JOURNAL OF
MATHEMATICAL
PHYSICS

VOLUME 6, NUMBER 10 OCTOBER 1965

Microcanonical Ensemble in Quantum Statistical Mechanics*

ROBERT B. GRIFFITHS
Department of Physics, Carnegie Institute of Technology, Pittsburgh, Pennsylvania
(Received 8 February 1965)

The infinite-volume limit of thermodynamic functions calculated in the quantum microcanonical ensemble is shown to exist for a fairly wide class of spin systems and quantum gases. The entropy is set equal to the logarithm of the number of eigenstates in an energy interval which increases linearly with the size of the system, but is otherwise arbitrary. The limiting entropy per unit volume agrees with that calculated in the canonical formalism, and possesses certain convexity properties required for thermodynamic stability. A precise criterion, in terms of the energy spectra of large systems, is given for determining the limit of the thermodynamic entropy as the temperature approaches zero. This is not determined by the degeneracy of the ground state, contrary to the discussion of the "third law of thermodynamics" found in some textbooks.

I. INTRODUCTION

A NATURAL question is: "What distribution of energy levels for a large system gives rise to a phase transition as the temperature is varied?"[1] One answer is obtained by calculating the partition function

$$Z = \sum_i \exp(-E_i/T) \tag{1}$$

where E_i is the energy of the ith level and $T(>0)$ the absolute temperature in units such that Boltzmann's constant is unity. Phase transitions appear as discontinuities[2] in the first-, second-, or higher-order derivatives with respect to temperature of the free energy

$$F = -T \log Z. \tag{2}$$

However, F is an analytic[3] function of T over the range of temperature for which the sum (1) converges, and thus discontinuities in the derivatives only appear if one considers the limiting function

$$f(T) = \lim_{V\to\infty} F_V(T)/V \tag{3}$$

obtained by dividing F by some extensive parameter, for example the volume V of the system, and taking the limit of an infinite system in a suitable way. Proofs[4-7] that the limit (3) exists for quantum systems of any complexity[8] have only appeared

* Supported in part by the National Science Foundation.

[1] First brought to the author's attention by Professor Robert Rhodes.

[2] We shall refer to a discontinuity in the first derivative of the free energy with respect to temperature as a "first-order phase transition"; if the first derivative is continuous but the second is discontinuous or approaches infinity, we call it a "second-order phase transition", etc.

[3] Because the series (1) is uniformly convergent in the interior of region where it converges. See E. C. Titchmarsh, *The Theory of Functions* (Oxford University Press, New York, 1939), 2nd ed., p. 95.

[4] D. Ruelle, Helv. Phys. Acta **36**, 789 (1963).

[5] M. E. Fisher, Arch. Ratl. Mech. Anal. **17**, 377 (1964).

[6] R. B. Griffiths, J. Math. Phys. **5**, 1215 (1964).

[7] Classical systems have been considered by C. N. Yang and T. D. Lee, Phys. Rev. **87**, 404 (1952); D. Ruelle, Helv. Phys. Acta **36**, 183 (1963).

[8] If a system consists of a large number of noninteracting subsystems, or an aggregate of noninteracting particles, the proofs become much simpler. This is the case considered, for example, by A. Y. Khinchin, Mathematical Foundations of Quantum Statistics (Graylock Press, Albany, New York, 1960). See also J. van der Linden and P. Mazur, Physica **27**, 609 (1961).

very recently. The relation between $f(T)$ and the limiting distribution of energy levels for a large system is not particularly transparent since the limiting procedure (3) is preceded by a Laplace transform (1).

The quantum microcanonical ensemble provides a more direct approach in answering the question posed above. It is presented in many textbooks[9] somewhat as follows: Let n_E be the number of levels in the energy interval $[E, E + \Delta E]$. Define the entropy S as

$$S(E) = \log n_E. \tag{4}$$

Once $S(E)$ is known, other thermodynamic quantities are obtained by differentiation; for example,

$$T^{-1} = dS/dE. \tag{5}$$

Equation (4), though it provides a direct connection between thermodynamic quantities and the energy spectrum, is logically unsatisfactory in at least two respects. First, S depends on the magnitude of ΔE. This is a very weak dependence for a large system. Nevertheless, an uncomfortable ambiguity remains in the definition. Second, $S(E)$ is a discontinuous function of E, since n_E can change only by integral values. A "smoothing" operation of some sort must be applied to $S(E)$ before taking derivatives, as in (5). The *average* density of levels for a system of macroscopic size is very large, but there is no reason to expect *a priori* that the distribution is smooth on the scale of the ΔE we have arbitrarily chosen.

Both of these conceptual difficulties disappear, as we shall demonstrate, when one takes the infinite-volume ("thermodynamic") limit in a suitable fashion. As a byproduct, the question posed at the beginning of this introduction receives a precise answer.

One sometimes encounters the following objection to computing thermodynamic functions by the limiting procedure (3): "Why should experimental measurements, always performed on *finite* systems, be compared with results calculated for the limiting case of an *infinite system*?" The objection is best answered by considering a particular thermal experiment. The heat capacity of a crystal is measured by heating it in a calorimeter. When the experiment is complete, the investigator divides the heat capacity by the mass of the crystal and publishes the result as a specific heat in calories/gram°K or other suitable units. He does this because he is confident that, under the conditions of the experiment and within the accuracy of the measurement, the heat capacity is proportional to the mass but otherwise independent of the size and shape of the crystal. In brief, the heat capacity is extensive. Naturally, this confidence can be (and sometimes is) checked through measurements on samples of various sizes and shapes. The theoretical quantity to compare with the experimental result is clearly that portion of the heat capacity strictly proportional to the size of a system for a large system. The convenient mathematical procedure for obtaining this is the limiting process (3). If the limit exists, one has a sensible quantity to compare with experiment.

Various definitions of the microcanonical ensemble are introduced in Sec. II. Section III contains a proof that the limits exist for a spin system. The case of a quantum gas of bosons or fermions is treated in less detail in Sec. IV. That the various definitions in Sec. II lead to identical functions in the thermodynamic limit is shown in Sec. V, and that these agree with those calculated in the canonical formalism is shown in Sec. VI. As T approaches zero, the energy always approaches the ground-state energy, but the limiting entropy does *not* necessarily bear any relation to the degeneracy of the ground state, contrary to the statements found in many textbooks. These matters are discussed in Sec. VII.

II. DEFINITIONS OF THE ENTROPY, AND A THEOREM

Let H be a Hamiltonian with a discrete spectrum bounded from below, and let $\mu(E)$ be the number of eigenstates with energy not exceeding E[10]. Define the entropy S by

$$S(E) = \log \mu(E) = \log \mathrm{Tr}\,[\theta(E - H)] \tag{6a}$$

where Tr stands for trace and $\theta(x)$ is 1 for $x \geq 0$ and 0 for $x < 0$. For any $\delta > 0$ define

$$S^-(\delta; E) = \log n^-(\delta; E), \tag{6b}$$

where $n^-(\delta; E)$ is the number of levels in the interval $[E - V\delta, E]$. Here V is the volume or some other parameter (such as the number of particles) which gives a measure of the size of the system.

[9] For example, K. Huang, *Statistical Mechanics* (John Wiley & Sons, Inc., New York, 1963), p. 188; L. D. Landau and E. M. Lifschitz, *Statistical Physics* (Pergamon Press, Ltd., London, 1958), p. 22; C. Kittel, *Elementary Statistical Physics* (John Wiley & Sons, Inc., New York, 1958), p. 18; J. E. Mayer and M. G. Mayer, *Statistical Mechanics* (John Wiley & Sons, Inc., New York, 1940), p. 53.

[10] For all systems we wish to consider, $\mu(E)$ is finite for all finite values of E. See Refs. 4 and 5.

The inverse[11] function to $\mu(E)$, $\xi(n)$, is defined as follows. Let the normalized eigenstates ϕ_i of H, $i = 1, 2, 3, \cdots$, be numbered in the order of increasing energy (within a degenerate multiplet, the order may be chosen arbitrarily), and let $\xi(n)$ be the energy of the state ϕ_n. For x between $n-1$ and n, $\xi(x)$ is equal to $\xi(n)$.

Define the normalized quantities

$$\sigma = V^{-1}S; \qquad \epsilon = V^{-1}E \tag{7}$$

in terms of which (6a) may be written as

$$\sigma(\epsilon) = V^{-1} \log \mu(V\epsilon) \tag{8a}$$

or the inverse[11] function

$$\epsilon(\sigma) = V^{-1}\xi(e^{V\sigma}). \tag{8b}$$

The function $\sigma(\epsilon)$ is only defined for $V\epsilon$ greater than or equal to the ground-state energy, whereas $\epsilon(\sigma)$ is defined for $-\infty < \sigma < \sigma_m$. For a quantum gas, σ_m is $+\infty$, and for a spin system it is some constant less than infinity. For $\sigma \leq 0$, $V\epsilon(\sigma)$ is equal to the ground-state energy and hence is not very interesting.

The function $\sigma^-(\delta; \epsilon)$ is defined in analogous fashion to $\sigma(\epsilon)$. We shall show (in Sec. V) that σ and σ^- are identical in the thermodynamic limit.

The entropy in (6a) may also be written as

$$S(E) = -\mathrm{Tr}\,[\rho \log \rho] \tag{9a}$$

where the density matrix ρ is given by

$$\rho_{ii} = (\phi_i, \rho\phi_i) = [\mu(E)]^{-1} \tag{9b}$$

for $i \leq \mu(E)$, and all other matrix elements are zero. Let

$$\bar{E} = \mathrm{Tr}\,[\rho H] \tag{10a}$$

denote the average energy associated with this density matrix. The corresponding normalized quantity is

$$\bar{\epsilon}(\sigma) = V^{-1}\bar{E} = (VN)^{-1} \sum_{n=1}^{N} \xi(n), \tag{10b}$$

where N is the smallest integer not less than $e^{V\sigma}$.

An analogous density matrix corresponds to the definition (6b). We shall not in this paper make any use of the average energy (10b). For $\sigma < \sigma_m$ it coincides with $\epsilon(\sigma)$ in the thermodynamic limit (see Appendix F) for the systems considered in Secs. III and IV.

[11] We define the "inverse" of a discontinuous, monotone function as follows. Draw the graph of the function and connect the points of discontinuity with vertical lines. This is the graph of the inverse function if ordinate and abscissa are interchanged. The inverse of a function continuous to the right is continuous to the left, and vice versa.

The proofs in Sec. III utilize the following theorem: Let H be a self-adjoint operator in a finite-dimensional vector space $\mathfrak{V}$. If

$$(\phi, H\phi) \leq E \tag{11}$$

for every normalized ϕ in an m-dimensional linear subspace $\mathfrak{M}$ of $\mathfrak{V}$, then

$$\mu(E) \geq m \tag{12}$$

or the equivalent:

$$\xi(m) \leq E. \tag{13}$$

Since this is merely one form of the "minimax principle",[12] we shall not give the proof here. Section IV requires an extended form of the theorem in which H is a self-adjoint operator with a discrete spectrum bounded from below, and $\mathfrak{V}$ is an infinite-dimensional Hilbert space. The theorem is still valid provided the (finite-dimensional) subspace $\mathfrak{M}$ lies within the domain where H is defined.[13]

III. THE LIMITING ENERGY FOR A SPIN SYSTEM

The Heisenberg model of ferromagnetism with Hamiltonian

$$H = -2J \sum_{(ij)} \mathbf{S}_i \cdot \mathbf{S}_j, \tag{14}$$

where S_i is the spin operator for the ith atom and the sum extends over all nearest-neighbor pairs of atoms in a regular lattice, provides a typical example of a "spin system"[14]: note that only the spin degree of freedom is considered.

We shall require that an acceptable Hamiltonian H for a spin system

(a) have the translational symmetry of the lattice

(b) consist of a sum of terms, each a Hermitian operator (all matrix elements finite) involving a group of spins[15] no two of which are separated by a distance greater than r (where r does not depend on the term considered)

(c) contain only a finite number of terms involving a given spin.

The norm $|A|$ of a Hermitian operator A shall be the maximum of the absolute values of its eigenvalues. It has the property

$$|A + B| \leq |A| + |B|. \tag{15}$$

[12] P. R. Halmos, *Finite-Dimensional Vector Spaces* (D. Van Nostrand, Inc., Princeton, New Jersey, 1958), 2nd ed., p. 181.

[13] N. Dunford and J. T. Schwartz, *Linear Operators* (John Wiley & Sons, Inc., New York, 1963), Vol. II, p. 1543.

[14] See Ref. 6 for a more precise definition.

[15] Since only the spin degree of freedom is considered, we employ interchangeably the terms "atom" and "spin."

Let h_i for $i = 1,2, \cdots, n$ [n is finite by (c)] be all the terms in H which involve a particular, say the pth, spin. By (15) we have

$$|\sum_i h_i| \leq C = \sum_i |h_i|. \tag{16}$$

It is clear from (16) that if H' involves at most M spins and is the sum of certain terms in H, then

$$|H'| \leq CM. \tag{17}$$

In particular (setting $H' = H$), all the eigenvalues of H for a lattice containing V spins lie in the interval $[-CV, CV]$.

A. The Basic Inequality

Consider two spin systems 1 and 2 (for example, two halves of a cube) which together constitute a system with Hamiltonian

$$H = H_1 + H_2 + H', \tag{18}$$

where H_1 is the sum of all terms in H involving only the spins in system 1, H_2 is similarly defined for system 2, and H' is the sum of all terms which simultaneously involve spins in both 1 and 2. Let $\psi_1, \psi_2, \cdots, \psi_n$ be the eigenfunctions of H_1 associated with the n lowest eigenvalues, and let $\chi_1, \chi_2, \cdots, \chi_m$ be a similar set for H_2. The product functions $\psi_i\chi_j$ span a subspace $\mathfrak{M}$ of dimension nm. If ξ_1 and ξ_2 (defined in Sec. II) refer to H_1 and H_2, respectively, the inequality

$$(\phi, H\phi) \leq \xi_1(n) + \xi_2(m) + (\phi, H'\phi) \tag{19}$$

holds for any normalized ϕ in $\mathfrak{M}$. Now $(\phi, H'\phi)$ cannot exceed $|H'|$, and thus by the theorem of Sec. II (ξ refers to H),

$$\xi(nm) \leq \xi_1(n) + \xi_2(m) + |H'|. \tag{20}$$

This inequality holds for n, m integers and therefore also (see the definition of ξ in Sec. II) when n (or m) is any real number greater than zero and less than or equal to the total number of levels for system 1 (or 2).

The generalization of (20) to a system consisting of l subsystems is

$$\xi\left(\prod_{i=1}^{l} n_i\right) \leq \sum_{i=1}^{l} \xi_i(n_i) + |H'|, \tag{21}$$

where H' includes all terms in the Hamiltonian involving spins in two or more subsystems. This inequality is the basis of the arguments below.

B. Limiting Energy for a Special Sequence

Consider a simple cubic lattice with lattice constant = 1. More general three-dimensional lattices may be reduced to this case by elementary manipulation.[6] The cube Ω_k for $k = 2, 3, 4, \cdots$ of volume V_k contains $V_k = 2^{3k}$ spins. It is made up of eight cubes of the type Ω_{k-1}. Let us apply (21) to Ω_k with $l = 8$, four of the n_i's equal to $\exp(V_{k-1}\sigma_1)$, and four equal to $\exp(V_{k-1}\sigma_2)$. In terms of normalized quantities (8b), (21) becomes

$$\epsilon_k(\tfrac{1}{2}\sigma_1 + \tfrac{1}{2}\sigma_2) \leq \tfrac{1}{2}\epsilon_{k-1}(\sigma_1) + \tfrac{1}{2}\epsilon_{k-1}(\sigma_2) + V_k^{-1}|H'|. \tag{22}$$

Since each term in H' involves spins in at least two of the smaller cubes, it can, by condition (b) above, involve only the spins within a distance r of the surface of one of the smaller cubes. There are at most $12r \cdot 2^{2k}$ such spins and therefore, by (17),

$$V_k^{-1}|H'| \leq C'2^{-k}, \tag{23}$$

where $C' = 12r\,C$ is a constant independent of k.

If we set $\sigma_1 = \sigma_2 = \sigma$ and insert (23) in (22), it is evident that

$$\epsilon_k(\sigma) + C'2^{-k} \tag{24}$$

is monotone decreasing as $k \to \infty$, and must, since bounded below by $-C$ [see the remark following (17)], approach a limit $\epsilon(\sigma)$. The inequality (22) in the limit $k \to \infty$ together with the fact that $\epsilon(\sigma)$ is bounded from above by (24) implies that $\epsilon(\sigma)$ is convex-downwards and continuous, except, perhaps, at the upper limit σ_m of the interval where ϵ is defined.[16] Thus $T = d\epsilon/d\sigma$ is defined except at most on a denumerable set of points, and is a monotone increasing function of σ. Further, $\epsilon(\sigma)$ is monotone increasing, since each ϵ_k has this property. Figure 1 shows its general appearance.

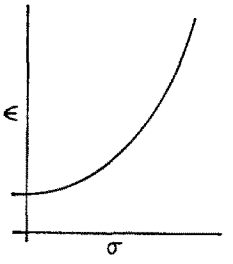


FIG. 1. The function $\epsilon(\sigma)$ in a typical case. For negative σ, ϵ is equal to its value at $\sigma = 0$.

The convergence of $\epsilon_k(\sigma)$ to $\epsilon(\sigma)$ is uniform on an interval $0 \leq \sigma \leq \sigma_1 < \sigma_m$. For suppose there were a $\delta > 0$ and an increasing sequence[17] $\{K\}$ such that

[16] Convex functions are discussed by G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities* (Cambridge University Press, New York, 1959), 2nd ed., Chap. III. By "convex" we mean *continuous* convex functions.

[17] "Increasing sequence" is an abbreviation for "a sequence of integers, not necessarily consecutive, increasing to infinity."

$$\epsilon_K(\sigma_K) \geq \epsilon(\sigma_K) + \delta \tag{25}$$

for σ_k in the interval converging to some σ'. The inequality (25) together with the monotonicity in σ of ϵ_k and the continuity of ϵ would then imply that for any $\sigma'' > \sigma'$, $\epsilon_K(\sigma'')$ would converge to a number $[\epsilon(\sigma'')]$ not less than $\epsilon(\sigma') + \delta$. This cannot occur, since $\epsilon(\sigma)$ is continuous. The possibility $\epsilon_K(\sigma_K) \leq \epsilon(\sigma_K) - \delta$ need not be considered since (24) is monotone decreasing in k.

C. A General Sequence

Next we shall show that an arbitrary sequence of cubes of increasing volume yields the same limit $\epsilon(\sigma)$ as the particular sequence Ω_k. If this were not true, we could find a sequence of cubes ω_L with volume V_L increasing to infinity and chosen so that either

$$\epsilon_L(\sigma) \geq \epsilon(\sigma) + \delta \tag{26}$$

or

$$\epsilon_L(\sigma) \leq \epsilon(\sigma) - \delta \tag{27}$$

for some $\delta > 0$.

Consider the case (26). Provided $2^{3k} < V_L$, we may imagine ω_L to be made up of m_k cubes of type Ω_k (close packed, starting in one corner of ω_L), plus a system ω_L^* of volume $V_L^* = V_L - m_k V_k$. The basic inequality (21), with appropriate n_i and written in terms of normalized quantities, becomes

$$\epsilon_L(\sigma) \leq \epsilon_k(\sigma) + V_L^* V_L^{-1}[\epsilon_L^*(\sigma) - \epsilon_k(\sigma)] + V_L^{-1}\,|H'|. \tag{28}$$

The second term on the right is bounded by $2C\,V_L^* V_L^{-1}$ and the third by $C'2^{-k}$. Provided V_L is sufficiently large we may, by a judicious choice of k, ensure that neither term exceeds $\delta/4$, and also that $|\epsilon_k(\sigma) - \epsilon(\sigma)| \leq \delta/4$. Then (28) contradicts (26).

By considering a cube Ω_k with $V_k \gg V_L$ as made up of m_L cubes of type ω_L plus a system Ω_k^*, we may prove, using the arguments above, that (27) also leads to a contradiction.

This completes the proof that the normalized energies for a sequence of cubes of increasing volume and for a Hamiltonian satisfying properties (a) to (c) converge to a well-defined limit. The result may be extended in various ways.

(i) Other shapes of domain. There is no difficulty in extending the result to (rectangular) parallelepipeds in which all three linear dimensions increase to infinity. Presumably Fisher's[5] techniques could be employed for more general domains.

(ii) Periodic boundary conditions. It is easily shown[18] that $\epsilon_k(\sigma)$ is altered at most by a term proportional to surface divided by volume if periodic boundary conditions are applied on a parallelepiped. This correction is negligible if the linear dimensions are large.

(iii) Interaction terms of unbounded range. The arguments previously employed in the canonical case[6] for a particular class of these interactions may be modified in an obvious fashion in the microcanonical case.

IV. THE LIMITING ENERGY FOR A QUANTUM GAS

We consider a gas of identical particles, either bosons or fermions, with a Hamiltonian of the form

$$H = -(\hbar^2/2m)\sum_i \Delta_i + \sum_{i<j} v(\mathbf{r}_i - \mathbf{r}_j), \tag{29}$$

where Δ_i is the Laplacian for the ith particle. The potential v may depend on internal coordinates (e.g., spin) of the interacting atoms, in which case these coordinates must also be included in the wavefunctions and summed over when taking inner products, etc.

There are some complications in the case of gases which are not encountered in spin systems: (a) The Hamiltonian is unbounded, and it requires some care to obtain a self-adjoint operator in the Hilbert space of square-integrable functions. Ruelle[4] has discussed the case of particles confined in a (cubic) box with "hard" walls. That is, the wavefunction vanishes at the walls. We restrict our considerations to this case. (b) The energy depends on the density of particles ρ as well as the entropy. (c) Some restriction must be placed upon the potential v, even if it is of finite range, to ensure *stability*, by which we mean that H has a lower bound proportional to the size of the system. It is sufficient to require that the total potential energy W for N particles satisfy

$$W(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N) \geq -CN \tag{30}$$

for all allowed values of the coordinates, where C is a constant independent of N. Conditions on v sufficient to insure stability have been published.[4,5,19]

In addition, restrictions must be placed on the long-range behavior of $v(r)$. For brevity of exposition, we shall use the condition employed by Ruelle[4,19]: there is a constant R_o such that

$$v(\mathbf{r}) \leq 0 \quad \text{for} \quad |\mathbf{r}| \geq R_o. \tag{31}$$

[18] See the analogous arguments in Ref. 6.
[19] D. Ruelle, Ref. 7.

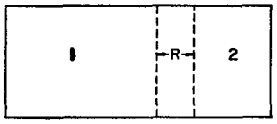


FIG. 2. Closed box divided into two subsystems.

A. The Basic Inequality

Consider a closed box (Fig. 2) divided into two boxes 1 and 2 by a partition of thickness $R \geq R_o$. The Hamiltonian H_1 of box 1 containing N particles with coordinates $\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N$ is defined using (29) plus the requirement that the wavefunction have the correct symmetry and that it vanish if any particle is located outside or on the boundary of the box. (We assume the box is of "reasonable" shape, so that a suitable self-adjoint operator may be defined.) Let $\psi_1, \psi_2, \cdots$ be the normalized eigenfunctions of H_1 in order of increasing energy. Similarly define H_2 for M particles in box 2 with eigenstates $\chi_1, \chi_2, \cdots$ functions of the particle coordinates $\mathbf{r}_{N+1}, \mathbf{r}_{N+2}, \cdots, \mathbf{r}_{N+M}$.

Consider a normalized function ϕ in the mn-dimensional subspace $\mathfrak{M}$ spanned by the product functions

$$\phi_{jk}(\mathbf{r}_1, \cdots, \mathbf{r}_{N+M}) = \psi_j\chi_k, \tag{32}$$

for $j = 1, 2, \cdots n$ and $k = 1, 2, \cdots m$. The Hamiltonian H (as a differential operator) for the entire box with partition removed is equal to

$$H = H_1 + H_2 + H', \tag{33}$$
$$H' \equiv \sum_{i\leq N} \sum_{j\geq N+1} v(\mathbf{r}_i - \mathbf{r}_j),$$

whence we obtain the result

$$(\phi, H\phi) = (\phi, H_1\phi) + (\phi, H_2\phi) + (\phi, H'\phi). \tag{34}$$

The function ϕ is nonzero only when the first N coordinates lie in box 1 and the last M in box 2. Under these conditions H' is negative, and thus we conclude that

$$(\phi, H\phi) \leq (\phi, H_1\phi) + (\phi, H_2\phi) \leq \xi_1(n) + \xi_2(m), \tag{35}$$

where ξ_1, ξ_2 refer to H_1 and H_2, respectively.

Two problems remain before we may apply the theorem of Sec. II. In the first place, ϕ is not symmetric (antisymmetric) in all the particle coordinates. The normalized functions ϕ^s_{jk} of proper symmetry constructed by permuting the coordinates in $\psi_j\chi_k$, summing the permutations with appropriate sign, and multiplying the result by a suitable factor[20] span an mn-dimensional subspace $\mathfrak{M}^s$. It is easily verified that for any φ^s a linear combination of the φ^s_{jk},

$$(\phi^s, H\phi^s) = (\phi, H\phi), \tag{36}$$

where ϕ is the corresponding linear combination of the ϕ_{jk}.

A second problem is that ϕ^s will not, in general, lie within the domain of definition of the self-adjoint operator H due to discontinuities in the first derivatives at the boundaries of the partition. To get around this difficulty, one may (we shall not give the details here) approximate the ϕ_{jk} with twice continuously differentiable functions in imitation of Ruelle's discussion on p. 791 of Ref. 4.

Let ξ refer to H; we conclude from (36) and (35) that

$$\xi(N + M; nm) \leq \xi_1(N; n) + \xi_2(M; m). \tag{37}$$

Here the number of particles appears as an additional parameter. The reader may verify that (37) also holds if ξ_1 is defined for nonintegral N by linear interpolation between consecutive integers, and the same procedure is used for ξ_2 and ξ.[21] The obvious extension of (37) to a box broken up into l regions (of reasonable shape), each separated by at least a distance R_o from all other regions, is

$$\xi(\sum_i N_i; \prod_i n_i) \leq \sum_i \xi_i(N_i; n_i). \tag{38}$$

B. Limiting Energy for a Special Sequence

Consider a sequence of cubes Ω_k with volume $V_k = 2^{3k}v_o$ ($v_o \geq R_0^3$) for $k = 1, 2, \cdots$. The wavefunctions are required to vanish if any particle lies outside the cube or on the inside within a distance $\frac{1}{2}R_o$ of any boundary.[22] The cube Ω_k consists of eight cubes of type Ω_{k-1}. Let us suppose that each of the first four contains $\rho_1 V_{k-1}$ particles and each of the second four, $\rho_2 V_{k-1}$ particles. For $i = 1$–4 let $n_i = \exp(V_{k-1}\sigma_1)$ and for $i = 5$–8, $\exp(V_{k-1}\sigma_2)$. The inequality (38) applied to this case and expressed in terms of normalized quantities ($\epsilon = V^{-1}\xi$) becomes

$$\epsilon_k(\rho, \sigma) \leq \tfrac{1}{2}\epsilon_{k-1}(\rho_1, \sigma_1) + \tfrac{1}{2}\epsilon_{k-1}(\rho_2, \sigma_2), \tag{39}$$
$$\rho = \tfrac{1}{2}(\rho_1 + \rho_2); \qquad \sigma = \tfrac{1}{2}(\sigma_1 + \sigma_2),$$

where ϵ depends on both the entropy (per unit volume) σ and the density ρ.

For $\rho_1 = \rho_2 = \rho$ and $\sigma_1 = \sigma_2 = \sigma$, (39) and (30) imply that $\epsilon_k(\rho, \sigma)$ is monotone decreasing in k and bounded below by $-C\rho$.[23] Thus

$$\epsilon(\rho, \sigma) = \lim_{k\to\infty} \epsilon_k(\rho, \sigma) \tag{40}$$

[20] See, for example, Sec. 11 of Ref. 5.

[21] Compare Sec. 5 of Ref. 5.

[22] That is, the particles all lie within a "free volume" less than the volume of the cube.

[23] The first term in (29), the kinetic energy, is always nonnegative.

exists for σ in $(-\infty, \infty)$ and ρ in the interval $(0, \rho_m)$ where ρ_m is the maximum density (finite or infinite). Furthermore, (39) in the limit $k \to \infty$ together with the fact that ϵ is bounded from above by any ϵ_k implies that $\epsilon(\rho, \sigma)$ is a continuous convex downwards function of both arguments together. That is, for $0 < \lambda < 1$,

$$\epsilon[\lambda\rho_1 + (1-\lambda)\rho_2, \lambda\sigma_1 + (1-\lambda)\sigma_2] \leq \lambda\epsilon(\rho_1, \sigma_1) + (1-\lambda)\epsilon(\rho_2, \sigma_2) \quad (41)$$

for any σ_1, σ_2 and any ρ_1, ρ_2 in $(0, \rho_m)$. The convergence of ϵ_k to ϵ is uniform on any closed rectangle in the (ρ, σ) plane which lies within the domain of definition. Appendix A contains the proof.

C. Extension to Other Sequences, etc.

The above arguments may be generalized in at least two ways: to more general sequences of domains and to more general types of potential. A considerable generalization has already been carried out by Fisher[5] for the canonical and grand canonical formalisms. Fortunately, analogous arguments apply in the microcanonical case: Fisher's proofs require but minor modifications; for the most part only a change in notation. Details of this analogy are given in Appendix B. The conclusion is that for all cases of quantum mechanical systems for which Fisher has shown the existence of the canonical thermodynamic limit, the microcanonical limit also exists. [The equivalence of the thermodynamic functions derived in the two schemes is the subject of Sec. VI below.] Therefore we shall omit the details and only point out the generalizations possible.

(i) Any sequence of cubes with volume increasing to infinity yields the same result as (40), provided the density of particles remains constant. Much more general shapes of domain are possible, though in all the cases considered by Fisher essentially the same "hard walls" boundary conditions apply: the wavefunction vanishes when any particle lies outside the domain. In particular, periodic boundary conditions are not considered.

(ii) The condition (31) may be replaced by one of the "weak tempering" conditions discussed by Fisher. Many particle potentials are also permissible (with certain restrictions, of course). It appears difficult, perhaps impossible, to weaken the stability condition (30).

In the above derivations we considered, as a matter of convenience, the energy and entropy per unit volume; no particular difficulty arises if one or both of these quantities is normalized on a per-particle basis. Thus let u, s, and $v = \rho^{-1}$ be the energy, entropy, and volume per particle, respectively. From the relation

$$u_k(v, s) = v\epsilon_k(v^{-1}, v^{-1}s), \quad (42)$$

it is clear that the convergence of ϵ_k to ϵ at fixed (ρ, σ) is equivalent to the convergence of u_k to a limit u at fixed (v, s). The convexity of ϵ implies that the corresponding limiting function u is convex downwards in v and s together.

V. ENTROPY AS A FUNCTION OF ENERGY

In Sec. III [IV] we showed that the $\epsilon_k(\sigma)[\epsilon_k(\rho, \sigma)]$ converge to a well-defined $\epsilon(\sigma)[\epsilon(\rho, \sigma)]$ given certain conditions on the interaction terms in the Hamiltonian and on the sequence of domains Ω_k of increasing volume. This implies that the inverse[11] functions $\sigma_k(\epsilon)[\sigma_k(\rho, \epsilon)]$ converge to $\sigma(\epsilon)[\sigma(\rho, \epsilon)]$, the inverse function to $\epsilon(\sigma)[\epsilon(\rho, \sigma)]$, for any ϵ greater than

$$\epsilon_o = \epsilon(\sigma = 0), \quad (43)$$

the (normalized) ground-state energy[24]. For suppose there is some $\epsilon_1 \geq \epsilon_0[\sigma(\epsilon)$ is not defined for $\epsilon < \epsilon_0]$ and a number $\delta > 0$ such that

$$\sigma_K(\epsilon_1) \leq \sigma(\epsilon_1) - \delta \quad (44)$$

for an increasing sequence[17] $\{K\}$. Now $\epsilon_k(\sigma)$ is monotone increasing, and thus (44) implies that

$$\epsilon_K(\sigma_1 - \delta) \geq \epsilon_1, \quad (45)$$

where σ_1 denotes $\sigma(\epsilon_1)$; or, in the limit $K \to \infty$

$$\epsilon(\sigma_1 - \delta) \geq \epsilon_1 = \epsilon(\sigma_1). \quad (46)$$

Since $\epsilon(\sigma)$ is also montone increasing, (46) means that it must equal ϵ_1 for all σ between $\sigma_1 - \delta$ and σ_1. However, $\epsilon(\sigma)$ is also convex downwards, so that such a horizontal portion of its graph must extend to $\sigma = 0$. Or, in other words, $\epsilon_1 = \epsilon_0$. An analogous argument yields the same conclusion if in place of (44) we consider

$$\sigma_K(\epsilon_1) \geq \sigma(\epsilon_1) + \delta. \quad (47)$$

Thus the convergence of $\sigma_k(\epsilon)$ to $\sigma(\epsilon)$ for all $\epsilon > \epsilon_0$ is assured. Similarly, in the case of gases, $\sigma_k(\rho, \epsilon)$ converges to $\sigma(\rho, \epsilon)$ for

$$\epsilon > \epsilon_0(\rho) = \epsilon(\rho, \sigma = 0). \quad (48)$$

Simple geometrical considerations show that since $\epsilon(\sigma)$ is monotone increasing and convex downwards

[24] The limiting function $\sigma(\epsilon)$ and, in qualitative terms, its connection with the energy spectrum, have been considered by H. B. G. Casimir, Z. Physik 171, 246 (1962), in a discussion of the third law of thermodynamics.

(Fig. 1), $\sigma(\epsilon)$ is monotone increasing and convex upwards. Similarly, $\sigma(\rho, \epsilon)$ is a convex upwards function of both arguments together, and monotone increasing in the second. It should be noted in this last case that both σ and ϵ are normalized per unit volume. Of course, the entropy per particle s is a convex upwards function of v and u, the volume and energy per particle, together (see Sec. IV).

Define ϵ_∞ as the smallest energy such that

$$\sigma(\epsilon) = \sigma(\epsilon_\infty) \tag{49}$$

for all $\epsilon > \epsilon_\infty$—that is, the energy associated with the left edge of the horizontal portion (if any) of the $\sigma(\epsilon)$ curve. It is properly identified (see Sec. VI) as the energy in the limit $T \to \infty$. For a spin system, ϵ_∞ is finite since all the energy eigenvalues (Sec. III) fall within an interval which increases linearly with the size of the system. For a quantum gas, on the other hand, ϵ is defined for values of σ extending to $+\infty$, and thus ϵ_∞ is infinite.

We are now in a position to prove the assertions in Sec. II regarding the entropy S^-. For a system Ω_k of volume V_k define[25]

$$\sigma_k^-(\delta; \epsilon) = V_k^{-1} S_k^{-1}(\delta; V\epsilon). \tag{50}$$

We shall show that for a fixed $\delta > 0$ and provided $\epsilon_0 < \epsilon \leq \epsilon_\infty$

$$\lim_{k\to\infty} \sigma_k^-(\delta; \epsilon) = \sigma(\epsilon), \tag{51}$$

where the right-hand side is the limit of $\sigma_k(\epsilon)$ as $k \to \infty$. In the proof we assume δ is less than $\epsilon - \epsilon_0$; the modifications required for larger δ are obvious.

From the definition of S^- we obtain the inequalities

$$\sigma_k(\epsilon) \geq \sigma_k^-(\delta; \epsilon) \geq \sigma_k(\epsilon) + V_k^{-1} \log \{1 - \exp V_k[\sigma_k(\epsilon - \delta) - \sigma_k(\epsilon)]\}. \tag{52}$$

The quantity

$$\alpha = \sigma(\epsilon) - \sigma(\epsilon - \delta) \tag{53}$$

is greater than zero since $\epsilon \leq \epsilon_\infty$ and $\sigma(\epsilon)$ is convex upwards. Choose k large enough so that

$$|\sigma_k(\epsilon) - \sigma(\epsilon)| \leq \tfrac{1}{3}\alpha, \qquad |\sigma_k(\epsilon - \delta) - \sigma(\epsilon - \delta)| \leq \tfrac{1}{3}\alpha \tag{54}$$

and hence

$$\sigma_k(\epsilon - \delta) - \sigma_k(\epsilon) \leq -\tfrac{1}{3}\alpha, \tag{55}$$

and therefore

$$\sigma_k(\epsilon) \geq \sigma_k^-(\delta; \epsilon) \geq \sigma_k(\epsilon) + V_k^{-1} \log [1 - \exp(-\tfrac{1}{3}\alpha V_k)]. \tag{56}$$

[25] Since the density is held constant in the following argument, it is not explicitly indicated.

Since $V_k \to \infty$ as $k \to \infty$, it is apparent that $\sigma_k(\epsilon)$ and $\sigma_k^-(\delta; \epsilon)$ approach the same limit.

VI. EQUIVALENCE OF THE CANONICAL AND MICROCANONICAL FORMALISMS

Before stating the formal theorems which constitute the principle results of this section, we shall discuss in general terms the relation between the canonical and microcanonical formalisms. Let ϵ, σ, f be the normalized energy, entropy, and free energy (per unit volume). The well-known thermodynamic relationship

$$f(T) = \epsilon(\sigma) - \sigma T, \qquad \sigma = -df/dT; \quad T = d\epsilon/d\sigma \tag{57}$$

is often called a Legendre transformation.

In the canonical formalism[4-6] one considers a sequence of normalized free energies (convex-upwards functions of T):

$$f_k(T) = V_k^{-1} F(T), \tag{58}$$

where F is defined by (1) and (2) using the energy spectrum of the system Ω_k. Under certain conditions this sequence approaches a limiting function $f(T)$. Under similar conditions the microcanonical formalism yields a sequence $\epsilon_k(\sigma)$ converging to a limiting function $\epsilon(\sigma)$. If the limiting functions are related by the Legendre transformation (57), the thermodynamic predictions of the two formalisms are identical.

Equation (57) is not quite adequate for our purposes. A simple extension discussed in Appendix C, which we shall also call a "Legendre transformation," eliminates the ambiguities in cases where ϵ (or f) has a discontinuous first or vanishing second derivative. The essential condition is that $\epsilon(\sigma)$ be convex downwards, or, equivalently, that $f(T)$ be convex upwards.

Our equivalence proof employs the following device. As in Appendix C, but with a slight change in notation, define

$$f_k^*(T) = \inf_\sigma [\epsilon_k(\sigma) - \sigma T] \tag{59}$$

the convex-upwards function "associated with" $\epsilon_k(\sigma)$. The function

$$\epsilon_k^*(\sigma) = \sup_T [f_k^*(T) + \sigma T] \tag{60}$$

which is convex-downwards and related to f_k^* by a Legendre transformation, is the "convex cover" of ϵ_k, that is

$$\epsilon_k^*(\sigma) \leq \epsilon_k(\sigma) \tag{61}$$

and for any σ is greater than or equal to any other convex-downwards function satisfying (61) [See Appendix C]. Theorem 2 of Appendix C shows that the convergence of the f_k^* to a function $f(T)$ implies the convergence of the ϵ_k^* to the function $\epsilon^*(\sigma)$ associated with $f(T)$ by a Legendre transformation. Thus to demonstrate the equivalence of canonical and microcanonical formalisms it suffices to show that (i) $f_k(T)$ and $f_k^*(T)$ approach the same limit as $k \to \infty$; (ii) $\epsilon_k(\sigma)$ and $\epsilon_k^*(\sigma)$ approach the same limit as $k \to \infty$.

We consider first the relation of f_k and f_k^* by means of the following inequality:

$$\exp\left[-V_k f_k(T)/T\right] = Z = \sum_i e^{-E_i/T} = \sum_n e^{-\xi(n)/T}$$
$$\geq \sum_{n=1}^{N} e^{-\xi(n)/T} \geq N e^{-\xi(N)/T}, \qquad (62)$$

where N is any integer in the domain of definition of ξ. [The final term is still a lower bound for Z when N is any real number in that domain.] Equation (59) may be rewritten as

$$\exp\left[-V_k f_k^*(T)/T\right] = \sup_N \left[N e^{-\xi(N)/T}\right], \qquad (63)$$

which upon comparison with (62) yields the inequality

$$f_k^*(T) \geq f_k(T). \qquad (64)$$

When V_k is large there is always some N which makes the right side of (62) a "good approximation" for Z (in a suitable sense), and in consequence we have the following two theorems.

Theorem 1. If as $V_k \to \infty$, the $f_k(T)$ converge to a function $f(T)$ on some interval $0 \leq T_1 < T < T_2 \leq \infty$, then the $f_k^*(T)$ converge to $f(T)$ on the same interval.

Theorem 2. Assume that the $\epsilon_k^*(\sigma)$ converge to a function $\epsilon^*(\sigma)$ on the interval $\sigma_1 < \sigma < \sigma_2 \leq \infty$ as $V_k \to \infty$. Let T_1 be the limit of $d\epsilon^*/d\sigma$ as σ approaches σ_1 from above, and T_2 the limit as σ approaches σ_2 from below.[26] Then on the interval $T_1 < T < T_2$ the sequences $f_k(T)$ and $f_k^*(T)$ both converge to the same limiting function $f(T)$.

The straightforward, but tedious, proofs of these theorems will be found in Appendix D. For the systems considered in Secs. III and IV, it has previously been shown[4-6] that the $f_k(T)$ converge to $f(T)$ on $(0, \infty)$, and thus by Theorem 1 the $f_k^*(T)$ also converge to $f(T)$. The only remaining problem is to show that ϵ_k and ϵ_k^* approach a common limit as $k \to \infty$. The function $\epsilon(\sigma)$ for the systems considered in Secs. III, IV is always convex downwards and the limit[27] of a sequence

$$\epsilon_k(\sigma) + \delta_k \qquad (65)$$

which is monotone decreasing in k, with δ_k approaching 0 as $k \to \infty$. Thus, because $\epsilon_k^* + \delta_k$ is the "convex cover" of $\epsilon_k + \delta_k$,

$$\epsilon(\sigma) \leq \epsilon_k^*(\sigma) + \delta_k \leq \epsilon_k(\sigma) + \delta_k \qquad (66)$$

and we conclude that the ϵ_k^* converge to $\epsilon(\sigma)$. The proof of equivalence of canonical and microcanonical formalisms under the conditions given in Secs. III and IV is therefore complete.

There are some systems for which the conditions of Sec. III are violated with the result that ϵ_k and ϵ_k^* do not converge to the same function. An instructive, though highly artificial example, is found in Part 1 of Appendix E. For this case, $f_k^*(T)$ and $f_k(T)$ still converge to the same function in agreement with Theorem 1. In fact, Theorems 1 and 2 are quite general and are independent of the conditions used in Secs. III and IV.

VII. PHASE TRANSITIONS, ZERO TEMPERATURE LIMITS, AND THE THIRD LAW OF THERMODYNAMICS

The limiting curve[25] $\epsilon(\sigma)$ is monotone increasing and convex downwards (as in Fig. 1). If a straight line of slope T is drawn tangent to (touches, but does not cross) the curve $\epsilon(\sigma)$, the values of ϵ and σ at the point of contact are the energy and entropy associated with the temperature T (See Appendix C). Should the tangent line and the curve $\epsilon(\sigma)$ have an entire segment (rather than one point) in common, the monotone increasing functions $\epsilon(T)$, $\sigma(T)$ show jump discontinuities indicating a first order phase transition.[2] If the tangent line touches the curve at a single point, but at this point $d^2\epsilon/d\sigma^2$ vanishes, then the specific heat $[d\epsilon/dT]$ becomes infinite at this temperature, as in certain types of second order phase transition. A similar analysis may be made for other types of phase transition. The question posed in the introduction thus receives a precise answer in terms of Eq. (8b) in the limit as $V \to \infty$.

[26] The existence of these limits is guaranteed by the fact that $\epsilon^*(\sigma)$ is convex downward and therefore its derivative is monotone increasing in σ. T_2 may, of course, be infinite. T_1 cannot be less than zero since $\epsilon^*(\sigma)$ is monotone increasing.

[27] This is clear for the special sequences of cubes considered in Sec. III, IV. In the latter, the "strong tempering" condition (31) allows one to set $\delta_k = 0$. Fisher's argument (Sec. 6 of Ref. 5) employs in place of δ_k a t_k which approaches a finite limit t. The argument works just as well if t_k is replaced by $\delta_k = t_k - t$. The case of a sequence of domains other than the special sequence of cubes is not essentially different, since one uses the ϵ_k from the latter to set lower bounds on the ϵ_K from the former.

The zero temperature limit merits a careful discussion because a certain amount of confusion exists in the literature. For $\sigma < 0$ we know that

$$\epsilon(\sigma) = \epsilon_0 = \epsilon(\sigma = 0). \tag{67}$$

The graphical procedure given above shows that

$$\lim (T \to 0+)\epsilon(T) = \epsilon_0 \tag{68}$$

Since by definition $\epsilon(\sigma = 0)$ is the limit as $k \to \infty$ of V_k^{-1} times the ground-state energy E_{0k} of the system Ω_k, (68) shows that $\epsilon(T)$ approaches the ground-state energy per unit volume as T approaches zero. This is not a trivial result for, as we shall see, the corresponding equality is not in general true for the entropy.

Again using the geometrical construction, one finds that

$$\lim (T \to 0+)\sigma(T) = \sigma_0 = \lim (\epsilon \to \epsilon_0+)\sigma(\epsilon), \tag{69}$$

or, in other words, σ_0 is the largest value of σ for which $\epsilon(\sigma) = \epsilon_0$. We may write σ_0 in terms of the energy spectrum:

$$\sigma_0 = \lim (\delta \to 0+) \lim_{k\to\infty} V_k^{-1} \log \mu(E_{0k} + V_k\delta), \tag{70}$$

where use has been made of (69) and (8a) together with the remarks following (68). Note that the argument of the logarithm is the number of states lying within an energy interval $V_k\delta$ above the ground state.

If the order of limits in (70) is interchanged, one obtains

$$\sigma' = \lim_{k\to\infty} V_k^{-1} \log d_k, \tag{71}$$

where d_k is the degeneracy of the ground state of Ω_k. The limit (71) may or may not exist; if it does exist, σ' may be equal to or less than σ_0.

The arguments of Sec. I provide some justification for associating σ_0 [see (69)] with the "entropy at zero temperature" obtained by extrapolating to $T = 0$ measurements made on macroscopic systems at finite temperatures.[28] The latter appears in discussions of the "Third Law of Thermodynamics", one form of which states that the entropy in the limit $T \to 0$ is independent of pressure, magnetic field, etc.[29]

Many textbooks[30], on the other hand, discuss the third law on the basis of the degeneracy of the ground state, that is, in terms of σ' rather than σ_0. Such discussions may be based on a naive application of the canonical formalism to finite systems, and perhaps in some cases[31] reflect the erroneous assumption that at very low temperatures only the energy levels lying within an interval of order kT above the ground state make a significant contribution to the partition sum (1).[32] The reader may find instructive the examples given in Appendix E, Part 2, of spin systems satisfying the postulates of Sec. III, but for which $\sigma' = 0$ and $\sigma_0 > 0$. These examples are artificial, but we feel that they nonetheless illustrate a significant point: There is no reason to assume, on the basis of statistical mechanics alone, that the properties of a macroscopic system at low temperatures bear any resemblance to those calculated for the ground state. The energy forms an exception [see (67), (68)]. But for other quantities, including the entropy, one must know something about other low-lying levels in addition to the ground state. For many systems it is probably the case that the vast majority of low-lying levels "resemble" the ground state in certain respects, but this does not seem to be true in general.

VIII. CONCLUSION

The main result of this paper has been to establish rigorously the connection between thermodynamic quantities and the distribution of energy levels in the limit of an infinite system. The thermodynamic results depend only on the coarse distribution of levels. Roughly speaking, the number of levels in an energy interval which increases in direct proportion to the volume serves to determine the entropy. It would be of interest to know whether under certain conditions a smaller interval could be chosen; for example one of constant magnitude or increasing as (say) $V^{\frac{1}{2}}$, and the same results obtained.

[28] Our approach to the problem is similar to that of Casimir, Ref. 24.

[29] A discussion together with applications to practical problems is found in J. Wilks, *The Third Law of Thermodynamics* (Oxford University Press, New York, 1961).

[30] For example, Ref. 29, p. 87; K. Huang, Ref. 9, p. 191; L. D. Landau and E. M. Lifschitz, Ref. 9, p. 66; T. L. Hill, *Statistical Mechanics* (McGraw-Hill Book Company, Inc., New York, 1956), p. 76; A. H. Wilson, *Thermodynamics and Statistical Mechanics* (Cambridge University Press, New York, 1960), p. 190.

[31] An argument of this type has been presented and criticized by M. J. Klein, Scuolo Internazionale di Fisica "Enrico Fermi" Corso X (Varenna, Italy, 1959), p. 17. Our criticism (footnote 32) is of a slightly different character.

[32] For most systems of the type discussed in Secs. III, IV (exceptions may occur if $d\epsilon/d\sigma$ is discontinuous at σ_0) the exact opposite is the case; given a temperature T, the partition sum for a system of sufficient size is not appreciably altered if contributions from the ground state and all levels within an energy interval kT above the ground state are *omitted*. [The rapid increase with energy of the density of levels more than compensates for the decrease of the exponential weighting factor.]

The principle utility of our rigorous formulation is probably found in answering questions of principle rather than in applications to exact or approximate calculations on model systems, for which the the canonical and grand canonical approaches seem better suited. In many cases we can be certain, from the arguments of Sec. VI, that the thermodynamic results from canonical and microcanonical calculations will be identical.

A proper statistical-mechanical discussion of the "third law of thermodynamics" ought certainly to start from a correct expression for the entropy as the temperature approaches zero; that is, from Eq. (70) rather than (71). One may well ask: Under what conditions will these two expressions be equal? And, more generally, what conditions on the Hamiltonian insure that the properties of a macroscopic system at low temperatures resemble those of the ground state? Both questions are of some interest since it is frequently far easier to calculate properties of the ground state of a system than it is to obtain thermal averages at finite temperatures.

ACKNOWLEDGMENTS

The author wishes to acknowledge a number of helpful comments by Dr. M. E. Fisher. He is also grateful for the hospitality shown him while a National Science Foundation Postdoctoral Fellow at the University of California in San Diego and the Summer Institute for Theoretical Physics of the University of Colorado, where a portion of the research here reported was carried out.

