_1)^{\text{out}} \mid V\Theta_{\mathbf{k}'}\rangle \, d^3k'. \end{aligned}$$

The last matrix element can be evaluated from our $V\Theta$-state vector:

$$\langle (V\Theta_1)^{\text{out}} \mid V\Theta_{\mathbf{k}'}\rangle = Z^{\frac{1}{2}} \delta(\mathbf{k}_1 - \mathbf{k}') + Z^{\frac{1}{2}} \psi_1(\mathbf{k}', \mathbf{k}).$$

Therefore, by (56b)

$$T(\omega) = \frac{f^2(\omega)}{\omega} + f(\omega) \int \frac{f(\omega')}{\omega'} \psi_1(\mathbf{k}'; \mathbf{k}) \, d^3k'. \quad \text{(II.2)}$$

The integral can be done. Using (I') and (3) it is

$$f(\omega)^2 \Big[C_1(\omega) \frac{1}{\pi} \int_\mu^\infty \frac{\operatorname{Im} G(\omega') \, d\omega'}{\omega'(\omega' - \omega - i\epsilon)} + C_2(\omega) \frac{1}{\pi} \int_\mu^\infty \frac{\operatorname{Im} G(\omega')}{\omega'} I(\omega - \omega') \, d\omega' \Big].$$

The first integral is $\omega^{-1}[G(\omega) - 1]$. In the second we insert the definition (29) and change orders of integration, leading to

$$-\frac{1}{\pi} \int_\mu^\infty \left(\operatorname{Im} \frac{1}{H(\omega'')} \right) \frac{d\omega''}{G(\omega - \omega'')} \times \frac{1}{\pi} \int_\mu^\infty \frac{\operatorname{Im} G(\omega') \, d\omega'}{\omega'(\omega' - \omega + \omega'' - i\epsilon)}.$$

Doing the last integral again

$$\begin{aligned} &-\frac{1}{\pi} \int \left(\operatorname{Im} \frac{1}{H(\omega')} \right) \frac{d\omega'}{\omega - \omega' + i\epsilon} \\ &\qquad + \frac{1}{\pi} \int \left(\operatorname{Im} \frac{1}{H(\omega')} \right) \frac{d\omega'}{H(\omega - \omega')} \\ &\qquad = [H^{-1}(\omega) - \omega^{-1}] + A(\omega) \end{aligned}$$

where we have just inserted the definitions.

Collecting, we find for $T(\omega)$

$$T(\omega) = f^2(\omega) \Big[\frac{1}{\omega} + \frac{1}{G} \frac{1 + H\bar{H}}{1 - H\bar{H}} \frac{1}{\omega} (G - 1) + \frac{2}{1 - H\bar{H}} \left(\frac{1}{H} - \frac{1}{\omega} + \bar{H} \right) \Big].$$

This can be simplified to

$$T(\omega) = \frac{f^2(\omega)}{H(\omega)} \frac{1 + H\bar{H}}{1 - H\bar{H}}$$

or

$$T(\omega) = [f^2(\omega)/\omega] C_1(\omega)$$

which is nothing other than the residue of $\psi_1(\text{I}')$ at $\omega_0 = \omega$.

Similar derivations can be given for $T_{N2\Theta}$ and $P_{V\Theta, N\Theta\Theta}$.

Integrals of the Second-Order Linear Differential Equation*

J. KANE
California Institute of Technology, Pasadena, California
AND
E. R. SURYNARAYAN
University of Rhode Island, Kingston, Rhode Island

(Received 5 October 1964)

In this study we describe procedures for the numerical solution of the second-order linear differential equation which have either continuous or discontinuous coefficients. Our motivation is a well-known technique in the theory of inhomogeneous transmission lines: the treatment of a continuously varying line by considering it to be comprised of various sections of *uniform* lines. Although this procedure is very suggestive physically, there are difficulties with its applications to second-order equations which do not describe wave propagation. First, the language of the circuit engineer is such that a pair of first-order equations describing some analogs of the complex quantities voltage and current seem to be required. Whereas these equations appear naturally in transmission line theory, we show that it is an unnecessary burden to find their counterparts when the problem is but to solve a second-order equation. The second objection to this approach is that it is not apparent that a piecewise constant partition of the coefficients in a differential equation will yield the rigorous solution if the subdivision is carried out to an arbitrary degree. Indeed, we show that the limit of the quantization scheme can *not* yield the rigorous solution. On the other hand, we are led to a well-defined technique which generates the solution by a method suggested by the procedures used in the discrete case. This enables the second-order equation to be solved rigorously by iteration with no complications, and in a form ideally suited for computer programs.

I. INTRODUCTION

A. Basic Equations

THE class of problems, mathematical and physical, which require an explicit knowledge of the solutions of the linear second-order differential equation

$$d^2U(t)/dt^2 + A(t)dU(t)/dt + B(t)U(t) = 0 \qquad (1)$$

is legion. Only rarely can exact solutions of this equation be found in closed form. While local approximations are easily constructed, the task of finding global representations offers a major problem. This difficulty is particularly vexing inasmuch as the first-order equation,

$$dV/dt - P(t)V(t) = 0 \qquad (2)$$

possesses an immediate integral

$$V(t) = \exp\left\{\int^t P(\tau)\, d\tau\right\}, \qquad (3)$$

but no similar representation for Eq. (1) has been found. The solution (3) has the interpretation of defining the present value of $V(t)$ in terms of the history of its logarithmic derivative $P(t) = (d/dt)(\ln V)$. If we specify an initial condition on the possible pair of solutions of (1) at say $t = t_0$, then Eq. (1) describes a unique function with that property. From this initial condition and a knowledge of Eq. (1), the behavior of the second-order $U(t)$ should likewise be discernible in terms of the prior trajectory of its logarithmic derivative: the topic of the present analysis.

Following a change of unknown

$$u(t) = U(t) \exp\left\{\frac{1}{2}\int^t A(\tau)\, d\tau\right\} \qquad (4)$$

and an introduction of

$$k^2(t) = B(t) - \tfrac{1}{2}dA(t)/dt - \tfrac{1}{4}A^2(t), \qquad (5)$$

Eq. (1) can be recast into the canonical form

$$d^2u/dx^2 + k^2(x)u(x) = 0 \qquad (6)$$

in which we replace t by x to eliminate confusion between the forms (1) and (6). At some point $x = x_0$, we may suppose that we have a given constraint

$$\frac{d}{dx}\ln u = \frac{du}{dx}/u = g_0 \qquad (x = x_0), \qquad (7)$$

which singles out one solution; another linearly independent solution can be specified by choosing a different value for the parameter g_0.

The form of the solution (3) to the first-order equation (2), suggests that it might be more profitable to concentrate our attention on

$$\int_{x_0}^{x} \Gamma(\xi)\, d\xi = \ln u(x), \quad u(x) = \exp\int_{x_0}^{x} \Gamma(\xi)\, d\xi, \qquad (8)$$

* Contribution No. 1344 from the Division of Geological Sciences, California Institute of Technology, Pasadena, California.

that is, the logarithmic derivative

$$\Gamma(x) = \frac{du/dx}{u}, \tag{9}$$

rather than on $u(x)$. Indeed, from the point of view of a mathematical physicist interested in eigenvalue problems, an explicit knowledge of the logarithmic derivative provides more insight than would formulas for two distinct solutions of (1), therefore we shall pay special attention to the logarithmic derivative. The continuity of Γ is assured if we restrict our attention to continuous solutions $u(x)$ possessing a continuous first derivative $u'(x)$. If these functions are bounded, then we note that the poles and zeros of $\Gamma(x)$ correspond to the zeros and maxima respectively of $u(x)$. In terms of $\Gamma(x)$, the equations (6) and (7) become

$$d\Gamma/dx = -(k^2 + \Gamma^2), \qquad x_0 \leq x \leq x_m \tag{10}$$

$$\Gamma = g_0, \qquad x = x_0 \tag{11}$$

wherein we have adopted the convention that the initial condition is given at $x = x_0$, and we wish to propagate the solution to the right of x_0 up to some arbitrary point x_m. Formula (10) describes $\Gamma(x)$ in terms of a Ricatti equation, and even though this expression is nonlinear, it can easily be solved by iterative techniques if $k^2(x)$ is negative. On the other hand, if $k^2(x)$ is positive, then $\Gamma(x)$ behaves like an ill-tempered function with numerous singularities. Even high-speed computers have difficulty in coping with such a wildly fluctuating function, and numerical work under such circumstances offers major difficulties. One conclusion of the present work is the need to isolate and describe explicitly the bad manners of $\Gamma(x)$ so that if numerical methods must be used, they should treat slowly-varying functions.

B. Summary of Results

Equation (6) can be thought of as a reduced wave equation: subject to severe limitations on the behavior of $k^2(x)$, approximate solutions can be found by the phase-integral[1] method. This technique yields reasonably simple solutions of, say, exponential type. However, the character of the solution must change in the vicinity of a turning point of $k^2(x)$: in such a region, the solution must be oscillatory rather than exponential which means that the phase integral needs severe correction. Loosely speaking, the major disadvantage of the phase-integral method is that it involves an argument of the form $\int^x k(\xi)\, d\xi$ in trigonometric or exponential functions. As a result, when $k^2(x)$ changes sign, the argument does not change character correctly because of the integral's inertia in storing the past values of $k(x)$. That is, the phase-integral solution *must* involve complex quantities in the presence of turning points even if the original differential equation's coefficients are real. We know that such equations have real solutions, but the phase integral's hysteresis prevents its providing a valid representation of them. For this reason, we avoid using formulas which depend upon integrals of the form $\int^x k(\xi)\, d\xi$.

One means of avoiding turning-point difficulties is to consider $k^2(x)$ as being piecewise constant $k^2(x) = k_i^2$ over sufficiently small intervals $x_{i-1} \leq x \leq x_i$. Within each quantum cell a pair of trial solutions can be inserted: trigonometric or exponential depending upon the sign of k_i^2. By this process, the task becomes an algebraic one of determining the correct coefficients of the trial function in each cell so that it joins smoothly with its neighbors. The difficulties with this procedure are twofold. First, while the procedure appears to converge to a solution of the basis of physical arguments, the mathematical justification for this procedure has neither been simple nor convincing, since the limiting process involves a sequence of step-function approximations. Second, for fine-grained quantization, the large number of algebraic equations to be solved can be overwhelming. In diffraction theory, an approximate means of surmounting this hurdle is the heuristic device of neglecting certain contributions within each cell that can be identified with multiple reflections of a wave at a boundary which has a very small reflection coefficient.[2-4] Indeed, the phase integral is but the leading term in such an expansion.[5] In Sec. II, we consider this discretization process and show how the algebraic difficulties can be sidestepped by introducing the proper variables and format of the trial solution without making any assumption concerning the behavior of the field within each cell. This yields an expression for $\Gamma(x)$ in terms of a product of n 2×2 unimodular matrices for an nth-order quantization. In the limit, as $n \to \infty$, each matrix operator becomes an infinitesimal trans-

[1] We prefer this adjective to any acronym based upon some perturbation of the initials of Liouville, Green, Carlini, Stokes, Horn, Rayleigh, Birkhoff, Jeffreys, Langer, Wentzel, Kramers, Brillouin, *et al.*

[2] L. M. Brekhovskikh, *Waves in Layered Media* (Academic Press Inc., New York, 1960).

[3] K. G. Budden, *The Wave-Guide Mode Theory of Wave Propogation*, (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1961).

[4] J. R. Wait, *Electromagnetic Waves in Stratified Media*, (The Macmillan Company, New York, 1962).

[5] H. Bremmer, Comm. Pure Appl. Math. 4, 105 (1951).

formation; we analyze this differential change in Γ and find a geometric characterization of the trajectory of the logarithmic derivative. We then find that the quantization scheme cannot yield the rigorous solution if the coefficients in the original differential equation are analytic. Nonetheless, the procedures suggest a transformation which enables the rigorous solution to be found by a related technique. This analysis will be found in Sec. III together with an illustrative numerical example.

II. DISCRETE FORMALISM AND SOLUTION

A. Introduction

We assume that g_0 and $k^2(x)$ are real [but not $k(x)$] so that we can confine our attention to real solutions of (6). Of course, once the real solutions of (6) are known, complex solutions can easily be constructed by the superposition principle.

Let $\{\delta_i\}^m$ be any monotone sequence of $(m+1)$ points which divides any portion of the x axis, say, the interval

$$-\infty \leq x_0 \leq x \leq x_m \leq +\infty \tag{12}$$

into subintervals

$$\Delta x_i = \delta_i - \delta_{i-1}, \quad (\delta_0 = x_0, \delta_{i+1} > \delta_i, \delta_m = x_m). \tag{13}$$

If this quantization is so fine that within the ith interval Δx_i we can replace $k^2(x)$ by k_i^2 to any preassigned degree of accuracy, then the equation

$$d^2u_i/dx^2 + k_i^2 u_i = 0, \quad \delta_{i-1} < x < \delta_i \tag{14}$$

has the general solution

$$u_i = A_i[\Gamma_i k_i^{-1} \sin k_i(x - \delta_{i-1}) + \cos k_i(x - \delta_{i-1})], \tag{15}$$

wherein the discrete parameters A_i and Γ_i remain to be specified. From a theoretical point of view, any two-parameter combination of $\sin k_i x$ and $\cos k_i x$ could serve as well as the particular form selected in Eq. (15). However, from a practical viewpoint, major advantages are introduced by the selection (15). First, no assumption has been made concerning the character of u_i in the ith interval: oscillatory or damped. For, should k_i^2 be negative, or $k_i = i\kappa_i$ be imaginary, then

$$u_i = A_i[-\Gamma_i \kappa_i^{-1} \sinh \kappa_i(x - \delta_{i-1}) + \cosh \kappa_i(x - \delta_{i-1})] \tag{16}$$

will continue to be real provided only that A_i and Γ_i are chosen real. Note that at $x = \delta_{i-1}$ we have

$$u_i = A_i, \quad \frac{du_i/dx}{u_i} = \Gamma_i, \quad x = \delta_{i-1}, \tag{17}$$

and so A_i and Γ_i can be identified as the amplitude and logarithmic derivative of the solution at the left endpoint. Most important however, is that the form of (15) is such that both the boundary conditions (17) are *decoupled*, and it becomes possible to solve for A_i and Γ_i separately: a considerable reduction of computational effort.

B. Solution

The equations (6) and (7) are homogeneous, and so we can normalize A_0 to any convenient value, but we must choose $\Gamma_0 = g_0$ if the boundary condition (11) is to be satisfied. Even though k_i is a jump function, this initial data can be continued as a smooth function for all x, provided we choose

$$\Gamma_{i+1} = \frac{\Gamma_i \cos k_i\Delta x_i - k_i \sin k_i\Delta x_i}{\Gamma_i k_i^{-1} \sin k_i\Delta x_i + \cos k_i\Delta x_i} \tag{18}$$

and

$$A_{i+1} = A_i[\Gamma_i k_i^{-1} \sin k_i\Delta x_i + \cos k_i\Delta x_i]. \tag{19}$$

In this manner we obtain a function which satisfies (6) everywhere except that $k^2(x)$ must be replaced by some local average value k_i^2. Furthermore, the iterations required to extend $u_i(x)$ are very easily performed since (18) has the form of a Möbius transformation whose coefficient matrix

$$\mathsf{M}_i = \begin{bmatrix} \cos k_i\Delta x_i & -k_i \sin k_i\Delta x_i \\ k_i^{-1} \sin k_i\Delta x_i & \cos k_i\Delta x_i \end{bmatrix} \tag{20}$$

has its determinant equal to unity. As a result, the coefficient matrix T_m required to transform Γ_0 into some final value Γ_m at the arbitrary point $x = x_m$ is the product of m factors

$$\mathsf{T}_m = \mathsf{M}_m \cdot \mathsf{M}_{m-1} \cdot \ \ldots \ \cdot \mathsf{M}_i \cdot \ \ldots \ \cdot \mathsf{M}_1, \tag{21}$$

and is likewise unimodular.

C. Check

It is now but a simple calculation to show that if $m \to \infty$, and $\Delta x_i \to 0$ for any sequence $\{\delta_i\}^m$, then the process yields a function which converges to the solution of an equation consistent with assumed behavior of $k^2(x)$. Since $\Gamma(x)$ defines $u(x)$ uniquely apart from a multiplicative constant, it will be sufficient to show that the continuous limit of the Γ_i's defined (18) satisfies the Ricatti equation (10) on any interval Δx_i. We calculate that

$$\frac{\Gamma_{i+1} - \Gamma_i}{\Delta x_i} = \frac{-k_i \sin k_i\Delta x_i - (\Gamma_i^2 k_i^{-1}) \sin k_i\Delta x_i}{\Delta x_i[(\Gamma_i k_i^{-1}) \sin k_i\Delta x_i + \cos k_i\Delta x_i]}, \tag{22}$$

and if we pass to the limit $\Delta x_i \to 0$ we obtain

$$\lim_{\Delta x_i \to 0} \frac{\Gamma_{i+1} - \Gamma_i}{\Delta x_i} = \frac{d\Gamma}{dx} = -(k^2 + \Gamma^2), \qquad (23)$$

which is precisely Eq. (10). As a double check, we can repeat the limiting process for Eq. (19),

$$\frac{A_{i+1} - A_i}{\Delta x_i} = \frac{A_i[(\Gamma_i k_i^{-1}) \sin k_i \Delta x_i + \cos k_i \Delta x_i - 1]}{\Delta x_i} \qquad (24)$$

and derive the defining equation for $\Gamma(x)$ in terms of $u(x)$

$$\lim_{\Delta x_i \to 0} \frac{A_{i+1} - A_i}{\Delta x_i} = \frac{du}{dx} = u\Gamma(x), \qquad (25)$$

if we note that in the limit, the amplitude A_i becomes $u(x_i)$.

Since we have always assumed that $dk/dx = 0$ locally, what we have just shown is that *to the extent that the derivative k' can be neglected on any quantum cell the scheme converges to a solution of* (6).

D. The Infinitesimal Transformation

The preceding analysis furnishes a simple illustration of the algebraic structure of a particular Ricatti equation: one example of the general theory of such equations[6] which shows that its solution can be characterized by matrix operators. In the limit, each matrix $\mathbf{M}_i$ in the product (21) approaches an infinitesimal transformation, and the total transformation $\mathbf{T}_m$ giving $\Gamma(x_m)$ in terms of $\Gamma(x_0)$ can be expressed either as a matrizant[7] or a product integration[8]: the finite result of compounding a sequence of infinitesimal transformations. We can think of this operation as defining a curve traced out by the repeated reaction of the original point $\Gamma(x_0)$ to each infinitesimal product $\mathbf{M}_i$. If we can find some analytic means of specifying this trajectory, then we might be spared the need of calculating matrizants, or product integrals to describe the limit of the iterative process. For this purpose, we consider the geometry of the transformation implied by any matrix $\mathbf{M}_i$.

No doubt some readers have been struck by the resemblance of $\mathbf{M}_i$ to the transmission-line matrices of network theory. However, we caution that this resemblance is deceiving: $\mathbf{M}_i$ is a *real* transformation unlike the transmission-line matrix which is a *complex* one. In the context of transmission-line theory, voltage and current, or **E** and **H** are out of phase, and the complex entries in the transmission-line matrix emphasize this feature. In our analysis however, we have avoided introducing any analogs of the first-order transmission-line equations since we feel that their use is artificial and contrived for applications other than those described by Maxwell's equations. Furthermore, unlike the $\mathbf{M}_i$'s, their geometric interpretation requires a complicated hyperbolic model rather than the more familiar Euclidian plane.

We emphasize that Eq. (20) represents a *real* transformation: in more general terms, it represents a mapping of the complex Γ plane upon itself such that the real axis is left invariant. Since Möbius transformations describe elementary geometric deformations of the complex plane, this provides us with a clue to the proper characterization of $\mathbf{M}_i$. By direct multiplication, it is easily checked that each $\mathbf{M}_i$ can be re-expressed as the product of three terms

$$\mathbf{M}_i = \mathbf{C}_i^{-1}\mathbf{R}_i\mathbf{C}_i, \qquad (26a)$$

where

$$\mathbf{C}_i = \begin{bmatrix} k_i^{-\frac{1}{2}} & 0 \\ 0 & k_i^{+\frac{1}{2}} \end{bmatrix} \qquad (26b)$$

$$\mathbf{R}_i = \begin{bmatrix} \cos k_i \Delta x_i & -\sin k_i \Delta x_i \\ \sin k_i \Delta x_i & \cos k_i \Delta x_i \end{bmatrix}. \qquad (26c)$$

Each factor has a simple geometric interpretation. The first, $\mathbf{C}_i$ represents a uniform contraction of the Γ'-plane by a factor $1/k_i$; on the other hand, its inverse $\mathbf{C}_i^{-1}$ represents a uniform dilation by the inverse factor k_i. Between these two operations, we need perform $\mathbf{R}_i$ which represents a rotation of the stereographic projection of the Γ'-plane on the Riemann sphere.

The sequence of operations is illustrated in Fig. 1 which only shows the plane intersecting the real Γ-axis and the north pole N of the Riemann sphere. First, we lay off the initial point Γ_i on the Γ-axis

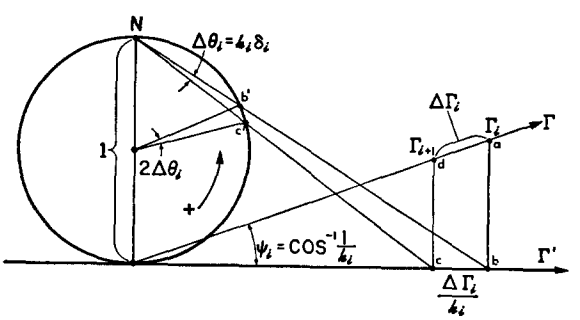


FIG. 1. The geometry of the matrix transformation $\mathbf{M}_i$.

[6] R. Redheffer, J. Ratl. Mech. Anal. 5, 835 (1956).
[7] R. A. Frazer, W. J. Duncan, and A. R. Collar, *Elementary Matrices* (Cambridge University Press, New York, 1959).
[8] G. Birkhoff, J. Math & Physics, 16, 104 (1937).

which has been drawn obliquely for convenience. The Γ'-axis is drawn horizontally and intersects the Γ-axis at an angle $\psi_i = \cos^{-1}(1/k_i)$. If we project the Γ-axis onto the Γ'-axis, then all lengths on this new axis will be uniformly contracted. We continue by finding the sterographic projection of this image point, and then we rotate the Riemann sphere by a *negative* (clockwise) angle $-\theta_i = \tan^{-1}(k_i\Delta x_i)$ if $k_i\Delta x_i$ is positive. We retrace our steps to the original Γ-axis, and it is a simple matter to show that the aforementioned sequence is mimicked by the matrix product $\mathsf{M}_i = \mathsf{C}_i^{-1}\mathsf{R}_i\mathsf{C}_i$. The proof of this assertion is left to the reader, and is easily reproduced once it is recalled that the angle subtended by an arc on a circle from the center is twice the angle subtended by that arc from the north pole.

However, the geometric construction also shows that any transformation $\Delta\Gamma_i$, can be obtained by differentiating the function $\Gamma_i = k_i \tan(-k_i x)$, that is

$$\Delta\Gamma_i = (d/dx)[k_i \tan(-k_i x)]_{x=x_i}\Delta x_i. \tag{27}$$

If we recognize that this transformation must evolve continuously from some initial value $\Gamma_i = g_i$ at $x = x_i$, we thus find the general integral

$$\Gamma(x) = k_i \tan[k_i(\gamma_i - x)] \tag{28}$$

representing the locus of $\Gamma(x)$ on the ith interval where the constant γ_i is fixed by the initial condition at $x = x_i$.

$$\gamma_i = x_i - k_i^{-1}\tan^{-1}(g_i/k_i). \tag{29}$$

In the limit, as $\Delta x_i \to 0$ it would seem that k_i could be replaced by $k(x)$ and g_i by its initial value g_0 since $g_i - g_{i-1} \to 0$. In other words, we expect that the net transformation could be described as

$$\Gamma(x) = k(x)\tan[(\gamma - x)k(x)]$$
$$\gamma = x_0 + [1/k(x_0)]\tan^{-1}[g_0/k(x_0)], \tag{30}$$

where g_0 is the initial value of Γ at $x = x_0$. However, the promise of this anticipation is broken for if we substitute (30) into the Ricatti equation we find.

$$\frac{d\Gamma}{dx} = -(k^2 + \Gamma^2) + (dk/dx)\{\tan[(\gamma - x)k] + k(\gamma - x)\sec^2[(\gamma - x)k]\}. \tag{31}$$

In other words, (30) satisfies the Ricatti equation only to the extent that the first derivative of $k(x)$ can be neglected. Where is the fallacy in the preceding argument that can explain our disappointment, and perhaps lead us to a correction?

The resolution of this conundrum depends upon a subtle point in the character of the analysis. The basic assumption was that a continuous $k^2(x)$ could be approximated arbitrarily well by a staircase approximation. Whereas any such approximation is discontinuous, it does approach a continuous limit as the treads and risers get finer and finer. Of the eventual continuity, there is no doubt, but it is a peculiar type of continuity. Think for example of a staircase approximation to $k^2(x) = x$ on the interval $0 \leq x \leq 1$. Each riser represents the jump from one tread to the next, and the height of any one riser goes to zero as the approximation gets better. However, no matter how fine an approximation we care to make, *the sum of all the risers, or the total jump will always be constant.* In precise terminology, this is the difference between continuity and *absolute continuity*. The fallacy is now evident: while $\gamma_i - \gamma_{i-1} \to 0$, we can not conclude that γ_0 will be unchanged as we pass through an unbounded number of quantum cells. That is, as we proceed from the ith cell to the $(i + 1)$th one, we can not neglect the differential jump $\gamma_{i+1} - \gamma_i$ at the end of the ith interval. In other words, we can think of Eq. (30) as a representation of the continuous or *principal part* of the transformation (21) which overlooks the denumerable collection of minute jumps.

We can now see that the quantization scheme can not yield the rigorous solution even if carried out to the continuous limit, for the result of a staircase approximation to $k^2(x)$ can not be absolutely continuous. From the differential equation (6) we see that the ratio u''/u must be equal to $-k^2(x)$, hence in the limit, u''/u can not be absolutely continuous no matter how smooth the original $k^2(x)$ might have been. In general, we know that whenever $k^2(x)$ is analytic, the differential equation (6) must have an analytic solution. As a result, for such equations the fact that the limiting u''/u is not absolutely continuous contradicts the required analyticity of the solution. As a result, we conclude that the limit of the quantization process in general will not converge to the rigorous solution. That the limit of the iterative process might be a weak solution in terms of some norm remains an open question. On physical grounds, there can be no doubt that the discretization approach is a useful procedure when $k^2(x)$ can be considered to be well approximated by a piecewise constant function. Perhaps if we retain information concerning the derivatives of $k^2(x)$, we can construct a rigorous solution? This question is answered affirmatively in the next section.

III. THE CONTINUOUS TRANSFORMATION

In the preceding paragraphs, we have discussed the hazards that obstruct an extension of the finite quantization theory to the continuous case. In such a situation, it is always well to pause, and reflect upon the possiblity that the infinitesimal case might better be handled by other methods. Basically, we are considering the second-order equation as an initial-value problem. That is, we wish some means of knowing the direction the solution or its logarithmic derivative will take once an initial value is specified. This question is of course immediately answered by the Ricatti equation (10) for it describes the derivative explicity in terms of the function. However, we know that when $k^2(x)$ is positive, $\Gamma(x)$ behaves like the tangent function and has numerous poles and zeros. Any numerical method which attempts to pursue the path of such a spirited function will almost of necessity be doomed from the outset. On the other hand, the poles and zeros of the logarithmic derivative have no more significance than locating the maxima and zeros of the desired solution of the differential equation and should not present any intrinsic hurdle.

In the previous section, we have introduced the notion of the *principal part* of the limit of the quantization process Eq. (30). In the initial stages of our investigation we had had the idea that by considering the constant γ to be in fact a slowly varying function of x, we could find the actual solution by solving for $\gamma(x)$. This turned out to be a fruitful pursuit and in this fashion we were able to find excellent numerical replicas of the solutions to the differential equation without any trouble except in the neighborhood of a turning point. This difficulty had to do with choosing a trial function for the logarithmic derivative of the particular form (30); the problems vanish by choosing a simpler starting point.

The essential ingredient is to find a slowly varying function which can characterize the solution. One possibility that we have just discussed is to use a component of the argument of the tangent function, that is, some element describing the *phase*. In quantum mechanics, the scattering of a particle by a spherically symmetric potential can be described in terms of phase shifts. By introducing a nonlinear first-order ordinary differential equation, it is possible to find the relevant information without solving the second-order Schrödinger equation. This approach, introduced by Morse and Allis,[9] and more recently discussed by Calogero[10] and Levy and Keller,[11] is a powerful technique for finding the phase shift. However, these authors have been content with specialized information and have overlooked the possibilities inherent in this approach which can lead to solutions of the second-order differential equation.

Consider the function $\Phi(x)$ defined by

$$\Gamma(x) = -k_0 \tan [k_0\Phi(x)], \tag{32}$$

where k_0 is some convenient, but arbitrary reference value, for example, the mean value of $k^2(x)$ over some interval. Note that even if $\Gamma(x)$ behaves wildly, the corresponding $\Phi(x)$ will reflect a much calmer character. Since $\Gamma(x)$ is specified once $\Phi(x)$ is known, we can ask what equation must $\Phi(x)$ obey if $\Gamma(x)$ is to be a solution of the Ricatti equation. The answer is found by straightforward substitution into (10), and after simplification we find that $\Phi(x)$ must satisfy

$$d\Phi/dx = [K(x) - 1] \cos^2 (k_0\Phi) + 1, \tag{33}$$

where the function $K(x)$ is

$$K(x) = k^2(x)/k_0^2.$$

The simple formula (33) provides all the information necessary to specify a precise solution of the second-order differential equation with virtually no complication for two important reasons. First, *if $k^2(x)$ is bounded then the derivative Φ' is always bounded.* Second, *this formula only involves $K(x) = k^2(x)/k_0^2$ so that imaginary quantities will never be introduced if the original differential equation's coefficients are real.* As a result, it is a trivial matter to solve equation (33) by an iterative process. That is, given $\Phi_0 = \Phi(x_0)$ at $x = x_0$, then $\Phi(x)$ is given by the expansion

$$\Phi(x) = \Phi_0 + \Phi_0'\Delta x + \tfrac{1}{2}\Phi_0''(\Delta x)^2 + \cdots + \frac{1}{n!}\Phi_0^{(n)}(\Delta x)^n + \cdots, \tag{34}$$

where the higher-order derivatives are easily calculated from (33),

$$\begin{aligned}
\Phi'' &= -k_0\Phi'(K - 1) \sin (2k_0\Phi) + K' \cos^2 (k_0\Phi) \\
\Phi''' &= -k_0\Phi''(K - 1) \sin (2k_0\Phi) - 2k_0\Phi' \\
&\quad \times [K' \sin (2k_0\Phi) + k_0(K - 1) \\
&\quad \times \cos (2k_0\Phi)] + K'' \cos^2 (k_0\Phi) \\
&\vdots \\
\Phi^{(n)} &= \text{etc.}
\end{aligned} \tag{35}$$

[9] P. Morse and W. P. Allis, Phys. Rev. 44, 269 (1933).

[10] F. Calogero, Nuovo Cimento, 27, 261 (1963).

[11] B. R. Levy and J. B. Keller, J. Math. Phys. 4, 54 (1963).

In other words we have a recurrance relation expressing the higher-order derivatives in terms of *known* lower-order ones. Once we are given Φ_0, it is no problem to determine what $\Phi_1 = \Phi(x_1)$ is at some nearby point x_1 to any degree of accuracy by truncating the expansion (34). A repetition of the process allows one to calculate $\Phi_2 = \Phi(x_2)$ and so on. No doubt some purists will argue that this is not a closed form representation of the solution. While this is true, it is in large part a meaningless objection. Any number of closed form expressions, e.g. contour integrals, are very elegant and concise but quite intractable when they are interrogated for numerical data. The important question to be answered is how readily can desired information be obtained. Before we can give any demonstrations of the utility of Eq. (33), we must digress and examine the intractability of a closed form solution, namely Eq. (3).

We have implied that once the logarithmic derivative $\Gamma(x)$ is known, we could reconstruct the solution $u(x)$ from the integral (3). While theoretically this formula should provide the desired answers, the practical situation is otherwise since the integral diverges at the poles of $\Gamma(x)$. If $u'(x)$ is bounded, then any such infinity in $\Gamma(x)$ is just another way of saying that $u(x)$ has a zero at that point, but nonetheless numerical integration of Eq. (3) to obtain $u(x)$ is to a great extent out of the question. In other words, even with an explicit $\Gamma(x)$ at our disposal, we may have all sorts of grief in translating this knowledge into numbers describing $u(x)$ by use of (3)—so much for closed form expressions.

The scheme we prefer to adopt neglects formula (3) and instead interpolates $u(x)$ by a polynomial at each stage of the iterative process. At some x_i, suppose we know Γ_i and an initial value u_i; at $x_{i+1} > x_i$ we are given Γ_{i+1}. These three pieces of information determine a unique second-order polynomial which must assume the value

$$u_{i+1} = u_i \frac{2 + \Gamma_i(x_{i+1} - x_i)}{2 - \Gamma_{i+1}(x_{i+1} - x_i)} \tag{36}$$

at $x = x_{i+1}$ (the derivation of this formula is given in the Appendix). The ability of this rational function to predict the future value u_{i+1} in terms u_i, Γ_i, and Γ_{i+1} is rather remarkable. For example, with an increment $x_{i+1} - x_i = 0.1$, Eq. (36) reproduces $\sin x$ from its logarithmic derivative $\cot x$ to within four significant figures. In addition, Eq. (36) automatically yields a zero of $u(x)$ should $\Gamma(x)$ have a pole at a point where $u(x)$ is regular, so that the role of the singularities of Γ is correctly reproduced.

With these preliminaries, we can find numerical solutions of second-order differential equations to any desired degree of accuracy. The iterative scheme which carries the solution from one point to another is to truncate the Taylor expansion (34) to a polynomial of suitable order for the increment chosen. A numerical example may serve to illustrate the procedure. The differential equation

$$d^2u/dx^2 + (a^2 - 6/x^2)\, u = 0 \tag{37}$$

is distinguished by having a particularly simple pair of solutions

$$u_1 = (3/ax)\cos ax + [1 - (3/a^2x^2)] \sin ax$$

$$u_2 = \frac{3}{ax}\cos\left(ax + \frac{\pi}{2}\right) + \left(1 - \frac{3}{a^2x^2}\right)\sin\left(ax + \frac{\pi}{2}\right). \tag{38}$$

Since (37) has a turning piont at $x = 6^{\frac{1}{2}}/a$ as well as a singularity at the origin it poses a rather severe computational challenge. We have used this equation as a test of our procedure for various choices of the parameter a. The larger this parameter, the more rapid are the oscillations of the solution for $|x| > 6^{\frac{1}{2}}/a$, and the more likely are errors to grow in any computational scheme. By use of a quadratic truncation of (35) we have found that the formulas (35) and (36) reproduce the solution very satisfactorily. As mentioned, the errors get worse with increasing a and Fig. 2, illustrates one of the poorer cases we have considered for which the parameter a had the value 3.5 and $6^{\frac{1}{2}}/a = 0.70$. The increment chosen was $\Delta x = 0.0142$, a rather odd number which resulted from choosing an increment inversely proportional to a and equal to 0.1 if $a = 0.5$. At $x = -10$, the initial logarithmic derivative was adjusted to match u_2, the solution with a singularity at the origin. From this initial value a solution was computed in the direction of increasing x using a quadratic formula for $\Phi(x)$ to compute $\Gamma(x)$. This was followed by use of Eq. (36) to determine the solution. The calculation and the actual solution agree so well that it is not worthwhile to draw curves of each. Instead, the absolute error curve is drawn with a *magnified* scale. As might be expected in any initial-value calculation, the errors tend to grow as the computation progresses. However, except in the neighborhood of the singularity we note that the amplitude of the error increases but linearly, and furthermore it oscillates about zero. The divergence at the origin is more apparent than real since the percentage error remains

bounded. By choosing a smaller increment or adding an additional term to the expansion for $\Phi(x)$, the error can be reduced by an additional order of magnitude. We might add that the computations are sufficiently simple and proceed so fast that they can be performed with but the aid of a desk calculator. Even on a relatively slow IBM 1410 computer, the calculations were performed as fast as the machine could print out the answers.

The preceding example illustrates the ease with which a second-order equation can be solved without any complications or need of any sophisticated numerical techniques. Of course, certain tricks of the trade can be very beneficial in reducing errors. For example we recommend use of a variable increment which is smaller at the initial stages of the calculation to reduce a buildup of error. In addition the increment $x_{i+1} - x_i$ should be scaled so that $\Phi_{i+1} - \Phi_i$ does not exceed some preassigned bound. In this fashion, a computer program can be both speeded up and made more reliable by using a floating increment whose size varies inversely with the size of the change in $\Phi(x)$ from x_i to x_{i+1}. Finally, we note that if the logarithmic derivative Γ_u of Eq. (6) is known, then the logarithmic derivative Γ_U of equation (1) which inspired the analysis is given by

$$\Gamma_U = \Gamma_u - \tfrac{1}{2}A(x) \tag{39}$$

so that without any further ado, the solution of Eq. (1) can be found.

As a last remark we should like to point out that the results of this section complement rather than supersede the discrete analysis of Sec. II for the case of a differential equation with analytic coefficients. For whereas this section is concerned with accurate representations which need a relatively fine interval for their calculation, the analysis in Sec. II yields good approximation for coarser intervals. While this may be of no matter when an equation is to be solved but once, the distinction becomes important when the differential equation must be solved many times for some set of perturbations in its coefficients. The formulas in Sec. II furnish a quick method of obtaining insight into the effect of a large number of variations and locating regions demanding further investigation. Following this exploration, the formulas of Sec. III can chart these unknown domains. To cite a nautical analogy: it is the difference between using dead reckoning and an inertial guidance system. One is cheap and often adequate, the other is more expensive but arbitrarily accurate.

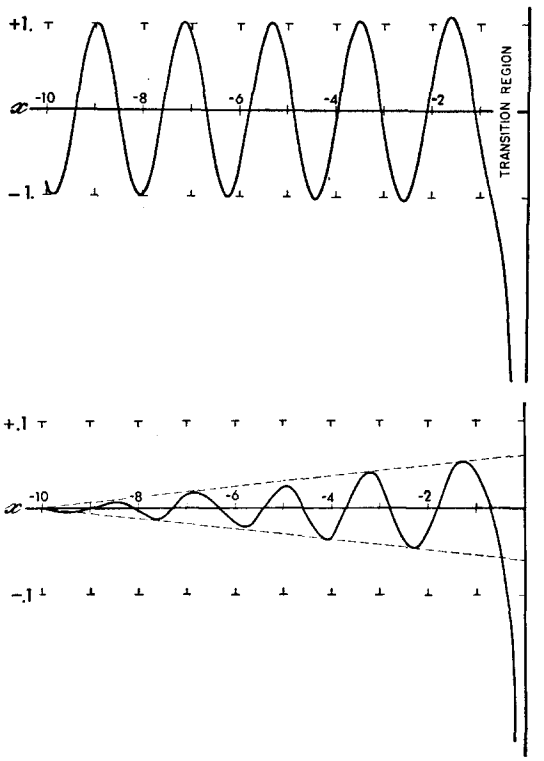


FIG. 2. The upper curve is the solution $u_2(x)$ of Eq. (37), and the lower curve is the computational error of the numerical procedure described in the text. Note that there is a 10:1 ratio between the two scales.

ACKNOWLEDGMENT

The research described in this report has been sponsored by the Advanced Research Projects Agency, Project Vela-Uniform, under Contract AF 19(628)-319 of the Geophysics Research Directorate, Air Force Cambridge Research Laboratories, Office of Aerospace Research, U. S. Air Force, Bedford, Massachusetts.

APPENDIX

In this Appendix we shall derive the formula (36) in Sec. III. From the equation (9) we have

$$u\Gamma = u'. \tag{40}$$

With no loss of generality suppose $x_i = 0$. Let us assume that the function $u(x)$ can be approximated by some unknown quadratic expression which can be chosen in the completely general form

$$u(x) = \alpha x(x - \Delta) + \beta(x - \Delta)(x + \Delta) + \gamma x(x + \Delta), \tag{41}$$

where α, β, γ are constants to be determined and Δ is the length of the increment.

Substituting (41) into (40) we find that

$$[x^2(\alpha+\beta+\gamma)+x\Delta(-\alpha+\gamma)-\beta\Delta^2]\Gamma = 2x(\alpha+\beta+\gamma)+\Delta(-\alpha+\gamma). \quad (42)$$

Let u_+, u_0, u_- and Γ_+, Γ_0, Γ_- denote the values of u and Γ at the points $x=\Delta$, $x=0$, $x=-\Delta$, respectively. Then, from (42) we obtain

$$\begin{aligned} u_+\Gamma_+ &= (\alpha+2\beta+3\gamma)\Delta = 2\gamma\Delta^2\Gamma_+, \\ u_0\Gamma_0 &= (-\alpha+\gamma)\Delta = -\beta\Delta^2\Gamma_0, \\ u_-\Gamma_- &= -(3\alpha+2\beta+\gamma)\Delta = 2\Delta^2\alpha\Gamma_-. \end{aligned} \quad (43)$$

The last two equations imply that

$$\frac{u_-}{u_0} = -\frac{2\alpha}{\beta} = \frac{2-\Delta\Gamma_0}{2+\Delta\Gamma_-} \quad (44)$$

or

$$u_0 = u_-(2+\Delta\Gamma_-)/(2-\Delta\Gamma_0) \quad (45)$$

By similar reasoning more sophisticated formulas can be derived whose accuracy is of arbitrarily high order.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 6 JUNE 1965

Mathematical Study of the Cutoff Procedure for Divergent Integrals in Bound-State Problems

J. MANDELBROJT
Faculté des Sciences, Marseilles, France
AND
S. MANDELBROJT
Collège de France, Paris, France and Department of Mathematics, Rice University, Houston, Texas
(Received 15 October 1964)

We consider the class of conveniently normalized functions having a given rate of decrease at infinity on the real axis. The lower bound of the length of the supports of the Fourier transforms of such a given class of functions is evaluated. This could be considered as a refined form of the uncertainty relation. The condition, which is necessary and sufficient for this lower bound to be finite is indicated. In the physically important case when this lower bound is infinite, a length, which we call the "natural cutoff of this class of functions" is nevertheless defined, it characterizes the interval outside of which the value of the Fourier transform of a function of the given class may be negligible. A construction which gives the cutoff, or the natural cutoff, is given; it is valid for practically any given rate of decrease. We carry out the calculation explicitly in the case of a decrease of the functions characterized by $|F(x)| \leq A\, e^{-\mu|x|}$ (in this case the support of the Fourier transform is necessarily infinite). A canonical function is built such that the abscissa for which it begins to become "small" should be reasonably considered to be the natural cutoff of the Fourier transform of all the functions satisfying this inequality. The same method can be used for the class of functions satisfying an inequality of the form $|F(x)| \leq A \exp[-C(x)]$, with $\int_1^\infty C(u)\, u^{-3}\, du < \infty$.

INTRODUCTION

THIS paper deals with an evaluation of the support of the Fourier transform of a function when, roughly speaking, the behavior of the function can be evaluated at infinity on the x axis and when a normalization condition is given. The functions we are concerned with will be either wavefunctions or scattering amplitudes.

The properties we give are particularly well adapted to the problem of evaluating cutoffs in momentum space in field theory, due to the following feature of all the approximations used to calculate interactions in field theory: when one stops at a given order of approximation, the interactions between particles are always known when the particles are far apart and not well known for small distances. We give three examples.

(1) In the case of nuclear forces due to the exchange of π mesons, the exchange of one meson gives the tail of the potential, and as the number of virtual mesons exchanged increases one gets nearer to the origin. After a calculation to a finite order of perturbation, the nuclear forces are known only for a large enough distance—between the nucleons.

(2) In dispersion theory, in bootstrap calculations for instance, solving the coupled integral equations

Substituting (41) into (40) we find that

$$[x^2(\alpha+\beta+\gamma)+x\Delta(-\alpha+\gamma)-\beta\Delta^2]\Gamma = 2x(\alpha+\beta+\gamma)+\Delta(-\alpha+\gamma). \quad (42)$$

Let u_+, u_0, u_- and Γ_+, Γ_0, Γ_- denote the values of u and Γ at the points $x=\Delta$, $x=0$, $x=-\Delta$, respectively. Then, from (42) we obtain

$$\begin{aligned} u_+\Gamma_+ &= (\alpha+2\beta+3\gamma)\Delta = 2\gamma\Delta^2\Gamma_+, \\ u_0\Gamma_0 &= (-\alpha+\gamma)\Delta = -\beta\Delta^2\Gamma_0, \\ u_-\Gamma_- &= -(3\alpha+2\beta+\gamma)\Delta = 2\Delta^2\alpha\Gamma_-. \end{aligned} \quad (43)$$

The last two equations imply that

$$\frac{u_-}{u_0} = -\frac{2\alpha}{\beta} = \frac{2-\Delta\Gamma_0}{2+\Delta\Gamma_-} \quad (44)$$

or

$$u_0 = u_-(2+\Delta\Gamma_-)/(2-\Delta\Gamma_0) \quad (45)$$

By similar reasoning more sophisticated formulas can be derived whose accuracy is of arbitrarily high order.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 6 JUNE 1965

Mathematical Study of the Cutoff Procedure for Divergent Integrals in Bound-State Problems

J. MANDELBROJT
Faculté des Sciences, Marseilles, France
AND
S. MANDELBROJT
Collège de France, Paris, France and Department of Mathematics, Rice University, Houston, Texas
(Received 15 October 1964)

We consider the class of conveniently normalized functions having a given rate of decrease at infinity on the real axis. The lower bound of the length of the supports of the Fourier transforms of such a given class of functions is evaluated. This could be considered as a refined form of the uncertainty relation. The condition, which is necessary and sufficient for this lower bound to be finite is indicated. In the physically important case when this lower bound is infinite, a length, which we call the "natural cutoff of this class of functions" is nevertheless defined, it characterizes the interval outside of which the value of the Fourier transform of a function of the given class may be negligible. A construction which gives the cutoff, or the natural cutoff, is given; it is valid for practically any given rate of decrease. We carry out the calculation explicitly in the case of a decrease of the functions characterized by $|F(x)| \leq A\, e^{-\mu|x|}$ (in this case the support of the Fourier transform is necessarily infinite). A canonical function is built such that the abscissa for which it begins to become "small" should be reasonably considered to be the natural cutoff of the Fourier transform of all the functions satisfying this inequality. The same method can be used for the class of functions satisfying an inequality of the form $|F(x)| \leq A \exp[-C(x)]$, with $\int_1^\infty C(u)\, u^{-3}\, du < \infty$.

INTRODUCTION

THIS paper deals with an evaluation of the support of the Fourier transform of a function when, roughly speaking, the behavior of the function can be evaluated at infinity on the x axis and when a normalization condition is given. The functions we are concerned with will be either wavefunctions or scattering amplitudes.

The properties we give are particularly well adapted to the problem of evaluating cutoffs in momentum space in field theory, due to the following feature of all the approximations used to calculate interactions in field theory: when one stops at a given order of approximation, the interactions between particles are always known when the particles are far apart and not well known for small distances. We give three examples.

(1) In the case of nuclear forces due to the exchange of π mesons, the exchange of one meson gives the tail of the potential, and as the number of virtual mesons exchanged increases one gets nearer to the origin. After a calculation to a finite order of perturbation, the nuclear forces are known only for a large enough distance—between the nucleons.

(2) In dispersion theory, in bootstrap calculations for instance, solving the coupled integral equations

by successive iterations amounts to introducing in a given channel "potentials,"[1] such as they can be defined in field theory, with shorter and shorter ranges; and one supposes that "the effect of short-range forces can be imitated by a simple cutoff procedure."[2]

(3) In the problem of finding bound states of a fermion field interacting with itself via a Fermi interaction, calculating the creation operator of a bound boson at rest, expressed in the "one-pair approximation" in terms of the creation operators of the fermion and antifermion, we have

$$A^* = \int \{\lambda(u)b^*(u)d^*(-u) + \mu(u)b(u)d(-u)\}\, d\mathbf{u},$$

where A^* is the creation operator of the bound boson at rest, $b^*(u)$, $b(u)$ are the creation and annihilation operators of the fermion with momentum $\mathbf{u}$, $d^*(u)$ and $d(u)$ the creation and annihilation operators of the antifermion with momentum $\mathbf{u}$.

One obtains the eigenvalue equation for the mass ω of the bound state in the form[3]

$$\text{(i)} \qquad \int (|\lambda^2(u)| + |\mu^2(u)|)\, d^3u = 1,$$

$$\text{(ii)} \qquad \int (\lambda(u) - \mu(u))\, d^3u = 0,$$

with

$$\lambda = 1/(2E_u - \omega), \qquad E_u = (u^2 + m^2)^{\frac{1}{2}},$$

where m is the fermion's mass.

Here $\lambda(u)$ can be interpreted as the "wavefunction" representing the probability for the presence of a pair $b^*(u)d^*(-u)$ in the bound state A^*. Thus one expects that $\lambda(u)$ should have a Fourier transform $\hat{\lambda}(x)$ which behaves like $\exp(-\mu|x|)$ for x large, *μ being the binding energy of the fermions.* [To see the asymptotic behavior of $\hat{\lambda}(x)$ one has simply to look at the pole of $\lambda(u)$ which is nearest the real axis. This is the way one should proceed in general when the function is given in momentum space.]

It is indeed the case, and the problem we want to solve is to determine what cutoff on the momentum variable u can one introduce so as to make the integrals (i) and (ii) convergent and *so as still to have* $|\hat{\lambda}(x)| \leq \exp(-\mu|x|)$ *for x large.* In this problem again we notice that, due to the one pair approximation, only the expression of $\hat{\lambda}(x)$ for x large has a direct physical meaning.

The solution of the Schrödinger equation with the potential calculated in a certain order of perturbation in the first example, or the integrand in an integral equation after a certain number of iterations in the second example, or the function λ in the one pair approximation in the third example is called "the perturbative function f_p." The function obtained from this one by further introducing a cutoff in momentum space is called the "modified function f_m"; a function calculated from the perturbative function by any process, such as integration, is also called a perturbative function, and the function obtained by the same process from the modified function is called modified function. We also use the terminology "exact potential V_e" "exact function" etc. ... in the obvious sense: the potential calculated to *all* orders of perturbation or a function deduced from this one. *The philosophy of the cutoff procedure is that the "modified function" is hoped to be a better approximation to the exact function than the perturbative function.*

Now, as we have said, the mathematical properties we will give indicate, roughly speaking, what smoothed cutoff can be put on a perturbative function in momentum space so as to respect in a certain sense its asymptotic behavior in x space and so as to have a normalization condition such as (i). (In the case of dispersion theory the normalization condition consists in giving the value of the scattering amplitude for a given value of the variable s.) Thus they indicate what is the smallest cutoff compatible with the knowledge of the interaction that one has acquired after having stopped at a given order of perturbation, the idea being that the cutoff procedure remedies somewhat the imperfection due to the fact that we use functions calculated only in a finite order of perturbation. But it is then essential that the cutoff procedure should not spoil the correct features of the perturbative function, and among these correct features an important one is as we have shown, the behavior of $f(x)$ for x large.

We now consider, to be more specific, the perturbative wavefunction ψ_p representing an S-wave bound state

$$\psi_p(\mathbf{x}) = v_p(r)/(4\pi)^{\frac{1}{2}}r, \qquad r = |\mathbf{x}| = (x_1^2 + x_2^2 + x_3^2)^{\frac{1}{2}},$$

where $v_p(r)$ is the perturbative radial wavefunction.

The three-dimensional Fourier transform of $\psi_p(\mathbf{x})$ is

$$\hat{\psi}_p(\mathbf{u}) = -\pi^{\frac{1}{2}} \int_{-\infty}^{+\infty} F_p(x) \exp(iux)\, dx,$$

[1] G. Chew, "S-Matrix Theory of Strong Interactions" in *Frontiers in Physics* (W. A. Benjamin, Inc., New York, 1962).
[2] E. Abers and C. Zemach, Phys. Rev. **131**, 2305 (1963).
[3] C. Ihara and S. Hatano: Progr. Theoret. Phys. (Kyoto) **20**, 356 (1958).

where $F_p(x)$ is the primitive which tends to zero as $x \to \infty$ of the function $w_p(x)$ defined by $w_p(x) = v_p(x)$ for $x > 0$, $w_p(x) = -v_p(-x)$ for $x < 0$.

$w_p(x)$ is, for instance, the bound-state solution of a one-dimensional Schrödinger equation

$$-\hbar^2(2m)^{-1}\, d^2w_p/dx^2 + (V_p - E)w_p = 0,$$

where $V_p(x)$ is a perturbative potential (i.e., an approximate potential in the sense already given).

We start by making the following remark which is fundamental for the problem we consider: classical theorems show that, no matter how much the perturbative function $F_p(x)$ approaches the exact function $F_e(x)$, as long as they are not actually equal on the whole x axis. the behavior for $|u| \to \infty$ of the Fourier transform $\hat{F}_p(u)$ of $F_p(x)$ gives no indication on the behavior of $\hat{F}_e(u)$ for $|u| \to \infty$. Thus we must look elsewhere for an indication on the asymptotic behavior of $\hat{F}_e(u)$ for $u \to \infty$.

Such an indication is given by the fact that one gets good numerical results in many problems when one introduces a damping function or a cutoff for u, that is to say a restriction of the Fourier transform on an interval around the origin. We therefore assume that $\hat{F}_e(u)$ decreases with great rapidity, i.e., that *it is well approximated by a* $\hat{F}_m(u)$ *corresponding to a certain well chosen cutoff introduced in* $\hat{F}_p(u)$ and the problem is: *what can be said about the value* u_0 *of the variable* u *such that for* $u > u_0$, $\hat{F}_e(u)$ *is small enough to be neglected* ($|\hat{F}_e(u)|$ small for $|u| > u_0$) *when the behavior of* $F_e(x)$ *for* $|x| \to \infty$ *is known.* (From the considerations in the introduction one can reasonably assume that $F_e(x) \sim F_p(x)$ for $|x| \to \infty$.)

We can sum up the situation by saying that $F_e(x)$ is well approximated by $F_p(x)$ for $|x| \to \infty$ (property of approximation methods used in field theory), but $\hat{F}_e(u)$ is suitably approximated by $\hat{F}_m(u)$, for $|u| \to \infty$.

II. PRELIMINARY MATHEMATICAL CONSIDERATIONS

It follows from well-known theorems of mathematical analysis, that the Fourier transform of a function $F(x)$, not identically zero, satisfying an inequality of the form

$$F(x) \leq \exp[-C(|x|)] \qquad (\alpha)$$

with

$$\int C(u)u^{-2}\, du = \infty \qquad \text{(for instance } C(u) = \mu\,|u|)$$

cannot have a finite support.[4] But it can be shown that, $C(u)$ being nondecreasing and such that

$$\int C(u)u^{-2}\, du < \infty, \qquad (\beta)$$

then there exists a quantity δ^* with the property that no function F satisfying (α) [with C satisfying (β)] and properly normalized (by its L_2 norm, or simply by the absolute value of its Fourier transform at the origin) can have its support inside $(-\delta^*, \delta^*)$ (see for instance Ref. 5).

For a general class of functions $C(u)$, for instance when $C(u)$ is a power of u, δ^* is of the form $\delta^* = \gamma\Delta$ where γ is a constant independent of the norm, and where $[-\Delta, \Delta]$ is the support of a function $f(u)$, even and not negative, the construction of which is useful for our purpose. If we denote by $C_\alpha(u)$ the restriction of $C(u)$ in the interval (α, ∞) $(\alpha > 0)$ and $[C_\alpha(u)]$ its integer part [i.e., the largest integer not larger than $C_\alpha(u)$] we consider the points of abscissa $1/\nu_n$ where $[C_\alpha(u)]$ has discontinuities. k_n being the corresponding jump of $[C_\alpha(u)]$, we denote by $\{\mu_n\}$ the sequence of all the ν_n each one repeated k_n times.

If we denote then by $p_\beta(u)$ the function equal to $1/2\beta$ for $|u| \leq \beta$ and 0 elsewhere, we construct the infinite convolution

$$a \lim_{n=\infty} p_a * p_a * p_{\mu_1} * p_{\mu_2} * \cdots * p_{\mu_n};$$

that is to say the function

$$p(u) = \lim_{n=\infty} \pi(\mu_1\mu_2 \cdots \mu_n 2^n)^{-1} \times \int_{-\alpha}^{\alpha} dv \int_{-\mu_n}^{\mu_n} dt_n \int_{-\mu_{n-1}}^{\mu_{n-1}} dt_{n-1} \cdots \times \int_{-\mu_1}^{\mu_1} p_a(u + v + t_1 + \cdots + t_n)\, dt_1.$$

This limit exists whenever $\int_1^\infty C(u)u^{-2}$ exists and p is the Fourier transform of the restriction on the real axis R of the entire function

$$p(z) = \frac{(\sin az)^2}{az^2} \prod_1^\infty \frac{\sin \mu_n z}{\mu_n z} \qquad (z = x + iy).$$

With a suitable choice of α and a, the support of the corresponding function p gives the value of Δ.

[4] This follows from Paley–Wiener's theorem and a theorem solving the famous Watson's problem. See S. Mandelbrojt *Series adherentes. Régularisation des suites. Applications.* (Gauthier-Villars, Paris, 1952), and S. Mandelbrojt "Classes of infinitely differentiable functions" The Rice Institute Pamphlet Vol. XXIX (1942).

[5] S. Mandelbrojt, J. Anal. Math. (Jerusalem) **10**, 381 (1962–1963).

If we take $\hat{\Phi}(u) = p(u)\ ||p||^{-1}$, $\hat{\Phi}$ is the Fourier transform of a function Φ such that

$$|\Phi(x)| \leq A \exp[-C(|x|)], \qquad A = ||p||^{-1},$$
$$||\hat{\Phi}|| = 1, \qquad \hat{\Phi}(u) = 0 \quad \text{for} \quad |u| > \Delta,$$

and for every function F of which the Fourier transform has its support in $(-\delta, \delta)$ and which satisfies (α) {F is roughly speaking the restriction of an entire function $F(z)$ satisfying $|F(z)| \leq A \exp[\delta\ |\ y| - C(|x|)]$}, we have $\delta \geq \gamma\Delta$.

Our purpose is to show that in the case of divergence of the integral in (β) [for instance if $C(u) = \mu\ |u|$], an analogous construction yields no longer a sharp cutoff, but a quantity which, to our understanding, should replace it.

III. AN EXAMPLE USEFUL IN PHYSICS

We now turn to the simple example where $\hat{w}_p(u)$ corresponds to a bound state of binding energy μ [$\hat{w}_p(u)$ is, for instance, the function $\lambda(u)$ of example 3 in the introduction, the normalization condition we will use is the one corresponding to that example]. We have taken for simplicity the function $\mu(u) = 0$ (not to be mixed up with the binding energy μ). Then $\hat{w}_p(u)$ has a pole on the imaginary axis at a distance μ from the real axis.

The functions $w_p(x)$ are defined up to a multiplicative constant. We suppose that one of them is asymptotic to $\exp(-\mu\ |x|)$ at $x = \pm\infty$. This is the solution we choose. (The choice of this solution can usually be made without having to go over to the x space. One has just to look at the residue of the pole μ.) The essence of the cutoff procedure is that

$$\hat{w}_m(u) \approx A'\hat{w}_p(u) \quad \text{for} \quad -\delta < u < \delta,$$
$$\hat{w}_m(u) \quad \text{negligible for} \quad |u| > \delta,$$

A' being a multiplicative constant which corresponds to the fact that the functions w are solutions of an homogeneous equation, δ being the hitherto unknown cutoff.

The constant A' is related to the cutoff δ by the normalization relationship

$$\int_{-\infty}^{+\infty} \hat{w}_m^2(u)\, du \approx \int_{-\delta}^{\delta} \hat{w}_m^2(u)\, du = A'^2 \int_{-\delta}^{\delta} \hat{w}_p^2(u)\, du = 1.$$

This is a first relationship between A' and δ [let us recall that $\hat{w}_p(u)$ is known, for instance in the example 3, it is $\lambda(u)$]. The mathematical theory which we use will give us a second relationship as follows.

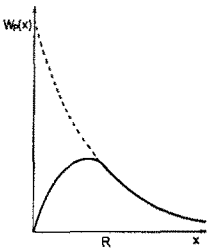


FIG. 1. Curve representing $w_p(x)$. The dotted curve represents $\exp A'(-\mu|x|)$.

[The first relationship depends specifically on the function $\hat{w}_p(u)$, whereas the second relationship which we derive later will be the same for all the functions having the same asymptotic behavior for $|x| \to \infty$.]

According to the remarks given in the introduction $w_m(x)$ will behave for $|x| \to \infty$ like $A'w_p(x)$ that is like $A' \exp(-\mu\ |x|)$. We are thus led to the study of the more general class of functions which satisfy the following two conditions:

$$\limsup_{|x|\to\infty} |w_m(x)| \exp(\mu\ |x|) \leq A',$$
$$\int_{-\infty}^{+\infty} |\hat{w}_m(u)|^2\, du = 1.$$

We now replace these conditions by conditions on $F_m(x)$ [which is defined as the primitive of $w_m(x)$ which tends to zero for $x \to \infty$]:

$$\limsup_{|x|=\infty} \exp(\mu\ |x|)\ |F_m(x)| \leq A, \quad \text{with} \quad A = A'/\mu$$
$$\int_{-\infty}^{+\infty} |u^2\hat{F}^2(u)|\, du = 2\pi. \qquad (\gamma)$$

We now go one step further and replace the asymptotic condition (γ) by the inequality

$$|F_m(x)| \leq A \exp(-\mu\ |x|) \qquad (1)$$

on the whole axis. The fact that this condition is imposed on the whole x axis is the weaker point of our argument. However, if one bears in mind that $w_p(x)$ corresponds to a bound state, and thus, the interaction being attractive, has its curvature at the inflexion point R as indicated (see Fig. 1, R being roughly the range of the potential in the case of a Schrödinger equation) then one sees that $w_p(x)$ and also $F_p(x)$, do satisfy the inequality on the whole x axis. Now recalling that $w_m(x)$ is obtained from $w_p(x)$ by modifying the interaction at *short* distances [if one has gone far enough in the perturbation series for the calculation of $w_p(x)$] and recalling furthermore that $F_m(x)$ is a primitive of $w_m(x)$ so that the modifications on $w_p(x)$ to obtain $w_m(x)$ are somewhat smoothed out by the process of integration [since $w_m(x)$ is different from $w_p(x)$ only in a small

region around the origin for x] it is seen that it is natural to impose the condition $|F(x)| \leq A \exp(-\mu |x|)$ on the whole x axis. We are thus led to the study of the Fourier transform of functions $F(x)$ which satisfy the inequality

$$|F(x)| \leq A \exp(-\mu |x|) \tag{2}$$

and the normalization relationship

$$\int_{-\infty}^{\infty} (u^2 \hat{F}^2(u))\, du = ||u\hat{F}(u)||_2^2 = 2\pi. \tag{3}$$

Now recalling that the Fourier transform of a function $F(x)$ satisfying (2) cannot have a finite support, i.e., $F(x)$ cannot be the restriction on the real axis of an entire function $F(z)$ of exponential type, we thus discard the condition that $F(z)$ is of exponential type and replace it by the weaker hypothesis that it is an entire function, which condition is natural enough, since we want $\hat{F}(u)$ to decrease rapidly when $|u| \to \infty$.

But even with only these hypotheses the construction mentioned at the end of II still provides a function [denoted below by $K(\mu, A; u)$] of which the *inflexion point* (for $u > 0$) generates a "natural cutoff" of the Fourier transform of all functions F satisfying (2) and (3). We feel it necessary to go through the mathematical analysis which follows, since it justifies, on one hand, the point of view just mentioned, and permits us, on the other hand, to carry out the numerical applications of our principle.

Let us begin by some preliminary considerations. Denote by

$\Delta_c(u)$ the function defined for $c > 0$,

$$\Delta_c(u) = 1, \quad \text{for} \quad |u| \leq c,$$

$$\Delta_c(u) = 0, \quad \text{for} \quad |u| > c;$$

and set

$$\mathfrak{M}_n(y) = \pi n!\, 2^{-n} \int_{-1/n}^{1/n} dt_n \int_{-1/(n-1)}^{1/(n-1)} dt_{n-1} \cdots \times \int_{-1}^{1} \Delta_c(y + t_1 + t_2 + \cdots + t_n)\, dt_1$$

$$\mathfrak{M}(y) = \lim_{n=\infty} \mathfrak{M}_n(y).$$

This limit exists and represents an even positive function. We shall also write, $\psi(y)$ being a (locally integrable) function:

$$M_c(\psi(u)) = \frac{1}{2c} \int_{-c}^{+c} \psi(u + y)\, dy.$$

It can be shown that $M_c(\mathfrak{M}(u))$ is the Fourier transform of

$$F_c(x) = \frac{\sin^2(cx)}{(cx^2)} \prod_{n=1}^{\infty} \sin\left(\frac{x}{n}\right) \frac{n}{x},$$

the product $\prod (\sin (z/n)n/z)$ converging uniformly in each compact of the complex plane, $F_c(z)$ is thus an even entire function.

It satisfies, moreover, the following inequality:

$$|F_c(z)| \leq A_c \exp[\varphi_c(|y|) - |z|] \qquad (z = x + iy) \tag{4}$$

where A_c depends only on c, and where (see Ref. 5)

$$\varphi_c(u) = \sum \left[\min\left(\frac{u}{n}, \frac{u^2}{2n^2}\right) + \min\left(cu, \frac{c^2u^2}{2}\right)\right].$$

We have, moreover,

$$|F_c(x)| \leq B_c \exp(-|x|),$$

where

$$B_c = c \exp(p_c), \tag{5}$$

the quantity p_c being defined by $p_c = \max(p, 1)$ with p given by

$$c^2 p^{\frac{1}{2}} \exp(p) = (2\pi)^{\frac{1}{2}} \exp(13/12).$$

This follows from the obvious inequalities

$$|F_c(x)| \leq c \qquad (-\infty < x < \infty),$$

$$|F_c(x)| \leq (cx)^{-2} \inf_n n!\, x^{-n} \leq cx^{-2}\Gamma([x])x^{1-[x]} \qquad (x > 0),$$

where $[x]$ denotes the integral part of x, the second inequality furnishing, by Stirling's formula, for $x \geq 1$:

$$|F_c(x)| \leq (cx^2)^{-1}(2\pi x)^{\frac{1}{2}} \exp\left(-[x] + \frac{\theta}{12[x]}\right), \qquad 0 < \theta < 1.$$

We set now for $a > 0$, $\mu > 0$:

$$F_{\mu,a}(z) = \frac{\sin^2(az)}{az^2} \prod_{n=1}^{\infty} \sin\left(\frac{\mu z}{n}\right) \frac{n}{\mu z}.$$

It is clear that $F_{\mu,a}(z) = \mu\,]_{a/\mu}(\mu z)$ and

$$[F_{\mu,a}(x)]\hat{} = \mu[F_{a/\mu}(\mu x)]\hat{} = \hat{F}_{a/\mu}(u/\mu).$$

With B_c defined as in (5) we get

$$|F_{\mu,a}(x)| \leq \mu B_{a/\mu} \exp(-\mu |x|). \tag{6}$$

We have also

$$|F_{\mu,a}(z)| \leq A_{\mu,a} \exp[\varphi_{\mu,a}(|y|) - \mu |z|], \tag{7}$$

where $A_{\mu,a} = \mu A_{a/\mu}$, A_c being defined as in (4), and where

$$\varphi_{\mu,a}(u) = \sum_n \min(\mu u/n, \mu^2 u^2/2n^2) + \min(au, a^2u^2/2) \qquad (u > 0).$$

$A > 0$ being given, let us define $a = a(A)$, by $a = \inf a'$, where

$$A\,||u\hat{F}_{\mu,a'}(u)||_2 \geq \mu B_{a'/\mu}(2\pi)^{\frac{1}{2}}. \tag{8}$$

This inequality amounts to

$$A_\mu^{\frac{1}{2}}\,||u\hat{F}_{a'/\mu}(u)||_2 \geq (2\pi)^{\frac{1}{2}}B_{a'/\mu}. \tag{9}$$

Let us set, with this choice of a,

$$K_0(x) = K_0(\mu, A; x) = F_{\mu,a}(x)/||F'_{\mu,a}(x)||_2. \tag{10}$$

We have then obviously

$$||K_0'(x)||_2 = 1$$

and

$$|K_0(x)| \leq A \exp(-\mu|x|). \tag{11}$$

Denote by $K(\mu, A; u)$ the Fourier transform of $K_0(\mu, A; u)$:

$$K(u) = K(\mu, A; u) = [K_0(\mu, A; x)]\hat{}.$$

The inflection point of the function $K(u)$ so determined furnishes the natural cutoff of functions $F(x)$ satisfying (2) and (3).

We shall now explain why this way of introducing the cutoff seems natural to us.

As it has been said before, if (β) holds and if $|F(x)| \leq \exp[-C(|x|)]$, $|F(0)| \geq B > 0$ (without having to suppose *a priori* an inequality in the complex plane), then $\hat{F}(u)$ cannot be zero for $|u| > \delta^* - \epsilon$ ($\epsilon > 0$), δ^* being a number depending on the quantity B and the function C [at least if we suppose that $C(u)$ is a convex function of log u ($u > 0$)].

And, as we said, there is nothing to change in the reasoning when the condition $|\hat{F}(0)| \geq B$ is replaced by a condition of the type $||\hat{F}(u)||_2 \geq B$ or $||u\hat{F}(u)||_2 \geq B$; the quantity δ^* has only to be replaced by a corresponding quantity δ^{**} or δ^{***}. The function furnishing the quantity δ^*, when (β) holds, is given by the function (here we set $a = 1$, and K_0^* is then the corresponding p of Sec. II)

$$K_0^*(z) = \sin^2\left(\frac{z}{2}\right)z^{-2}\prod\frac{\sin(\mu_n z)}{\mu_n z}, \tag{12}$$

where the μ_n are defined above. And we have then

$$K_0^*(u) = \lim_{n=\infty}\pi(2^n\mu_1\mu_2\cdots\mu_n)^{-1}\int_{-a}^{a}dv\int_{-\mu_n}^{\mu_n}dt_n\cdots \times\int_{-\mu_1}^{\mu_1}\Delta(y+v+t_1+t_2\cdots t_n)\,dt_1,$$

where $\Delta(u) = 1$ for $|u| \leq \frac{1}{2}$, $\Delta(u) = 0$ for $|u| > \frac{1}{2}$. If

$$\int_1^\infty C(u)u^{-2}\,du = \infty,$$

and if

$$|F(x)| \leq \exp[-C(|x|)],$$

there is no δ such that $\hat{F}(u) = 0$ for $|u| > \delta$. But, provided the condition

$$\int_1^\infty C(u)u^{-3}\,du < \infty \tag{14}$$

is satisfied, the previous constructions are still possible: the expressions (12) and (13) still converge and represent correspondingly an even entire function and the Fourier transform of its restriction on the real axis.

The Fourier transform $\hat{K}^*(u)$ [if (14) holds but (β) does not hold] is not any more zero outside an interval, but it decreases rapidly to zero when $u > 0$ tends to ∞ (and, of course, when u tends to $-\infty$, since the function is even), most of its decreasing being realized on the part of the axis stretching from the origin to its inflection point; afterwards its derivative is virtually stabilized—it is, of course, still negative but it is near to zero.