APPENDIX A: UNIFORM CONVERGENCE OF (40)

The convergence of $\epsilon_k(\rho, \sigma)$ to $\epsilon(\rho, \sigma)$ on the closed rectangle $\rho' \leq \rho \leq \rho''$, $\sigma' \leq \sigma \leq \sigma''$ is monotone. If it is not uniform, there is a number $\delta > 0$, an increasing sequence[17] $\{K\}$, and points (ρ_K, σ_K) in the rectangle with the property

$$\epsilon_K(\rho_K, \sigma_K) \geq \epsilon(\rho_K, \sigma_K) + \delta. \tag{A1}$$

Without loss of generality, assume that the points (ρ_k, σ_k) converge to some point (ρ_0, σ_0). Choose σ^* in the interval (σ_0, σ''), or equal to σ'' if $\sigma_0 = \sigma''$, such that [note that ϵ is continuous]:

$$\epsilon(\rho_0, \sigma^*) \leq \epsilon(\rho_0, \sigma_0) + \tfrac{1}{4}\delta. \tag{A2}$$

Choose K_0 large enough so that if $K \geq K_0$ belongs to our sequence,

$$\sigma_K \leq \sigma^* \tag{A3}$$

$$|\epsilon(\rho_K, \sigma_K) - \epsilon(\rho_0, \sigma_0)| \leq \tfrac{1}{4}\delta \tag{A4}$$

and therefore

$$\begin{aligned}\epsilon_K(\rho_K, \sigma^*) &\geq \epsilon_K(\rho_K, \sigma_K) \geq \epsilon(\rho_K, \sigma_K) + \delta \\ &\geq \epsilon(\rho_0, \sigma^*) + \tfrac{1}{2}\delta,\end{aligned} \tag{A5}$$

where we have used the monotonicity of ϵ_K in σ together with (A3), (A1), (A2), and (A4).

If K, L both belong to our sequence and satisfy $K \geq L \geq K_0$ we have

$$\epsilon_L(\rho_K, \sigma^*) \geq \epsilon_K(\rho_K, \sigma^*) \geq \epsilon(\rho_0, \sigma^*) + \tfrac{1}{2}\delta \tag{A6}$$

since ϵ_k is monotone decreasing in k. Now ϵ_L is a continuous function of ρ [see the remark following (37)]; thus the limit of (A6) as $K \to \infty$,

$$\epsilon_L(\rho_0, \sigma^*) \geq \epsilon(\rho_0, \sigma^*) + \tfrac{1}{2}\delta, \tag{A7}$$

since it holds for all $L \geq K_0$ in the sequence $\{K\}$, contradicts the assumption that ϵ_k converges to ϵ at (ρ_0, σ^*).

APPENDIX B: QUANTUM GASES WITH GENERAL DOMAINS AND WEAKLY-TEMPERED POTENTIALS

We enclose the equation and section numbers from Fisher's paper[5] in square brackets. "Weakly-tempered" potentials satisfy the condition [3.10] in place of the "strong tempering" condition (31). The methods of Sec. IV yield a generalization of (37) in the case of weak tempering[33]:

$$\begin{aligned}&\xi(N + M; nm) \\ &\quad \leq \xi_1(N; n) + \xi_2(M; m) + NMw_B/R^{3+\epsilon},\end{aligned} \tag{B1}$$

where systems 1 and 2 are separated by a partition of minimum thickness R. In terms of normalized quantities, (B1) may be written as

$$\begin{aligned}V\epsilon(\rho, \sigma) \leq\ & V_1\epsilon_1(\rho_1, \sigma_1) + V_2\epsilon_2(\rho_2, \sigma_2) \\ &+ NMw_B/R^{3+\epsilon}\end{aligned} \tag{B2}$$

with

$$\begin{aligned}V\rho &= V_1\rho_1 + V_2\rho_2, \\ V\sigma &= V_1\sigma_1 + V_2\sigma_2.\end{aligned} \tag{B3}$$

Now (B2) is precisely the equivalent of [5.3] if in the latter each g[34] is replaced by $-\beta = -T^{-1}$ times the corresponding ϵ. The relation [5.4] becomes the condition (B3) on both density and entropy.

The arguments following [5.4] are largely geometrical in nature and may be applied to (B2)

[33] The ϵ appearing in the exponent of R in Eqs. (B1) and (B2) is a constant greater than zero and should not be confused with the energy.

[34] Fisher's g is $-\beta$ times the free energy f.

with the appropriate substitution. The only complication is that one must keep track of σ as well as ρ. This may be done by regarding ρ where it appears as an argument of g, as in [5.5], or in equations such as [5.6], as a two-component vector (ρ, σ). However, in certain cases, as where ρ^2 appears in [5.5] and [6.12], it is the density alone (first component of the "vector") which appears in the analogous equations for the microcanonical ensemble.

Note that we obtain a lower bound to ϵ solely by the use of (30), whereas the analogous upper bound for g, Sec. [11b], requires additional considerations. In this respect the proof for a microcanonical ensemble is formally more simple than that for the canonical case.

The arguments of Sec. [6] when applied to the microcanonical case yield the limiting $\epsilon(\rho, \sigma)$ for a standard sequence of cubes. The convergence is uniform on rectangular domains lying within the region of definition and the proof is an obvious generalization of that in Appendix A. This uniform convergence plus the arguments of Sec. [9], suitably modified to apply to the microcanonical case, show that the same limiting $\epsilon(\rho, \sigma)$ is obtained for a fairly general sequence of dimains.

APPENDIX C: LEGENDRE TRANSFORMATIONS

Let $g(x)$ be defined and finite for x a real number in the connected interval $\mathfrak{A}$. Mandelbrojt[35] calls the

$$g^*(y) = \sup\,(x \in \mathfrak{A})[yx - g(x)] \tag{C1}$$

convex function *associated*[36] with $g(x)$. It is defined and finite for y in some set of real numbers $\mathfrak{B}$. Assume that $y_1 < y_3$ are both in $\mathfrak{B}$, and y_2 is a number lying between them. Define the function

$$f_x(y) = xy - g(x). \tag{C2}$$

It is evident that

$$f_x(y_2) \geq (y_3 - y_1)^{-1} \times [(y_3 - y_2)f_x(y_1) + (y_2 - y_1)f_x(y_3)] \tag{C3}$$

and therefore, from (C1)

$$g^*(y_2) \leq (y_3 - y_1)^{-1}[(y_3 - y_2)g^*(y_1) + (y_2 - y_1)g^*(y_3)]. \tag{C4}$$

Thus $g^*(y)$ is defined and finite for all y between y_1 and y_3, and therefore $\mathfrak{B}$ is a connected interval.

[35] The results (C1) to (C10), with the exception of the geometrical interpretation, are taken from the paper by S. Mandelbrojt, Compt. Rend. (Paris) 209, 977 (1939), with minor modifications.

[36] In Sec. VI, $-g^*(x)$ is sometimes used in place of $g^*(x)$ and called the function associated with $g(x)$.

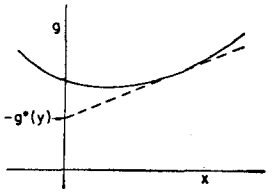


FIG. 3. Geometrical construction for $g^*(y)$ when $g(x)$ is convex. The slope of the tangent (dotted) line is equal to y.

Further, (C4) shows that $g^*(y)$ is convex downwards on $\mathfrak{B}$.

Consider

$$g^{**}(x) = \sup\,(y \in \mathfrak{B})[xy - g^*(y)], \tag{C5}$$

the convex function associated with $g^*(y)$. Note that for x in $\mathfrak{A}$ and y in $\mathfrak{B}$,

$$g^*(y) \geq xy - g(x) \tag{C6}$$

and thus, from (C5), we obtain for x in $\mathfrak{A}$

$$g^{**}(x) \leq g(x). \tag{C7}$$

A stronger result holds when $g(x)$ is convex downwards. A line drawn tangent to the curve $g(x)$ at the point x_1 in $\mathfrak{A}$, as in Fig. 3, lies everywhere on or below the curve $g(x)$[16]. If the slope of the tangent line is y_1, its intercept on the g axis is $-g^*(y_1)$ given by

$$g^*(y_1) = x_1y_1 - g(x_1), \tag{C8}$$

which is just the formula for a Legendre transformation. Now by (C5) and (C8),

$$g^{**}(x_1) \geq x_1y_1 - g^*(y_1) = g(x_1). \tag{C9}$$

This together with (C7) implies that

$$g^{**}(x_1) = g(x_1) \tag{C10}$$

for $g(x)$ convex downwards and x_1 in $\mathfrak{A}$. If $g(x)$ is not convex downwards, (C7) and (C10) together imply that g^{**} is its "convex cover," the largest convex function less than or equal to $g(x)$ for all x in $\mathfrak{A}$.

Convergence of Sequences

Theorem C1. If $g_n(x)$ is a sequence converging uniformly to $g(x)$ on the interval $\mathfrak{A}$, then the $g_n^*(y)$ converge uniformly to $g^*(y)$ on $\mathfrak{B}$, the interval where $g^*(y)$ is defined.

Proof. If $g_n(x)$ differs from $g(x)$ by at most δ for x in $\mathfrak{A}$, (C1) implies that $g_n^*(y)$ and $g^*(y)$ cannot differ by more than δ.

We also need the stronger result of Theorem C2 below, the proof of which is more difficult and utilizes a lemma: If the sequence $g_n(x)$ converges to $g(x)$ on $\mathfrak{A}$, then

$$\varliminf_n g_n^*(y) \geq g^*(y). \tag{C11}$$

Proof. Given any $\epsilon > 0$, we can choose an x_0 (depending on y) in $\mathfrak{a}$ such that

$$x_0 y - g(x_0) \geq g^*(y) - \tfrac{1}{2}\epsilon. \tag{C12}$$

Choose k large enough so that

$$g_k(x_0) < g(x_0) + \tfrac{1}{2}\epsilon. \tag{C13}$$

Inequalities (C12) and (C13) together with the definition of g_k^* yield the result

$$g_k^*(y) \geq g^*(y) - \epsilon. \tag{C14}$$

Theorem C2. Let $g_n(x)$ be a sequence of convex-downwards functions converging on the interval $\mathfrak{a}$ to a function $g(x)$. Then the $g_n^*(y)$ converge to $g^*(y)$ for all y in the *interior* of $\mathfrak{B}$, the interval where g^* is defined.

Proof. Assume the theorem is false. Then there is (see the preceding lemma) a y_1 in the interior of $\mathfrak{B}$, some $\epsilon > 0$, and an increasing sequence[17] $\{K\}$ for which

$$g_K^*(y_1) > g^*(y_1) + 2\epsilon. \tag{C15}$$

We shall show this leads to a contradiction. Choose x_1 in $\mathfrak{a}$ so that

$$x_1 y_1 - g(x_1) \leq g^*(y_1) \leq x_1 y_1 - g(x_1) + \epsilon. \tag{C16}$$

For each K for which (C15) holds we can find an x_K in $\mathfrak{a}$ such that

$$y_1 x_K - g_K(x_K) > g^*(y_1) + 2\epsilon. \tag{C17}$$

The x_K all lie within some finite interval even when $\mathfrak{a}$ is infinite. For suppose that $\mathfrak{a}$ extends to $+\infty$. Were it true that

$$g(x) \leq g(x_1) + y_1(x - x_1) \tag{C18}$$

for all $x \geq x_1$, $g^*(y)$ would not exist for $y > y_1$ and y would not lie in the interior of $\mathfrak{B}$. Thus there must be some $x_2 > x_1$ in $\mathfrak{a}$ for which

$$\delta = g(x_2) - g(x_1) - y_1(x_2 - x_1) > 0. \tag{C19}$$

The g_K converge to g on $\mathfrak{a}$; therefore

$$g_K(x_1) \leq g(x_1) + \tfrac{1}{2}\delta, \tag{C20}$$

$$g_K(x_2) \geq g(x_2) - \tfrac{1}{2}\delta = g(x_1) + y_1(x_2 - x_1) + \tfrac{1}{2}\delta$$

for K sufficiently large. The convexity of g_K implies that

$$g_K(x) \geq g(x_1) + y_1(x - x_1) + \tfrac{1}{2}\delta \tag{C21}$$

for all $x \geq x_2$. Upon comparison of (C16), (C17), and (C21) we see that $x_K \leq x_2$ for K sufficiently large. Similarly, if $\mathfrak{a}$ extends to $-\infty$, there is some x_2' which is a lower bound to the x_K for K sufficiently large.

We conclude that the x_K have at least one point of accumulation x_0 (which will not, in general, lie in $\mathfrak{a}$). Assume first that any interval (x^*, x_0) with $x^* < x_0$ contains infinitely many of the x_K. Choose x_a, x_b in $\mathfrak{a}$ satisfying

$$x_a < x_b < x_0 \tag{C22}$$

and choose $M > 0$ so that

$$y_1 - M < [g(x_b) - g(x_a)]/(x_b - x_a). \tag{C23}$$

Choose $\delta > 0$ sufficiently small so that

$$x_0 - \delta = x^* > x_b, \tag{C24}$$

$$M < \epsilon/2\delta, \tag{C25}$$

and choose K large enough so that

$$g_K(x^*) > g(x^*) - \tfrac{1}{2}\epsilon. \tag{C26}$$

From (C1) and (C17) we obtain the inequality

$$g(x^*) \geq (x^* - x_1)y_1 + g(x_1) - \epsilon. \tag{C27}$$

Upon combining (C17), (C26), (C16), and (C27), we have

$$g_K(x_K) - g_K(x^*) < y_1(x_K - x^*) - \tfrac{1}{2}\epsilon. \tag{C28}$$

When x_K lies in the interval (x^*, x_0), the convexity of g_K together with (C24) implies that[37]

$$\frac{g_K(x_b) - g_K(x_a)}{x_b - x_a} \leq \frac{g_K(x_K) - g_K(x^*)}{x_K - x^*}. \tag{C29}$$

As K approaches infinity, the left side of (C29) approaches the right side of (C23). This observation combined with (C28), (C24), (C25) and the fact that x_K lies in (x^*, x_0) leads to a contradiction.

An entirely analogous argument works for the case where any interval (x_0, x^*) contains infinitely many x_K. If infinitely many x_K coincide with x_0, (C17) combined with (C1) implies that the g_K do not converge to g at this point.

APPENDIX D: PROOFS FOR THE THEOREMS OF SEC. VI

Theorem 1. The convergence of the convex-upwards, monotone-decreasing functions $f_k(T)$ to $f(T)$ implies that the positive and monotone-increasing functions

$$s_k(T) = -df_k/dT \tag{D1}$$

[37] See the footnote on p. 327 of R. Courant, *Differential and Integral Calculus* (Interscience Publishers, Inc., New York, 1936), Vol. II.

converge to $s(T) = -df/dT$ at every point where the latter is continuous.[38] Thus for T in (T_1, T_2) there is a finite function $p(T)$ such that

$$s_k(T) \leq p(T) \tag{D2}$$

for all k. For a particular temperature let M be the smallest integer[39] satisfying

$$M \geq \exp [2V_k p(T)] \tag{D3}$$

and let $\sum_1$, $\sum_2$ denote sums for which $n < M$ and $n \geq M$, respectively. The inequality

$$\begin{aligned} ZV_k s_k(T) &= \sum_n (\log Z + \xi(n)/T) e^{-\xi(n)/T} \\ &\geq \sum_2 (\log Z + \xi(M)/T) e^{-\xi(n)/T} \end{aligned} \tag{D4}$$

combined with (62) in the form

$$\log Z \geq \log M - \xi(M)/T \tag{D5}$$

and with (D3) and (D2) yields the result

$$\sum_2 e^{-\xi(n)/T} \leq \tfrac{1}{2} Z. \tag{D6}$$

We may use (D6) and (63) to obtain the estimate

$$\begin{aligned} Z &\leq 2 \sum_1 e^{-\xi(n)/T} \\ &\leq 2(\log M) \exp [-V_k f_k^*(T)/T]. \end{aligned} \tag{D7}$$

Upon taking the logarithm of both sides of (D7), we see that $f_k(T)$ can exceed $f_k^*(T)$ by at most a term of order $V_k^{-1} \log V_k$. This observation together with (64) completes the proof.

Theorem 2. Since the convex-downwards functions $\epsilon_k^*(\sigma)$ converge to $\epsilon(\sigma)$, Theorem C2 of Appendix C shows that the $f_k^*(T)$ converge to a function $f(T)$ for T in (T_1, T_2). Given a particular T in (T_1, T_2) choose some $T' > T$ in the same interval. Let M be the smallest integer satisfying

$$\log (M - 1) \geq \frac{T}{T' - T} \left\{ \log \frac{2T}{T' - T} + \frac{V_k [f_k^*(T) - f_k^*(T')]}{T} \right\} \tag{D8}$$

for a given value of k, and let $\sum_1$, $\sum_2$ denote sums for which $n < M$ and $n \geq M$, respectively. The inequality

$$\exp [-\xi(n)/T'] \leq n^{-1} \exp [-V_k f_k^*(T')/T'] \tag{D9}$$

obtained from (63) may be used to obtain the estimate

$$\sum_2 e^{-\xi(n)/T} \leq \exp [-V_k f_k^*(T')/T] \sum_2 n^{-T'/T}. \tag{D10}$$

If a bound for the sum on the right side of (D10) is obtained from the corresponding integral, we have

$$\sum_2 e^{-\xi(n)/T} \leq \tfrac{1}{2} \exp [-V_k f_k^*(T)/T] \leq \tfrac{1}{2} Z, \tag{D11}$$

where the second inequality makes use of (64). Combining (D11) and (63), we arrive at (D7) as in the previous proof. It is clear that $\log M$ is of order V_k for large k, and hence $|f_k(T) - f_k^*(T)|$ is at most of order $V_k^{-1} \log V_k$.

APPENDIX E: SPECIAL EXAMPLES OF SPIN SYSTEMS

1. A system for which ϵ_k and ϵ_k^* do not converge to the same limit. Consider a linear chain of atoms of spin 1 with Hamiltonian

$$H = \sum_i h(i, i + 1), \tag{E1}$$

where

$$h(i, i + 1) = D(\sigma_i^2 + \sigma_{i+1}^2) + \infty (\sigma_i^2 - \sigma_{i+1}^2)^2, \tag{E2}$$

D is a constant, and σ_i is the z component of the spin operator with possible values $0, \pm 1$. The lowest energy for a chain of length V is zero, corresponding to the nondegenerate eigenfunction

$$\sigma_i = 0, \qquad i = 1, 2, \cdots, V. \tag{E3}$$

At an energy $2VD$ there are 2^V eigenstates of the form

$$\sigma_i = \pm 1 \qquad i = 1, 2, \cdots V. \tag{E4}$$

In the limit $V \to \infty$, the $\epsilon_V(\sigma)$ converge to

$$\epsilon(\sigma) = \begin{cases} 0 & \text{for } \sigma < \log 2, \\ 2D & \text{for } \sigma \geq \log 2, \end{cases} \tag{E5}$$

whereas the $\epsilon_V^*(\sigma)$ converge to

$$\epsilon^*(\sigma) = \begin{cases} 0 & \text{for } \sigma \leq 0 \\ 2D\sigma/\log 2 & \text{for } 0 \leq \sigma \leq \log 2 \\ 2D & \text{for } \sigma \geq \log 2. \end{cases} \tag{E6}$$

The Hamiltonian (E1) does not satisfy the conditions of Sec. III due to the infinite matrix elements in (E2). Nevertheless, the reader may verify that the $f_V^*(T)$ and $f_V(T)$ both converge to the same limit in agreement with Theorem 2 of Sec. VI.

2. Systems for which σ_0 [defined by (70)] differs from σ' [defined by (71)]. These examples, unlike the one given above, satisfy all the conditions given in Sec. III.

[38] Reference 6, Appendix A.

[39] In cases where (D3) or (D9) yields a value of M exceeding the total number of energy levels N, M should be set equal to N.

2a. Consider a linear chain containing V atoms of spin 1 with Hamiltonian

$$H = J \sum_{i=1}^{V} \sigma_i^2 - J \sum_{i=1}^{V-1} (\sigma_i \sigma_{i+1})^2. \tag{E7}$$

A given configuration of the chain may have n_1 consecutive spins with $\sigma_i = \pm 1$ followed by m_1 consecutive spins with $\sigma_i = 0$ followed by n_2 consecutive spins with $\sigma_i = \pm 1$, etc. If there are p sections of the type $\sigma_i = \pm 1$, the energy is

$$E = pJ. \tag{E8}$$

Clearly the ground state, $\sigma_i = 0$ for all i, is nondegenerate, whereas the 2^V states of the form

$$\sigma_i = \pm 1, \qquad i = 1, 2, \cdots, V \tag{E9}$$

all have energy $E = J$. We conclude immediately, using (70), (71), that

$$\sigma_0 \geq \log 2; \qquad \sigma' = 0. \tag{E10}$$

In fact it is not hard to show that σ_0 is precisely log 2. The interested reader may wish to verify that the same value is obtained in the canonical formalism [using the left side of (69)]. Furthermore, it may be shown that the thermal average $\langle \sigma_i^2 \rangle$ in the thermodynamic limit approaches 1 as $T \to 0$ whereas it is zero in the ground state.

2b. The results are also interesting when the foregoing example is generalized to three dimensions. Consider a cube containing $V = L^3$ atoms of spin 1 arranged in a simple cubic lattice and having a Hamiltonian

$$H = 3J \sum_{i=1}^{V} \sigma_i^2 - J \sum_{\langle ij \rangle} (\sigma_i \sigma_j)^2, \tag{E11}$$

where the second sum goes over all $3(L^3 - L^2)$ pairs of nearest neighbors. The ground state with $E = 0$ is again given by $\sigma_i = 0$ for all i. The 2^V states where all σ_i are either $+1$ or -1 have energy equal to

$$E = 3JL^2 = 3JV^{\frac{2}{3}} \tag{E12}$$

and the analysis of Sec. VII shows us that (E10) is true once again. Further investigation shows that $\xi(2^V)$ is of order $V^{\frac{1}{3}}$, though of course less than (E12). Thus to obtain σ_0 correctly, it is necessary to consider all levels within an energy interval above the ground state increasing *at least* as rapidly as $V^{\frac{1}{3}}$.

In both cases 2a and 2b the discrepancy between σ_0 and σ' may be remedied by imposing periodic boundary conditions. There are other ways in which these examples are artificial, but at least they do illustrate the fallacy in supposing that the low-temperature thermal properties always resemble those of the system's ground state.

APPENDIX F: THE IDENTITY OF ϵ AND $\bar{\epsilon}$ IN THE THERMODYNAMIC LIMIT

From the definitions (8b) and (10b) it is evident that $\bar{\epsilon}_k(\sigma)$ cannot exceed $\epsilon_k(\sigma)$ for a system Ω_k of volume V_k, and for $\sigma \leq 0$ the two quantities coincide. Given $\sigma > 0$, choose $\sigma_1 < \sigma$, and let N_1 and N be the smallest integers not less than $\exp(V_k \sigma_1)$ and $\exp(V_k \sigma)$, respectively; then

$$\bar{\epsilon}_k(\sigma) = (V_k N)^{-1} \left\{ \sum_{n=1}^{N_1} \xi(n) + \sum_{n=N_1+1}^{N} \xi(n) \right\}$$

$$\geq N_1 E_{0k}/N V_k + (N - N_1 - 1)\epsilon_k(\sigma_1)/N, \tag{F1}$$

where $E_{0k} = \xi(1)$ is the ground-state energy. As $k \to \infty$, the ratio N_1/N decreases as $\exp -V_k(\sigma - \sigma_1)$, whereas E_{0k}/V_k is bounded from below for the systems discussed in Secs. III, IV. Thus as $k \to \infty$,

$$\epsilon(\sigma) \geq \lim_{k \to \infty} \bar{\epsilon}_k(\sigma) \geq \epsilon(\sigma_1) \tag{F2}$$

for any $\sigma < \sigma_1$. But $\epsilon(\sigma)$ is continuous for $\sigma < \sigma_m$ and therefore

$$\lim_{k \to \infty} \bar{\epsilon}_k(\sigma) = \epsilon(\sigma) \tag{F3}$$

provided $\sigma < \sigma_m$.

Magnetic Properties of Charged Ideal Quantum Gases in n Dimensions

ROBERT M. MAY
Theoretical Department, School of Physics, University of Sydney, Sydney, N.S.W., Australia
(Received 15 February 1965)

We consider the mathematical model of an ideal gas of charged bosons or fermions in an n-dimensional space, treating n as a continuous variable. The investigation shows the extent to which the magnetic behavior depends on the dimensionality of the system. In particular, the charged Bose gas in a homogeneous magnetic field does not condense unless $n > 4$, in contrast to the field-free gas which condenses for $n > 2$: however so long as $n > 2$ and $T < T_c$ this system completely expels homogeneous magnetic fields weaker than a certain critical field, H_c. (T_c is the field free transition temperature.) This field expulsion comes explicitly from the condensed ground-state bosons for $n > 4$; but it is still present, and H_c has the same form, for $4 > n > 2$ where there is no condensation.

1. INTRODUCTION AND SUMMARY

IT is well-known that a three-dimensional ideal gas of charged bosons exhibits a Meissner–Ochsen feld (M–O) effect below its transition temperature[1]. This system thus serves as a simple example of a superconductor. The mathematical model of an ideal charged Bose gas in two dimensions has also been explored[2]: this gas does not condense, but at sufficiently low temperatures it possesses a very large diamagnetism which leads to an "imperfect" M–O effect which is practically indistinguishable from a perfect London one.

In this paper we consider the abstract mathematical model[3] of an ideal charged Bose or Fermi gas in n dimensions, in the presence of a magnetic field. The number of spatial dimensions, n, is treated as a continuous variable. We believe that this rather "unphysical" mathematical investigation is worthwhile in that it helps to distinguish those properties which depend only on the existence of a condensation in the absence of an external field from the properties which are explicit to three dimensions.

A recent study[4] of the thermodynamics of ideal (uncharged) quantum gases in n dimensions shows that Fermi statistics leads to expressions in which the number of dimensions is unimportant. For Bose statistics there are two main regions: for $n \leqslant 2$ there is no condensation into the ground state at low temperatures; for $n > 2$ there is such a condensation below a critical temperature, T_c; and for $n > 4$ there is a discontinuity in C_v at T_c.

In the present study of charged ideal gases in n dimensions we get results analogous to those of Ref. 4. For fermions, we get the usual diamagnetism of a gas of charged particles for all (positive) n. A similar result (i.e., diamagnetism but no M–O effect) pertains for bosons with $n \leqslant 2$, although here the diamagnetism becomes very large as the temperature tends to zero. For $n > 2$, the ideal Bose gas exhibits a M–O effect below its transition temperature.

To make this last statement quantitative, we define a response function $K(q)$,

$$\mathbf{M}(\mathbf{q}) = K(q)\mathbf{B}(\mathbf{q}), \tag{1}$$

where $\mathbf{M}(\mathbf{q})$ is the magnetization corresponding to the applied inhomogeneous field $\mathbf{B}(\mathbf{q})$ with wave-number $\mathbf{q}$. Then for $T < T_c$ and $4 > n > 2$, the charged Bose gas gives in the limit $q \to 0$

$$K(q) = -\frac{e^2}{mc^2q^2}\frac{N}{V}\left[\left\{1 - \left(\frac{T}{T_c}\right)^{\frac{1}{2}n}\right\} + \frac{1}{\gamma}\left(\frac{\hbar^2q^2}{2mkT}\right)^{\frac{1}{2}n-1}\left(\frac{T}{T_c}\right)^{\frac{1}{2}n}\right] \tag{2}$$

[γ is a numerical constant, given by Eq. (33)]. For $n > 4$ we get

$$K(q) = -\frac{e^2}{mc^2q^2}\frac{N}{V}\left[\left\{1 - \left(\frac{T}{T_c}\right)^{\frac{1}{2}n}\right\} + \frac{\zeta(\frac{1}{2}n - 1)}{6\zeta(\frac{1}{2}n)}\left(\frac{\hbar^2q^2}{2mkT}\right)\cdot\left(\frac{T}{T_c}\right)^{\frac{1}{2}n}\right]. \tag{3}$$

In each case the $1/q^2$ singularity corresponds to a perfect or "London" M–O effect. On the other hand for $n < 2$, $K(q)$ is regular at the origin. Reference 2 is devoted to the special case $n = 2$.

A more significant difference between charged

[1] M. R. Schafroth, Phys. Rev. **100**, 463 (1955).
[2] R. M. May, Phys. Rev. **115**, 254 (1959).
[3] It is perhaps worth noting that even the two-dimensional uncharged Bose gas does *not* seem to be realized in practice as the limit of very thin films [for a detailed discussion see J. M. Ziman, Phil. Mag. **44**, 548 (1953) and references therein], nor does the two-dimensional charged Bose gas model[2] seem to be relevant to thin superconductors [see J. M. Blatt, *Theory of Superconductivity* (Academic Press, Inc., New York, 1964), p. 363]. It is to be emphasized that the present paper is devoted to a mathematical abstraction.
[4] R. M. May, Phys. Rev. **135**, A1515 (1964).

bosons with $n > 4$ and those with $4 > n > 2$ is found when one considers the application of a *homogeneous* rather than an inhomogeneous magnetic field.[5] For $n \leqslant 4$ the charged Bose gas *no longer condenses*. However provided the external field is less than some critical value H_c [given by Eq. (51)] it is completely expelled at temperatures below T_c. (T_c is the condensation temperature in the absence of a magnetic field.) This is so despite the absence of any phase transition or macroscopic occupation of the ground state. For $n > 4$ the presence of the external homogeneous field has no effect on the condensation; below the transition temperature T_c the condensed bosons in the ground state expel fields weaker than H_c.

In Sec. 2 we evaluate the response function $K(q)$ for an n-dimensional ideal gas of charged bosons or fermions in the presence of an inhomogeneous magnetic field, and consider the limit as $q \to 0$. In Sec. 3 we investigate the behavior of such n-dimensional gases in the presence of a homogeneous magnetic field, and justify the remarks made in the preceding paragraph.

2. INHOMOGENEOUS FIELDS: $K(q)$

To begin we write down an expression for the response function $K(q)$ for a system in an n-dimensional space.

For a gas of noninteracting ideal particles, with masses m and charges[6] e, the field-free Hamiltonian is

$$\mathcal{H} = \sum_{\mathbf{p}} (\hbar^2\mathbf{p}^2/2m). \tag{4}$$

For a weak inhomogeneous magnetic field we can use perturbation theory[5] to derive a relationship between the magnetization $\mathbf{M}$ and the applied field $\mathbf{B}$: expressed in wavenumber space this relationship has the form (1). The derivation of $K(q)$ is provided for 3 dimensions in Ref. 1 and for two dimensions in Ref. 2. Since the analysis is independent of the number of dimensions until a final obvious step, we will not recapitulate it here but will simply state the result:

$$K(q) = -\frac{e^2}{mc^2}\frac{1}{q^2}\frac{1}{(n-1)V} \times \sum_{\mathbf{p}} \left\{\frac{\mathbf{p}^2}{\mathbf{p}\cdot\mathbf{q}} - \frac{n\mathbf{p}\cdot\mathbf{q}}{\mathbf{q}^2}\right\}\left\{F_0\left(\frac{\hbar^2(\mathbf{p}+\tfrac{1}{2}\mathbf{q})^2}{2m}\right) - F_0\left(\frac{\hbar^2(\mathbf{p}-\tfrac{1}{2}\mathbf{q})^2}{2m}\right)\right\}; \tag{5}$$

$\mathbf{q}$ is a wavenumber vector in an n-dimensional space, and $F_0(E)$ is the field-free distribution function of the ideal quantum gas,

$$F_0(E) = \{\exp[\alpha(E-\mu)] \mp 1\}^{-1} \tag{6}$$

(μ is the chemical potential, $\alpha \equiv 1/kT$).

The evaluation of (5) in the limit $q \to 0$ proceeds differently for the two cases where (a) $F_0(E)$ is regular at $E = 0$, and (b) $F_0(E)$ is singular at $E = 0$.

Since F_0 is the *field-free* distribution function, we can immediately appeal to Ref. 4 to see that, for fermions, $F_0(E)$ is regular for all n. For bosons in $n \leqslant 2$ dimensions, $F_0(E)$ is again regular; but for bosons with $n > 2$ there exists a transition temperature,

$$T_c = \frac{4\pi\hbar^2}{2mk}\left(\frac{N}{V\zeta(\tfrac{1}{2}n)}\right)^{2/n}, \tag{7}$$

below which there is macroscopic occupation of the ground state by a number of bosons N_0:

$$N_0 = N\{1 - (T/T_c)^{\frac{1}{2}n}\}. \tag{8}$$

Thus for bosons in $n > 2$ dimensions and with $T < T_c$, $F_0(E)$ is not regular at $E = 0$.

A. $F_0(E)$ Regular at $E = 0$

Here we can use Eq. (2.18) of Ref. 1 to write

$$F_0\left(\frac{\hbar^2(\mathbf{p}+\tfrac{1}{2}\mathbf{q})^2}{2m}\right) - F_0\left(\frac{\hbar^2(\mathbf{p}-\tfrac{1}{2}\mathbf{q})^2}{2m}\right) = \frac{\hbar^2\mathbf{p}\cdot\mathbf{q}}{m}F_0'\left(\frac{\hbar^2\mathbf{p}^2}{2m}\right) + \frac{\hbar^4}{8m^2}\mathbf{p}\cdot\mathbf{q}\mathbf{q}^2F_0''\left(\frac{\hbar^2\mathbf{p}^2}{2m}\right) + \frac{1}{3}\left(\frac{\hbar^2}{2m}\mathbf{p}\cdot\mathbf{q}\right)^3F_0'''\left(\frac{\hbar^2\mathbf{p}^2}{2m}\right) + \mathcal{O}\mathbf{q}^5, \tag{9}$$

where the prime indicates differentiation with respect to the energy $E = \hbar^2p^2/2m$. Assuming the usual periodic boundary conditions, the sum over $\mathbf{p}$ in (5) can be replaced by an n-dimensional integral in the thermodynamic limit (i.e., $V \to \infty$, $N \to \infty$, $N/V \to$ finite):

$$\frac{1}{V}\sum_{\mathbf{p}} \to \frac{1}{(2\pi)^n}\int_0^\infty p^{n-1}\,dp\int d\Omega. \tag{10}$$

The $\int d\Omega$ represents an integration over the surface of an n-dimensional unit sphere.

---

[5] Perturbation theory is valid for inhomogeneous fields, but for homogeneous fields an arbitrarily small field can produce qualitative changes in the particle wavefunctions, once the volume is big enough. Another way of putting this is to say that the homogeneous field can be derived from the inhomogeneous field as the limit $q \to 0$ only when $K(q)$ is regular at the origin.

[6] Note that in n dimensions, e has the dimensions $e^2 \sim ML^nT^{-2}$ and thus $e^2/(mc^2) \sim L^{n-2}$ [cf. Ref. 2, Eqs. (3.3) and (3.4)].

As $q \to 0$, Eq. (5) now reduces to

$$K(q) = -\frac{e^2\hbar^2}{m^2c^2}\frac{1}{q^2}\frac{1}{(n-1)(2\pi)^n} \times \int_0^\infty p^{n-1}\,dp \int d\Omega\,(1 - n\cos^2\theta) \times \left\{p^2F_0' + \frac{\hbar^2q^2p^2}{8m} \times \left[F_0'' + \frac{\hbar^2p^2}{3m}\cos^2\theta F_0'''\right] + \mathcal{O}q^4\right\}. \quad (11)$$

The surface area of the n-dimensional unit sphere, $S(n)$, is well known,

$$S(n) = 2\pi^{\frac{1}{2}n}/\Gamma(\tfrac{1}{2}n). \quad (12)$$

This result is easily generalized to get[7]

$$\int d\Omega\cos^2\theta = \frac{1}{n}S(n), \quad (13)$$

$$\int d\Omega\cos^4\theta = \frac{3}{n(n+2)}S(n). \quad (14)$$

We see that the leading term in (11) vanishes when the angular integral is performed, and so does the term in F_0''. For the remaining term, in F_0''',we can perform successive integration by parts to finally get

$$K(q) = -\frac{(n-2)}{24}\frac{e^2\hbar^2}{m^2c^2}\frac{N\langle E^{-1}\rangle}{V}(1 + \mathcal{O}q^2). \quad (15)$$

By $\langle E^{-1}\rangle$ we mean[8]

$$\langle E^{-1}\rangle = \frac{1}{N}\sum_{\mathbf{p}}\left(\frac{2m}{\hbar^2\mathbf{p}^2}\right)F_0\left(\frac{\hbar^2\mathbf{p}^2}{2m}\right) = \frac{2}{(n-2)kT}\frac{F_+(-\alpha\mu;\,\frac{1}{2}n-1)}{F_+(-\alpha\mu;\,\frac{1}{2}n)}. \quad (16)$$

The generalized Riemann zeta function $F_+(z;\,m)$ and its relative $F_-(z;\,m)$ are defined as

$$F_\pm(z;\,m) = \sum_{j=1}^\infty (\pm 1)^{j+1}\frac{e^{-jz}}{j^m}. \quad (17)$$

μ can be replaced by N/V in (16) by use of the relationship[4]

$$N/V = (2mkT/4\pi\hbar^2)^{\frac{1}{2}n}F_\pm(-\alpha\mu;\,\tfrac{1}{2}n). \quad (18)$$

We observe that at high temperatures, $4\pi\hbar^2/(2mkT) \ll d^2$ (d is the interparticle spacing, $d^n = V/N$), we must have $\alpha\mu$ large and negative to satisfy Eq. (18). This is just the limit of Boltzmann statistics, and we get

$$K(q) = -\frac{e^2\hbar^2}{12m^2c^2}\frac{1}{kT}\frac{N}{V}(1 + \mathcal{O}q^2) \quad (19)$$

for both Bose and Fermi statistics, regardless of the dimensionality of the system.

At low temperatures $[4\pi\hbar^2/(2mkT) \gg d^2]$ in n-dimensional Fermi gases, we get the familiar occupation of all states up to the level of the Fermi sea, and hence

$$\langle E^{-1}\rangle = \frac{n}{n-2}\frac{1}{E_F}\left\{1 + \mathcal{O}\left(\frac{kT}{E_F}\right)^2\right\} \quad (20)$$

where E_F is the Fermi energy, $E_F = \mu(T = 0)$.

The Bose gas at low temperatures $[4\pi\hbar^2/(2mkT) \gg d^2]$ condenses for $n > 2$ and is treated separately below. For $n < 2$, however, we can use (16) together with the expansion[9]

$$F_+(z;\,m) = z^{m-1}\Gamma(1-m) + \sum_{i=0}^\infty \frac{(-z)^i\zeta(m-i)}{i!} \quad (21)$$

(m not an integer) to get

$$(n-2)\langle E^{-1}\rangle = \frac{1}{kT}\left(\frac{4\pi\hbar^2}{2mkT\,d^2}\right)^{n/(2-n)}\frac{(2-n)}{[\Gamma(1-\frac{1}{2}n)]^{2/(2-n)}}. \quad (22)$$

Thus although $K(q)$ is always regular in the limit $q \to 0$ for bosons in $n < 2$ dimensions, this diamagnetic term can be very large as the temperature becomes very small.

B. Bosons in $n > 2$ dimensions ($T < T_c$).

For temperatures below a critical temperature T_c, defined by (7), the Bose gas condenses for $n > 2$. In this event, $F_0(E)$ is singular at the origin and (9) is no longer useful. The ground state is highly occupied [cf. Eq. (8)] and must be treated separately; for the other states we can use the Bose distribution function with $\mu = 0$, and replacement of sums by integrals is justified.

If we call the contribution to $K(q)$ from the condensed bosons $K_0(q)$, and the contribution from the remaining states $K_1(q)$, we can immediately use (5) to write

$$K_0(q) = -\frac{e^2}{mc^2}\frac{1}{q^2}\frac{N_0}{V}. \quad (23)$$

[7] These expressions are derived from the identity

$$\int e^{-r^2}f(\mathbf{r}\cdot\mathbf{a})\,d\mathbf{r} = \left(\int e^{-x^2}\,dx\right)^{n-1}\left(\int e^{-y^2}f(ya)\,dy\right).$$

[8] For $0 < n < 2$ we must be careful in the partial integrations; the resulting $F_\pm(-\alpha\mu;\,\frac{1}{2}n - 1)$ are the usual analytic continuation of the generalized Riemann zeta functions to such values of n.

[9] J. E. Robinson, Phys. Rev. **83**, 678 (1951); J. Clunie, Proc. Phys. Soc. (London) **A67**, 682 (1954).

N_0 is given as a fraction of the total number of bosons, N, by Eq. (8).

On the other hand, for the noncondensed particles we have

$$K_1(q) = -\frac{e^2}{mc^2}\frac{1}{q^2}\frac{1}{(n-1)(2\pi)^n} \times {}^P\!\int dp \left\{(n-2)+\frac{2(p^2-q^2/4)}{(\mathbf{p}-\frac{1}{2}\mathbf{q})\cdot\mathbf{q}}\right\} \times \{\exp(\alpha\hbar^2p^2/2m)-1\}^{-1}. \quad (24)$$

It is convenient at this point to introduce the dimensionless variables $s=\hbar p/(2mkT)^{\frac{1}{2}}$ and $t=\hbar q/(2mkT)^{\frac{1}{2}}$, and to expand the distribution function in Eq. (24), to arrive at

$$K_1(q) = -\frac{e^2}{mc^2q^2(n-1)}\left(\frac{2mkT}{4\pi^2\hbar^2}\right)^{\frac{1}{2}n} \times \sum_{j=1}^{\infty} {}^P\!\int ds\, e^{-js^2} \times \left\{(n-2)+\frac{2(s^2-t^2/4)}{(\mathbf{s}-\frac{1}{2}\mathbf{t})\cdot\mathbf{t}}\right\}. \quad (25)$$

There is no question as to the validity of taking the principal value in the angular integration, nor of interchanging the order of summation and integration, because the integrand in (25) has been obtained by rewriting the summand in (5) which is perfectly regular. Equation (25) can be reduced to get[7]

$$K_1(q) = -\frac{e^2}{mc^2q^2}\left(\frac{2mkT}{4\pi\hbar^2}\right)^{\frac{1}{2}n}\sum_{j=1}^{\infty}\frac{\pi^{-\frac{1}{2}}}{j^{(n-1)/2}} \times {}^P\!\int_{-\infty}^{\infty} du\, e^{-ju^2}\left\{1+\frac{1}{jt(u-\frac{1}{2}t)}\right\}. \quad (26)$$

The integral here can be rewritten to give

$$K_1(q) = -\frac{e^2}{mc^2q^2}\left(\frac{2mkT}{4\pi\hbar^2}\right)^{\frac{1}{2}n} \times \sum_{j=1}^{\infty}\frac{1}{j^{\frac{1}{2}n}}\int_0^1 dv\,\{1-\exp[-jt^2(1-v^2)/4]\}. \quad (27)$$

Recalling the definitions (7) and (17), we can finally write

$$K_1(q) = -\frac{e^2}{mc^2q^2}\frac{N}{V\zeta(\frac{1}{2}n)}\left(\frac{T}{T_c}\right)^{\frac{1}{2}n} \times \int_0^1 dv\,\{\lim_{\tau\to 0}[F_+(\tau;\tfrac{1}{2}n)-F_+(t^2(1-v^2)/4;\tfrac{1}{2}n)]\}. \quad (28)$$

Now we can proceed to the limit of very small q (i.e., $t \ll 1$). Equation (21) will be used to expand the F_+ functions in this limit; inspection of (21) shows that we must discuss two cases, namely $n > 4$ and $n < 4$. For $n > 4$ and $t < 1$ we have

$$\lim_{\tau\to 0}[F_+(\tau;\tfrac{1}{2}n)-F_+(t^2(1-v^2)/4;\tfrac{1}{2}n)] = \tfrac{1}{4}t^2(1-v^2)\zeta(\tfrac{1}{2}n-1)(1+\mathcal{O}t^{(n-4)/2}), \quad (29)$$

whereas for $2 < n < 4$ and small t we get

$$\lim_{\tau\to 0}[F_+(\tau;\tfrac{1}{2}n)-F_+(t^2(1-v^2)/4;\tfrac{1}{2}n)] = -\{t^2(1-v^2)/4\}^{\frac{1}{2}n-1}\Gamma(1-\tfrac{1}{2}n)(1+\mathcal{O}t^{(4-n)/2}). \quad (30)$$

These two expressions are divided by the special case of $n = 4$ where one can show

$$\lim_{\tau\to 0}[F_+(\tau;2)-F_+(t^2(1-v^2)/4;2)] = -\tfrac{1}{4}t^2(1-v^2)\log\left(\frac{t^2(1-v^2)}{4e^2}\right)(1+\mathcal{O}t^2). \quad (30')$$

The integration over v is now trivial and we finally have for $n > 4$

$$K_1(q) \underset{q\to 0}{\to} -\frac{e^2}{6mc^2}\left(\frac{\hbar^2}{2mkT}\right)\frac{\zeta(\frac{1}{2}n-1)}{\zeta(\frac{1}{2}n)}\cdot\frac{N}{V}\left(\frac{T}{T_c}\right)^{\frac{1}{2}n}, \quad (31)$$

and for $2 < n < 4$ [using the relation $\Gamma(z)\Gamma(1-z) = \pi/\sin(\pi z)$] we get

$$K_1(q) \underset{q\to 0}{\to} -\frac{e^2}{\gamma mc^2q^2}\left(\frac{\hbar^2q^2}{2mkT}\right)^{\frac{1}{2}n-1}\frac{N}{V}\left(\frac{T}{T_c}\right)^{\frac{1}{2}n}. \quad (32)$$

γ is a numerical constant:

$$\gamma = \pi^{-3/2}2^{n-1}\Gamma[\tfrac{1}{2}(n+1)]\zeta(\tfrac{1}{2}n)\sin[\tfrac{1}{2}\pi(n-2)]. \quad (33)$$

The special case $n = 4$ leads to

$$K_1(q) \underset{q\to 0}{\to} -\frac{e^2}{6mc^2}\left(\frac{\hbar^2}{2mkT}\right)\frac{N}{\zeta(2)V}\left(\frac{T}{T_c}\right)^2\log\left(\frac{2mkT}{\hbar^2q^2}\right). \quad (34)$$

Combination of these results for the noncondensed bosons with the contribution (23) for the condensed bosons leads to the results (2) and (3) stated in the Introduction.

3. HOMOGENEOUS MAGNETIC FIELDS

In this section we restrict attention to the charged Bose gas in ($n > 2$) dimensions (see Sec. 2B). The reason for this restriction is simple: all other cases have a response function $K(q)$ which is regular at $q = 0$, and so the behavior in a homogeneous field can be found by taking $q = 0$ in the results of Sec. 2. For the Bose gas with $n > 2$ below T_c, $K(q)$ is singular at the origin and to obtain the behavior in a homogeneous field it is necessary to include the field in zeroth order.[5]

For an ideal charged Bose gas in an acting homogeneous field B in n-dimensional space ($n > 2$), the grand canonical partition function $\exp(-\alpha\Omega)$ may be derived in a standard way[10]:

$$\Omega = kT\frac{eB}{\hbar c}\frac{L^2}{2\pi}\sum_{\mathbf{k}}\sum_{\nu=0}^{\infty}\log\{1-\exp[-\alpha(\mu_0 B(2\nu+1)+\hbar^2k^2/(2m)-\mu)]\}. \tag{35}$$

Here we have assumed periodic boundary conditions, so that $\mathbf{k} = L\boldsymbol{\kappa}/2\pi$ with $\boldsymbol{\kappa}$ being the points on a unit cubic lattice in $(n-2)$ dimensions. The other 2 dimensions are involved with the magnetic field to give the eigenvalues with the summation index ν. μ_0 is the Bohr magneton, $\mu_0 = e\hbar/2mc$; to avoid negative occupation numbers we must have $\mu < \mu_0 B$.

Replacing sums by integrals, and expanding the logarithm to sum over ν, we get the n-dimensional generalization of the expressions in Ref. 1 and 2:

$$\omega \equiv \frac{\Omega}{V} = -\frac{kT}{\lambda^n}\sum_{j=1}^{\infty}\frac{e^{-2jxz}}{j^{\frac{1}{2}n+1}}\left(\frac{2jx}{1-e^{-2jx}}\right), \tag{36}$$

$$\eta \equiv \frac{N}{V} = \frac{1}{\lambda^n}\sum_{j=1}^{\infty}\frac{e^{-2jxz}}{j^{\frac{1}{2}n}}\left(\frac{2jx}{1-e^{-2jx}}\right). \tag{37}$$

For notational convenience we have introduced the variables

$$x \equiv \alpha\mu_0 B, \tag{38}$$

$$z \equiv \tfrac{1}{2}(\mu_0 B - \mu)/\mu_0 B, \tag{39}$$

and the thermal wavelength λ,

$$\lambda^2 \equiv 4\pi\alpha\hbar^2/(2m). \tag{40}$$

We can immediately make an important comment about the expression (37) for N/V. Bearing in mind that $\mu < \mu_0 B$, i.e., $z > 0$, the series in (37) cannot diverge more strongly than the series $\Sigma j^{1-\frac{1}{2}n}$; that is to say, the series in (37) has an upper bound for $n > 4$, but not for $n < 4$. *The Bose gas does not condense in a fixed homogeneous magnetic field for $n < 4$, but does so condense for $n > 4$.*

Just as in the familiar three-dimensional field-free Bose gas, so here for $n > 4$ our replacement of sums by integrals in (35) is incorrect below the critical temperature: the ground state must be treated separately, and expression (37) represents only the remaining bosons outside the ground state.

We now wish to evaluate the magnetization M, which is derived from $\Omega(B)$ by use of the thermodynamic relation

$$M = -1/V(\partial\Omega(B)/\partial B)_{\alpha,\mu}. \tag{41}$$

The B–H relationship then follows from

$$B = H + S(n)M, \tag{42}$$

with $S(n)$ as the surface area of the unit sphere in n dimensions [Eq. (12)]. (Notice that the "acting" field has been identified with the average microscopic field in the gas, B; the argument is again as given in Refs. 1 and 2.) Applying (41) to (36), the magnetization is written as

$$M = \frac{\mu_0}{\lambda^n}\sum_{j=1}^{\infty}\frac{e^{-2jxz}}{j^{\frac{1}{2}n+1}}\left(\frac{2jx}{1-e^{-2jx}}\right) \times\left(\frac{1}{x} - j\left[\frac{e^{2jx}+1}{e^{2jx}-1}\right]\right). \tag{43}$$

We consider separately the two cases of $4 > n > 2$ and $n > 4$, since the analysis will obviously be rather different. In each case it is also necessary to treat separately the opposite extremes of zero temperature, $x \gg 1$, and finite temperatures, $x \ll 1$.