The abscissa of the inflection point of $\hat{K}^*(u)$ (for $u > 0$), multiplied by the numerical constant γ introduced above, seems to be a natural substitute for the nonexisting δ^* (or δ^{**} or δ^{***} when adequate normalizations are performed).

Of course the case $C(u) = \mu u$ is a particular one corresponding to (14); the quantities μ_n are then equal to μ/n and the construction we dealt with above corresponds exactly to what has to be done in that case. (The constant a is introduced for normalization purposes).

There is also another reason why the graph of $K(u) = K(\mu, A; u)$ should be considered as furnishing a "natural cutoff" for functions satisfying the condition $|F(x)| \leq A \exp(-\mu|x|)$ and normalized in a proper way.

To construct $K(u)$ one begins by considering the function $\Delta_c(u)$, with $\Delta_c(u) = 1$ for $|u| \leq c$, $\Delta_c(u) = 0$ for $|u| > c$. Now $\pi\Delta_c(u)$ is the Fourier transform of $\sin(cx)/x$. The real cutoff of its Fourier transform, represented by $u = c$, is then successively increased by $1/n$ ($n \geq 1$), this Fourier transform being diminished in parts (mostly at the ends) of the newly obtained interval (that is to say the Fourier transform being "swept" by an averaging process to all the points of the new support), the function

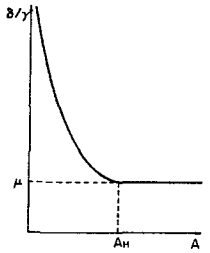


FIG. 2. Curve representing δ/γ.

itself being each time multiplied by $(\sin x/n)\ n/x$; these successive increases of the support of the Fourier transform seem to be the smallest possible consistent with the fact that the obtained function [here

$$\sin \frac{cx}{x} \prod \sin \left(\frac{x}{n}\right) \frac{n}{x}]$$

should decrease as $A \exp(-|x|)$.In other words, a deformation as "economical" as possible on a "true cutoff" is made in order to obtain a function with a given behavior at infinity.

Simple changes of variables and normalizations lead to the function $K(u)$ itself.

It is now our purpose to show how fast $K(\mu, A; u)$ decreases when $|u| \to \infty$. We have

$$\hat{F}_c(u) = \int_{-\infty}^{\infty} F_c(x) \exp(-ixu)\, dx,$$

and, y being any real number, we see easily by Cauchy's theorem that we have also

$$\hat{F}_c(u) = \int_{-\infty+iy}^{\infty+iy} F_c(z) \exp(-iuz)\, dx \qquad (z = x + iy).$$

It follows then from the inequality (4) that

$$|\hat{F}_c(u)| \leq P_c \exp[\varphi_c(|y|) + uy].$$

If we set $u > 0$, we see that

$$|\hat{F}_c(u)| \leq P_c \exp[\inf_{y>0} (\varphi_c(y) - uy)].$$

But, it follows from the definition of φ_c that for every $\epsilon > 0$

$$\varphi_c(y) < y(\log y + \tfrac{1}{2} + \epsilon + c)$$

provided $y > y_\epsilon$. We find then that

$$|\hat{F}_c(u)| \leq N_\epsilon P_c \exp[-\exp(u - \epsilon - c - \tfrac{3}{2}],$$

where N_ϵ depends only on ϵ. Therefore,

$$|[F_{\mu,a}(x)]\hat{}\,| = \left|\hat{F}_{a/\mu}\left(\frac{u}{\mu}\right)\right| \leq N_\epsilon P_{a/\mu} \times \exp\left[-\exp\left(\frac{u}{\mu} - \epsilon - \frac{a}{\mu} - \frac{3}{2}\right)\right].$$

Let us notice that the curve

$$y = \exp[-\exp(u - d)], \qquad d > 0$$

has its inflection point (for $u > 0$) at $u = d$. One can also see that for $u \geq 2\mu$

$$|\hat{F}_{\mu,a}(u)| \leq P_{a/\mu} \exp[-\exp(u - a - 2\mu)/\mu].$$

We thus have an indication for the rate of decrease of $F_{a/\mu}(u/\mu)$ for $u \to \infty$. We see, in particular, that after a certain point it decreases extremely rapidly (more rapidly than the inverse of an exponential of an exponential) so that a well-chosen straight cutoff gives a good approximation to its behavior as long as we stay in momentum space.

IV. RESULTS AND INTERPRETATIONS

The coupled equations of Sec. III which give the minimum cutoff δ as a function of A and μ have been solved numerically in a first approximation [where only the first three factors in the infinite product giving $F_{\mu,a}(z)$ are considered] by Bourrely.[6] The results are best visualized with the following curve (Fig. 2). The physical meaning of this curve is very significative. We see that for $A > A_H$ (A_H being a constant determined once μ is given), δ/γ is, in this approximation, simply μ [a calculation using more terms in the infinite product $F_{\mu,a}(z)$ shows that $\delta/\gamma \to \mu$ as $A \to \infty$]. This is what should be expected from the uncertainty relation: indeed for a function $|F(x)| \leq A \exp(-\mu|x|)$ (A large enough), the "standard deviation" Δx is of the order $1/\mu$, and Δu is 2δ, thus the uncertainty relation, $\Delta x \cdot \Delta u \approx$ const, gives $\delta/\mu = $ const. This interval $A > A_H$ will be called the "Heisenberg interval H."

If however A is very small, the relation $|F(x)| \leq A \exp(-\mu|x|)$ together with the normalization condition $\int [F'(x)]^2\, dx = 1$ implies that $F(x)$ has many large oscillations, this is indeed the only way to have $|F'(x)|$ large enough to have the normalization condition while having $|F(x)|$ small. In this region (non H), Δx has no longer any link with the asymptotic behavior of $F(x)$ given by $A \exp(-\mu|x|)$. In fact in this region Δx has no longer any significant physical meaning and δ can only be calculated with our method.

ACKNOWLEDGMENT

We wish to thank Professor W. V. Houston for the interest he has shown in this work.

[6] For more precise numerical results see C. Bourrely, Compt. Rend. **260**, 439 (1965).

On the Quantum Field Theories Leading to the Corben Equations*

RICHARD E. NORTON
Quantum Physics Laboratory, Space Technology Laboratories, Redondo Beach, California
and
Department of Physics, University of California, Los Angeles, California
(Received 30 November 1964)

Formulations are presented for the second quantized versions of the field theories which lead to Corben's equations of motion. It is demonstrated that an indefinite metric is required to guarantee positive energies for all the particles, but that otherwise the theories are physically unambiguous. The number and properties of the resulting particles are studied and compared with the conclusions from previous work. Alternative formulations are also discussed.

I. INTRODUCTION

A SET of relativistic wave equations have recently been proposed whose solutions yield the mass spectra of free particles and their spins in the range 0, $\frac{1}{2}$, 1, $\frac{3}{2}$, or 2.[1,2] It has been argued[1,2] that the quantum theories which incorporate these equations connect in the correspondence limit to the classical, relativistic theory of spinning particles derived some time ago by Bhabha and Corben.[3] A few recent reports have been devoted to showing that the quantum mechanical version of this theory—i.e., a set of wave equations—implies a collection of particles whose masses, charges, and spins show an impressive similarity to those of the stable particles and resonances presently observed.[4-7]

In this paper the structure of the Corben theory is studied in detail. The field theories which lead to these equations are shown to allow a physically consistent second quantization, and the number and properties of the particles which emerge are reexamined from this viewpoint.

The four Corben equations, together with the spins of the associated particles, are

$$(i^{-1}\gamma_\mu \partial_\mu + m + \tfrac{1}{4}m_0\sigma_{\mu\nu}\sigma'_{\mu\nu})\psi = 0 \text{ (spin 0, 1)}, \quad (1)$$

$$(i^{-1}\gamma_\mu \partial_\mu + m + \tfrac{1}{2}m_0\sigma_{\mu\nu}\beta_{\mu\nu})\psi = 0 \text{ (spin } \tfrac{1}{2}, \tfrac{3}{2}), \quad (2)$$

$$(i^{-1}\beta_\mu \partial_\mu + m + \tfrac{1}{2}m_0\beta_{\mu\nu}\sigma_{\mu\nu})\psi = 0 \text{ (spin } \tfrac{1}{2}, \tfrac{3}{2}), \quad (3)$$

$$(i^{-1}\beta_\mu \partial_\mu + m + m_0\beta_{\mu\nu}\beta'_{\mu\nu})\psi = 0 \text{ (spin 0, 1, 2)}. \quad (4)$$

The γ_μ, γ'_μ, β_μ, β'_μ are all to be considered as acting on separate vector spaces, and hence the dimension of each ψ is the product of the dimensions of the two such operators which occur in the equation. The γ_μ and γ'_μ are four dimensional and fulfill the Dirac anticommutation rules,

$$\{\gamma_\mu, \gamma_\nu\} = \{\gamma'_\mu, \gamma'_\nu\} = -2\delta_{\mu\nu}, \quad (5)$$

whereas the β_μ and β' are either one-, five-, or ten-dimensional and satisfy the Duffin–Kemmer–Petiau relation typified by[8]

$$\beta_\mu\beta_\nu\beta_\sigma + \beta_\sigma\beta_\nu\beta_\mu = -\beta_\mu\delta_{\nu\sigma} - \beta_\sigma\delta_{\nu\mu}. \quad (6)$$

The γ_μ, γ'_μ, β_μ, β'_μ are all chosen to be anti-Hermitian, and $\sigma_{\mu\nu} = \frac{1}{2}i[\gamma_\mu, \gamma_\nu]$, $\beta_{\mu\nu} = i[\beta_\mu, \beta_\nu]$ with similar relations for the primed matrices. Each of the Eqs. (1–4) has two parameters, m and m_0. Although the best "fit" to the observed particle spectra is obtained by giving these parameters somewhat different values in the four cases, we do not make this distinction. Each of the four equations will be discussed separately, and it will always be clear from the context to which equation the parameters refer.

In Secs. II–V, we discuss in turn each of the theories which incorporate Eqs. (1)–(4). The theory corresponding to Eq. (1) is studied in Sec. II. Two alternative formulations are presented, but the first formulation is considered the more natural; all the solutions to Eq. (1) are retained, and a parity suggests itself for all the resulting particles. The charges also are determined if the electric current proposed by Corben is employed. In contrast to this approach, we then note in the second formulation that almost none of these conclusions are actually necessary if all we require is the existence of a field ψ satisfying Eq. (1). There exists the freedom, consistent with this requirement, to retain arbitrarily few of the solutions to Eq. (1) and to designate independently the charges and parities of

* Research supported partly by the Office of Naval Research, the Space Technology Laboratories Company Independent Research Program, and the National Science Foundation.

[1] H. C. Corben, Proc. Nat. Acad. Sci. **48**, 1559, 1746 (1962).
[2] H. C. Corben, Nuovo Cimento **28**, 202 (1963).
[3] H. J. Bhabha and H. C. Corben, Proc. Roy. Soc. (London) **A178**, 273 (1941).
[4] H. C. Corben, Phys. Rev. Letters **10**, 555 (1963).
[5] H. C. Corben, Phys. Rev. **131**, 2219 (1963).
[6] H. C. Corben, Phys. Rev. **134**, **B832** (1964).
[7] H. C. Corben, J. Math. Phys. **5**, 1664 (1964).
[8] N. Kemmer, Proc. Roy. Soc. (London) **A173**, 91 (1939).

the remaining particles. When this freedom is utilized, however, the field ψ assumes a rather remote role, which is difficult to understand if Eq. (1) is to be the basis of the theory.

In Secs. III–V we display the structure of the theories corresponding to Eqs. 2–4 from a viewpoint analogous to the first formulation in Sec. II. The properties of the solutions are presented in detail, and they are shown to be different, in some respects, from what was believed previously. In particular, we shall see that for each mass and spin the multiplicity of the solutions is not in agreement with the observed particle spectrum, and that it will be difficult to rule out only the unwanted solutions in a general manner. However, we emphasize that these theories could also be formulated in analogy with the second approach in Sec. II. If this were done, and the unpleasant feature referred to above were accepted, it would then be possible to retain only those solutions which can be made to correspond to observed particles. From either point of view, the spin of a solution is fixed by the theory.

II. FIELD THEORY OF EQUATION (1)

A. First Formulation

In analogy with the conventional Dirac theory, the field equation (1) suggests a Lagrangian density

$$\mathfrak{L} = \bar{\psi}(\gamma_\mu i^{-1}\partial_\mu + \mathfrak{M}_1)\psi, \tag{7}$$

with

$$\mathfrak{M}_1 = m + \tfrac{1}{4}m_0\sigma_{\mu\nu}\sigma'_{\mu\nu}, \tag{8}$$

and

$$\bar{\psi} = \psi^+\gamma_0\gamma'_0.$$

From the Lagrangian density in Eq. (7), the expressions for the energy–momentum 4-vector, and for the generalized angular momentum tensor can be derived in the standard manner. There results

$$P_i = :\int d\mathbf{x}\ \bar{\psi}\gamma_0 i^{-1}\ \partial_i\psi:, \tag{9}$$

$$H = :\int d\mathbf{x}\ \bar{\psi}(\boldsymbol{\gamma}\cdot i^{-1}\nabla + \mathfrak{M}_1)\psi:, \tag{10}$$

and

$$J_{\mu\nu} = :\int d\mathbf{x}\ \bar{\psi}\gamma_0(x_\mu i^{-1}\ \partial_\nu - x_\nu i^{-1}\ \partial_\mu + \tfrac{1}{2}\sigma_{\mu\nu} + \tfrac{1}{2}\sigma'_{\mu\nu})\psi:. \tag{11}$$

The double dots on both sides of these expressions indicate the normal ordered products obtained by moving all destruction operators to the right. From Eq. (11) we see that the spin of the particles contained in the field ψ is $\frac{1}{2}(\boldsymbol{\sigma} + \boldsymbol{\sigma}')$ and it is therefore either 0 or 1. Complying with the TCP theorem,[9] we postulate the commutation rules

$$[\psi(x),\ \bar{\psi}(y)\gamma_0]_{x_0=y_0} = \delta(\mathbf{x} - \mathbf{y}). \tag{12}$$

To see more clearly the decomposition of the field ψ into its normal modes (i.e., particles), we describe the theory in momentum space. Once this is accomplished, it will also be easy to check that all the particles yield a positive contribution to the energy in Eq. (10). We write

$$\psi(x) = \frac{1}{(2\pi)^3}\int d^4p\ \psi(p)e^{ip\cdot x}, \tag{13}$$

and observe from Eq. (1) that $\psi(p)$ satisfies

$$(\gamma\cdot p + \mathfrak{M}_1)\psi(p) = 0, \tag{14}$$

which, for zero 3-momentum, reduces to

$$\gamma_0\mathfrak{M}_1\psi(0, W) = W\psi(0, W). \tag{15}$$

If we multiply Eq. (15) with the matrix $\gamma_0\mathfrak{M}_1$, we obtain

$$W^2\psi(0, W) = m(m + m_0\boldsymbol{\sigma}\cdot\boldsymbol{\sigma}')\psi(0, W), \tag{16}$$

so that [1,6]

$$W^2 = m(m - 3m_0), \quad (\boldsymbol{\sigma}\cdot\boldsymbol{\sigma}' = -3, \quad \text{spin } 0) \tag{17a}$$

$$W^2 = m(m + m_0), \quad (\boldsymbol{\sigma}\cdot\boldsymbol{\sigma}' = \ 1, \quad \text{spin } 1) \tag{17b}$$

represents the mass spectrum.

In case (17a), there are four linearly independent states corresponding to the four combinations of the signs of W and of the eigenvalue (either $+1$ or -1) of $\gamma_0\gamma'_0$. In case (17b) there are twelve such states which reflect the same alternatives separately for each of the three orientations of the total spin. If we let α indicate the mass and spin orientation and display the sign of $\gamma_0\gamma'_0$ explicitly, we can write

$$\psi(0, W) = \sum_\alpha \{[a_+(\alpha)u_+(\alpha) + a_-(\alpha)u_-(\alpha)]\theta(W) + [b_+^\dagger(\alpha)v_+(\alpha) + b_-^\dagger(\alpha)v_-(\alpha)]\theta(-W)\}\,\delta(W^2 - m_\alpha^2), \tag{18}$$

where $\theta(W)$ is one for W positive and zero otherwise. The m_α^2 designates either of the two solutions in Eq. (17), and the $a_\pm(\alpha)$ and $b_\pm(\alpha)$ are destruction operators in the Hilbert space for the corresponding particle and antiparticle states.

In terms of $u_\pm(\alpha)$ and $v_\pm(\alpha)$, which describe the particles and antiparticles at rest, we can construct the solutions for arbitrary momenta by applying

[9] G. Luders, Ann. Phys. (N. Y.) 2, 1 (1957).

the appropriate Lorentz transformations. That is, defining

$$u_\pm(\mathbf{p}, \alpha) \equiv \exp\,[\tfrac{1}{2}\hat{p}_i(\sigma_{i4} + \sigma'_{i4})\theta_\alpha(p)]u_\pm(\alpha), \quad (19a)$$

$$v_\pm(\mathbf{p}, \alpha) \equiv \exp\,[\tfrac{1}{2}\hat{p}_i(\sigma_{i4} + \sigma'_{i4})\theta_\alpha(p)]v_\pm(\alpha), \quad (19b)$$

where $\theta_\alpha(p) \equiv \sinh^{-1} p/m_\alpha$ and $\hat{p} = \mathbf{p}\,|\mathbf{p}|^{-1}$, there follows

$$\begin{aligned}(\gamma\cdot p_\alpha + \mathfrak{M}_1)u_\pm(\mathbf{p}, \alpha) &= 0\\ (-\gamma\cdot\mathbf{p}_\alpha + \mathfrak{M}_1)v_\pm(\mathbf{p}, \alpha) &= 0,\end{aligned} \quad (20)$$

where $p_\alpha = (\mathbf{p}, [\mathbf{p}^2 + m_\alpha^2]^{\frac{1}{2}})$. We can now write

$$\begin{aligned}\psi(p, p_0) = \sum_\alpha \{&[a_+(\mathbf{p}, \alpha)u_+(\mathbf{p}, \alpha)\\ &+ a_-(\mathbf{p}, \alpha)u_-(\mathbf{p}, \alpha)]\theta(p_0) + [b_+^\dagger(-\mathbf{p}, \alpha)v_+(-\mathbf{p}, \alpha)\\ &+ b_-^\dagger(-\mathbf{p}, \alpha)v_-(-\mathbf{p}, \alpha)]\theta(-p_0)\}\,\delta(p^2 - m_\alpha^2),\end{aligned} \quad (21)$$

and by substituting this expression into Eq. (13)

$$\begin{aligned}\psi(x) = \frac{1}{(2\pi)^3}\sum_\alpha \int \frac{d\mathbf{p}}{2\omega_\alpha(p)} \{&[a_+(\mathbf{p}, \alpha)u_+(\mathbf{p}, \alpha)\\ &+ a_-(\mathbf{p}, \alpha)u_-(\mathbf{p}, \alpha)]e^{ip\cdot x}\\ &+ [b_+^\dagger(\mathbf{p}, \alpha)v_+(\mathbf{p}, \alpha) + b_-^\dagger(\mathbf{p}, \alpha)v_-(\mathbf{p}, \alpha)]e^{-ip\cdot x}\}.\end{aligned} \quad (22)$$

where $\omega_\alpha(p) = (\mathbf{p}^2 + m_\alpha^2)^{\frac{1}{2}}$.

The orthogonality properties of the $u_\pm(p, \alpha)$ and $v_\pm(p, \alpha)$ are derived in the Appendix. With a convenient normalization, they can be expressed as

$$\begin{aligned}u_-^\dagger(\mathbf{p}, \alpha)\gamma'_0 u_-(\mathbf{p}, \alpha') &= 0,\\ u_+^\dagger(\mathbf{p}, \alpha)\gamma'_0 v_\pm(-\mathbf{p}, \alpha') &= 0,\\ v_-^\dagger(\mathbf{p}, \alpha)\gamma'_0 v_-(\mathbf{p}, \alpha') &= 0,\\ u_-^\dagger(\mathbf{p}, \alpha)\gamma'_0 v_\pm(-\mathbf{p}, \alpha') &= 0,\end{aligned} \quad (23a)$$

and

$$\begin{aligned}u_\pm^\dagger(\mathbf{p}, \alpha)\gamma'_0 u_\pm(\mathbf{p}, \alpha') &= \pm 2\omega_\alpha(p)\delta_{\alpha\alpha'},\\ v_\pm^\dagger(\mathbf{p}, \alpha)\gamma'_0 v_\pm(\mathbf{p}, \alpha') &= \mp 2\omega_\alpha(p)\delta_{\alpha\alpha'}.\end{aligned} \quad (23b)$$

These relations then allow Eq. (22) to be inverted

$$a_\pm(\mathbf{p}, \alpha) = \pm\int d\mathbf{x}\, e^{-ipx}u_\pm^\dagger(\mathbf{p}, \alpha)\gamma'_0\psi(x) \quad (24a)$$

$$b_\pm(\mathbf{p}, \alpha) = \mp\int d\mathbf{x}\, \psi^\dagger(x)\gamma'_0 v_\pm(\mathbf{p}, \alpha)e^{-ipx}, \quad (24b)$$

so that from Eq. (12) we have the commutation rules

$$\begin{aligned}[a_\pm(\mathbf{p}, \alpha), a_\pm^\dagger(\mathbf{p}', \alpha')] &= [b_\pm(\mathbf{p}, \alpha), b_\pm^\dagger(\mathbf{p}', \alpha')]\\ &= \pm 2\omega_\alpha(p)(2\pi)^3\delta_{\alpha\alpha'}\delta(\mathbf{p} - \mathbf{p}').\end{aligned} \quad (25)$$

All other commutators are zero. We note in particular that the operators for the modes with the negative signature (sign of $\gamma_0\gamma'_0$) demonstrate the "wrong" sign in their commutation rules. Before we discuss this feature, let us look at the expression for the energy in Eq. (10) when expressed in terms of these elementary creation and destruction operators. By substituting Eq. (22) into Eq. (10) and employing the relations (23), we obtain

$$\begin{aligned}H = \frac{1}{(2\pi)^3}\sum_\alpha \int \frac{d\mathbf{p}}{2\omega_\alpha(p)} &[a_+^\dagger(\mathbf{p}, \alpha)a_+(\mathbf{p}, \alpha)\\ &+ b_+^\dagger(\mathbf{p}, \alpha)b_+(\mathbf{p}, \alpha) - a_-^\dagger(\mathbf{p}, \alpha)a_-(\mathbf{p}, \alpha)\\ &- b_-^\dagger(\mathbf{p}, \alpha)b_-(\mathbf{p}, \alpha)]\omega_\alpha(p).\end{aligned} \quad (26)$$

We observe that in both Eqs. (25) and (26) the terms involving the modes with the negative signature appear with a sign opposite to what is conventional.[10] In both instances this situation would be corrected if we could simply interpret $a_-^\dagger(\mathbf{p}, \alpha)$ and $b_-^\dagger(\mathbf{p}, \alpha)$ as not the Hermitian adjoints of $a_-(\mathbf{p}, \alpha)$ and $b_-(\mathbf{p}, \alpha)$, but the negative of the Hermitian adjoints.[11] Formally this can be realized if we think of the theory as quantized with an indefinite metric. That is, if the usual Hermitian adjoint is denoted by an asterisk (*), we can define

$$\psi^\dagger = \eta\psi^*\eta, \quad (27)$$

where the metric η is

$$\begin{aligned}\eta = \eta^\dagger = \eta^{-1} = \exp\Big\{&\frac{\pi i}{(2\pi)^3}\sum_\alpha \int \frac{d\mathbf{p}}{2\omega_\alpha(p)}\\ &\cdot[a_-^\dagger(p, \alpha)a_-(p, \alpha) + b_-^\dagger(p, \alpha)b_-(p, \alpha)]\Big\}.\end{aligned} \quad (28)$$

Let us think of the theory as quantized with this indefinite metric, but realize that it is simply a formal device for reinterpreting the † adjoint. For the positive signature modes, the † adjoint and the Hermitian adjoint are the same.

With this reinterpretation, the theory defined by the Lagrangian density (7) is characterized by a positive definite energy, and with a positive norm for all states. As a free-field theory, it is therefore physically consistent.

We now discuss the charges and parities of the particles. We have seen that there are eight particles in the theory, four with spin 0 and four with spin 1. Equation (17) shows that the masses of the particles depend only on their spin and that the spin-1 particles are the more massive. As has been suggested,[6]

[10] Similar problems have arisen in other theories in which one wave equation describes more than one particle. A. Pais and G. E. Uhlenbeck, Phys. Rev. 79, 145 (1950).

[11] To conserve probability, any extension of the theory to include interactions would have to be carefully constructed to guarantee that the interaction Hamiltonian was Hermitian [in the * sense].

the parameters m and m_0 can be adjusted to fit the masses of the spin-0 and spin-1 particles to the observed masses of the K and K^* mesons.

The electric current which has been proposed for the theory of Eq. (1) is

$$j_\mu = e :\bar{\psi}\gamma_\mu\psi:, \tag{29}$$

and the operator for the charge is therefore

$$Q = e : \int d\mathbf{x}\ \bar{\psi}\gamma_0\psi. \tag{30}$$

In order to make clear the charges carried by the various particles, we express Eq. (30) explicitly in terms of the elementary creation and destruction operators. By decomposing the fields ψ and $\bar{\psi}$ in Eq. (30) according to Eq. (22), and making use of the relations (23), we obtain

$$\begin{aligned} Q = \frac{e}{(2\pi)^3} \sum_\alpha \int \frac{d\mathbf{p}}{2\omega_\alpha(p)} \\ \cdot[a_+^\dagger(p,\alpha)a_+(p,\alpha) - b_+^\dagger(p,\alpha)b_+(p,\alpha) \\ - a_-^\dagger(p,\alpha)a_-(p,\alpha) + b_-^\dagger(p,\alpha)b_-(p,\alpha)]. \end{aligned} \tag{31}$$

The charges carried by all the particles are now apparent, if we compare Eq. (31) with Eq. (26) and remember that all the modes must contribute a positive energy. The particles destroyed by the operators $a_\pm(p,\ \alpha)$ carry the charge e, whereas those destroyed by the $b_\pm(p,\ \alpha)$ have the opposite charge. In particular, none of the particles are electrically neutral [if the electric current is given by (29)], and thus it is inconsistent to conclude that the theory (as formulated here) describes the K *and* K^* mesons.

If space inversion (i.e., parity) is to be represented by a simple transformation of the field ψ, the natural choice which suggests itself is

$$P : \psi(\mathbf{x},\ t) \rightarrow \gamma_0\gamma_0'\psi(-\mathbf{x},\ t). \tag{32}$$

From Eq. (24), it follows from this definition that

$$\begin{aligned} P : a_\pm(\mathbf{p},\ \alpha) &\rightarrow \pm a_\pm(-\mathbf{p},\ \alpha) \\ b_\pm(\mathbf{p},\ \alpha) &\rightarrow \pm b_\pm(-\mathbf{p},\ \alpha), \end{aligned} \tag{33}$$

and the eight particles divide into two scalars (these with positive signature), two pseudoscalars, two vectors (with negative signature), and two pseudovectors. As before, these conclusions do not allow the theory, as it stands, to describe the K and K^*.

B. Alternative Formulation of the Theory of Equation (1)

We now construct the theory described by Eq. (1) in a manner which will make more clear the connection between the formulation in part A and the conventional field theory of particles with spin 0 and 1. We consider Eq. (1) as the requisite feature of the theory and see what freedom is allowed.

Let us write the field ψ as $\psi_{\alpha\beta}$, a four-by-four matrix, where the unprimed Dirac operators act on the first subscript and the primed operators act on the second. It can readily be verified, after considerable algebra, that if we define

$$\begin{aligned} \psi(x) \equiv \frac{1}{4}\Big[&(\tfrac{1}{2}m)^{\frac{1}{2}}\phi(x) + (\tfrac{1}{2}m)^{\frac{1}{2}}\theta(x)\gamma_5 \\ &- [\tfrac{1}{2}(m+m_0)]^{\frac{1}{2}}\Big[V_\mu(x) - \frac{i}{m}\,\partial_\mu\phi(x)\Big]\gamma_\mu \\ &+ [\tfrac{1}{2}(m+m_0)]^{\frac{1}{2}}\Big[A_\mu(x) + \frac{i}{m}\,\partial_\mu\theta(x)\Big]\gamma_\mu\gamma_5 \\ &- [8(m+m_0)]^{-\frac{1}{2}}[\partial_\mu V_\nu(x) - \partial_\nu V_\mu(x)] \\ &+ \epsilon_{\mu\nu\alpha\beta}\,\partial_\alpha A_\beta(x)\sigma_{\mu\nu}\Big], \end{aligned} \tag{34}$$

so that

$$\phi(x) = (2/m)^{\frac{1}{2}}\ \mathrm{Tr}\ [\psi\sigma_2] \tag{35a}$$

$$\theta(x) = (2/m)^{\frac{1}{2}}\ \mathrm{Tr}\ [\psi\sigma_2\gamma_5] \tag{35b}$$

$$V_\mu(x) = [2/(m+m_0)]^{\frac{1}{2}}\ \mathrm{Tr}\ [\psi\sigma_2\gamma_\mu\gamma_5 + i\,\partial_\mu\psi\sigma_2/m] \tag{35c}$$

$$A_\mu(x) = [2/(m+m_0)]^{\frac{1}{2}}\ \mathrm{Tr}\ [\psi\sigma_2\gamma_\mu\gamma_5 - i\,\partial_\mu\psi\sigma_2\gamma_5/m], \tag{35d}$$

where Tr means trace, then, to within four divergences, the Lagrangian in Eq. (7) is equal to

$$\begin{aligned} \mathfrak{L} = &-\tfrac{1}{2}[\partial_\mu\phi^\dagger\,\partial_\mu\phi + m(m-3m_0)\phi^\dagger\phi] \\ &+ \tfrac{1}{2}[\partial_\mu\theta^\dagger\,\partial_\mu\theta + m(m-3m_0)\theta^\dagger\theta] \\ &+ [\tfrac{1}{4}(\partial_\mu V_\nu^\dagger - \partial_\nu V_\mu^\dagger)(\partial_\mu V_\nu - \partial_\nu V_\mu) \\ &+ \tfrac{1}{2}m(m+m_0)V_\mu^\dagger V_\mu] \\ &- [\tfrac{1}{4}(\partial_\mu A_\nu^\dagger - \partial_\nu A_\mu^\dagger)(\partial_\mu A_\nu - \partial_\nu A_\mu) \\ &+ \tfrac{1}{2}m(m+m_0)A_\mu^\dagger A_\mu]. \end{aligned} \tag{36}$$

Except for the signs[10] of the terms involving the fields θ and V_μ, this is a conventional Lagrangian for four uncoupled fields, two of spin 0 and two of spin 1, whose masses agree with Eq. (17).

An alternative point of view can now be adopted which still allows Eq. (1) to be satisfied. We take the Lagrangian as given by Eq. (36), but with the terms involving θ and V_μ changed in sign. These changes of sign have the same effect as the introduction of the indefinite metric in Sec. IIA. The equations of motion are not altered, and hence the field

ψ as defined in Eq. (34) still satisfies Eq. (1). In fact, if all we require is Eq. (1), we can ignore in Eq. (36) the terms involving as many of the fields ϕ, θ, etc. as we please, providing that ψ in Eq. (34) is expressed only in terms of the retained fields. It is also clear that the charges and parities of the remaining particles can be fixed arbitrarily. The electric current and the parity transformations would then not necessarily be given by Eq. (1). In particular, it is possible to assign the charges and the parities of the eight particles described by Eq. (36) to conform to those of the K and K^* mesons.

III. FIELD THEORY OF EQUATION (2)

Here we formulate the field theory of Eq. (2) in analogy with Sec. IIA. The Lagrangian is given by

$$\mathfrak{L} = -\bar{\psi}(\gamma_\mu i^{-1}\partial_\mu + \mathfrak{M}_2)\psi, \tag{37}$$

with

$$\mathfrak{M}_2 \equiv m + \tfrac{1}{2}m_0\sigma_{\mu\nu}\beta_{\mu\nu}$$
$$\bar{\psi} = \psi^\dagger\gamma_0\eta_4 \tag{38}$$

and

$$\eta_\mu = -1 - 2\beta_\mu^2 \qquad (\mu = 1, 2, 3, 4). \tag{39}$$

The operators $\sigma_{\mu\nu}$ and $\beta_{\mu\nu}$ have been defined in Sec. I. From Eq. (37) we obtain the expressions for the energy–momentum 4-vector and for the relativistic angular momentum tensor

$$P_i = :\int d\mathbf{x}\bar{\psi}\gamma_0\frac{1}{i}\partial_i\psi: \tag{40}$$

$$H = :\int d\mathbf{x}\bar{\psi}\left(\boldsymbol{\gamma}\cdot\frac{1}{i}\nabla + \mathfrak{M}_2\right)\psi: \tag{41}$$

$$J_{\mu\nu} = :\int d\mathbf{x}\bar{\psi}\gamma_0\left(\mathbf{x}_\mu\frac{1}{i}\partial_\nu - \mathbf{x}_\nu\frac{1}{i}\partial_\mu + \tfrac{1}{2}\sigma_{\mu\nu} + \beta_{\mu\nu}\right)\psi:. \tag{42}$$

If, in analogy with the conventional definition of $\mathfrak{d}$, we define

$$\Sigma_i = \epsilon_{ijk}\beta_{jk} \quad \text{(no sum)}, \tag{43}$$

then the particle spins are given by

$$\mathbf{S} = \tfrac{1}{2}\mathfrak{d} + \boldsymbol{\Sigma}. \tag{44}$$

From Eqs. (6) and (43), it follows that $\Sigma^2(\Sigma^2-2)=0$, and hence the Σ spin can be either 0 or 1. The total spin in Eq. (44) is therefore either $\frac{1}{2}$ or $\frac{3}{2}$, and we thus choose the anticommutation rule[9]

$$\{\psi(x), \bar{\psi}(y)\gamma_0\}_{x_0=y_0} = \delta(\mathbf{x}-\mathbf{y}). \tag{45}$$

We again describe the theory in momentum space in order to check that all of the particles yield a positive contribution to the energy, and to observe more clearly the features of the implied particles. For the $\mathbf{p} = 0$ Fourier components, we have in analogy with Eq. (15)

$$\gamma_0\mathfrak{M}_2\psi(0, W) = W\psi(0, W). \tag{46}$$

The solutions to Eq. (46) are listed in Table I of Ref. 6. As in Sec. II, we denote the positive and negative W solutions of Eq. (46) by $u_\pm(\alpha)$ and $v_\pm(\alpha)$. Here α refers to the choice of mass $|W|$, and to the choice of spin and its orientation—both of which can be diagonalized simultaneously with the operator $\gamma_0\mathfrak{M}_2$ in Eq. (46). The $\pm$ sign refers to the eigenvalue of $\gamma_0\eta_4$. For each mass and spin, there are solutions corresponding to both the eigenvalues $+1$ and -1 of this operator.

As in Sec. II, the wavefunctions $u_\pm(p, \alpha)$ and $v_\pm(p, \alpha)$ for finite 3-momentum can be constructed from the $u_\pm(\alpha)$ and $v_\pm(\alpha)$ by Lorentz transformation,

$$u_\pm(p, \alpha) = \exp\left[\hat{p}_i(\tfrac{1}{2}\sigma_{i4} + \beta_{i4})\theta_\alpha(p)\right]u_\pm(\alpha) \tag{47a}$$

$$v_\pm(p, \alpha) = \exp\left[\hat{p}_i(\tfrac{1}{2}\sigma_{i4} + \beta_{i4})\theta_\alpha(p)\right]v_\pm(\alpha), \tag{47b}$$

and hence

$$(\boldsymbol{\gamma}\cdot p_\alpha + \mathfrak{M}_2)u_\pm(\mathbf{p}, \alpha) = 0$$
$$(-\boldsymbol{\gamma}\cdot p_\alpha + \mathfrak{M}_2)v_\pm(\mathbf{p}, \alpha) = 0. \tag{48}$$

The unit vector $\hat{p}$ and the angle $\theta_\alpha(p)$ are defined after Eq. (19). We could now write the general solution of Eq. (2) in the form of Eq. (22), except that here the sum on α would refer to the masses and spin appropriate to Eq. (2).

From the Appendix, we have the orthonormality relations

$$u_+^\dagger(\mathbf{p}, \alpha)\eta_4 u_-(\mathbf{p}, \alpha') = 0, \quad u_+^\dagger(\mathbf{p}, \alpha)\eta_4 v_\pm(-\mathbf{p}, \alpha') = 0,$$
$$v_+^\dagger(\mathbf{p}, \alpha)\eta_4 v_-(\mathbf{p}, \alpha') = 0, \quad u_-^\dagger(\mathbf{p}, \alpha)\eta_4 v_\pm(-\mathbf{p}, \alpha') = 0. \tag{49a}$$

and

$$u_\pm^\dagger(\mathbf{p}, \alpha)\eta_4 u_\pm(\mathbf{p}, \alpha') = \pm 2\omega_\alpha(p)\delta_{\alpha\alpha'}$$
$$v_\pm^\dagger(\mathbf{p}, \alpha)\eta_4 v_\pm(\mathbf{p}, \alpha') = \mp 2\omega_\alpha(p)\delta_{\alpha\alpha'}. \tag{49b}$$

The equivalent of Eq. (22) can now be inverted to yield

$$a_\pm(\mathbf{p}, \alpha) = \pm\int d\mathbf{x}e^{-ip\cdot x}u_\pm^\dagger(\mathbf{p}, \alpha)\eta_4\psi(\mathbf{x}) \tag{50a}$$

$$b_\pm(\mathbf{p}, \alpha) = \mp\int d\mathbf{x}\psi^\dagger(x)\eta_4 v_\pm(\mathbf{p}, \alpha)e^{-ip\cdot x}, \tag{50b}$$

so that from Eqs. (45), (49), and (50)

$$\{a_\pm(\mathbf{p},\alpha), a_\pm^\dagger(\mathbf{p}',\alpha')\} = \{b_\mp(\mathbf{p},\alpha), b_\mp^\dagger(\mathbf{p}',\alpha')\} = \pm 2\omega_\alpha(p)(2\pi)^3\delta_{\alpha\alpha'}\delta(\mathbf{p}-\mathbf{p}'). \quad (51)$$

All other pairs of these operators anticommute. Observe, that due to the anticommutation rules, the relationship between the sign of the anticommutor and the signature of $\gamma_0\eta_4$ differs for the operators $b_\pm(\mathbf{p},\alpha)$ and $b_\pm^\dagger(\mathbf{p},\alpha)$ from the corresponding expression in Eq. (25).

Next, we employ the equivalent of Eq. (22) and Eqs. (41), (48), and (49) to obtain

$$H = \frac{1}{(2\pi)^3}\sum_\alpha \int \frac{d\mathbf{p}}{2\omega_\alpha(p)} \times [a_+^\dagger(\mathbf{p},\alpha)a_+(\mathbf{p},\alpha) + b_-^\dagger(\mathbf{p},\alpha)b_-(\mathbf{p},\alpha) - a_-^\dagger(\mathbf{p},\alpha)a_-(\mathbf{p},\alpha) - b_+^\dagger(\mathbf{p},\alpha)b_+(\mathbf{p},\alpha)]\omega_\alpha(p). \quad (52)$$

We note in this case that the terms involving $a_-(\mathbf{p},\alpha)$ and $b_+(\mathbf{p},\alpha)$ occur with the wrong sign in Eqs. (51) and (52). As in Sec. II, we resolve this difficulty by reinterpreting the † adjoint of these operators to be the negative of the * Hermitian adjoint.[10,11] Formally, this can be accomplished by employing the definition (27) with η given by

$$\eta = \exp\left\{\frac{\pi i}{(2\pi)^3}\sum_\alpha \int \frac{d\mathbf{p}}{2\omega_\alpha(p)}[a_-^\dagger(\mathbf{p},\alpha)a_-(\mathbf{p},\alpha) + b_+^\dagger(\mathbf{p},\alpha)b_+(\mathbf{p},\alpha)]\right\}. \quad (53)$$

Once this reinterpretation is understood, all the particles yield a positive contribution to the energy in Eq. (52). To this extent the theory is physically admissible.

The electric current proposed by Corben[6] is

$$j_\mu = \pm\tfrac{1}{2}e :\bar\psi\gamma_\mu(1+\eta_5)\psi:, \quad (54)$$

where $\eta_5 = \eta_1\eta_2\eta_3\eta_4$, and the $\pm$ sign depends only upon the representation of the β_μ. Each of the solutions to Eq. (46), and therefore the $u_\pm(\mathbf{p},\alpha)$ and $v_\pm(\mathbf{p},\alpha)$ defined in Eq. (47) are eigenstates of η_5. We denote by $\delta_\alpha = \pm 1, 0$ the eigenvalue of $\pm\frac{1}{2}(1+\eta_5)$ and write the operator for the total change,

$$Q = \pm\tfrac{1}{2}e :\int d\mathbf{x}\bar\psi\gamma_0(1+\eta_5)\psi: \quad (55)$$

in the form

$$Q = \frac{e}{(2\pi)^3}\sum_\alpha \delta_\alpha \int \frac{d\mathbf{p}}{2\omega_\alpha(p)} \times [a_+^\dagger(\mathbf{p},\alpha)a_+(\mathbf{p},\alpha) - b_-^\dagger(\mathbf{p},\alpha)b_-(\mathbf{p},\alpha) - a_-^\dagger(\mathbf{p},\alpha)a_-(\mathbf{p},\alpha) + b_+^\dagger(\mathbf{p},\alpha)b_+(\mathbf{p},\alpha)]. \quad (56)$$

By comparing Eq. (56) with Eq. (52), and remembering that all the modes yield a positive contribution to the energy, it is evident that the $a_\pm(\mathbf{p},\alpha)$ destroy particles of charge $e\delta_\alpha$ and the $b_\pm(\mathbf{p},\alpha)$ destroy particles of charge $-e\delta_\alpha$.

We have already mentioned that all the solutions are composed of doublets, both members of which are characterized by the same mass and spin but by different eigenvalues (either $+1$ or -1) of $\gamma_0\eta_4$. This situation is analogous to the occurrence of both signs $\gamma_0\gamma_0'$ in the theory discussed in Sec. II. In that case, the natural interpretation was that every particle had its counterpart differing only in parity. A similar interpretation suggests itself here. Under space inversion, the field ψ would then transform as

$$P:\psi(\mathbf{x},t) \to \gamma_0\eta_4\psi(-\mathbf{x},t). \quad (57)$$

It has been suggested that this theory describes the nucleons, the Ξ particles, and the N_{13}^* pion–nucleon resonances.[6] This definition of the parity would suggest, that for every one of these particles, there should exist another with the same mass and spin but with opposite parity. Since such counterparts apparently do not exist, we must conclude that half of the solutions are unaccounted for physically.

Finally, let us emphasize that the theory discussed in this Section can be formulated in a manner analogous to Part B in Sec. II. It is thus possible to satisfy Eq. (2), retaining only half its solutions and adjusting arbitrarily the charges and parities of the retained particles to conform to the nucleons, the Ξ baryons, and the N_{13}^* resonances.

IV. FIELD THEORY OF EQUATION (3)

The field Eq. (3) follows from the Lagrangian

$$\mathcal{L} = -\bar\psi(\beta_\mu i^{-1}\partial_\mu + \mathfrak{M}_3)\psi \quad (58)$$

with

$$\mathfrak{M}_3 = m + \tfrac{1}{2}m_0\beta_{\mu\nu}\sigma_{\mu\nu} \quad (59)$$

and $\bar\psi$ the same as in Eq. (38). The energy–momentum 4-vector is

$$P_i = :\int d\mathbf{x}\bar\psi\beta_0\frac{1}{i}\partial_i\psi: \quad (60)$$

$$H = :\int d\mathbf{x}\bar\psi\left(\boldsymbol{\beta}\cdot\frac{1}{i}\nabla + \mathfrak{M}_3\right)\psi:, \quad (61)$$

and the relativistic angular momentum tensor is

$$J_{\mu\nu} = :\int d\mathbf{x}\bar\psi\beta_0\left(\mathbf{x}_\mu\frac{1}{i}\partial_\nu - \mathbf{x}_\nu\frac{1}{i}\partial_\mu + \beta_{\mu\nu} + \tfrac{1}{2}\sigma_{\mu\nu}\right)\psi:. \quad (62)$$

As in Sec. III, the spin matrices are given by Eq. (44), and the theory describes particles of spin $\frac{1}{2}$ and $\frac{3}{2}$.

In Eq. (3) a new feature appears. The coefficient of the time derivative is the singular matrix β_0, and as a consequence, the solutions of this equation do not constitute a complete set. At each instant, the quantity $(\boldsymbol{\beta}\cdot i^{-1}\nabla + \mathfrak{M}_3)\psi$ must be orthogonal to the subspace belonging to the null eigenvalue of β_0. That is, it must be an eigenstate of β_0^2 corresponding to the eigenvalue $+1$. The anticommutation rules satisfied by ψ and $\bar{\psi}$ must be constructed to allow arbitrary variations of only those fields which satisfy this constraint. We therefore require

$$\{(\boldsymbol{\beta}\cdot i^{-1}\nabla + \mathfrak{M}_3)\psi(x),\ \bar{\psi}(y)\beta_0\}_{x_0=y_0} = \beta_0^2(\boldsymbol{\beta}\cdot i^{-1}\nabla + \mathfrak{M}_3)\,\delta(\mathbf{x}-\mathbf{y}). \tag{63}$$

In analogy to our procedure in the two previous theories, we first look at the Fourier transform of Eq. (3) restricted to zero 3-momentum,

$$\mathfrak{M}_3\psi(0, W) = W\beta_0\psi(0, W). \tag{64}$$

The solutions of Eq. (64) are distinguished by the representation of the β_μ and by the eigenvalues of W, the spin and its orientation, and $\gamma_0\eta_4$. There are no solutions in the 1×1 representation of the β_μ.

In the 5×5 representation, there are solutions corresponding to four particles of spin $\frac{1}{2}$ and with mass $m_\alpha = |W|$ given by [6]

$$W = \pm m[(1+b)(1-3b)/(1-2b)]^{\frac{1}{2}}, \tag{65}$$

where $b = m_0/m$. These solutions differ in the sign of W and in the eigenvalue (either $+1$ or -1) of $\gamma_0\eta_4$. In this representation of the β_μ, the solutions to Eq. (64) or Eq. (3) occupy only eight of the 20 dimensions in the direct-product space of the β_μ and γ_μ.

In the 10×10 representation of the β_μ, 24 of the 40 dimensions are spanned by solutions of Eq. (3). Eight dimensions correspond to solutions with spin $\frac{1}{2}$ and 16 with spin $\frac{3}{2}$. The eigenvalues of W[6] are

$$W = \pm m[(1+b)(1-3b)(1-4b)/(1-2b)]^{\frac{1}{2}} \quad (\text{spin } \tfrac{1}{2}), \tag{66}$$

and

$$W = \pm m(1+2b)^{\frac{1}{2}} \quad (\text{spin } \tfrac{3}{2}). \tag{67}$$

Both signs of the eigenvalue of $\gamma_0\eta_4$ occur in this case also.

We denote the positive (or negative) W solutions of Eq. (40) by $u_\pm(\alpha)$ [or $v_\pm(\alpha)$] where the $\pm$ sign indicates the eigenvalue of $\gamma_0\eta_4$ and the symbol α denotes the other distinguishing features of the solutions. The wavefunctions for arbitrary three momentum can be constructed from these $u_\pm(\alpha)$ and $v_\pm(\alpha)$ in accordance with Eq. (47), and the field ψ can be expanded in the form of Eq. (22). From the orthogonality rules discussed in the Appendix,

$$u_+^\dagger(\mathbf{p},\alpha)\gamma_0\beta_0u_-(\mathbf{p},\alpha') = 0,\quad u_+^\dagger(\mathbf{p},\alpha)\gamma_0\beta_0v_\pm(-\mathbf{p},\alpha') = 0$$
$$v_+^\dagger(\mathbf{p},\alpha)\gamma_0\beta_0v_-(\mathbf{p},\alpha') = 0,\quad u_-^\dagger(\mathbf{p},\alpha)\gamma_0\beta_0v_\pm(-\mathbf{p},\alpha') = 0 \tag{68a}$$

and

$$u_+^\dagger(\mathbf{p},\alpha)\gamma_0\beta_0u_\pm(\mathbf{p},\alpha') = \pm 2\omega_\alpha(p)\,\delta_{\alpha\alpha'},$$
$$v_\pm^\dagger(\mathbf{p},\alpha)\gamma_0\beta_0v_\pm(\mathbf{p},\alpha') = \mp 2\omega_\alpha(p)\,\delta_{\alpha\alpha'}, \tag{68b}$$

we obtain, by inverting the equivalent of Eq. (22),

$$a_\pm(\mathbf{p},\alpha) = \pm\int d\mathbf{x}\, e^{-ipx}u_\pm^\dagger(\mathbf{p},\alpha)\gamma_0\beta_0\psi(x) \tag{69a}$$

$$b_\pm(\mathbf{p},\alpha) = \mp\int d\mathbf{x}\, \psi^\dagger(x)\gamma_0\beta_0v_\pm(\mathbf{p},\alpha)e^{-ip\cdot x}. \tag{69b}$$

From Eqs. (63), (68), and (69), and by integrating by parts and making use of the equations

$$(\boldsymbol{\beta}\cdot p_\alpha + \mathfrak{M}_3)u_\pm(\mathbf{p},\alpha) = 0,$$
$$(-\boldsymbol{\beta}\cdot p_\alpha + \mathfrak{M}_3)v_\pm(\mathbf{p},\alpha) = 0, \tag{70}$$

we obtain

$$\{a_\pm(\mathbf{p},\alpha),\ a_\pm^\dagger(\mathbf{p}',\alpha')\} = \{b_\mp(\mathbf{p},\alpha),\ b_\mp^\dagger(\mathbf{p}',\alpha')\} = \pm 2\omega\alpha(p)(2\pi)^3\delta_{\alpha\alpha'}\,\delta(\mathbf{p}-\mathbf{p}'). \tag{71}$$

All other combinations anticommute. Finally, Eq. (61) and the equivalent of Eq. (22) allow us to write the total energy as

$$H = \frac{1}{(2\pi)^3}\sum_\alpha\int\frac{d\mathbf{p}}{2\omega_\alpha(p)} \times [a_+^\dagger(\mathbf{p},\alpha)a_+(\mathbf{p},\alpha) + b_-^\dagger(\mathbf{p},\alpha)b_-(\mathbf{p},\alpha) - a_-^\dagger(\mathbf{p},\alpha)a_-(\mathbf{p},\alpha) - b_+^\dagger(\mathbf{p},\alpha)b_+(\mathbf{p},\alpha)]\omega_\alpha(p). \tag{72}$$

Equations (71) and (72) are the same as Eqs. (51) and (52) in Sec. III. The need for reinterpretation of the $\dagger$ adjoint and the method for accomplishing this with the metric in Eq. (53) are applicable here exactly as in Sec. III. All the particles then give positive contributions to the energy in Eq. (72).

The total baryonic current proposed for this theory is

$$b_\mu = :\bar{\psi}\beta_\mu\psi:, \tag{73}$$

and hence the baryon number B is

$$B = :\int d\mathbf{x}\bar{\psi}\beta_0\psi:. \tag{74}$$

By substituting the equivalent of Eq. (22) into Eq. (74), and employing the relations (68), we obtain

$$B = \frac{1}{(2\pi)^3}\sum\int\frac{d\mathbf{p}}{2\omega_\alpha(p)} \times [a_+^\dagger(\mathbf{p},\alpha)a_+(\mathbf{p},\alpha) - b_-^\dagger(\mathbf{p},\alpha)b_-(\mathbf{p},\alpha) - a_-^\dagger(\mathbf{p},\alpha)a_-(\mathbf{p},\alpha) + b_+^\dagger(\mathbf{p},\alpha)b_+(\mathbf{p},\alpha)]. \tag{75}$$

Comparison of Eq. (75) with Eq. (72) reveals that the particles destroyed by the $a_\pm(\mathbf{p},\alpha)$ carry positive baryonic number, whereas those destroyed by the $b_\pm(\mathbf{p},\alpha)$ carry negative baryonic number.

The electric current which has been suggested[6] for this theory is $j_\mu = e\,:\bar{\psi}\gamma_5\beta_\mu\psi:$. Although it is conserved, this form is not a plausible candidate in the second quantized version for two reasons: (1) If, as seems most natural, the field $\psi(x,t) \rightarrow \gamma_0\eta_4\psi(-x,t)$ under parity, then this current is a pseudovector. Of course this definition of space inversion is not required; the theory discussed here could also be formulated in analogy with Sec. IB, and this apparent problem could then be avoided. However, (2), the "charge" $Q = e\,:\int d\mathbf{x}\bar{\psi}\gamma_5\beta_0\psi:$ is not diagonal in the one-particle states of this theory. The particles corresponding to the operators $a_\pm(\mathbf{p},\alpha)$ and $b_\pm(\mathbf{p},\alpha)$ would not carry definite amounts of charge, and this feature is of course physically inadmissible.[12]

It has been suggested that the theory of Eq. (3) describes the Λ, the Y_0^* (1405 MeV, spin $\frac{1}{2}$), and the Y_{03}^* (1520 MeV, spin $\frac{3}{2}$). Although the neutrality of all the particles was concluded from the above definition of the electric current, nothing stops us from retaining this conclusion but rejecting the current.

Once the neutrality of all the particles is adopted, it becomes possible to assert that three of the six particles contained in the theory are those mentioned. The particles corresponding to the other solutions are identical to those in mass and spin, but apparently do not occur physically.

[12] We should note, however, that it is possible to redefine the one-particle states as linear combinations of the two degenerate modes at each mass in such a way that the wavefunctions for the new states are eigenfunctions of γ_5. Since the eigenvalues of γ_5 are ± 1, all the particles would then be electrically charged, and it would not be possible to associate this theory with only neutral particles as has been suggested.

V. FIELD THEORY OF EQUATION (4)

The Lagrangian density is

$$\mathfrak{L} = -\bar{\psi}(\beta_\mu i^{-1}\partial_\mu + \mathfrak{M}_4)\psi, \tag{76}$$

with

$$\mathfrak{M}_4 = m + m_0\beta_{\mu\nu}\beta'_{\mu\nu}, \tag{77}$$

and

$$\bar{\psi} = \psi^\dagger\eta_4\eta'_4. \tag{78}$$

The energy–momentum 4-vector takes the form

$$P_i = :\int d\mathbf{x}\bar{\psi}\beta_0\frac{1}{i}\partial_i\psi: \tag{79}$$

$$H = :\int d\mathbf{x}\bar{\psi}\left(\boldsymbol{\beta}\cdot\frac{1}{i}\nabla + \mathfrak{M}_3\right)\psi: \tag{80}$$

and

$$J_{\mu\nu} = :\int d\mathbf{x}\bar{\psi}\beta_0\left(x_\mu\frac{1}{i}\partial_\nu - x_\nu\frac{1}{i}\partial_\mu + \beta_{\mu\nu} + \beta'_{\mu\nu}\right)\psi:. \tag{81}$$

We note again that the matrix β_0 is singular and that the solutions to Eq. (4) do not constitute a complete set. We can repeat here the arguments leading to Eq. (63), with the exception that since the spin is integral we postulate commutation rules,[9]

$$[(\boldsymbol{\beta}\cdot i^{-1}\nabla + \mathfrak{M}_4)\psi(\mathbf{x}),\ \bar{\psi}(\mathbf{y})\beta_0]_{x_0=y_0} = \beta_0^2(\boldsymbol{\beta}\cdot i^{-1}\nabla + \mathfrak{M}_4)\delta(\mathbf{x}-\mathbf{y}). \tag{82}$$

The Fourier transform of Eq. (4), restricted to zero 3-momentum, is

$$\mathfrak{M}_4\psi(0,W) = W\beta_0\psi(0,W). \tag{83}$$

The solutions of Eq. (83) are distinguished by the representations of β_μ and β'_μ and by the eigenvalues of W, of the spin, and of $\eta_4\eta'_4$ (either $+1$ or -1). All of these possibilities occur and have been tabulated extensively in Refs. 4, 6, and 7.

If the sign of the eigenvalue of $\eta_4\eta'_4$ is indicated explicitly, we can write the positive and negative W solutions of Eq. (83) as $u_\pm(\alpha)$ and $v_\pm(\alpha)$. We can then construct the wavefunctions for arbitrary 3-momentum,

$$u_\pm(\mathbf{p},\alpha) = \exp[\hat{p}_i(\beta_{i4} + \beta'_{i4})\{_\alpha(p)]u_\pm(\alpha) \tag{84a}$$

$$v_\pm(\mathbf{p},\alpha) = \exp[\hat{p}_i(\beta_{i4} + \beta'_{i4})\theta_\alpha(p)]v_\pm(\alpha), \tag{84b}$$

where $\hat{p}$ and $\theta_\alpha(p)$ have the same meaning as stated after Eq. (19). It follows that

$$(\boldsymbol{\beta}\cdot p_\alpha + \mathfrak{M}_4)u_\pm(\mathbf{p},\alpha) = 0,$$
$$(-\boldsymbol{\beta}\cdot p_\alpha + \mathfrak{M}_4)v_\pm(\mathbf{p},\alpha) = 0. \tag{85}$$

As discussed in the Appendix, the orthogonality relations are

$$u_+^\dagger(\mathbf{p},\alpha)\eta_4'\beta_0 u_-(\mathbf{p},\alpha') = 0, \quad u_+^\dagger(\mathbf{p},\alpha)\eta_4'\beta_0 v_\pm(-\mathbf{p},\alpha') = 0,$$
$$v_+^\dagger(\mathbf{p},\alpha)\eta_4'\beta_0 v_-(\mathbf{p},\alpha') = 0, \quad u_-^\dagger(\mathbf{p},\alpha)\eta_4'\beta_0 v_\pm(-\mathbf{p},\alpha') = 0, \tag{86a}$$

and

$$u_\pm^\dagger(\mathbf{p},\alpha)\eta_4'\beta_0 u_\pm(\mathbf{p},\alpha') = \pm 2\omega_\alpha(p)\delta_{\alpha\alpha'},$$
$$v_\pm^\dagger(\mathbf{p},\alpha)\eta_4'\beta_0 v_\pm(\mathbf{p},\alpha') = \mp 2\omega_\alpha(p)\delta_{\alpha\alpha'}. \tag{86b}$$

When the field ψ is decomposed in the form of Eq. (22), the relations (86) can be employed to obtain

$$a_\pm(\mathbf{p},\alpha) = \pm\int d\mathbf{x}\, e^{-ip\cdot x} u_\pm^\dagger(\mathbf{p},\alpha)\eta_4'\beta_0\psi(\mathbf{x}) \tag{87a}$$

$$b_\pm(\mathbf{p},\alpha) = \mp\int d\mathbf{x}\, \psi^\dagger(x)\eta_4'\beta_0 v_\pm(\mathbf{p},\alpha)e^{-ip\cdot x}, \tag{87b}$$

from which, making use also of Eqs. (82), (85), and (86),

$$[a_\pm(\mathbf{p},\alpha), a_\pm^\dagger(\mathbf{p}',\alpha')] = [b_\pm(\mathbf{p},\alpha), b_\pm^\dagger(\mathbf{p}',\alpha')]$$
$$= \pm 2\omega_\alpha(p)(2\pi)^3\delta_{\alpha\alpha'}\delta(\mathbf{p}-\mathbf{p}'). \tag{88}$$

Finally, by substituting the equivalent of (22) into Eq. (80), the total energy can be written as

$$H = \frac{1}{(2\pi)^3}\sum_\alpha \int \frac{d\mathbf{p}}{2\omega_\alpha(p)}$$
$$\times [a_+^\dagger(\mathbf{p},\alpha)a_+(\mathbf{p},\alpha) + b_+^\dagger(\mathbf{p},\alpha)b_+(\mathbf{p},a)$$
$$- a_-^\dagger(\mathbf{p},\alpha)a_-(\mathbf{p},\alpha) - b_-^\dagger(\mathbf{p},\alpha)b_-(\mathbf{p},\alpha)]\omega_\alpha(p), \tag{89}$$

We observe that Eqs. (88) and (89) are the same as Eqs. (25) and (26). The theories of Eqs. (1) and (4) should hence, formally, be quantized in the same manner. The discussion after Eq. (26) and the form of the metric in Eq. (28) apply to both theories.

The electric current which has been proposed[4,6,7] for this theory is

$$j_\mu = \tfrac{1}{2}e :\bar\psi\beta_\mu(1+\eta_5')\psi:. \tag{90}$$

All the $u_\pm(\mathbf{p},\alpha)$ and $v_\pm(\mathbf{p},\alpha)$ are eigenstates of η_5' with the eigenvalue $+1$ or -1, and for a given mass and spin, the four particles corresponding to positive and negative W solutions, and to the two signs of $\eta_4\eta_4'$, all have the same sign of this eigenvalue. From Eq. (90), by decomposing the fields ψ and $\bar\psi$ according to Eq. (22) and employing the orthogonality relations (86), we can write the charge Q as

$$Q = \frac{e}{(2\pi)^3}\sum_\alpha{}' \int \frac{d\mathbf{p}}{2\omega_\alpha(p)}$$
$$\times [a_+^\dagger(\mathbf{p},\alpha)a_+(\mathbf{p},\alpha) - b_+^\dagger(\mathbf{p},\alpha)b_+(\mathbf{p},\alpha)$$
$$- a_-^\dagger(\mathbf{p},\alpha)a_-(\mathbf{p},\alpha) + b_-^\dagger(\mathbf{p},\alpha)b_-(\mathbf{p},\alpha)], \tag{91}$$

where the primed sum includes only those terms with the eigenvalue $+1$ of η_5'. Comparing Eqs. (91) and (89), we see that of the particles included in the primed sum, those destroyed by the $a_\pm(\mathbf{p},\alpha)$ have the charge e whereas those destroyed by the $b_\pm(\mathbf{p},\alpha)$ carry charge $-e$.

Table II of Ref. 7 displays the essential features of the particle spectrum arising from Eq. (4). It has been emphasized there that the two parameters m and m_0 can be adjusted to fit the masses and spins to those of a large number of the observed strangeness zero bosons. The multiplicity of the solutions, however, does not agree with experiment. The theory actually contains four particles for each mass and spin; two corresponding to both eigenvalues of $\eta_4\eta_4'$, and another factor of 2 because of both positive- and negative-frequency solutions.

The experimental situation requires that for charged particles there should be two degenerate solutions corresponding to both signs of the charge. The theory thus contains twice as many particles as are actually observed. For neutral particles no degeneracy is required, and hence there are four times as many solutions as can be accounted for physically. Exceptions to this latter conclusion, reducing the factor of 4 to a factor of 2, occur in two cases when the solutions of this theory are related to the observed particles as in Ref. 7. There, the ω and ρ_0 correspond to the two degenerate, positive-frequency solutions which differ only in their eigenvalue of $\eta_4\eta_4'$. The two negative-frequency solutions are hence the only two with this mass and spin which are unaccounted for physically. Similar remarks apply for the f and B_0 resonances.

Finally, let us note that if in analogy to the previous theories [see, for example, Eq. (32)], space inversion is given by

$$P\colon \psi(\mathbf{x},t) \to \eta_4\eta_4'\psi(-\mathbf{x},t), \tag{92}$$

then the ρ_0 and ω would have opposite parities, as would also the f and B_0. Since these conclusions are in contradiction to experiment (at least for the ρ and ω) the definition of parity in Eq. (92) is incompatible with the interpretation of these solutions in Ref. 7.

VI. SUMMARY

We have seen that the field theories which lead to Corben's equations of motion can be made physically consistent, but that the multiplicity of the solutions imply considerably more particles than have been observed. It is possible to adjust the parameters m and m_0 to fit the masses and spins

to experiment surprisingly well. On the other hand, the number of particles predicted at each such set of values does not agree with the experimental situation.

Our viewpoint has been to take seriously the field theories which lead to Corben's equations of motion and to examine in detail the number and properties of the particles which result. Let us emphasize again that this is not the only possible approach. We could, for example, simply demand that the Corben equations be satisfied without retaining all the solutions. In Sec. IIB, we saw how this procedure could be formalized for the theory of Eq. (1). It is clear that similar methods could be applied to the other equations. We should keep in mind also, that in addition to the four equations discussed here, the Corben point of view actually suggests many more equations reflecting the fact that the term involving m_0 can couple together arbitrary spin matrices.[13] If the theories corresponding to these additional equations are studied, most of the solutions will describe particles of high spin ($s > 2$) and with masses above the present experimental observations. There will be some solutions, however, which refer to particles with spins and masses in the range considered in this paper. It seems likely that when these additional solutions are taken into account, a reinterpretation of the particles predicted by Eqs. (1)–(4) will be suggested, and it is possible that the multiplicity of the solutions existing then will suggest a simple, plausible scheme for ruling out those which do not fit in the observed particle spectrum.

ACKNOWLEDGMENTS

The author expresses his gratitude to the Physics Division of the Aspen Institute for Humanistic Studies for its hospitality during part of the summer of 1964. Although most of the work represented in this paper was done while the author was a consultant at the Space Technology Laboratories, the actual writing was done at Aspen. He would also like to thank Dr. Herbert C. Corben for many helpful discussions.

APPENDIX

In this Appendix we indicate the derivation of the orthogonality relations employed in Eqs. (23), (49), (68), and (86), which refer, respectively, to the theories of Eqs. (1)–(4).

A. Equation (23)

It follows from the definition of $u_\pm(\alpha)$ and $v_\pm(\alpha)$ in Sec. II that

$$\gamma_0 m_\alpha u_\pm(\alpha) = \mathfrak{M}_1 u_\pm(\alpha), \qquad \gamma_0 m_\alpha v_\pm(\alpha) = -\mathfrak{M}_1 v_\pm(\alpha)$$
$$\gamma_0' m_\alpha u_\pm(\alpha) = \pm\mathfrak{M}_1 u_\pm(\alpha), \qquad \gamma_0' m_\alpha v_\pm(\alpha) = \mp\mathfrak{M}_1 v_\pm(\alpha). \tag{A1}$$

If we apply the exponential operator in Eq. (19b) to the relations in (A1) we obtain after some algebra

$$\begin{aligned}(\gamma\cdot p_\alpha + \mathfrak{M}_1)u_\pm(\mathbf{p}, \alpha) &= 0,\\ (-\gamma\cdot p_\alpha + \mathfrak{M}_1)v_\pm(\mathbf{p}, \alpha) &= 0,\\ (\pm\gamma'\cdot p_\alpha + \mathfrak{M}_1)u_\pm(\mathbf{p}, \alpha) &= 0,\\ (\mp\gamma'\cdot p_\alpha + \mathfrak{M}_1)v_\pm(\mathbf{p}, \alpha) &= 0.\end{aligned} \tag{A2}$$

To prove the orthogonality relation involving $u_+^\dagger(\mathbf{p}, \alpha)$ and $u_-(\mathbf{p}, \alpha)$, consider

$$\begin{aligned}u_+^\dagger(\mathbf{p}, \alpha)(-\gamma'\cdot\mathbf{p} + \mathfrak{M}_1)u_-(\mathbf{p}, \alpha') &= -\omega_{\alpha'}(p)u_+^\dagger(\mathbf{p}, \alpha)\gamma_0' u_-(\mathbf{p}, \alpha')\\ &= \omega_\alpha(p)u_+^\dagger(\mathbf{p}, \alpha)\gamma_0' u_-(\mathbf{p}, \alpha'),\end{aligned} \tag{A3}$$

where both equalities arise from (A3); the first expression from applying $(-\gamma'\cdot\mathbf{p} + \mathfrak{M}_1)$ to the right, and the lower expression form applying it to the left. Clearly, both equalities can only be valid if $u_+^\dagger(\mathbf{p}, \alpha)\gamma_0' u_-(\mathbf{p}, \alpha') = 0$, which is the first of the relations in Eq. (23). The corresponding relation for the negative frequency solutions, $v_+^\dagger(\mathbf{p}, \alpha)\gamma_0' v_-(\mathbf{p}, \alpha') = 0$, can be proved similarly.

Equation (A2) can also easily be employed to prove the orthogonality relations involving $u_+^\dagger(\mathbf{p}, \alpha)$ and $v_\pm(-\mathbf{p}, \alpha')$. Consider

$$\begin{aligned}u_+^\dagger(\mathbf{p}, \alpha)\gamma_0\gamma_0'(\gamma\cdot\mathbf{p} + \mathfrak{M}_1)v_\pm(-\mathbf{p}, \alpha') &= -\omega_{\alpha'}(p)u_+^\dagger(\mathbf{p}, \alpha)\gamma_0' v_\pm(-\mathbf{p}, \alpha')\\ &= \omega_\alpha(p)u_+^\dagger(\mathbf{p}, \alpha)\gamma_0' v_\pm(-\mathbf{p}, \alpha'),\end{aligned} \tag{A4}$$

where again the first equality arises from applying $(\gamma\cdot\mathbf{p} + \mathfrak{M}_1)$ to the right and the second from applying it to the left. It follows obviously that $u_+^\dagger(\mathbf{p}, \alpha)\gamma_0' v_\pm(-\mathbf{p}, \alpha') = 0$, which is another of the relations in Eq. (23a). The final orthogonality relation in Eq. (23a) can be proved similarly.

To show that $u_\pm^\dagger(\mathbf{p}, \alpha)\gamma\, u_\pm(\mathbf{p}, \alpha') \propto \delta_{\alpha\alpha'}$, we consider

$$\begin{aligned}u_\pm^\dagger(\mathbf{p}, \alpha)\gamma_0\gamma_0'(\gamma\cdot\mathbf{p} + \mathfrak{M}_1)u_\pm(\mathbf{p}, \alpha') &= \omega_{\alpha'}(p)u_\pm^\dagger(\mathbf{p}, \alpha)\gamma_0' u_\pm(\mathbf{p}, \alpha')\\ &= \omega_\alpha(p)u_\pm^\dagger(\mathbf{p}, \alpha)\gamma_0' u_\pm(\mathbf{p}, \alpha'),\end{aligned} \tag{A5}$$

[13] See, for example, L. Castell, University College (London) preprint.

where again the upper and lower terms on the right come from acting with $\boldsymbol{\gamma}\cdot\mathbf{p} + \mathfrak{M}_1$, to the right or left, respectively. Except for the factor $\pm 2\omega_\alpha(p)$, the first version of Eq. (23b) follows from comparing the two forms of (A5). That $v_\pm^\dagger(\mathbf{p}, \alpha)\gamma_0' v_\pm(\mathbf{p}, \alpha') \propto \delta_{\alpha\alpha'}$ can be shown analogously.

Finally we must verify the $\pm$ signs appearing in Eq. (23b). Take, for example,

$$u_\pm^\dagger(\mathbf{p}, \alpha)\gamma_0' u_\pm(\mathbf{p}, \alpha) = u_\pm^\dagger(\alpha) \exp\left[\tfrac{1}{2}\hat{p}_i(\sigma_{i4} + \sigma_{i4}')\theta_\alpha(p)\right]\gamma_0' \times \exp\left[\tfrac{1}{2}\hat{p}_i(\sigma_{i4} + \sigma_{i4}')\theta_\alpha(p)\right]u_\pm(\alpha), \quad \text{(A6)}$$

where we have made use of Eq. (19). Taking into account the definition of $\theta_\alpha(p)$ after Eq. (19), Eq. (A6) can be rewritten as

$$\begin{aligned} &u_\pm^\dagger(\mathbf{p}, \alpha)\gamma_0' u_\pm(\mathbf{p}, \alpha) \\ &\quad = u_+^\dagger(\alpha)\gamma_0'[\omega_\alpha(p) + \sigma_{i4}p_i]m_\alpha^{-1}u_\pm(\alpha) \\ &\quad = \omega_\alpha(p)m_\alpha^{-1}u_\pm^\dagger(\alpha)\gamma_0' u_\pm(\alpha) \\ &\quad = \pm\omega_\alpha(p)m_\alpha^{-2}u_\pm^\dagger(\alpha)\mathfrak{M}_1 u_\pm(\alpha). \end{aligned} \quad \text{(A7)}$$

The second version of (A7) follows from the first since $\{\sigma_{i4}, \gamma_0\gamma_0'\} = 0$ and hence σ_{i4} does not connect two states belonging to the same eigenvalue of $\gamma_0\gamma_0'$. The third version follows from the second by (A1). The possibility of choosing the normalization of the $u_\pm(\mathbf{p}, \alpha)$ according to the $\pm$ sign in the first form of (23b) is now apparent since $\mathfrak{M}_1 = m + \frac{1}{4}m_0\sigma_{\mu\nu}\sigma_{\mu\nu}'$ is a positive definite operator (for the values of m and m_0 actually employed). This positive definiteness can be easily proved (it is obviously true for small enough m_0), but we do not discuss the proof here. It turns out that $\mathfrak{M}_1$ is positive unless m_0 is large enough to produce complex m_α in Eq. (A1).