A. $4 > n > 2$. No Condensation in Homogeneous Field

First we consider the zero-temperature limit, $x \gg 1$. Here we can immediately write

$$M = -(2x\mu_0/\lambda^n)F_+(2xz: \tfrac{1}{2}n - 1). \tag{44}$$

Since there is no condensation, Eq. (37) relates μ to N/V, and we get for $x \gg 1$

$$M = -\mu_0\eta. \tag{45}$$

Secondly we consider the less trivial case where $x \ll 1$. Here it is shown in Appendix A that if $T > T_c$, where T_c is the condensation temperature in the absence of a field [Eq. (7)], then

$$M = -\chi B \tag{46}$$

with

$$\chi = \mu_0^2\eta/3kT, \quad \text{if} \quad T \gg T_c, \tag{47}$$

and

$$\chi = \frac{\mu_0^2\eta}{6kT}\left\{(n-2)\left[\left(\frac{T}{T_c}\right)^{\frac{1}{2}n} - 1\right]\right\}^{(n-4)/(n-2)} \times\left\{\frac{2\Gamma(2-\frac{1}{2}n)}{\zeta(\frac{1}{2}n)}\left(\frac{T}{T_c}\right)^{\frac{1}{2}n}\right\}^{2/(n-2)}, \quad \text{if} \quad T \gtrsim T_c. \tag{48}$$

On the other hand, if $T < T_c$, then

$$M = -\mu_0\eta\{1 - (T/T_c)^{\frac{1}{2}n}\}. \tag{49}$$

This latter result joins to the zero-temperature result (45).

[10] W. Pauli, *Proceedings of the Solvay Congress 1930* (Gauthier-Villars, Paris, 1932) pp. 183–190.

Above T_c, the magnetization amounts to a small diamagnetism in agreement with (15). Below T_c, even though there is no condensation and no macroscopic occupation of the ground state, there is, nevertheless, a spontaneous magnetization (negative in sense) given by Eq. (49). The B–H law below T_c can be written

$$\begin{aligned} B &= H - H_c; \qquad H > H_c \\ \dot{B} &= 0; \qquad\quad H < H_c \end{aligned} \tag{50}$$

This corresponds to expulsion of applied fields which are less than a certain critical value, H_c,

$$H_c = S(n)\mu_0\eta\{1 - (T/T_c)^{\frac{1}{2}n}\}. \tag{51}$$

B. $n > 4$. Condensation in Homogeneous Field

We consider first finite temperatures, $x \ll 1$. Now for $n > 4$, the series in (37) converges as $z \to 0$, and so it can be treated more simply than was the case for $n < 4$. For $x \ll 1$ we have

$$\eta = (1/\lambda^n)F_+(2xz; \tfrac{1}{2}n)\{1 + \mathcal{O}x\}. \tag{52}$$

Ignoring the correction terms, we have an expression for N/V which is identical with that for the n dimensional Bose gas in the absence of a magnetic field (except that $\mu \to \mu - \mu_0 B$). Thus below a critical temperature, given by Eq. (7), we have $z \to 0$ and a macroscopic number of bosons, $\eta_0 = N_0/V$ [given by Eq. (8)], occupy the singular ground state. The other states, which accommodate the remaining $\eta_1 = \eta - \eta_0$ bosons, obey Eq. (52):

$$\eta_1 = \eta(T/T_c)^{\frac{1}{2}n}\{1 + \mathcal{O}x\}. \tag{53}$$

For $n > 4$ and $x \ll 1$ it is similarly a straightforward matter to find the magnetization *above* T_c:

$$M = -(\mu_0 x/3\lambda^n)F_+(2zx; \tfrac{1}{2}n - 1). \tag{54}$$

Below T_c it is necessary to treat the ground state separately, and the result is

$$M = \lim_{s\to 0}\left\{-\frac{\mu_0}{V}\left(\frac{1}{e^{2xs} - 1}\right) - \frac{\mu_0 x}{3\lambda^n}F_+(2xz; \tfrac{1}{2}n - 1)\right\}. \tag{55}$$

Comparing this with the corresponding expression for η, we get

$$M = -\mu_0\eta_0 - \tfrac{1}{3}\mu_0 x\eta_1\zeta(\tfrac{1}{2}n - 1)/\zeta(\tfrac{1}{2}n). \tag{56}$$

That is, for $T < T_c$,

$$M = -\mu_0\eta\left\{\left[1 - \left(\frac{T}{T_c}\right)^{\frac{1}{2}n}\right] + \frac{\mu_0 B}{3kT}\left(\frac{T}{T_c}\right)^{\frac{1}{2}n}\frac{\zeta(\frac{1}{2}n - 1)}{\zeta(\frac{1}{2}n)}\right\}. \tag{57}$$

The above analysis of the case $x \ll 1$ has been predicated upon the assumption $x_c \ll 1$, i.e., $\mu_0 B \ll kT_c$. Retaining this assumption (see Appendix B), it is easy to see that for "zero" temperature, $x \gg 1$, we again get the result (45) which joins to Eq. (57).

As in Part (A) of this section, we have found that above T_c there is a diamagnetism in accord with Eq. (15); below T_c there is a spontaneous magnetization which leads again to the B–H law (50), with H_c given by (51).

The noteworthy feature is that for $n > 4$ the field expulsion comes explicitly from the condensed ground-state bosons [the noncondensed bosons contribute the small diamagnetic term in (57)], whereas for $2 < n < 4$ the same field expulsion is achieved without any condensation.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge stimulating discussions with Professor M. J. Buckingham, and the excellent facilities provided by Professor H. Messel. This work was supported in part by the Nuclear Research Foundation within the University of Sydney.

APPENDIX A

In this Appendix we calculate M as a function of N/V for $4 > n > 2$ and $x \ll 1$.

For $x \ll 1$ and $xz < 1$ we can replace the sum in (43) by an integral to get

$$M = -(\mu_0/\lambda^n)(2x)^{\frac{1}{2}n-1}J(z) \tag{58}$$

with

$$J(z) = \int_0^\infty \frac{ds e^{-ss}}{s^{\frac{1}{2}n}}\left(\frac{s}{1 - e^{-s}}\right)\left(\frac{e^s + 1}{e^s - 1} - \frac{2}{s}\right). \tag{59}$$

We want a similar relation between η and z. Before replacing the sum in (37) by an integral, we first arrange that the integral will converge at its lower limit:

$$\eta = (1/\lambda^n)\{\zeta(\tfrac{1}{2}n) + (2x)^{\frac{1}{2}n-1}I(z)\} \tag{60}$$

with

$$I(z) = \int_0^\infty \frac{ds}{s^{\frac{1}{2}n}}\left(\frac{se^{-ss}}{1 - e^{-s}} - 1\right). \tag{61}$$

Using the identity (7) which defines T_c in the absence of a field, Eq. (60) becomes

$$\eta\lambda^n\{1 - (T/T_c)^{\frac{1}{2}n}\} = (2x)^{\frac{1}{2}n-1}I(z). \tag{62}$$

Limiting forms for $I(z)$ are easily obtained. For $z > 1$,

$$I(z) = \int_0^\infty \frac{ds}{s}\{e^{-sz}(1 + \mathcal{O}s) - 1\}s^{1-\frac{1}{2}n}, \tag{63}$$

which for $4 > n > 2$ gives

$$I(z) = -\frac{2\Gamma(2 - \frac{1}{2}n)}{(n-2)}z^{\frac{1}{2}n-1}\left(1 + \mathcal{O}\frac{1}{z}\right). \tag{64}$$

For $z < 1$,

$$I(z) = \int_0^\infty dse^{-sz}s^{1-\frac{1}{2}n} + \mathcal{O}z^{\frac{1}{2}n-1}, \tag{65}$$

which for $4 > n > 2$ gives

$$I(z) = \Gamma(2 - \tfrac{1}{2}n)z^{\frac{1}{2}n-2}(1 + \mathcal{O}z). \tag{66}$$

The corresponding limiting forms of $J(z)$ are

$$J(z) \overset{z\geq 1}{\to} \tfrac{1}{6}\Gamma(2 - \tfrac{1}{2}n)z^{\frac{1}{2}n-2}(1 + \mathcal{O}1/z), \tag{67}$$

$$J(z) \overset{z\leq 1}{\to} \Gamma(2 - \tfrac{1}{2}n)z^{\frac{1}{2}n-2}(1 + \mathcal{O}z). \tag{68}$$

Substitution of (64) and (66) into (62) shows that $z < 1$ corresponds to $T < T_c$, and $z > 1$ to $T > T_c$. Elimination of z between (58) and (62) leads to

$$M = -\mu_0\eta\{1 - (T/T_c)^{\frac{1}{2}n}\} \tag{69}$$

for $T < T_c$, and to

$$M = -\frac{\mu_0 x\eta}{6}\frac{(n-2)\{(T/T_c)^{\frac{1}{2}n} - 1\}}{(2xz)} \tag{70}$$

with $(2xz)$ to be found from (64) and (62) for $T > T_c$. These are the results quoted in (48) and (49).

As we noted en route, the replacement of sums by integrals is valid when $x \ll 1$ if $xz < 1$. Thus we must distinguish between the situation $x^{-1} > z > 1$ which gives (64), (67), and hence (70), and the situation where z is so large that $z > x^{-1}$. Inspection of the results above shows that $x^{-1} > z > 1$ corresponds to $1 > (T - T_c)/T_c > 0$, and $z > x^{-1}$ to $T \gg T_c$. In this latter case, both the series (37) and (43) are simply dominated by their leading term, and we get Eq. (47) for $T \gg T_c$.

APPENDIX B

The assumption made in Sec. 3 is most certainly true for any reasonable assumption about e, m, η, and B in three dimensions. However, since we are concerned with an abstract mathematic model, we should also mention the results if $\mu_0 B \gg kT_c$ in $n > 4$ dimensions. In this event the Bose gas condenses below a temperature T_0 which is related to the field free condensation temperature T_c by

$$\frac{T_0}{T_c} = \left(\frac{\zeta(\frac{1}{2}n)}{2\Gamma(\frac{1}{2}n - 1)x_c}\right)^{2/(n-2)} < 1 \tag{71}$$

($x_c \equiv \mu_0 B/kT_c \gg 1$). Comparison of the expressions for M and η both above and below this new transition temperature, T_0, shows that

$$M = -\mu_0\eta \tag{72}$$

so long as $x \gg 1$, i.e., so long as $T \ll T_c x_c$. Since $T_c x_c \gg T_c \gg T_0$, Eq. (72) will be valid even far above T_0. Once the temperature is sufficiently large ($x < 1$), the system will again be diamagnetic, described by Eqs. (52) and (54).

A Proof of Nakanishi's Inequality*

J. B. BOYLING
Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, England

(Received 11 September 1964)

Nakanishi's conjectured inequality for the coefficients of the external masses and invariants in the Feynman denominators for scattering processes is proved by a determinantal method suggested by the analogy between Feynman diagrams and electrical networks.

IN the absence of spin, the contribution to a given invariant amplitude of a connected Feynman diagram G with N external lines, n internal lines, v vertices and $l = n - v + 1$ independent closed loops may be written[1-3] in the form

$$\text{const} \lim_{\epsilon \to 0+} \left(\prod_{i=1}^{n} \int_0^1 d\alpha_i \right) \frac{\delta(1 - \sum_{i=1}^{n} \alpha_i)}{[\mathrm{U}(\alpha)]^2 [\mathrm{V}(p, \alpha) - i\epsilon]^{n-2l}}, \quad (1)$$

where the integration variables α_i are the Feynman parameters associated with the internal lines i of G.

The function $U(\alpha)$ may be defined by

$$\mathrm{U}(\alpha) = \sum \alpha_{i_1} \alpha_{i_2} \cdots \alpha_{i_l}, \quad (2)$$

summed over all sets of l internal lines $\{i_1, i_2, \cdots, i_l\}$ having at least one member in every closed loop of G. When G contains no closed loops, U is defined to be unity.

$V(p, \alpha)$ is of the form[3]

$$\mathrm{V}(p, \alpha) = \sum_{i=1}^{n} \alpha_i m_i^2 - Q(p, \alpha), \quad (3)$$

where m_i is the mass of the internal line i and

$$Q(p, \alpha) = \sum_h \zeta_h(\alpha) s_h. \quad (4)$$

The summation in (4) is over all the channels h of the given amplitude, where by a channel $h = (h_1 \mid h_2)$ we mean a partitioning of the external lines, $\beta = 1, \cdots, N$, into two disjoint classes h_1 and h_2 each having at least one member. The quantity s_h is the corresponding channel invariant, given in terms of the inwardly measured external momenta p_β satisfying

$$\sum_{\beta=1}^{N} p_\beta = 0 \quad (5)$$

by

$$s_h = \left(\sum_{\beta \in h_1} p_\beta \right)^2 = \left(\sum_{\beta \in h_2} p_\beta \right)^2. \quad (6)$$

When h is such that h_1 or h_2 consists of a single external line β of mass m_β, s_h takes the constant value m_β^2. The functions $\zeta_h(\alpha)$ are of the form[3]

$$\zeta_h(\alpha) = \mathrm{W}_h(\alpha)/\mathrm{U}(\alpha), \quad (7)$$

where

$$\mathrm{W}_h(\alpha) = \sum_{S \in \mathbf{S}(h)} \mathrm{U}_{H_1}(\alpha) \mathrm{U}_{H_2}(\alpha) \prod_{i \in S} \alpha_i. \quad (8)$$

The symbol S in (8) stands for an intermediate state, i.e., a set S of internal lines such that G separates into two parts H_1 and H_2 when all $i \in S$ are opened up, while every diagram obtained from G by opening up the lines of a proper subset of S is connected. $\mathbf{S}(h)$ denotes the set of all S such that $(h_1 \mid h_2) = h$, where h_1 and h_2 are the sets of external lines of G attached to H_1 and H_2, respectively.

In the case of scattering ($N = 4$), it is natural to write

$$Q(p, \alpha) = Q(s, t, u; \alpha) = \sum_{\beta=1}^{4} \zeta_\beta m_\beta^2 + \zeta_s s + \zeta_t t + \zeta_u u$$

$$= [U(\alpha)]^{-1} \left[\sum_{\beta=1}^{4} W_\beta m_\beta^2 + W_s s + W_t t + W_u u \right], \quad (9)$$

where s, t, and u are the usual invariants

$$\begin{aligned} s &= (p_1 + p_2)^2 = (p_3 + p_4)^2, \\ t &= (p_1 + p_3)^2 = (p_2 + p_4)^2, \\ u &= (p_1 + p_4)^2 = (p_2 + p_3)^2, \end{aligned} \quad (10)$$

which satisfy

$$s + t + u = \sum_{\beta=1}^{4} m_\beta^2. \quad (11)$$

* The research reported in this document has been sponsored by the Air Force Office of Scientific Research, OAR, under Grant No. AF EOAR 63-79 with the European Office of Aerospace Research, U. S. Air Force.

[1] Y. Nambu, Nuovo Cimento **6**, 1064 (1957).

[2] K. Symanzik, Progr. Theoret. Phys. (Kyoto) **20**, 690 (1958).

[3] N. Nakanishi, Progr. Theoret. Phys. (Kyoto) Suppl. **18**, 1 (1961); Progr. Theoret. Phys. (Kyoto) **26**, 337 and 927, (1961).

It has been conjectured by Nakanishi[3] that the W_k functions in (9) satisfy the inequalities

$$\begin{aligned} W_1W_4 &\geq W_sW_t, & W_2W_3 &\geq W_sW_t, \\ W_1W_2 &\geq W_tW_u, & W_3W_4 &\geq W_tW_u, \\ W_1W_3 &\geq W_uW_s, & W_2W_4 &\geq W_uW_s, \end{aligned} \tag{12}$$

for all nonnegative values of the α's. No counterexamples disproving (12) have been found, and it has already been shown that (12) holds for a very wide class of diagrams.[3,4] We shall give a general proof of (12) based on the analogy with electric circuit theory.[5,6] Only the first inequality will be proved, as the other five may be obtained from it by relabeling the external lines.

In the general case, the diagram G less it N external lines may be regarded as a passive electrical network, in which the internal lines are branches of resistance α_i and the vertices, $a = 1, 2, \cdots, v$, are the nodes of the network.[6] For all positive values of the α's, we define the symmetric $v \times v$ matrix B by

$$B_{ab} = \begin{cases} -\sum_{\omega} (\alpha_{ab\omega})^{-1} & \text{for } a \neq b, \\ -\sum_{c \neq a} B_{ac} & \text{for } a = b, \end{cases} \tag{13}$$

where $\sum_\omega$ stands for a summation over all the internal lines $i \leftrightarrow (a, b, \omega)$ which join the vertices a and b directly, and $B_{ab} = 0$ when there are no such lines. The determinant of B vanishes and all its principal minors of order $v - 1$ are equal, their common value Δ being related[2] to $U(\alpha)$ by

$$U(\alpha) = \left(\prod_{i=1}^{n} \alpha_i\right)\Delta. \tag{14}$$

It is clear from (2) and (14) that $\Delta > 0$ provided the α's are all positive.

We now treat an arbitrarily chosen vertex d of G asymmetrically, evaluating Δ as the determinant of the submatrix of B obtained by striking out the row and column corresponding to d. The cofactor of B_{ab} in Δ is denoted by Δ_{ab}, and Δ_{ab} is defined to be zero when a or b equals d. The fundamental equations of electric circuit theory show that the power dissipated in the network when unit current enters at the node a and unit current leaves at the node b is

$$P_{ab} = \Delta^{-1}(\Delta_{aa} + \Delta_{bb} - 2\Delta_{ab}). \tag{15}$$

From (14) and (15) it is not hard to show that

$$P_{ab} = (\prod_{\omega} \alpha_{ab\omega})U^{(ab)}(\alpha)/U(\alpha), \tag{16}$$

where $U^{(ab)}$ is the U function for the diagram $G^{(ab)}$ obtained from G by removing all lines of the type (a, b, ω) and identifying the vertices a and b.

For any pair of external lines, β and γ, we define $Q_{\beta\gamma} = P_{bc}$, where b and c are the vertices at which β and γ are attached. In terms of these functions, $Q(p, \alpha)$ may be expressed in the form[6]

$$Q(p, \alpha) = -\sum_{\beta<\gamma} p_\beta p_\gamma Q_{\beta\gamma}(\alpha). \tag{17}$$

To prove the inequality, $W_1W_4 \geq W_sW_t$, for all connected diagrams G with four external lines, it is sufficient to consider only those diagrams in which the external lines 1, 2, 3, and 4 are attached to four different vertices, which we label 1, 2, 3, and 4, respectively. For, when this is not the case, both sides of the inequality vanish identically. Comparison of (9) and (17) shows that

$$\begin{aligned} \zeta_1 + \zeta_u &= \tfrac{1}{2}(Q_{12} + Q_{13} - Q_{23}), \\ \zeta_4 + \zeta_u &= \tfrac{1}{2}(Q_{24} + Q_{34} - Q_{23}), \\ \zeta_s - \zeta_u &= \tfrac{1}{2}(Q_{14} + Q_{23} - Q_{12} - Q_{34}), \\ \zeta_t - \zeta_u &= \tfrac{1}{2}(Q_{14} + Q_{23} - Q_{13} - Q_{24}), \\ \zeta_1 + \zeta_4 + \zeta_s + \zeta_t &= Q_{14}. \end{aligned} \tag{18}$$

Our labeling convention makes $Q_{\beta\gamma} = P_{\beta\gamma}$ ($\beta, \gamma = 1, 2, 3, 4$), and, taking $d = 1$, we see that, for positive α's, (15) gives

$$\begin{aligned} \zeta_1 + \zeta_u &= \Delta^{-1}\Delta_{23}, \\ \zeta_4 + \zeta_u &= \Delta^{-1}(\Delta_{23} + \Delta_{44} - \Delta_{24} - \Delta_{34}), \\ \zeta_s - \zeta_u &= \Delta^{-1}(\Delta_{34} - \Delta_{23}), \\ \zeta_t - \zeta_u &= \Delta^{-1}(\Delta_{24} - \Delta_{23}). \end{aligned} \tag{19}$$

Hence

$$\begin{aligned} \zeta_1\zeta_4 - \zeta_s\zeta_t &= (\zeta_1 + \zeta_u)(\zeta_4 + \zeta_u) \\ &\quad - (\zeta_s - \zeta_u)(\zeta_t - \zeta_u) - \zeta_u(\zeta_1 + \zeta_4 + \zeta_s + \zeta_t) \\ &= \Delta^{-2}(\Delta_{23}\Delta_{44} - \Delta_{24}\Delta_{34}) - P_{14}\zeta_u. \end{aligned} \tag{20}$$

Now Jacobi's theorem shows that $\Delta_{23}\Delta_{44} - \Delta_{24}\Delta_{34}$ equals Δ times the cofactor of the minor

$$\begin{vmatrix} B_{23} & B_{24} \\ B_{43} & B_{44} \end{vmatrix}$$

in Δ, which is the same as the cofactor of B_{23} in Δ_{44}. Applying (15) to the diagram $G^{(14)}$, we see that

$$\begin{aligned} &\Delta_{23}\Delta_{44} - \Delta_{24}\Delta_{34} \\ &= \Delta\Delta_{44}\tfrac{1}{2}\{Q_{12}^{(14)} + Q_{13}^{(14)} - Q_{23}^{(14)}\} = \Delta\Delta_{44}\zeta_u^{(14)}. \end{aligned} \tag{21}$$

[4] D. B. Fairlie, Proc. Cambridge Phil. Soc. **59**, 157 (1963).
[5] J. Mathews, Phys. Rev. **113**, 381 (1959).
[6] T. T. Wu, Phys. Rev. **123**, 678 (1961).

It also follows from (15) that

$$P_{14} = \Delta^{-1}\Delta_{44}. \tag{22}$$

Substitution of (21) and (22) into (20) now gives

$$\zeta_1\zeta_4 - \zeta_s\zeta_t = P_{14}\{\zeta_u^{(14)} - \zeta_u\}. \tag{23}$$

Multiplying both sides by U^2 and using (7) and (16), we obtain

$$W_1W_4 - W_sW_t = (\prod_\omega \alpha_{14\omega})[UW_u^{(14)} - U^{(14)}W_u]. \tag{24}$$

Since (24) holds for all positive values of the α's and both sides of the equation are polynomials, it must hold identically for all values of the α's.

To show that the right-hand side of (24) is nonnegative when the α's are all nonnegative, we expand W_u and $W_u^{(14)}$ by (8), observing that every intermediate state S contributing to the sum for W_u also contributes to that for $W_u^{(14)}$. The converse is not in general true, as the example of Fig. 1 shows. Here the planar diagram G has no intermediate states in the u channel, but $G^{(14)}$ has one consisting of the internal lines, 1, 3, 5 and 7. Since all the terms in both summations are nonnegative, it is sufficient to show that, for all S contributing to the sum for W_u,

$$UU_{H_2}^{(14)}U_{H_2}\prod_{i\in S}\alpha_i \geq U^{(14)}U_{H_1}U_{H_2}\prod_{i\in S}\alpha_i,$$

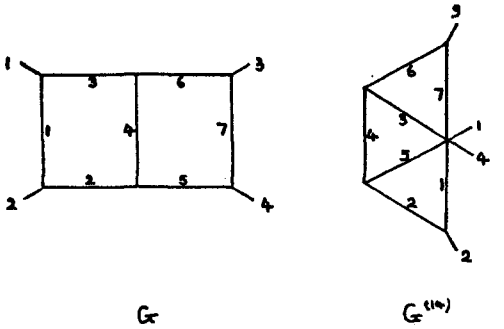


FIG. 1. A case where $G^{(14)}$ has an intermediate state in the u channel though G has none.

where

$$h_1 = \{1, 4\}, \qquad h_2 = \{2, 3\}.$$

By (16) we see that this will hold provided $P_{14}(H_1) \geq P_{14}(G)$. That this is indeed the case follows at once from the theorem that the power dissipated in a passive electrical network supplied by fixed external currents can only increase when the resistances of some of its branches are increased.[6] For H_1 is obtained from G by removing those internal lines i which belong to S and H_2, and this is equivalent to increasing to $+\infty$ the corresponding resistances α_i.

ACKNOWLEDGMENTS

I wish to thank Dr. J. C. Polkinghorne for help and encouragement, and the Department of Scientific and Industrial Research for a grant.

Mixing of Regge Poles and Cuts in a Field-Theory Model*

ARTHUR R. SWIFT†
Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, England

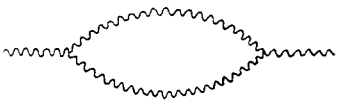
(Received 22 February 1965)

The Regge poles generated by ladder diagrams in a $\lambda\phi^3$ theory are mixed with the moving cuts generated by a class of nonplanar graphs containing an internal ladder. Since the contribution from the cut-generating diagram is of order $t^{-2}\ln^m t$, where t is the asymptotic variable, the t^{-2} behavior of the pure ladder graphs is examined and the trajectory of the Regge pole near $l = -2$ is calculated. The Mellin transform method is used throughout. The transformed amplitude corresponding to a single cut insertion is given by a product of the form pole–cut–pole, where it is only the second Regge trajectory that mixes with the cut. The cut itself depends on the leading trajectory. This result substantiates the predictions as to form of other work based on unitarity, but differs in that the cut and pole depend on different trajectory functions. Finally, multiple insertions of the cut diagrams are shown to generate an amplitude with two moving poles on each sheet of the cut.

1. INTRODUCTION

ANALYTICITY in the angular momentum has come to be recognized as a powerful tool in the analysis of high-energy (or momentum-transfer) behavior of scattering amplitudes; however, the original hypothesis that only a few Regge poles[1] in the crossed channel are necessary to describe the asymptotic behavior[2] has had to be modified by the introduction of a more complicated set of singularities. In particular it seems that not only moving poles but also moving cuts must be considered. The presence of these cuts was deduced by Mandelstam[3] from arguments relating to the Gribov–Pomeranchuk phenomenon[4], and their existence in a class of perturbation theory diagrams was explicitly exhibited by Polkinghorne[5]. Recently Gribov, Pomeranchuk, and Ter-Martirosyan[6] discussed a possible form for the discontinuity associated with these cuts, which they derived heuristically from unitarity. Their form leads to a suggested mixing of Regge poles and cuts near $l = 1$ which would determine the form of the diffraction peak. The dominant contribution was associated with terms, pictorially represented by Fig. 1, which give as a typical contribution

FIG. 1. Cut-generating contribution to a reggeon propagator.



$$\frac{1}{l-\alpha(t)}\int\frac{dt_1'\,dt_2'\,\rho(t,t_1',t_2')}{l-\alpha(t_1')-\alpha(t_2')+1}\frac{1}{l-\alpha(t)}. \tag{1}$$

However, whether the form (1) can actually be obtained depends upon assumptions about the nature of the Reggeon vertices in Fig. 1. It seems worthwhile to attempt to test these assumptions in the "laboratory" of Feynman integrals.

As is well known, the simplest Feynman diagrams leading to a Regge-pole-like a symptotic behavior are the ladder diagrams[7]. Regge cuts, on the other hand, have been shown to arise from a class of nonplanar diagrams containing an internal ladder[5]. The mechanism for generating these cuts is discussed in I. In this paper we insert the cut generating diagrams into a general ladder graph. If a ladder is identified with the propagator of a reggeon, we are analyzing a model of two-particle scattering where the particles coalesce to form a reggeon, which, after propagating, decays into two particles. The reggeon Green's function contains not only a contribution from the bare reggeon, but also a term from an intermediate state of a reggeon and a particle. This class of Feynman diagrams corresponds then to the simplest case considered by Gribov *et al.*,[6] except that our reggeon is in the region of $l = -1$, rather than $l = +1$.

The contribution from a cut generating diagrams is known to be of order $t^{-2}\ln^m t$, where t is the asymptotic variable[5]. Thus, an analysis of ladder graphs containing such an insertion will require a knowledge of the t^{-2} behavior of the external ladders. As far

* The research reported in this document has been sponsored in part by the U. S. Air Force Office of Scientific Research, OAR, under Grant AF EOAR 63-79 with the European Office of Aerospace Research, U. S. Air Force.

† NATO Postdoctoral Fellow.

[1] T. Regge, Nuovo Cimento **14**, 951 (1959); *ibid.* **18**, 947 (1960).

[2] G. F. Chew and S. C. Frautschi, Phys. Rev. Letters **7**, 394 (1961). S. C. Frautschi, M. Gell-Mann, and F. Zachariasen, Phys. Rev., **126**, 2704 (1962).

[3] S. Mandelstam, Nuovo Cimento **30**, 1113, 1127, 1148 (1963).

[4] V. N. Gribov and I. Ya. Pomeranchuk, Phys. Letters **2**, 239 (1962).

[5] J. C. Polkinghorne, J. Math. Phys. **4**, 1396 (1963). (Hereafter referred to as I.)

[6] V. N. Gribov, I. Ya. Pomeranchuk, K. A. Ter-Martirosyan, to be published in Soviet Phys.—JETP.

[7] J. C. Polkinghorne, J. Math. Phys. **5**, 431 (1964).

as the t^{-1} behavior is concerned, the insertion will represent a higher-order correction to the leading Regge pole near $l = -1$. Therefore, we must be able to extract the t^{-2} behavior of a simple ladder diagram. This problem is interesting in its own right since it will involve contributions from a recurrence of the leading Regge pole, due to the fact that we use Mellin transforms rather than Legendre transforms, as well as contributions from a new pole that approaches $l = -2$ in the weak coupling limit. The presence of a second pole is expected by analogy with potential theory, but its exact nature has not heretofore been examined. Once the problem of extracting the t^{-2} terms is solved for the ladder diagram, the insertion of an internal reggeon proves comparatively simple. In fact, we can consider not only a single insertion which leads to a form analogous to (1), but also the sum of all possible insertions of the cut generating diagram; the result is analogous to that obtained from a sum of bubble diagrams and leads to an amplitude of a totally different form than that in (1).

Of the several techniques developed for extracting the asymptotic behavior of Feynman integrals, the most powerful has proved to be that utilizing the Mellin transform of the scattering amplitude[7-9]. If α is the transform variable, a pole at α_0 in the transform corresponds to an asymptotic behavior of t^{α_0}. In Sec. 2 we show that although the leading Regge pole contributes to a pole in α near $\alpha = -2$, its effect can be separated from that of the second pole. In Sec. 3 we actually calculate the singularities of the ladder diagrams near $\alpha = -2$, and show that the expected separation does indeed occur. In the next section we insert the reggeon-particle intermediate state into the ladder and show that the result sums into the desired form. In the final section we briefly discuss the iteration of the cut generating graphs as well as other possible insertions. The appendices contain some of the algebraic details needed in the body of the paper.

2. MELLIN TRANSFORM OF A REGGE POLE

The Mellin transform $F(\alpha)$ of a function $f(t)$ is defined to be

$$F(\alpha) = \int_0^\infty t^{-\alpha-1} f(t)\, dt. \tag{2}$$

Equation (1) may be inverted to give

$$f(t) = \frac{1}{2\pi i} \int_C t^\alpha F(\alpha)\, d\alpha, \tag{3}$$

where the contour C is a line parallel to the imaginary α axis from $-i\infty$ to $+i\infty$. For the purposes of this paper, the contour will be taken such that $-1 < \mathrm{Re}\ \alpha < 0$. From (3) $f(t)$ is seen to vanish as $t^{re\alpha}$, or faster. The contour C may be shifted to the left until it meets any singularities of $F(\alpha)$. If $F(\alpha)$ has a pole at α_0 as well as a cut in α, then $f(t)$ will go as

$$f(t) = \frac{1}{2\pi i} \int_C t^\alpha \left[\frac{F(\alpha_0)}{\alpha - \alpha_0(s)} + \int \frac{\rho(s')\, ds'}{\alpha - \alpha_1(s')} \right] d\alpha$$

$$= F(\alpha_0) t^{\alpha_0(s)} + \int \rho(s') t^{\alpha_1(s')}\, ds'. \tag{4}$$

Thus, terms in $f(t)$ of order $t^{-1+\epsilon}$ come from poles or branch points in $F(\alpha)$ near $\alpha = -1$, and terms which fall off faster will be given by poles, or other singularities further to the left in the α plane.

If the scattering amplitude $f(s, t)$ has a Regge pole at $l = l_p(s)$, then if $l_p < -\frac{1}{2}$, the contribution of the pole to $f(s, t)$ has the form[10]

$$f(s, t) = -\frac{(2l_p + 1)\beta_p(s) Q_{-l_p-1}[-1 - 2t/(s - 4m^2)]}{\pi \cos \pi l_p}. \tag{5}$$

$Q_\nu(z)$ is the Legendre function of the second kind. Since the scattering amplitude has cuts for positive real t, we examine the limit $t \to -\infty$.[7] The Mellin transform of $f(s, -\tau)$, where $\tau = -t$, is given by[11]

$$F(s, \alpha) = -\frac{(2l_p + 1)\beta_p(s)}{2\pi \cos \pi l_p} \frac{(s - 4m^2)^{-\alpha} [\Gamma(-\alpha)]^2}{\Gamma(-l_p - \alpha)} \times \Gamma(\alpha - l_p) \cos \pi(\alpha + l_p). \tag{6}$$

Due to the factor $\Gamma(\alpha - l_p)$, $F(s, \alpha)$ will have a series of poles in α when $\alpha - l_p(s)$ is equal to a negative integer; all other terms are regular for negative α and l_p. In the weak coupling limit the leading Regge pole from the ladder diagrams has $l_1(s) = -1 + \delta_1(s)$, where $\delta_1(s)$ is small. We then find that if $\alpha \to -1$, -2, $F(s, \alpha)$ takes the form

$$\underset{\alpha \to -1}{F(s, \alpha)} = -\beta_1(s)(s - 4m^2)/2\pi(\alpha + 1 - \delta_1(s)), \tag{7}$$

$$\underset{\alpha \to -2}{F(s, \alpha)} = -\beta_1(s)(s - 4m^2)^2/4\pi[\alpha + 2 - \delta_1(s)], \tag{8}$$

[8] J. D. Bjorken and T. T. Wu, Phys. Rev. **130**, 2566 (1963).
[9] T. L. Trueman and T. Yao, Phys. Rev. **132**, 2741 (1963).
[10] S. Mandelstam, Ann. Phys. (N. Y.) **19**, 254 (1962).
[11] *Tables of Integral Transforms*, edited by A. Erdelyi (McGraw-Hill Book Company, Inc., New York, 1954), Vol. II, pp. 316, 204.

while if $l_p \to -2$, and $\alpha \to -2$, we find

$$F(s, \alpha) = \beta_2(s)(s - 4m^2)^2/4\pi[\alpha + 2 - \delta_2(s)]. \quad (9)$$

Hence, near $\alpha = -2$, the two poles in the angular momentum plane contribute in the same way to $F(s, \alpha)$, except that their residues and trajectory functions will be different. For the ladder diagrams $\beta_1(s)$ and $\delta_1(s)$ are well known. Hence, a calculation of $F(s, \alpha)$ near $\alpha = -2$ should produce two terms corresponding to (8) and (9) from which $\beta_2(s)$ and $\delta_2(s)$ can easily be determined.

3. BEHAVIOR AT $\alpha = -2$

The contribution to the total scattering amplitude of a ladder diagram with $N + 1$ rungs can be written in the form[3]

$$f_N(s, -\tau) = \lambda^2\left(\frac{\lambda^2}{16\pi^2}\right)^N \int_0^\infty \frac{\prod_{i=1}^{N} dy_i\, dz_i \prod_{i=1}^{N+1} dx_i}{(\Delta_N)^2} \times \exp\left[-\frac{g\tau}{\Delta_N} - Q_N\right], \quad (10)$$

where $g = \prod_{i=1}^{N+1} x_i$ and Q_N and Δ_N are the same functions as defined in Ref. 7. The parameters x_i, y_i, z_i are as labeled in Fig. 2. We have again replaced t by $-\tau$. Let us define the following quantities:

$$\begin{aligned} \sigma_i &= y_i + z_i, \\ \theta_i &= x_i + \sigma_i + x_{i+1} - x_i^2/\theta_{i+1}, \\ \theta_1 &= x_1 + \sigma_1 + x_2, \\ \phi_i(y) &= y_i + x_i\phi_{i-1}(y)/\theta_{i-1}, \qquad \phi_1(y) = y_1. \end{aligned} \quad (11)$$

Then, as discussed in Appendix A, Q_N and Δ_N are given by

$$\Delta_N = \prod_{i=1}^{N} \theta_i \quad (12)$$

$$\begin{aligned} Q_N = {} & m^2 \sum_{i=1}^{N} \sigma_i + m^2 \sum_{i=2}^{N} x_i \\ & + m^2\left(\frac{x_{N+1}^2}{\theta_N} + \frac{2x_1 \cdots x_N}{\theta_1 \cdots \theta_N}\right) \\ & + m^2 \sum_{i=1}^{N} \left(\frac{x_1 \cdots x_i}{\theta_1 \cdots \theta_{i-1}}\right)^2 \frac{1}{\theta_i} \\ & - s \sum_{i=1}^{N} \frac{\phi_i(y)\phi_i(z)}{\theta_i} = m^2\mathfrak{M}_N - s\mathfrak{S}_N. \end{aligned} \quad (13)$$

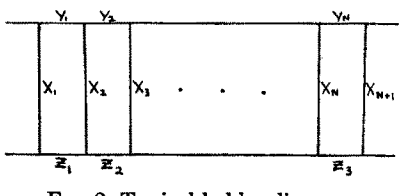


FIG. 2. Typical ladder diagram.

The representations of Δ_N and Q_N given in (12) and (13), while unfamiliar, prove to be essential for the following discussion.

The Mellin transform of (10) is

$$F_N(s, \alpha) = \lambda^2\left(\frac{\lambda^2}{16\pi^2}\right)^N \Gamma(-\alpha) \times \int_0^\infty dy_i\, dz_i\, dx_i \frac{(x_1 \cdots x_{N+1})^\alpha}{\Delta_N^{2+\alpha}} e^{-Q_N}. \quad (14)$$

$F_N(s, \alpha)$ is well defined for $-1 < \mathrm{Re}\ \alpha < 0$, but contains a singularity as $\alpha \to -1$ due to the $(x_i)^\alpha$ factors. We apply the now standard trick of integration by parts to isolate this singularity[7,9]; in fact, since we are interested in (14) near $\alpha = -2$, we integrate by parts twice to obtain

$$\begin{aligned} F_N(s, \alpha) = {} & \lambda^2\left(\frac{\lambda^2}{16\pi^2}\right)^N \frac{\Gamma(-\alpha)}{(\alpha + 1)^{N+1}(\alpha + 2)^{N+1}} \\ & \times \int_0^\infty dx_i\, (x_1 \cdots x_{N+1})^{\alpha+2} \\ & \times \left[\frac{\partial^{2N+2}}{\partial x_1^2 \cdots \partial x_{N+1}^2} \frac{e^{-Q_N}}{\Delta_N^{2+\alpha}}\right]. \end{aligned} \quad (15)$$

If we now set $\alpha = -2$, except in the pole terms, the x_i integrations can be done. Thus, the most singular part of $F_N(s, \alpha)$ at $\alpha = -2$ is

$$F_N(s, \alpha) = \lambda^2\left(\frac{\lambda^2}{16\pi^2}\right)^N \frac{1}{(\alpha + 2)^{N+1}} \times \int_0^\infty dy_i\, dz_i \left[\frac{\partial^{N+1}}{\partial x_1 \cdots \partial x_{N+1}} e^{-Q_N}\right]_{x_i=0}. \quad (16)$$

The point $\alpha = -2$ is unique in that the dependence on Δ disappears. This results in considerable simplification.

To evaluate the integral in (16) we need to examine the detailed structure of Q_N. If there were no derivatives present, then Q_N would become $Q_N(0)$ where

$$Q_N(0) = m^2 \sum_{i=1}^{N} \sigma_i - s \sum_{i=1}^{N} \frac{y_i z_i}{\sigma_i}. \quad (17)$$

For the present problem all those terms of Q_N which are linear in each x_i must be retained; all higher powers do not contribute. Thus, the coefficient of m^2, $\mathfrak{M}_N$, becomes

$$\mathfrak{M}_N = \sum_{i=1}^{N} \sigma_i + \sum_{i=2}^{N} x_i + \frac{2x_1 \cdots x_{N+1}}{\sigma_1 \cdots \sigma_N}, \quad (18)$$

since $\theta_i = \sigma_i$ in this limit. Consider next $\mathfrak{S}_N$:

$$\mathfrak{S}_N = \mathfrak{S}_{N-1} + \left(y_N + \frac{x_N\phi_{N-1}(y)}{\theta_{N-1}}\right)\frac{1}{\theta_N}\left(z_N + \frac{x_N\phi_{N-1}(z)}{\theta_{N-1}}\right). \quad (19)$$

If the second term of (19) is expanded by use of (11) and only terms linear in each x_i are kept, we find that

$$\begin{aligned}\mathcal{S}_N &= \mathcal{S}_{N-1} + \frac{y_N z_N}{\sigma_N}\left(1 - \frac{x_N}{\sigma_N} - \frac{x_{N+1}}{\sigma_N} + \frac{2x_N x_{N+1}}{\sigma_N^2}\right)\\ &+ \frac{x_N \alpha_{NN-1}}{\sigma_N \sigma_{N-1}}\left(1 - \frac{x_{N+1}}{\sigma_N} - \frac{x_{N-1}}{\sigma_{N-1}} + \frac{x_{N-1}x_{N+1}}{\sigma_N \sigma_{N-1}}\right)\\ &+ \frac{x_N x_{N-1}\alpha_{NN-2}}{\sigma_N \sigma_{N-1}\sigma_{N-2}}\left(1 - \frac{x_{N+1}}{\sigma_N} - \frac{x_{N-2}}{\sigma_{N-2}} + \frac{x_{N-2}x_{N+1}}{\sigma_N \sigma_{N-2}}\right)\\ &+ \cdots\\ &+ \frac{x_N \cdots x_2 \alpha_{N1}}{\sigma_N \cdots \sigma_1}\left(1 - \frac{x_{N+1}}{\sigma_N} - \frac{x_1}{\sigma_1} + \frac{x_{N+1}x_1}{\sigma_N \sigma_1}\right), \qquad (20)\end{aligned}$$

where $\alpha_{Ni} = y_N z_i + z_N y_i$. The derivative with respect to x_N in (16) will generate a sum of terms of the form

$$\int dy_i\, dz_i \int dy_N\, dz_N\, \alpha_{Ni} f_{Ni}(y_N, z_N). \qquad (21)$$

In as much as $\mathcal{S}_{N-1}$ is independent of y_N, z_N, all the terms in (20) that remain after x_N is set equal to zero are symmetric in y_N and z_N. Thus, regardless of what other properties $f_{Ni}(y_N, z_N)$ may have, it is surely symmetric in y_N and z_N. Using this symmetry property, we can replace α_{Ni} by $\frac{1}{2}\sigma_N \sigma_i$. This is equivalent to making the same replacement in (20). When such a change is made in (20), a large number of cancellations occur between successive terms of the sum to leave

$$\begin{aligned}\mathcal{S}_N &= \mathcal{S}_{N-1} + \frac{y_N z_N}{\sigma_N} + x_N\left(\frac{1}{2} - \frac{y_N z_N}{\sigma_N^2}\right)\\ &- x_{N+1}\frac{y_N z_N}{\sigma_N^2} + x_N x_{N+1}\left(\frac{2y_N z_N}{\sigma_N^3} - \frac{1}{2\sigma_N}\right)\\ &- \frac{1}{2}\frac{x_1 \cdots x_N}{\sigma_1 \cdots \sigma_N} + \frac{1}{2}\frac{x_1 \cdots x_{N+1}}{\sigma_1 \cdots \sigma_N}. \qquad (22)\end{aligned}$$

Equation (22), together with the value of $\mathcal{S}_1$,

$$\mathcal{S}_1 = \frac{y_1 z_1}{\sigma_1}\left(1 - \frac{x_1}{\sigma_1} - \frac{x_2}{\sigma_1} + \frac{2x_1 x_2}{\sigma_1^2}\right), \qquad (23)$$

determine $\mathcal{S}_N$ completely.

$$\begin{aligned}\mathcal{S}_N &= \sum_1^N \frac{y_i z_i}{\sigma_i} - \frac{x_i y_i z_i}{\sigma_1^2} - \frac{x_{N+1} y_N z_N}{\sigma_N^2}\\ &+ \sum_2^N x_i\left(\frac{1}{2} - \frac{y_i z_i}{\sigma_i^2} - \frac{y_{i-1}z_{i-1}}{\sigma_{i-1}^2}\right)\\ &+ \sum_1^N x_i x_{i+1}\left(\frac{2y_i z_i}{\sigma_i^2} - \frac{1}{2\sigma_i}\right) + \frac{x_1 \cdots x_{N+1}}{2\sigma_1 \cdots \sigma_N}. \qquad (24)\end{aligned}$$

The integral in (16), $\mathcal{I}_N$, may now be evaluated. From (17), (18), and (24) Q_N is seen to have the form

$$Q_N = Q_N(0) + Q'_N + \frac{x_1 \cdots x_{N+1}}{2\sigma_1 \cdots \sigma_N}(4m^2 - s), \qquad (25)$$

$$\begin{aligned}Q'_N &= Q'_{N-1} + x_N m^2 + (x_{N+1} y_N z_N/\sigma_N^2)s\\ &- x_N\left(\frac{1}{2} - \frac{y_N z_N}{\sigma_N^2}\right)s - x_N x_{N+1}\left(\frac{2y_N z_N}{\sigma_N^3} - \frac{1}{2\sigma_N}\right)s. \qquad (26)\end{aligned}$$

The third term in (25) will contribute only when all $N + 1$ derivatives act upon it. Hence, $\mathcal{I}_N$ is given by

$$\begin{aligned}\mathcal{I}_N &= -\frac{(4m^2 - s)}{2}[A(s)]^N\\ &+ \int_0^\infty dy_i\, dz_i e^{-Q_N(0)}\left[\frac{\partial^{N+1}}{\partial x_1 \cdots \partial x_{N+1}} e^{-Q_N}\right]_{x_i = 0}, \qquad (27)\end{aligned}$$

$A(s)$ is the trajectory function for the leading Regge pole. For $0 < s < 4m^2$, its explicit functional form is

$$\begin{aligned}A(s) &= \int_0^\infty \frac{dy\, dz}{y + z}\exp\left[-m^2(y + z) + \frac{syz}{1 + z}\right]\\ &= \frac{4}{[s(4m^2 - s)]^{\frac{1}{2}}}\tan^{-1}\left(\frac{s}{4m^2 - s}\right)^{\frac{1}{2}}. \qquad (28)\end{aligned}$$

The remaining integral in (27), $\mathcal{I}'_N$, is given in terms of a recursion formula by explicitly evaluating the derivative with respect to x_{N+1} and commuting the remaining derivatives through the result. In this way we obtain

$$\begin{aligned}\mathcal{I}'_N &= -sD(s)\mathcal{I}'_{N-1} + s(2C(s) - \tfrac{1}{2}A(s)\\ &+ (m^2 - \tfrac{1}{2}s)D(s) + sE(s))\mathcal{I}''_{N-1}, \qquad (29)\end{aligned}$$

where

$$\begin{aligned}\mathcal{I}''_N &= \int_0^\infty dy_i\, dz_i e^{-Q_N(0)}\\ &\times\left[\frac{\partial^N}{\partial x_1 \cdots \partial x_N} e^{-Q_N'(x_{N+1}=0)}\right]_{x_i = 0} \qquad (30)\\ &= \beta(s)\mathcal{I}'_{N-1} - ([m^2 - s/2]\beta(s) + sD(s)\mathcal{I}''_{N-1}. \qquad (31)\end{aligned}$$

The various functions of s in (29) and (31) are closely related to $A(s)$,

$$\begin{aligned}C(s) &= \int_0^\infty \frac{dy\, dz\, yz}{(y + z)^2}\exp\left[-m^2(y + z) + s\frac{yz}{y + z}\right]\\ B(s) &= -\frac{\partial A(s)}{\partial m^2},\\ D(s) &= \frac{\partial A(s)}{\partial s} = -\frac{\partial C(s)}{\partial m^2},\\ E(s) &= \frac{\partial C(s)}{\partial s}. \qquad (32)\end{aligned}$$

Equations (29) and (31) form a coupled set of recurrence relations for g'_N and g''_N. If they are written as a two-component column matrix, a solution in matrix form is easily obtained.

$$\begin{pmatrix} g'_N \\ g''_N \end{pmatrix} = \hat{\alpha}\begin{pmatrix} g'_{N-1} \\ g''_{N-1} \end{pmatrix} = \hat{\alpha}^{N-1}\begin{pmatrix} g'_1 \\ g''_1 \end{pmatrix}, \tag{33}$$

$$\hat{\alpha} = \begin{pmatrix} -[2/(4m^2 - s)][1 + \frac{1}{2}(s - 2m^2)A(s)] & 0 \\ [2/(4m^2 - s)][m^2 + A(s)] & -m^{-2} \end{pmatrix}. \tag{34}$$

The functions $B(s)$, $C(s)$, etc., have been explicitly evaluated in terms of $A(s)$; g'_1, g''_1 are easily determined from the defining equations to be

$$\begin{aligned} g'_1 &= (s^2E + 2sC - \tfrac{1}{2}sA) = \tfrac{1}{2}s(s - 2m^2)D(s) \\ &= [(s - 2m^2)/(4m^2 - s)][1 + \tfrac{1}{2}(s - 2m^2)A], \\ g''_1 &= -sD. \end{aligned} \tag{35}$$

Equations (33), (34), and (35) constitute the solution to our problem. Since α_{12} turns out to be zero, we could solve for g'_N without reference to g''_N. However, it is g''_N we shall want in the next section, so we carry through the matrix method here.

If $F(s, \alpha)$ is defined as the sum of all $F_N(s, \alpha)$, $N \geq 1$, together with the Born term, it is given by the following expression:

$$\begin{aligned} F(s, \alpha) &= -\frac{\lambda^2 m^2}{\alpha + 2} \\ &- \frac{\lambda^2(4m^2 - s)}{2(\alpha + 2)}\left[\frac{1}{1 - \gamma A(s)/(\alpha + 2)} - 1\right] \\ &+ \frac{\lambda^2\gamma}{(\alpha + 2)^2}\hat{P}\left(\frac{1}{1 - \gamma\hat{\alpha}/(\alpha + 2)}\right)\begin{pmatrix} g'_1 \\ g''_1 \end{pmatrix}, \end{aligned} \tag{36}$$

where $\gamma = \lambda^2/16\pi^2$ and $\hat{P}$ is a projection operator that selects the upper component of a column matrix. The first term on the right of (36) is the Mellin transform of the Born amplitude near $\alpha = -2$. Inversion of the matrix in (36) is a simple problem in matrix algebra; it is especially simple since $\alpha_{12} = 0$.