B. Equation (49)

To demonstrate the orthogonality relations involving $u_+(\mathbf{p}, \alpha)$ and $v_\pm(-\mathbf{p}, \alpha')$ we make use of Eq. (48) to write

$$\begin{aligned} &u_+^\dagger(\mathbf{p}, \alpha)\gamma_0\eta_4(\boldsymbol{\gamma}\cdot\mathbf{p} + \mathfrak{M}_2)v_\pm(-\mathbf{p}, \alpha') \\ &\quad = -\omega_{\alpha'}(p)u_+^\dagger(\mathbf{p}, \alpha)\eta_4 v_\pm(-\mathbf{p}, \alpha') \\ &\quad = \omega_\alpha(p)u_+^\dagger(\mathbf{p}, \alpha)\eta_4 v_\pm(-\mathbf{p}, \alpha'), \end{aligned} \quad \text{(A8)}$$

where the first and second equalities arise from acting with $(\boldsymbol{\gamma}\cdot\mathbf{p} + \mathfrak{M}_2)$ on the right and left, respectively. The desired relation follows immediately, and the corresponding expression involving $u_-^\dagger(\mathbf{p}, \alpha)$ is proved in an identical manner.

The proof of Eq. (49b) follows in complete analogy with that of Eq. (23b). Instead of (A5) we have

$$\begin{aligned} &u_\pm^\dagger(\mathbf{p}, \alpha)\gamma_0\eta_4(\boldsymbol{\gamma}\cdot\mathbf{p} + \mathfrak{M}_2)u_\pm(\mathbf{p}, \alpha') \\ &\quad = \omega_{\alpha'}(p)u_\pm^\dagger(\mathbf{p}, \alpha)\eta_4 u_\pm(\mathbf{p}, \alpha') \\ &\quad = \omega_\alpha(p)u_\pm^\dagger(\mathbf{p}, \alpha)\eta_4 u_\pm(\mathbf{p}, \alpha'), \end{aligned} \quad \text{(A9)}$$

which demonstrates that the left side of Eq. (49b) is proportional to $\delta_{\alpha\alpha'}$. To show that the $u_\pm(\mathbf{p}, \alpha)$ can be normalized according to Eq. (49b), consider

$$u_\pm^\dagger(\mathbf{p}, \alpha)\eta_4 u_\pm(\mathbf{p}, \alpha) = u_\pm^\dagger(\alpha) \exp\left[\hat{p}_i(\tfrac{1}{2}\sigma_{i4}\beta_{i4})\theta_\alpha(p)\right]\eta_4 \times \exp\left[\hat{p}_i(\tfrac{1}{2}\sigma_{i4} + \beta_{i4})\theta_\alpha(p)\right]u_\pm(\alpha) \quad \text{(A10)}$$

in analogy with (A6). If γ_0' in Eq. (A7) is replaced by η_4, the various versions of (A7) follow identically, and we obtain

$$u_\pm^\dagger(\mathbf{p}, \alpha)\eta_4 u_\pm(\mathbf{p}, \alpha) = \pm\omega_\alpha(p)m_\alpha^{-2}u_\pm^\dagger(\alpha)\mathfrak{M}_2 u_\pm(\alpha). \quad \text{(A11)}$$

The positive definiteness of the operator $\mathfrak{M}_2$ (which can easily be proved) then implies our result. The corresponding expression involving $v_\pm(\mathbf{p}, \alpha)$ can be demonstrated similarly.

Finally we must show that $u_+^\dagger(\mathbf{p}, \alpha)\eta_4 u_-(\mathbf{p}, \alpha) = 0$, and also that $v_+^\dagger(\mathbf{p}, \alpha)\eta_4 v_-(\mathbf{p}, \alpha) = 0$. We note that these relations are trivially true for all the solutions except for the pair corresponding to the proton and the pair corresponding to the N_{13}^{*+} in Table I of Ref. 6, since the two members of every other pair occur in different representations of the β_μ. Consider the matrix element of η_4 written in the form of (A10). Since η_4 anticommutes with β_{i4}, this expression can be written as

$$\begin{aligned} &u_+^\dagger(\mathbf{p}, \alpha)\eta_4 u_-(\mathbf{p}, \alpha) \\ &\quad = u_+^\dagger(\alpha)\eta_4[\omega_\alpha(p) + \sigma_{i4}p_i]m_\alpha^{-1}u_-(\alpha). \end{aligned} \quad \text{(A12)}$$

The term involving $\omega_\alpha(p)$ does not contribute, since it commutes with $\gamma_0\eta_4$ and hence doesn't connect eigenstates of $\gamma_0\eta_4$ belonging to different eigenvalues. Further, the η_4 can be replaced by γ_0 [since $\eta_4^2 = 1$ and $u_+^\dagger(\alpha) = u_+^\dagger(\alpha)\gamma_0\eta_4$] to yield

$$\begin{aligned} &u_+^\dagger(\mathbf{p}, \alpha)\eta_4 u_-(\mathbf{p}, \alpha) \\ &\quad = p_i m_\alpha^{-1}u_+^\dagger(\alpha)\gamma_0\sigma_{i4}u_-(\alpha) \\ &\quad = p_i(2m_\alpha)^{-1}u_+^\dagger(\alpha)[\gamma_0, \sigma_{i4}]u_-(\alpha). \end{aligned} \quad \text{(A13)}$$

We now use the fact that the $u_\pm(\alpha)$ are solutions of Eq. (46) with $W = m_\alpha$ to rewrite (A13) as

$$\begin{aligned} &u_+^\dagger(\mathbf{p}, \alpha)\eta_4 u_-(\mathbf{p}, \alpha) \\ &\quad = p_i(2m_\alpha^2)^{-1}u_+^\dagger(\alpha)[\mathfrak{M}_2, \sigma_{i4}]u_-(\alpha). \end{aligned} \quad \text{(A14)}$$

If we define $\Lambda_i \equiv \beta_{i4}$, $\tau_i \equiv \sigma_{i4}$, and employ $\boldsymbol{\Sigma}_i$ defined in Eq. (43) to express

$$\mathfrak{M}_2 = m + \tfrac{1}{2}m_0\sigma_{\mu\nu}\beta_{\mu\nu} = m + m_0(\boldsymbol{\delta}\cdot\boldsymbol{\Sigma} + \boldsymbol{\tau}\cdot\boldsymbol{\Lambda}), \quad \text{(A15)}$$

we can make use of the commutation rules

$$\begin{aligned} &[\tau_i, \tau_j] = i\epsilon_{ijk}\sigma_k, \quad [\Sigma_i, \Sigma_j] = [\Lambda_i, \Lambda_j] = i\epsilon_{ijk}\Sigma_k, \\ &[\tau_i, \sigma_j] = i\epsilon_{ijk}\tau_k, \quad [\Lambda_i, \Sigma_j] = i\epsilon_{ijk}\Lambda_k, \end{aligned} \quad \text{(A16)}$$

to rewrite (A14) as

$$\begin{aligned} &u_+^\dagger(\mathbf{p}, \alpha)\eta_4 u_-(\mathbf{p}, \alpha) = m_0 p_i(2m_\alpha^2)^{-1}u_+^\dagger(\alpha) \\ &\quad \times \{[\boldsymbol{\delta}\cdot\boldsymbol{\Sigma}, \tau_i] + [\boldsymbol{\tau}\cdot\boldsymbol{\Lambda}, \tau_i]\}u_-(\alpha). \end{aligned} \quad \text{(A17)}$$

From (A16) we can readily verify that

$$[\boldsymbol{\tau}\cdot\boldsymbol{\Lambda}, \tau_i] = -[\boldsymbol{\vartheta}\cdot\boldsymbol{\Sigma}, \Lambda_i].$$

Therefore,

$$u_+^\dagger(\mathbf{p}, \alpha)\eta_4 u_-(\mathbf{p}, \alpha) = m_0 p_i (2m_\alpha^2)^{-1} u_+^\dagger(\alpha) \times [\boldsymbol{\vartheta}\cdot\boldsymbol{\Sigma}, (\tau_i - \Lambda_i)]u_-(\alpha). \quad \text{(A18)}$$

Now it turns out by explicitly looking at the solutions of Eq. (46), that for the two pairs of solutions for which the left side of (A18) is not trivially zero (the pairs for the proton and for the N_{13}^{*+} in Table I of Ref. 6), the solutions are also eigenstates of the operator $\boldsymbol{\vartheta}\cdot\boldsymbol{\Sigma}$. The commutator in (A18) does not therefore contribute to the matrix element.

C. Equation (68)

The orthogonality relations involving the $u_\pm(\mathbf{p}, \alpha)$ and the $v_\pm(-\mathbf{p}, \alpha)$ are proved easily as in the two previous cases. We observe, for example that

$$\begin{aligned} u_+^\dagger(\mathbf{p}, \alpha)\gamma_0\eta_4(\boldsymbol{\beta}\cdot\mathbf{p} + \mathfrak{M}_3)v_\pm(-\mathbf{p}, \alpha') &= -\omega_{\alpha'}(p)u_+^\dagger(\mathbf{p}, \alpha)\gamma_0\beta_0 v_\pm(-\mathbf{p}, \alpha') \\ &= \omega_\alpha(p)u_+^\dagger(\mathbf{p}, \alpha)\gamma_0\beta_0 v_\pm(-\mathbf{p}, \alpha'), \end{aligned} \quad \text{(A19)}$$

where again the first and second forms on the right follow from acting with $(\boldsymbol{\beta}\cdot\mathbf{p} + \mathfrak{M}_3)$ on the $v_\pm$ and on the u_+, respectively. As before, the desired orthogonality relation is proved by equating the two equivalent forms of (A19).

To prove that $u_\pm^\dagger(\mathbf{p}, \alpha)\gamma_0\beta_0 u_\pm(\mathbf{p}, \alpha)$, or the corresponding expressions involving $v_\pm(\mathbf{p}, \alpha)$, is proportional to $\delta_{\alpha\alpha'}$, merely repeat the discussion of (A9) but with $(\boldsymbol{\gamma}\cdot\mathbf{p} + \mathfrak{M}_2)$ replaced by $(\boldsymbol{\beta}\cdot\mathbf{p} + \mathfrak{M}_3)$. The proof that the $\pm$ signs in Eq. (68b) can be chosen as shown is also similar to our procedure in the two previous cases. Since $\eta_4\beta_0 = \beta_0$, we can write

$$\begin{aligned} &u_\pm^\dagger(\mathbf{p}, \alpha)\gamma_0\beta_0 u_\pm(\mathbf{p}, \alpha) \\ &= u_\pm^\dagger(\alpha)\gamma_0\eta_4 \exp[-\hat{p}_i(\tfrac{1}{2}\sigma_{i4} + \beta_{i4})\theta_\alpha(p)]\beta_0 \\ &\quad \times \exp[\hat{p}_i(\tfrac{1}{2}\sigma_{i4} + \beta_{i4})\theta_\alpha(p)]u_\pm(\mathbf{p}, \alpha), \end{aligned} \quad \text{(A20)}$$

where we have used the fact that $\gamma_0\eta_4$ anticommutes with both σ_{i4} and β_{i4}. Evaluating the product of β_0 and the two exponentials in (A20), and replacing $\gamma_0\eta_4$ by its eigenvalues ± 1, we obtain

$$\begin{aligned} &u_\pm^\dagger(\mathbf{p}, \alpha)\gamma_0\beta_0 u_\pm(\mathbf{p}, \alpha) \\ &\quad = \pm\omega_\alpha^{-1}(p)u_\pm^\dagger(\alpha)[\boldsymbol{\beta}\cdot\mathbf{p} + \beta_0\omega_\alpha(p)]u_\pm(\alpha). \end{aligned} \quad \text{(A21)}$$

Only the second term on the right-hand side contributes, since β_i anticommutes with $\gamma_0\eta_4$ and therefore can't connect two eigenstates of this operator belonging to the same eigenvalue. The second term can be rewritten by realizing that $u_\pm(\alpha)$ is a solution of Eq. (64) with $W = m_\alpha$. We obtain

$$u_\pm^\dagger(\mathbf{p}, \alpha)\gamma_0\beta_0 u_\pm(\mathbf{p}, \alpha) = \pm m_\alpha^{-1} u_\pm^\dagger(\alpha)\mathfrak{M}_3 u_\pm(\alpha), \quad \text{(A22)}$$

and the positive definiteness of $\mathfrak{M}_3$ (which we again do not prove in detail) assures us that the choice of signs in the first form of Eq. (68b) is correct. The second form can be verified similarly.

Finally, we must prove the orthogonality relation involving the $u_+(\mathbf{p}, \alpha)$ and $u_-(\mathbf{p}, \alpha)$. Equation (A21) also is valid if the $\pm$ sign on one of the $u_\pm(\alpha)$ are reversed. In that case only the first term on the right of (A21) contributes (since $[\gamma_0\eta_4, \beta_0] = 0$), and we have

$$u_\pm^\dagger(\mathbf{p}, \alpha)\gamma_0\beta_0 u_\mp(\mathbf{p}, \alpha) = \pm u_\pm^\dagger(\alpha)\beta_i u_\mp(\alpha)p_i\omega_\alpha^{-1}(p). \quad \text{(A23)}$$

We have not been able to find a simple proof that the right-hand side of (A22) vanishes. We have, however, verified that it is zero by calculating straightforwardly with the solutions to Eq. (64).

D. Equation (86)

We prove first the orthogonality relations involving the $u_\pm(\mathbf{p}, \alpha)$ and the $v_\pm(-\mathbf{p}, \alpha)$. We note

$$\begin{aligned} u_+^\dagger(\mathbf{p}, \alpha)\eta_4\eta_4'(\boldsymbol{\beta}\cdot\mathbf{p} + \mathfrak{M}_4)v_\pm(-\mathbf{p}, \alpha') &= -\omega_{\alpha'}(p)u_+^\dagger(\mathbf{p}, \alpha)\eta_4'\beta_0 v_\pm(-\mathbf{p}, \alpha') \\ &= \omega_\alpha(p)u_+^\dagger(\mathbf{p}, \alpha)\eta_4'\beta_0 v_\pm(-\mathbf{p}, \alpha'), \end{aligned} \quad \text{(A24)}$$

where, as before, the first term results from acting with $(\boldsymbol{\beta}\cdot\mathbf{p} + \mathfrak{M}_4)$ to the right, and the second from acting with this operator to the left. Clearly, both forms are compatible only if $u_+^\dagger(\mathbf{p}, \alpha)\eta_4'\beta_0 v_\pm(-\mathbf{p}, \alpha') = 0$. It is also clear that a similar relation holds with u_+ replaced by u_-.

The proof of Eq. (86b) is essentially identical to our previous procedures. In (A9) replace γ_0 by η_4, $\boldsymbol{\gamma}\cdot\mathbf{p}$ by $\boldsymbol{\beta}\cdot\mathbf{p}$, and $\mathfrak{M}_2$ by $\mathfrak{M}_4$. We then have immediately that the left sides of (86b) are proportional to $\delta_{\alpha\alpha'}$. The $\pm$ signs in (86b) are also verified as before. If we replace in (A20), (A21), and (A22) γ_0 by η_4', $\frac{1}{2}\sigma_{i4}$ by β_{i4}', and $\mathfrak{M}_3$ by $\mathfrak{M}_4$, the signs in (86b) follow as a result of the positive definiteness of $\mathfrak{M}_4$.

The proof of the orthogonality relations involving $u_+(\mathbf{p}, \alpha)$ and $u_-(\mathbf{p}, \alpha)$ proceeds as in the previous case, i.e., Eq. (68). We can readily obtain that

$$u_\pm(\mathbf{p}, \alpha)\eta_4'\beta_0 u_\mp(\mathbf{p}, \alpha) = \pm u_\pm^\dagger(\alpha)\beta_i u_\mp(\alpha)p_i. \quad \text{(A25)}$$

Again, we have not been able to construct a simple proof that the right side of (A25) is zero, but that this is indeed true can be verified by looking directly at the solutions of Eq. (83).

Momentum Transfer Cross-Section Theorem*

E. GERJUOY†
Joint Institute for Laboratory Astrophysics, Boulder, Colorado
(Received 1 October 1964)

The momentum transfer cross section is expressed in terms of a matrix element of grad V, where V is the potential which need not be spherically symmetric. The result may be useful for estimating the momentum transfer cross section in circumstances where the usual expansion in partial waves is inconvenient or inapplicable as, e.g., when V is noncentral.

FIRST, I consider the scattering of a particle by a potential, with real Hamiltonian

$$H = T + V = (-\hbar^2/2m)\nabla^2 + V. \tag{1}$$

It is well known that the scattering solutions ψ to

$$(H - E)\psi = 0 \tag{2}$$

obey a cross-section theorem

$$\sigma = \int d\mathbf{n}' \, |A(\mathbf{n} \to \mathbf{n}')|^2 = \frac{4\pi}{k} \operatorname{Im} A(\mathbf{n} \to \mathbf{n}), \tag{3}$$

where σ is the total cross section, and $A(\mathbf{n} \to \mathbf{n}')$ is the amplitude for scattering from initial direction $\mathbf{n}$ to final direction $\mathbf{n}'$. The significance of the cross-section theorem is that the total cross section, which is the integral of the differential scattering cross section over all angles, can be re-expressed in terms of a single matrix element only, *linear* rather than quadratic in the interaction V. It is my objective to derive a similar result for the so-called momentum transfer cross section

$$\sigma_d = \int d\mathbf{n}' \, (1 - \mathbf{n}\cdot\mathbf{n}') \, |A(\mathbf{n} \to \mathbf{n}')|^2. \tag{4}$$

From (2)

$$-\int d\mathbf{r} \, (H\psi)^* p\psi + \int d\mathbf{r} \, \psi^* p H\psi = 0, \tag{5}$$

where p is any component of momentum, the z-component say; to keep the integrals in (5) convergent, the integration volume may be supposed to equal the interior of a sphere of finite though very large radius. Using (1), Eq. (5) becomes

$$\frac{-\hbar^2}{2m} \int d\mathbf{r} \left[\psi^* \nabla^2 \frac{\partial \psi}{\partial z} - (\nabla^2 \psi^*) \frac{\partial \psi}{\partial z} \right] + \int d\mathbf{r} \, \psi^* \frac{\partial V}{\partial z} \psi = 0. \tag{6}$$

The solution is

$$\psi = e^{ik\mathbf{n}\cdot\mathbf{r}} + \Phi, \tag{7}$$

where Φ is everywhere outgoing at infinity

$$\lim_{r\to\infty \| \mathbf{n}'} \Phi = A(\mathbf{n} \to \mathbf{n}') e^{ikr} r^{-1}. \tag{8}$$

Substituting (7) into (6),

$$\frac{-\hbar^2}{2m} \int d\mathbf{r} \left\{ e^{-ik\mathbf{n}\cdot\mathbf{r}} \nabla^2 \frac{\partial \Phi}{\partial z} - \frac{\partial \Phi}{\partial z} \nabla^2 e^{-ik\mathbf{n}\cdot\mathbf{r}} + \Phi^* \nabla^2 \frac{\partial}{\partial z} e^{ik\mathbf{n}\cdot\mathbf{r}} - \left(\frac{\partial}{\partial z} e^{ik\mathbf{n}\cdot\mathbf{r}} \right) \nabla^2 \Phi^* + \Phi^* \nabla^2 \frac{\partial \Phi}{\partial z} - \frac{\partial \Phi}{\partial z} \nabla^2 \Phi^* \right\} + \int d\mathbf{r} \, \psi^* \frac{\partial V}{\partial z} \psi = 0, \tag{9}$$

or

$$\frac{-\hbar^2}{2m} \int d\mathbf{S}\cdot \left\{ e^{-ik\mathbf{n}\cdot\mathbf{r}} \nabla \frac{\partial \Phi}{\partial z} - \frac{\partial \Phi}{\partial z} \nabla e^{-ik\mathbf{n}\cdot\mathbf{r}} + \Phi^* \nabla \frac{\partial}{\partial z} e^{ik\mathbf{n}\cdot\mathbf{r}} - \left(\frac{\partial}{\partial z} e^{ik\mathbf{n}\cdot\mathbf{r}} \right) \nabla \Phi^* + \Phi^* \nabla \frac{\partial \Phi}{\partial z} - \frac{\partial \Phi}{\partial z} \nabla \Phi^* \right\} + \int d\mathbf{r} \, \psi^* \frac{\partial V}{\partial z} \psi = 0, \tag{10}$$

where the integration volume in (10) can be extended to infinite r.

Now I specifically choose $\mathbf{n}$ along the positive z-direction. Also in (10)

$$\frac{\partial}{\partial z} = \frac{\partial r}{\partial z}\frac{\partial}{\partial r} + \frac{\partial \theta}{\partial z}\frac{\partial}{\partial \theta} = \cos\theta \frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}, \tag{11}$$

where $d\mathbf{S} = r^2 d\mathbf{n}'$, and $\mathbf{n}\cdot\mathbf{n}' = \cos\theta$. Moreover, it can be shown[1] that

$$\lim_{r\to\infty \| \mathbf{n}'} e^{ik\mathbf{n}\cdot\mathbf{r}} = \frac{2\pi i}{k} \left[\delta(\mathbf{n} + \mathbf{n}') \frac{e^{-ikr}}{r} - \delta(\mathbf{n} - \mathbf{n}') \frac{e^{ikr}}{r} \right]. \tag{12}$$

* This research was supported by the Advanced Research Projects Agency (Project Defender), and was monitored by the U. S. Army Research Office—Durham, under Contract DA-31-124-ARO-D-139.

† Permanent address: Department of Physics, University of Pittsburgh, Pittsburgh, Pennsylvania.

[1] E. Gerjuoy and D. S. Saxon, Phys. Rev. 94, 1445 (1954).

With the asymptotic forms (8) and (12), it is evident that the last term on the right side of (11), being of higher order in $1/r$, makes a negligible contribution to (10). Thus, in (10) we can replace $\partial/\partial z$ by $(\mathbf{n}\cdot\mathbf{n}')\partial/\partial r$.

Substituting now Eqs. (8) and (12) in (10), there results

$$\frac{-\hbar^2}{2m}\int d\mathbf{n}'(\mathbf{n}\cdot\mathbf{n}')[-4\pi ik\delta(\mathbf{n}-\mathbf{n}')A(\mathbf{n}\to\mathbf{n}') + 4\pi ik\delta(\mathbf{n}-\mathbf{n}')A^*(\mathbf{n}\to\mathbf{n}') - 2k^2|A(\mathbf{n}\to\mathbf{n}')|^2] + \int d\mathbf{r}\,\psi^*\frac{\partial V}{\partial z}\psi = 0, \quad (13a)$$

or

$$\int d\mathbf{n}'(\mathbf{n}\cdot\mathbf{n}')\,|A(\mathbf{n}\to\mathbf{n}')|^2 = \frac{4\pi}{k}\,\mathrm{Im}\,A(\mathbf{n}\to\mathbf{n}) - \frac{1}{2E}\int d\mathbf{r}\,\psi^*\frac{\partial V}{\partial z}\psi. \quad (13b)$$

In the event there is any question about employment of (12), as well as dropping the last term in (11), which seems to imply

$$\frac{\partial}{\partial z}e^{ikz} = \frac{\partial}{\partial z}e^{ikr\cos\theta} = ik\cos^2\theta e^{ikz}, \quad (14)$$

I note that with somewhat more effort (13a) can be derived from (10) without direct recourse to (12). The reason use of (14) introduces no contradiction is, of course, the presence of the δ-functions in (12), which effectively make $\cos^2\theta = 1$.

Using (3) and (4), Eq. (13b) implies

$$\sigma_d = \frac{1}{2E}\int d\mathbf{r}\,\psi^*\frac{\partial V}{\partial z}\psi. \quad (15)$$

Equation (15) is the desired result. For spherically symmetric potentials, it is known[2] that

$$\sigma_d = \frac{4\pi}{k^2}\sum_l (l+1)\sin^2(\delta_l - \delta_{l+1}). \quad (16)$$

A direct derivation of (16) from (15), avoiding the forgoing surface integration over the sphere at infinity, is given in the Appendix.

At low energies, where beam techniques tend to be impractical, it is σ_d—not σ—which experiments typically measure.[3] Although at very low energies one expects the differential cross section to be spherically symmetric, implying $\sigma_d = \sigma$, nevertheless there are data suggesting many cross sections, e.g., electron–atom scattering cross sections,[4] may be nonspherically symmetric down to quite low energies. Thus, it becomes questionable whether one assuredly can compare low-energy measurements of σ_d with theoretical estimates of σ. In fact very recent electron–atom collision data at quite low energies[5] indicate σ_d does not wholly agree with effective range theories[6] of σ.

The significance of (15) is that, with recent developments of variational techniques[7] applicable to matrix elements of the form (15) involving continuum eigenfunctions, it may be possible to make accurate estimates of σ_d directly from (15), rather than from (4). In this way we should be able to sidestep the conventional means of evaluating (4), namely, via (16) from the phase shifts for the individual partial waves; it is just this procedure which is likely to fail when the actual angular distribution persists in asymmetry down to very low energies. To put it differently, Equation (15) offers the possibility of developing an "effective range theory" directly for the momentum transfer cross section which under appropriate circumstances, e.g., when $\partial V/\partial z$ is very large, may be appreciably different from the effective range expansion for σ, even at quite low energies. Equation (15) will be particularly useful if the interaction is noncentral, as it can be in electron scattering from atoms or molecules.

It is doubtless possible to generalize (15) to more complicated collisions, including the effects of particle spin, particle interchange and (when they can occur) inelastic collisions. However, these complications customarily are neglected in effect at the essentially thermal energies where σ_d actually is measured. The main approximation is to suppose the scattering can be expressed in terms of a set of amplitudes (e.g., singlet and triplet amplitudes) each of which can be thought to represent scattering by an effectively one-particle Hamiltonian of form (1). If this approximation is not justified, the conventional effective-range theories of σ require generalization no less than Eq. (15).

I remark that this momentum transfer cross section theorem is related to the so-called hypervirial theorems[8] which recently have excited some interest.

[2] D. R. Bates, *Atomic and Molecular Processes* (Academic Press Inc., New York, 1962), p. 645.

[3] H. S. W. Massey and E. H. S. Burhop, *Electronic and Ionic Impact Phenomena* (Oxford University Press, New York, 1952), pp. 366 ff.

[4] Reference 3, Chaps. 2 and 3, especially p. 125.

[5] L. S. Frost and A. V. Phelps, Phys. Rev. **136**, A1538 (1964).

[6] T. F. O'Malley, Phys. Rev. **130**, 1020 (1963).

[7] L. M. Delves, Nucl. Phys. **45**, 313 (1963).

[8] J. O. Hirschfelder, J. Chem. Phys. **33**, 1462 (1960); P. D. Robinson and J. O. Hirschfelder, Phys. Rev. **129**, 1391 (1963); M. B. McElroy and J. O. Hirschfelder, *ibid.* **131**, 1589 (1963).

APPENDIX

To further confirm the correctness of (15), I shall derive (16) directly from (15), using essentially only the properties of the radial eigenfunctions. As will be seen, the derivation is less trivial than might have been expected. The difficulty stems from the fact that (15) involves a matrix element of grad V, whereas customary expressions for scattering amplitudes or scattering phase shifts involve matrix elements of V.

For spherically symmetric V the solution (7) has the expansion

$$\psi = \frac{1}{k}\sum_l (2l+1) i^l e^{i\delta_l} P_l(\cos\theta) R_l(r,k), \quad (17)$$

where

$$\left[\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} - \frac{l(l+1)}{r^2}\right] R_l = \frac{2m}{\hbar^2}(V-E)R_l \quad (18)$$

and

$$\lim_{r\to\infty} R_l = r^{-1}\sin(kr - \tfrac{1}{2}l\pi + \delta_l) \equiv r^{-1}\sin\eta_l. \quad (19)$$

Substitute (17) in (15), and make use of $\partial V/\partial z = \cos\theta\, \partial V/\partial r$, as well as

$$\int_0^\pi d\theta \sin\theta\cos\theta P_l P_{l'} = \frac{2l}{(2l-1)(2l+1)}\delta_{l',l-1} + \frac{2(l+1)}{(2l+1)(2l+3)}\delta_{l',l+1}. \quad (20)$$

It then follows that

$$\frac{1}{2E}\int d\mathbf{r}\, \psi^* \frac{\partial V}{\partial z}\psi = \frac{2\pi i}{Ek^2}\int_0^\infty dr\, r^2 \frac{dV}{dr} \times \left\{\sum_{l=0}^\infty (l+1)\{\exp[i(\delta_{l+1}-\delta_l)]\} R_l R_{l+1} - \sum_{l=1}^\infty l\{\exp[i(\delta_{l-1}-\delta_l)]\} R_l R_{l-1}\right\}. \quad (21)$$

Replacing l by $l+1$ in the last term of (21), so that the sum runs from 0 to ∞, Eq. (21) simplifies to

$$\frac{1}{2E}\int d\mathbf{r}\, \psi^* \frac{\partial V}{\partial z}\psi = \frac{4\pi}{k^2 E}\int_0^\infty dr\, r^2 \frac{dV}{dr} \times \sum_{l=0}^\infty (l+1) R_l R_{l+1}\sin(\delta_l - \delta_{l+1}). \quad (22)$$

Equation (22) will imply (16) if

$$\int_0^\infty dr\, r^2 R_l R_{l+1}\frac{dV}{dr} = E\sin(\delta_l - \delta_{l+1}). \quad (23)$$

I now shall demonstrate that the rather remarkable relation (23) indeed holds. I first claim that

$$R_l R_{l+1}\frac{dV}{dr} = \frac{\hbar^2}{2m} R_{l+1}\left[\left(\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} - \frac{(l+1)(l+2)}{r^2}\right)\left(\frac{d}{dr} - \frac{l}{r}\right)R_l\right] - \frac{\hbar^2}{2m}\left[\left(\frac{d^2R_{l+1}}{dr^2} + \frac{2}{r}\frac{dR_{l+1}}{dr} - \frac{(l+1)(l+2)}{r^2}R_{l+1}\right)\left(\frac{dR_l}{dr} - \frac{lR_l}{r}\right)\right]. \quad (24)$$

Equation (24) is most easily proved noting the identities

$$\left[\frac{d}{dr} + \frac{(l+2)}{r}\right]\left[\frac{d}{dr} - \frac{l}{r}\right] = \frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} - \frac{l(l+1)}{r^2}, \quad (25)$$

$$\left[\frac{d}{dr} - \frac{l}{r}\right]\left[\frac{d}{dr} + \frac{(l+2)}{r}\right] = \frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} - \frac{(l+1)(l+2)}{r^2}.$$

Using (25), the first term on the right side of (24) is

$$\frac{\hbar^2}{2m}R_{l+1}\left[\left(\frac{d}{dr} - \frac{l}{r}\right)\left(\frac{d}{dr} + \frac{(l+2)}{r}\right)\left(\frac{d}{dr} - \frac{l}{r}\right)R_l\right] = \frac{\hbar^2}{2m}R_{l+1}\left(\left[\frac{d}{dr} - \frac{l}{r}\right)\left(\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} - \frac{l(l+1)}{r^2}\right)R_l\right]. \quad (26)$$

Thus using (18), the entire right side of (24) is

$$\frac{\hbar^2}{2m}R_{l+1}\left[\left(\frac{d}{dr} - \frac{l}{r}\right)\frac{2m}{\hbar^2}(V-E)R_l\right] - \frac{\hbar^2}{2m}\left[\frac{2m}{\hbar^2}(V-E)R_{l+1}\left(\frac{dR_l}{dr} - \frac{lR_l}{r}\right)\right],$$

which reduces to the left side of (24).

Having established (24), define

$$u_l = (d/dr - l/r)R_l \quad (27)$$

so that

$$\int_0^\infty dr\, r^2 R_l R_{l+1}\frac{dV}{dr} = \frac{\hbar^2}{2m}\int_0^\infty dr\, r^2 \times \left\{R_{l+1}\left[\left(\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} - \frac{(l+1)(l+2)}{r^2}\right)u_l\right] - u_l\left[\left(\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} - \frac{(l+1)(l+2)}{r^2}\right)R_{l+1}\right]\right\}. \quad (28)$$

Consequently, integrating by parts the terms involving derivatives of u_l,

$$\int_0^\infty dr\, r^2 R_l R_{l+1} \frac{dV}{dr} = \frac{\hbar^2}{2m}\left[r^2\left(R_{l+1}\frac{du_l}{dr} - u_l\frac{dR_{l+1}}{dr}\right)\right]\Bigg|_0^\infty. \quad (29)$$

Even in the worst $l = 0$ case, there is no contribution to the right side of (29) at $r = 0$, where $R_l \sim r^l$, $u_l \sim r^{l-1}$. Using (19), one sees that

$$\lim_{r\to\infty} r^2\left(R_{l+1}\frac{du_l}{dr} - u_l\frac{dR_{l+1}}{dr}\right) = -k^2[\sin\eta_{l+1}\sin\eta_l + \cos\eta_{l+1}\cos\eta_l] = -k^2\cos(\eta_l - \eta_{l+1}). \quad (30)$$

Substituting for η_l in terms of δ_l, Eqs. (29) and (30) now immediately yield (23).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 6 JUNE 1965

Note on Bég's Approach To Peratization*

ELLIOT LEADER
Lawrence Radiation Laboratory, University of California, Berkeley, California
(Received 19 November 1964)

It is shown that Bég's elegant formulation of the peratization result of Feinberg and Pais can be rigorously justified.

I. INTRODUCTION

RECENTLY Bég[1] has given a very elegant derivation of the main result of the peratization theory of Feinberg and Pais.[2] However, Bég's method involves a certain amount of juggling with divergent series, and also the introduction of a regulator mass which is allowed to go to infinity after the completion of the manipulations.

We show in the following that Bég's result (with a minor qualification) is rigorously correct and can be obtained without the use of a regulator.

II. MAIN RESULT OF PERTIZATION THEORY AND BÉG'S METHOD

Consider leptons interacting via the exchange of a massive, charged vector boson, with leading order matrix element (Born approximation) given by

$$B_{\mu\nu}(p) = [-ig^2/(p^2 + m^2)](g_{\mu\nu} + p_\mu p_\nu/m^2). \quad (1)$$

Here p is the 4-momentum transfer, m the boson mass, and g the boson–lepton coupling constant, and the metric is $p^2 = \mathbf{p}^2 - p_0^2$. The indices μ, ν are to be contracted with the usual V–A lepton currents.