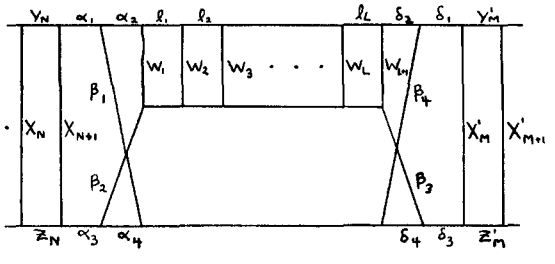


FIG. 3. Ladder diagram with a cut-generating insertion.

when all terms in (36) are combined in an appropriate fashion, $F(s, \alpha)$ becomes

$$\begin{aligned} F(s, \alpha) &= \frac{\lambda^2(s - 4m^2)}{2(\alpha + 2 - \gamma A(s))} \\ &- \frac{\lambda^2(s - 2m^2)}{2\{\alpha + 2 - 2\gamma[1 + \frac{1}{2}(s - 2m^2)A(s)]/(4m^2 - s)\}}. \end{aligned} \tag{37}$$

Equations (37) is to be compared with the sum of (8) and (9); it indeed has the desired form. The parameters of the first two Regge poles are determined to be

$$\begin{aligned} \beta_1(s) &= -2\pi\lambda^2/(s - 4m^2), \qquad \delta_1(s) = (\lambda^2/16\pi^2)A(s), \\ \beta_2(s) &= \frac{-2\pi\lambda^2(s - 2m^2)}{(s - 4m^2)^2}, \\ \delta_2(s) &= -\frac{\lambda^2}{8\pi^2}\frac{[1 + \frac{1}{2}(s - 2m^2)A(s)]}{4m^2 - s}. \end{aligned} \tag{38}$$

Notice that $\delta_2(0) = 0$, whether any significance should be attached to the precise form of $\delta_2(s)$ is not known at this point. Trueman and Yao[9] calculated the $\alpha = -2$ contribution of the ladder diagrams at $s = 0$; as can be seen from (37), they evaluated only the recurrence of the leading pole.

4. REGGE CUTS

We next turn our attention to an analysis of ladder diagrams containing an internal ladder. Since we are looking for Regge-cut behavior, we are forced to consider a diagram such as that in Fig. 3, rather than the corresponding one with uncrossed lines[5]. As shown in I, planar diagrams with internal ladders do not give cuts on the physical sheet. The amplitude corresponding to Fig. 3 may be written as in Eq. (10), but with different values for Q, g, and Δ. The power of $\lambda^2/16\pi^2$ is $N + M + L + 5$ corresponding to ladders of $N + 1$ and $M + 1$ rungs surrounding an internal ladder of $L + 1$ rungs. The appropriate Feynman parameters are labeled in Fig. 3. The coefficient of t is

$$g = L(\Delta'[X + P_2/\Delta'][Y + P_1/\Delta'] + \bar{g})J, \tag{39}$$

where

$$\begin{aligned} L &= x_1 \cdots x_{N+1}, & J &= x'_1 \cdots x'_{m+1}, \\ X &= \alpha_2\alpha_4 - \beta_1\beta_2, & Y &= \delta_2\delta_4 - \beta_3\beta_4, \end{aligned}$$

and Δ' is the same function of the parameters of the internal ladders as Δ_N in (10). The precise forms of Q, Δ, $\bar{g}$, P_1, and P_2 are discussed in Appendix A; it is worth noting that $\bar{g}$ is proportional to $w_1 \cdots w_{L+1}$.

As shown in I, the Regge cut is generated by that part of the integration hypercontour where $X \sim -P_2/\Delta'$ and $Y \sim -P_1/\Delta'$ and, in addition, $\bar{g}$ vanishes. Before evaluating this cut contribution, we must Mellin-transform the amplitude. Since g can be either positive or negative, the Mellin transform is not well defined for either positive or negative t. This is a reflection of the fact that the diagrams under consideration have a nonzero tu spectral function and, therefore, a right- and a left-hand cut in t. Accordingly we let t become $i\tau$ and let τ approach infinity; the amplitude for real t can be obtained by analytic continuation. For pure imaginary t the Mellin transform is well defined independent of the sign of g and is given by

$$F(s,\alpha) = \lambda^2\left(\frac{\lambda^2}{16\pi^2}\right)^{N+M+L+5}(i)^{-\alpha}\Gamma(-\alpha)\int_0^\infty \frac{d\alpha_i g^\alpha}{\Delta^{2+\alpha}}\,e^{-Q} \tag{40}$$

$$= \lambda^2\left(\frac{\lambda^2}{16\pi^2}\right)^{N+M+L+5}(i)^{-\alpha}\Gamma(-\alpha)\int_0^\infty \frac{d\alpha_i (LJ)^\alpha e^{-Q}}{\Delta^{2+\alpha}}$$
$$\times\,\delta(X+P_2/\Delta')\delta(Y+P_1/\Delta')$$
$$\times\int_{X_1}^{X_2}\int_{Y_1}^{Y_2} dX'\,dY'\,[\Delta' X' Y' + \bar{g}]^\alpha, \tag{41}$$

where in going from (40) to (41) we have isolated the cut-generating part of the integral. The limits of the X' integration are such that $X_1 < 0 < X_2$, and similarly for Y'. The X', Y' integral in (41) can be done by breaking it up into a sum of four terms such that the product $X'Y'$ does not change sign within each region of integration. Then, if only those terms which become singular near $\alpha = -2$ are retained, we have from Appendix B

$$\int_{X_1}^{X_2}\int_{Y_1}^{Y_2}[\Delta' X' Y' + \bar{g}]^\alpha = -\frac{2\pi i}{\Delta'}\,\bar{g}^{\alpha+1} + R. \tag{42}$$

The remainder term R does not become singular at $\alpha = -2$, while the $\bar{g}^{\alpha+1}$ term is just the factor we are looking for. Note that if g consists of a product of terms representing a number of cut-generating insertions, the Mellin-transform method allows the problem to be factorized and each insertion will contribute two delta functions and a $\bar{g}^{\alpha+1}$ factor.

Upon combining (42) and (41), we find for the most singular part of $F(s,\alpha)$

$$F(s,\alpha) = 2\pi i\lambda^2\left(\frac{\lambda^2}{16\pi^2}\right)^{N+M+L+5}\int_0^\infty d\alpha_i\,\frac{\bar{g}^{\alpha+1}(LJ)^\alpha}{\Delta'}\,e^{-Q}$$
$$\times\,\delta(X+P_2/\Delta')\delta(Y+P_1/\Delta'). \tag{43}$$

The integrand of (43) is singular at $\alpha = -2$ due to both $\bar{g}^{\alpha+1}$ and $(LJ)^\alpha$. The delta functions may be used to carry out the α_4, δ_4 integrations,

$$\delta(X+P_2/\Delta') = \delta\left(\alpha_4 - \frac{\beta_1(\beta_2+E)}{\alpha_2+B-E}\right)\frac{1}{\alpha_2+B-E},$$
$$\delta(Y+P_1/\Delta') = \delta\left(\delta_4 - \frac{\beta_4(\beta_3+F)}{\delta_2+C-F}\right)\frac{1}{\delta_2+C-F}. \tag{44}$$

B, C, E, and F are functions of the internal ladder parameters and are given in Appendix A; it suffices to say that they vanish when all the w_i vanish. In addition, we know from Appendix A that, under the conditions (44), $\bar{g}$ is equal to $\beta_1\beta_4 w_1 \cdots w_{L+1}\bar{g}'$. Therefore, in order to extract the singular part of (43) near $\alpha = -2$, it is necessary to integrate by parts twice over each x_i and x_i' due to $(LJ)^\alpha$, and once over β_1, β_4 and the w_i. The procedure is the same as in the previous section. After setting $\alpha = -2$ in the resulting integrand, we do the x_i, x_i', β_1, β_4, and ω_i integrations. Q becomes a sum of three independent terms; one can be associated with the internal ladder and the other two with the external ladders. Thus, the integral for $F(s,\alpha)$ factorizes.

$$F(s,\alpha) = 2\pi i\lambda^2\left(\frac{\lambda^2}{16\pi^2}\right)^{N+M+L+5}\frac{1}{(\alpha+2)^{N+M+L+5}}$$
$$\times\left[\int_0^\infty\prod_{i=1}^{N+1}dy_i\,dz_i\right.$$
$$\left.\times\left\{\frac{\partial^{N+1}}{\partial x_1\cdots\partial x_{N+1}}\exp\left[-Q'_{N+1}(x_{N+2}=0)\right]\right\}\right]$$
$$\times\left[\int_0^\infty\prod_0^{L+1}\frac{dl_i\,dn_i e^{-Q_{N''}}}{\Delta_L''}\right]\left[\frac{N\to M}{x_i', y_i, z_i \to x_i', y_i', z_i'}\right]. \tag{45}$$

We have set α_1, $\alpha_3 = y_{N+1}$, z_{N+1}, δ_1, $\delta_3 = y'_{M+1}$, z'_{M+1} and α_2, β_2, δ_2, $\beta_3 = l_0$, n_0, 1_{L+1}, n_{L+1}. The reduction of Q is discussed in Appendix A and the explicit form of Q'_L is given there; Δ'_L is defined as $\bar{g}'\Delta'\alpha_2\delta_2$ under the restrictions (44) and $w_i = \beta_1 = \beta_4 = 0$. Q'_{N+1} is the same function encountered in the previous section, hence; the integrals in (45) referring to the external ladders are just g''_{N+1} and g''_{M+1}, where g''_N is the supplementary integral introduced in the previous section. We can evaluate them directly from (33), (34), and (35) by using a projection operator $\hat{P}'$ that retains only the lower components

of the two-component matrix. If $\gamma = \lambda^2/16\pi^2$, the sum over N in (45) gives a factor of the form

$$\gamma \sum_{N=0}^{\infty} \left(\frac{\gamma}{\alpha+2}\right)^N \mathcal{J}''_{N+1} = \hat{P}' \frac{\gamma}{1-\gamma\hat{a}/(\alpha+2)} \begin{pmatrix} \mathcal{J}'_1 \\ \mathcal{J}''_1 \end{pmatrix}$$
$$= (\alpha+2)\delta_2(s)/[\alpha+2-\delta_2(s)], \qquad (46)$$

where $\delta_2(s)$ is the trajectory function of the second pole which appeared in (37). As can be seen from (25) and (26), setting $x_{N+2} = 0$ in Q_{N+1} removes those terms which would introduce a contribution from the leading pole into (45).

The middle factor in (45) is identical to the integral appearing in Eq. (20) of I, after scaling all the parameters by $(l_i, n_i) = \rho(\bar{l}_i, \bar{n}_i)$. Hence, using the results of I, we can perform the sum over L and obtain

$$\gamma^2 \sum_{L=0}^{\infty} \left(\frac{\gamma}{\alpha+2}\right)^L \left[\int_0^{\infty} \prod_{i=0}^{L+1} \frac{dl_i\, dn_i}{\Delta''_L} e^{-q_L''}\right]$$
$$= 2(\alpha+2) \int_{\lambda\leq 0} \frac{ds_1\, ds_2 \rho(s, s_1, s_2)}{\alpha+2-\delta_1(s_1)}, \qquad (47)$$

where

$$\rho(s, s_1, s_2) = [\delta_1(s)]^2/\lambda(s, s_1, s_2)[s_2 - m^2],$$

and $\lambda = s^2 + s_1^2 + s_2^2 - 2ss_1 - 2ss_2 - 2s_1s_2$; $\delta_1(s)$ is the trajectory function of the leading Regge pole.

Upon putting the various factors together and taking into account the contributions from diagrams with only one external ladder, or no external ladders, we arrive at a final form for $F(s, \alpha)$:

$$F(s, \alpha) = \frac{i\lambda^4}{4\pi} \left\{\frac{1}{\alpha+2-\delta_2(s)}\right\}^2 \int \frac{\rho(s, s_1, s_2)\, ds_1\, ds_2}{\alpha+2-\delta_1(s_1)}. \qquad (48)$$

Equation (48) constitutes the final result. It shows a mixing of pole and cut near $\alpha = -2$ of the form suggested by (1). However, the pole which mixes with the cut generated by the leading Regge trajectory is not the leading pole itself, but rather the second Regge trajectory $\delta_2(s)$. If this model is relevant to the problem of combining cuts and poles near $l = 1$, it does not confirm the conjecture of Gribov, Pomeranchuk, and Ter-Martirosyan[6], for in order to determine the diffraction peak they need to combine a pomeranchukon-generated cut with pomeranchukon poles. The absence of a recurrence of the leading pole in (48) suggests that the leading singularity remains a pure pole.

5. ITERATION OF THE CUT

Equation (48) represents the result of inserting a single cut-generating diagram into a ladder. Although such an amplitude gives the lowest-order cut contribution, there is no reason not to consider the effect of an arbitrary number of insertions. In fact, in a complete description of the amplitude, we must include all possible insertions of the cut-generating graph, not to mention all other possible graphs. Besides, it is always interesting to sum up as large a class of graphs as possible. As will be seen, the problem is very similar to the summing of bubble graphs. From the work of the previous section and the results of Appendix B, we know that a diagram containing p cut insertions will factor into p integrals representing the insertion, $p-1$ integrals for the internal ladders, and two integrals for the external ladders. The factorization property is crucially dependent on the fact that in each cut diagram we set, in the notation of Fig. 3, β_1, α_4, β_4, δ_4 equal to zero. Therefore, a diagram with p insertions each of which contains a ladder of $L_i + 1$ rungs, $p-1$ internal ladders with M_i rungs, and external ladders with N_1 and N_2 rungs, will have an amplitude given by

$$F_p(s, \alpha) = \lambda^2(\gamma)^{r+p-1} \frac{(2\pi i)^p}{(\alpha+2)^r} (\mathcal{J}''_{N_1}\mathcal{J}''_{N_2})$$
$$\times \left(\prod_{i=1}^{p-1} \mathcal{J}''_{M_i+1}\right)\left(\prod_{1}^{p} \mathcal{J}_{L_i}\right), \qquad (49)$$

where $r = N_1 + N_2 + \sum L_i + \sum M_i + 3p$. The integrals $\mathcal{J}''_{N_i}$, $\mathcal{J}''_{N_2}$ were evaluated in the previous section; $\mathcal{J}''_{M_i+1}$ is the same integral as $\mathcal{J}'_{M_i+1}$ except that the corresponding Q'_{M_i+1} has both x_{M_i+2} and x_1 equal to zero. This reflects the fact that the internal ladders are closed at both ends, while the external ladders are closed at only one end. The recursion relation determining $\mathcal{J}''_{M_i+1}$ is unchanged; there is an integral $\mathcal{J}'_{M_i+1}$ differing from $\mathcal{J}'_{M_i+1}$ only in that $x_1 = 0$. The difference occurs in $\mathcal{J}'_1$ and $\mathcal{J}''_1$. In terms of the functions defined in (12), $\mathcal{J}'_1$ and $\mathcal{J}''_1$ are given by

$$\mathcal{J}'_1 = -sD(s),$$
$$\mathcal{J}''_1 = B(s). \qquad (50)$$

$\mathcal{J}''_1$ is the integral for an internal ladder of zero rungs. The internal ladders sum up into the following expression:

$$\sum_{m=0}^{\infty} \left(\frac{\gamma}{\alpha+2}\right)^m \mathcal{J}''_{m+1} = \hat{P}' \left[1 - \frac{\gamma}{\alpha+2}\hat{a}\right]^{-1} \begin{pmatrix} \mathcal{J}'_1 \\ \mathcal{J}''_1 \end{pmatrix},$$
$$= B(s)(\alpha+2)^2/[\alpha+2-\delta_2(s)](\alpha+2+\gamma/m^2). \qquad (51)$$

We have again used the explicit forms of $B(s)$ and $D(s)$. Notice the surprising result that the internal

ladder contributes two poles, one of which is fixed. This is a basic difference between the $l = -1$ and $l = -2$ behavior.

The sums over N_1, N_2, and L_i can be carried out as in the previous section. If $\Omega(s)$ is defined as $2\pi i$ times the integral in (47), $F_p(s, \alpha)$ becomes

$$F_p(s, \alpha) = \frac{\lambda^2\gamma^{2p-1}}{(\alpha+2)^{2p}}\left[\frac{\alpha+2}{\alpha+2-\delta_2(s)}\right][(\alpha+2)\Omega(s)]^p \times\left[\frac{B(s)(\alpha+2)^2}{(\alpha+2-\delta_2(s))(\alpha+2+\gamma/m^2)}\right]^{p-1}. \quad (52)$$

Upon summing (52) over p and adding in the second term of (37) we find

$$F(s, \alpha) = \lambda^2\frac{[-\frac{1}{2}(s-2m^2)(\alpha+2+\gamma/m^2)+\gamma\Omega(s)]}{[(\alpha+2-\delta_2(s))(\alpha+2+\gamma/m^2)-\gamma^2B(s)\Omega(s)]}. \quad (53)$$

In Eq. (53) the single Regge pole has been completely obliterated; in its place we have an amplitude that has a cut in α and two moving poles on each sheet of the cut. Thus, the insertion of higher-order diagrams into the second Regge trajectory splits it into two trajectories plus a moving cut; the only difference between the two poles is that the residue of one of them vanishes to lowest-order in λ.

A natural question at this point is whether the introduction of other graphs into the ladder diagrams will restore the single pole that would be expected on heuristic grounds. There are two kinds of possible insertions; those which close the ends of the ladder at $\alpha = -2$, as the cut insertion does, and those which do not. All diagrams which factor will contribute in exactly the same way as $\Omega(s)$ and thus will generate two poles at $\alpha = -2$. Among such insertions are ones which, in the notation of Halliday[12], are bounded by "two" lines at both ends. These graphs can be planar, so that the presence of two poles is not an effect of a *tu* spectral function. The far larger class of diagrams which do not allow factorization will be coupled to the ladder diagrams through an integral equation. Their effect will be very complicated, but it seems unlikely that they can combine to restore the single pole. Undoubtedly the nonfactorizing insertions generate a whole new class of singularities in α.

In conclusion then the mixing of poles and cuts near $\alpha = -2$ is, in lowest order, very similar in form to (1); however, the cut generated by the leading Regge trajectory mixes only with the second Regge pole. In addition, an analysis of higher-order iterations of the cut-generating graphs indicate that there are two moving poles which mix with the cut, where only one would be expected. The appearance of the second pole is an essential difference between the behavior of ladder diagrams at $l = -1$ and at $l = -2$.

[12] I. G. Halliday, Nuovo Cimento 30, 177 (1963).

ACKNOWLEDGMENT

I wish to thank Dr. J. C. Polkinghorne for suggesting this problem, for several helpful discussions, and for reading the manuscript.

APPENDIX A

For appropriate values of the external masses and momenta, any Feynman diagram can be expressed in the form

$$f(s, t) = \int_0^\infty \frac{d\alpha: e^{-M}}{\Delta^2}. \quad (A1)$$

The complete set of integration variables is denoted by α_i. In terms of the conventional C and D functions, Δ is just C expressed in terms of the new set of variables and M is D/C.[7] A ladder diagram of $N+1$ rungs, with the external particles off the mass shell, and carrying a total 4-momentum Q, has for M_N the following expression:

$$M_N = m^2A_N + B_Nk_1^2 + C_Nk_2^2 + D_NQ^2 - 2E_NQ\cdot k_1 - 2F_NQ\cdot k_2 - 2G_Nk_1\cdot k_2, \quad (A2)$$

where k_1 and $Q - k_1$ are the incoming momenta and k_2 and $Q - k_2$ are the outgoing. An Euclidean metric with an imaginary fourth component has been used in writing (A2). The functions A, B, C, etc., are given by

$$\begin{aligned} A_N &= \sum_1^N \sigma_i + \sum_{i=1}^{N+1} x_i, \\ B_N &= x_1 - \sum_{i=1}^N \frac{(x_1\cdots x_i)^2}{(\theta_1\cdots\theta_{i-1})^2\theta_i}, \\ C_N &= x_{N+1} - x_{N+1}^2/\theta_N, \\ D_N &= \sum_{i=1}^N z_i - \sum_1^N \frac{\phi_i^2(z)}{\theta_i}, \\ E_N &= \sum_{i=1}^N \frac{\phi_i(z)(x_1\cdots x_i)}{\theta_1\cdots\theta_i}, \\ F_N &= \phi_N(z)x_{N+1}/\theta_N, \\ G_N &= x_1\cdots x_{N+1}/\theta_1\cdots\theta_N. \end{aligned} \quad (A3)$$

The variables in (A3) correspond to the labeling in Fig. 2, and σ_i, θ_i, ϕ_i were defined in (11). On the

mass shell, $k_1^2 = k_2^2 = -m^2$, $Q^2 = 2Q\cdot k_1 = 2Q\cdot k_2 = -s$, and $k_1\cdot k_2 = t/2 - m^2$. We then find that

$$\begin{aligned} \mathfrak{M}_N &= A_N - B_N - C_N + 2G_N, \\ \mathfrak{S}_N &= D_N - E_N - F_N, \end{aligned} \tag{A4}$$

where a few algebraic manipulations are necessary to put $\mathfrak{S}_N$ in the form used in (12). This representation of M for a ladder diagram is generated by successively completing the square and integrating out the N internal momenta of the graph.

The procedure used to derive (A2) and (A3) may be used to evaluate Fig. 3. It is simpler, however, to insert (A2) for each of the three ladders, with the appropriate mass-shell constraints, and consider the diagram to have only five independent internal momenta. The final result of carrying out all the integrations is extremely cumbersome. The coefficient of t can be written relatively concisely in the form

$$g/\Delta = LJ(\Delta'(X + P_2/\Delta')(Y + P_1/\Delta') + \bar{g}). \tag{A5}$$

X, Y, L, J, and Δ' are defined in Sec. 3, and Δ is just the C function for the whole graph. The other factors in (A5) are

$$\begin{aligned} \bar{g} = w_1 \cdots w_{L+1}[&x_{N+2}\beta_4(\beta_3 + \delta_4) \\ &+ x'_{m+2}\beta_1(\alpha_4 + \beta_2) + x'_{m+2}x_{N+2}(D + \gamma) \\ &- E\alpha_4 x'_{m+2} - F\delta_4 x_{N+2} - G\alpha_4\delta_4], \end{aligned} \tag{A6}$$

$$P_2/\Delta' = B\alpha_4 - Ex_{N+2}, \qquad P_1/\Delta' = C\delta_4 - Fx'_{m+2},$$

where $x_{N+2} = \alpha_4 + \beta_1$ and $x'_{M+2} = \delta_4 + \beta_4$. B, C, D, E, F are functions of the internal parameters such that as all the w_i vanish, P_1 and P_2 also vanish. The remaining part of M, referred to as Q in the paper, is given by the following expression:

$$Q = J_N + J_M + s\mathfrak{S}' + M^2\mathfrak{M}'. \tag{A7}$$

J_N and J_M depend only on the external ladders.

$$\begin{aligned} J_N = m^2(A_N - B_N) \\ - s(D_N - E_N) = Q_N(x_{N+1} = 0). \end{aligned}$$

In the limit of interest in this paper ($\beta_1 = \alpha_4 = \beta_4 = \delta_4 = w_i = 0$) the other terms in (A7) can be evaluated with relative ease. The result is that

$$\begin{aligned} \mathfrak{M}' = &\sum_{i=1}^{L}(l_i + n_i) + (\alpha_2 + \beta_2) \\ &+ (\delta_2 + \beta_3) + \gamma + (\alpha_1 + \alpha_3) + (\delta_1 + \delta_3), \end{aligned} \tag{A8}$$

$$\begin{aligned} \mathfrak{S}' = &-\gamma - \left(\alpha_3 - \frac{\phi_{N+1}^2}{\theta_{N+1}} - \frac{x_1 \cdots x_{N+1}}{\theta_1 \cdots \theta_{N+1}}\phi_{N+1}\right) \\ &+ \frac{\gamma^2}{H} - \left(\delta_3 - \frac{\phi_{M+1}'^2}{\theta'_{M+1}} - \frac{x'_1 \cdots x'_{M+1}}{\theta'_1 \cdots \theta'_{M+1}}\phi'_{M+1}\right), \end{aligned} \tag{A9}$$

where

$$\theta_{N+1} = \alpha_1 + \alpha_3 + y_{N+1} \quad \phi_{N+1} = \alpha_3 + x_{N+1}\phi_N/\theta_N,$$

$$\theta'_{M+1} = \delta_1 + \delta_3 + x'_{M+1} \quad \phi'_{M+1} = \delta_3 + x'_{M+1}\phi'_M/\theta'_M,$$

and

$$H = \gamma + \frac{\beta_2\alpha_2}{\alpha_2 + \beta_2} + \frac{\beta_3\delta_2}{\delta_2 + \beta_3} + \sum_{i=1}^{L}\frac{l_i n_i}{l_i + n_i}.$$

We have dropped all terms in $\mathfrak{M}'$ and $\mathfrak{S}'$ which are quadratic in x_i or x'_i. Hence, if α_i, β_i, δ_i are relabeled as indicated below (45), we can combine the various terms in (A7) to write

$$\begin{aligned} Q = &\, m^2(A_{N+1} - \beta_{N+1}) - s(D_{N+1} - E_{N+1}) \\ &+ m^2(A'_{M+1} - \beta'_{M+1}) - s(D'_{M+1} - E'_{M+1}) \\ &+ m^2\left(\gamma + \sum_{i=0}^{L+1}(l_i + n_i)\right) - s(\gamma - \gamma^2/H) \\ = &\, Q_{N+1}(x_{N+1} = 0) + Q_{M+1}(x'_{M+2} = 0) + Q''_L. \end{aligned} \tag{A10}$$

Finally we look at Δ''_L which is defined to be $\bar{g}'\Delta'\alpha_2\delta_2$ with all the appropriate variables vanishing. From (A6), when $E = F = G = 0$, $\alpha_4 = (\beta_1\beta_2)/\alpha_2$, and $\delta_4 = (\beta_3\beta_4)/\delta_2$, Δ''_L is given by

$$\begin{aligned} \Delta''_L = \bar{g}'\Delta'\alpha_2\delta_2 &= \prod_{i=1}^{L}(l_i + n_i)(\alpha_2 + \beta_2)(\delta_2 + \beta_3)H \\ &= H\prod_{i=0}^{L+1}(l_i + n_i). \end{aligned} \tag{A11}$$

APPENDIX B

In the course of evaluating the amplitude corresponding to Fig. 3, we needed the integral

$$I = \int_{x_1}^{x_2}\int_{y_1}^{y_2}[axy + b]^\alpha, \tag{B1}$$

where x_1, $y_1 < 0$ and x_2, $y_2 > 0$. In particular, we were interested in the dominant terms as α tended to -2 and b to zero. Consider first the following integral:

$$\begin{aligned} I' &= \int_0^{x_2}\int_0^{y_2}[axy + b]^\alpha \\ &= \lim_{\epsilon\to 0}\frac{1}{a(\alpha + 1)} \\ &\quad \times \int_\epsilon^{ax_2y_2} dz\left[(t + b)^\alpha \sum_{n=0}^{\infty}\left(\frac{b}{z + b}\right)^n - b^{\alpha+1}\right]. \end{aligned}$$

$$I' = \frac{1}{a(\alpha+1)}\left\{\frac{(ax_2y_2+b)^{\alpha+2}}{b}\ln\left(\frac{ax_2y_2+b}{ax_2y_2}\right)\right.$$
$$+ (ax_2y_2+b)^{\alpha+1}(\alpha+2)F\left(\frac{b}{ax_2y_2+b}\right)$$
$$+ b^{\alpha+1}\ln\left(\frac{ax_2y_2}{b}\right)$$
$$\left. - \tfrac{1}{6}(\alpha+2)b^{\alpha+1}\pi^2 + O[(\alpha+2)^2]\right\}, \qquad \text{(B2)}$$

where

$$F(x) = \sum_{n=0}^{\infty}\frac{x^n}{(n+1)^2}.$$

Only the terms proportional to $b^{\alpha+1}$ are singular near $\alpha = -2$.

$$I' = \frac{b^{\alpha+1}}{a(\alpha+1)}\left[\ln\left(\frac{ax_2y_2}{b}\right) - (\alpha+2)\frac{\pi^2}{b}\right]. \qquad \text{(B3)}$$

Equation (B3) is valid for x_2 and y_2 both positive, or both negative, since the argument of the logarithm is then positive. To obtain the value of the integral for $x_2 > 0$ and $y_2 < 0$, we analytically continue (B3) in y_2 to obtain

$$I = [b^{\alpha+1}/(\alpha+1)a][\ln(ax_2|y_2|/b) - i\pi - \tfrac{1}{6}(\alpha+2)\pi^2]. \qquad \text{(B4)}$$

where we have arbitrarily set $\ln(-1) = -i\pi$. The integral in (B1) can now be evaluated:

$$I = \left[\int_0^{x_2}\int_0^{y_2} + \int_0^{x_1}\int_0^{y_1} - \int_0^{x_2}\int_0^{y_1} - \int_0^{x_1}\int_0^{y_2}\right]$$
$$\times [axy+b]^{\alpha} = [2\pi i/(\alpha+1)^2]b^{\alpha+1}. \qquad \text{(B5)}$$

This is the result used in (42). Clearly the same kind of manipulation leads to

$$\prod_{i=1}^{N}\iint dx_i\,dy_i[a_ix_iy_i + b_i]^{\alpha} = \frac{(2\pi i)^N}{(\alpha+1)^N}\prod_{i=1}^{N}\frac{b_i^{\alpha+1}}{a_i}. \qquad \text{(B6)}$$

A Covariant Theory of Relativistic Brownian Motion I. Local Equilibrium

RÉMI HAKIM*
Laboratoire de Physique Théorique et Hautes Energies, Orsay, France
(Received 9 June 1964; final manuscript received 24 December 1964)

This paper is devoted to the study of some relativistic methods in statistical mechanics and their applications to the construction of a covariant theory of Brownian movement. In this work we deal only with times which are long compared with the relaxation time β^{-1}. As a consequence of this covariant theory we derive relativistic Wiener integrals.

INTRODUCTION

THE probable appearance in the future of plasmas with very high temperatures necessary to obtain thermonuclear fusion has increased the need for a systematic study of relativistic plasmas (perhaps already existing in white dwarfs). Though actual plasmas now encountered in physics are not very hot, it is necessary to treat them with a covariant method if one deals with radiation phenomena for which relativistic effects are very important[1] (more than quantal ones) at least in a large domain of frequencies. Accordingly, it is necessary to have a covariant theory; because of the qualitative change introduced by the special theory of relativity, we cannot take into account relativistic effects by merely adding corrective terms in classical expressions. Furthermore there is no systematic study of relativistic statistical mechanics and we hope to derive methods appropriate to such a study. More particularly we hope to give a general formalism for relativistic irreversible processes based on methods used in the theory of Brownian motion (Onsager relations).[2]

The first section of this paper is devoted to some relativistic concepts: description of the evolution of the observable quantities, probabilities and relativistic mean values. Section 2 deals with the basis of the method used. Starting from a "maximum entropy principle," we derive the transition density by assuming that the first two moments are given. This entropy is discussed here. We find at least three kinds of possible entropies which lead to the same results. In Sec. 3, the transition density is derived and next specified by means of a covariant diffusion equation (derived from the assumption of the markovian character of the stochastic process associated to the covariant Brownian motion). In Sec. 4, we define the probabilistic axioms of a covariant Brownian motion and as a consequence, we derive the covariant Wiener integral.

Finally, as direct applications of the theory hereafter presented, we think that we are able to treat several problems of radiation and this will be done in another paper.

The classical study[3] of Brownian motion involved (a) the derivation of the conditional density of probability of a particle within a random environment, (b) the derivation of the diffusion equation with or without an external field of force, (c) the study of the approach to local equilibrium by means of the Langevin and the Fokker–Planck equations, (d) the derivation of the Wiener integral (defined by a measure on the set of continuous trajectories), and last, (e) applications of the methods used to analogous problems. Our program is to derive a covariant theory of the above points. In this paper we treat only points (a), (b), and (d), the other points being treated in our next paper. Of course we do not claim that our theory is the only one: it is merely a possibility.

Throughout this paper we treat only the case of local equilibrium, i.e., the case where the stochastic process associated with the covariant Brownian motion is stationary. The general case—the approach to local equilibrium—is treated in the second part of this paper.

NOTATION

$\otimes$: tensorial product
$\theta(x)$: heaviside step function
$\mathfrak{M}^n$: n-dimensional pseudo-Euclidian space with the metric: $+ - - -$.

* Postal address: Laboratoire de Physique Théorique et Hautes Energies, Bâtiment 211, Faculté des Sciences, Orsay (Seine-et-Oise), France.

[1] J. L. Delcroix, in *La théorie des gaz neutres et ionisés*, des Houches Summer School (Dunod & Cie, Paris, 1959).

[2] P. Mazur, in *Termodinamica dei processi irreversibili*, Varenna Summer School, 1959 (N. Zanichelli, Bologna, 1960).

[3] S. Chandrasekhar, Rev. Mod. Phys. **15**, 1 (1943).

$\Gamma^+(x_0)$: future null cone, the origin of which is at x_0^μ [$\Gamma^+ \equiv \Gamma^+(0)$]
$\prec$: order
$\wedge$: operation of an exterior algebra
R^+: positive real numbers
1_A: characteristic function of the set A.
δ_Σ: superficial measure $+1$ on the surface Σ
$||\alpha_{\mu\nu}||$: matrix, the elements of which are $\alpha_{\mu\nu}$
$g_{\mu\nu}$: metric tensor: $g_{00} = 1$, $g_{ii} = -1$; $g_{\mu\nu} = 0$ for $\mu \neq \nu$
$\delta_{(4)}(x_\mu)$: $\delta(x_0) \otimes \delta(x_1) \otimes \delta(x_2) \otimes \delta(x_3)$
A^B: (where A and B are sets) represents the sets of all mappings of B into A.
c: speed of light taken as unity throughout this paper.
$\mathfrak{L}(E, E)$: designates the set of all linear applications of the vectorial space E onto itself.

I. SOME RELATIVISTIC CONCEPTS

Before we treat Brownian motion in a covariant manner, we need some relativistic methods hereafter presented. Some ideas given in this section are explained in another work.[4] In order to deal with problems of statistical mechanics we ask for a covariant notion of evolution, covariant probabilistic methods and covariant mean values.

Location in Minkowski Space–Time

In all that follows, we shall consider only the flat space-time of *special relativity*: $\mathfrak{M}^4$. Let us first introduce the notion of σ-partition.

Definition 1. We shall call σ-partition $\{S\}$, a family of spacelike three-dimensional manifolds S constituting a partition of $\mathfrak{M}^4$ and satisfying the following axioms:

S_1: Each manifold S belonging to the σ-partition is defined by a function $S(x_0x_1x_2x_3) \equiv_{\text{def}} S(x_\nu)$ and an invariant parameter s by the equation: $S(x_\nu) = s$. The parameter s will be called the pseudotime relative to the σ-partition $\{S\}$. In what follows we shall call S_s the surface corresponding to the value s of the pseudotime.

S_2: if $s_1 \neq s_2$ then $S_{s_1} \cap S_{s_2} = \phi$,

S_3: $\mathfrak{M}^4 = \bigcup_{s \in R} S_s$,

or the weakened form of S_3,

S_3': $\Gamma^+ = \bigcup_{s \in R^+} S_s$ where Γ^+ is the future null cone,

S_4: S_s is indefinitely differentiable,

S_5: S_s has no singularities: i.e. $\forall x_\nu \in S_s$ one has $\partial_\mu S(x_\nu) \neq 0$,

S_6: S_s is orientable.

Let us now explain the reasons for which we introduce this definition. In relativity there exists no privileged notion of simultaneity so that the physical space at a given "instant" is no longer a 3-plane $t = \text{const}$, but rather an *arbitrary spacelike* 3-surface: $S \subset \mathfrak{M}^4$. By taking an arbitrary spacelike 3-surface we give a sense to physical space but the word "instant" is not clear because of the arbitrariness of the surface. Let us try to give a precise meaning to the word "instant". To do this we should give a fully ordered set structure to the set of all spacelike 3-manifolds of Minkowski space–time. This is hardly possible and accordingly we introduce a partial order, and only on a subset (the subset of all spacelike 3-surfaces satisfying the regularity axioms S_4, S_5, S_6): because of S_1, $\{S\}$ is ordered by

$$S_{s_1} \prec S_{s_2} \quad \text{if and only if} \quad s_1 < s_2.$$

Then, the value taken by a pseudo-instant s determines a unique cut of space–time, on which one can evaluate global (i.e. nonlocal) quantities, for example relative to a given fluid. For instance, we have

$$\tilde{\varphi}_S \equiv \tilde{\varphi}_\alpha^{\mu\nu\cdots}(s) = \int_{S(x_\rho)=s} \varphi_\alpha^{\mu\nu\cdots}(x_\rho)\, dS$$

with $dS = n_\lambda dS^\lambda$, where n_λ is the normal unit to S and dS_λ is the symmetric differential form with vectorial values:

$$dS_\lambda = \frac{1}{3!}\, \epsilon_{\lambda\mu\nu\rho}\, dx^\mu \wedge dx^\nu \wedge dx^\rho$$

and

$$\epsilon_{\lambda\mu\nu\rho} = \begin{cases} +1 & \text{if } (\lambda\mu\nu\rho) \text{ is an even permutation of } (0\ 1\ 2\ 3), \\ -1 & \text{if } (\lambda\mu\nu\rho) \text{ is an odd permutation of } (0\ 1\ 2\ 3), \\ 0 & \text{otherwise.} \end{cases}$$

We shall point out that when we take another σ-partition $\{\Sigma\}$, there is *a priori* no reason why $\tilde{\varphi}_\Sigma = \tilde{\varphi}_S$, when $s = \sigma$, if we admit that we have fixed an origin and a suitable scale for s and σ.

As a consequence of these definitions, we see that the temporal evolution of a global quantity is perfectly defined in a covariant manner when a σ-partition is given. But this method, which permits the determination of the evolution of a physical system (a fluid or a statistical collection of physical objects . . .), depends greatly on the σ-partition used.

[4] R. Hakim (to be published).

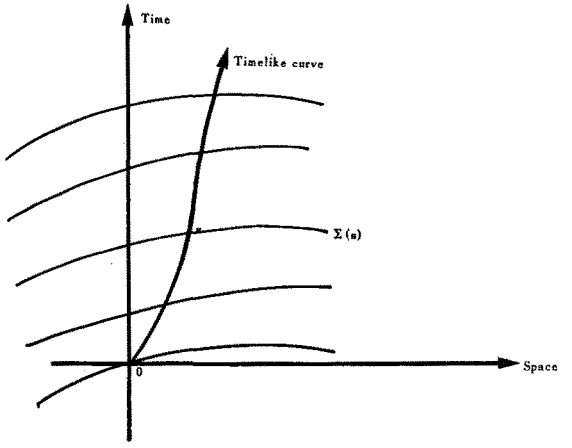


FIG. 1. A way to number a σ-partition. The number associated with the surface Σ, is given by the proper time of the above timelike line at the intersecting point (the origin is chosen as indicated on the figures).

However, since there is no privileged notion of simultaneity (all the σ-partitions are equivalent from the point of view of the description of physical space) there must not exist a privileged notion of temporal evolution (as a temporal counterpart of the aforementioned equivalence of all σ-partitions); i.e., all pseudotimes are equivalent (we assume implicitly that we have chosen a suitable scale for s and σ, i.e., we have numbered in the same way the two σ-partitions $\{S\}$ and $\{\Sigma\}$): $\tilde{\varphi}_{\Sigma\in\{\Sigma\}}(\sigma)$ has the same form as $\tilde{\varphi}_{S\in\{S\}}(s)$ and the difference between the two σ-partitions used $\{\Sigma\}$ and $\{S\}$ is reflected *only* in the initial conditions. This means that when the global quantity $\tilde{\varphi}$ satisfies a given equation, the precise form of the solutions of this equation will depend on the initial conditions through the element number zero of an arbitrary σ-partition. For instance, if the global quantity $\tilde{\varphi}(s)$ satisfies the differential equation

$$(d/ds)\tilde{\varphi}(s) + \beta\tilde{\varphi}(s) = 0,$$

we will have $\tilde{\varphi}(s) = \tilde{\varphi}(0)\exp[-\beta s]$; but as we have already said, there is *a priori* no reason why

$$\tilde{\varphi}_1(0) \equiv \int_{\Sigma_0} \varphi\, d\Sigma = \tilde{\varphi}_2(0) \equiv \int_{S_0} \varphi\, dS$$

$$\text{with} \quad \Sigma_0 \in \{\Sigma\}, \quad S_0 \in \{S\},$$

and consequently the initial conditions reflect the difference between the various possible σ-partitions.

If we define some method for constructing a pseudotime s relative to a given σ-partition, all the pseudotimes constructed by the same method are equivalent for the description of the evolution of global quantities. For example, a possible way to number a given σ-partion can be determined by using a timelike line with a fixed arbitrary origin on it. Then we attribute to a surface of the σ-partition a number defined by the proper time of the intersecting point (see Fig. 1). We can thereby introduce the concept of statistical time (which appears as an independent parameter) defined as a kind of class of equivalence of the pseudotimes corresponding to a given way of numeration. The word "statistical" is used to designate this extra parameter because of its use in the description of statistical collections: according to its definition, it is not a mean value.

In practice, we take a well-determined σ-partition, chosen to suit the problem under study and therefore a given notion of simultaneity. Generally it may be of interest to define a σ-partition by a particular property of the physical system. For example, in a relativistic fluid, a useful σ-partition is determined by the surfaces orthogonal to the stream lines. Unfortunately, these surfaces do not always exist (e.g., when one considers the curly motion of a relativistic fluid).

For many people a Lorentz observer is constituted by a σ-partition, for others a Lorentz observer is a congruence of timelike curves (Cattaneo[5]). For us, a Lorentz observer implies (a) a system of coordinates; (b) a notion of simultaneity, and a change of Lorentz observer implies two things: (a) a change of coordinates; (b) a change of the notion of simultaneity.

We must point out that to every Lorentz observer 0, is attached a σ-partition; and when we change the Lorentz observer (i.e., when we have $0' = L0$), we have $\{S\}' = L\{S\} = \{LS\}$ (since the Lorentz group is a continuous group, it transforms a σ-partition into another σ-partition), where L is a Lorentz transformation.

If the system has a privileged point (which is the case when we deal with a diffusion phenomenon where the origin is a privileged point; this is also the case when we must use a transition density) there exists a very useful σ-partition, geometrically and analytically invariant, which consists of the family of hyperboloids $\mathcal{H}(\tau)$ relative to this point taken as the origin:

$$\{\mathcal{H}(\tau)\}: \begin{cases} x^\mu x_\mu = \tau^2 \\ \tau^2 > 0. \end{cases}$$

[5] C. Cattaneo, Seminar given at Institut Henri Poincaré, Paris, 1961.

The orthogonal trajectories of such a σ-partition are the straight lines beginning at the origin. $\{\mathcal{K}(\tau)\}$ is useful[6] in our problem where the origin plays a special role (the Brownian particles being initially concentrated at the origin). The principlal advantage of such a σ-partition is the following: to all density-like quantities (i.e., local quantities) having a determined variance correspond global (integral) quantities having the same variance. The disadvantages of $\{\mathcal{K}(\tau)\}$ are

(a) $\mathcal{K}(\tau)$ does not reduce to $\{0\}$ when $\tau \to 0$.

(b) $\mathcal{K}(\tau)$ is not compact but only locally compact.

(c) The measure of $\mathcal{K}(\tau)$ is infinite.

(d) $\mathcal{K}(\tau)$ is not invariant under the translations of $\mathfrak{M}^4$.

Local and Global Quantities

A relativistic fluid is often defined as a congruence of timelike curves directed towards the positive x^0. This definition covers two types of elements characterizing the fluid: the first is rather kinematic (the field of four speeds), while the second is rather dynamic (the invariant world density). These quantities are *local quantities* in contrast to the observable macroscopic quantities which are, especially in statistical mechanics, *global quantities*. Global quantities are generally integral quantities, and consequently are relative either to a domain of $\mathfrak{M}^4$, or (in the cases with which we deal) to a spacelike cut of Minkowski space. Examples of global quantities are: the electric charge of a fluid at a given pseudotime (i.e., on a spacelike 3-surface $S \in \{S\}$), the number of particles at the same pseudotime, the center of gravity $X^\mu(S)$, etc. $\cdots$. For an arbitrary fluid scheme,[7] a global quantity has not, in general, a well-defined relativistic variance[8,9]; in order to explain this point, let us take the case of a global scalar. One has, in a given fluid,

$$\bar{\varphi}(S) = \int_S \varphi(x_\nu) j^\mu \, dS_\mu \quad \text{with} \quad S \in \{S\}, \qquad (1)$$

where j^μ is the four-current density of the fluid. If we perform a homogeneous Lorentz transformation L on our physical system, i.e., when we change the Lorentz observer, then (where we have assumed that we have given a correct definition of LS) $S \to LS$ and generally

$$\bar{\varphi}(LS) \neq \bar{\varphi}(S)$$

since $\bar{\varphi}(LS)$ and $\bar{\varphi}(S)$ are integral quantities relative to different cuts of space–time. Thus the left-hand side of (1) is generally not a scalar, but appears to be a numerical function defined on the Lorentz group. However, there exists a particular kind of local quantities $\varphi(x_\nu)$ for which we have $\bar{\varphi}(LS) = \bar{\varphi}(S)$; they are the quantities j^μ-conservative, i.e., such that

$$\partial\mu(\varphi j^\mu) = 0. \qquad (2)$$

These quantities obviously form a vectorial space. Expression (2) is equivalent to the following: $(\delta/\delta S)\bar{\varphi}(S) = 0$, where $\delta/\delta S$ is a functional derivative.

There is another case where $\bar{\varphi}(LS) = \varphi(S)$; this is the case of the σ-partition defined by the family of hyperboloids $\mathcal{K}(\tau)$, since $L\mathcal{K}(\tau) = \mathcal{K}(\tau)$.

From a general point of view, global quantities appear as functions defined on the Lorentz group, the values of which are tensorial. In fact, if one contracts a local quantity with a constant dual tensor, one obtains a local scalar from which arises a global scalar after an integration over an $S \in \{S\}$. Then from this global scalar we obtain a global tensor, i.e., a function on the Lorentz group taking its values in a certain tensorial space.

From a probabilistic point of view, and consequently from the point of view of relativistic statistical mechanics, this circumstance is unfortunate, because all the observable macroscopic quantities, which are mean values of local quantities, have a well defined variance. Therefore we must find, whenever possible, a way of obtaining mean values having the same variance as the associated density.

Probability and Mean Values

It is easy to show that a probabilistic fluid scheme (i.e., a so-called fluid of probability defined by a conserved four-current) is not well adapted to a covariant description of various probabilistic notions: entropy, stationarity, markovian processes etc. ... The simplest way to describe the probabilistic notions in Minkowski space–time is to fix a σ-partition $\{\Sigma\}$, once and for all, to call ρ_Σ the probability density on a given element of the σ-partition used, and finally to transcribe into this scheme all the properties we need for our theory. In doing this, mean values must be defined in such a way that they have the same variance as the associated densities and reduce, in the classical limit, to the usual mean values. We now try to give a solution to this problem.

---

[6] R. Hakim, Institut du Radium Report IRPO-T 64-01.

[7] A. Lichnerowicz, *Les théories relativistes de l'électromagnétisme et de la gravitation* (Masson et Cie, Paris, 1955).

[8] P. G. Bergmann, in *Handbuch der Physik* edited by S. Flügge (Springer-Verlag, Berlin, 1962), Vol. 4, p. 135.

[9] F. Halbwachs, *Théorie relativiste des fluides à spin*, (Gauthier-Villars, Paris, 1960).

First, let E be the tensorial algebra generated by $\mathfrak{M}^4$. Let us call $E(\mathfrak{M}^4)$ a set of applications of $\mathfrak{M}^4$ into E, these applications being of physical interest: i.e., $E(\mathfrak{M}^4)$ is the set of densities on $\mathfrak{M}^4$ having a certain tensorial variance. Let us call $E(L^{++})$ the set of all continuous applications of L^{++} (homogeneous orthochronous proper Lorentz group) onto E. A star designates the algebraic dual of a given vectorial space.

Now let I_L be the application of $E(\mathfrak{M}^4)$ onto $E(L^{++})$ defined as follows:

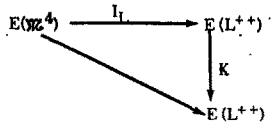
$$\forall\varphi \in E(\mathfrak{M}^4) \quad \text{and} \quad \forall\chi \in E^*;$$

let us form the local scalar quantity ($\forall\chi \in E^*$ and such that $[\chi, \varphi] \in R$ or C); let us form the local scalar quantity $\alpha = [\chi, \varphi]$, where the brackets denote a "scalar product," which is constituted by the contraction of χ and φ. Next, let us form the global scalar:

$$[\chi, \tilde{\varphi}(L)] = \tilde{\alpha} = \int_{LS} \alpha\rho \, dS, \quad \text{with} \quad \tilde{\alpha} \in E(L^{++});$$

then I_L is the operation defined by $\tilde{\varphi}(L) = I_L\varphi$.

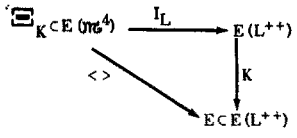
We can now search for the desired averaged operation. If we know how to write a linear operation K such that one has the following (commutative) diagram:



and if we restrict ourselves to the elements of $E(\mathfrak{M}^4)$ such that

$$K \circ I_L(\varphi) \in E \subset E(L^{++}),$$

then our goal is reached for *these elements*. Because of the hereafter-assumed linearity of K, these elements form a vectorial subspace of $E(\mathfrak{M}^4)$, namely Ξ_K, so that one has the following (commutative) diagram:



where $\langle\ \rangle$ is the average operation, one wants to find

$$\langle\ \rangle = K \circ I_L.$$

Obviously this is not a general method since we define an average only for a certain subspace of $E(\mathfrak{M}^4)$. However this is more general than, for instance, integrability (Ref. 8) conditions. A generalization of the above method can be obtained if we are able to find a family of K_i, $i \in I$ being given, such that

$$\forall i, j \in I \quad \text{with} \quad i \neq j; \quad \text{we have } \Xi_{K_i} \cap \Xi_{K_j} = E.$$

Then, an average operation may be defined on

$$\Xi(I) = \bigcup_{i\in I} \Xi_{K_i}$$

by taking

$$K = \sum_{i\in I} K_i P(\Xi_{K_i}),$$

where $P(\Xi_{K_i})$ is the projection on Ξ_{K_i}. This method will be developed in a later paper. Let us now return to our operation K: one can see that K possesses all the properties of a Reynolds's operator[10,11]:

R_1: K is a linear operation: $K \in \mathcal{L}[E(L^{++}), E(L^{++})]$.
R_2: $\forall f \in E(L^{++})$: $f \geq 0$, then $Kf \geq 0$.
R_3: K must be continuous [in a sense, to be made more precise, according to the topological structure of $E(L^{++})$].
R_4: K const = const.
R_5: $\forall f, g \in E(L^{++})$, one has: $K(f\cdot Kg) = Kf\cdot Kg$.