Feinberg and Pais[2] have shown that it is possible to give a meaning to the sum of the uncrossed ladder graphs in this theory, with the result that the leading order matrix element becomes modified or peratized to

$$B^P_{\mu\nu}(p) = \frac{-ig^2}{p^2 + m^2}\left(g_{\mu\nu} + \frac{p_\mu p_\nu}{m^2}\right) + \frac{ig^2}{4m^2}g_{\mu\nu}. \quad (2)$$

For simplicity we shall pretend in what follows that we are dealing with neutral vector bosons and that the leptons are massless. The relationship between this and the realistic situation is discussed in detail in Bég's paper.[1]

Following Bég we write, for the scattering amplitude,

$$T_{\mu\nu} = T_a g_{\mu\nu} + T_b[\tfrac{1}{4}p^2 g_{\mu\nu} - p_\mu p_\nu]. \quad (3)$$

From (1) the scalar functions T_a, T_b have Born approximations given by

$$B_a(p) = -\frac{3}{4}\frac{ig^2}{p^2 + m^2} - \frac{ig^2}{4m^2} \quad (4)$$

and

$$B_b(p) = ig^2/m^2(p^2 + m^2).$$

They satisfy the integral equations

$$T_a(p^2) = B_a(p) - \frac{16}{(2\pi)^4}\int B_a(p - q)\frac{1}{q^2}T_a(q^2)\,d^4q \quad (5)$$

* This work was performed under the auspices of the U. S. Atomic Energy Commission.
[1] M. A. B. Bég, Ann. Phys. (N.Y.) 27, 183 (1964).
[2] G. Feinberg and A. Pais, Phys. Rev. 131, 2724 (1963), and Phys. Rev. 133, B 477 (1964).

Consequently, integrating by parts the terms involving derivatives of u_l,

$$\int_0^\infty dr\, r^2 R_l R_{l+1} \frac{dV}{dr} = \frac{\hbar^2}{2m}\left[r^2\left(R_{l+1}\frac{du_l}{dr} - u_l\frac{dR_{l+1}}{dr}\right)\right]\Bigg|_0^\infty. \quad (29)$$

Even in the worst $l = 0$ case, there is no contribution to the right side of (29) at $r = 0$, where $R_l \sim r^l$, $u_l \sim r^{l-1}$. Using (19), one sees that

$$\lim_{r\to\infty} r^2\left(R_{l+1}\frac{du_l}{dr} - u_l\frac{dR_{l+1}}{dr}\right) = -k^2[\sin\eta_{l+1}\sin\eta_l + \cos\eta_{l+1}\cos\eta_l] = -k^2\cos(\eta_l - \eta_{l+1}). \quad (30)$$

Substituting for η_l in terms of δ_l, Eqs. (29) and (30) now immediately yield (23).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 6 JUNE 1965

Note on Bég's Approach To Peratization*

ELLIOT LEADER
Lawrence Radiation Laboratory, University of California, Berkeley, California
(Received 19 November 1964)

It is shown that Bég's elegant formulation of the peratization result of Feinberg and Pais can be rigorously justified.

I. INTRODUCTION

RECENTLY Bég[1] has given a very elegant derivation of the main result of the peratization theory of Feinberg and Pais.[2] However, Bég's method involves a certain amount of juggling with divergent series, and also the introduction of a regulator mass which is allowed to go to infinity after the completion of the manipulations.

We show in the following that Bég's result (with a minor qualification) is rigorously correct and can be obtained without the use of a regulator.

II. MAIN RESULT OF PERTIZATION THEORY AND BÉG'S METHOD

Consider leptons interacting via the exchange of a massive, charged vector boson, with leading order matrix element (Born approximation) given by

$$B_{\mu\nu}(p) = [-ig^2/(p^2 + m^2)](g_{\mu\nu} + p_\mu p_\nu/m^2). \quad (1)$$

Here p is the 4-momentum transfer, m the boson mass, and g the boson–lepton coupling constant, and the metric is $p^2 = \mathbf{p}^2 - p_0^2$. The indices μ, ν are to be contracted with the usual V–A lepton currents.

Feinberg and Pais[2] have shown that it is possible to give a meaning to the sum of the uncrossed ladder graphs in this theory, with the result that the leading order matrix element becomes modified or peratized to

$$B^P_{\mu\nu}(p) = \frac{-ig^2}{p^2 + m^2}\left(g_{\mu\nu} + \frac{p_\mu p_\nu}{m^2}\right) + \frac{ig^2}{4m^2}g_{\mu\nu}. \quad (2)$$

For simplicity we shall pretend in what follows that we are dealing with neutral vector bosons and that the leptons are massless. The relationship between this and the realistic situation is discussed in detail in Bég's paper.[1]

Following Bég we write, for the scattering amplitude,

$$T_{\mu\nu} = T_a g_{\mu\nu} + T_b[\tfrac{1}{4}p^2 g_{\mu\nu} - p_\mu p_\nu]. \quad (3)$$

From (1) the scalar functions T_a, T_b have Born approximations given by

$$B_a(p) = -\frac{3}{4}\frac{ig^2}{p^2 + m^2} - \frac{ig^2}{4m^2} \quad (4)$$

and

$$B_b(p) = ig^2/m^2(p^2 + m^2).$$

They satisfy the integral equations

$$T_a(p^2) = B_a(p) - \frac{16}{(2\pi)^4}\int B_a(p - q)\frac{1}{q^2}T_a(q^2)\,d^4q \quad (5)$$

* This work was performed under the auspices of the U. S. Atomic Energy Commission.
[1] M. A. B. Bég, Ann. Phys. (N.Y.) 27, 183 (1964).
[2] G. Feinberg and A. Pais, Phys. Rev. 131, 2724 (1963), and Phys. Rev. 133, B 477 (1964).

and

$$T_b(p^2) = B_b(p) + \frac{4}{3(2\pi)^4 p^4}\int B_b(p-q) \times [3p^4 - p^2q^2 - 6p^2q\cdot p + 4(p\cdot q)^2]T_b(q^2)\,d^4q. \quad (6)$$

In Eqs. (5) and (6) we have further specialized to the situation in which the leptons are scattering at zero energy. The final momenta are zero, and the initial momenta, p and $-p$, are of course off the mass shell. They are to be put onto the mass shell after the equations are solved.

In this situation the on-shell amplitude is determined entirely by $T_a(0)$ (see Footnote 4 of Ref. 1), and we see that we will recover the result of Feinberg and Pais [Eq. (2)] if we can show that to leading order in g^2 Eq. (5) has the solution

$$T_a(p^2) \simeq B_a^P(p), \quad (7)$$

where

$$B_a^P(p) = -\frac{3}{4}\frac{ig^2}{p^2+m^2} \equiv B_a(p) + \frac{ig^2}{4m^2}. \quad (8)$$

Let us write Eq. (5) symbolically as

$$T_a = B_a + B_aGT_a. \quad (9)$$

Consider now the continuous family of integral equations

$$T(c) = \mathfrak{B} + c + [\mathfrak{B} + c]GT(c), \quad (10)$$

where c is an arbitrary constant and $\mathfrak{B}$ is a Born term such that

$$\mathfrak{B}(p)_{p^2\to\infty} = O(p^{-2}). \quad (11)$$

It is clear that Eq. (9) is precisely of this form, with

$$\mathfrak{B}(p) = -\tfrac{3}{4}ig^2/(p^2+m^2) = B_a^P(p) \quad (12)$$

and

$$c = -ig^2/4m^2.$$

Bég's derivation of the peratization result hinges on the following remarkable result derived by him: There exists a single function

$$T_{\mathfrak{B}}(p^2) \equiv T(p^2; c=0), \quad (13)$$

which is independent of c and which satisfies Eqs. (10) for *all* values of c. Moreover $T_{\mathfrak{B}}(p^2)$ is given by the *iterative* solution of Eq. (10) with $c = 0$.

Once this result is accepted we immediately get the peratized result [Eqs. (7) and (8)] by writing Eq. (9) in the form of Eq. (10), with $\mathfrak{B}(p)$ given by Eq. (12), i.e., in leading order we have

$$T_a(p^2) \equiv T_{\mathfrak{B}}(p^2) \simeq \mathfrak{B}(p) \equiv B_a^P(p). \quad (14)$$

In Bég's work, the above-mentioned result is derived with the use of a regulator by invoking the self-damping properties of the Fermi interaction in the chain approximation.

We shall see in the next section that this result can be obtained directly in a very simple manner and without need of a regulator. However, it will turn out that the validity of the result depends upon the *sign* of $\mathfrak{B}(p)$ as $p^2 \to \infty$, and that the sign in Eq. (12) is "good" in this sense. In this respect the result reached in Ref. 1 is not quite correct, since there the proof was given for the "singular" part of T_a, called T_s, whose $\mathfrak{B}(p)$ term has the opposite sign to that in Eq. (12).

It is thus interesting to note that the result is valid only when $\mathfrak{B}(p)$ corresponds to a potential which is repulsive at small distances.

III. SOLUTION OF INTEGRAL EQUATION FOR $T(c)$

We shall solve Eq. (10) only for the case of direct interest, in which $\mathfrak{B}(p)$ is given by Eq. (12). Writing out the equation in detail we have

$$T(p^2; c) = -\frac{3}{4}\frac{ig^2}{p^2+m^2} + c - \frac{16}{(2\pi)^4} \times \int\left[c - \frac{3}{4}\frac{ig^2}{q^2+m^2}\right]\frac{T(q^2;c)}{q^2}\,d^4q. \quad (15)$$

If a solution exists to Eq. (15) then *a fortiori* the integral

$$\int\left[c - \frac{3}{4}\frac{ig^2}{q^2+m^2}\right]\frac{T(q^2;c)}{q^2}\,d^4q$$

converges, and since the terms in parenthesis cannot possibly cancel against each other, we see that the integral

$$\int\frac{T(q^2;c)}{q^2}\,d^4q$$

also converges.

We can now rewrite Eq. (15) as

$$T(p^2; c) = -\frac{3}{4}\frac{ig^2}{p^2+m^2} - \frac{16}{(2\pi)^4}\int\left[-\frac{3}{4}\frac{ig^2}{q^2+m^2}\right] \times \frac{T(q^2;c)}{q^2}\,d^4q + c - \frac{16c}{(2\pi)^4}\int\frac{T(q^2;c)}{q^2}\,d^4q. \quad (16)$$

The function $T(p^2; c = 0)$, which is the solution of (15) when $c = 0$, will now satisfy (15) for *all* c provided only that

$$c - \frac{16c}{(2\pi)^4}\int\frac{T(q^2; c=0)}{q^2}\,d^4q = 0,$$

i.e., provided that

$$\frac{1}{(2\pi)^4}\int \frac{T(q^2;c=0)}{q^2}\,d^4q = \frac{1}{16}. \tag{17}$$

The condition (17) takes on a less formidable aspect if we look at it in coordinate space.

Defining

$$\mathfrak{L}(x) = \frac{1}{(2\pi)^4}\int e^{-iq\cdot x}\frac{T(q^2;c=0)}{q^2-i\epsilon}\,d^4q, \tag{18}$$

we see that condition (17) is nothing more than the boundary condition

$$\mathfrak{L}(0) = \tfrac{1}{16}. \tag{19}$$

We must not examine the Fourier transform of Eq. (15) and show that it possesses a solution with the required properties.

By taking the Fourier transform of Eq. (15) we get that $\mathfrak{L}(x)$ satisfies the differential equation

$$\Box^2\mathfrak{L}(x) = -\tfrac{3}{4}ig^2\Delta_F(x)[16\mathfrak{L}(x) - 1], \tag{20}$$

where $\Delta_F(x)$ is the usual Feynman propagator function

$$\Delta_F(x) = \frac{1}{(2\pi)^4}\int \frac{e^{-iq\cdot x}}{q^2+m^2-i\epsilon}\,d^4q. \tag{21}$$

Introducing the complex variable η defined by $\eta^2 = x^2$, we can reduce Eq. (20) to

$$\left(\frac{d^2}{d\eta^2}+\frac{3}{\eta}\frac{d}{d\eta}\right)\mathfrak{L}(\eta) = -\tfrac{3}{4}ig^2\Delta_F(\eta)[16\mathfrak{L}(\eta) - 1]. \tag{22}$$

This equation is of the type studied by Pwu and Wu,[3] so we may take over their results directly. For convenience we outline here the essential steps.

Firstly the Feinberg–Pais reduction formula[2] for taking Fourier transforms allows us to consider Eq. (22), restricting η to lie only in the first quadrant of the complex η plane. We may then replace $\Delta_F(\eta)$ by

$$(im/4\pi^2\eta)K_1(m\eta),$$

so that Eq. (22) becomes

$$\left(\frac{d^2}{d\eta^2}+\frac{3}{\eta}\frac{d}{d\eta}\right)\mathfrak{L}(\eta) = \frac{3g^2m}{\pi^2\eta}K_1(m\eta)[\mathfrak{L}(\eta) - \tfrac{1}{16}]. \tag{23}$$

As a result of the analyticity properties of (23) we can ignore the contour integral in the reduction formula and we may thus further restrict ourselves to η real and positive.

Secondly we see that for η real and positive the solutions of (23) have the behavior

$$\mathfrak{L}(\eta) \underset{\eta\to 0}{\longrightarrow} \tfrac{1}{16} + \eta^{\lambda_\pm}, \tag{24}$$

where

$$\lambda_\pm = -1 \pm (1 + 3g^2/\pi^2)^{\frac{1}{2}} \tag{25}$$

and

$$\mathfrak{L}(\eta) \underset{\eta\to\infty}{=} O(\eta^\mu), \tag{26}$$

where

$$\mu = 0 \quad \text{or} \quad -2. \tag{27}$$

Finally, it can be shown that an *iterative* solution exists provided that

$$3g^2/\pi^2 < 1, \tag{28}$$

and which has the behavior given by taking λ_+ in Eq. (24) and $\mu = -2$ in Eq. (26).

Since $\lambda_+ > 0$, the condition (19) is satisfied.

Since $\mu = -2$, the Fourier transform $T(p^2;c=0)$ exists.

This completes the proof of Bég's assertion and justifies his method of obtaining the Feinberg–Pais peratization result.

ACKNOWLEDGMENTS

The author is indebted to M. A. Baqi Bég for many enlightening discussions. He wishes to thank Professor G. C. Wick for his kind hospitality at the Brookhaven National Laboratory, where this work was performed.

[3] Y. Pwu and T. T. Wu, Phys. Rev. **133**, B 778 (1964).

Disconnected Groups as Higher Symmetry Groups*

W. H. KLINK
Johns Hopkins University, Baltimore, Maryland
(Received 16 June 1964; final manuscript received 17 December 1964)

Disconnected groups are investigated to see whether they can be used as higher symmetry groups. Some disconnected groups are given which satisfy certain minimal physical conditions. Two disconnected groups are analyzed in detail using little group techniques, and it is shown how a doubling of certain multiplets results, so that particles and their antiparticles are in the same multiplets—rather than being in separate, though equivalent multiplets as in the usual SU_2 or SU_3 schemes.

SINCE SU_3[1] has thus far been the most successful group in classifying the strongly interacting particles, it is of some interest to list the assumptions which lead to SU_3 and to see whether new classes of groups become available when some of these assumptions are relaxed.

Denote by G a group which is to be used as a "higher symmetry" group. Roughly speaking there are three "physical" assumptions and two "mathematical" assumptions which, when imposed on G, lead to SU_3 (or SU_3/C, where C is the center of SU_3). They are:

(1) SU_2 is a subgroup of G;
(2) G allows the assignment of two additive quantum numbers;
(3) G has an eight-dimensional irreducible representation;
(4) G is a connected Lie group;
(5) G is semisimple.

Assumption 1 means that the strong interactions are charge independent. One may question Assumption 2, for it is well known that it is possible to assign three additive quantum numbers to strongly interacting particles, T_3, the third component of isospin, Y, the hypercharge, and B, baryon number, rather than just two, T_3 and Y, as implicitly assumed in Assumption 2. If G is a connected Lie group, Assumption 2 means G is of rank 2.[2] If G were a connected Lie group of rank 3, its supermultiplets would have different values of B. For example, in such a group there would exist a supermultiplet in which baryons and antibaryons were placed with other particles or resonances which had zero baryon number, while still having the same space–time characteristics as the baryons and antibaryons—and such particles have not been observed. Millerd[3] has investigated rank-3 groups in some detail and concludes that it seems impossible to use them as a "higher symmetry" group because none of them seems to have the correct supermultiplet structure. Another argument against rank-3 groups is that if quarks[4] do not exist, then, as Fairbairn[5] has shown, there exist only rank-4 groups which contain SU_3/C as a subgroup. It will be shown in Sec. III that disconnected groups exist which do not have these difficulties and therefore may be suitable groups for incorporating baryon number.

The prejudice of eight-dimensional irreducible representations in Assumption 3 serves to eliminate the connected rank-2 Lie groups $C_2(B_2)$ and G_2^2 which thus far do not seem to have alternative representations in which to fit the baryons and pseudoscalar mesons. Assumption 5 is needed because all rank-2 connected Lie groups having invariant Abelian subgroups have not been classified.

The purpose of this paper is to relax Assumption 4—a requirement of mathematical convenience—and see whether disconnected groups are possible candidates for "higher symmetry" groups. In Sec. I a general outline of the properties of disconnected groups will be given, including a discussion of how observables appear. In Sec. II, several groups satisfying Assumptions 1 and 2 (or an assumption 2′: G allows the assignment of three additive quantum numbers) will be given. Finally in Sec. III two particular disconnected groups, which incorporate baryon number, will be analyzed in detail, and in particular their irreducible representations will be given.

* Supported in part by the National Science Foundation.
[1] M. Gell-Mann, "The Eightfold Way," CTSL-20 (1961); Phys. Rev. **125**, 1067 (1962); Y. Ne'eman, Nucl. Phys. **26**, 222 (1961).
[2] R. E. Behrends, J. Dreitlein, C. Fronsdal, and B. W. Lee, Rev. Mod. Phys. **34**, 1 (1962).
[3] W. Millerd (private communication).
[4] M. Gell-Mann, Phys. Letters, **8**, 214 (1964).
[5] W. Fairbairn (private communication).

I. DISCUSSION OF DISCONNECTED GROUPS

In this paper, the letter "D" will designate a disconnected group, where by disconnected group is meant any group in which it is impossible to pass continuously from the identity element to every other element of the group. The infinite dihedral group, D_∞, the group generated by rotations about the z axis in a real three-dimensional space plus a discrete 180° rotation about the y axis forms a disconnected group, as does the group G_0 of Lee and Yang.[6]

In both of these disconnected groups there exist subgroups which are connected Lie groups. These subgroups are the set of elements which can be reached continuously from the identity. It can be shown[7] that such a subgroup, called the component of the identity, always forms an invariant subgroup of the disconnected group. In the case of D_∞ the invariant subgroup is the group of rotations about the z axis. The quotient groups of both D_∞ and G_0 are of order 2; disconnected groups which have quotient groups of order greater than 2 are the disconnected subgroups of SU_3, called $\Delta(3\infty^2)$ and $\Delta(6\infty^2)$.[8]

The fact that the component of the identity forms an invariant subgroup of a disconnected group means that if the irreducible representations of the component of the identity are known, the irreducible representations of the disconnected group can always be found by using the techniques of little groups.[9] All the disconnected groups considered in this paper have, as components of the identity, Lie groups whose irreducible representations are known.

One of the reasons that connected Lie groups have been used as higher symmetry groups is that a powerful mathematical theory exists[2] which generates all possible semisimple connected Lie groups. No such comparable theory exists for generating disconnected groups. If the connected Lie group which forms the invariant subgroup of D is denoted by L, then D/L is isomorphic to a finite group. Since there are an enormous number of finite groups, it would seem to be a hopeless task to classify the disconnected groups. However, if certain physical requirements are imposed on the disconnected groups, the number of admissible groups is sharply reduced. For example, in the disconnected groups considered in this paper it will always be possible to assign two (T_3 and Y) or three (T_3, Y, and B) additive quantum numbers. If Assumptions 1 and 3 above are also imposed, only a small number of disconnected groups remain. Several disconnected groups satisfying the above requirements are given in Sec. II.

It is possible to label the elements of a disconnected group D by

$$D(\alpha_1, \alpha_2, \alpha_3 \cdots \alpha_n, m) = D(\alpha, m), \qquad (1)$$

where the α_i are parameters which vary continuously between zero and 2π, and m is a discrete parameter which labels the various disconnected parts of the group—and therefore takes on as many different values as the order of the finite group D/L. The invariant subgroup of D is $D(\alpha, 1)$.

To obtain observables from disconnected groups, it is necessary to pick a complete set of commuting elements of the group. The most natural choice of a complete set of commuting observables is obtained by considering elements in the component of the identity, $D(\alpha, 1)$. Since this subgroup forms a connected Lie group, it is possible to consider its Lie algebra and obtain the observables from a complete set of commuting elements of the Lie algebra. However, in contrast to connected Lie groups, it may be possible to adjoin to these observables, observables coming from the disconnected parts of the group. It is always possible to write elements of the disconnected parts of the group as

$$D(\alpha, m) = D(\alpha, 1)D(0, m). \qquad (2)$$

It is not difficult to show that only elements of the type $D(0, m)$ would give rise to new observables. Such elements would have the property that the additive quantum numbers arising from them would be conserved only modulo an integer, as is the case generally with finite groups.[8]

Finally it should be noted that since the disconnected groups being considered are compact, it is possible to define invariant integration, and thus evaluate Kronecker products. Clebsch–Gordan coefficients can be obtained for the disconnected groups if the Clebsch–Gordan coefficients for the component of the identity are known.

II. EXAMPLES OF DISCONNECTED GROUPS

In order to satisfy Assumption 1—namely that D contain isospin invariance—it is necessary to have the component of the identity contain SU_2 as a subgroup. As an example of this, consider

[6] T. D. Lee and C. N. Yang, Phys. Rev. **122**, 1954 (1961).
[7] L. S. Pontriagin, *Topological Groups* (Princeton University Press, Princeton, New Jersey, 1939).
[8] W. Fairbairn, T. Fulton, and W. H. Klink, J. Math. Phys. 5, 1478 (1964).
[9] J. S. Lomont, *Applications of Finite Groups* (Academic Press Inc., New York, 1959), Chap. V.

$$D(\alpha_1, \alpha_2, m) = \left\{\begin{bmatrix} SU_2(\alpha_1) & 0 \\ 0 & SU_2(\alpha_2) \end{bmatrix}, \begin{bmatrix} 0 & SU_2(\alpha_1) \\ SU_2(\alpha_2) & 0 \end{bmatrix}\right\} = \left\{\begin{matrix} D(\alpha_1, \alpha_2, 1) \\ \\ D(\alpha_1, \alpha_2, 2) \end{matrix}\right\}, \quad (3)$$

where $SU_2(\alpha_i)$ is the group of unitary, unimodular 2×2 matrices with three parameters α_i. The group $SU_2(\alpha_3) \otimes D(\alpha_1, \alpha_2, m)$ is the global symmetry group G_0 considered by Lee and Yang.[6] Notice that

$$D(\alpha_1, \alpha_2, m)/D(\alpha, {}_1\alpha_2, 1) \cong S_2, \quad (4)$$

where S_2 is the permutation group on two letters. Since the component of the identity $D(\alpha_1, \alpha_2, 1)$ is the direct product of two SU_2 groups, the multiplets may be labeled by T_3 and T_3' and will have the structure of an isospin multiplet within an isospin multiplet. Clearly many disconnected groups can be formed along the lines of (3). Since all finite groups are subgroups of some permutation group and the permutation group S_n can always be written in the regular representation as

$$S_n = \left\{ I_n = \begin{bmatrix} 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \\ 0 & 0 & 1 & \\ \vdots & & & \ddots \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 & \cdots \\ 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 \cdots \\ \vdots & & & \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 & 0 \cdots \\ 0 & 1 & 0 & 0 \cdots \\ 1 & 0 & 0 & 0 \cdots \\ 0 & 0 & 0 & 1 \cdots \\ \vdots & & & \end{bmatrix}, \cdots \right\} \quad (5)$$

disconnected groups can be generated by replacing the 1's in the identity element of (5) by a suitable connected group.[10] However, it should be pointed out that for S_n, $n > 3$, there are always more than three additive quantum numbers, which violates Assumption 2′.

Instead of considering only SU_2 it is possible to consider $U_2(\alpha) = SU_2(\alpha_1, \alpha_2, \alpha_3) \otimes e^{i\alpha_4}$ and form disconnected groups in exactly the same manner as above, everywhere replacing SU_2 by U_2. However, these groups all violate Assumption 2′.

It is also possible to form disconnected groups in the following way:

$$D(\alpha, m) = \left\{\begin{bmatrix} SU_2(\alpha) & 0 \\ 0 & SU_2^*(\alpha) \end{bmatrix}, \begin{bmatrix} 0 & SU_2(\alpha) \\ SU_2^*(\alpha) & 0 \end{bmatrix}\right\} = \left\{\begin{matrix} D(\alpha, 1) \\ D(\alpha, 2) \end{matrix}\right\}, \quad (6)$$

[10] D. R. Speiser and J. Tarski, J. Math. Phys. 4, 588 (1963) (preprint version).

where * means complex conjugation. Such a group does not have as many additive quantum numbers as groups like (3). However, it seems to be possible to form groups like (6) only when $D(\alpha, m)/D(\alpha, 1)$ is of order 2. The two groups which are analyzed in some detail in Sec. III are of the same type as (6). They are

$$D(\alpha, m) = \left\{\begin{bmatrix} U_2(\alpha) & 0 \\ 0 & U_2^*(\alpha) \end{bmatrix}, \begin{bmatrix} 0 & U_2(\alpha) \\ U_2^*(\alpha) & 0 \end{bmatrix}\right\} \equiv D_{U_2}, \quad (7)$$

$$D(\alpha, m) = \left\{\begin{bmatrix} U_3(\alpha) & 0 \\ 0 & U_3^*(\alpha) \end{bmatrix}, \begin{bmatrix} 0 & U_3(\alpha) \\ U_3^*(\alpha) & 0 \end{bmatrix}\right\} \equiv D_{U_3}, \quad (8)$$

where $U_3 = SU_3(\alpha_1 \cdots \alpha_8) \otimes e^{i\alpha_9}$. The irreducible representations of these two groups are obtained and it is shown how they naturally bring in the additive quantum number baryon number. Thus, D_{U_2} will have the quantum numbers T_3 and B, while D_{U_3} will have T_3, Y, and B.

III. IRREDUCIBLE REPRESENTATIONS OF D_{U_2} AND D_{U_3}

In this section the irreducible representations of D_{U_2} and D_{U_3} will be obtained. The characteristic feature of these two groups is that d-dimensional multiplets of the component of the identity having nonzero baryon number will become $2d$-dimensional multiplets with baryon number B and $-B$, while those with baryon number zero will remain d dimensional. Such a statement is not quite true for D_{U_3} since SU_3 has representations for which the representation and its complex conjugate are inequivalent. D_{U_3} is analyzed in more detail than D_{U_2} since the analysis of the more complicated group D_{U_3} is a straightforward generalization of that done for D_{U_2}.

We thus consider the group D_{U_2} of (7). Now it is possible to write the group $SU_2(\alpha)$ as

$$\begin{aligned} SU_2(\alpha) &= I_2 \cos \alpha_1 - i\tau \cdot \mathbf{n} \sin \alpha_1 \\ &= I_2 \cos \alpha_1 - i\tau \cdot \mathbf{l}, \end{aligned} \quad (9)$$

where $\mathbf{l} = \mathbf{n} \sin \alpha_1$, τ are the 2×2 Pauli matrices, α_1 is an angle of "rotation," and $\mathbf{n}$ is a unit vector specifying the axis of "rotation."

Notice that $l^2 = \sin^2 \alpha_1$, so that $\mathbf{l}$ specifies the direction and, except for a double valuedness, the magnitude of the "rotation." If now a "rotation"

specified by 1′ is followed by a "rotation" specified by 1, one gets

$$SU_2(\mathbf{L}) = SU_2(1) \times SU_2(1'), \tag{10}$$

where

$$\mathbf{L} = 1 \cos \alpha_1' + 1' \cos \alpha_1 + 1 \times 1' \tag{11}$$

and gives the resultant of the two "rotations."

Now parametrize $U_2(\alpha)$ by

$$U_2(\alpha) \to SU_2(1)e^{i\alpha_3} = (I_2 \cos \alpha_1 - i\tau \cdot 1)e^{i\alpha_3}. \tag{12}$$

Then it is possible to write D_{U_3} as

$$\begin{aligned} D(1, \alpha_2, 1) &= \begin{bmatrix} SU_2(1)e^{i\alpha_3} & 0 \\ 0 & SU_2^*(1)e^{-i\alpha_3} \end{bmatrix}, \\ D(1, \alpha_2, 2) &= \begin{bmatrix} 0 & SU_2(1)e^{i\alpha_3} \\ SU_2^*(1)e^{-i\alpha_3} & 0 \end{bmatrix}. \end{aligned} \tag{13}$$

By multiplying the matrices (13) we get the following group table:

	$D(1', \alpha_2', 1)$	$D(1, \alpha_2', 2)$
$D(1, \alpha_2, 1)$	$D(1 \circ 1', \alpha_2 + \alpha_2', 1)$	$D(1 \circ 1', \alpha_2 + \alpha_2', 2)$
$D(1, \alpha_2, 2)$	$D(1 \circ 1'^*, \alpha_2 - \alpha_2', 2)$	$D(1 \circ 1'^*, \alpha_2 - \alpha_2', 1)$

(14)

where $1 \circ 1'$ means $1 \cos \alpha_1' + 1' \cos \alpha_1 + 1 \times 1'$. It is clear from (9) that 1^* means $l_1 \to -l_1$, $l_2 \to l_2$, $l_3 \to -l_3$.

Notice that the "*" and "$\circ$" operations commute. Also

$$\begin{aligned} D^{-1}(1, \alpha_2, 1) &= D(-1, -\alpha_2, 1), \\ D^{-1}(1, \alpha_2, 2) &= D(-1^*, \alpha_2, 2). \end{aligned} \tag{15}$$

Finally

$$\begin{aligned} &D(1, \alpha_2, 2)D(1', \alpha_2', 1)D^{-1}(1, \alpha_2, 2) \\ &\quad = D(1 \circ 1'^*, \alpha_2 - \alpha_2', 2)D(-1^*, \alpha_2, 2) \\ &\quad = D(1 \circ 1'^* \circ -1, -\alpha_2', 1), \end{aligned} \tag{16}$$

indicating that indeed $D(1, \alpha_2, 1)$ is an invariant subgroup of D_{U_3}.

The irreducible representations of $D(1, \alpha_2, 1)$ are

$$D(1, \alpha_2, 1) \to R^{T,B}(1, \alpha_2) = \mathfrak{D}^T(1)e^{iB\alpha_3}, \tag{17}$$

where $\mathfrak{D}^T(1)$ are the Wigner matrices,[11] $T = 0, \frac{1}{2}, 1, \cdots$ and $B = 0, \pm 1, \pm 2, \cdots$. The letter B is used because it will be seen later to correspond to baryon number.

[11] E. P. Wigner, *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra* (Academic Press Inc., New York, 1959), Chap. 15.

Consider now those elements $D(1, \alpha_2, 2)$ which satisfy the relation

$$\begin{aligned} &R^{T,B}[D(1, \alpha_2, 2)D(1', \alpha_2', 1)D^{-1}(1, \alpha_2, 2)] \\ &\qquad \text{is equivalent to } R^{T,B}(1', \alpha_2') \end{aligned} \tag{18}$$

for all 1′ and α_2'. These elements plus all the elements $D(1, \alpha_2, 1)$ form what is called the little group of the second kind, $L^{II}(T, B)$.[9]

The kernel $K(T, B)$ is that set of elements $D(1, \alpha_2, 1)$ which are mapped into the identity of the representation $R^{T,B}$. There are several little groups to be distinguished:

$$\begin{aligned} &(1)\quad T = 0, \quad B = 0, \quad L^{II}(0, 0) = D_{U_3}, \\ &\qquad K(0, 0) = D(1, \alpha_2, 1), \frac{L^{II}(0, 0)}{K(0, 0)} \cong S_2, \\ &(2)\quad T, B \neq 0, \quad L^{II}(T, B \neq 0) = D(1, \alpha_2, 1), \\ &\qquad K(T, B \neq 0) = D(0, 0, 1), \\ &(3)\quad T \neq 0, \quad B = 0, \quad L^{II}(T \neq 0, 0) = D_{U_3}, \\ &\qquad K(T \neq 0, 0) = D(0, \alpha_2, 1), \\ &\qquad L^{II}(T \neq 0, 0)/K(T \neq 0, 0) \cong D_{SU_3}, \end{aligned} \tag{19}$$

where D_{SU_3} is D_{U_3} with $\alpha_2 = 0$.

The theory of little groups states that some of the irreducible representations of $L^{II}(T, B)/K(T, B)$ can be used to generate the irreducible representations of D_{U_3}. These representations of $L^{II}(T, B)/K(T, B)$ are called allowable by Lomont[9] and it is easy to check that all of the irreducible representations of $L^{II}(T, B)/K(T, B)$ given in the next paragraph are allowable.

The irreducible representations of $L^{II}(0, 0)/K(0, 0)$ are trivial. From them two irreducible representations of D_{U_3} can be immediately given:

$$\begin{Bmatrix} D(1, \alpha_2, 1) \\ D(1, \alpha_2, 2) \end{Bmatrix} \to 1, \qquad \begin{Bmatrix} D(1, \alpha_2, 1) \\ D(1, \alpha_2, 2) \end{Bmatrix} \to \begin{matrix} 1 \\ -1. \end{matrix} \tag{20}$$

The irreducible representations of $L^{II}(T, B \neq 0)/K(T, B \neq 0)$ are given in (17) while the irreducible representations of D_{SU_3} are given in Ref. 11, p. 342 and can be summarized as

$$\begin{aligned} D(1, 0, 1) &\to \mathfrak{D}^T(1) \\ D(1, 0, 2) &\to \mathfrak{D}^T(1)M^T. \end{aligned} \tag{21}$$

M^T is a $(2T + 1)$-dimensional matrix which satisfies $M^T\mathfrak{D}^T(1)(M^T)^{-1} = \mathfrak{D}^T(1^*)$ and $[M^T]^2 = I_{2T+1}$, where I_{2T+1} is the $(2T + 1)$-dimensional identity matrix. The irreducible representations of D_{U_3} generated by $L^{II}(T \neq 0, 0)/K(T \neq 0, 0)$ are

$$D(1, \alpha_2, 1) \rightarrow \mathfrak{D}^T(1),$$
$$D(1, \alpha_2, 2) \rightarrow \mathfrak{D}^T(1)M^T. \tag{22}$$

To obtain the irreducible representations of D_{U_3} generated by the allowable representations of $L^{II}(T, B \neq 0)$, it is necessary to evaluate the 2×2 matrix

$$\sigma_{ij}(1', \alpha_2'; 1, \alpha_2, m) = \begin{cases} 1 \text{ when } D(0,0,i)D(1', \alpha_2', 1)D^{-1}(0,0,j) \\ \qquad = D(1, \alpha_2, m) \\ 0 \text{ otherwise.} \end{cases} \tag{23}$$

One obtains

$$\sigma_{ij} = \begin{bmatrix} \delta_{m,1}\delta_{1,1'}\delta_{\alpha_2,\alpha_2'} & \delta_{m,2}\delta_{1,1'}\delta_{\alpha_2,\alpha_2'} \\ \delta_{m,2}\delta_{1^*,1'}\delta_{-\alpha_2,\alpha_2'} & \delta_{m,1}\delta_{1^*,1'}\delta_{-\alpha_2,\alpha_2'} \end{bmatrix}. \tag{24}$$

Then the irreducible representations generated by $L^{II}(T, B \neq 0)$ are

$$S^{T,B}(1, \alpha_2, m) = \sum_{\substack{\text{all elements} \\ \text{of } D(1', \alpha_2', 1)}} \sigma_{ij} \otimes R^{T,B}(1', \alpha_2') \tag{25}$$

or

$$D(1, \alpha_2, 1) \rightarrow S^{T,B}(1, \alpha_2, 1) = \begin{bmatrix} \mathfrak{D}^T(1)e^{iB\alpha_2} & 0 \\ 0 & \mathfrak{D}^T(1^*)e^{-iB\alpha_2} \end{bmatrix},$$
$$D(1, \alpha_2, 2) \rightarrow S^{T,B}(1, \alpha_2, 2) = \begin{bmatrix} 0 & \mathfrak{D}^T(1)e^{iB\alpha_2} \\ \mathfrak{D}^T(1^*)e^{-iB\alpha_2} & 0 \end{bmatrix}. \tag{26}$$

Multiplets of the group D_{U_3} can be labeled by $|T, B; T_3, \pm B\rangle$ where T and B label the irreducible representation and T_3 and $\pm B$ label the possible states in a given multiplet. Thus the group SU_2 has the doublets

$$\begin{pmatrix} p \\ n \end{pmatrix} = \left|\tfrac{1}{2}; \begin{matrix} \tfrac{1}{2} \\ -\tfrac{1}{2} \end{matrix}\right\rangle, \quad B = 1$$

and

$$\begin{pmatrix} \bar{n} \\ \bar{p} \end{pmatrix} = \left|\tfrac{1}{2}; \begin{matrix} \tfrac{1}{2} \\ -\tfrac{1}{2} \end{matrix}\right\rangle, \quad B = -1.$$

These combine in D_{U_3} to form one quartet $|\frac{1}{2}, 1; \pm\frac{1}{2}, \pm 1\rangle$. However, for particles with $B = 0$, no such doubling occurs, as can be seen from (22); thus the pions, labeled

$$\left|1, \begin{matrix} 1 \\ 0 \\ -1 \end{matrix}\right\rangle$$

remain

$$\left|1, 0; \begin{matrix} 1 \\ 0, 0 \\ -1 \end{matrix}\right\rangle$$

in the group D_{U_3}. Notice that since B can take on any integer value, multiplets with arbitrarily high baryon number are allowed.