Accordingly, one can verify that such an operation K may be given by

$$Kf = \lim_{B(L)\to L^{++}} \left\{ \frac{1}{\text{mes}\, B(L)} \int_{B(L)\subset L^{++}} f(L')\, d\mu(L') \right\}, \tag{3}$$

where L, L' belong to L^{++}, $\mu(L')$ is the Haar measure on the homogeneous, orthochronous proper Lorentz group L^{++} and $B(L)$ is a compact subset of L^{++} containing the element L in a consistent manner. Of course, this is not the only possible operation. Nevertheless it appears to be natural since it constitutes also a kind of mean value. In formula (3),

$$\lim_{B(L)\to L^{++}} \{\cdots\}$$

must be understood as follows: the characteristic function of $B(L)$ converges simply towards 1 on L^{++} ($1_B \to 1_{L^{++}}$).

As we have already said, the operation K in (3) only gives us an average function on the Lorentz group, instead of a constant. However for an important class of functions on the Lorentz group, Kf = const and we shall restrict ourselves to this class. In the other cases, we shall say that f has no K average.

Now, let us give an explicit form of Kf, dealing

[10] K. De Feriet in *Théorie des fonctions aléatoires* (Masson et Cie., Paris, 1953).
[11] B. Brainerd, J. Math. Anal. & Appl. 5, 347 (1962).

with the simpler case of the one-parameter Lorentz group operating in the space $\mathfrak{M}^2$. In this case, we have

$$L(\theta) = \begin{vmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{vmatrix} \quad \text{and} \quad \begin{matrix} d\mu(\theta') = d\theta' \\ B(\theta) = [\theta - \tau,\ \theta + \tau] \end{matrix}$$

with $\tanh\theta = -v$. Consequently we have

$$Kf = \lim_{\tau\to\infty} \frac{1}{2\tau} \int_{\theta-\tau}^{\theta+\tau} f(\theta')\, d\theta'; \tag{4}$$

One can easily verify that a class of functions having a constant average consists of the locally integrable functions bounded above or below by a constant, and such that the limit (4) exists. In the case of the space $\mathfrak{M}^4$, where the Lorentz group is a six-parameter group, a class of functions with a constant average is not built in such a simple way.

Let us now prove that our covariant mean value $\langle\ \rangle = K \circ I_L$ reduces to the classical one at the nonrelativistic limit. It is sufficient to show this consistency property for the scalars. At the classical limit

$$I_L\varphi \to \int_{t=\text{const}} \varphi\rho\, d_3x$$

since all notions of simultaneity (i.e., all spacelike surfaces) reduce to the one of instantaneity (i.e., $t = \text{const}$). Consequently $\bar{\varphi}$ is no longer a function on the Galilei group; $\bar{\varphi}$ is a constant. The operation K reduces to a similar operation on the functions defined on the Galilei group, and satisfies again the given axioms $R_1 \cdots R_5$. Then

$$K\bar{\varphi} \to \bar{\varphi}. \qquad \text{Q.E.D.}$$

II. BASIC METHOD

Outline of the Method

Though a relativistic theory of Brownian motion can be built from several hypotheses, classical methods are unfortunately very difficult to set up in a covariant way, so that we shall start with an *analogy*, shown hereafter, which will help us to imagine a method more appropriate to the covariant case.

First, let us consider a nonrelativistic uniform gas of particles of mass m, without interactions between them, at thermodynamic equilibrium and consider (for a Brownian particle) only the case of interest in this paper: $t \gg \beta^{-1}$ (β is the friction coefficient). The conditional momentum density of a Brownian particle within the gas, is

$$\varphi(\mathbf{p}_0, 0 \mid \mathbf{p}, t) = (4\pi mkT)^{-\frac{3}{2}} \exp[-\mathbf{p}^2/2mkT], \tag{5}$$

while the conditional position density is

$$\rho(\mathbf{x}_0 \mid \mathbf{x}, t) = (4\pi Dt)^{-\frac{3}{2}} \exp[-\mathbf{x}^2/2Dt]; \tag{6}$$

if we compare expressions (5) and (6) one can see that (6) follows from (5) by the use of the substitution

$$\mathbf{p} \to (kT/D)\mathbf{x} \tag{7a}$$

to which

$$m \to (kT/D)t \tag{7b}$$

must be added.

The substitution (7a) can be justified by the use of the Langevin equation (with $t \gg \beta^{-1}$).

The analogy between expressions (5) and (6) is not accidental. It is due to the fact that the phenomena of molecular impacts are the basis leading to expressions (5) and (6). Indeed, (5) and (6) are the limit laws for the addition of a great number of independant random variables [(5) is the limit law for the addition of momenta while (6) is the limit law for the addition of $\mathbf{x}$].

As a consequence (5) and (6) may be found by the same statistical methods.

If we now use the same statistical methods as those used in deriving the Maxwell–Boltzmann law, in order to look for the conditional spatial distribution of the Brownian particles we should proceed as follows.

First we have a kind of equilibrium, to be specified later, which gives the condition

$$\delta S = 0,$$

where S is the entropy (also to be specified later). Next, we write that the first two moments of the distribution are given and therefore independant of the distribution to be found:

$$\delta\langle\mathbf{x}\rangle = 0,$$

$$\delta\langle\mathbf{x} \otimes \mathbf{x}\rangle = 0$$

by analogy with

$$\delta\langle\mathbf{p}\rangle = 0,$$

$$\delta\langle\mathbf{p} \otimes \mathbf{p}/2m\rangle = 0.$$

It is clear that the main problem is to give a meaning to the condition $\delta S = 0$ and in particular we must answer the question: what is entropy?

Entropy

Because of the aforementioned analogy, we start with an entropy derived from the most probable state of our system. Let us consider a system of

N independent Brownian particles and study their repartition when we place them in a certain number of cells. To do this, let us prove the following proposition:

Proposition 1. Let $\rho(x_\nu, s)$ be the transition density from the origin[12] ($x\nu = 0$, $s = 0$) to points of a surface $\Sigma(s)$, element of a given σ-partition. Then we have to maximize an expression of the form

$$S = -X \int_{\Sigma(s)} \rho(x_\nu, s) \log \rho(x_\nu, s)\, d\Sigma(s), \quad X > 0. \tag{8}$$

Proof. To obtain our "cells," let us take a partition of $\Sigma(s)$. Let Λ be this partition. It satisfies

$$\Sigma(s) = \bigcup_{i_\lambda \in I_\lambda} \sigma_{i_\lambda}$$

$$\text{with} \quad \sigma_{i_\lambda} \in \Lambda \quad \text{and} \quad \sigma_{i_\lambda} \cap \sigma_{j_\lambda} = \varnothing, \qquad i \neq j,$$

$$0 < \sup_{i_\lambda \in I_\lambda} [d(\sigma_{i_\lambda})] \leq \lambda,$$

where $d(\sigma_{i_\lambda})$ is the diameter of σ_{i_λ}.

To prevent the possibility of a $\Sigma(s)$, the measure of which is infinite, let us introduce a family of measures on $\Sigma(s)$ having the properties indicated hereafter. Let $\mu_{\alpha,s}$ be these measures ($\alpha \in R^+$). They are such that: $L\mu_{\alpha,s} = \mu_{\alpha,s}$ with $L \in L^{++}$ (Where L^{++} is the orthochronous proper homogeneous Lorentz group); for any positive s and for any finite α, $\mu_{\alpha,s}[\Sigma(s)] < \infty$; for any $\varphi(x_\nu) > 0$ and $\mu_{\alpha,s}$-integrable: $\int \varphi(x_\nu)\, d\mu_{\alpha,s} > 0$; For any s, $\mu_{\alpha,s}$ converges uniformly towards $1_{\Sigma(s)}$, the characteristic function of $\Sigma(s)$.

For fixed s, these measures are finite, positive, Lorentz invariant (invariant under a change of coordinates but not in the mathematical sense (Ref. 13): only their *forms* are invariant) and converge towards 1, when $\alpha \to \infty$. The existence of such a family is obvious {e.g., let us take

$$\mu_{\alpha,s}(B) = \int_{B \subset \Gamma^+} \exp\left[-\xi^\mu x_\mu \alpha^{-1}\right] d\Sigma(s)$$

with ξ^μ timelike}. Since we fix s [i.e., we take a definite "physical space" $\Sigma(s)$] we can omit the index s of $\mu_{\alpha,s}$ and write $\mu_{\alpha,s} \equiv \mu_\alpha$.

Now, if we throw a Brownian particle on the surface $\Sigma(s)$, the probability that this particle will fall in the i_λth cell is

$$p^\alpha_{i_\lambda} = \frac{\mu_\alpha(\sigma_{i_\lambda})}{\mu_\alpha(\Sigma(s))} \equiv \frac{\mu^{i_\lambda}_\alpha}{\mu_\alpha(\Sigma(s))},$$

and in general the probability of having n_1 particles in the first cell, n_2 in the second one, $\cdots$ n_{i_λ} in the i_λth cell, is given by the well-known formula[14]

$$P_\alpha(\cdots n_{i_\lambda} \cdots) = N! \prod_{i_\lambda \in I_\lambda} \frac{(p^\alpha_{i_\lambda})^{n_{i_\lambda}}}{n_{i_\lambda}!},$$

$$N = \sum_{i_\lambda \in I_\lambda} n_{i_\lambda}.$$

Introducing now a "Boltzmann entropy",

$$S_{\alpha\lambda} = X \log P_\alpha(\cdots n_{i_\lambda} \cdots), \qquad X > 0,$$

and taking into account Stirling's formula and $n_{i_\lambda} \gg 1$, we find (after dropping irrelevant constants which play no role in the maximization)

$$S_{\alpha,\lambda} = -X \sum_{i_\lambda \in I_\lambda} n_{i_\lambda} \log \frac{n_{i_\lambda}}{p^\alpha_{i_\lambda}}.$$

Then, assuming that the measures μ_α have a density $m_\alpha(x_\nu)$ (which is possible as indicated in the preceding example), and writing the same symbol for the Lebesgue measure of the set σ_{i_λ} and for the set σ_{i_λ} we have

$$n_{i_\lambda} = \rho(x^\mu_{i_\lambda}, s)\sigma_{i_\lambda} \quad \text{with} \quad x^\mu_{i_\lambda} \in \sigma_{i_\lambda}$$
$$p^\alpha_{i_\lambda} = m_\alpha(x^\mu_{i_\lambda})\sigma_{i_\lambda}$$

and

$$S_{\alpha,\lambda} = -X \sum_{i_\lambda \in I_\lambda} \rho(x^\mu_{i_\lambda}, s) \log \frac{\rho(x^\mu_{i_\lambda}, s)}{m_\alpha(x^\mu_{i_\lambda})} \sigma_{i_\lambda}.$$

Assuming now the integrability of the function $\rho \log (\rho/m_\alpha)$ and "taking the limit" $\lambda \to 0$, we obtain

$$S_\alpha = \lim_{\lambda \to 0} S_{\alpha,\lambda} = -X \int_{\Sigma(s)} \rho(x_\nu, s) \log \frac{\rho(x_\nu, s)}{m_\alpha(x_\nu)}\, d\Sigma(s).$$

Then, making the parameter α increase indefinitely, and taking the most probable state of our collection, we have the expected result. Q.E.D.

In order to follow closely the classical derivation of expression (8), we have derived an entropy arising from the usual notion of the "most probable state" of a statistical system (Here the fictitious system of a large number of independent Brownian particles). In fact, this is not the only possibility. It would also have been correct to derive expression (8) from the consideration of a fictitious gas of Brownian particles at local equilibrium.[15] Another way to obtain the maximization of (8) consists in admitting the Jaynes' "maximum entropy prin-

[12] It is possible to take the origin as the starting point because of the invariance of the theory for translations, which is due to the fact that the gas is uniform.

[13] P. Methee, Commun. Math. Helv. **28**, 225 (1954).

[14] L. Schwartz, *Théorie des distributions* (Hermann & Cie., Paris, 1951).

[15] S. R. De Groot and P. Mazur, *Non-Equilibrium Thermodynamics* (North Holland Publishing Company, Amsterdam, 1960).

ciple''[16] the basis of which comes from considerations of information theory. Unfortunately the basis of this principle is not very convincing, though its rigorous justification certainly arises from information theory.

All that we have done up to now is to suggest strongly the maximization of (8). Perhaps a correct derivation could be found from considering the entropy properties of the limit law for the addition of a large number of independant random variables. Unfortunately this has not yet been done.

III. TRANSITION DENSITY

Let us now apply the results of the preceding sections. The first two moments are given by

$$X^\mu(s) = \langle x^\mu \rangle = K \int_{L\Sigma\cap\Gamma^+} x^\mu \rho(x_\alpha, s)\, d\Sigma(s), \tag{9}$$

$$X^{\mu\nu}(s) = \langle x^\mu x^\nu \rangle = K \int_{L\Sigma\cap\Gamma^+} x^\mu x^\nu \rho(x_\alpha, s)\, d\Sigma(s), \tag{10}$$

while the normalization condition of $\rho(x_\alpha, s)$ is

$$1 = \langle 1 \rangle = K \int_{L\Sigma\cap\Gamma^+} \rho(x_\nu, s)\, d\Sigma(s) \tag{11}$$

and the invariant entropy

$$S = \langle \log \rho(x_\alpha, s) \rangle = K \int_{L\Sigma\cap\Gamma^+} \rho(x_\alpha, s) \log \rho(x_\alpha, s)\, d\Sigma(s), \tag{12}$$

where K is the operation defined in Sec. 1 and L a Lorentz transformation. In Eqs. (9), (10), (11), and (12) we have restricted the domain of integration to the part of the surface Σ (or $L\Sigma$) within Γ^+ because of the requirement of causality which implies that x^μ should be timelike. Note that we have implicitly assumed that the Brownian particle is initially (i.e., at "time" $s = 0$) at the origin. Now we have a variational problem to solve: we must maximize expression (12) taking into account conditions (9), (10), and (11). To do this let us introduce fifteen Langrange multipliers: C, η'^μ, and $\eta'^{\mu\nu}$. We have now the variation equation

$$\delta S + C\delta\langle 1 \rangle + \eta'^\mu \delta\langle x_\mu \rangle + \eta'^{\mu\nu}\delta\langle x_\mu x_\nu \rangle = 0,$$

which reduces to

$$\delta \int_{\Sigma(s)\cap\Gamma^+} \{C\rho(x_\nu, s) - \rho(x_\nu, s) \log \rho(x_\nu, s) - \eta'^\mu x_\mu \rho(x_\nu, s) - \eta'^{\mu\nu} x_\mu x_\nu \rho(x_\alpha, s)\} = 0$$

or

$$C\,\delta\rho - \eta'^\mu x_\mu\,\delta\rho - \eta'^{\mu\nu} x_\mu x_\nu\,\delta\rho\,\delta\rho - \log\rho\,\delta\rho = 0. \tag{13}$$

From (13), we find the expression of $\rho(x_\nu, s)$,

$$\rho(x_\nu, s) = R(s) \exp[-\eta^\mu x_\mu - \eta^{\mu\nu} x_\mu x_\nu] \quad \text{and} \quad x_\mu \in \Sigma(s) \cap \Gamma^+. \tag{14}$$

It is clear that the fifteen constants $R(s)$, η^μ, $\eta^{\mu\nu}$, obtained from the fifteen Lagrange multipliers C, η'^μ, $\eta'^{\mu\nu}$, are to be considered as functions of s.

Before solving the principal problem concerning expression (14), which is to express the fifteen unknown constants $R(s)$, η^μ, $\eta^{\mu\nu}$, we shall give the expression for the transition density. We remember that expression (14) is simply

$$\rho(x_\nu, s) \equiv \rho(0, 0 \mid x_\nu, s) \equiv \rho(x_\nu - 0, s - 0).$$

Consequently, the general expression for the transition density will be

$$\rho(x_0^\nu, s_0 \mid x_1^\nu, s_1) = 1_{\Gamma^+(x_0^\nu)} R(s_1 - s_0) \times \exp[-\eta^\mu(x_\mu^1 - x_\mu^0) - \eta^{\mu\nu}(x_\mu^1 - x_\mu^0)(x_\nu^1 - x_\nu^0)] \cdot \delta_{\Sigma(s_1 - s_0)}. \tag{14}'$$

In the above formula, the factor 1_{Γ^+} simply expresses the causality condition.

If the first two moments of expression (14) are given, it is in principle possible to determine the unknown constants by solving the system

$$X^\mu(s) = K\left(\int_{\Sigma(s)} x^\mu R(s) \times \exp[-\eta^\alpha x_\alpha - \eta^{\alpha\beta} x_\alpha x_\beta]\, d\Sigma(s)\right),$$

$$X^{\mu\nu}(s) = K\left(\int_{\Sigma(s)} x^\mu x^\nu R(s) \times \exp[-\eta^\alpha x_\alpha - \eta^{\alpha\beta} x_\alpha x_\beta]\, d\Sigma(s)\right),$$

to which must be joined the normalization condition (11). K is the operation defined in Sec. 1.

The only thing which can be said *a priori* is that the matrix $||\eta^{\mu\nu}||$ is Hermitian and must be positive-definite so that the normalization condition makes sense for timelike or null vectors $\eta^{\mu\nu}x_\mu x_\nu > 0$; $x_\mu x^\mu \geq 0$.

The Covariant-Diffusion Equation

We shall derive the covariant diffusion in the usual way. In order to have the customary hypotheses for the stochastic nature of the Brownian

[16] E. T. Jaynes, Phys. Rev. **106**, 620 (1957).

movement, we first define what is a covariant Markov process, since this is the commonly admitted assumption. A stochastic process, the transition density of which satisfies the following Chapman–Kolmogorov property:

$$f(x_1, s_1 \mid x_2, s_2) = \int_{\Sigma(s_3)} f(x_1, s_1 \mid x_3, s_3) f(x_3, s_3 \mid x_2, s_2)\, d\Sigma(s_3) \quad (15)$$

and

$$s_1 < s_3 < s_2,$$

is Markovian by construction. It can be called a covariant Markovian process. It must be noted that the transition density $f(x_1 s_1 \mid x_2 s_2)$ is a superfical density *on* the surface $\Sigma(s_2)$. Unfortunately, definition (15) is not well suited to the description of stationarity

$$f(x_1, s_1 \mid x_2, s_2) = f(x_1, 0 \mid x_2, s_2 - s_1),$$

because of the following direct consequence of this equality which implies that

$$\Sigma(s_2) = \Sigma(s_2 - s_1).$$

Therefore only the σ-partitions, which can be used to describe stationarity, are constructed from one arbitrarity spacelike surface by taking the surfaces obtained by transformation of the given one by a one-parameter group of timelike translations. Though this last requirement seems to restrict the arbitrariness of the possible σ-partitions, in contrast to the spirit of the special theory of relativity, we shall limit ourselves to such a σ-partition in order to derive the covariant diffusion equation; however, in the following section, this equation will be shown to be valid whatever the σ-partition.

Now, we need other statistical assumptions. Let Δs denote an interval of pseudotime long enough for a particle to reach a local equilibrium state ($s \gg \beta^{-1} \sim \delta S = 0$) but still short enough for the net displacement δx^μ to be small, i.e., $\langle \delta x^\mu \rangle \ll 1$; $\langle \delta x^\mu \delta x^\nu \rangle \ll 1$. The Chapman–Kolmogorov equation then reads (because of the invariance of ρ under the translations of $\mathfrak{M}^4$).

$$\rho(x^\mu, s + \Delta s) = \rho(x^\mu, s) * \rho(x^\mu, \Delta s) = \int \rho(x^\mu - \delta x^\mu, s)\rho(\delta x^\mu, \delta s)\, d(\delta\Sigma).$$

Since δx^μ is assumed to be small, one can expand $\rho(x^\mu - \delta x^\mu, s)$ in powers of δx^μ. We have

$$\rho(x^\mu, s + \Delta s) = \int d(\delta\Sigma)\{\rho(x^\mu, s) - \delta x^\mu \partial_\mu \rho(x_\nu, s) + \tfrac{1}{2}\delta x^\mu \delta x^\nu \partial_{\mu\nu}\rho(x_\alpha, s) + \cdots\}\rho(\delta x_\nu, s)$$

which can be written

$$\rho(x^\mu, s + \Delta s) - \rho(x^\mu, s) = -\overline{\delta x^\mu}\partial_\mu \rho + \overline{\tfrac{1}{2}\delta x^\mu \delta x^\nu}\partial_{\mu\nu}\rho, \quad (16)$$

where the bar stands for the operation I_L defined in Sec. I, and where we have assumed that higher terms behave like $0(\Delta s^2)$. Then taking the K-transformation of the two sides of Eq. (16), we have

$$\rho(x^\mu, s + \Delta s) - \rho(x^\mu, s) = -\langle \delta x^\mu \rangle \partial_\mu \rho + \tfrac{1}{2}\langle \delta x^\mu \delta x^\nu \rangle \partial_{\mu\nu}\rho. \quad (17)$$

Assuming now that

$$\lim_{\Delta s \to 0} \frac{\langle \delta x^\mu \rangle}{\Delta s} = \alpha^\mu$$

and

$$\lim_{\Delta s \to 0} \frac{\langle \delta x^\mu \delta x^\nu \rangle}{\Delta s} = \alpha^{\mu\nu};$$

then, after taking the limit $\Delta s \to 0$, (17) reads

$$\partial\rho/\partial s = -\alpha^\mu \partial_\mu \rho + \tfrac{1}{2}\alpha^{\mu\nu}\partial_{\mu\nu}\rho \quad \text{for} \quad x^\mu \neq 0 \quad \text{and} \quad s \neq 0 \quad (18)$$

or

$$\partial\rho/\partial s + \alpha^\mu \partial_\mu \rho - \tfrac{1}{2}\alpha^{\mu\nu}\partial_{\mu\nu}\rho = \delta_{(4)} \otimes \delta_{(s)} \quad (18')$$

which is valid everywhere. These assumptions are more restrictive than the usual ones (see, e.g., Ref. 17) but are more clear from a physical point of view.

This is the covariant generalization of the usual diffusion equation. From its definition, α^μ appears to be the mean velocity of our collection of Brownian particles (or of our fictitious fluid), while $\alpha^{\mu\nu}$ may be called the inverse-diffusion tensor. The physically interesting solution of the diffusion equation (18′) must satisfy

$$\lim_{s \to 0} \rho(x_\nu, s) = \delta_{(4)}.$$

Here we shall not derive the diffusion equation for a Brownian particle in a fluid submitted to arbitrary forces; this will be done in the second part of this paper. The fact that $\rho(x_\nu, s)$ involves five variables instead of four, seems at first sight to be most surprising. However this is no longer the case when we remember that these five variables are not independant: they are interrelated by the equation

$$\Sigma(x_0, x_1, x_2, x_3) = s.$$

[17] J. L. Doob, *Stochastic Processes* (John Wiley & Sons, Inc., New York, 1962).

Note also that in order to obtain our diffusion equation, we have made two hypotheses concerning $\langle \delta x^\mu \rangle$ and $\langle \delta x^\mu \delta x^\nu \rangle$. We remark that the simplest form for $\langle \delta x^\mu \rangle$ and $\langle \delta x^\mu \delta x^\nu \rangle$ is

$$\langle \delta x^\mu \rangle \equiv \langle x^\mu \rangle_0 = \alpha^\mu s, \tag{19}$$

$$\langle \delta x^\mu \delta x^\nu \rangle \equiv \langle x^\mu x^\nu \rangle_0 = \alpha^{\mu\nu} s. \tag{20}$$

Equation (19) describes the motion of the center of gravity of our fictitious gas, and therefore can be considered as exact, while Eq. (20) could constitute a covariant generalization of the Einstein relation; in fact we shall see that the correct generalization of the Einstein relation is[18]

$$\langle x^\mu x^\nu \rangle_0 - \langle x^\mu \rangle_0 \langle x^\nu \rangle_0 = \alpha^{\mu\nu} s.$$

Calculation of the Coefficients η^μ and $\eta^{\mu\nu}$

Let us now introduce expression (14)′ where we have made $s_0 = 0$, $x_0 = 0$, into Eq. (18) and write that (14)′ satisfies (18) identically. Consequently we obtain a relation involving η^μ, $\eta^{\mu\nu}$, α^μ, $\alpha^{\mu\nu}$, and x^μ. Since this relation must hold for all values of x^μ, the coefficients of x^μ and $x^\mu x^\nu$ must vanish identically. There remains a constant term which must be equal to zero. This constant term is of no interest here since it gives only the coefficient $R(s)$ which depends on the surface $\Sigma(s)$. The coefficients of x^μ and $x^\mu \cdot x^\nu$ vanishing, we find the two relations

$$\dot{\eta}^\mu = -2(\eta^{\mu\nu}\alpha_\nu + \alpha_{\rho\sigma}\eta^\sigma \eta^{\mu\rho}), \tag{21}$$

$$\dot{\eta}^{\mu\nu} = -2\alpha_{\rho\sigma}\eta^{\mu\rho}\eta^{\nu\sigma}, \tag{22}$$

where the dot denotes the pseudotime derivative. Setting

$$N = ||\eta^{\mu\nu}||,$$

$$A = ||\alpha^{\mu\nu}||,$$

and

$$\eta^\mu = n, \qquad \alpha^\mu = a,$$

Eqs. (21) and (22) then read

$$\dot{n} + 2(Na + NAn) = 0, \tag{23}$$

$$\dot{N} + 2NAN = 0. \tag{24}$$

First let us solve (24). Its solution is obviously $N = \frac{1}{2}A^{-1}s^{-1}$ or $\eta^{\mu\nu} = \frac{1}{2}(A^{-1})^{\mu\nu}s^{-1}$. The simplest solution of (23), consistent with the nonrelativistic limit is such that $\dot{n} = 0$. Consequently, $n = -A^{-1}a$ or $\eta^\mu = -(A^{-1})^{\mu\nu}\alpha_\nu$. Now the transition density reads

$$\rho(x_0' s_0 \mid x_1' s_1) = 1_{\Gamma^+(x_0')} R(s_1 - s_0) \exp\left[(A^{-1})^{\mu\nu}\alpha_\mu x_\nu - \tfrac{1}{2}(A^{-1})^{\mu\nu}x_\mu x_\nu (s_1 - s_0)^{-1}\right] \delta_{\Sigma(s_1 - s_0)}, \tag{25}$$

where

$$x_\mu \equiv x_\mu^1 - x_\mu^0.$$

IV. AXIOMATICS OF THE RELATIVISTIC BROWNIAN MOTION AND THEIR APPLICATION TO THE COVARIANT WIENER INTEGRAL

As a first application to the previous section, we shall derive a covariant Wiener integral. With this covariant Wiener integral we hope to treat different problems of radiation of a relativistic plasma at local equilibrium or problems of binary collisions, etc. The need for a covariant Wiener integral can be explained as follows. Since many physical quantities are functionals of the trajectories, we need an integral over the set of trajectories in order to obtain measurable quantities (which are mean values). (An example of such a physical quantity depending on the trajectory is the following: the electromagnetic field of a relativistic particle. For instance, it is interesting to compute the mean radiation emitted by a relativistic test particle with the help of a covariant Wiener integral.) These trajectories must be continuous in order to be physically interesting. The main difference between the covariant Wiener integral and the classical one is that the former is an integral over the set of *timelike curves*. Note also that this set of trajectories can either be constructed from infinite curves or from finite parts of them [i.e., trajectories beginning at some point (e.g., the origin) and finishing at some other (namely: x_n^μ)]. However, before defining such a covariant Wiener integral, we must specify the axiomatics of a covariant Brownian motion and more particularly its associated stochastic process.

The Stochastic Process Associated with the Relativistic Brownian Motion

From the expression (25) for the transition density, it is easy to derive a covariant Wiener measure (and hence a covariant Wiener integral) with the use of well-known techniques.[19–22] However this derivation is entirely formal and cannot be rigorously justified (essentially because of the presence in ex-

[18] R. Hakim, "A covariant Theory of Relativistic Brownian Motion II." Orsay, TH. 68.

[19] P. Courrege, *Le processus stochastique du mouvement brownien* (Centre de Documentation Universitaire, Paris, 1963).

[20] M. Kac, *Probability and Related Topics in Physical Sciences* (Interscience Publishers, London, 1959).

[21] E. W. Montroll, "On the statistical Mechanics of Transport Processes" in *Termodinamica dei Processi irreversibili* Varenna Summer School, 1959 (N. Zanichelli, Bologna, 1960).

[22] I. M. Guelfand and A. M. Yaglom, J. Math. Phys. 1, 48 (1960).

pression (25) of an *arbitrary surface* $\Sigma(s)$. Therefore we should proceed on the basis of the following remarks.

Let us first consider the expression (25) for the transition density. It has the form

$$\rho = F(x_1^\mu - x_2^\mu, s_1 - s_2)\delta_\Sigma,$$

and therefore can be considered as the projection of F on the arbitrary 3-surface Σ. Then it seems that the main role is played by F rather than by the density itself. This remark suggests strongly the association of a stochastic process with the covariant Brownian motion, which has the following properties: (1) its transition density is $F(x_1^\mu - x_0^\mu, s_1 - s_0)$ which is normalized in $\Gamma^+(x_0^\mu)$; (2) the values of this stochastic process are in $\mathfrak{M}^4$ rather than on the Σ's (3) its trajectories are timelike. At this point we must note that $F(x, s)$ describes an "unphysical" stochastic process. In fact, we show in Part II that F is as "physical" as ρ. The difference between ρ and F is that the former has been derived on the basis of an assumption chosen so as to follow closely the usual definition of physical space, while the latter appear to be more specifically relativistic.[4,18] $F(x_\nu, s)$ will arise in a natural manner from a covariant generalization of the Langevin equation.

Before giving the axioms of a covariant Brownian motion, let us give the definition of a Lorentz–Gaussian stochastic process.

Definition 2. A stochastic process is said to be Lorentz–Gaussian if all its joint distributions are (a) invariant under a change of coordinates defined by a Lorentz transformation; (b) derived from the maximization of an entropy (in the sense of information theory) once the first two moments are given. The entropy is given by

$$S = -X \int F \log F \, d_4x,$$

where X is an arbitrary positive constant.

Let us now give the axioms of a covariant Brownian motion.

Definition 3. A stochastic process X_s^μ defined on R^+ (or R) and the values of which are in $\mathfrak{M}^4$ is said to be a covariant Brownian motion if the three following properties are satisfied.

(B_1)—X_s^μ is a process with independant increments. This means that for any finite sequence $s_0 < s_1 < s_2 < \cdots < s_p$, with $s_i \in R^+$, the random variables $X_{s_{k+1}}^\mu - X_{s_k}^\mu$, $(0 \leq k < p)$ are independent.

(B_2)—The stochastic process X_s^μ is Lorentz–Gaussian (see Def. 2).

(B_3)—The trajectories of X_s^μ satisfy: (B_{31}) they are almost certainly timelike; (B_{32}) they are almost certainly continuous.

Axiom B_1 arises from the fact that our gas is uniform, which implies the stationarity of the process and the invariance of the transition density under the translations of $\mathfrak{M}^4$. B_1 also implies that the process is Markovian. As a consequence of B_1, the transition density can be written

$$F(x_1^\mu s_1 \mid x_2^\mu s_2) = F(0, 0 \mid x_2^\mu - x_1^\mu, s_2 - s_1) \equiv F(x_2^\mu - x_1^\mu, s_2 - s_1).$$

This equality implies that the Chapman–Kolmogorov property, expressing analytically the Markovian character of the process is given by

$$F(x_\mu, s_1 + s_2) = F(x_\mu, s_1) * F(x_\mu, s_2). \tag{26}$$

Equation (26) is always defined since, as we shall see below, $F(x_\mu, s)$ belongs to the convolution algebra $\mathfrak{D}'_{\Gamma+}$.

Axiom B_2 is only the covariant generalization of the usual assumption of the Gaussian character of the Brownian movement. The Lorentz–Gaussian character of the stochastic process used with the Chapman–Kolmogorov property (26) yields the diffusion equation

$$\{\partial/\partial s + \alpha^\mu \partial_\mu - \tfrac{1}{2}\alpha^{\mu\nu}\partial_{\mu\nu}\}F = 0. \tag{27}$$

Axiom B_{32} is only the requirement that the stochastic process describe in fact a physical process: a Brownian particle can, in principle, be followed continuously; its trajectory is continuous.

Axiom B_{31} is an axiom of causality. Its interest comes from the following obvious proposition.

Proposition 2. A covariant Markovian process X_s^μ has its trajectories almost certainly timelike if and only if its transition density $F(x_1^\mu s_1 | x_2^\mu s_2)$ is zero out of the null cone $\Gamma^+(x_1^\mu)$.

Proof. (1) if $F(x_1^\mu s_1 | x_2^\mu s_2)$ is zero out of the null cone $\Gamma^+(x_1^\mu)$, there cannot exist transitions such that $x_2^\mu - x_1^\mu$ be spacelike. Consequently the trajectories of the process are almost certainly timelike; (2) if the Markovian process has timelike trajectories, there are no possible transitions such that $x_2^\mu - x_1^\mu$ be spacelike and consequently F must be zero out of the null cone $\Gamma^+(x_1^\mu)$. Q.E.D.

From a physical point of view, the expression "almost certainly" occuring in Axiom B_3 is sufficient. In fact, it is known[23] that there exists an equivalent

[23] P. Courrege (private communication).

stochastic process which has its trajectories exactly timelike and continuous. As a consequence of the causality condition, F is normalized in $\Gamma^+(x_1^\mu)$:

$$\int_{\Gamma^+(x_1,\mu)} F(x_2^\mu - x_1^\mu, s_2 - s_1)\, d_4x = 1.$$

An Explicit Expression for F

Maximizing the entropy,

$$S = -X \int_{\Gamma^+} F(x^\mu, s) \log F(x^\mu, s)\, d_4x,$$

and taking into account the fact that the first two moments are given [Note that, because of the definition of F (which no longer involves an arbitrary surface through the factor δ_Σ), mean values now have a good variance.[4]], we obtain

[using the same techniques as in Sec. 3: we first introduce fifteen Lagrange multipliers; next we solve our variational problem; finally introducing the derived expression into the diffusion equation (27), we find expression (28)]

$$F(x^\mu, s) = 1_{\Gamma^+} R(s) \exp\left[\alpha^\mu x_\mu - \tfrac{1}{2}(A^{-1})^{\mu\nu} x_\mu x_\nu s^{-1}\right], \quad (28)$$

which is, as expected, expression (25) (modulo its domain of definition).

We therefore have[18]

$$\langle x^\mu \rangle = \alpha^\mu s$$

$$\langle x^\mu x^\nu \rangle = \alpha^{\mu\nu} s + \alpha^\mu \alpha^\nu s^2$$

In expression (28) we see that $F(x^\mu, s)$ is different from zero *on* the frontier of Γ^+, and is zero outside Γ^+. Consequently F verifies equation (27) only *in* the null cone Γ^+, and not everywhere. From a physical point of view, this is sufficient since a Brownian particle cannot reach the null cone (if it is massive). From a mathematical point of view, this difficulty gives rise to interesting problems[23] not yet solved.

The diffusion equation (27) can now allow us to obtain the expression for the factor $R(s)$. Introducing expression (28) in Eq. (27) and letting the coefficients of x^μ and $x^\mu x^\nu$ vanish, we obtain the preceding expressions (see Sec. 3) for the coefficients η^μ and $\eta^{\mu\nu}$; there remains a constant term which must also be identically equal to zero

$$\dot{R}(s) + R(s)\{\tfrac{1}{2}\alpha^{\mu\nu}[(A^{-1})^\rho_\mu (A^{-1})^\beta_\nu \alpha_\rho \alpha_\beta] + 2s^{-1}\} = 0.$$

This first-order linear homogeneous equation can be rewritten as

$$\dot{R}(s) + R(s)\{\tfrac{1}{2}\langle a, A^{-1}a\rangle + 2s^{-1}\} = 0,$$

the solution of which is

$$R(s) = R_0 s^{-2} \exp\left[-\tfrac{1}{2}\langle a, A^{-1}a\rangle s\right],$$

where R_0 is a constant of integration to be determined by the normalization condition. It is found (Ref. 24) that

$$R_0 = (2\pi)^{-2}(\det A)^{-1/2}.$$

Therefore, we can write the transition density as

$$F(x_1^\mu s_1 \mid x_2^\mu s_2) = R_0 s^{-2} \exp\{-(A^{-1})^{\mu\nu}[2(s_2 - s_1)]^{-1} \times [x_\mu^2 - x_\mu^1 - \alpha_\mu(s_2 - s_1)] \times [x_\nu^2 - x_\nu^1 - \alpha_\nu(s_2 - s_1)]\}(1_{\Gamma^+}). \quad (29)$$

This last expression is the same as the classical one except that it involves: (a) one more dimension (4 instead of 3); (b) a causality factor (1_{Γ^+}); (c) a quadratic form which is not positive and definite everywhere.

The Covariant Wiener Integral

Once we have defined the stochastic process associated with the covariant Brownian motion, we can derive correctly the covariant Wiener measure (which allows the construction of a canonical covariant Brownian motion[19]).

Let us first recall the following lemma.

Lemma. For any sequence $0 \leq s_1 \leq s_2 \leq \cdots \leq s_n$ and for each measurable and bounded numerical function on $\mathfrak{M}^{4n}$,

$$(x_1^\mu, x_2^\mu \cdots x_n^\mu) \to f(x_1^\mu \cdots x_n^\mu),$$

the expectation value of $f(X_1^\mu \cdots X_n^\mu)$ is given by

$$E(f) = \int \cdots \int f(x_1^\mu \cdots x_n^\mu) \times \prod_i F(x_i^\mu - x_{i-1}^\mu, s_i - s_{i-1})\, d_4x_i$$

and this property is equivalent to the Markovian character of the covariant Brownian motion X_s^μ.

Proof. see Ref. 19, p.40.

This lemma allows us to construct a probability π_U on the space $(\mathfrak{M}^4)^U$ where $U = \{s_1, s_2, \cdots, s_n\}$ and $0 \leq s_1 < s_2 < \cdots < s_n$, for the random variable:

$$\omega \to [X_s^\mu(\omega)]_{s \in U}$$

(ω belongs to the sample space). We have

$$d\pi_U = \prod_{i=1}^{i=n} F(x_i^\mu - x_{i-1}^\mu, s_i - s_{i-1})\, d_4x_i$$

and this is true for any finite n.

It is now easy to show with the help of the Chapman–Kolmogorov property, that the $\{\pi_U\}$ con-

[24] F. Lurcat & P. Mazur, Nuovo Cimento **31**, 140 (1964).

stitutes a projective system[25] of Radon probabilities.[26] Then, according to the

Kolmogorov Theorem: Let $(\Omega_i, \mathfrak{I}_i)_{i\in I}$ be a family of locally compact topological spaces, and let $F_i = \sigma(\mathfrak{I}_i)$ be the Borelian field of the topological space Ω_i. Let Φ be the set of all finite parts of I. Then, if for all $J \in \Phi$, π_J is a Radon probability on the locally compact product space, $\prod_{i\in J} \Omega_i = \Omega_J$, then the projective system $\{\pi_J\}_{J\in\Phi}$ has a unique projective limit π.

Proof. See Ref. 17, p. 21.

The projective system $\{\pi_U\}$ has a unique projective limit π. π is defined on the set of all continuous mappings of R^+ into $\mathfrak{M}^4$, namely $(\mathfrak{M}^4)^{R+}$ which is nothing but the trajectories of the stochastic process X^μ_s. Because of proposition 2, and because of the support of F, we are sure that π is a measure on the continuous timelike trajectories and that the set of all other trajectories (discontinuous or spacelike) has a null measure.

We have therefore obtained a covariant Wiener measure. Next, because of a well-known theorem (see, e.g., Ref. 26, p.46), it is easy to construct a covariant Wiener integral from the covariant Wiener measure.

Connected Questions

(1) Let us now give an explicit expression for $d\pi_U$, and hence for $d\pi$,

$$d\pi_U = \prod_{i=1}^{i=n} F(x_i^\mu - x_{i-1}^\mu, s_i - s_{i-1})\, d_4x_i$$

$$= \prod_{i=1}^{1=n} \frac{R_0}{(s_i - s_{i-1})} \times \exp\left[-(A^{-1})^{\mu\nu}\{(x_\mu^i - x_\mu^{i-1}) - \alpha_\mu(s_i - s_{i-1})\} \times \{(x_\nu^i - x_\nu^{i-1}) - \alpha_\nu(s_i - s_{i-1})\}\right] d_4x_i.$$

Introducing now the reduced variables $X_i^\mu = x_i^\mu - \alpha^\mu s_i$, we have

$$d\pi_U = \exp\left[-\tfrac{1}{2}(A^{-1})^{\mu\nu} \sum_{i=1}^{i=n} \frac{X_\mu^i - X_\mu^{i-1}}{(s_i - s_{i-1})} \cdot \frac{X_\nu^i - X_\nu^{i-1}}{(s_i - s_{i-1})} \times (s_i - s_{i-1})\right] \prod_{i=1}^{i=n} \frac{R_0}{(s_i - s_{i-1})^2}\, d_4x_i.$$

Now taking the limit $n \to \infty$, $d\pi_U$ becomes

$$d\pi = \lim_{n\to\infty} d\pi_U = \exp\left[-\tfrac{1}{2}(A^{-1})^{\mu\nu} \int_0^s \frac{dX_\mu}{ds} \cdot \frac{dX_\nu}{ds}\, ds\right] \prod_0^s d[x^\mu(s)]$$

where we implicitly limit ourselves to the parts of the trajectories corresponding to (0, s). We must note that the expression

$$\frac{1}{2} \int_0^s (A^{-1})^{\mu\nu} \dot{X}_\mu \dot{X}_\nu\, ds$$

is *never* the relativistic action of the particle, because the tensor $A^{\mu\nu}$ is never the metric tensor (in order that the normalization integral have a sense). This will be shown for the explicit form of $A^{\mu\nu}$ (see Part II).

(2) Let $G[x^\mu(s)]$ be a functional of the trajectory of a Brownian particle, and assume that it is π-integrable. The mean value of G will be[21]

$$\bar{G} = \int G[x^\mu(s)]\, d\pi = \lim_{n\to\infty} \int \cdots \int G(x_1^\mu \cdots x_n^\mu)\, d\pi_U$$

where we have "discretized" the functional G[19],

$$G[x^\mu(s)] = \lim_{n\to\infty} G(x_1^\mu \cdots x_n^\mu).$$

As an example, let us compute the mean value of the functional

$$G[x^\mu(s)] = \exp\left\{-\xi \int_{s_0}^{s_1} V[x^\mu(s)]\, ds\right\}$$

which is interesting relative to sources problems or to a generalization of the Bloch equation.[19] Once "discretized", this functional is

$$G[x^\mu(s)] = \lim_{n\to\infty} \exp\left\{-[\sum_k \xi V(x_k^\mu)]\left[\frac{s_1 - s_0}{n}\right]\right\} = \lim_{n\to\infty} \prod_k \exp\left[-\xi V(x_k^\mu)\left(\frac{s_1 - s_0}{n}\right)\right]$$

so that the calculation of

$$\bar{G}(x_1^\mu s_1 \mid x_0^\mu s_0) = \int_{(x_0^\mu(s_0))}^{(x_1^\mu(s_1))} G[x^\mu(s)]\, d\pi$$

reduces to the calculation of the fundamental solution of a partial differential equation [the limits of the integral symbolically designate the fact we integrate G over the set of the trajectories beginning at $x^\mu(s_0)$ and finishing at $x^\mu(s_1)$].[20-22] This solution must satisfy

$$\lim_{s_0\to s_1} \bar{G}(x_1^\mu s_1 \mid x_0^\mu s_0) = \delta_{(4)}(x_1^\mu - x_0^\mu).$$

---

[25] N. Bourbaki, *Théorie des Ensembles* (Hermann & Cie, Paris 1963), Chap. 3.

[26] P. Courrege, *Théorie de la mesure* (Centre de Documentation Universitaire, Paris, 1963).

Using the same methods as in the nonrelativistic case,[20-22] it is found that this partial differential equation is

$$\{\partial/\partial s_1 + \alpha^\mu \partial_\mu - \tfrac{1}{2}\alpha^{\mu\nu}\partial_{\mu\nu} - \xi V(x_1^\mu)\}\bar{G}(x_1^\mu s_1 \mid x_0^\mu s_0) = \delta_{(4)}(x_1^\mu - x_0^\mu) \otimes \delta(s_1 - s_0). \quad (30)$$

As it is expected, Eq. (30) becomes the diffusion equation (27) when $\xi = 0$. Let us note that we have implicitly assumed that V is bounded and measurable (for a V not bounded, see the procedure used in Ref. 20).

V. CONCLUSION

In this paper we have obtained several results which we should now discuss. First, we have defined a notion of evolution in statistical relativistic mechanics, introducing the concept of σ-partition of space–time which clarifies the idea of statistical time needed to locate a global quantity. This statistical time is an independent parameter and it is fully used only in Sec. 4, where we consider the associated stochastic process of the Brownian particle. Next we have given a method permitting the calculation of the covariant mean values of local quantities on an element of a given σ-partition. We must point out that such a mean value does not always exist, so that we need, at least in the case of a conditional expectation value, an invariant σ-partition: because of axiom R_4, a global quantity has always a covariant mean value. This is a reason for using the σ-partition $\mathfrak{IC}(\tau)$.[6]

Another remark must be made. It might be surprising at first sight that our covariant mean value plays a very small role in the derivation of the transition density of the Brownian particle, but when we look more closely at the derivation of this transition density, we see that only local quantities are involved, and therefore there is no need for a covariant mean value at this point (except in a rigorous derivation).

It must also be pointed out that all we have said about Lorentz transformation, Lorentz observers, Lorentz frame, simultaneity, etc. would require a special discussion. This will be done in a further paper.

As we have already said, the method used (i.e. the maximization of an entropy) is only strongly suggested by the classical results and cannot be considered as proved. However, we believe that the transition density obtained is correct, as it could be shown by studying the addition of independent random variables, or from a covariant Langevin equation.

The fact that $F(x_\nu, s)$ is more fundamental than $F(x_\nu, s)\cdot\delta_S$ clarifies the meaning of the diffusion equation (27). $F(x_\nu, s)$ must be regarded as a density in $\mathfrak{M}^4$ normalized in the future null cone, which at first sight has no sense by itself, except when projected on an element of a σ-partition. We shall return to this point in Part II.

The second part of this paper (Ref. 18) will be devoted to the approach to equilibrium: Focker–Planck's equations, Langevin's equations, Brownian particle in a force field. This second part will clarify the results here obtained.

ACKNOWLEDGMENTS

We are deeply indebted to Professor R. Nataf for the interest shown in this work, and profoundly grateful to Professor P. Mazur and Professor J. Yvon for illuminating criticisms and active interest. We wish also to express our thanks to Professor F. Lurçat, Dr. R. Bkouche, and Dr. P. Leruste for many valuable discussions. We are grateful to Dr. M. M. Alexander and Dr. P. Loève (Orsay), and to the referees for a large number of corrections to this manuscript.

On the Solution of the BBGKY Equations for a Dense Classical Gas*

FRANK C. ANDREWS†
Theoretical Chemistry Institute, University of Wisconsin, Madison, Wisconsin
(Received 22 February 1965)

Formal solutions of the BBGKY equations of classical statistical mechanics are obtained in the form of integral equations. This form makes it particularly straightforward to obtain the density corrections to the functional equation for f_2 and to the evolution equation for f_1. The physical implications of the derivation, in particular the validity of Bogoliubov's functional hypothesis, are discussed in detail.

INTRODUCTION

RECENT years have seen a very large number of treatments of the problem of the approach to equilibrium in dense classical gases starting from the Liouville equation.[1-9] These methods are being shown to agree in their results[10] so that considerable faith may now be placed in them. Unfortunately, the successful theories have been sufficiently complicated that their physical implications have not been obvious. Their use in numerical problems (like computing density corrections to Chapman–Enskog transport coefficients[11]) has been limited.

This paper resulted from a search for a way to treat the problem with minimum mathematical complication and maximum physical insight.

In Part I, the low-order BBGKY equations, viewed as simple inhomogeneous equations, are formally solved by expressing them as integral equations. The advantages arising from integral equation solutions in statistical mechanics have been stressed previously by Montroll.[12] Also, a density expansion is introduced and discussed. The resulting first density corrections to the functional equation for f_2 and to the Boltzmann equation are derived in Part II. In Part III these results are compared with the results of other theories. The physical implications of the theory are discussed in Part IV.

I. SOLUTION OF THE BBGKY HIERARCHY

The BBGKY Equations

This theory starts with the BBGKY equations for low-order reduced distribution functions [f_s normalized to $N!/(N-s)!$] in the case of no external field. Incorporation of external fields should pose no serious problem later. The BBGKY equations arise from a partial integration of the Liouville equation.[13] Expressed in notation commonly used, these equations have the form

$$\frac{\partial f_1(1)}{\partial t} + \mathbf{v}_1 \cdot \frac{\partial f_1(1)}{\partial \mathbf{r}_1} = J_1(1), \qquad \frac{\partial f_s}{\partial t} + L_s f_s = J_s, \quad (1)$$

where

$$L_s = \sum_{i=1}^{s} \mathbf{v}_i \cdot \frac{\partial}{\partial \mathbf{r}_i} - \sum_{i<j=1}^{s} \theta_{ij}, \quad \theta_{ij} = -\mathbf{F}_{ij} \cdot \frac{\partial}{\partial \mathbf{p}_{ij}},$$

$$J_s = \int d\mathbf{r}_{s+1}\, d\mathbf{p}_{s+1} \sum_{i=1}^{s} \theta_{i,s+1} f_{s+1}. \quad (2)$$

* Supported in part by National Science Foundation Grant NSF-G20725. Portions of this paper were presented at the Midwest Conference on Theoretical Physics, Ames, Iowa, June, 1964.

† Alfred P. Sloan Fellow.

[1] N. N. Bogoliubov, *"Problems of a Dynamical Theory in Statistical Physics"* 1946, English translation in *Studies in Statistical Mechanics*, edited by J. de Boer and G. E. Uhlenbeck (North-Holland Publishing Company, Amsterdam, 1962), Vol. 1.

[2] S. T. Choh, "The Kinetic Theory of Phenomena in Dense Gases" thesis, University of Michigan (1958).