Exactly the same type of analysis can be carried out for $D_{U_3/C}$. $D_{U_3/C}$ is analyzed rather than D_{U_3} because at present there is no evidence for the existence of quarks. The only difference between D_{U_3} and $D_{U_3/C}$ occurs because in SU_3/C some representations and their complex conjugate may be inequivalent; for example, the 10 and 10* representations are inequivalent. Thus, even when B is zero, a doubling of representations can occur.

The little groups are now labeled $[(p, q)_d, B]$, where d is the dimension and (p, q) the Weyl indices[2] which label the irreducible representations of SU_3 The little groups, $L^{II}[(p, q)_d, B]$, are

(1) $L^{II}[(0, 0)_1, 0]$—generates two 1-dimensional irreducible representations with the same structure as (20).

(2) $L^{II}[(p, q)_d, B \neq 0]$—generates irreducible representations of dimension $2d$.

(3a) $L^{II}[(p, q)_d, B = 0]$, (p, q) and (q, p) are inequivalent—generates irreducible representations of dimension $2d$.

(3b) $L^{II}[(p, q)_d B = 0]$, (p, q) and (q, p) are equivalent—generates irreducible representations of same dimension as $(p, q)_d$.

Only the little groups $L^{II}[(p, q)_d, B \neq 0]$ generate multiples containing baryon number B and $-B$. The other little groups generate multiplets with baryon number zero. In analogy to D_{U_3}, states of $D_{U_3/C}$ may be labeled by $|(p, q)_d, B; T, T_3, Y, \pm B\rangle$.

IV. CONCLUSION

It has been shown how disconnected groups may be used in the study of higher symmetries. Though a careful analysis was only done for D_{U_3}, the most interesting group is $D_{U_3/C}$. In this group the strongly interacting particles are placed in the following multiplets:

(1) Baryons and antibaryons—$|(1, 1)_8, 1; T, T_3, Y, \pm 1\rangle$—16 dimensions, contains two eight-dimensional representations of SU_3/C.

(2) Baryon and antibaryon resonances—$|(3, 0)_{10}, 1; T, T_3, Y, \pm 1\rangle$—20 dimensional, contains the 10 and 10* representations of SU_3/C.

(3) Mesons—$|(1, 1)_8, 0; T, T_3, Y, 0\rangle$—eight dimen-

sional, similar to the eight-dimensional representation of SU_3/C.

(4) Meson resonances—$|(3, 0)_{10}, 0; T, T_3, Y, 0\rangle$—20 dimensional, contains the 10 and 10* representations of SU_3/C.

The Clebsch–Gordan coefficients obtained from reducing Kronecker products of $D_{U_3/C}$, such as those needed for baryon–meson scattering, is the same as in SU_3/C, since no new overlapping states result in $D_{U_3/C}$. Thus, any predictions made by SU_3/C, such as comparisons of cross sections of various reactions, is the same in $D_{U_3/C}$.

ACKNOWLEDGMENTS

The author gratefully acknowledges the help and advice of Dr. Thomas Fulton; he also wishes to thank Dr. H. Yoshizawa for helpful mathematical discussions.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 6 JUNE 1965

Generalized Boltzmann Equation for a Quantum Gas Obeying Classical Statistics

S. FUJITA*
Department of Physics, Pennsylvania State University, University Park, Pennsylvania
(13 November 1964)

By means of the connected-diagram expansion a generalized Boltzmann equation is obtained for the Wigner distribution function f describing distinguishable quantum particles subject to constant electric and magnetic fields. This equation is exact and meaningful in the thermodynamic limit provided that the main interaction processes which cause the change in f involve finite numbers of particles. If the interaction processes are assumed to be localized in time, and long-time solutions are looked for, the equation could be simplified so that its collision term can be described in terms of f only just like the conventional Boltzmann equation for a dilute gas.

1. INTRODUCTION

RECENTLY Severne[1] generalized the diagram techniques developed by Prigogine and his coworkers[2] to discuss an inhomogeneous classical gas. In particular he obtained the generalized Boltzmann equation valid to the general order in coupling constant, and showed that in the kinetic stage the collision term can be expressed in terms of the one-body distribution function only. In the present paper[3] we deal with a quantum gas of distinguishable particles subject to constant electric and magnetic fields.

In some earlier theories[1,2,4] developed for nonequilibrium problems, use is made of the representation in which a chosen unperturbed Hamiltonian is diagonal. This makes the theories appear much dependent on the splitting of the Hamiltonian. In a previous note,[5] the present author developed a representation-independent connected-diagram expansion treatment for an electron-impurity system, which allows one-body description. Generalizing this treatment, we shall, in the present article, derive a non-Markoffian evolution equation (3.11) for the one-body density matrix n (or Wigner distribution function f) describing a quantum gas. This equation is exact and meaningful in the thermodynamic limit (3.8) provided that the main interaction processes which cause the change in n involve finite numbers of particles. In many situations, the equation could be simplified so that its collision term can be described in terms of n instead of the two-body density matrix which appears in the mechanical evolution equation (2.14).

In the absence of electromagnetic fields and in the classical mechanical limit the equation becomes equivalent to the equation derived by Severne.[1]

The present treatment shows that in the kinetic stage, the integral of the product of two-body density matrix and potential in the mechanical evolution

* On leave from l'Université Libre de Bruxelles, Brussels, Belgium.

[1] G. Severne, Physica (to be published).

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

[3] The content of this paper is published in Phys. Letters 12, 300 (1964).

[4] L. Van Hove, Physica **21**, 517 (1955); **23**, 441 (1957).

[5] S. Fujita, Phys. Letters **10**, 175 (1964).

sional, similar to the eight-dimensional representation of SU_3/C.

(4) Meson resonances—$|(3, 0)_{10}, 0; T, T_3, Y, 0\rangle$—20 dimensional, contains the 10 and 10* representations of SU_3/C.

The Clebsch–Gordan coefficients obtained from reducing Kronecker products of $D_{U_3/C}$, such as those needed for baryon–meson scattering, is the same as in SU_3/C, since no new overlapping states result in $D_{U_3/C}$. Thus, any predictions made by SU_3/C, such as comparisons of cross sections of various reactions, is the same in $D_{U_3/C}$.

ACKNOWLEDGMENTS

The author gratefully acknowledges the help and advice of Dr. Thomas Fulton; he also wishes to thank Dr. H. Yoshizawa for helpful mathematical discussions.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 6, NUMBER 6 JUNE 1965

Generalized Boltzmann Equation for a Quantum Gas Obeying Classical Statistics

S. FUJITA*
Department of Physics, Pennsylvania State University, University Park, Pennsylvania
(13 November 1964)

By means of the connected-diagram expansion a generalized Boltzmann equation is obtained for the Wigner distribution function f describing distinguishable quantum particles subject to constant electric and magnetic fields. This equation is exact and meaningful in the thermodynamic limit provided that the main interaction processes which cause the change in f involve finite numbers of particles. If the interaction processes are assumed to be localized in time, and long-time solutions are looked for, the equation could be simplified so that its collision term can be described in terms of f only just like the conventional Boltzmann equation for a dilute gas.

1. INTRODUCTION

RECENTLY Severne[1] generalized the diagram techniques developed by Prigogine and his coworkers[2] to discuss an inhomogeneous classical gas. In particular he obtained the generalized Boltzmann equation valid to the general order in coupling constant, and showed that in the kinetic stage the collision term can be expressed in terms of the one-body distribution function only. In the present paper[3] we deal with a quantum gas of distinguishable particles subject to constant electric and magnetic fields.

In some earlier theories[1,2,4] developed for nonequilibrium problems, use is made of the representation in which a chosen unperturbed Hamiltonian is diagonal. This makes the theories appear much dependent on the splitting of the Hamiltonian. In a previous note,[5] the present author developed a representation-independent connected-diagram expansion treatment for an electron-impurity system, which allows one-body description. Generalizing this treatment, we shall, in the present article, derive a non-Markoffian evolution equation (3.11) for the one-body density matrix n (or Wigner distribution function f) describing a quantum gas. This equation is exact and meaningful in the thermodynamic limit (3.8) provided that the main interaction processes which cause the change in n involve finite numbers of particles. In many situations, the equation could be simplified so that its collision term can be described in terms of n instead of the two-body density matrix which appears in the mechanical evolution equation (2.14).

In the absence of electromagnetic fields and in the classical mechanical limit the equation becomes equivalent to the equation derived by Severne.[1]

The present treatment shows that in the kinetic stage, the integral of the product of two-body density matrix and potential in the mechanical evolution

* On leave from l'Université Libre de Bruxelles, Brussels, Belgium.

[1] G. Severne, Physica (to be published).

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

[3] The content of this paper is published in Phys. Letters 12, 300 (1964).

[4] L. Van Hove, Physica **21**, 517 (1955); **23**, 441 (1957).

[5] S. Fujita, Phys. Letters **10**, 175 (1964).

equation (2.14) can be expressed in terms of one-body density matrix. This is of relevance to, but different from, Bogoliubov's conjecture[6] on the functional dependence of the two-body distribution function through one-body distribution function. The difference may be resolved by saying that those two-particle correlations in the two-body distribution function which make the conjecture approximate are effectively destroyed in the thermodynamic limit by the integration with respect to the coordinate of one of the two particles in (2.14).

The applications and the quantum statistical generalization of the present theory will be discussed in separate publications.

2. DENSITY MATRIX AND ITS MECHANICAL EVOLUTION

Let us consider a system of distinguishable particles characterized by the Hamiltonian

$$H \equiv \sum_i h_0^{(i)} + \lambda \sum_{i>j} v^{(ij)}, \tag{2.1}$$

where the upper indices denote particles and λ the coupling constant. The single-particle Hamiltonian h_0 may contain the energies due to the electric and magnetic fields (**E**, **B**), and may be written as

$$h_0 \equiv \frac{1}{2M}\left[\mathbf{p} - \frac{e}{c}\mathbf{A}(\mathbf{r}, t)\right]^2 + e\Phi(\mathbf{r}, t), \tag{2.2}$$

where the customary notations are used. For simplicity we assume that the external fields are constant in space and time:

$$\begin{aligned} \mathbf{E} &\equiv -\nabla\Phi = \text{constant}, \\ \mathbf{B} &\equiv \nabla \times \mathbf{A} = \text{constant}. \end{aligned} \tag{2.3}$$

It will not be difficult to extend the following treatment to more complicated situations.

The position density matrix is defined by

$$\begin{aligned} n(\mathbf{r}_1\mathbf{r}_2, t) &\equiv \mathrm{Tr}\,\{\hat{n}_{12}\rho(t)\} \\ &[\equiv n_r(\mathbf{R}, t)], \end{aligned} \tag{2.4}$$

$$\hat{n}_{12} \equiv \sum_i |\mathbf{r}_2^{(i)}\rangle\langle\mathbf{r}_1^{(i)}|, \tag{2.5}$$

where $\rho(t)$ is the density operator, and $|\mathbf{r}_k^{(i)}\rangle$ and $\langle\mathbf{r}_k^{(i)}|$ are ket and bra vectors associated with the position value $\mathbf{r}_k$ of the particle j. We shall specify the $\xi' - \xi''$ element

$$\xi' \equiv (\mathbf{r}_1'^{(1)}, \mathbf{r}_2'^{(2)}, \cdots, \mathbf{r}_N'^{(N)}) \tag{2.6}$$

of an arbitrary N-body operation A by the set of numbers (q, Q):

$$\begin{aligned} \langle\xi'|\, A\, |\xi''\rangle &\equiv A_q(Q), \\ q \equiv \xi' - \xi'', &\qquad Q \equiv \tfrac{1}{2}(\xi' + \xi''), \end{aligned} \tag{2.7}$$

or

$$\mathbf{r}_i^{(j)} \equiv \mathbf{r}_i'^{(j)} - \mathbf{r}_i''^{(j)}, \quad \mathbf{R}_i^{(j)} \equiv \tfrac{1}{2}(\mathbf{r}_i'^{(j)} + \mathbf{r}_i''^{(j)}). \tag{2.8}$$

In this specification the equation of motion for $n_r(\mathbf{R}, t) \equiv n(\mathbf{r}_1\mathbf{r}_2, t)$ can be written as

$$\begin{aligned} &[\partial/\partial t + i\mathfrak{h}_0(\mathbf{r}, \mathbf{R})]n_r(\mathbf{R}, t) \\ &= -i\lambda \sum_{ik} \int \cdots \int d^{3N}q\, d^{3N}Q \hat{n}_{12q}^{(ik)}(Q)\mathfrak{v}^{(ki)} \\ &\quad \times \exp\,[-i\mathfrak{K}(q, Q)t]\rho_q(Q, 0), \end{aligned} \tag{2.9}$$

$$\mathfrak{K}(q, Q) \equiv \sum_i \mathfrak{h}_0^{(i)}(q, Q) + \lambda \sum_{i>j} \mathfrak{v}^{(ij)}(q, Q), \tag{2.10}$$

$$\begin{aligned} \mathfrak{h}_0^{(j)}(q, Q) &\equiv \mathfrak{h}_0(\mathbf{r}, \mathbf{R}) \\ &= \frac{1}{M}\left(-i\frac{\partial}{\partial\mathbf{r}} - \frac{e}{c}\mathbf{A}(\mathbf{R})\right) \\ &\quad \cdot\left(-i\frac{\partial}{\partial\mathbf{R}} - \frac{e}{c}\mathbf{A}(\mathbf{r})\right) - e\mathbf{r}\cdot\mathbf{E}, \\ \mathfrak{v}^{(ij)}(q, Q) &\equiv \mathfrak{v}(\mathbf{rs}, \mathbf{RS}) \\ &= v(\mathbf{R} + \tfrac{1}{2}\mathbf{r} - \mathbf{S} - \tfrac{1}{2}\mathbf{s}) \\ &\quad - v(\mathbf{R} - \tfrac{1}{2}\mathbf{r} - \mathbf{S} + \tfrac{1}{2}\mathbf{s}), \end{aligned} \tag{2.11}$$

where (2.3) is used in deriving the first equations of (2.11).

If we introduce the two-body density matrix by

$$n(\mathbf{r}_1\mathbf{r}_3, \mathbf{r}_2\mathbf{r}_4, t) \equiv \mathrm{Tr}\,\{\hat{n}_{13,24}\rho(t)\} \tag{2.12}$$

$$\hat{n}_{13,24} \equiv \sum_{i>j} |\mathbf{r}_2^{(j)}\rangle\,|\mathbf{r}_4^{(i)}\rangle\langle\mathbf{r}_3^{(i)}|\,\langle\mathbf{r}_1^{(j)}|, \tag{2.13}$$

we may write (2.9) as

$$\begin{aligned} &\left\{\frac{\partial}{\partial t} + \frac{i}{M}\left[-i\frac{\partial}{\partial\mathbf{r}} - \frac{e}{c}\mathbf{A}(\mathbf{R})\right]\right. \\ &\quad \left.\cdot\left[-i\frac{\partial}{\partial\mathbf{R}} - \frac{e}{c}\mathbf{A}(\mathbf{r})\right] - ie\mathbf{r}\cdot\mathbf{E}\right\}n_r(\mathbf{R}, t) \\ &= -i\lambda \int d^3S[v(\mathbf{R} + \tfrac{1}{2}\mathbf{r} - \mathbf{S}) \\ &\quad - v(\mathbf{R} - \tfrac{1}{2}\mathbf{r} - \mathbf{S})]n(\mathbf{R} + \tfrac{1}{2}\mathbf{rS}, \mathbf{R} - \tfrac{1}{2}\mathbf{rS}, t). \end{aligned} \tag{2.14}$$

In this equation we see that the evolution of the one-body density matrix is describable in terms of one- and two-body density matrices. In general we can see that the mechanical evolution of the k-body density matrix can be described in terms of density

[6] N. Bogoliubov, J. Phys. (URSS) **10**, 256, 265 (1946); English translation, in *Studies in Statistical Mechanics*, edited by G. E. Uhlenbeck and J. de Boer (North-Holland Publishing Company, Amsterdam, 1962), Vol. I., pp. 1–118.

matrices involving at most $k+1$ particle coordinates, $k+1 \leq N$ (the number of particles). The chain of evolution equations, $k = 1, 2, \cdots, N-1$, is known as the hierarchy equations of Bogoliubov–Born–Green–Kirkwood–Yvon.

If we introduce the gauge-invariant Wigner distribution function f by

$$f(\mathbf{R}, \boldsymbol{\pi}, t) \equiv \int d^3r \times \exp\left[-i\mathbf{r}\cdot\left(\boldsymbol{\pi} + \frac{e}{c}\mathbf{A}(\mathbf{R})\right)\right] n_r(\mathbf{R}, t), \tag{2.15}$$

$$n_r(\mathbf{R}, t) = \int \frac{d^3\pi}{(2\pi)^3} \times \exp\left[i\mathbf{r}\cdot\left(\boldsymbol{\pi} + \frac{e}{c}\mathbf{A}(\mathbf{R})\right)\right] f(\mathbf{R}, \boldsymbol{\pi}, t),$$

we can easily show that the lhs (left-hand side) of (2.14) becomes

$$\left[\frac{\partial}{\partial t} + \frac{1}{M}\boldsymbol{\pi}\cdot\frac{\partial}{\partial \mathbf{R}} + \left(e\mathbf{E} + \frac{e}{M}\boldsymbol{\pi}\times\mathbf{B}\right)\cdot\frac{\partial}{\partial \boldsymbol{\pi}}\right] \times f(\mathbf{R}, \boldsymbol{\pi}, t), \tag{2.16}$$

which is just the expected form of the time variation term subtracted by the flow term in the conventional Boltzmann equation. In this equation, which is valid for the kinetic description of a dilute gas, the collision term $\partial f/\partial t|_{\text{collision}}$ is expressed in terms of one-body distribution function only, thus making the equation *closed* in contrast to the mechanical equation (2.14). Does this characteristic of the collision term still hold in more general situations? If so, in what conditions? These questions will be answered in the following section.

3. GENERALIZED BOLTZMANN EQUATION

We may expand $e^{i\mathcal{H}t}$ in the perturbation series:

$$e^{-i\mathcal{H}t} = e^{-i\mathcal{H}_0 t}\left[1 + \sum_1^\infty (-i\lambda)^k \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 \times \int_0^{\tau_{k-1}} d\tau_k\, \mathcal{V}(\tau_1)\mathcal{V}(\tau_2)\cdots\mathcal{V}(\tau_k)\right] \tag{3.1}$$

$$\mathcal{V}(\tau) \equiv e^{i\mathcal{H}_0 t}\mathcal{V}e^{-i\mathcal{H}_0 t}. \tag{3.2}$$

We represent terms in the perturbation expansion of the rhs of (2.9) by diagrams.

We draw N horizontal solid lines for the N particles. Corresponding to $\mathfrak{v}^{(ij)}(\tau)$ we draw a vertical dotted line, called *potential bond*, connecting the particle lines i and j at t (time) $= \tau$, where the time is measured from the right to the left. Corresponding to those components of $\rho(0)$ involving potentials $\mathfrak{v}^{(ij)}$ we draw dotted lines connecting the lines i and j at $t = 0$.

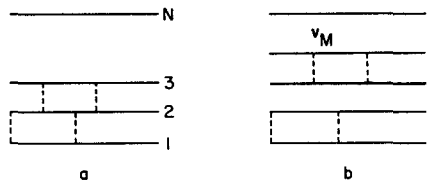


FIG. 1. Diagrams representing components of the rhs of (2.9). The diagram (a) is connected and (b) disconnected.

In this way we can represent all the perturbation terms in the one-to-one correspondence. Some typical diagrams are drawn in Fig. 1.

A diagram is said to be *connected* if the set of particles describing the potential bonds cannot be separated into two or more subsets. Otherwise the diagram will be called a *disconnected* one. For example the diagram (a) in Fig. 1 is connected while the diagram (b) is disconnected.

Certain diagrams have a potential bond of the *type M*, like the v_M in the diagram (b), which sees nothing but the two free particle lines on its left. A particle line segment is said to be *free* if the diagram breaks into two by cutting it. Denoting the coordinates of the two particles with $(\mathbf{r}\mathbf{r}', \mathbf{R}\mathbf{R}')$, we may write the contribution of the diagram in the following form:

$$\int_0^t d\tau \int\cdots\int d^3R\, d^3r\, d^3R'\, d^3r'\, \delta^{(3)}(\mathbf{r})\delta^{(3)}(\mathbf{r}') \times \exp\{-i[\mathfrak{h}_0(\mathbf{r}, \mathbf{R}) + \mathfrak{h}_0(\mathbf{r}', \mathbf{R}')](t-\tau)\} \times [v(\mathbf{R} + \tfrac{1}{2}\mathbf{r} - \mathbf{R}' - \tfrac{1}{2}\mathbf{r}') - v(\mathbf{R} - \tfrac{1}{2}\mathbf{r} - \mathbf{R}' + \tfrac{1}{2}\mathbf{r}')]g(\mathbf{r}\mathbf{R}, \mathbf{r}'\mathbf{R}', \tau), \tag{3.3}$$

where g is a certain function of the arguments shown, and the two delta functions arise from $n_{12\alpha}(Q)$. We immediately notice that the integral (3.3) will vanish if the exponential function is replaced with 1. Introducing the following expression for the delta function

$$\delta^{(3)}(\mathbf{r}) = \frac{1}{(2\pi)^3}\int d^3\pi \exp\left\{i\mathbf{r}\cdot\left[\boldsymbol{\pi} + \frac{e}{c}\mathbf{A}(\mathbf{R})\right]\right\}, \tag{3.4}$$

we may write (3.3) as

$$\int_0^t d\tau \int\cdots\int d^3R\, d^3R' \frac{d^3\pi}{(2\pi)^3}\frac{d^3\pi'}{(2\pi)^3} \times [\exp\{-i[h_0(\boldsymbol{\pi}, \mathbf{R}) + h_0(\boldsymbol{\pi}', \mathbf{R}')](t-\tau)\} - 1] \times \left[v\left(\mathbf{R} + \frac{i}{2}\frac{\partial}{\partial\boldsymbol{\pi}} - \mathbf{R}' - \frac{i}{2}\frac{\partial}{\partial\boldsymbol{\pi}'}\right) - v\left(\mathbf{R} - \frac{i}{2}\frac{\partial}{\partial\boldsymbol{\pi}} - \mathbf{R}' + \frac{i}{2}\frac{\partial}{\partial\boldsymbol{\pi}'}\right)\right] g'(\boldsymbol{\pi}\mathbf{R}, \boldsymbol{\pi}'\mathbf{R}', \tau), \tag{3.3a}$$

$$h_0(\pi, \mathrm{R}) \equiv -i\frac{\pi}{M}\cdot\frac{\partial}{\partial \mathrm{R}} + \left(e\mathbf{E} + \frac{e}{c}\frac{\pi}{M}\times\mathbf{B}\right)\cdot\frac{\partial}{\partial \pi}. \quad (3.5)$$

We thus see that (3.3a) can be written as a surface integral which may vanish in many situations of physical interest (notably excluding the case of long-range forces).

This means that *any diagram having a M-type potential bond contributes nothing.*

All the disconnected diagrams have M-type potential bonds except those which contain some connected parts involving potential bonds exclusively at $t = 0$. These parts can be shown to contribute only to the normalization, and therefore these are irrelevant to the temporal change of $n_r(\mathrm{R}, t)$.

Thus, we have only to deal with connected diagrams containing the potential bond at $t = t$. We may simplify the drawings by omitting particle lines without potential connections. For example, we may represent the diagram (a) in Fig. 1 as a component of the diagram (a) in Fig. 2 where we leave out indices for particle lines. By such an *unindexed diagram* we may imply a collection of particle-indexed diagrams of the same structure.

A connected diagram will in general contain several free-particle line segments. Some free segments are indicated by check marks $\surd$ in Fig. 2. A diagram will contain a certain number of those parts which consist of nonfree line segments and of potential bonds, and which are connected with each other by free segments. Such a part will be called a *d-part* or *g-part* according to whether it contains or not a potential bond at $t = 0$. The diagram (a) in Fig. 2 has a g-part and the diagram (c) a d-part. The diagram (b) contains two g-parts, one of which has the potential bond at $t = t$.

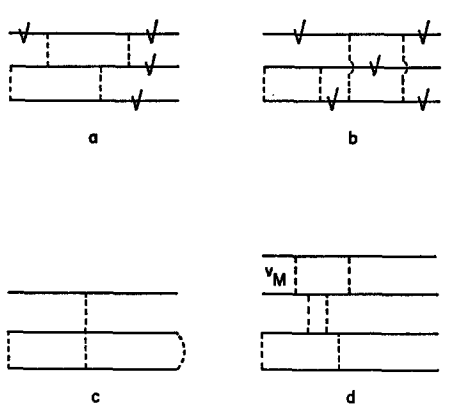


FIG. 2. The unindexed diagram (a) corresponds to the set of indexed diagrams of the same structure as that in the diagram (a) in Fig. 1. The reducible diagram (b) contains two g-parts. The diagram (c) contains a d-part. The diagram (d) containing a M-type potential v_M contributes nothing.

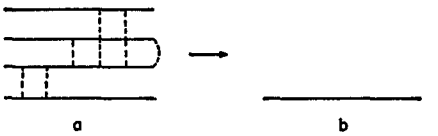


FIG. 3. Diagrams representing terms in the expansion of the density matrix (3.6). The diagram (a) is reducible to the diagram (b).

If a diagram contains a g-part suspended by two free segments corresponding to the same particle or a d-part standing to the right of a free-line segment, it can be simplified or *reduced* by suppressing the g- or d-part. Otherwise the diagram is called *irreducible.* In the reduction, the only particle line representing a particle at the fixed values (r, R) should not be suppressed. By this rule, the reduction becomes unique as illustrated in Fig. 3.

We have so far considered diagrams representing the rhs of (2.9). We may represent by diagrams the perturbation expansion of the density matrix

$$n_r(\mathrm{R}, t) \equiv \sum_i \int \cdots \int d^{3N}q\, d^{3N}Q\hat{n}^{(i)}_{12q}(Q) \times e^{-i\mathcal{H}(q,Q)t}\rho_q(Q, 0). \quad (3.6)$$

Analyzing in a similar manner, we find that any diagram giving a nontrivial contribution is a connected one containing a certain number of g- and/or d-parts. A typical diagram is drawn in Fig. 3(a). It is immediately seen that every diagram (except one) is reducible to the unique diagram (b) in Fig. 3 which is free from any potential bond.

The diagrams drawn here seem to represent the particles evolution contributing to the density matrix $n_r(\mathrm{R}, t)$ at $t = t$, and g- and d-parts describe the effect of interaction processes, each involving a few particles and extending over a region of the order of the force range in linear position.

The same sort of interpretation will hold for diagrams representing the rhs of (2.9).

Let us consider an irreducible diagram containing a g-part. Its contribution will be contained in

$$\int \cdots \int e^{-i\mathcal{H}_0 t}\Big\{-i\lambda\mathcal{V}(t) + \sum_1^\infty (-i\lambda)^{k+1}\int_0^t d\tau_1 \times \int_0^{\tau_1} d\tau_2 \cdots \int_0^{\tau_{k-1}} d\tau_k[\mathcal{V}(t)\mathcal{V}(\tau_1)\cdots\mathcal{V}(\tau_k)]^{(ic)}\Big\} \times \mathrm{Tr}\left\{\prod_{l=2}^{m+1} d^3r^{(l)}\, d^3R^{(l)}\hat{n}_{r^{(l)}}(\mathrm{R}^{(l)})\hat{n}_r(\mathrm{R})\rho(0)\right\}, \quad (3.7)$$

where the upper index ic means that the contribution of only irreducible connected parts (g-parts in the present case) should be taken; $m + 1$ is the number of particles involved in a chosen part.

Consider now a reducible diagram which upon reduction gives rise to the irreducible diagram. By the construction of the diagram the two sets of particles involved in the evolution of two particles of the $m+1$ particles are separate from each other. Therefore, the evolution of one particle may be followed independently of that of the other. Furthermore, the structure of all those subdiagrams which upon reduction give rise to the free line segment for one particle is seen to be identical with the structures of all the diagrams for $n_r(\mathbf{R}, t)$ in the thermodynamic limit:

$$N \to \infty, \quad \Omega \to \infty \quad \text{such that}$$
$$N/\Omega = \text{constant}. \tag{3.8}$$

Noticing these points we may write for the contribution of all the irreducible diagrams containing g-part and their associated reducible diagrams

$$\begin{aligned} G(\mathbf{r}, \mathbf{R}, t; n) \equiv & \int \cdots \int \prod_{l=2}^{m+1} \delta^{(3)}(\mathbf{r}^{(l)}) e^{-i\mathfrak{K}_0 t} \\ & \times \left\{ -i\lambda \mathcal{U}(t) + \sum_1^\infty (-i\lambda)^{k+1} \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 \cdots \right. \\ & \left. \times \int_0^{\tau_{k-1}} d\tau_k [\mathcal{U}(t)\mathcal{U}(\tau_1) \cdots \mathcal{U}(\tau_k)]^{(ic)} \right\} e^{i\mathfrak{h}_0 \tau^{(1)}} \\ & \times \prod_{l'=2}^{m+1} d^3r^{(l')}\, d^3R^{(l')} \exp(i\tau^{(l')}\mathfrak{h}_0^{(l')}) n_{r^{(l')}} \\ & \times (R^{(l')}, \tau^{(l')}), \end{aligned} \tag{3.9}$$

where $\tau^{(l)}$ denotes the smallest of those τ_i which characterize the interactions $\mathcal{U}$ involving the particle l.

The contribution of all irreducible diagrams containing d-parts and their associated reducible diagrams may be written as

$$\begin{aligned} D(\mathbf{r}, \mathbf{R}, t, \rho(0)) \equiv & \iint e^{-i\mathfrak{K}_0 t} \left\{ -i\lambda [\mathcal{U}(t)\rho_q(Q, 0)]^{(ic)} \right. \\ & + \sum_1^\infty (-i\lambda)^{k+1} \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 \cdots \int_0^{\tau_{k-1}} d\tau_k \\ & \left. \times [\mathcal{U}(t)\mathcal{U}(\tau_1) \cdots \mathcal{U}(\tau_k)\rho_q(Q, 0)]^{(ic)} \right\}, \end{aligned} \tag{3.10}$$

which is not expressible in terms of $n_r(\mathbf{R}, t)$.

Thus we may write for (2.9)

$$\begin{aligned} & [\partial/\partial t + i\mathfrak{h}_0(\mathbf{r}, \mathbf{R})] n_r(\mathbf{R}, t) \\ & \quad = G(\mathbf{r}, \mathbf{R}, t; n) + D(\mathbf{r}, \mathbf{R}, t; \rho(0)) \end{aligned} \tag{3.11}$$

in the thermodynamic limit.

Using the relations (2.15) and (3.4) ,we may rewrite (3.11) as

$$\begin{aligned} & \left\{ \frac{\partial}{\partial t} + M^{-1}\boldsymbol{\pi} \cdot \frac{\partial}{\partial \mathbf{R}} + \left[e\mathbf{E} + \frac{e}{Mc} \boldsymbol{\pi} \times \mathbf{B} \right] \cdot \frac{\partial}{\partial \boldsymbol{\pi}} \right\} f(\mathbf{R}, \boldsymbol{\pi}, t) \\ & \quad = G(\mathbf{R}, \boldsymbol{\pi}, t, f) + D(\mathbf{R}, \boldsymbol{\pi}, t; \rho(0)). \end{aligned} \tag{3.12}$$

Here in view of the transformations (2.15) functions of $(\mathbf{R}, \boldsymbol{\pi})$ should be constructed with the aid of the equivalence relations:

$$\begin{aligned} & i \frac{\partial}{\partial \boldsymbol{\pi}} \bigg|_{\mathbf{R}} \leftrightarrow \mathbf{r}, \\ & \boldsymbol{\pi} + \frac{e}{c}\mathbf{A}(\mathbf{R}) \leftrightarrow -i \frac{\partial}{\partial \mathbf{r}} \bigg|_{\mathbf{R}}, \\ & -i \frac{\partial}{\partial \mathbf{R}} \bigg|_{\boldsymbol{\pi}} + \frac{e}{c}\mathbf{A}(\mathbf{r}) \leftrightarrow -i \frac{\partial}{\partial \mathbf{R}} \bigg|_{\mathbf{r}}, \\ & \mathbf{R} \leftrightarrow \mathbf{R}. \end{aligned} \tag{3.13}$$

If we assume that the change of $f(\mathbf{R}, \boldsymbol{\pi}, t)$ is due to interaction processes *localized* in space and time the term D coming from d-parts describing localized interactions may be dropped out a long time compared with the average collision duration after initial time point $t = 0$. In this case, the equation (3.12) becomes a *closed* equation with respect to f just like the usual Boltzmann equation for a dilute gas. The equation can be further simplified in practical applications which will be dealt with in separate publications.

ACKNOWLEDGMENTS

The author wishes to express his sincere thanks to Professor I. Prigogine and Dr. P. Résibois for valuable discussions.

Fluctuations in Equilibrium and Nonequilibrium Ensembles*

GERALD L. JONES AND MICHAEL DAVIES†
Department of Physics and Radiation Laboratory, University of Notre Dame, Notre Dame, Indiana
(Received 9 November 1964)

Using the results of a previous paper, the probability of a fluctuation occurring in a given time interval, of certain dynamical variables, is considered for a rather general nonequilibrium situation and also a more specialized equilibrium case. In the first case we show that this probability vanishes at least as fast as $1/N$. In the second case we obtain a much stronger decrease of the probability with increasing N.

I. INTRODUCTION

IN a previous paper[1] an upper bound has been found for the probability that a dynamical variable will suffer a fluctuation from its ensemble mean value, of a given size in a given time interval. The theorem is applicable to fluctuations in either equilibrium or nonequilibrium situations. In Sec. II of this paper we give an application to certain nonequilibrium situations. In Sec. III we repeat an application to an equilibrium case (fluctuation of kinetic energy in a canonical ensemble) given in the previous paper, but we obtain a much stronger result by the use of more specialized techniques.

The notation in this paper will be the same as in Ref. 1.

II. NONEQUILIBRIUM CASE

The fluctuation problem in the nonequilibrium case arises in the following way. One has an ensemble distribution function which is not constant in time. From this one can compute the mean values of variables $\langle A\rangle_t$. This is most often done by deriving from the Liouville equation for the distribution function in Γ space $f(x, t)$, some kind of an irreversible kinetic equation, and then using the kinetic equation to find the time dependence of various dynamical variables. To be specific let us consider a distribution such that a particular $\langle A\rangle_t$ approaches some stationary value A_s in a time t_r, for example, some kind of relaxation to equilibrium process. Even supposing that $\langle A\rangle_t$ is known, a question arises as to what the time dependence of $\langle A\rangle_t$ has to do with the actual variation of A for the single system in which we are watching the relaxation. What we would like is that most or all of the members of the ensemble have A behaving nearly the same way as $\langle A\rangle_t$ at least over a time interval long enough to follow the relaxation process, i.e., longer than t_r. To say it loosely, this will guarantee that the relaxation is characteristic of a single system and is not a characteristic of only the ensemble mean values.