[3] G. E. Uhlenbeck and G. W. Ford, *Lectures in Statistical Mechanics* (American Mathematical Society, Providence, Rhode Island, 1963), Chap. VII.

[4] E. D. G. Cohen, *Fundamental Problems in Statistical Mechanics* (North-Holland Publishing Company, Amsterdam, 1962), pp. 110–156; Physica **28**, 1025, 1045, 1060 (1962); J. Math. Phys. **4**, 183 (1963).

[5] I. Prigogine, *Non-Equilibrium Statistical Mechanics* (Interscience Publishers, Inc., New York, 1962); and references therein (Balescu, Résibois, Henin, etc.); and P. Résibois, J. Math. Phys. **4**, 166 (1963).

[6] J. G. Kirkwood, reviewed with references in S. Rice and H. Frisch, Ann. Rev. Phys. Chem. **11**, 187 (1960).

[7] M. S. Green, in *Lectures in Theoretical Physics*, edited by W. E. Brittin, B. W. Downs, and J. Downs (Interscience Publishers, Inc., New York, Vol. III, 1961), pp. 195–220, and references therein. M. S. Green and R. A. Piccirelli, Phys. Rev. **132**, 1388 (1963).

[8] H. B. Hollinger and C. F. Curtiss, J. Chem. Phys. **33**, 1386 (1960) D. K. Hoffman and C. F. Curtiss, Phys. Fluids, **8**, 667 (1965).

[9] R. M. Lewis, J. Math. Phys. **2**, 222 (1961).

[10] J. Stecki, Phys. Fluids **7**, 33 (1964). P. Résibois, Phys. Letters **9**, 139 (1964).

[11] R. F. Snider and C. F. Curtiss, Phys. Fluids **1**, 122 (1958). C. F. Curtiss, M. McElroy, and D. K. Hoffman, Intern. J. Eng. Soc., **3**, 269 (1965).

[12] E. W. Montroll, in *Fundamental Problems in Statistical Mechanics*, Ref. 4, pp. 230–249.

[13] Reference 3 and references therein. These works, however, impose boundary conditions on f_N that are more stringent than needed. Equations 1 are valid for any condition on f_N or f_{s+1} at the walls which conserves number of particles in the system.

The L_s is the streaming operator in the phase space of the s particles, J_s is the collision integral which always involves knowing the distribution function for $s + 1$ particles, and the molecular force term is θ_{ij} where the $\mathbf{F}_{ij}$ are the intermolecular forces.

Formal Solution of the BBGKY Equations

We use the following formal solution of Eq. (1):

$$f_s(\mathbf{r}^s, \mathbf{p}^s, t) = e^{-tL_s} f_s(\mathbf{r}^s, \mathbf{p}^s, 0) + \int_0^t d\tau e^{-\tau L_s} J_s(\mathbf{r}^s, \mathbf{p}^s, t - \tau). \quad (3)$$

The meaning of the operators is discussed in Appendix A, where also Eq. (3) is formulated from Eq. (1) by standard Green's function techniques. Proof of Eq. (3) may also be established by differentiating it with respect to time and recovering Eq. (1). The first term of Eq. (3) represents the effect of the s particles interacting among themselves, and the last term represents the effect of "outside" particles interacting with the s particles through all times.

It proves helpful later in eliminating non-Markovian time dependence from these equations to express $f_1(0)$ in terms of $f_1(t)$. This is done by applying $\exp(tL_s)$ to Eq. (3) for $s = 1$:

$$f_1(0) = e^{tL_1} f_1(t) - \int_0^t d\tau e^{(t-\tau)L_1} J_1(t - \tau). \quad (4)$$

The Initial Correlations

In any mechanical problem, the nature of the initial conditions profoundly affects the solution. In our problem the crucial question is the following: Over what range of distances of separation of the s particles is there *correlation* of the particles' positions and momenta at time zero? Correlation is defined as any information which prevents the factorization,

$$f_s(\mathbf{r}^s, \mathbf{p}^s, 0) = \prod_{i=1}^{s} f_1(\mathbf{r}_i, \mathbf{p}_i, 0). \quad (5)$$

Initial conditions can be imagined for which correlation extends over even infinite distances. However, so long as initial correlation plays an important role in a problem, the solution depends intimately on the nature of the correlation; every problem is different; and no one would expect a general treatment such as we are seeking here (e.g., like a Boltzmann equation) to be useful.

We stress here that one's interest in f_s is in that region of s-particle phase space in which the s particles are engaged in a collision (at time t). If one rules out the chance of bound states, then the positions $\exp(-tL_s)\mathbf{r}^s$ eventually (after a time τ_{coll} characteristic of the duration of an s-body collision) get farther and farther apart with increasing time. We must limit ourselves to times long compared to a time τ_{corr} characteristic of the initial correlations, where τ_{corr} is defined as being representative of the minimum time beyond which the following expression is correct to whatever order is being treated in the given smallness parameter:

$$e^{-tL_s} f_s(\mathbf{r}^s, \mathbf{p}^s, 0) = e^{-tL_s} \prod_{i=1}^{s} f_1(\mathbf{r}_i, \mathbf{p}_i, 0). \quad (6)$$

This limitation is discussed in detail in Part IV.

The Density Expansion

Casting Eq. (1) into the form of an integral equation still leaves the infinite hierarchy of coupled equations, since J_s contains f_{s+1}. For a moderately dense gas, the obvious small expansion parameter to break this hierarchy is Nr_0^3/V, where r_0 is a length characteristic of the range of the intermolecular forces (or of the "size" of a molecule). The problem now is to determine the order in Nr_0^3/V of the various terms encountered in Eqs. (3) and (4).

First, we observe that τJ_s is of order $N\tau\bar{v}r_0^3/V$ compared to f_s. The N arises from the normalization of f_{s+1} compared to f_s. The $\tau\bar{v}r_0^2$ is the effective volume swept out by the s particles during the time τ. This arises because of the presence of $\mathbf{F}_{ij}$ in θ_{ij}, which restricts the effective configuration integral. It must be compared with the V of the normalization integral.

In addition to the characteristic times τ_{corr} and τ_{coll}, it is convenient to discuss a time $\tau_{\text{mfp}} = O(\Lambda/\bar{v})$, where Λ is the mean free path. Thus, τ_{mfp} characterizes the average duration of free flight of particles between collisions. By the definition of τ_{mfp}, it follows that $N\tau_{\text{mfp}}\bar{v}r_0^2/V$ is of the order of unity. Also, by the very definition of a relatively dilute gas, it follows that $\tau_{\text{coll}} \ll \tau_{\text{mfp}}$.

Until the discussion of Part IV, at which time the following restrictions are discussed and drastically reduced, we assume simply that $\tau_{\text{corr}} = O(\tau_{\text{coll}})$ and that $\tau_{\text{coll}} < t \ll \tau_{\text{mfp}}$. We omit bound states from consideration. In this case the last term of Eq. (3) is of order Nr_0^3/V compared to the others, and the hierarchy may be closed.

Next, consider the first BBGKY equation,

$$\tau_{\text{coll}} \frac{\partial f_1(1)}{\partial t} + \tau_{\text{coll}} \mathbf{v}_1 \cdot \frac{\partial f_1(1)}{\partial \mathbf{r}_1} = \tau_{\text{coll}} J_1(1). \quad (7)$$

The term on the rhs is $O(Nr_0^3/V)$ compared to f_1.

It is possible at this point, as most theories do, to make the other terms small by limiting the problem to slowly varying conditions and to make an expansion in the gradients of the properties as well as in density. We simply limit our consideration to cases in which the first and second terms of Eq. (7) are separately of the same order in Nr_0^3/V as the collision term. For a homogeneous system this is valid, because there the second term is zero and the first equals the third. In a steady state this is also valid, because there the first term is zero and the second equals the third. Our limitation omits from consideration systems for which the first two terms are each large, but whose sum is small, e.g., a gas rushing into a vacuum.

By combining Eqs. (3), (4), and (6), one obtains an expression for f_s valid to lowest order in Nr_0^3/V:

$$f_s(\mathbf{r}^s, \mathbf{p}^s, t) \simeq \mathcal{S}_{12\cdots s} \prod_{i=1}^{s} f_1(i, t), \tag{8}$$

where

$$\mathcal{S}_{12\cdots s} \equiv e^{-tL_s} \prod_{i=1}^{s} e^{tL_1(i)}. \tag{9}$$

The order of application of successive exponential operators, as in Eq. (9), is from left to right. The streaming operators $\mathcal{S}$ thus demand that the phase point $\mathbf{r}^s$, $\mathbf{p}^s$ first be changed into what would be found if the s particles moved backward along exact s-body trajectories a time t, then forward along straight-line noninteracting trajectories for a time t. Clearly, if the initial $\mathbf{r}^s$ are indeed close enough together to represent an s-particle collision, then for $t > \tau_{coll}$, the $\mathcal{S}$ are independent of t.

Since the only t-dependence in Eq. (8) rests in the $f_1(t)$'s, it is clear from the discussion of Eq. (7) that not only is $\tau_{coll}J_s$ of higher order in Nr_0^3/V than f_s, but so also is $\tau_{coll}\partial f_s/\partial t$. In fact, to lowest order in Nr_0^3/V for $s > 1$, the BBGKY equation is simply $L_s f_s = 0$. Even this includes one term of higher order, since L_s could be rewritten so as to express explicit change in f_s due to motion of the center of mass of the s particles and also to relative motions among the s particles. The term in the center of mass variables would then be of one higher order in Nr_0^3/V, but this does not affect the fact that $L_s f_s = 0$ is correct to lowest order for $s > 1$. For later reference we write three particular examples of Eq. (1) to lowest order using Eq. (8):

$$L_3(1, 2, 3)\mathcal{S}_{123}f_1(1)f_1(2)f_1(3) = [L_2(1, 2) + L_1(3) - \theta_{13} - \theta_{23}]\mathcal{S}_{123}f_1(1)f_1(2)f_1(3) = 0, \tag{10}$$

$$L_2(2, 3)\mathcal{S}_{23}f_1(2)f_1(3) = [L_1(2) + L_1(3) - \theta_{23}]\mathcal{S}_{23}f_1(2)f_1(3) = 0, \tag{11}$$

$$L_2(1, 3)\mathcal{S}_{13}f_1(1)f_1(3) = [L_1(1) + L_1(3) - \theta_{13}]\mathcal{S}_{13}f_1(1)f_1(3) = 0. \tag{12}$$

We also note that the difference between $J(t - \tau)$ and $J(t)$ in Eqs. (3) and (4) is of order $\tau_{coll}\partial f_1/\partial t$. Thus, $J(t)$ may be used in these integrands with results correct to the two lowest orders in Nr_0^3/V.

II. EVOLUTION EQUATIONS CORRECTED FOR DENSITY

Functional Equation for f_2 with Complete Density Correction

Substitution from Eq. (6) into Eq. (3) for f_2 with use of Eq. (4), replacement of $J(t - \tau)$ by $J(t)$, and use of Eq. (8) in the J's yields

$$\begin{aligned} f_2(1, 2) \simeq \mathcal{S}_{12}f_1(1)f_1(2) + \lambda \int d\mathbf{r}_3\, d\mathbf{p}_3 \int_0^t d\tau \\ \times [e^{-tL_s(1,2)}e^{(t-\tau)L_1(2)}\theta_{23}e^{tL_1(1)}\mathcal{S}_{23} \\ - e^{-tL_s(1,2)}e^{(t-\tau)L_1(1)}\theta_{13}e^{tL_1(2)}\mathcal{S}_{13} \\ + e^{-\tau L_s(1,2)}(\theta_{13} + \theta_{23})\mathcal{S}_{123}]f_1(1)f_1(2)f_1(3). \end{aligned}$$

A counting parameter λ ($\lambda = 1$) has been used to display the relative order in Nr_0^3/V of the terms. Performance of the τ-integral formally and substitution of L's for θ's by using Eqs. (10)–(12) yields

$$\begin{aligned} f_2(1, 2) \simeq \mathcal{S}_{12}f_1(1)f_1(2) + \lambda \int d\mathbf{r}_3\, d\mathbf{p}_3 \{-e^{-tL_s(1,2)} \\ \times e^{tL_1(2)}(1 - e^{-tL_1(2)})L_1^{-1}(2)[L_1(2) + L_1(3)]e^{tL_1(1)}\mathcal{S}_{23} \\ - e^{-tL_s(1,2)}e^{tL_1(1)}(1 - e^{-tL_1(1)})L_1^{-1}(1)[L_1(1) + L_1(3)] \\ \times e^{tL_1(2)}\mathcal{S}_{13} + (1 - e^{-tL_s(1,2)})L_2^{-1}(1, 2) \\ \times [L_2(1, 2) + L_1(3)]\mathcal{S}_{123}\}f_1(1)f_1(2)f_1(3). \end{aligned}$$

Since no function of $\mathbf{r}_3$ precedes the $L_1(3)$'s of each bracket, these terms all yield surface integrals over the boundary which vanish by particle conservation at the boundaries. In each case, the remaining operator in the bracket is canceled by the preceding inverse operator. This leaves simply

$$\begin{aligned} f_2(1, 2) \simeq \mathcal{S}_{12}f_1(1)f_1(2) + \lambda \int d\mathbf{r}_3\, d\mathbf{p}_3 \\ \times [-\mathcal{S}_{12}(1 - e^{-tL_1(2)})\mathcal{S}_{23} - \mathcal{S}_{12}(1 - e^{-tL_1(1)})\mathcal{S}_{13} \\ + (1 - e^{-tL_s(1,2)})\mathcal{S}_{123}]f_1(1)f_1(2)f_1(3). \end{aligned} \tag{13}$$

This may be simplified by noting that

$$\int d\mathbf{r}_3\, d\mathbf{p}_3 e^{-tL_3(1,2)}[e^{-tL_3(2,3)} + e^{-tL_3(1,3)} - e^{-tL_3(1,2,3)} - e^{-tL_1(3)}] \times e^{tL_1(1)}e^{tL_1(2)}e^{tL_1(3)}f_1(1)f_1(2)f_1(3) = 0$$

for $t \gg \tau_{\text{coll}}$ if $\mathbf{r}_1$ and $\mathbf{r}_2$ are within approximately r_0 of each other. The reason is that the exp $[-tL_2(1, 2)]$ takes particles 1 and 2 a long way apart. Then the contribution of the first two terms in the bracket is the same as that of the last two terms and they cancel. This may be rewritten as

$$\int d\mathbf{r}_3\, d\mathbf{p}_3[\mathcal{S}_{12}e^{-tL_1(2)}\mathcal{S}_{23} + \mathcal{S}_{12}e^{-tL_1(1)}\mathcal{S}_{13} - e^{-tL_3(1,2)}\mathcal{S}_{123} - \mathcal{S}_{12}]f_1(1)f_1(2)f_1(3) = 0, \quad (14)$$

which may be used to replace the three terms in Eq. (13) which have exponential operators remaining:

$$f_2(1, 2) \simeq \mathcal{S}_{12}f_1(1)f_1(2) + \lambda \int d\mathbf{r}_3\, d\mathbf{p}_3 \times (\mathcal{S}_{123} - \mathcal{S}_{12}\mathcal{S}_{23} - \mathcal{S}_{12}\mathcal{S}_{13} + \mathcal{S}_{12})f_1(1)f_1(2)f_1(3). \quad (15)$$

Equation (15) is the generally accepted form of f_2, correct to the first two orders in density.[2-4,7,8]

Boltzmann Equation for f_1 with Complete Density Correction

The evolution equation for f_1 may be obtained simply by substituting Eq. (15) directly into Eq. (1):

$$\frac{\partial f_1(1)}{\partial t} + \mathbf{v}_1\cdot\frac{\partial f_1(1)}{\partial \mathbf{r}_1} = \int d\mathbf{r}_2\, d\mathbf{p}_2\theta_{12}\mathcal{S}_{12}f_1(1)f_1(2) + \lambda \int d\mathbf{r}_2\, d\mathbf{p}_2\, d\mathbf{r}_3\, d\mathbf{p}_3\theta_{12}(\mathcal{S}_{123} - \mathcal{S}_{12}\mathcal{S}_{23} - \mathcal{S}_{12}\mathcal{S}_{13} + \mathcal{S}_{12})f_1(1)f_1(2)f_1(3). \quad (16)$$

We use the notation

$$\mathcal{S}_{12}\mathbf{r}_1 = \mathbf{r}_1'; \quad \mathcal{S}_{12}\mathbf{r}_2 = \mathbf{r}_2'; \quad \mathcal{S}_{12}\mathbf{p}_1 = \mathbf{p}_1'; \quad \mathcal{S}_{12}\mathbf{p}_2 = \mathbf{p}_2' \quad (17)$$

and expand $\mathcal{S}_{12}f_1(1)f_1(2)$ in an effort to obtain a Boltzmann collision integral plus corrections:

$$f_1(\mathbf{r}_1', \mathbf{p}_1')f_1(\mathbf{r}_2', \mathbf{p}_2') \simeq f_1(\mathbf{r}_1, \mathbf{p}_1')f_1(\mathbf{r}_1, p_2') + [\lambda f_1(\mathbf{r}_1, \mathbf{p}_2')(\mathbf{r}_1' - \mathbf{r}_1)\cdot\partial f_1(\mathbf{r}_1, \mathbf{p}_1')/\partial\mathbf{r}_1 + \lambda f_1(\mathbf{r}_1, \mathbf{p}_1')(\mathbf{r}_2' - \mathbf{r}_1)\cdot\partial f_1(\mathbf{r}_1, \mathbf{p}_2')/\partial\mathbf{r}_1]. \quad (18)$$

This assumes that f_1 is essentially a linear function of $\mathbf{r}$ over distances of order $\bar{v}\tau_{\text{coll}}$.

The first term of the expansion is

$$f_1(\mathbf{r}_1, \mathbf{p}_1', t)f_1(\mathbf{r}_1, \mathbf{p}_2', t),$$

which we write $\bar{f}_2$. The primed momenta depend only on $\mathbf{r}_{21}$, $\mathbf{p}_1$, and $\mathbf{p}_2$, and all points $\mathbf{r}_{21}$, $\mathbf{p}_1$, and $\mathbf{p}_2$ on the same two-body trajectory lead to the same $\mathbf{p}_1'$ and $\mathbf{p}_2'$. Thus, $\mathbf{p}_1'$ and $\mathbf{p}_2'$ are constants of the motion on such a trajectory; thus $\bar{f}_2$ satisfies the equation for such constants:

$$\theta_{12}\bar{f}_2 = \mathbf{v}_{21}\cdot\left(\frac{\partial\bar{f}_2}{\partial\mathbf{r}_{21}}\right)_{\mathbf{r}_2+\mathbf{r}_1} \equiv \mathbf{v}_{21}\cdot\left(\frac{\partial\bar{f}_2}{\partial\mathbf{r}_{21}}\right)_{\mathbf{r}_1} - \tfrac{1}{2}\mathbf{v}_{21}\cdot\left(\frac{\partial\bar{f}_2}{\partial\mathbf{r}_1}\right)_{\mathbf{r}_{21}}. \quad (19)$$

Since θ_{12} is odd in $\mathbf{r}_{21}$, $\bar{f}_2$ must be odd in $\mathbf{r}_{21}$. Thus, the last term in Eq. (19) is odd in $\mathbf{r}_{21}$ and vanishes in the integral over $\mathbf{r}_{21}$. The $\bar{f}_2$ part of $\mathcal{S}_{12}f_1(1)f_1(2)$ therefore gives the classical Boltzmann surface integral, as is seen by choosing as integration surface a large cylinder about $\mathbf{r}_1$ with axis parallel to $\mathbf{v}_{21}$. On the surface where $d\mathbf{S}_{21}\cdot\mathbf{v}_{21} < 0$, $f_1(1')f_1(2')$ equals simply $f_1(1)f_1(2)$ because the particles have not yet interacted. This surface yields the unprimed term in the conventional Boltzmann equation. The other surface where $d\mathbf{S}_{21}\cdot\mathbf{v}_{21} > 0$ yields the primed term.

The bracketted correction terms in Eq. (18), written $[\cdots]$, may be rewritten by use of the following identity:

$$\begin{aligned}[\cdots] &= f_1(2')(\mathbf{r}_1' - \mathbf{r}_2')\cdot\frac{\partial f_1(1')}{\partial\mathbf{r}_1} + f_1(1')(\mathbf{r}_2' - \mathbf{r}_1')\cdot\frac{\partial f_1(2')}{\partial\mathbf{r}_1} \\ &\quad + f_1(2')(\mathbf{r}_2' - \mathbf{r}_1)\cdot\frac{\partial f_1(1')}{\partial\mathbf{r}_1} + f_1(1')(\mathbf{r}_1' - \mathbf{r}_1)\cdot\frac{\partial f_1(2')}{\partial\mathbf{r}_1} \\ &= \tfrac{1}{2}(\mathbf{r}_1' - \mathbf{r}_1 + \mathbf{r}_2' - \mathbf{r}_1)\cdot\frac{\partial}{\partial\mathbf{r}_1}[f_1(1')f_1(2')] \\ &\quad + \tfrac{1}{2}(\mathbf{r}_2' - \mathbf{r}_1')\cdot\left[f_1(1')\frac{\partial f_1(2')}{\partial\mathbf{r}_1} - f_1(2')\frac{\partial f_1(1')}{\partial\mathbf{r}_1}\right] \\ &= \tfrac{1}{2}(\mathbf{r}_2 - \mathbf{r}_1)\cdot\frac{\partial}{\partial\mathbf{r}_1}[f_1(1')f_1(2')] \\ &\quad + \tfrac{1}{2}(\mathbf{r}_2' - \mathbf{r}_1')\cdot\left[f_1(1')\frac{\partial f_1(2')}{\partial\mathbf{r}_1} - f_1(2')\frac{\partial f_1(1')}{\partial\mathbf{r}_1}\right]. \quad (20)\end{aligned}$$

In the first step the same quantities were added and subtracted to $[\cdots]$; in the second step $[\cdots]$ was equated to half the sum of itself and the rhs of the first step; in the third step use was made of the fact that $\mathbf{r}_1 + \mathbf{r}_2 = \mathbf{r}_1' + \mathbf{r}_2'$.

On using Eq. (18) in Eq. (16), then applying Eq. (19) and making a surface integral out of the

volume integral, and using Eq. (20) to rewrite the correction terms, one obtains

$$\frac{\partial f_1(1)}{\partial t} + \mathbf{v}_1\cdot\frac{\partial f_1(1)}{\partial \mathbf{r}_1}$$
$$= \int d\mathbf{p}_{21}\int d\mathbf{S}_{21}\cdot\mathbf{v}_{21} f_1(\mathbf{r}_1, \mathbf{p}_1', t) f_1(\mathbf{r}_1, \mathbf{p}_2', t)$$
$$+ \frac{\lambda}{2}\int d\mathbf{r}_2\, d\mathbf{p}_2\theta_{12}\Big\{(\mathbf{r}_2 - \mathbf{r}_1)\cdot\frac{\partial}{\partial \mathbf{r}_1}[f_1(\mathbf{r}_1, \mathbf{p}_1', t)$$
$$\times f_1(\mathbf{r}_1, \mathbf{p}_2', t)] + (\mathbf{r}_2' - \mathbf{r}_1')\cdot\int f_1(\mathbf{r}_1, \mathbf{p}_1', t)$$
$$\times \frac{\partial f_1(\mathbf{r}_1, \mathbf{p}_2', t)}{\partial \mathbf{r}_1} - f_1(\mathbf{r}_1, \mathbf{p}_2', t)\frac{\partial f_1(\mathbf{r}_1, \mathbf{p}_1', t)}{\partial \mathbf{r}_1}\Big]\Big\}$$
$$+ \lambda\int d\mathbf{r}_2\, d\mathbf{p}_2\, d\mathbf{r}_3\, d\mathbf{p}_3\theta_{12}(\mathcal{S}_{123} - \mathcal{S}_{12}\mathcal{S}_{23}$$
$$- \mathcal{S}_{12}\mathcal{S}_{13} + \mathcal{S}_{12})f_1(1)f_1(2)f_1(3). \quad (21)$$

The first term on the rhs of Eq. (21) is the classical Boltzmann integral written in a slightly unconventional form, as noted after Eq. (19). The remaining terms give the complete correction to the next order in Nr_0^3/V.

III. COMPARISON WITH RESULTS OF OTHERS

Theory of Enskog

The first major work on correcting the Boltzmann equation for higher densities was done for a hard-sphere fluid by Enskog in 1921.[14] He ruled out "three-particle contributions" on physical grounds, stating that only two-body collisions play a role in hard-sphere fluids. There is, however, great difficulty in defining three-particle contributions, since by their very order in density one is led to suspect that *all* the correction terms present in Eq. (21) are three-body in origin.

In order to obtain Enskog's result we note that to the first two orders in Nr_0^3/V, f_2 is a solution of Eq. (1):

$$\lambda\frac{\partial f_2^\circ}{\partial t} + \mathbf{v}_{21}\cdot\left(\frac{\partial f_2^\circ}{\partial \mathbf{r}_{21}}\right)_{\mathbf{r}_1} + \lambda\mathbf{v}_{21}\cdot\left(\frac{\partial f_2'}{\partial \mathbf{r}_{21}}\right)_{\mathbf{r}_1} + \lambda\mathbf{v}_1\cdot\left(\frac{\partial f_2^\circ}{\partial \mathbf{r}_1}\right)_{\mathbf{r}_{21}} - \theta_{12}f_2 = \lambda J_2^\circ, \quad (22)$$

where we have changed to the variables $\mathbf{r}_1$ and $\mathbf{r}_{21} = \mathbf{r}_2 - \mathbf{r}_1$. We have also written Eq. (15) in the form $f_2 = f_2^\circ + \lambda f_2'$. If the $\theta_{12}f_2$ in first BBGKY equation is replaced by its value from Eq. (22) and the surface integral about $\mathbf{r}_1$ obtained, one gets

$$\frac{\partial f_1(1)}{\partial t} + \mathbf{v}_1\cdot\frac{\partial f_1(1)}{\partial \mathbf{r}_1}$$
$$= \int d\mathbf{p}_{21}\int d\mathbf{S}_{21}\cdot\mathbf{v}_{21} f_2^\circ + \lambda\int d\mathbf{p}_{21}\int d\mathbf{S}_{21}\cdot\mathbf{v}_{21} f_2'$$
$$+ \lambda\int d\mathbf{r}_{21}\, d\mathbf{p}_{21}\left[\frac{\partial f_2^\circ}{\partial t} + \mathbf{v}_1\cdot\left(\frac{\partial f_2^\circ}{\partial \mathbf{r}_1}\right)_{\mathbf{r}_{21}} - J_2^\circ\right]. \quad (23)$$

The time derivative term yields

$$f_1(\mathbf{r}_1, \mathbf{p}_1')\frac{\partial f_1(\mathbf{r}_1, \mathbf{p}_2')}{\partial t} + f_1(\mathbf{r}_1, \mathbf{p}_2')\frac{\partial f_1(\mathbf{r}_1, \mathbf{p}_1')}{\partial t}$$
$$= f_1(\mathbf{r}_1, \mathbf{p}_1')J_1^\circ(\mathbf{r}_1, \mathbf{p}_2') + f_1(\mathbf{r}_1, \mathbf{p}_2')J_1^\circ(\mathbf{r}_1, \mathbf{p}_1')$$
$$- f_1(\mathbf{r}_1, \mathbf{p}_1')\mathbf{v}_2'\cdot\frac{\partial f_1(\mathbf{r}_1, \mathbf{p}_2')}{\partial \mathbf{r}_1} - f_1(\mathbf{r}_1, \mathbf{p}_2')\mathbf{v}_1'\cdot\frac{\partial f_1(\mathbf{r}_1, \mathbf{p}_1')}{\partial \mathbf{r}_1}, \quad (24)$$

where use was made of Eq. (1). Since

$$\left(\frac{\partial}{\partial \mathbf{r}_1}\right)_{\mathbf{r}_{21}} = \left(\frac{\partial}{\partial \mathbf{r}_1}\right)_{\mathbf{r}_2} + \left(\frac{\partial}{\partial \mathbf{r}_2}\right)_{\mathbf{r}_1} = \left(\frac{\partial}{\partial \mathbf{r}_1'}\right)_{\mathbf{r}_2'} + \left(\frac{\partial}{\partial \mathbf{r}_2'}\right)_{\mathbf{r}_1'}, \quad (25)$$

the other terms in the bracket of Eq. (23) yield

$$f_1(\mathbf{p}_1')\mathbf{v}_1\cdot\frac{\partial f_1(\mathbf{p}_2')}{\partial \mathbf{r}_1} + f_1(\mathbf{p}_2')\mathbf{v}_1\cdot\frac{\partial f_1(\mathbf{p}_1')}{\partial \mathbf{r}_1} - J_2^\circ(1, 2). \quad (26)$$

We can now rewrite Eq. (23) as Boltzmann collision integral plus corrections, an expression equivalent to Eq. (21) but easier to discuss from the point of view of Enskog:

$$\frac{\partial f_1(1)}{\partial t} + \mathbf{v}_1\cdot\frac{\partial f_1(1)}{\partial \mathbf{r}_1} = \int d\mathbf{p}_{21}\int d\mathbf{S}_{21}\cdot\mathbf{v}_{21} f_1(1')f_1(2')$$
$$+ \lambda\int d\mathbf{p}_{21}\int d\mathbf{S}_{21}\cdot\mathbf{v}_{21}\Big[f_1(2')(\mathbf{r}_1' - \mathbf{r}_1)\cdot\frac{\partial f_1(1')}{\partial \mathbf{r}_1}$$
$$+ f_1(1')(\mathbf{r}_2' - \mathbf{r}_1)\cdot\frac{\partial f_1(2')}{\partial \mathbf{r}_1} + f_2'(1, 2)\Big]$$
$$+ \lambda\int d\mathbf{r}_{21}\, d\mathbf{p}_{21}\Big[f_1(1')(\mathbf{v}_1 - \mathbf{v}_2')\cdot\frac{\partial f_1(2')}{\partial \mathbf{r}_1}$$
$$+ f_1(2')(\mathbf{v}_1 - \mathbf{v}_1')\cdot\frac{\partial f_1(1')}{\partial \mathbf{r}_1} + f_1(1')J_1^\circ(2')$$
$$+ f_1(2')J_1^\circ(1') - J_2^\circ(1, 2)\Big]. \quad (27)$$

The positions of all f_1's appearing in Eq. (27) are $\mathbf{r}_1$, and the times are all t. The $^\circ$ superscript means to lowest order in density; f_2° and f_2' are given by Eq. (15).

Enskog's hard-sphere result is equivalent to the surface integral correction in Eq. (27) only, omitting the term in f_2'. His surface is chosen to be infinitesimally outside the surface of particle 1, thus $\mathbf{r}_2' = \mathbf{r}_2$

[14] D. Enskog, Kungl. Svenska Vetenskaps Akad. Handl. **63**, No. 4, 1921, given in S. Chapman and T. G. Cowling *The Mathematical Theory of Non-Uniform Gases* (Cambridge University Press, London, 1952), 2nd ed., Chap. 16.

and $\mathbf{r}_1' = \mathbf{r}_1$, since the two-body collision time is zero. He made an *ad hoc* correction for all further density effects by multiplying his result by χ, a function of $\frac{1}{2}(\mathbf{r}_1 + \mathbf{r}_2)$ only. This form of correction was known then to be an oversimplification, but it was hoped to yield a useful semi-empirical relation. When one considers Enskog's theory in the light of Eq. (27), it is clear how difficult it is to say which terms have two-body and which terms three-body origin. Only the f_2' and the J_2° demand knowledge of three-particle trajectories.

Theory of Born and Green

A major contribution to the study of dense gases was made in 1946 by Born and Green.[15] They initially limited their work to considering only two-body effects, which permitted them to use an expression which gives the same results as simply

$$f_2(1, 2) = e^{-tL_s(1,2)} f_1(1, t = 0) f_1(2, t = 0). \quad (28)$$

They then change the time dependence in the f_1's from 0 to t by using what amounts to Eq. (4) with the last term neglected. This leads to the approximation $f_2(1, 2) = \mathcal{S}_{12} f_1(1) f_1(2)$ and thus to the evolution equation given by Eq. (21) without the last integral with the $\mathcal{S}$'s.

It is a matter of definition whether this theory of Born and Green is consistent at even the two-body level. The neglected corrections like the last term of Eq. (4) are truly of two-body origin. In fact, only the $\mathcal{S}_{123}$ term of Eq. (28) involves true three-body collisions. The difficulty with the Born–Green theory is of course that it sets out to get only *some* of the correction terms at a given order in a perturbation series.

Theory of Bogoliubov, Choh, Uhlenbeck, and Cohen

The first complete nonequilibrium theory of the statistical mechanics of dense gases was suggested by Bogoliubov in 1946,[1] elaborated by Choh and Uhlenbeck in 1958,[2,3] and later by Cohen.[4] This theory rested on certain assumptions which proved very difficult to analyze. These are discussed in Part IV of this paper. The results of the Bogoliubov theory for the density corrections are identical to Eqs. (15) and (21).

Theory of Prigogine, Balescu, Résibois, *et al.*

Since 1959, the group in Brussels under the guidance of Professor I. Prigogine has developed a formal theory of irreversible processes[5] which has had profound influence. The Brussels theory was not especially suited to a density expansion, and only recently has it been shown[10] to give identical results to the Bogoliubov theory, and thus to Eqs. (15) and (21). Some of the physical implications of the Brussels theory are discussed in light of this paper in Part IV.

[15] M. Born, and H. S. Green, Proc. Roy. Soc. (London) **A188**, 10 (1946).

IV. DISCUSSION OF THE RESULTS

List of Physical Restrictions Imposed in order to Obtain Some of the Results

In the subsequent discussion it is convenient to refer to a numbered list of requirements used in obtaining some of the results of Parts I and II:

1. System composed of classical point particles interacting through pairwise, short-range forces.
2. System volume large compared to any length of interest; enough particles that unity may be neglected compared to N. Fixed number of particles in system.
3. Region of configuration space many mean free paths from the walls.
4. No external forces.
5. Time long compared to τ_{corr} and τ_{coll}.
6. No bound states.
7. Convergent "density" expansion in powers of Nr_0^3/V, with the physical restriction imposed in the discussion of Eq. (7) regarding rate of change of f_1 with time and space.
8. Use of causality to determine the "direction of time"; i.e., correlation now is due to interaction in the past.
9. Approximately linear form for $f_1 = f_1(\mathbf{r})$ over distances of the order of a mean free path.

Validity and Significance of the Formal Solution, Eq. (3).

The BBGKY equations are identical in content with the formal solution, Eq. (3). As we have used them, they demand requirements 1–4 inclusive. Requirement No. 1 could be weakened with no difficulty by formally incorporating an explicit three-body term into the intermolecular potential. This term would arise in the first density corrections and would have the effect of changing the trajectories and energies during three-body collisions.

Requirements Nos. 1 and 3 are coupled. The system must be large so the effects of the boundaries are negligible in the interior. *Any* choice of boundary condition that is consistent with the physical situation is satisfactory. Requirement No. 3 eliminates explicit dependence in Eq. (3) on the particular

boundary condition chosen. Usual ways of eliminating explicit concern about boundary conditions are either to let $N \to \infty$, $V \to \infty$, $N/V \to$ finite; or else to work on a time scale $t \ll V^{\frac{1}{3}}/\bar{v}$.

If one desired to relax requirement No. 4 and formally incorporate the effects of external forces, no serious problem would arise. Added terms would appear in the operator L giving the effect of the external force on the free motion. Furthermore, in computing the trajectories during collisions, the paths of all particles would be modified by the forces.

The Initial Correlations and the "Initial Period" of Bogoliubov

The problem of statistical mechanics starts with a physical system about which, as a result of measurements, certain information is known. By using this information and the fact that the system evolves through the mechanical motion of its constituent particles, statistical mechanics sets out to compute the probabilities of the various possible results of subsequent measurements that might be made on the system.[16] Since the initial information is almost never complete mechanical information, one constructs a statistical ensemble into which one builds the initial information and nothing more.

Information on initial correlations might be of two kinds. The first is precise knowledge of how pairs, triples, etc. are correlated. The second is a less precise knowledge that whatever correlations the particles have, it is similar to the way they would have been correlated had they evolved mechanically from some earlier state at negative time (like $t = -\infty$).

If precise information of the first kind is available, then it is clear whether or not τ_{corr} is finite and therefore how long one must wait for times long compared to τ_{corr} for our general equations to become valid. Information of the first kind is rarely available, and in many cases, of course, τ_{corr} is of the same order as τ_{coll}. However, it is easy to imagine situations for which τ_{corr} is infinite. For example, if the initial positions and momenta of all the particles were known quite closely, this would imply that even very complicated correlations were physically accessible. One would then be forced to study the exact evolution of f_N as an exercise in classical mechanics. The simplications of statistical mechanics would not arise. However, only in special uses of statistical mechanics does there seem any likelihood of devising experiments to yield such complete information.

Another example of initial conditions leading to a large value of τ_{corr} is the following: Suppose a nonequilibrium system is allowed to evolve during the time period $-\tau$ to 0. Then at 0 the momenta of all the particles are simultaneously reversed. Clearly, during the period 0 to τ, the system must evolve away from equilibrium. In this case, the ensemble one constructs at $t = 0$ must reflect the information that with all momenta reversed it evolved from a certain nonequilibrium condition at $t = -\tau$. If this is built into the ensemble, all members will return to the nonequilibrium condition by the time τ. The statistical mechanics agrees with what one knows must happen, but during the period 0 to τ, no simple, general evolution equation describes the ensemble. The special initial conditions dominate the evolution.

The second kind of information on initial correlations is less specific. For a given value of t, one has instead of Eq. (6),

$$e^{-tL^*}f_s(0) = e^{-tL^*}\prod_{i=1}^{s} f_1(i, 0) + \text{correction terms.} \tag{30}$$

The question is, of what order in density are the correction terms? Consider, for example, f_2. After time t ($t > \tau_{coll}$), one is concerned with factoring $f_2(0)$ for r_{21} of the order of $\bar{v}_{21}t$. Particles 1 and 2 are on incoming precollision trajectories. The only way they could be correlated is by one or both of them having bounced off other particles of the fluid after a previous direct interaction. If one other particle is involved, the effect is $O(Nr_0^3/V)$ compared to $f_1(1, 0)f_1(2, 0)$. If ν other particles are involved, the effect is $O(Nr_0^3/V)^\nu$. As $\bar{v}_{21}t$ increases, the most probable number of other particles involved in the correlating collision increases rapidly. Since these initial correlations are completely analogous to those that develop in the gas naturally through mechanical interaction, the following may be stated: For t greater than a typical two-body collision time, the correction terms are at least of one higher order in Nr_0^3/V. For t greater than a typical three-body collision time, the correction terms are at least of two higher order in Nr_0^3/V. The generalization is immediate.

It is clear that general evolution equations arise only for times long compared to both τ_{corr} and τ_{coll}. The formulations in Part I were made under the assumption that τ_{corr} and τ_{coll} were of the same order of magnitude. As seen above, this is by no means

[16] F. C. Andrews, *Equilibrium Statistical Mechanics* (John Wiley & Sons, Inc., New York, 1963), Sec. 3.

necessarily true. However, if in order for t to exceed τ_{corr}, $N\bar{v}tr_0^2/V$ is no longer of order Nr_0^3/V, the work of Part I appears to fail.

This question may be resolved simply by redefining the origin of the time scale, since the absolute value of t is meaningless. With the presence of explicit correlation of the first kind, one merely needs to wait for $t_1 \gg \tau_{corr}$, then choose t_1 as the time origin. The only kind of correlation present then would be of the second kind. The second kind of correlation, however, is always present in the fluid. It is correctly treated by the theory of this paper. So long as it is known that the system evolved mechanically for a time long compared to τ_{coll}, then one may be assured that this source of correlation is accounted for by this theory, whatever origin of time is chosen. This is true even if t is formally equal to zero.

It is clear that the "initial period" of Bogoliubov is the period τ_{corr} during which the initial correlations play a significant role in the problem. During this period, each problem is a special case. Bogoliubov suggested that this period would be of order τ_{coll}, which indeed it may be for many initial conditions. But of course it might be of any length, depending on the initial information.

This also is the period during which in the Brussels theory the presence of "destruction fragments" dominates the evolution of f_s. Only after the natural elimination of these diagrams by the mechanical processes discussed above, does the Brussels theory lead to closed expressions.

In summary, during the time τ_{corr}, the particular evolution demanded by the knowledge of explicit initial correlations loses its dominant role. Then, the only correlations arise from the mechanical motion of the particles. So long as the gas is relatively dilute, i.e., $\tau_{coll} \ll \tau_{mfp}$, general closed equations are then obtained to describe the subsequent evolution.

Validity of Functional Equations for $f_2, f_3, \cdots$

Under requirements 1–8, it is proved above that f_s becomes a time-independent functional of f_1's, with all the time dependence lying in the f_1's. To lowest order, f_s is given by Eq. (8). To the first two orders, f_2 is given by Eq. (15). This establishes the physical requirements under which Bogoliubov's hypothesis that f_s becomes a functional of f_1 is true.

The first five requirements and the possibilities of weakening them are discussed above. Requirement No. 6 is a severe limitation, since at even moderate densities bound states affect gas properties profoundly at low temperatures. Perhaps the increasing simplicity of the physical picture presented by this theory will lead at least to a useful approximation for treating bound states.

One might ask how important requirement No. 7 is to the general theory. In this paper there appears to be a vital need for $\tau_{coll} \ll \tau_{mfp}$ or else the entire analysis fails. In presentations of the Bogoliubov and the Brussels schools, the theory appears to be valid for all densities, including those of liquids. However, the analysis is difficult for the nonconvergent virial-type series that results. Without a well-defined smallness parameter, there is no way of truncating the series. The meaning of τ_{coll} is unclear. The Bogoliubov theory[1-4] avoids the entire question by applying the factorization of Eq. (7) at $t = \infty$. There is some question as to the meaning of infinite time. It is also uncertain how long one must wait for such equations, even if they were tractable, to describe satisfactorily a liquid system.

The density correction terms to even higher orders in Nr_0^3/V could be obtained by methods similar to those used here if ever one wanted them. However, it seems likely that if a gas is so dense that four-body collisions play a significant role in its evolution, then the requirement $\tau_{coll} \ll \tau_{mfp}$ is no longer met, the gas is not reasonably dilute, and the entire analysis fails.

The restriction imposed in the discussion of Eq. (7) could be lifted for gases known to be evolving rapidly with sharp gradients in physical properties. A reordering of the terms would result. This is another case of special initial conditions or special boundary conditions dominating the solution.

Requirement No. 8 is more philosophical than scientific.[17] It is universally accepted that correlation between two events implies that the events interacted at some time in the past. This is simply a statement of causality, and it represents the way the direction of time is always established.

Validity of Boltzmann Equation with Density Correction

The transport or evolution equation for f_1, Eq. (21), has been derived in the form of a classical Boltzmann equation plus correction terms in the density. The result has all nine requirements as necessary and sufficient conditions for its validity. The irreversibility displayed by Eq. (21) is not "merely an illusion introduced by the imperfections of the statistical method" nor does it arise through any approximation. The information available about the system at time zero is built into an ensemble,

[17] H. Reichenbach, *The Direction of Time* (University of California Press, Berkeley, California 1956), especially Sec. 18.

each member of which is in a completely specified mechanical state. Then, using mechanics, the evolution of the ensemble is followed. The system about which one wishes to make predictions is actually represented by only one phase point in the ensemble. but since one does not know which, he must make his predictions from the entire ensemble. If he wants to predict a quantity dependent on only single particles, it is reasonable to predict the average of that quantity weighted by $f_1 d\mathbf{r}\, d\mathbf{p}$. The fact that f_1 goes to equilibrium reflects simply the following truth: A reasonable man's predictions of properties will with time proceed to the essentially time-independent values associated with the equilibrium condition. This is so if the predictions are based on the available initial information and on knowledge that the system evolves through the mechanical motion of its constituent particles. Once the predicted values of physical properties reach the equilibrium values, they stay there forever, even though the system may actually be undergoing very unlikely fluctuations. The likelihoods of various fluctuations are correctly given by the final "equilibrium" form of the function $f_1(\mathbf{r}, \mathbf{p})$. The initial information available is incomplete, and the final information considered of importance is only partial (i.e., f_1 as opposed to f_N), and thus irreversibility is perfectly natural.[18]

ACKNOWLEDGMENTS

The author gratefully acknowledges valuable discussions with P. L. Livingston and C. F. Curtiss. He also acknowledges the National Aeronautics and Space Administration for Grant NsG-275-62 which provided partial support.

APPENDIX (A) THE GREEN'S FUNCTION SOLUTION OF THE BBGKY EQUATIONS.

The Green's functions of the Liouville operator have been previously investigated,[19] but the results applicable to this paper are derived below much more concisely. For brevity we drop the s subscripts and denote the vectors $\mathbf{r}^s$ and $\mathbf{p}^s$ by r and p.

The solution of the homogeneous Liouville equation,

$$\partial f(r, p, t)/\partial t + Lf(r, p, t) = 0 \qquad \text{(A1)}$$

with the initial condition $f = f'(r, p, t')$ at $t = t'$ is

$$f(r, p, t) = \exp\,[-(t - t')L]\, f'(r, p, t'), \qquad \text{(A2)}$$

as may be proved by substitution into Eq. (A1). The fact that $\exp\,[-(t - t')L]$ is the phase-space transformation operator for the s particles may be proved by showing that a general solution of the homogeneous equation is any arbitrary function of the variables

$$\mathbf{r}_i' = \mathbf{r}_i - \int_{t'}^{t} \mathbf{v}_i(t_1)\, dt_1 \quad \text{and} \quad \mathbf{p}_i' = \mathbf{p}_i - \int_{t'}^{t} \mathbf{F}_i(t_1)\, dt_1,$$

where $\mathbf{v}_i(t_1)$ is the exact velocity of particle i at time t_1 and $\mathbf{F}_i(t_1)$ the exact force on it at time t_1, when all s particles have been interacting among themselves. The proof is immediate on substituting such a function into Eq. (A1) and performing the differentiations.

Now, we seek the Green's function which is a solution of the equation,

$$(\partial/\partial t + L)G(r, p, t \mid r', p', t') = \delta(r - r')\delta(p - p')\delta(t - t'). \qquad \text{(A3)}$$

Each delta function in a phase-space variable is of course $(3s)$-dimensional. If g is any solution of the homogeneous equation and $\eta(x)$ is the Heaviside or unit step function defined by $\eta(x) = 1$ for $x \geq 0$, $\eta(x) = 0$ for $x < 0$, then

$$\left(\frac{\partial}{\partial t} + L\right)[\eta(t - t')g] = g\left(\frac{\partial}{\partial t} + L\right)\eta(t - t') = g\delta(t - t'). \qquad \text{(A4)}$$

If we choose $g = \exp\,[-(t - t')]\delta(r - r')\delta(p - p')$ in Eq. (A4), the effect of the operator's acting on $\eta(t - t')g$ is to leave

$$\delta(t - t')\, \exp\,[-(t - t')L]\delta(r - r')\delta(p - p')$$

or simply

$$\delta(t - t')\delta(r - r')\delta(p - p').$$

Thus such a choice of g gives a causal Green's function:

$$G(r, p, t \mid r', p', t') = \eta(t - t')e^{-(t-t')L}\delta(r - r')\delta(p - p') \qquad \text{(A5)}$$

$$= \eta(t - t')e^{(t-t')L'}\delta(r - r')\delta(p - p'). \qquad \text{(A6)}$$

The L' acts on primed quantities. The presence of the Heaviside function is the only difference between the Green's function and the solution of the homogeneous equation, which is really a phase-space transformation function.[20] Which of the two equiv-

[18] An extensive discussion of this viewpoint is given by F. C. Andrews, Proc. Natl. Acad. Sci. U. S. **53**, 1284 (1965); **54**, 13 (1965).

[19] F. C. Andrews, Bull. Classe Sci. Acad. Roy. Belg. **46**, 475 (1960).

[20] H. Jeffreys, Phil. Mag. **33**, 815 (1942); Moyal, J. E., Proc. Cambridge Phil. Soc. **45**, 99 (1949), Ross, J. and Kirkwood, J. G., J. Chem. Phys. **22**, 1094 (1954); Ross, J., J. Chem. Phys. **24**, 375 (1956).

alent forms for G is used depends on the coordinates over which the integration is performed. One does not want the coordinates of integration to be acted on by the exponential operator. So the operator in the first expression may be viewed as acting on the unprimed coordinates, moving them back along the trajectory a time $t - t'$. It is completely equivalent to consider the second form where the operator moves the primed coordinates forward along the trajectory a time $t - t'$.

The inhomogeneous problem, Eq. (1), may now be solved by standard Green's-function techniques.[21] The equation for the adjoint Green's function is multiplied by f and integrated over r, p, and t. Also, Eq. (1) is multiplied by $\tilde{G}$ and integrated over r, p, and t. The two resulting equations are subtracted and use is made of Green's theorem in its general form. The integrals over the p- and r-surfaces vanish, the $t = \infty$ contribution is zero because of the Heaviside function, and only the $t = 0$ contribution from the bilinear concomitant remains. Use is made of the symmetry:

$$\tilde{G}(r, p, t \mid r', p', t') = G(r', p', t' \mid r, p, t)$$

and the primed and unprimed variables may be interchanged. When the delta functions are integrated over and the Heaviside function used to restrict the range of the t'-integration, the expression, Eq. (3), is obtained for f.[22]

[21] P. M. Morse, and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill Book Company, New York, 1953), Sec. 7.5.

[22] Similar integral equations have been derived before, e.g., Hollinger and Curtiss, Ref. 8.

Variational Principle for the Time-Dependent Schrödinger Equation*

MORTON E. GURTIN
Division of Applied Mathematics, Brown University, Providence, Rhode Island
(Received 7 January 1965)

In this paper a variational principle is established which fully characterizes the initial-value problem associated with the time-dependent Schrödinger equation.

INTRODUCTION

In a previous paper[1] variational principles were established for the initial-value problems associated with the heat conduction and wave operators. Subsequently, Gurtin and Leitman[2] demonstrated the usefulness of these principles in obtaining approximate solutions by direct variational methods.

In this note I show that the initial-value problem for the time-dependent Schödinger equation can also be characterized by means of a variational principle.

In the interest of brevity I omit all smoothness hypotheses and use the formal notation of the calculus of variations.