Now in Ref. 1 we have proven that if $P(\alpha, \tau)$ is the probability that a member of the ensemble has $|A(x, t) - \langle A\rangle_t| > \alpha$ for at least one t in the interval 0 to τ, then

$$P(\alpha, \tau) \leq 4\alpha^{-2}C + 8\alpha^{-3}\int_0^\tau D(\tau')\,d\tau', \tag{1}$$

where

$$C = \langle[A(x) - \langle A\rangle_0]^2\rangle_0 \tag{2}$$

$$D(\tau) = \langle|\{A(x), H(x)\}$$

$$- \langle\{A(x), H(x)\}\rangle_\tau|\,[A(x) - \langle A\rangle_\tau]^2\rangle_\tau. \tag{3}$$

We use the notation $\langle\ \rangle_\tau$ to mean the phase average using the distribution function $f(x, \tau)$.

Now if $P(\alpha, \tau) \ll 1$ then for most of the members of the ensemble $|A(x, t) - \langle A\rangle_t| < \alpha$ for all t in the interval 0 to τ, and this is just the property we mentioned in the discussion at the beginning of this section. Whether or not $P(\alpha, \tau) \ll 1$ will depend on α, τ, $H(x)$, $A(x)$, and $f(x)$ and we must therefore make some assumptions about these quantities. It is of particular interest when $\alpha^{-2}C$ and $\alpha^{-3}D(\tau)$ turn out to be of order $1/N$, where N is the number of degrees of freedom of the system, since one can then make $P(\alpha, \tau)$ small by considering large systems.

We shall consider a classical gas of N particles in a container of volume Ω interacting with smooth two-body forces of finite range. Therefore $x = \{\mathbf{r}_1, \mathbf{p}_1 \cdots \mathbf{r}_N, \mathbf{p}_N\}$ and

$$H(x) = \sum_i -p_i^2/2m + \tfrac{1}{2}\sum_{i,j}{}' V(|\mathbf{r}_i - \mathbf{r}_j|). \tag{4}$$

We shall consider fluctuations of any dynamical variable of the form

$$A(x) = \sum_i a(\mathbf{p}_i), \tag{5}$$

that is, a sum of functions of the individual particle momenta. The class of distribution functions we wish

* Supported in part by the U. S. Atomic Energy Commission.
† Supported in part by the A. E. C. Radiation Laboratory of the University of Notre Dame.
[1] G. L. Jones, J. Math. Phys. 6, 106 (1965).

to allow is, in general, the same as those considered by Prigogine and co-workers[2] in their treatment of the nonequilibrium behavior of a gas. They assume, roughly, that the generic reduced distribution functions are bounded in the limit $N \to \infty$, $\Omega \to \infty$, and N/Ω constant, and that all correlations are of finite range in this limit. To be specific we make the following requirements $f(x, t) = f(\mathbf{r}_1, \mathbf{p}_1 \cdots \mathbf{r}_N, \mathbf{p}_N, t)$ shall be a normalized ensemble distribution function, symmetric under interchange of particle indices. It shall be such that the following functions are finite in the thermodynamic limit.

$$\varphi(\mathbf{p}_1, t) = \int f(\mathbf{r}_1, \mathbf{p}_1 \cdots \mathbf{r}_N, \mathbf{p}_N, t)\, d\mathbf{r}_1 \cdots d\mathbf{r}_N\, d\mathbf{p}_2 \cdots d\mathbf{p}_N, \tag{6a}$$

$$f_s(\mathbf{r}_1, \mathbf{p}_1 \cdots \mathbf{r}_s, \mathbf{p}_s, t) = [N!/(N-s)!] \int f(\mathbf{r}_1, \mathbf{p}_1 \cdots \mathbf{r}_N, \mathbf{p}_N, t)\, d\mathbf{r}_{s+1} \cdots d\mathbf{r}_N\, d\mathbf{p}_{s+1} \cdots d\mathbf{p}_N, \tag{6b}$$

$$f_{s,r}(\mathbf{r}_1, \mathbf{p}_1 \cdots \mathbf{r}_s, \mathbf{p}_s, \mathbf{p}_{s+1} \cdots \mathbf{p}_r, t) = [N!/(N-s)!] \int f(\mathbf{r}_1, \mathbf{p}_1 \cdots \mathbf{r}_N, \mathbf{p}_N, t)\, d\mathbf{r}_{s+1} \cdots d\mathbf{r}_N\, d\mathbf{p}_{r+1} \cdots d\mathbf{p}_N. \tag{6c}$$

We also assume the factorization property

$$f_{s,r}(\mathbf{r}_1, \mathbf{p}_1 \cdots \mathbf{r}_s, \mathbf{p}_s, \mathbf{p}_{s+1} \cdots \mathbf{p}_r, t) = f_s(\mathbf{r}_1, \mathbf{p}_1 \cdots \mathbf{r}_s, \mathbf{p}_s, t)\varphi(\mathbf{p}_{s+1}, t) \cdots \varphi(\mathbf{p}_r, t). \tag{7}$$

With these assumptions we note that

$$\langle A \rangle_t = N\langle a \rangle_t = N \int \varphi(\mathbf{p}, t) a(\mathbf{p})\, d\mathbf{p}, \tag{8}$$

where $\langle a \rangle_t$ is supposed to depend only on the ratio N/Ω. Hence A is an extensive variable. Now we shall consider only macroscopic fluctuations, that is, we choose

$$\alpha = \gamma \langle A \rangle_0 = \gamma N \langle a \rangle_0, \tag{9}$$

where we suppose γ to be independent of the size of the system and much less than 1. (We note that in certain cases where $\langle a \rangle_0$ may vanish identically, usually because of some symmetry, one must measure α by some other criterion.)

Now the dependence of the quantity C on the size of the system can be determined from the assumptions we have made. Using (2) and (5) we have

$$C = \sum_{i,j} \int [a(\mathbf{p}_i) - \langle a \rangle_0][a(\mathbf{p}_j) - \langle a \rangle_0] \times f(\mathbf{r}_1, \mathbf{p}_1 \cdots \mathbf{r}_N, \mathbf{p}_N, 0)\, d\mathbf{r}_1 \cdots d\mathbf{p}_N.$$

Using the symmetry of f under interchange of particle index and splitting the sum into a part with $i = j$ and a part $i \neq j$ we have

$$C = \int [a(\mathbf{p}_1) - \langle a \rangle_0][a(\mathbf{p}_2) - \langle a \rangle_0] \times f_2(\mathbf{r}_1, \mathbf{p}_1, \mathbf{r}_2, \mathbf{p}_2, 0)\, d\mathbf{r}_1\, d\mathbf{r}_2\, d\mathbf{p}_1\, d\mathbf{p}_2 + \int [a(\mathbf{p}) - \langle a \rangle_0]^2 f_1(\mathbf{r}, \mathbf{p}, 0)\, d\mathbf{r}\, d\mathbf{p}. \tag{10}$$

[2] I. Prigogine, *Non-Equilibrium Statistical Mechanics* (John Wiley & Sons, Inc., New York, 1962).

Now from (6) and (7) we see that

$$\int f_1(\mathbf{r}, \mathbf{p}, t)\, d\mathbf{r} = N\varphi(\mathbf{p}, t), \tag{11a}$$

$$\int f_2(\mathbf{r}_1, \mathbf{p}_1, \mathbf{r}_2, \mathbf{p}_2, t)\, d\mathbf{r}_1\, d\mathbf{r}_2 = N(N-1)\varphi(\mathbf{p}_1, t)\varphi(\mathbf{p}_2, t). \tag{11b}$$

Hence,

$$C = N(N-1) \int [a(\mathbf{p}_1) - \langle a \rangle_0] \times [a(\mathbf{p}_2) - \langle a \rangle_0]\varphi(\mathbf{p}_1, 0)\varphi(\mathbf{p}_2, 0)\, d\mathbf{p}_1\, d\mathbf{p}_2 + N \int [a(\mathbf{p}) - \langle a \rangle_0]^2 \varphi(\mathbf{p}, 0)\, d\mathbf{p}.$$

The first term vanishes identically because of the independence of the integrations. The expression for C can then be written

$$C = N\langle [a - \langle a \rangle_0]^2 \rangle_0 = N[\langle a^2 \rangle_0 - \langle a \rangle_0^2], \tag{12}$$

where both $\langle a^2 \rangle_0$ and $\langle a \rangle_0^2$ do not depend on the size of the system. From (12) and (9) we find

$$\alpha^{-2} C = (N\gamma^2)^{-1}(\langle a^2 \rangle_0 \langle a_0 \rangle^{-2} - 1), \tag{13}$$

and this we can make as small as we wish, for fixed γ, by increasing the size of the system.

The treatment of the quantity $D(\tau)$ is merely a slightly more complicated version of the above argument. From (4) and (5) we have

$$\{A(x), H(x)\} = -\sum_{i,j}{}' \frac{\partial a(\mathbf{p}_i)}{\partial \mathbf{p}_i} \cdot \mathbf{F}(\mathbf{r}_i, \mathbf{r}_j), \tag{14}$$

where $\mathbf{F}(\mathbf{r}_i, \mathbf{r}_j)$ is the force exerted by the jth particle on the ith particle, and we have used $\mathbf{F}(\mathbf{r}_i, \mathbf{r}_j) = -\mathbf{F}(\mathbf{r}_j, \mathbf{r}_i)$. For convenience we shall write (14) in the form

$$\{A(x), H(x)\} = \sum_{i,j}{}' \mathbf{F}(\mathbf{r}_i, \mathbf{r}_j) \cdot [\partial a(\mathbf{p}_i)/\partial \mathbf{p}_i - \partial a(\mathbf{p}_j)/\partial \mathbf{p}_j] = \sum_{i,j}{}' K(\mathbf{r}_i, \mathbf{p}_i, \mathbf{r}_j, \mathbf{p}_j). \tag{15}$$

Note that K is invariant under interchange of particle index. We shall define

$$K(t) = \int |K(\mathbf{r}_1, \mathbf{p}_1, \mathbf{r}_2, \mathbf{p}_2)| \times f_2(\mathbf{r}_1, \mathbf{p}_1, \mathbf{r}_2, \mathbf{p}_2, t)\, d\mathbf{r}_1\, d\mathbf{r}_2\, d\mathbf{p}_1\, d\mathbf{p}_2. \quad (16)$$

Since $K(\mathbf{r}_1, \mathbf{p}_1, \mathbf{r}_2, \mathbf{p}_2)$ vanishes for $|\mathbf{r}_1 - \mathbf{r}_2|$ greater than the range of the intermolecular force, $K(t)$ is proportional to Ω; that is, it is an extensive quantity. We define an intensive quantity by

$$K(t) = Nk(t). \quad (17)$$

From the definition (3) of $D(\tau)$ and the triangle inequality one can obtain

$$D(\tau) \leq \sum_{i,j}{}' \sum_{k,l} \langle |K(\mathbf{r}_i, \mathbf{p}_i, \mathbf{r}_j, \mathbf{p}_j)| \times [a(\mathbf{p}_k) - \langle a\rangle_\tau][a(\mathbf{p}_l) - \langle a\rangle_\tau]\rangle_\tau + \sum_{k,l} Nk(\tau)\langle [a(\mathbf{p}_k) - \langle a\rangle_\tau][a(\mathbf{p}_l) - \langle a\rangle_\tau]\rangle_\tau.$$

The second term can be written as

$$N^2k(\tau)[\langle a^2\rangle_\tau - \langle a\rangle_\tau^2]$$

by the same argument that was used to derive (12). We split the first term, involving four summations into a number of contributions.

(a) Those terms in the sum for which no two indices are the same. This contribution can be written as

$$\int |\Lambda(\mathbf{r}_1, \mathbf{p}_1, \mathbf{r}_2, \mathbf{p}_2)|\, [a(\mathbf{p}_3) - \langle a\rangle_\tau][a(\mathbf{p}_4) - \langle a\rangle_\tau] \times f_4(\mathbf{r}_1, \mathbf{p}_1 \cdots \mathbf{r}_4, \mathbf{p}_4, \tau)\, d\mathbf{r}_1 \cdots d\mathbf{p}_4. \quad (18)$$

The r_3 and r_4 variables appear only in f_4 and we see from (6) and (7) that

$$\int f_4(\mathbf{r}_1, \mathbf{p}_1 \cdots \mathbf{r}_4, \mathbf{p}_4, \tau)\, d\mathbf{r}_3\, d\mathbf{r}_4 = (N-2)(N-3) f_{2,4}(\mathbf{r}_1, \mathbf{p}_1, \mathbf{r}_2, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4, \tau) = (N-2)(N-3) f_2(\mathbf{r}_1, \mathbf{p}_1, \mathbf{r}_2, \mathbf{p}_2, \tau)\varphi(\mathbf{p}_3, \tau)\varphi(\mathbf{p}_4, \tau).$$

Using this result (18) must vanish identically because of either the $\mathbf{p}_3$ or $\mathbf{p}_4$ integration.

(b) Those terms in the sums for which only one pair of indices are the same, that is; $i = k$ and $j \neq i \neq l \neq j$ and three other combinations substituting j for i and/or l for k. All of these terms vanish by exactly the same kind of argument as in part (a), namely there will exist in these contributions a factor of the form $\int [a(\mathbf{p}) - \langle a\rangle_\tau] \times \varphi(\mathbf{p}, \tau)\, d\mathbf{p} = 0$. However, terms of the form $k = l$, $i \neq k \neq j$ give a nonzero contribution of

$$\int |K(\mathbf{r}_1, \mathbf{p}_1, \mathbf{r}_2, \mathbf{p}_2)|\, [a(\mathbf{p}_3) - \langle a\rangle_\tau]^2 \times f_3(\mathbf{r}_1, \mathbf{p}_1 \cdots \mathbf{r}_3, \mathbf{p}_3, \tau)\, d\mathbf{r}_1 \cdots d\mathbf{p}_3. \quad (19)$$

If we do the r_3 integration and again use (6) and (7) as in part (a), we can put (19) in the form

$$N(N-2)k(\tau)[\langle a^2\rangle_\tau - \langle a\rangle_\tau^2], \quad (20)$$

where we have also used (16) and (17).

(c) Those terms with two pairs of indices equal. There are two possibilities, $i = k \neq l = j$ and $i = l \neq k = j$, and together they contribute

$$2\int |K(\mathbf{r}_1, \mathbf{p}_1, \mathbf{r}_2, \mathbf{p}_2)|\, [a(\mathbf{p}_1) - \langle a\rangle_\tau][a(\mathbf{p}_2) - \langle a\rangle_\tau] \times f_2(\mathbf{r}_1, \mathbf{p}_1, \mathbf{r}_2, \mathbf{p}_2, \tau)\, d\mathbf{r}_1 \cdots d\mathbf{p}_2.$$

Because K vanishes for $|\mathbf{r}_1 - \mathbf{r}_2|$ greater than the range of intermolecular force, this contribution is proportional to Ω.

(d) Those terms with three indices the same. The possibilities are $i \neq j = k = l$ and $j \neq i = k = l$. The contribution is

$$2\int |K(\mathbf{r}_1, \mathbf{p}_1, \mathbf{r}_2, \mathbf{p}_2)| \times [a(\mathbf{p}_1) - \langle a\rangle_\tau]^2 f_2(\mathbf{r}_1, \mathbf{p}_1, \mathbf{r}_2, \mathbf{p}_2, \tau)\, d\mathbf{r}_1 \cdots d\mathbf{p}_2.$$

This contribution is proportional to Ω for the same reason as in (c).

If we collect all of these contributions together we obtain

$$D(\tau) \leq 2N^2k(\tau)[\langle a^2\rangle_\tau - \langle a^2\rangle_\tau] + O(N). \quad (21)$$

Where we have indicated those terms proportional to N or Ω by $O(N)$. We will assume them to be negligible compared to the term proportional to N^2.

Using (21), (9), and (12) in (1) we obtain

$$P(\gamma\langle A\rangle_0, \tau) \leq 4(N\gamma^2)^{-1}[\langle a^2\rangle_0\langle a\rangle_0^{-2} - 1] + 16(N\gamma^3)^{-1} \times \int_0^\tau k(t)[\langle a^2\rangle_t\langle a\rangle_0^{-3} - \langle a\rangle_t^2\langle a\rangle_0^{-3}]\, dt + O(N^{-2}). \quad (22)$$

All of the quantities appearing in this expression are intensive, i.e; independent of the size of the system for large systems, except for N, therefore $P(\alpha, \tau)$ can be made arbitrarily small for large-enough systems.

This result can only be regarded as suggestive. In any macroscopic system N, although large, is finite and therefore one should compute the right-hand side of (22) for a specific system. One also must verify that the terms of order $1/N^2$ are negligible. We shall leave this argument now and proceed to an example of a specific nature.

III. FLUCTUATIONS OF KINETIC ENERGY

In Ref. 1, a calculation of the bound for $P(\alpha, \tau)$ was done for the case where $f(x)$ was a canonical ensemble and $A(x)$ was the kinetic energy. In this calculation, as in Sec. II of this paper, the bound

was seen to vary as $1/N$ for large N. We wish to show that a smaller bound, which decreases much more rapidly with increasing N can be found for this specific example. In order to do this we repeat the calculation of Ref. 1 using different techniques.

For this calculation we use the bound on $P(\alpha, \tau)$ given by (36), (37), and (39) of Ref. 1, that is

$$P(\alpha, \tau) \leq (\alpha - \alpha_0)^{-1} \times \left[\int_{\alpha_0}^{\infty} C'(\alpha')\, d\alpha' + \int_0^{\tau} D'(\alpha_0, \tau')\, d\tau'\right]. \quad (23)$$

Where α_0 is any number less than α, and

$$C'(\alpha) = \int_{|A(x)-\langle A\rangle_0|>\alpha} f(x)\, dV, \quad (24a)$$

$$D'(\alpha, \tau) = \int_{|A(x)-\langle A\rangle_\tau|>\alpha} |\{A(x), H(x)\} - \langle\{A(x), H(x)\}\rangle_\tau|\, f(x, \tau)\, dV. \quad (24b)$$

We shall take

$$f(x) = z^{-1}e^{-\beta H(x)}, \quad (25a)$$

$$z = \left(\frac{\beta}{2\pi m}\right)^{\frac{3}{2}N} \int e^{-\beta V}\, d\mathbf{r}_1 \cdots d\mathbf{r}_N,$$

$$H(x) = K + V = \sum_i p_i^2/2m + \tfrac{1}{2}\sum_{i,j}{}' V(|\mathbf{r}_i - \mathbf{r}_j|), \quad (25b)$$

$$A(x) = K = \sum_i p_i^2/2m, \quad (25c)$$

$$\{A(x), H(x)\} = -\sum_{i,j}{}' \mathbf{F}_i(|\mathbf{r}_i - \mathbf{r}_j|)\cdot\mathbf{p}_i/m, \quad (25d)$$

$$\mathbf{F}_i(|\mathbf{r}_i - \mathbf{r}_j|) = \partial V(|\mathbf{r}_i - \mathbf{r}_j|)/\partial \mathbf{r}_i, \quad (25e)$$

$$\langle A\rangle = 3N/2\beta, \quad (25f)$$

$$\alpha = \gamma\langle A\rangle, \qquad \alpha_0 = \gamma_0\langle A\rangle, \qquad \gamma_0 < \gamma. \quad (25g)$$

Equations (24) can be written

$$C'(\alpha) = z^{-1}\int_{|K-\langle K\rangle|>\alpha} e^{-\beta H(x)}\, dV \quad (26a)$$

$$D'(\alpha, \tau) = z^{-1}\int_{|K-\langle K\rangle|>\alpha} |\sum_{i,j}{}' \mathbf{F}_i(|\mathbf{r}_i - \mathbf{r}_j|)\cdot\mathbf{p}_i/m|\, e^{-\beta H(x)}\, dV. \quad (26b)$$

In both (26a) and (26b) the integration goes over all of configuration space since K involves only the momenta. In (26a) the coordinate integration just cancels a factor in Z^{-1} and we are left with

$$C'(\alpha) = \left(\frac{\beta}{2\pi m}\right)^{\frac{3}{2}N}\int_{|K-\langle K\rangle|>\alpha} e^{-\beta K}\, d\mathbf{p}_1 \cdots d\mathbf{p}_N. \quad (27)$$

If we define the pair distribution function by

$$f_2(\mathbf{r}_1, \mathbf{r}_2) = N(N-1) \times \int e^{-\beta V}\, d\mathbf{r}_3 \cdots d\mathbf{r}_N \Big/ \int e^{-\beta V}\, d\mathbf{r}_1 \cdots d\mathbf{r}_N$$

and use the triangle inequality in (26b) we obtain

$$D'(\alpha, \tau) \leq z^{-1}\sum_{i,j}{}' \int_{|K-\langle K\rangle|>\alpha} |\mathbf{F}_i(|\mathbf{r}_i - \mathbf{r}_j|)\cdot\mathbf{p}_i/2m|\, e^{-\beta H(x)}\, dV$$
$$= (\beta/2\pi m)^{\frac{3}{2}N}(N^2 - N)^{-1} \times \sum_{i,j}{}' \int_{|K-\langle K\rangle|>\alpha}\left[\int \left|\mathbf{F}_i(|\mathbf{r}_i - \mathbf{r}_j|)\cdot\mathbf{p}_i/m\right| \times f_2(\mathbf{r}_i, \mathbf{r}_j)\, d\mathbf{r}_i\, d\mathbf{r}_j\right] e^{-\beta K}\, d\mathbf{p}_1 \cdots d\mathbf{p}_N. \quad (28)$$

If the intermolecular potential is spherically symmetic, and of finite range, and if Ω is large then f_2 is a function of $|\mathbf{r}_i - \mathbf{r}_j|$ alone. In this case

$$\int f_2(\mathbf{r}_i, \mathbf{r}_j)\left|\mathbf{F}_i(|\mathbf{r}_i - \mathbf{r}_j|)\cdot\mathbf{p}_i/m\right| d\mathbf{r}_i\, d\mathbf{r}_j = 2\pi\Omega p_i Q/m, \quad (29)$$

where

$$Q = \int_0^{\infty} |\mathbf{F}(r)|\, f_2(r) r^2\, dr. \quad (30)$$

Combining (30) and (28) we have

$$D'(d, \tau) \leq (\beta/2\pi m)^{\frac{3}{2}N}[2\pi\Omega Q/m] \times \int_{|K-\langle K\rangle|>\alpha} p_1 e^{-\beta K}\, d\mathbf{p}_1 \cdots d\mathbf{p}_N. \quad (31)$$

We see that both (27) and (31) involve integrals over only part of momentum space. One can remove the restriction on the region of integration by using the Dirchlet discontinuous factor.[3] We consider the function

$$S_\alpha(x) = 1 \quad \text{for} \quad |x| < \alpha,$$
$$S_\alpha(x) = 0 \quad \text{for} \quad |x| \geq \alpha, \quad (32)$$

which has the integral representation

$$S_\alpha(x) = (2\pi i)^{-1} \times \int_{-\infty}^{+\infty} y^{-1}[e^{-iy(x-\alpha)} - e^{-iy(x+\alpha)}]\, dy. \quad (33)$$

Now let us consider the integral which appears in (27),

[3] See, for example, C. Kittel, *Elementary Statistical Physics* (John Wiley & Sons, Inc., New York, 1958), Appendix B.

$$\int_{|K-\langle K\rangle|>\alpha} e^{-\beta K}\, d\mathbf{p}_1 \cdots d\mathbf{p}_N = \int_{-\infty}^{+\infty} e^{-\beta K}\, d\mathbf{p}_1 \cdots d\mathbf{p}_N \qquad (34)$$

$$-\int_{|K-\langle K\rangle|<\alpha} e^{-\beta K}\, d\mathbf{p}_1 \cdots d\mathbf{p}_N = \int_{-\infty}^{+\infty} e^{-\beta K}\, d\mathbf{p}_1 \cdots d\mathbf{p}_N$$

$$-\int_{-\infty}^{+\infty} S_\alpha(K-\langle K\rangle) e^{-\beta K}\, d\mathbf{p}_1 \cdots d\mathbf{p}_N = I_1 - I_2.$$

I_1 is a standard integral,

$$I_1 = (2\pi m/\beta)^{\frac{3}{2}N}. \qquad (35)$$

To treat I_2 we use (33) and assume we can interchange the order of integration

$$I_2 = (2\pi i)^{-1} \int_{-\infty}^{+\infty} y^{-1} \Big\{ \int e^{-\beta K} [e^{-iy(K-\langle K\rangle-\alpha)} - e^{-iy(K-\langle K\rangle+\alpha)}]\, d\mathbf{p}_1 \cdots d\mathbf{p}_N \Big\}\, dy. \qquad (36)$$

Since $\langle K\rangle$ and α are not functions of the momenta, we need only evaluate

$$\int e^{-(\beta+iy)K}\, d\mathbf{p}_1 \cdots d\mathbf{p}_N = \left[4\pi \int_0^\infty e^{-(\beta+iy)p^2/2m} p^2\, dp\right]^N = \left(\frac{2\pi m}{\beta+iy}\right)^{\frac{3}{2}N}.$$

With this result I_2 becomes

$$I_2 = (2\pi i)^{-1} \int_{-\infty}^{+\infty} y^{-1} \left(\frac{2\pi m}{\beta+iy}\right)^{\frac{3}{2}N} \times \{e^{iy(\langle K\rangle+\alpha)} - e^{iy(\langle K\rangle-\alpha)}\}\, dy. \qquad (37)$$

This integral can be evaluated exactly by the method of residues, however, the result comes out in series form and is inconvenient for large N. Since we are only interested in large values of N a more manageable result is obtained by the method of steepest descent. To do this let us deform the path of integration in (37) so that it passes below the origin in the complex y plane (Fig. 1) and write I_2 as the sum of two contributions

$$I_2 = I_2^+ - I_2^-,$$

where

$$I_2^\pm = (2\pi i)^{-1} \int_C \left(\frac{2\pi m}{\beta+iy}\right)^{\frac{3}{2}N} e^{iy(\langle K\rangle\pm\alpha)} y^{-1}\, dy. \qquad (38)$$

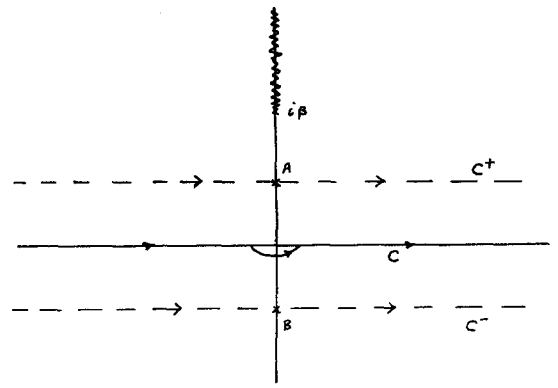


FIG. 1. Complex y plane.

In the integrand there is a simple pole at the origin and a multiple pole at $y = i\beta$. If $3N/2$ is not integer there will be a cut starting at $y = i\beta$ also.

Using (25f) and (25g) one can write

$$I_2^\pm = (2\pi i)^{-1} \int_C y^{-1} \exp\left\{\tfrac{3}{2}N\left[\ln\left(\frac{2\pi m}{\beta+iy}\right)^{\frac{3}{2}N} + iy(1\pm\gamma)/\beta\right]\right\} dy. \qquad (39)$$

For large N, the saddle-point method gives an approximate value for these integrals. The saddle points occur at $i\beta\gamma(\gamma\pm1)^{-1}$ and since $\gamma > 0$ both of these points lie below $i\beta$ on the imaginary axis (points A and B of Fig. 1). Now let C^+ and C^- be paths passing through points A and B and parallel to the real axis. Let us consider I^+, with saddle point $i\beta\gamma(\gamma+1)^{-1}$. For all positive values of γ the integrand vanishes at $y = \pm\infty$ therefore by Cauchy's theorem we can set the integral along C equal to the integral along C^+ plus the contributions from the pole at the origin. The integral along C^+ can be done by the saddle-point method. Without going into the details,

$$I_2^+ = (2\pi m/\beta)^{\frac{3}{2}N}\{1 - (3\pi N\gamma^2)^{-\frac{1}{2}}\} \times \exp\{-\tfrac{3}{2}N[\ln(1+\gamma) - \gamma]\}. \qquad (40)$$

Let us consider I^- for the case where $\gamma < 1$. The saddle point is $i\beta\gamma(\gamma-1)^{-1}$ which is below the origin (point B in Fig. 1). Since the integrand vanishes at $y = \pm\infty$, the integrals along C and C^- are the same and again by saddle-point integration we obtain

$$I_2^- = (2\pi m/\beta)^{\frac{3}{2}N}(3\pi N\gamma^2)^{-\frac{1}{2}} \times \exp\{\tfrac{3}{2}N[\ln(1-\gamma) + \gamma]\}. \qquad (41)$$

For $\gamma > 1$, $I_2^- = 0$ and this can be seen in (39). Since $1 - \gamma < 0$, the contour C can be closed in an

infinite semicircle below the real axis. As there are no singularities in this region the integral must vanish. We can therefore write for any γ

$$I_2^- = (2\pi m/\beta)^{\frac{3}{2}N}(3\pi N\gamma^2)^{-\frac{1}{2}} \times \exp\{-\tfrac{3}{2}N[\ln(1-\gamma)+\gamma]\}\theta(1-\gamma). \quad (42)$$

Where $\theta(x) = 0$ for $x < 0$ and $\theta(x) = 1$ for $x > 0$. Combining (42), (40), (38), (35), (34), and (27) we obtain

$$C'(\alpha) = (3\pi N\gamma^2)^{-\frac{1}{2}}\{\exp(\tfrac{3}{2}N[\ln(1+\gamma)-\gamma]) + \theta(1-\gamma)\exp(\tfrac{3}{2}N[\ln(1-\gamma)+\gamma])\} \quad (43)$$

where $\alpha = \gamma 3N/2\beta$.

By the same technique the integral appearing in (31) can be evaluated. We give only the result for the case where $\gamma_0 \ll 1$ and we have set $\ln(1 \pm \gamma_0) \mp \gamma_0 = \frac{1}{2}\gamma_0^2$

$$D(\gamma_0\langle K\rangle, \tau) \leq 2\Omega Q(\tfrac{3}{2}Nm\beta\gamma_0^2)^{-\frac{1}{2}}\exp(-3N\gamma_0^2/4) \quad (44)$$

From (23) we see that it remains to integrate $C'(\alpha)$ given by (43), from α_0 to ∞.

$$\int_{\alpha_0}^{\infty} C'(\alpha)\,d\alpha = \frac{3N}{2\beta}\int_{\gamma_0}^{\infty} C'(\gamma\langle K\rangle)\,d\gamma = [(3\pi N)^{-\frac{1}{2}}3N/2\beta] \times \left[\int_{\gamma_0}^{\infty}\exp\{\tfrac{3}{2}N[\ln(1+\gamma)-\gamma]\}\gamma^{-1}\,d\gamma + \int_{\gamma_0}^{1}\exp\{\tfrac{3}{2}N[\ln(1-\gamma)+\gamma]\}\gamma^{-1}\,d\gamma\right]. \quad (45)$$

Instead of attempting an exact evaluation of these integrals we shall simply find upper bounds for them assuming $0 < \gamma_0 \ll 1$. In the second integral the integrand is a positive decreasing function of γ in the interval γ_0 to 1, therefore

$$\int_{\gamma_0}^{1}\exp\{\tfrac{3}{2}N[\ln(1-\gamma)+\gamma]\}\gamma^{-1}\,d\gamma < \gamma_0^{-1}\exp\tfrac{3}{2}N[\ln(1-\gamma_0)+\gamma_0]. \quad (46)$$

For the first integral we have

$$\int_{\gamma_0}^{\infty}\exp\{\tfrac{3}{2}N[\ln(1+\gamma)-\gamma]\}\gamma^{-1}\,d\gamma < \gamma_0^{-1}\int_{\gamma_0}^{\infty}(1+\gamma)^{\frac{3}{2}N}e^{-\frac{3}{2}N\gamma}\,d\gamma = \gamma_0^{-1}e^{\frac{3}{2}N}\int_{1+\gamma_0}^{\infty}s^{\frac{3}{2}N}e^{-\frac{3}{2}Ns}\,ds = -\gamma_0^{-1}e^{M}\sum_{r=0}^{M}(-1)^r(1+\gamma_0)^{M-r} \times e^{-M(1+\gamma_0)}M!/(-M)^{r+1}(M-r)!.$$

Where we have taken $M = \frac{3}{2}N$ to be integer. Since $M!/M^r(M-r)! < 1$, we can write

$$\int_{\gamma_0}^{\infty}\exp\{\tfrac{3}{2}N[\ln(1+\gamma)-\gamma]\}\gamma^{-1}\,d\gamma < (M\gamma_0)^{-1}e^{-M\gamma_0}\sum_{r=0}^{M}(1+\gamma_0)^{M-r} = (M\gamma_0)^{-1}e^{-M\gamma_0}(1+\gamma_0)^M\left\{\frac{1-(1+\gamma_0)^{-M}}{1-(1+\gamma_0)^{-1}}\right\} \approx (M\gamma_0^2)^{-1}e^{-M\gamma_0}(1+\gamma_0)^{M+1}. \quad (47)$$

The sum of (46) and (47) can be written, for $\gamma_0 \ll 1$,

$$\gamma_0^{-1}[1 + (M\gamma_0)^{-1}]\exp(-M\gamma_0^2/2),$$

so (45) becomes

$$\int_{\alpha_0}^{\infty} C'(\alpha)\,d\alpha \leq [(3\pi N)^{-\frac{1}{2}}3N/2\beta] \times \gamma_0^{-1}[1 + 2/3N\gamma_0]\exp(-3N\gamma_0^2/4). \quad (48)$$

We are interested in the case where $N\gamma_0^2 \gg 1$. We can drop the $\frac{2}{3}N\gamma_0$ term in (48); then (23) becomes

$$P(\gamma\langle K\rangle, \tau) < \{(3\pi N)^{-\frac{1}{2}} + (4/3)(3N^3m/2\beta)^{-\frac{1}{2}} \times \Omega Q\tau\}(\gamma-\gamma_0)^{-1}\gamma_0^{-1}\exp(-3N\gamma_0^2/4). \quad (49)$$

There is no doubt that the bound (49) could be reduced by more careful evaluation of some of the integrals involved in this treatment. One can also pick γ_0 to minimize (49). For $N\gamma^2 \gg 1$ this is given to lowest order in $(N\gamma^2)^{-1}$ by

$$\gamma_0 = \gamma(1 - 1/3N\gamma^2),$$

and with this value of γ_0 (49) becomes very nearly

$$P(\gamma\langle K\rangle, \tau) < \{(3\pi N)^{-\frac{1}{2}} + (\tfrac{4}{3})(3N^3m/2\beta)^{-\frac{1}{2}} \times \Omega Q\tau\}3N\exp(-\tfrac{3}{2}N\gamma^2 + 1). \quad (50)$$

The result (50) shows the exponential decrease of the bound with increasing N, in contrast to the result in Ref. 1 where the bound varied as $1/N$. We have assumed in the derivation of (50) that $N\gamma^2 \gg 1$ and $\gamma \ll 1$. Actually the replacement of $\exp\{\frac{3}{2}N[\ln(1\pm\gamma)\mp\gamma]\}$ by $\exp(-3N\gamma^2/4)$ requires also that $N\gamma^3 \gg 1$. If γ does not satisfy these conditions one still can write a formula similar to (50) but involving the functions $\exp\{\frac{3}{2}N[\ln(1\pm\gamma)\mp\gamma]\}$.

ACKNOWLEDGMENT

We would like to thank Professor Walter Johnson for helpful discussions on some integration techniques.