THE INITIAL-VALUE PROBLEM

Let R denote a region in n-dimensional Euclidean space, $x = (x_1, x_2, \cdots, x_n)$ a point of R, B the boundary of R, $t \geq 0$ the time,

$$\nabla^2 = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2}$$

the Laplacian operator, and

$$\nabla = (\partial/\partial x_1, \partial/\partial x_2, \cdots, \partial/\partial x_n)$$

the gradient operator. We state the initial-value problem for Schrödinger's equation as follows: assume that we are *given* a constant $k \neq 0$, an energy function $V(x)$ on R, initial data $\varphi_0(x)$ on R, and boundary data $\hat{\varphi}(x, t)$ on $B \times (0, \infty)$; we are to *find* a function $\varphi(x, t)$ on $R \times [0, \infty)$ which satisfies *Schrödinger's equation*

$$i\,\partial\varphi/\partial t = -k\nabla^2\varphi + V\varphi, \tag{1}$$

the *initial condition*

$$\varphi(x, 0) = \varphi_0(x) \quad \text{on} \quad R, \tag{2}$$

and the *boundary condition*

$$\varphi(x, t) = \hat{\varphi}(x, t) \quad \text{on} \quad B \times (0, \infty). \tag{3}$$

Our variational principle is based on a reformulation of this problem as a boundary-value problem for a certain integrodifferential equation. With this in mind we integrate (1) with respect to time, use (2), and arrive at the relation

$$i[\varphi - \varphi_0] = -k \int_0^t \nabla^2\varphi\, ds + V \int_0^t \varphi\, ds. \tag{4}$$

If we define the convolution $f * g(x, t)$ of two functions $f(x, t)$ and $g(x, t)$ in the usual manner:

$$f * g(x, t) = \int_0^t f(x, t - s)g(x, s)\, ds, \tag{5}$$

and notice that, since V is a function of x alone and k is a constant (function),

$$\begin{aligned} k * g(x, t) &= k \int_0^t g(x, s)\, ds, \\ V * g(x, t) &= V(x) \int_0^t g(x, s)\, ds, \end{aligned} \tag{6}$$

we see that (4) can be written in the following form:

$$i[\varphi - \varphi_0] = -k * \nabla^2\varphi + V * \varphi. \tag{7}$$

It is a trivial matter to verify that φ *satisfies* (1) *and* (2) *if and only if* φ *satisfies* (7).

THE VARIATIONAL PRINCIPLE

Before we state our variational principle it is convenient to define the convolution $\nabla f * \nabla g(x, t)$ of the vector-valued functions $\nabla f(x, t)$ and $\nabla g(x, t)$ as follows:

$$\nabla f * \nabla g = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} * \frac{\partial g}{\partial x_i}. \tag{8}$$

Theorem. Let K denote the space of all functions $\varphi(x, t)$ on $R \times [0, \infty)$ which satisfy the boundary

* These results were obtained in the course of an investigation sponsored by the Advanced Research Projects Agency, Department of Defence under ARPA Contract Sd-86 with Brown University.

[1] M. E. Gurtin, Quart. Appl. Math. 22, 3 (1964).

[2] M. E. Gurtin and M. J. Leitman, *"On the Use of Variational Principles for the Approximate Solution of Linear Initial-Value Problems"* Rept. No. C11-96, Contract Nonr 562(10), Brown University, December 1964.

condition (3). For each fixed time $t \geq 0$ let $\Omega_t(\cdot)$ be the functional defined on K by

$$\Omega_t(\varphi) = \int_R \cdot [(V * \varphi + 2i\varphi_0 - i\varphi) * \varphi + k * \nabla\varphi * \nabla\varphi](x, t)\, dx. \quad (9)$$

Then

$$\delta\Omega_t(\varphi) = 0 \quad \text{over } K \text{ for every fixed } \quad t \geq 0 \quad (10)$$

at a particular function φ if, and only if, φ satisfies Schrödinger's equation (1), the initial condition (2), and the boundary condition (3).

Proof: By the definition of K and the remark following Eq. (7) it suffices to show that (10) holds at a particular function φ in K if, and only if, φ satisfies (7). Suppose φ belongs to K. By (9) and the associativity and commutivity of the convolution

$$\delta\Omega_t(\varphi) = 2 \int_R [(V * \varphi + i\varphi_0 - i\varphi) * \delta\varphi + k * \nabla\varphi * \nabla\delta\varphi](x, t)\, dx \quad (11)$$

for $t \geq 0$. Since every φ in K satisfies the boundary condition (3), it follows that

$$\delta\varphi = 0 \quad \text{on} \quad B \times (0, \infty) \quad (12)$$

and hence, if we apply the divergence theorem to (11), we find that

$$\delta\Omega_t(\varphi) = -2 \int_R [i\varphi - i\varphi_0 + k * \nabla^2\varphi - V * \varphi] * \delta\varphi(x, t)\, dx \quad (13)$$

for every $t \geq 0$. From (13) it is obvious that if φ in K is a solution of (7), then φ satisfies (10). To establish the converse assertion we assume (10) holds. Then (13) implies

$$\delta\Omega_t(\varphi) = -2 \int_0^t \int_R \psi(x, t - s)\delta\varphi(x, s)\, dx\, ds,$$

$$\psi = i\varphi - i\varphi_0 + k * \nabla^2\varphi - V * \varphi \quad (14)$$

for every $t \geq 0$ and every variation $\delta\varphi$ which satisfies (12). But this fact and the "fundamental lemma" of the calculus of variations imply that φ must be a solution of (7). This completes the proof.

We remark that this theorem is also valid when the region R equals the entire n-dimensional Euclidean space; of course, in this instance, the boundary condition (3) is removed, but the functions φ in K, as well as the data φ_0 and V, are required to satisfy certain regularity conditions[3] as $|x| \to \infty$.

[3] Sufficient conditions are that φ_0, V, and the functions φ in K, together with their gradients $\nabla\varphi$, are all $0(|x|^{-2})$, as $|x| \to \infty$, uniformly on every closed time interval.

Hydromagnetic Switch-on Shocks*

LEONARD SARASON†
Courant Institute of Mathematical Sciences, New York University, New York, New York
(Received 14 December 1964)

In one-dimensional hydromagnetic flow, switch-on shocks have been shown to be nonevolutionary, e.g., by Syrovatskii. Such shocks, however, may be considered as a superposition of a switch-on shock and an Alfvén shock, and in this paper some results of such an assumption are examined. If the two shocks remain superimposed, the shock pair is, in fact, evolutionary. If the pair splits under the influence of incoming waves, a first-order nonlinear analysis indicates that highly unstable situations can occur. But there is a non-uniqueness implicit in the ansätz, and the author has not been able to exclude the possibility that a solution always exists which is stable except between shocks.

1. INTRODUCTION

THE one-dimensional time-dependent equations of magnetohydrodynamics may be represented in conservation form

$$u_t + f_x = 0, \tag{1.1}$$

where u is a vector and f a vector function of u. With $\mathfrak{A} = \partial f/\partial u = \text{grad } f$, (1.1) takes the quasilinear form

$$u_t + \mathfrak{A}(u)u_x = 0, \tag{1.2}$$

which is an equation of hyperbolic type, for $\mathfrak{A}$ has real eigenvalues and only simple eigenspaces.

Weak solutions of conservation systems (1.1) which are piecewise-continuous (i.e., solutions with shocks), satisfy across each shock the generalized Rankine–Hugoniot relations

$$s[u] = [f], \tag{1.3}$$

where [ ] denotes the jump across a shock, and s the shock speed.

Even when (1.1) is hyperbolic, such solutions are not in general uniquely determined by their initial values, and some further condition at the shock is needed if initial or mixed initial and boundary value problems are to be well posed. In compressible fluid flow, any of several conditions may be assumed, e.g., that entropy does not decrease across a shock, that the flow is supersonic ahead of and subsonic behind the shock, or the evolutionary condition described below. In magnetohydrodynamics, however, the above conditions are not equivalent, and the evolutionary condition is preferred, since the others may lead to unstable situations.[1]

The hydromagnetic switch-on shock, which is needed for the solution of certain piston problems,[2] is not evolutionary in the usual sense; this raises questions on the validity of a one-dimensional analysis, and motivates the present work.

The switch-on shock can be considered, however, as a superposition of a switch-on and an Alfvén shock which may separate under the influence of perturbing waves, the Alfvén shock falling slowly behind. In the region between them, Eq. (2.1) is assumed to hold.

Under small perturbations, the switch-on shock becomes an almost switch-on shock traveling faster than the Alfén shock or wave behind it. A perturbing Alfvén wave ahead rotating the transverse component of magnetic field through a given angle, is magnified by the switch-on shock into an Alfvén wave behind the shock with the same angular rotation. Thus a small perturbation of the state ahead of the almost switch-on shock may produce a large and steep change in the region between shocks.

If the perturbation is a pure Alfvén wave, a nonlinear analysis shows that there is no resulting perturbation behind the "magnified" wave. However, the shock-wave–shock configuration may now be unstable to non-Alfvénic disturbances.

If in the remaining region the solution, e.g., of the Cauchy problem depends continuously on the data, we shall consider the switch-on shock to be evolutionary in a generalized sense. In the remainder of this paper we shall propose and investigate a plau-

* The work presented here was done at the Courant Institute of Mathematical Sciences, New York University, under contract AT(30–1)1480 with the U. S. Atomic Energy Commission.
† Present address: Department of Mathematics, Stanford University, Stanford, California.

[1] A. I. Akhiezer, G. Ia. Liubarskii, and R. V. Polovin, Zh. Eksperim. i Teor. Fiz. **35**, 731 (1958) [English transl.: Soviet Phys.—JETP **8**, 507 (1959)].
[2] J. Bazer, Astrophys. J. **128**, 686 (1958).

sible necessary condition for evolutionarity in this sense of the switch-on shock.

The routine computation of the coefficients listed in Sec. 9 and in Appendix 1 has been omitted.

2. THE EVOLUTIONARY CONDITION FOR A GENERAL SYSTEM AND FOR A SWITCH-ON SHOCK

Let u_0, u_1 be two constant states connected by a shock. The system has first-order stability [satisfies the *generalized entropy condition* (Lax[3]), or the *evolutionary condition* (Polovin[4])] if the following condition obtains:

Linearize Eq. (1.1) on either side of the shock about the states u_0, u_1, and linearize the Rankine–Hugoniot relation (1.3) about the given jump and shock speed. The resulting system should be uniquely solvable for arbitrary initial conditions, and the solution should be bounded in terms of the initial conditions.

The linearized equations are given by

$$w_t^i + \mathcal{A}_i w_x^i = 0, \qquad i = 0, 1, \tag{2.1}$$

where w^i represents a small perturbation on u_i, and $\mathcal{A}_i = \mathcal{A}(u_i)$. Let Q_i^{-1} be the matrix of (right) eigenvectors of $\mathcal{A}_i$, so that

$$Q_i \mathcal{A}_i Q_i^{-1} = \Lambda_i \tag{2.1a}$$

is the diagonal matrix of eigenvalues $\lambda_i^{(j)}$ of $\mathcal{A}_i$. The change of variables $a_i = Q_i w^i$ reduces (2.1) to characteristic form

$$a_{i_t} + \Lambda_i a_{i_x} = 0, \tag{2.2}$$

or

$$\partial a_i^{(j)} / \partial s_i^{(j)} = 0, \tag{2.2}'$$

where $a_i = (a_i^{(1)}, a_i^{(2)}, \cdots)$, and $\partial/\partial s_i^{(j)}$ is the directional derivative $\partial/\partial t + \lambda_i^{(j)} \partial/\partial x$. Equation (2.2)′ represents a wave traveling to the left or right according as $\lambda_1^{(j)}$ is > 0 or < 0.

Without loss of generality, we assume the independent variable to be transformed so that the unperturbed shock speed vanishes; this leads to a simplification of the resulting formulas.

With a dot representing differentiation with respect to a small parameter ϵ, we find from (1.3), using $s = 0$ at $\epsilon = 0$,

$$\dot{s}[u] = [\dot{f}] = [\mathcal{A}\dot{u}], \quad \text{or} \quad \dot{s}[u] = \mathcal{A}_1 \dot{u}_1 - \mathcal{A}_0 \dot{u}_0. \tag{2.3}$$

Making the change of variable $u_i = Q_i a_i$, (2.3) implies

$$s[u] = Q_1^{-1} \Lambda_1 a_1 - Q_0^{-1} \Lambda_0 a_0. \tag{2.4}$$

We take (as subscripts) the state 0 to the right of the shock, the state 1 to the left.

Then the a_1^i corresponding to positive (negative) λ_1^i and the a_0^i corresponding to negative (positive) λ_0^i represent incoming (outgoing) waves at the shock. The incoming waves are determined by the initial conditions, and the shock is evolutionary if the outgoing a^i and $\dot{s}$ may be chosen uniquely to satisfy (2.4) for each choice of incoming wave. A sufficient condition is that the column vectors of Q_1^{-1} and Q_0^{-1} corresponding to outgoing waves together with $[u]$ are linearly independent, and span the space generated by the columns of Q^{-1} and Q_0^{-1} corresponding to incoming waves. Clearly the number of outgoing waves must be less than the number of variables, for otherwise they form, together with $[u]$, a dependent set.

For shocks on the borderline of evolutionarity, the above condition may not be satisfied; one may, on approaching the borderline shock from evolutionary shocks (in a shock diagram) lose an outgoing wave, or the "outgoing" eigenvectors together with $[u]$ may become linearly dependent. The former case occurs when a switch-on shock is approached from a fast shock or weaker strength; the latter case occurs when a 180° Alfvén shock is approached by Alfvén shocks across which the magnetic field rotates through a smaller angle.

Our claim is that the failure of these shocks to satisfy the above sufficient condition for evolutionarity is not sufficient to damn them as nonevolutionary; because of their borderline nature, one must take some account at least of quadratic as well as of linear effects.

For instance, we should exclude those outgoing waves which perturb the shock to an adjacent nonevolutionary shock; it may be assumed that such adjacent shocks will immediately, under the effects of second-order perturbations, resolve into the given borderline shock, plus small outgoing disturbances, and that the excluded outgoing waves will, therefore, not reach first-order strength.

If the above exclusion is accepted, 180° Alfvén shocks are in fact evolutionary. For the 180° Alfvén shock is connected to a nonevolutionary slow shock. Hence the condition that the shock remain Alfvén ($[B^2] = 0$) will provide an extra linear condition

$$[B_y \delta B_y + B_z \delta B_z] = 0 \tag{2.5}$$

sufficient to determine the outgoing waves uniquely. If all unperturbed components of B, u lie in an x–y

[3] P. Lax, Commun. Pure Appl. Math. **10**, 537 (1957).
[4] R. V. Polovin, Zh. Eksperim. i Teor. Fiz. **39**, 1005 (1960) [English transl.: Soviet Phys.—JETP **12**, 699 (1961)].

plane, and if the incoming waves lie in the same plane, we may replace (2.5) by the simpler restriction

$$\delta B_y(2) + \delta B_y(1) = 0. \tag{2.5a}$$

It should be noted that as the angle through which the magnetic field rotates across an Alfvén shock approaches $\pm\pi$, the ratio of the sizes of outgoing to incoming waves is uniformly bounded.

In the case of a switch-on shock, there are not enough outgoing waves, and the above "exclusion" does not help. One may, however, consider the switch-on shock as a superposition of switch-on and Alfvén shocks, with a different free parameter associated with each shock speed. This leads to a "linearized" resolution of incoming waves that would seem to indicate stability. However, if the shocks have different speed they cannot remain superimposed, and the differential equation must be satisfied in the region between shocks. We describe now the specific assumptions and simplifications made.

Let a switch-on shock separate two constant states, u_0 ahead of and u_2 behind the shock. The shock is assumed to split as described above into two shocks. Ahead of the switch-on and behind the Alfvén shock, the equation is linearized about u_0 and u_2 respectively, and the linearized jump conditions are assumed for each shock.

The Alfvén shock may be replaced by a combination of an Alfvén shock and a (steep) Alfvén wave between the shocks, conveniently chosen. We shall say that such a system is evolutionary if for sufficiently small disturbances impinging on the shocks from the regions "0" and "2", the outgoing disturbances in "0" and "2" depend continuously on them in an appropriate sense, which may involve comparing states at neighboring rather than at indentical points.

We abandon immediately the attempt to settle the evolutionarity of such systems, and pose a simpler question, the answer to which may indicate how things go. Suppose at a given time the state between shocks is to lowest order an Alfvén wave. In the region between shocks, the equation is linearized about this wave as carried out in Sec. 7 (we shall see that because of the nonlinearity of the geometry, this procedure is justified only if the perturbations considered are of second or higher order in the reciprocal of the gradients in the wave). We ask if a small perturbation in the state between shocks will tend to decay. If so, we shall call the system weakly evolutionary.

In Sec. 7, it is shown that if the Alfvén wave and shock each have small amplitude, the system is weakly evolutionary. In Sec. 10, it is shown that if the Alfvén shock is sufficiently large and the switch-on shock sufficiently small, the constant state between shocks is unstable and such a system is not weakly evolutionary (this does not imply that the system is not evolutionary in the above sense).

The latter result suggests that an appropriate question to ask is this: given the state u_2 and the state just behind the switch-on shock (determined to lowest order by the incoming waves ahead of the shock), can they be connected in such a way that the whole system is weakly evolutionary? A negative answer would provide strong grounds for believing that one-dimensional switch-on shocks do not exist; a positive answer would be a indication in their favor.

We have not settled the last question. It can probably be done by numerical experiment if not by analysis, and we hereby bequeath the problem to the interested reader.

3. THE SWITCH-ON SHOCK

The vector u occurring in (1.1) in one-dimensional magnetohydrodynamics is given as

$$\{B_y, B_z, \rho u_x, \rho u_y, \rho u_z, \rho, \epsilon\},$$

where B_y, B_z are components of magnetic field, ρ is the density, u_x, u_y, u_z are the components of fluid velocity $\mathbf{u}$, $\epsilon = \frac{1}{2}\rho\mathbf{u}^2 + \rho e + B^2/2\mu$ is the energy density (e is internal energy). The magnetic component B_x is constant and is treated as a parameter. We will denote the magnitudes of the transverse parts of B and u by B_t and u_t, i.e., $B_t = (B_y^2 + B_z^2)^{\frac{1}{2}}$, and $u_t = (u_y^2 + u_z^2)^{\frac{1}{2}}$. For convenience we shall assume the fluid flow is in the negative x-direction ($u_x < 0$). For the vector f we refer the reader to Bohachevsky.[5]

If, in the state 0, $B_y = B_z = 0$, and in the state 1, $B_y^2 + B_z^2 > 0$, the shock is called a "switch-on" shock. Such shocks present a more complicated picture not only because the state 0 has a multiple eigenvalue (the fast Alfvén and speeds here coincide), but because the Alfvén speed after the shock equals the shock speed.

If the sound speed a_0 in the state 0 is greater than the Alfvén speed $A_0 = |B_x|^{\frac{1}{2}}(\mu\rho)^{-\frac{1}{2}}$, a small switch-on shock cannot be evolutionary, for there are too many outgoing waves. Further, its place in helping to resolve an initial discontinuity is taken by a centered switch-on simple wave. (The vector $a = Qw$, and

[5] I. Bohachevsky, Phys. Fluids 5, 1456 (1962).

the gas-dynamical sound speed a will be distinguishable by context).

In the case $A_0 > a_0$, the switch-on shock is not evolutionary according to the criterion of Sec. 2, as there are not enough outgoing waves (the shock travels at the Alfvén speed). If an initial value problem with an initial switch-on shock and small incoming waves is to be well posed, there must be some stability relative to these disturbances, and, in at least a very limited way, there is. For, the unperturbed switch-on shock may be considered as a switch-on shock in an undetermined direction followed by an Alfvén shock. The effect of an incoming wave from the right (ahead of the shock), and containing a transverse magnetic field, is to determine uniquely the state immediately behind the switch-on shock. The switch-on shock is then perturbed into a fast (nearly switch-on) shock, which is closely followed by an Alfvén shock or wave-and-shock combination which "rotates" the state to one neighboring the unperturbed state on the left, which we call u_2.

A small variation in the incoming wave from the right may change the direction associated with the unperturbed "switch-on" shock, and hence may produce a gross change in the state immediately behind the fast shock. However, this is rectified to the left by the Alfvén wave and/or shock, so that outside of the thin region between the shocks, the resulting perturbations may be small (there are no outgoing waves to the right).

Considering the fast-shock–Alfvén wave-and-shock pair as a kind of "compound" shock, reflection and transmission coefficients may be computed, but they depend on the direction of switch on, and hence there is a nonlinear dependence of the (linearized) perturbed state on one parameter of the unperturbed state.

4. DIRECTION OF SWITCHING ON

The direction of the magnetic field after a switch-on shock is determined by the transverse components of magnetic field and fluid velocity in the incoming waves from the right (ahead of the shock). For simplicity and without loss of generality, we will assume from this point that, in the unperturbed state ahead of the switch-on shock, $u_y = u_z = 0$.

We index the left side of the shock as 1, the right side as 0. To the state u_2 we associate an angle θ given by $\theta = \tan^{-1} B_y/B_x$. With the strength of the shock considered fixed (to zero order) $u_1 = u_1(\theta)$, $\mathfrak{A}_1 = \mathfrak{A}_1(\theta)$, and $Q_1 = Q_1(\theta)$.

Corresponding to Eq. (2.4) we have, with s the perturbed shock speed,

$$s[u] = Q_1^{-1}\Lambda_1 a_1 - Q_0^{-1}\Lambda_0 a_0. \qquad (2.4)'$$

Multiplying both sides of (2.4) by Q_1, and noting that $Q_1[u]$ is a constant independent of θ, which we call $\hat{a}$, we find

$$\Lambda_1 a_2 = s\hat{a} + Q_1(\theta)Q_0^{-1}\Lambda_0 a_0. \qquad (4.1)$$

Let the first component a_{1_1} of a_1 represent the Alfvén wave traveling at the shock speed (i.e., the corresponding eigenvalue $\lambda_1 = 0$). Note that $a_1 = 0$, for the jump in a switch-on shock is a linear combination of eigenvectors of $\mathfrak{A}_1$ lying in the plane of the switched-on state. Hence the first component of Eq. (4.1) can be satisfied only if $Q_1(\theta)Q_0^{-1}\Lambda_0 a_0$ has vanishing first component. The equation $(Q_1(\theta)Q_0^{-1}\Lambda_0 a_0)_1 = 0$ takes the form

$$\alpha_1 \cos\theta + \alpha_2 \sin\theta = 0, \qquad (4.2)$$

which will, in general, have two solutions, θ_1 and $\theta_1 + \pi$, with $0 \leq \theta_1 < \pi$.

Since the system with velocity and magnetic field components restricted to a plane satisfies the evolutionary condition, and because of the Galilean invariance of the equations, it is clear that the angle θ is determined modulo π by the angle $\tau = \tan^{-1} B_z/B_y$ associated with the incoming perturbation to the right.

A unique choice of θ is made by requiring that the fast shock connecting the perturbed states be itself evolutionary, i.e., that the Alfvén wave which traveled originally at the shock speed become outgoing. Since at the switch-on shock, $\dot{s} + \dot{A}_1 + \dot{u}_x \neq 0$ provided $B_y^2 + B_z^2 \neq 0$ in the incoming wave from the right, θ will in general be uniquely determined by the condition

$$\dot{s} + \dot{A}_1 + \dot{u}_x > 0. \qquad (4.3)$$

5. DECOMPOSITION OF THE SWITCH-ON SHOCK

We propose the following resolution of a switch-on shock impinged on by small waves:

At each instant the incoming waves from the right will determine an angle θ through Eqs. (4.2) and (4.3). A fast shock S_1 will connect the perturbed state to the right with a state 2 perturbed from $u_2(\theta)$.

Initially, the state 1 will be connected through an Alfvén shock S_2 with a state perturbed from state 2 (to the left, and associated with an angle θ_2). However, the fast shock speed will always exceed the Alfvén speed behind the shock by an amount which is to first order proportional to $B_t = (B_y^2 + B_z^2)^{\frac{1}{2}}$ in the incoming wave. Hence the two shocks will tend to separate and there will develop a region

between in which the differential equation must be satisfied.

We make the simplifying assumption that the state between the shocks is at any one moment to lowest order an exact Alfvén wave. (This assumption is probably necessary for stability). Moreover, if the incoming disturbance is all from the right, and is chosen so that the state on the right connects by a fast wave to some unperturbed $u(\theta)$, and if this incoming wave changes only by rotating, then the ansatz proposed is in fact an exact solution.

6. THE SHOCK PAIR, COMPOSITE JUMP CONDITION

It is easy to derive a composite linearized jump condition for the shock pair treated as a unit, assuming the state between the shocks to be constant, as is justified if the shocks are superimposed, or if the incoming waves vary very slowly in time; we need only eliminate the state between the shocks.

Let s_1, s_2 be the perturbed fast shock and Alfvén shock speeds, and $[u]_1$, $[u]_2$ the unperturbed jumps at the fast and Alfvén shocks. The jump conditions are

$$s_1[u]_1 = -Q^{-1}\Lambda_1 a_1 + Q_0^{-1}\Lambda_0 a_0, \tag{6.1}$$

$$s_2[u]_2 = -Q_2^{-1}\Lambda_2 a_2 + Q_1^{-1}\Lambda_1 a_1. \tag{6.2}$$

Adding (6.1) and (6.2),

$$s_1[u]_1 + s_2[u]_2 = -Q_2^{-1}\Lambda_2 a_2 + Q_0^{-1}\Lambda_0 a_0. \tag{6.3}$$

The linear condition for evolutionarity of such a shock pair is that the space spanned by "outgoing" waves *from the pair* together with the two jumps, is linearly independent and spans the space of "incoming" waves. (Note that u_1 may be computed either from a_2 and s_2 or from a_0 and s_1.)

In case the Alfvén shock is a 180° shock, we must again impose (2.5) to ensure uniqueness.

7. THE NONLINEAR REGION; EQUATION AND BOUNDARY CONDITIONS

In the case of time-dependent perturbations, we attempt to expand the solution about the steady-state solution at each time t. In case the first-order solution is an exact Alfvén wave, the second-order term is bounded in terms of the incoming perturbations, at least for small time, and perhaps for larger time. If the incoming perturbation is a general one, it is not clear whether the problem is well posed. We consider only the simplest case, in which the perturbations are of second order in the thickness of the Alfvén wave.

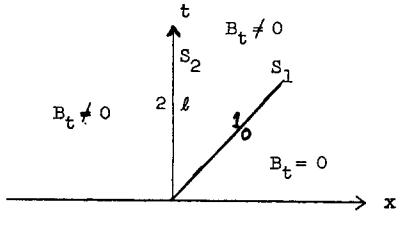


FIG. 1. Shock-splitting diagram.

We assume that an initial switch-on shock has split up as described above under some small incoming waves, which have died out, so that we now have switch-on (fast) shock followed by a slightly perturbed steep Alfvén wave, followed by an Alfvén shock. And we shall consider whether the perturbation in the Alfvén-wave region will tend to die out.

The states ahead of the switch-on shock, immediately behind it, and immediately ahead of and behind the Alfvén shock will be indexed by 0, 1, l, 2, respectively (see Fig. 1).

The region between the shocks is given to lowest order as an Alfvén wave, with coordinates chosen so that the Alfvén speed vanishes. We assign $u = u_0(\theta) + \epsilon^2 u^1$, $f \sim f_0 + \epsilon^2 a u^1$, $u^1 = Q^{-1}(\theta)a$, with initial conditions

$$\begin{aligned} \theta_0 &= \theta_0(x), \qquad x_0 < x < x_1, \\ a(\theta_0, 0) &= \bar{a}_0(\theta), \qquad a_{07} = 0, \end{aligned} \tag{7.1}$$

where a_{07} (the last component of a_0) corresponds to the nearly stationary Alfvén wave. The assumption that the initial disturbance is of second order in ϵ is consistent with the special assumption that there was an incoming switch-on shock of order ϵ from the right, which was followed by a rotation of this perturbed state; in this case we will have $\partial\theta_0/\partial x = p(x)/\epsilon$, where $p = \mathcal{O}(1)$, $p \not\equiv 0$.

Analysis of the more general case (incoming switch-on shock of order ϵ followed by incoming waves of order ϵ) is complicated by the nonlinearity of the geometry; a final answer to the question of stability must perhaps await such an analysis.

To account for the effect of higher-order terms on the geometry of the shock pair, we shall let $\theta = \theta(t, x)$ be a dependent variable, replacing a_{07} (we assume $a_{07} \equiv 0$). We shall assume further that $a_x = \mathcal{O}(\theta_x)$.

The unperturbed state behind the alfvén shock is given by $u = u_0(\theta_2)$, the state just behind the fast shock by $u_0(\theta_1)$, and that just ahead of the Alfvén shock (which depends on the solution) by $u = u_0(\theta_l)$, θ_l determined by the solution of the equation.

In writing the boundary condition for a, we take

advantage of the fact that only a single wave crosses the region between shocks from left to right, and represent the reflection matrices at the shocks (from the region between), simply as a vector.

Let $a = (a_1, 0, \cdots) + a'$, where a_1 corresponds to the fast wave moving to the right. With $\mathbf{S}$ a vector depending on $\theta_1 - \theta_l$, and $\mathbf{T}$ a fixed vector

$$\begin{aligned} a_1 &= a' \cdot \mathbf{S} \quad \text{ahead of the Alfvén shock,} \\ a' &= a_1 \mathbf{T} \quad \text{behind the fast shock.} \end{aligned} \tag{7.2}$$

As the unperturbed Alfvén wave is stationary, we have $f_{0_x} = 0$, and the equation $u_t + f_x = 0$ becomes, to the given order,

$$u_{0_t} + \epsilon^2 u_t' + \epsilon^2 (\mathfrak{A} u')_x = 0. \tag{7.3}$$

Substituting $\mathfrak{A} u' = Q^{-1} \Lambda a$ and multiplying (7.3) by Q, we find

$$Q u_{0_t} + \epsilon^2 Q (Q^{-1} a)_t + \epsilon^2 Q (Q^{-1} \Lambda a)_x = 0. \tag{7.4}$$

Since $u_0 = u_0(\theta)$, we have

$$u_{0_t} = \theta_t u_{0_\theta} = \theta_t c Q^{-1}(0, 0, \cdots, 0, 1), \tag{7.5}$$

where c is a constant and θ_t is to be found.

Since $Q_t^{-1} = \theta_t Q_\theta^{-1}$, $Q_x^{-1} = \theta_x Q_\theta^{-1}$, and $a_7 = 0$, the last row of (7.4) reads

$$\theta_t c + \epsilon^2 \theta_t (Q Q_\theta^{-1} a)_7 + \epsilon^2 \theta_x (Q Q_\theta^{-1} \Lambda a)_7 = 0. \tag{7.6}$$

To lowest order, then,

$$\theta_t = -(\epsilon^2 \theta_x / c)(Q Q_\theta^{-1} \Lambda a)_7, = \mathcal{O}(\epsilon) \tag{7.7}$$

and hence also $Q_t^{-1} = \mathcal{O}(\epsilon)$.

Since θ_t as given by (7.7) is independent of derivatives of a, and since an Alfvén wave may be connected to any state in which $B_y^2 + B_z^2 \neq 0$, $-\theta_t/\theta_x$ is in fact the perturbed Alfvén speed, which we denote simply by λ_7. Hence

$$\lambda_7 = (\epsilon^2/c)(Q Q_\theta^{-1} a)_7. \tag{7.8}$$

Let $a\hat{}$, $Q\hat{}$, $\Lambda\hat{}$, $\cdots\hat{}$, denote the restrictions of a, Q, Λ, $\cdots$, to the space spanned by the first six components of a. The leading terms in (7.7) give

$$a\hat{}_t + \Lambda\hat{} a\hat{}_x = -(Q Q_x^{-1} \Lambda)\hat{} a\hat{}. \tag{7.9}$$

It is convenient to replace the independent variable x by θ, i.e., t, x, by a pair T, θ with $t = T$. Then

$$\partial/\partial x = \theta_x \, \partial/\partial\theta, \qquad \partial/\partial t = \partial/\partial T + \theta_t \, \partial/\partial\theta.$$

Since $\theta_t = \mathcal{O}(\epsilon)$, we drop the term $\theta_t a\hat{}_\theta$, and find

$$a\hat{}_T + \theta_x \Lambda\hat{} a\hat{}_\theta = -\theta_x (Q Q_\theta^{-1} \Lambda)\hat{} a\hat{}. \tag{7.10}$$

We estimate θ_x as follows: Define $T_0(\theta)$ as the smallest $T \geq 0$ for which θ lies either on one of the shocks or between them. Then

$$x(T, \theta) = x(T_0(\theta), \theta) + \int_{T_0}^{T} \lambda_7(\theta, \tau) \, d\tau.$$

To lowest order,

$$x_\theta = x_\theta(T_0(\theta), \theta) + \int_{T_0}^{T} \frac{\partial \lambda_7}{\partial \theta}(\theta, \tau) \, d\tau,$$

for

$$[\partial x(T_0, \theta)/\partial T_0] \, \partial T_0/\partial\theta = \mathcal{O}(\epsilon).$$

Then

$$\frac{\partial}{\partial T} x_\theta = \frac{\partial \lambda^7}{\partial \theta} = \frac{\epsilon^2}{c} (Q Q_\theta^{-1} \Lambda a_\theta)_7.$$

Consequently

$$\frac{\partial}{\partial T} \theta_x = -\theta_x^2 \frac{\partial}{\partial T} x_\theta = -\theta_x^2 \frac{\epsilon^2}{c} (Q Q_\theta^{-1} \Lambda a_\theta)_7. \tag{7.11}$$

If $\theta_x = \mathcal{O}(1/\epsilon)$, then, according to (7.11), $\partial/\partial T \, \theta_x = \mathcal{O}(1)$. Hence, to lowest order,

$$\theta_x(T, \theta) = \theta_x(T_0, \theta) = \theta_x(\theta).$$

Since points $\theta =$ constant travel just at the Alfvén speed, it follows that to lowest order, θ_t (i.e., θ just ahead of the Alfvén shock) is constant. Hence, to the given order of approximation, the problem is defined by Eq. (7.10), with $\theta_x = \theta_x(\theta)$, boundary conditions (7.2), with $\mathbf{S}$, $\mathbf{T}$ fixed, and some initial values for $a\hat{}$.

Under the change of variable $T = \epsilon\tau$, a stretching of the time variable, (7.10) becomes

$$a\hat{}_\tau + \epsilon \theta_x \Lambda\hat{} a\hat{}_\theta = -\epsilon \theta_x (Q Q_\theta^{-1} \Lambda)\hat{} a\hat{}, \tag{7.12}$$

where now $\epsilon\theta_x = \mathcal{O}(1)$.

An explicit solution to (7.12) or to (7.10) with boundary condition given by (7.2), may be computed by taking a Laplace transform with respect to t or τ, and its properties may then be investigated (the result of Sec. 10 is derived with the help of transform techiques).

Here, we give a weak result, valid only for small Alfvén shock angles, and if the total variation of θ in the Alfvén wave is not too large.

In the following we shall write for $a\hat{}$ simply a.

Denote the Alfvén shock angle $\theta_1 - \theta_l$ by θ_s, and the size of the Alfvén wave $\theta_l - \theta_2$ by θ_A.

Theorem 7.1. There exists $\hat{\theta}_s > 0$ such that to each $|\theta_s| < \hat{\theta}_s$ there corresponds a $\hat{\theta}_A(|\theta_s|) > 0$, such that if $|\theta_A| < \hat{\theta}_A$, the solution to the initial-value problem given by (7.10), (7.12) with bounded initial

values decays exponentially fast in the maximum norm, with decay time of the order of ϵ.

To prove Theorem 7.1, we use the equation in the form (7.12). Let $\sup |\epsilon\theta_x| = p$. Define $\tau_1 = |\theta_A|\ p^{-1} \max_{i\neq 7} |\lambda_i^{-1}|$. Define the norms $|a|_\tau = \sup_\theta |a(\tau, \theta)|$, and $||a||_\tau = \sup_{\tau-\tau_1\leq t\leq\tau} |a|_\tau$, and define $M = p|(QQ_\theta^{-1}\Lambda)\hat{}|$.

Under suitable conditions, we shall derive an estimate of the form $|a|_\tau \leq C\ ||a||_\tau$, with $|C| < 1$; exponential decay of $|a|_\tau$ with respect to τ is an immediate consequence, and since $T = \epsilon\tau$, the decay time is $\mathcal{O}(\epsilon)$.

At the switch-on shock $a' = a_1\mathbf{T}$ [see (7.2)]. Hence $|a|^2 = a_1^2(1 + |\mathbf{T}|^2)$, and it follows that

$$|a_1(t, \theta)| \leq (1 + |\mathbf{T}|^2)^{-\frac{1}{2}}\ |a|_\tau.$$

For t such that $\tau - \tau_1 \leq t \leq \tau$, then

$$|a_i(t, \theta)| \leq |T_i|\ (1 + |\mathbf{T}|^2)^{-\frac{1}{2}}\ ||a||_\tau \quad i = 2, 3, \cdots. \tag{7.13}$$

Integrating the ith components of (7.12) ($i > 1$) along characteristics originating at the switch-on shock, we find

$$\sup_{\tau-\frac{1}{2}\tau_1\leq t\leq\tau} |a_i|_t \leq |T_i|\ (1 + |\mathbf{T}|^2)^{-\frac{1}{2}}\ ||a||_\tau + \tfrac{1}{3}(\tau_1)\ M\ ||a||_\tau. \tag{7.14}$$

The boundary condition (7.2), together with the inequality (7.14), implies

$$\sup_{\tau-\frac{1}{2}\tau_1\leq t\leq\tau} |a_1(t, \theta_1)| \leq \left\{\sum_{i>1} |S_i|\ |T_i|\ (1 + |\mathbf{T}|^2)^{-\frac{1}{2}} + \sum_{i>1} |S_i|\ \frac{\tau_1}{3}\ M\right\} ||a||_\tau. \tag{7.15}$$

Integration of the first component of (7.12) along appropriate characteristics originating at the Alfvén shock leads to the estimate

$$\sup_{\tau-\frac{1}{2}\tau_1\leq t\leq\tau} |a_1|_t \leq \left\{\sum_{i>1} |S_i|\ |T_i|\ (1 + |\mathbf{T}^2|)^{-\frac{1}{2}} + (1 + \sum |S_i|)\ \frac{\tau_1}{3}\ M\right\} ||a||_\tau. \tag{7.16}$$

In particular (7.16) applied at $\theta = \theta_1$, together with (7.2), and integration of appropriate components of (7.12) along characteristics originating at the switch-on shock leads to the estimate

$$|a_i|_\tau \leq \left[T_i\left\{\sum_{i>1} |S_i|\ |T_i|\ (1 + |\mathbf{T}|^2)^{-\frac{1}{2}} + (1 + \sum |S_i|)\ \frac{\tau_1}{3}\ M\right\} + \frac{\tau_1 M}{3}\right] ||a||_\tau. \tag{7.17}$$

Since $\tau_1 = \mathcal{O}\,|\theta_A|$, (7.16) applied to $t = \tau$, together with (7.17), implies

$$|a|_\tau \leq \left(\sum_{i>1} |S_iT_i| + \mathcal{O}\,|\theta_A|\right) ||a||_\tau.$$

Hence for fixed $\mathbf{S}$, $\mathbf{T}$, such that $\sum |S_iT_i| < 1$, θ_A sufficiently small implies

$$|a|_\tau/||a_\tau|| < c < 1.$$

It follows by an elementary argument that

$$||a||_{\tau+\tau_1} \leq c\ ||a||_\tau,$$

and hence

$$||a||_{\tau+r\tau_1} \leq c^{[r]}\ ||a||_\tau,$$

where $[r]$ is the integral part of τ, proving Theorem 7.1.

8. JUMPS ACROSS THE SWITCH-ON SHOCK AND ALFVÉN SHOCK

We recall that the state ahead of the switch-on shock is denoted by "0" and the state behind by "1." As mentioned above, in the case $a_0 > A_0$, the shock is unstable, and we shall consider only the case $A_0 > a_0$. We assume that behind the shock $u_x + A = 0$, i.e., there is a stationary Alfvén wave behind the shock.

With $\eta = \rho_1/\rho_0$, $\eta = \bar{\eta} - 1$, and $\sigma_0 = \mu\gamma P_0/B_0^2 = a_0^2/A_0^2$, the state behind the switch-on shock is given by

$$B_{t,1} = (2\bar{\eta})^{\frac{1}{2}}\{(1 - \sigma_0) - \tfrac{1}{2}(\gamma - 1)\bar{\eta}\}^{\frac{1}{2}}\ |B_0|, \tag{8.1a}$$

$$P_1 = \{1 + \gamma\bar{\eta}(1 - \tfrac{1}{2}(\gamma - 1)\bar{\eta}\sigma_0)\}P_0, \tag{8.1b}$$

$$u_{x,1} = u_{x,0} - |B_x|(\mu\rho_0)^{-\frac{1}{2}}(\eta^{\frac{1}{2}} - \eta^{-\frac{1}{2}}), \tag{8.1c}$$

$$u_{t,1} = |B_x|\ (\mu\rho_0)^{-\frac{1}{2}}\eta^{-\frac{1}{2}}(2\bar{\eta})^{\frac{1}{2}}\{1 - \sigma_0 - \tfrac{1}{2}(\gamma - 1)\bar{\eta}\}^{\frac{1}{2}}, \tag{8.1d}$$

where the transverse velocity $\mathbf{u}_t$ switches on in the direction similar to or opposite that of the magnetic field according as $B_x < 0$ or $B_x > 0$; and the shock speed s_1 is given by

$$s_1 = u_{x,0} - |B_x|\ (\mu\rho_1)^{-\frac{1}{2}}. \tag{8.1e}$$

[Relations (8.1) are slight rearrangements of formulas given by Bazer and Ericson.[6]]

The case of a small switch-on shock ($B_t \sim 0$) is of some interest, both because it gives insight into the more general case, and because such shocks may result from small disturbances. From this point, except where indicated, we shall confine our dis-

[6] J. Bazer and W. B. Ericson, Astrophys. J. 129, 758 (1959).

cussion to such small but finite switch-on shocks. It will be assumed that the perturbations are nonetheless small compared to the jumps across the switch-on shock.

The jump across a small switch-on shock is given, to lowest nonvanishing order in B_t, the transverse field after the shock, by

$$[u] \sim [B_t \cos\theta, B_t \sin\theta, 0, -\rho B_t(\mu\rho)^{-\frac{1}{2}} \cos\theta, -\rho B_t(\mu\rho)^{-\frac{1}{2}} \sin\theta, VB_t^2, WB_t^2], \tag{8.2}$$

where

$$V = [2\mu(A^2 - a^2)]^{-1}, \quad W = \tfrac{1}{2}V\{3A^2 - 4a^2 + 2a^2/(\gamma - 1)\}. \tag{8.3}$$

The formulas from this point on are greatly simplified if γ is taken, as usual, as $\frac{5}{3}$ and we shall therefore do so. With this value for γ, for instance, $W = 3A^2V/2$. Note that except for B_t and θ, the quantities in (8.2) and (8.3) may be evaluated on either side of the shock, to the given order of accuracy.

The jumps across the Alfvén shock depend on the angles θ_2, θ_1. Across the shock,

$$[B_t^2] = [u_x] = [p] = [\rho] = 0 \tag{8.4}$$

and

$$[u] = \{B_t(\cos\theta_2 - \cos\theta_1), \quad B_t(\sin\theta_2 - \sin\theta_1), 0, -B_t(\operatorname{sgn} B_x)(\cos\theta_2 - \cos\theta_1)/(\mu\rho)^{\frac{1}{2}}, -B_t(\operatorname{sgn} B_x)(\sin\theta_2 - \sin\theta_1)/(\mu\rho)^{\frac{1}{2}}, 0, 0\}.$$

A. Eigenvectors and Eigenvalues of $\mathfrak{A}$

For the sake of definiteness, we shall assume from this point that $B_x > 0$.

We give the eigenvectors of the matrix $\mathfrak{A}$ in a coordinate system chosen such that $B_t = (\mu\rho)^{\frac{1}{2}}u_t$. With $c_\pm$ the slow and fast characteristic speeds, and A the Alfvén speed, the eigenvectors of the matrix $\mathfrak{A}$ are ordered according to the following list of the associated disturbance speeds: $u_x + c_+$, $u_x - A$, $u_x - c_+$, u_x, $u_x + c_-$, $u_x - c_-$, $u_x + A$. Let $C = \cos\theta$, $S = \sin\theta$, $h_\pm = \mu\rho c_\pm^2 - B_x^2$. The eigenvectors of $\mathfrak{A}$ are given by

$$R_{2,7} = \{\pm S(\mu\rho)^{\frac{1}{2}}, \pm C(\mu\rho)^{\frac{1}{2}}, 0, -\rho S, \rho C, 0, 0\};$$

$$R_{1,3} = \{\mu c_+^2 CB_t, \mu c_+^2 SB_t, h_+(u_x \pm c_+), \pm c_+B_xB_tC + u_yh_+, \pm c_+B_xB_tS + u_zh_+, h_+, c_+B_t^2(c_+ \mp A) + h_+[2a^2 \pm c_+u_x]\};$$

$$R_4 = \{0, 0, u_x, u_y, u_z, 1, 0\}.$$

$R_{5,6}$ are given in the formulas for $R_{1,3}$ with c_+, h_+ replaced by c_-, h_-.

In order to compute the reflection coefficients at a small switch-on shock and in Alfvén shocks where $B_t \sim 0$, it is convenient to use a basis of eigenvectors which do not vanish as $B_t \downarrow 0$. The eigenvectors $R'_{2,4,7} = R_{2,4,7}$ are as given above; we replace $R_{1,3,5,6}$ by eigenvectors $R'_{1,3,5,6}$ given to second order in B_t by

$$R'_{1,3} = \left\{(\mu\rho)^{\frac{1}{2}}C, (\mu\rho)^{\frac{1}{2}}S, 0, C\left(\rho - \frac{3B_t^2}{2\mu(A^2 - a^2)}\right), S\left(\rho - \frac{3B_t^2}{2\mu(A^2 - a^2)}\right), \frac{(\rho/\mu)^{\frac{1}{2}}}{A^2 - a^2}, B_tE_\pm\right\},$$

where

$$E_+ = -\frac{(\rho/\mu)^{\frac{1}{2}}(A^2 - 2a^2)}{A^2 - a^2}, \qquad E_- = \frac{3(\rho/\mu)^{\frac{1}{2}}A^2}{A^2 - a^2},$$

and

$$R'_{5,6} = \left\{\frac{-a^2CB_t}{\rho(A^2 - a^2)}, -\frac{a^2SB_t}{\rho(A^2 - a^2)}, \pm a - A \mp \frac{aB_t^2}{2\mu\rho(A^2 - a^2)}, \left(-1 \mp \frac{aA}{A^2 - a^2}\right)\frac{CB_t}{(\mu\rho)^{\frac{1}{2}}}, \left(-1 \mp \frac{aA}{A^2 - a^2}\right)\frac{SB_t}{(\mu\rho)^{\frac{1}{2}}}, 1, 2a^2 \mp aA \pm \frac{a(3A \mp 2a)B_t^2}{2\mu\rho(A^2 - a^2)}\right\}.$$

The columns of the matrix Q^1 [partly defined by (2.1a)] are taken as the eigenvectors R'_t (if $B_t \sim 0$), or, more generally, R_t.

Assuming that in the state (1) $u_x = -A$, the characteristic speeds (eigenvalues of $\mathfrak{A}$) are, to second order in B_t,

$$JA, -2A, -(2 + J)A, -A, a - A - aJ, -a - A + aJ, 0,$$

where $J = B_t^2/2\mu\rho(A^2 - a^2)$. (To find the small disturbance speeds in a coordinate system traveling with the fluid, add A to each eigenvalue).

9. REFLECTION COEFFICIENTS

To lowest order in B_t, the reflection coefficients S_i, T_i at the Alfvén and switch-on shocks from the region between are given by

$$S_2 \sim -4\mu\rho S_\alpha \frac{(A^2 - a^2)}{C_\alpha B_t^2}, \quad S_3 \sim 1 - \frac{1}{C_\alpha}, \quad S_4 \sim \left(1 - \frac{1}{C_\alpha}\right)\left(\frac{\mu}{\rho}\right)^{\frac{1}{2}}\frac{(A^2 - a^2)}{B_t}, \quad S_5 \sim S_6 \sim \left(1 - \frac{1}{C_\alpha}\right)\left(\frac{\mu}{\rho}\right)^{\frac{1}{2}}\frac{(A^2 - a^2)}{B_t}, \tag{9.1}$$

where

$$C_\alpha = \cos(\theta_2 - \theta_1), \qquad S_\alpha = \sin(\theta_2 - \theta_1);$$

$$\begin{aligned} T_2 &\sim 0, \\ T_3 &\sim \frac{4(A^2 - a^2)}{9A^2}, \\ T_4 &\sim -\frac{1}{9}\left(\frac{\rho}{\mu}\right)^{\frac{1}{2}} B_t \frac{(4a^2 - 3A^2)}{a^2A^2}, \\ T_{5,6} &\sim -\frac{1}{6}\left(\frac{\rho}{\mu}\right)^{\frac{1}{2}} B_t \frac{(A \mp 4a)}{a^2(A \mp a)}. \end{aligned} \tag{9.2}$$

Here

$$\sum_{i=2}^{6} S_i T_i \sim 1 - \frac{1}{C_\alpha}. \tag{9.3}$$

It follows from (9.3) that (6.3) can, in fact, be solved for arbitrary $(\theta_2 - \theta_1)$ if B_t is sufficiently small. For, with a taken as constant in the state 1 between the shocks, the jump conditions (6.1), (6.2) imply the pair of equations

$$a_1 = k_1 + \sum_{i=2}^{6} a_i S_i, \tag{9.4a}$$

$$a_i = a_1 T_i + k_i \qquad (i > 1), \tag{9.4b}$$

where k_i are inhomogeneous terms depending on the incoming waves from outside the shock pair. Substitution of (9.4b) in (9.4a) yields

$$a_1 = \left(1 - \sum_{i=2}^{6} S_i T_i\right)^{-1}\left(k_1 + \sum_{i=2}^{6} S_i k_i\right),$$

which is uniquely solvable provided $1 - \sum S_i T_i \neq 0$. But $\lim_{B_t \to 0} 1 - \sum S_i T_i = C_\alpha^{-1}$, by (9.3). Since $C_\alpha^{-1} \neq 0$, a_1 is uniquely given. Now a_i $(i = 2, \cdots, 6)$ is determined by (9.4b), and the outgoing waves in the state (2) are uniquely determined by (6.2) and (2.5).

As regards the applicability of Theorem 7.1, we observe that if $4/3a^3 \leq A^2 \leq 16a^2$, then

$$\sum |S_i T_i| \sim |1 - 1/C_\alpha|,$$

so that in this case

$\lim_{B_t \downarrow 0} \sum |S_i T_i| < 1$ if and only if $|\theta_2 - \theta_1| < \frac{1}{3}\pi$.

10. AN UNSTABLE CASE

According to the analysis in Sec. 7, there are two conditions under which stability of the system (7.10) with boundary condition (7.2) and given initial condition is hard to prove; firstly, when the Alfvén wave is large (i.e., $|\theta_1 - \theta_l|$ large), and secondly, when $\sum |S_i T_i| \geq 1$. To study the latter problem we shall consider a situation in which the unperturbed state between the two shocks is constant $(\theta \equiv \theta_l \equiv \theta_1)$. We assume the incoming waves are of order ϵ^2, where ϵ is the thickness of the wave, so that the geometry of the setup may be considered as given. Let the unperturbed switch-on shock at $x = 0$ and the unperturbed Alfvén shock at $x = -\epsilon$ be steady.

The linearized equation and boundary conditions corresponding to (7.10), (7.2) are

$$a_t + \Lambda a_x = 0, \qquad -\epsilon < x < 0, \tag{10.1}$$

$$\begin{aligned} a_1 &= a'\mathbf{S} \quad \text{at} \quad x = -\epsilon, \\ a' &= a_1\mathbf{T} \quad \text{at} \quad x = 0, \end{aligned} \tag{10.2}$$

and the initial conditions are given by $a(0, x) = a_0(x)$ $(-\epsilon < x < 0)$.

Since the quantities a_i are propagated as invariants along characteristics, it is easy to convert the initial-value problem to a boundary-value problem with data of bounded support in t, by subtracting off the solution of (10.1) with the given initial conditions in $-\epsilon < x < 0$, vanishing initial conditions elsewhere in x, and no boundary conditions.

The boundary-value problem is explicitly solved through use of the Laplace transform. With $\hat{a} = \int_0^\infty e^{-t\tau} a(\tau) d\tau$, the equation takes the form

$$\hat{a}_x = -\Lambda^{-1}\tau\hat{a}, \tag{10.3}$$

with boundary condition

$$\hat{a}_1(-\epsilon) = \hat{a}' \cdot \mathbf{S} + k_1 \quad \text{at} \quad x = -\epsilon, \tag{10.4a}$$

$$\hat{a}'(0) = \hat{a}_1(0)\mathbf{T} = k' \quad \text{at} \quad x = 0, \tag{10.4b}$$

where (k_1, k') is the Laplace transform of a vector having finite support in t, $\mathbf{S} = (0, S_1, S_2, \cdots)$ and $\mathbf{T} = (0, T_1, T_2, \cdots)$.

With $\hat{S} = (-1, S_1, S_2, \cdots)$ and $\hat{T}) = (1, T_1, T_2, \cdots)$, the solution of the boundary-value problem (10.4) is given by

$$\hat{a}_1(0) = (\hat{S}e^{\Lambda^{-1}\tau\epsilon}\hat{T})^{-1}\{k_1 - \hat{S}e^{\Lambda^{-1}\tau\epsilon}k'\}, \tag{10.5}$$

together with (10.4b) and the equation

$$\hat{a}(x) = e^{-\Lambda^{-1}\tau x}\hat{a}(0).$$

Suppose for simplicity that $k' = 0$, so that

$$\hat{a}_1(0) = (\hat{S}e^{\Lambda^{-1}\tau\epsilon}\hat{T})^{-1}k_1. \tag{10.6}$$

Lemma 10.1. Suppose $|\theta_s| = |\theta_2 - \theta_1| > \frac{1}{2}\pi$. Then there exists $B_t^0 = B_t^0(\theta_\alpha) > 0$ such that to each $B_t < B_t^0$, there exists $\tau > 0$ such that $\hat{S}e^{\Lambda^{-1}\tau\epsilon}T = 0$.

Proof. If $\tau = 0$, $B_t \downarrow 0$, $\hat{S} \cdot \hat{T} \to -C_\alpha^{-1} > 0$. As $\tau \to \infty$, independently of B_t,

$$\hat{S}e^{\Lambda^{-1}\tau\epsilon}\hat{T} \sim -e^{\tau\epsilon\lambda_1^{-1}},$$

which approaches $-\infty$. By continuity, the lemma holds with $B_t^0 = \text{inf. } \{B_t \mid \hat{S}\cdot\hat{T} = 0\}$ (here B_t^0 may perhaps be infinite).

Theorem 10.1. If $|\theta_\alpha| > \frac{1}{2}\pi$, there exists $B_t^0(\theta_\alpha)$ such that if $B_t > B_t^0$, the solution of the boundary-value problem (10.3), (10.4) increases exponentially as $t \to \infty$ for almost all data.

Proof. The solution for $a_1(t, 0)$ is given by the inverse transform of (10.6). Since the zeros of $\hat{S}e^{A^{-1}\tau}\hat{T}$ have some maximum value R of Re τ, the inverse transform will represent the solution only if the path of integration is to the right of Re $\tau = R$. It follows that $a_1(t, 0)$ will grow faster than e^{ct} for all $c < R$, unless k_1 has a zero at each zero of $\hat{S}e^{A^{-1}\tau}\hat{T}$ on the line Re $\tau = R$ (note that R is proportional to ϵ^{-1}).

DISCUSSION

From the results of Secs. 9 and 10, it would appear that our ansätz leads in general to a configuration which is not weakly evolutionary. There is a possible remedy, however: it must be too much to ask that θ_t be constant to lowest order. If with given θ_1, θ_2, our setup is weakly evolutionary for at least one choice of θ_t and Alfvén wave, we may conjecture some dissipative or nonlinear mechanism under which the system will find such an appropriate state. Note here that even initially if $|\theta_1 - \theta_2| > \frac{1}{2}\pi$, we cannot have a stable constant state between shocks, because of Theorem 10.1.

Another conjecture is that the switch-on shock breaks up into a combination of other shocks which together resolve the initial discontinuity found across the switch-on shock. If true, it could mean that not all small initial discontinuities can be resolved by small shocks and centered waves, for the discontinuity across a small switch-on shock cannot be so resolved. In this case the hydromagnetic "piston problem" in which a magnetized piston is given a transverse impulse is not properly posed, for Bazer[1] required switch-on shocks to solve the problem.

A possible alternative conjecture is that through dissipative or higher-order nonlinear effects, the fast and Alfvén shocks will tend to coalesce and that the outgoing disturbances from the shock will remain small. In this case the composite jump condition (6.3) might yield a good approximation.

ACKNOWLEDGMENTS

I wish to thank Professors J. Bazer, H. Grad and H. Weitzner for many helpful and stimulating conversations.

APPENDIX 1

The matrix $QQ_\theta^{-1}\Lambda$ occurring in (7.9) *et seq.* is given to lowest order in B_t, with terms of higher-than-second order omitted, by

$$QQ_\theta^{-1}\Lambda = \begin{Vmatrix} 0 & 2A & 0 & 0 & 0 & 0 \\ AJ & 0 & 3AJ & \frac{AB_t}{2\rho(\mu\rho)^{\frac{1}{2}}} & \frac{AB_t}{2\rho(\mu\rho)^{\frac{1}{2}}} & \frac{AB_t}{2\rho(\mu\rho)^{\frac{1}{2}}} \\ 0 & -3AJ & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\left(\frac{\rho}{\mu}\right)^{\frac{1}{2}}\frac{AB_t}{a(A-a)} & 0 & 0 & 0 & 0 \\ 0 & \left(\frac{\rho}{\mu}\right)^{\frac{1}{2}}\frac{AB_t}{a(A+a)} & 0 & 0 & 0 & 0 \end{Vmatrix}.$$

Thus, in the steep Alfvén wave, the non-Alfvén disturbances interact with the leftward-moving Alfvén disturbance but not with each other.

APPENDIX 2

We shall derive two properties of switch-on shocks from general principles. We denote them by P_1, P_2. In particular,

P_1: The shock speed coincides with the Alfvén speed behind the shock.

P_2: To lowest order in the shock strength, the shock speed lies one third of the way between the fast disturbance speeds ahead of and behind the shock.

Proof of P_1. Let S_1 and S_2 be two switch-on shocks having ahead of the shock the same state u_0, in

which $u_y = u_z = 0$. Denote the states after the shock by u_1 and u_2 and suppose that u_1 and u_2 differ only in the orientation of the transverse components of magnetic field and fluid velocity. Denote by θ_i the quantity $\tan^{-1} B_z/B_y$ in u_i ($i = 1, 2$).

Subtracting the Rankine–Hugoniot relation for one shock from that for the other, we find that u_1 and u_2 may be connected by a shock traveling at the same speed s as S_1 and S_2. Letting $|\theta_2 - \theta_1| \to 0$, we find that s is in fact a characteristic speed. Further, s is the Alfvén speed, for the Alfvén speed is the only shock speed which is independent of the shock strength.

The proof of P_2 is a simple extension of the proof by Lax[3] that to first order, the speed of a shock is the arithmetic mean of the associated characteristic speeds before and after the shock.

Let the state before the shock be fixed, and denote it by u_0. In the state u_1 after the shock, $\theta = \tan^{-1}(B_z/B_y)$ is fixed, and u_1 depends on a single parameter ϵ of shock strength, i.e., $u_1 = u(\epsilon)$, such that $u(0) = u_0$.

For the fast disturbance speed λ and the matrix $\mathfrak{A} = \partial f/\partial u$ for the state u will also be functions of ϵ, as will the shock speed S. For convenience, coordinates are chosen such that $\lambda(0) = 0$. It will be assumed as known that at $\epsilon = 0$ and provided $0 < |\dot{u}| < \infty$, $s = \dot{s} = \dot{\lambda} = 0$, and $\ddot{s} \neq 0$. (Here the dot signifies $d/d\epsilon$.

Let r be the right eigenvector of $\mathfrak{A}$ such that

$$(\mathfrak{A} - \lambda)r = 0. \tag{A1}$$

Property P_2 is proved by comparing (A1) and its first two derivatives with respect to ϵ with the first three derivatives of the Rankine–Hugoniot condition

$$s(u - u_0) = f - f_0, \tag{A2}$$

all evaluated at $\epsilon = 0$. (Except for the fact that in each case we consider one more derivative, our procedure is identical with Lax's.[3]) In the following, all quantities are evaluated at $\epsilon = 0$.

The derivatives of (A2) yield, after substitution of 0 for λ, $\dot{\lambda}$, s, and $\dot{s}$.

$$\mathfrak{A}\dot{u} = 0, \tag{A3}$$

$$\mathfrak{A}\ddot{u} + \dot{\mathfrak{A}}\dot{u} = 0, \tag{A4}$$

$$\mathfrak{A}\frac{d^3u}{d\epsilon^3} + 2\frac{d\mathfrak{A}}{dt}\frac{d^2u}{d\epsilon^2} + \frac{d^2\mathfrak{A}}{d\epsilon^2}\frac{du}{d\epsilon} = 3\frac{d^2s}{d\epsilon^2}\frac{du}{d\epsilon}. \tag{A5}$$

(A.1) and its derivatives yield

$$Ar = 0, \tag{A6}$$

$$\dot{A}r + A\dot{r} = 0, \tag{A7}$$

$$(\ddot{A} - \lambda)r + 2\dot{A}\dot{r} + A\ddot{r} = 0. \tag{A8}$$

(A.3) implies that $\dot{u}$ is proportional to r. By multiplying ϵ by a fixed constant we can achieve

$$\dot{u} = r. \tag{A9}$$

Substituting r for $\dot{u}$ in (A4) and subtracting the result from (A7), we find

$$\mathfrak{A}(\ddot{u} - \dot{r}) = 0. \tag{A10}$$

Hence proper choice of the parameter ϵ ensures that

$$\ddot{u} = r. \tag{A11}$$

After substituting $\dot{r}$ for $\ddot{u}$ in (A5), we multiply (A5) and (A8) by the left null vector of $\mathfrak{A}$, and subtract one of the resulting equations from the other. There results

$$3\ddot{s} - \ddot{\lambda} = 0, \tag{A12}$$

from which property P_2 follows immediately.

On Some Unitary Representations of the Galilei Group I. Irreducible Representations*

J. VOISIN†
Physics Department, Syracuse University, Syracuse, New York
(Received 11 February 1965)

The true irreducible unitary representations of central extensions G_M of the Galilei universal covering group G and hence the physical representations of G are constructed by Mackey's method of induced representations. The elements of the representation space $\mathfrak{C}$ are obtained from functions defined on G_M and restricted to their values at one representative of each left coset of G_M modulo K where K is the induction subgroup. The physical interpretation of these functions is in terms of wave functions and comes from the definition of a basis in $\mathfrak{C}$. This interpretation depends on the choice of a fundamental frame of reference in space-time and on the physical meaning given to a fundamental state. To a change of the representatives corresponds a change of basis in $\mathfrak{C}$. By a suitable choice of these representatives, we obtain in particular the momentum-spin representation and the momentum-helicity representation. The zero mass case named class II by Inönü and Wigner is then obtained by the limit process $M \to 0$ applied to the helicity representation.

INTRODUCTION

SINCE the work of Wigner on the Poincaré group,[1] the method of induced representations has received more and more attention in theoretical physics. Its most complete formulation can be found in Mackey's Chicago lectures.[2] An introduction to Mackey's work has been given by several authors.[3] In particular, Wightman[4] has shown how the mathematical arguments go for the Euclidean group, emphasizing the measure theory aspects and the completeness of the corresponding induced representations. Although more concise, an essential account of the Galilei group can be found in Wightman's work from the point of view outlined above.

More recently another very elegant and useful introduction to Mackey's theory has been given by P. Moussa and R. Stora[5] who apply this technique for reducing the product of two irreducible representations of the Poincaré group and give at the same time general information about Mackey's theory (in particular, about functions defined on a group which, as recently shown by Lurçat,[6] are very important in the context of axiomatic field theory).

We thought it worthwhile to develop for the Galilei group the point of view taken by Moussa and Stora; hence the present work. In order to emphasize the parallelism between Poincaré and Galilei groups, we follow as much as possible the steps given by Moussa and Stora. We give however more importance to the irreducible representations, deducing the momentum–spin representation and the momentum–helicity representation from a unifying point of view and emphasizing the physical interpretation. We use also the helicity representation to deduce the mass-zero representations (named class II by Inönü and Wigner[3]) by a limit process $M \to 0$.

The Galilei group has been studied previously by several authors.[7] Especially close to our work is the article by Levy–Leblond[9] in which both the irreducible representations (for $M = 0$ and $M \neq 0$) and the l–S coupling are considered. With respect to this work, our originality lies (a) in the use of functions defined on the group and in the subsequent necessity of interpreting such functions, (b) in the new insight in the spectral decomposition of the representations with respect to the translations, (c) in the introduction of the helicity representation and

* Work supported by the U. S. Atomic Energy Commission.
† Aspirant du Fonds National de la Recherche Scientifique (Belgium). On leave from the University of Liege (Belgium).
[1] E. P. Wigner, Ann. Math. 40, 149 (1939).
[2] G. W. Mackey, *The Theory of Group Representations* (The University of Chicago Press, Chicago, Illinois, 1955).
[3] Besides the references 4 and 5 we mention the lectures on "The Theory of the Lorentz Group" given by Chr. Fronsdal, R. Hagedorn, J. M. Jauch, and R. Tolhoek (CERN, 1959) A useful introduction to the method of induced representations will be found in J. S. Lomont, *Applications of Finite Groups* (Academic Press, Inc., New York, 1959), Chap. V.
[4] A. S. Wightman, Rev. Mod. Phys. 34, 845 (1962).
[5] P. Moussa and R. Stora, preprint.
[6] F. Lurçat, Physics 1, 95 (1964).
[7] E. Inönü and E. P. Wigner[8] have considered the true representations and shown that they cannot be interpreted in a usual way. M. Hamermesh, Ann. Phys. 9, 518 (1960) has studied the physical interpretation both of the true representations and of the projective representations, using the Lie Algebra technique. E. C. G. Sudarshan in his lectures on *Principles of Classical Mechanics*, Report NYO-10250 (Rochester, 1963) gives some interpretation to the true representations in the frame work of classical mechanics. A. Loinger, Ann. Phys. 20, 132 (1962) shows that the only physical irreducible Hilbert representations of the proper Galilei group are those corresponding to the free particle Liouville Equation. See also references 4 and 9.
[8] E. Inönü and E. P. Wigner, Nuovo Cimento 9, 705 (1952).
[9] J. M. Levy-Leblond, J. Math. Phys. 34, 845 (1962).

its use in the study of the mass zero case and of the couplings, (d) in the unification of the different schemes (momentum–spin, momentum–helicity) obtained from Mackey's work.

The present article is devoted to the irreducible unitary representations of those central extensions G_M of the Galilei universal covering group which according to Bargmann's theorem[10] give the projective unitary representations of Galilei covering group. In the first section, we recall the usefulness of projective representations and Bargmann's theorem. In Secs. II and III, we construct the irreducible representations of G_M by Mackey's method. Section IV is devoted to the physical interpretation. In Sec. V we introduce the momentum–spin representation and the momentum–helicity representation and in Sec. VI we use this latter representation to study the mass-zero case.

The reduction of direct products of two representations will be considered in another article.

I. CENTRAL EXTENSIONS OF THE GALILEI COVERING GROUP.

Physically the Galilei group is the group of transformations relating the classical frames of reference. It contains the rotations R, the accelerations $\mathbf{v}$, the translations in space $\mathbf{a}$ and in time b. The elements of the group will be denoted by

$$(b, \mathbf{a}, \mathbf{v}, R)$$

and the multiplication law is

$$(b_2, \mathbf{a}_2, \mathbf{v}_2, R_2)(b_1, \mathbf{a}_1, \mathbf{v}_1, R_1)$$
$$= (b_2 + b_1, \mathbf{a}_2 + \mathbf{v}_2 \cdot b_1 + R_2\mathbf{a}_1, \mathbf{v}_2 + R_2\mathbf{v}_1, R_2R_1).$$

The covering group is obtained by replacing the rotation R by the elements of the group SU_2.

Since quantum mechanics does not fix the phase ω of the vector describing a pure state, and then associates to such a state a ray rather than a vector, the projective representations become essential when we study the symmetry properties of a quantum system. Wigner has shown[9] that the projective unitary representations of the Poincaré group $\mathcal{P}$ can be reduced to unitary representations up to a sign and thus in this case we may limit ourselves to the true representations of the universal covering group of $\mathcal{P}$. But in the case of the Galilei group, such a simplification does not occur. There is indeed a one-dimensional infinity of projective representation classes which cannot be reduced to the true representations.[10] They are characterized by the factors

$$\omega(b_2, \mathbf{a}_2, \mathbf{v}_2, R_2; b_1, \mathbf{a}_1, \mathbf{v}_1, R_1)$$
$$= \exp\,[i(M/\hbar)]\,(\mathbf{v}_2 \cdot R_2\mathbf{a}_1 + \tfrac{1}{2}\mathbf{v}_2^2 b_1) \qquad (1)$$

where the parameter M can run from $-\infty$ to $+\infty$. The important fact is that these projective representations are the only ones to which we can attribute a usual physical meaning.[7]

Another important fact is that the projective representations of type M can be deduced from the true unitary representations of a group G_M which is a central extension of the Galilei universal covering group by a one-dimensional Abelian group.[11] This eleven-parameter group G_M is defined as follows. Its elements are denoted by[12]

$$(\exp i\theta, b, \mathbf{a}, \mathbf{v}, R),$$

with the multiplication law

$$(\exp i\theta_2, b_2, \mathbf{a}_2, \mathbf{v}_2, R_2)(\exp i\theta_1, b_1, \mathbf{a}_1, \mathbf{v}_1, R_1)$$
$$= \{\exp i[\theta_2 + \theta_1 + (M/\hbar)(\mathbf{v}_2 \cdot R_2\mathbf{a}_1 + \tfrac{1}{2}\mathbf{v}_2 b_1)],$$
$$\times\, b_2 + b_1, \mathbf{a}_2 + \mathbf{v}_2 b_1, \mathbf{v}_2 + R_2\mathbf{v}_1, R_2R_1\}. \qquad (2)$$

The relation between the projective representations "M" of G and the representations of G_M is

$$U(\exp i\theta, b, \mathbf{a}, \mathbf{v}, R) = \exp i\theta \cdot U(b, \mathbf{a}, \mathbf{v}, R). \qquad (3)$$

We will often write $(\exp i\theta, b, \mathbf{a}, \mathbf{v}, R)$ as $(\exp i\theta, \mathbf{a}, \Gamma)$, $\mathbf{a}$ replacing both b and $\mathbf{a}$, and Γ replacing $\mathbf{v}$ and R. The multiplication law is then

$$(\exp i\theta_2, a_2, \Gamma_2)(\exp i\theta_1, a_1, \Gamma_1)$$
$$= \{\exp i(\theta_2 + \theta_1) \cdot \omega(\Gamma_2, a_1), a_2 + \Gamma_2 a_1, \Gamma_2\Gamma_1\}. \qquad (2')$$

The notation $\omega(\Gamma_2, a_1)$ for $\omega(a_2\Gamma_2; a_1\Gamma_1)$ is suited to express the fact that, M being given, ω depends on nothing but the homogeneous part of the left factor and the translation part of the right factor in the product.

In particular we have

$$(\exp i\theta, a, \Gamma)^{-1} = [\exp i(-\theta) \cdot \omega^{-1}(\Gamma, a^{-1}),$$
$$-b, -R^{-1}(\mathbf{a} - b\mathbf{v}), -R^{-1}\mathbf{v}, R^{-1}]$$

with

$$\omega^{-1}(\Gamma, a^{-1}) = \exp\,(-iM/\hbar)(-\mathbf{v} \cdot \mathbf{a} + \tfrac{1}{2}b\mathbf{v}^2). \qquad (4)$$

Let $\times$ be the symbol for the semidirect product.

[10] V. Bargmann, Ann. Math. 59, 1 (1954). A very useful account of that work can be found in M. Hamermesh, *Group Theory* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1962), Chap. 12. See also T. D. Newton in T. Kahan, *Théorie des Groupes en Physique Classique et Quantique* (Dunod & Cie., Paris, 1960), Vol. 1 Part II, Chap. III.

[11] V. Bargmann, Ref. 10.

[12] Here and in the following, we will denote by R the elements of SU^2 for sake of simplicity in the interpretation.

Then the structure of G_M may be written as[13]

$$G_M = T \times (V \times R) = T \times \Delta,$$

where T is the normal subgroup $\{(\exp i\theta,\ a,\ 1)\}$ formed by the subgroup of space-time translations and the one parameter central subgroup; V is the subgroup of pure Galilei transformations (V is the maximal invariant three-parameter subgroup of G_M/T); $R = SU_2 = G_M/T/V$ is the unitary unimodular two-dimensional group. It is simple.

We are going to study the representations of the G_M's. The general problem is to relate any representation to simpler ones. This problem is twofold involving:

1. The relation of the irreducible[14] unitary representations of G_M to those of simpler groups and
2. The relation of any given representation of G_M to the irreducible unitary representations of G_M itself. The first aspect is studied in the present paper. The latter one will be considered in a second article.

II. THE SUBGROUP $K = T \times R_{(q_0p_0)}$ AND ITS REPRESENTATIONS

The concept of induced representation of G_M is a relative one; it refers to a subgroup K of G_M and to a representation of this subgroup. Any subgroup of G_M can be used to induce a representation. However, if we want to get the irreducible representations of G_M, the subgroups to be taken are of the type $T \times R_{(q_0p_0)}$, i.e., they are semidirect products of the subgroup T with some little group $R_{(q_0p_0)}$.

1. The Abelian Normal Subgroup T and Its Unitary Irreducible Representations

The unitary irreducible representations of the Abelian subgroup T are one-dimensional and given by

$$\langle \exp i\theta,\ a \mid q,\ p\rangle = \exp i[q\theta + \hbar^{-1}(Eb - \mathbf{p}\cdot\mathbf{a})]. \quad (5)$$

Hence they are characterized by the numbers $(q,\ p)$.

We call any unitary irreducible representation of T a "character" of T. The characters of T form a group, the so-called character group $\bar{T}$ of T.

2. Automorphisms and Orbits of $\bar{T}$

Each element of the homogeneous subgroup Δ represents both an automorphism of T and an automorphism of $\bar{T}$.

The automorphism induced in T by the element Γ belonging to Δ is given by

$$(\exp i\theta,\ a,\ 1) \to \Gamma(\exp i\theta,\ a,\ 1)\Gamma^{-1}$$
$$= \{\exp i[\theta + (M/\hbar)(\mathbf{v}\cdot R\mathbf{a} + \tfrac{1}{2}\mathbf{v}^2 b)],\ \Gamma a,\ 1\}. \quad (6)$$

The corresponding transformation Γ^{-1} in T is defined by requiring that the transformed character evaluated at the point $(\exp i\theta,\ a,\ 1)$ be equal to the former character at the point $\Gamma(\exp i\theta,\ a,\ 1)\Gamma^{-1}$, i.e.,

$$\langle \exp i\theta,\ a \mid q'p'\rangle = \langle \exp i[\theta + (M/\hbar)$$
$$\times (\mathbf{v}\cdot R\mathbf{a} + \tfrac{1}{2}\mathbf{v}^2 b)],\ \Gamma a \mid q,\ p\rangle. \quad (7)$$

Hence Γ^{-1} induces the following transformation of the characters

$$\begin{aligned} q &\to q, \\ E &\to E - \mathbf{p}\cdot\mathbf{v} + \tfrac{1}{2}qM\mathbf{v}^2, \\ \mathbf{p} &\to R^{-1}(\mathbf{p} - qM\mathbf{v}), \end{aligned} \quad (8)$$

and obviously Γ induces the transformation

$$\begin{aligned} q &\to q, \\ E &\to E + R\mathbf{p}\cdot\mathbf{v} + \tfrac{1}{2}qM\mathbf{v}^2, \\ \mathbf{p} &\to R\mathbf{p} + qM\mathbf{v}. \end{aligned} \quad (8')$$

Let us point out that the transformed character has same q part as $(q,\ p)$ does.

These induced automorphisms in $\bar{T}$ allow us to split $\bar{T}$ into orbits in putting in the same orbit all the characters $(q,\ p)$ such that given any two of them (q_1p_1) and (q_2p_2) there exists an element Γ belonging to Δ transforming (q_1p_1) into (q_2p_2), i.e.,

$$(q_2p_2) = \Gamma[q_1p_1].$$

3. Little Groups

Let us call *little group* of G_M every subgroup of Δ leaving invariant at least one character (q_0p_0) and consider the little group $R_{(q_0p_0)}$ corresponding to (q_0p_0), i.e., the set of Γ's such that

$$\Gamma[q_0p_0] = (q_0p_0).$$

We thus have from (8)

$$\mathbf{p}_0 = R\mathbf{p}_0 + q_0M\mathbf{v}, \quad (9.1)$$

$$\mathbf{p}_0\cdot\mathbf{v} = \tfrac{1}{2}q_0M\mathbf{v}^2, \quad (9.2)$$

i.e.,

$$\begin{aligned} \mathbf{p}_0 &= R\mathbf{p}_0 + q_0M\mathbf{v}, \\ R\mathbf{p}_0\cdot\mathbf{v} &= -\tfrac{1}{2}q_0M\mathbf{v}^2, \end{aligned} \quad (9')$$

[13] See for instance J. M. Levy-Leblond Ref. 9.

[14] *Irreducible representation* will have in the present work the general meaning of representation where Schur's lemma holds, i.e. where any operator commuting with all the elements of the representation is a multiple of the identity. A more precise (and more correct) term would be primary representation.

as seen by multiplying (9.1) by $\mathbf{v}$ and taking (9.2) into account. From (9.1) we get $\mathbf{v}$ in terms of R and (q_0p_0). Hence the *little group of a* (q_0p_0) $p_0[(9')$ shows (q_0E) *does not matter*] *consists of all* $\Gamma = (\mathbf{v}, R)$ *of the form*

$$[q_0^{-1}M^{-1}(\mathbf{p}_0 - R\mathbf{p}_0), R] \tag{10}$$

and is isomorphic to the unitary unimodular SU_2.

4. The Subgroup $T \times R_{(q_0p_0)}$ and Its Representations

The semidirect product

$$K = T \times R_{(q_0p_0)} \tag{11}$$

is a subgroup of G_M and is closed. We can get a unitary irreducible representation of K if we are given an irreducible unitary representation of $R_{(q_0p_0)} = R_0$. Let D be such a representation of R_0. To the element[15] $(\exp i\theta, a) \times \Gamma_0 \in T \times R_0$, let us associate the matrix

$$L(\exp i\theta, a, \Gamma_0) = \langle \exp i\theta, a \mid q_0p_0\rangle D(R), \tag{12}$$

where $\langle \exp i\theta, a \mid q_0p_0\rangle$ is the representation of $(\exp i\theta, a)$ in the character (q_0p_0) and $D(\Gamma_0)$ is the representative of $[q_0^{-1}M^{-1}(\mathbf{p}_0 - R\mathbf{p}_0), R]$ in D. The identity (12) defines obviously a representation of K.

III. REPRESENTATION OF G_M INDUCED BY THE IRREDUCIBLE UNITARY REPRESENTATION L OF K.

Let us now determine the representation of $G_M = T \times \Delta$ induced by the representation (12) of the subgroup K.

1. Construction of the Induced Representation Hilbert Space.

We proceed in two steps.

Equivalence Classes or Left Cosets of G_M *mod* K

Two elements $g_1 = (q_1, a_1, \Gamma_1)$, $g_2 = (q_2, a_2, \Gamma_2)$ are in the same left coset of G_M mod K if and only if

$$g_2^{-1}g_1 \in K,$$

i.e.,

$$\Gamma_2^{-1}\Gamma_1 \in R_0$$

since belonging to K imposes no restriction on the translation part of the arguments.

Hence g_1 and g_2 are in the same left coset if and only if their components Γ_1 and Γ_2 transform the character $(\exp i\theta_0, p_0)$ in the same way. For

$$\Gamma_2^{-1}\Gamma_1 \in R_0$$

$$\Leftrightarrow \Gamma_1[\exp i\theta_0, p_0] = \Gamma_2[\exp i\theta_0, p_0] \tag{13}$$

The following theorem holds.

Theorem. There is a one-to-one correspondence between the characters belonging to the orbit of (q_0, p_0) and the left cosets of G_M mod K.

This theorem enables us to label the cosets by the characters equivalent to (q_0p_0). Let us point out that

$$E - \mathbf{p}^2/2Mq_0 = E_0 - \mathbf{p}_0^2/2Mq_0 = U_0 \tag{14}$$

so that the internal energy is the same in all the states of the corresponding primitive system. Note also that the independence of R_0 with respect to (q_0E_0) is related to the freedom left in the choice of the internal energy.

Definition of the representation space $\mathfrak{K}(G_M)$

Let us consider a set of functions of domain G_M satisfying the following covariance property.

$$f(gk) = f(g)L^*(k)$$

in which g is any element belonging to G_M and k any element of K. The symbol* means complex conjugation. With the previous notations, we have more precisely for the *covariance property*

$$f[(\exp i\theta, a, \Gamma)(\exp i\theta', a', R')]$$
$$= f(\exp i\theta, a, \Gamma)L^*(\exp i\theta', a', R') \tag{15}$$

where

$$(\exp i\theta', a', R') \in K.$$

As a result f will be completely characterized if we are given its value for one (arbitrary) element of each coset.[16]

The above discussion leads us to pick a particular element $g(q_0p)$ from each coset. For that purpose, let us first associate to each Γ a particular $\Gamma_{(q_0p)} = \Gamma_p$ such that

$$\Gamma[q_0p_0] = (q_0p) = \Gamma_p[q_0p_0]. \tag{16}$$

Γ_p is thus the same for all the Γ's such that $\Gamma[q_0p_0] = (q_0p)$. We then take (formally, see Sec. V)

$$(1, 0, \Gamma_p)$$

as the representative of the coset (q_0p) for any $p(q_0$ might be taken—and will be later taken—equal to one without any loss of generality).

On the other hand as $(q_0E\mathbf{p})$ on the orbit of $(q_0E_0\mathbf{p}_0)$ clearly indexes left cosets of G_M mod K we

[15] We write $(\exp i\theta, a)$ for $(\exp i\theta, a, 1)$.

[16] The domain of f is thus strongly connected with the orbit of (q_0p_0). Note however that f is not a constant on a given coset. Such a constant will be given, e.g., by the function $\langle f(g), f'(g)\rangle$.

may choose as invariant measure on the coset space the invariant measure

$$\delta[E - \mathbf{p}^2/2Mq_0 - (E_0 - \mathbf{p}_0^2/2Mq_0)]\, dE\, d\mathbf{p} = d\Omega_{U_0}(E, \mathbf{p}). \quad (17)$$

We then suppose that the functions f defined above have the following property: their restriction $f(\Gamma_p) = f(1, 0, \Gamma_p)$ to the representatives $(1, 0, \Gamma_p)$ form a Hilbert space with respect to the scalar product

$$\int d\Omega_{U_0}(p) f'^{\dagger}(\Gamma_p) f(\Gamma_p).$$

We take this space as representation Hilbert space $\mathcal{H}(G_M)$.

2. Representation of G_M in $\mathcal{H}(G_M)$

To define $U(\exp i\theta, a, \Gamma)$ we give the value of Uf at the points $(1, 0, \Gamma_p)$ according to the following law

$$[U(\exp i\theta, a, \Gamma)f](\Gamma_p) = f[(\exp i\theta, a, \Gamma)^{-1}(1, 0, \Gamma_p)] \quad (18)$$

Indeed as a result of the invariance property (15), the second member of (18) can be expressed in terms of the value of f at some point $(1, 0, \Gamma_{p'})$ For let us put the argument $(\exp i\theta, a, \Gamma)^{-1}(1, 0, \Gamma_p)$ into the product form $(1, 0, \Gamma_{p'}) \times k$ where $k \in R_0$. We first note that

$$(\exp i\theta, a, \Gamma)^{-1}(1, 0, \Gamma_p) = \{\exp i[-\theta + (M/\hbar) \times (\mathbf{v}\cdot\mathbf{a} - \tfrac{1}{2}b\mathbf{v}^2)], -\Gamma^{-1}a, \Gamma^{-1}\Gamma_p\}$$

and that this argument belongs to the coset $\Gamma^{-1}[q_0p]$. For

$$\Gamma^{-1}\Gamma_p[q_0p_0] = \Gamma^{-1}[q_0p] = [q_0p']. \quad (19)$$

Hence if we multiply $\Gamma^{-1}\Gamma_p$ from the left by $\Gamma_{p'}^{-1}$ the result will belong to the little group R_0. This suggests proceeding in the following way. We write

$$\begin{aligned}&(\exp i\theta, a, \Gamma)^{-1}(1, 0, \Gamma_p)\\ &= (1, 0, \Gamma_{p'})(1, 0, \Gamma_{p'}^{-1})(\exp i\theta, a, \Gamma)^{-1}(1, 0, \Gamma_p)\\ &= (1, 0, \Gamma_{p'})(1, 0, \Gamma_{p'}^{-1})\\ &\times \{\exp i[-\theta + (M/\hbar)(\mathbf{v}\cdot\mathbf{a} - \tfrac{1}{2}b\mathbf{v}^2), -\Gamma^{-1}a, 1\}\\ &\times (1, 0, \Gamma_{p'})(1, 0, \Gamma_{p'}^{-1}\Gamma^{-1}\Gamma_p). \quad (20)\end{aligned}$$

The second member of (20) is obviously of the form $(1, 0, \Gamma_{p'}) \times \varphi$ where $\varphi \in T \times R_0$ so that we get from (18)

$$[U(\exp i\theta, a, \Gamma)f](\Gamma_p) = f(\Gamma_{p'})L^*(\varphi) = L(\varphi^{-1})f(\Gamma_{p'}). \quad (21)$$

Now

$$\begin{aligned}L(\varphi^{-1}) &= \langle (1, 0, \Gamma_{p'}^{-1})\{\exp i[\theta - (M/\hbar)\\ &\times (\mathbf{v}\cdot\mathbf{a} - \tfrac{1}{2}b\mathbf{v}^2)], \Gamma^{-1}a, 1\}\\ &\times (1, 0, \Gamma_{p'}) \mid q_0p_0\rangle D(\Gamma_p^{-1}\Gamma\Gamma_{p'}) = \tau\cdot D(\Gamma_p^{-1}\Gamma\Gamma_{p'}).\end{aligned}$$

On the other hand according to (6) and (7)

$$\tau = \langle \exp i\theta, a \mid q_0p\rangle \quad (22)$$

so that we eventually get the following representation of G_M:

$$[U(\exp i\theta, a, \Gamma)f](\Gamma_p) = \langle \exp i\theta, a \mid q_0p\rangle D(\Gamma_p^{-1}\Gamma\Gamma_{p'})f(\Gamma_{p'}) \quad (23)$$

in complete analogy with the Poincaré group.

This result is true for any representative Γ_p. We will consider later two important choices for Γ_p. Such a consideration is not trivial from a physical point of view since a change of representatives leads to a new representation of G_M which is equivalent but not identical to the former one.

IV. PHYSICAL INTERPRETATION. CHANGE OF REPRESENTATIVES AND EQUIVALENT REPRESENTATIONS

1. Basis in $\mathcal{H}(G_M)$

The physical interpretation of $f(\Gamma_p)$ is in terms of wave functions.

Let us make a particular choice for the little group and take $(M \neq 0)$

$$(E_0, \mathbf{p}_0) = (E_0, 0).$$

As the particular value chosen for q_0 does not matter, we will omit writing it in this section. The little group R_0 is now the group SU_2 and D is an irreducible representation of SU_2. Let us then index the matrix elements of D by integer or half integer numbers s, s' and let us introduce a non-normalizable basis

$$|\Gamma_p s\rangle \quad (25)$$

of a Hilbert space and states (summation convention for s)

$$|\phi\rangle = \int d\Omega_{U_0}(p) f_s(\Gamma_p)\, |\Gamma_p s\rangle. \quad (26)$$

The orthonormality condition is

$$\langle \Gamma_p s \mid \Gamma_{p'} s'\rangle = \delta_{E_0}(p, p')\delta_{ss'} \quad (26)$$

where $\delta(p, p')$ is the invariant δ function on $E_0 = E - \mathbf{p}^2/2M$

$$\int d\Omega_{E_0}(p')\delta_{E_0}(p, p')f(p') = f(p). \quad (27)$$

This implies

$$\langle \psi \mid \phi \rangle = \int d\Omega_{E_0}(p')\, d\Omega_{E_0}(p) f'^{*}_{s's}(\Gamma_{p'}) f_{s}(\Gamma_p) \times \langle \Gamma_{p's'} \mid \Gamma_p s \rangle$$

$$= \int d\Omega_{E_0}(p) f'^{*}_{s's}(\Gamma_p) f_{s}(\Gamma_p).$$

Now $f'^{\dagger}(\Gamma_p) f(\Gamma_p)$ depends on p only since from (15)

$$f'^{\dagger}(\Gamma'_p) f(\Gamma'_p) = f'^{\dagger}(\Gamma_p R) f(\Gamma_p R) = f'^{\dagger}(\Gamma_p) f(\Gamma_p). \tag{28}$$

Hence

$$\langle \psi \mid \phi \rangle = \int d\Omega_{E_0}(p) \langle \tilde{\psi} \mid \tilde{\varphi} \rangle (p). \tag{29}$$

2. Covariance Property in Terms of the Basis and Equivalence of Representations

The interpretation of (28) is in terms of equivalence of the representations corresponding to different representatives.

To see this let us consider the covariance law (15) for the wavefunctions and its correspondent for the states. We get the latter one as follows. Let $|\phi\rangle$ be developed in terms of the $|\Gamma_p s\rangle$ and of the $|\Gamma'_p s'\rangle$ respectively where Γ'_p belongs to the same coset as Γ_p does.

We have

$$|\phi\rangle = \int d\Omega_{E_0}(p) f_s(\Gamma_p) \, |\Gamma_p s\rangle$$

$$= \int d\Omega_{E_0}(p) f_{s's}(\Gamma'_p) \, |\Gamma'_p s'\rangle \tag{30}$$

and

$$\Gamma'_p = \Gamma_p R,$$

where

$$R \in R_0 = SU_2.$$

Hence

$$f(\Gamma'_p) = f(\Gamma_p) D^{*}(R) = f(\Gamma_p) D^{*}(\Gamma_p^{-1} \Gamma'_p).$$

In terms of components

$$f_{s's}(\Gamma'_p) = f_s(\Gamma_p) D^{*}_{ss'}(\Gamma_p^{-1} \Gamma'_p) \tag{31}$$

(30) and (31) give then

$$|\phi\rangle = \int d\Omega_{E_0}(p) f_s(\Gamma_p) \, |\Gamma_p s\rangle$$

$$= \int d\Omega_{E_0}(p) f_s(\Gamma_p) D^{*}_{ss'}(\Gamma_p^{-1} \Gamma'_p) \, |\Gamma'_p s'\rangle.$$

Hence

$$|\Gamma_p s\rangle = D^{*}_{ss'}(\Gamma_p^{-1} \Gamma'_p) \, |\Gamma'_p s'\rangle = D^{-1}_{s's}(\Gamma_p^{-1} \Gamma'_p) \, |\Gamma'_p s'\rangle$$

and eventually

$$|\Gamma'_p s'\rangle = |\Gamma_p s\rangle D_{ss'}(\Gamma_p^{-1} \Gamma'_p). \tag{32}$$

This result may then be interpreted as follows:

A change $\Gamma_p \to \Gamma'_p$ in the representatives of the cosets results in the change of basis

$$|\Gamma_p s\rangle \to |\Gamma'_p s\rangle$$

given by (32). This transformation law (32) is the *covariance property for the states.*

Now we see that with this interpretation, (28) expresses the unitarity of the transformation law, as expected.

3. Physical Interpretation[17]

So far the functions f and the states $|\Gamma_p s\rangle$ are formal expressions. To get a physical interpretation, we have to give a meaning to the *symbols* Γ_p and s in $|\Gamma_p s\rangle$. For that purpose let us first of all choose a particular frame of reference F_0 (e.g., the laboratory system) and associate with the transformation $\Gamma = (\mathbf{v}, R)$ the frame ΓF_0 which has the velocity $\mathbf{v}$ with respect to F_0 and is parallel to the frame deduced from F_0 by the rotation R (Axiom I). As a result, $\mathbf{v}$ and R in Γ are now to be interpreted as a velocity and a rotation in the frame F_0. We note that the first axiom is very general and does not depend on M (either $\neq 0$ or $= 0$). A second general axiom will consider the characters $(E, \mathbf{p})$ as representing the linear momentum $\mathbf{p}$ and the energy E of the system (Axiom II). The further requirements depend on G_M and the representation of G_M associated with the particle.

We consider the case $M \neq 0$ and we first complete the physical meaning we give to $|\Gamma_p s\rangle$ when we write the symbol p. Let $(E_0, 0)$ be the character corresponding to the little group defining the representation we deal with. To Γ we associate the 4-momentum

$$(E, \mathbf{p}) = \Gamma[E_0, 0] = (E_0 + \mathbf{p}^2/2M, \mathbf{p} = M\mathbf{v})$$

and we then interpret $|\Gamma, s\rangle$ (for any s) as a state of momentum $\mathbf{p}$ and energy E (Axiom III). Now the state $|1, s\rangle$ (any s), in which, from the latter axiom,

[17] The following discussion is somewhat incorrect since we are dealing with the Galilei covering group and not directly with the Galilei group. The idea underlying the argument has been suggested to us by reading the article by Lurçat quoted in Ref. 6. See also Ref. 18.

the particle has momentum 0 and energy E_0, is also interpreted as a state where the particle is at rest in $F_0(\mathbf{v} = 0)$ (Axiom IV).

The representation of $\Gamma_{p'}$ in the basis $|\Gamma_p s\rangle$ is

$$U(\Gamma_{p'})\,|\Gamma_p s\rangle = |\Gamma_\pi s'\rangle D_{s's}(\Gamma_\pi^{-1}\Gamma_{p'}\Gamma_p) \qquad \pi = \Gamma_{p'}[p] \tag{33}$$

since

$$U(\Gamma_{p'})\,|\phi\rangle = \int d\Omega_{E_0}(p) f_s(\Gamma_p) U(\Gamma_{p'})\,|\Gamma_p s\rangle$$

$$= \int d\Omega_{E_0}(p) f'_{s'}(\Gamma_p)\,|\Gamma_p s'\rangle,$$

and from (23) with $\pi' = \Gamma_{p'}^{-1}[p]$

$$f'_{s'}(\Gamma_p) = D_{s's}(\Gamma_p^{-1}\Gamma_{p'}\Gamma_{\pi'}) f_s(\Gamma_{\pi'}).$$

It then results from (33) and our former axioms that if we impart a velocity $\mathbf{v}$ and a rotation R to the particle initially in the state $|1s\rangle$, the particle will get the momentum

$$\mathbf{p} = M\mathbf{v} \tag{34}$$

and the energy

$$E = E_0 + \tfrac{1}{2}M\mathbf{v}^2. \tag{34'}$$

Hence when the particle has the velocity $\mathbf{v}$ in F_0, its momentum and its energy are completely determined by (34), (34'). On the other hand as the frame $F = \Gamma_p F_0$ has the velocity $\mathbf{v}$ too, we see that F is a system in which the particle is at rest when in the state $|\Gamma_p s\rangle$ with respect to F_0.

The result of the above discussion is as follows: given a laboratory system F_0, $|\Gamma_{(E,p)} s\rangle$ is first of all to be considered as symbolizing a state in which the particle has energy E, momentum $\mathbf{p}$ in F_0.

We now consider the interpretation of s. It will depend on the form of Γ_p. Let us study successively the states $|1s\rangle$, $|Rs\rangle$, $|\Gamma s\rangle$ where Γ_p is a pure Galilei transformation, and $|\Gamma'_p s\rangle$ where Γ'_p contain a certain rotation.

(a) *The States* $|1, s\rangle$

We know that they correspond to a frame parallel to F_0 (it is F_0 itself) and to a particle at rest in F_0. Hence they give a basis for the possible states of the system when it is at rest in F_0.

On the other hand if we apply a rotation R to $|1s\rangle$ (i.e., if the system initially in the state $|1s\rangle$ is submitted to the rotation R) we get from (33)

$$U(R)\,|1s\rangle = |1s\rangle D^j_{s's}(R). \tag{35}$$

Hence the vectors $|1s\rangle$ constitute a basis for the j-representation of the rotation group associated with F_0. As a result *s can be interpreted as labeling the eigenvalues of the projection $\mathbf{S}\cdot\mathbf{e}_3$ on $\mathbf{e}_3$ of the angular momentum of the system at rest in F_0* (s then takes our choice of F_0 and in particular of $\mathbf{e}_3$ into account). Now the meaning of (35) is the following. If we apply a rotation to the system initially in the state $|1s\rangle$ the resulting state is no longer an eigenstate of S_3 but a superposition of such states. Computing the mean value of $\mathbf{S}\cdot R\mathbf{e}_3$ in the state $U(R)\,|1s\rangle$ we see that it is equal to s. In this state it is this projection and not $\mathbf{S}\cdot\mathbf{e}_3$ which is equal to s. In particular if s is the maximal value of S_3 we may say that in the state $|1s\rangle$ the spin is parallel to $\mathbf{e}_3$. Then in the state $U(R)\,|1s\rangle$, the spin is parallel to $R\mathbf{e}_3$. R has as the effect of rotating the spin through θ about $\mathbf{n}$ where $\mathbf{n}$ and θ are the axis and the angle of R respectively. These results are obviously well known. All that we do is to get them again from our point of view.

(b) *The States* $|Rs\rangle$

It results immediately from (32) that

$$|Rs\rangle = |1s'\rangle D_{s's}(R). \tag{36}$$

Comparing this result with (35) we thus have

$$|Rs\rangle = U(R)\,|1s\rangle. \tag{37}$$

Hence according to the discussion following (35), $|Rs\rangle$ can be interpreted as a state in which the system has momentum zero [since $(0, R)[E_0, 0] = (E_0, 0)$] and has s as projection of its F_0 -spin on $R\mathbf{e}_3$. In particular if s is maximum the spin is parallel to $R\mathbf{e}_3$ in the state $|Rs\rangle$

Let us note that if we adopt the passive point of view and change the fundamental frame from F_0 to RF_0, the state $|Rs\rangle$ will obviously become a state $|1s\rangle$.

We emphasize also the fact that to interpret $|Rs\rangle$, we reduce it to the states $|1s'\rangle$.

(c) *The States* $|\Gamma_p s\rangle$ *where* Γ_p *is a pure Galilei transformation*

In such states the particle has energy E and momentum $\mathbf{p}$; the associated frame is parallel to F_0 and moving with the velocity $\mathbf{v} = \mathbf{p}/M$.

According to (33) we have

$$U(\Gamma_p)\,|1s\rangle = |\Gamma_p s\rangle. \tag{38}$$

As Γ_p contains no rotation (pure case) the only possibility given by this formula (38) for the interpretation is to identify s in $|\Gamma_p s\rangle$ to s in $|1s\rangle$. We then adopt the following interpretation: *$|\Gamma_p s\rangle$ is a state in which the system has energy E, momentum*

p *and the same orientation in* F_0 *as it has in the state* $|1s\rangle$. (Axiom V).

From a passive point of view, $|\Gamma_p s\rangle$ is in F_0 that state which viewed from the frame $\Gamma_p F_0$ is of the type $|1s\rangle$.

(d) *The States* $|\Gamma'_p s\rangle$ *where* Γ'_p *contains a Rotation*

We restrict ourselves to the case where

$$\Gamma'_p = \Gamma_p R_p = R_p \Gamma_{p_s}, \tag{39}$$

where Γ_p is of the type discussed in (c), and R_p is the rotation of axis $[\mathbf{e}_3, \mathbf{p}]$, and angle θ equal to the smallest positive angle between $\mathbf{e}_3$ and **p**; p_3 is the momentum $(E, |\mathbf{p}|\mathbf{e}_3)$.

In the present case, Γ'_p involves the pure transformations (b) (rotation) and (c) (pure Galilei transformation). We may then hope to get the interpretation of $|\Gamma'_p s\rangle$ from the one we have adopted for the previous cases.

First of all Γ'_p is associated to a frame moving with the velocity $\mathbf{p}/M$ in F_0 and $|\Gamma'_p s\rangle$ corresponds to a particle having momentum **p** and energy E in F_0.

On the other hand we have from (33)

$$U(\Gamma'_p)\,|1s\rangle = |\Gamma'_p s\rangle \tag{40}$$

and thus from (39), (37)

$$U(\Gamma_p)U(R)\,|1s\rangle = U(\Gamma_p)\,|Rs\rangle = |\Gamma'_p s\rangle. \tag{41}$$

Hence since according to (c) we give the same interpretation (the same nature) to the quantity s appearing in two states $|Rs\rangle$, $|\Lambda s\rangle$ which are related by a pure Galilei transformation, we have as first result: *the system in the state* $|\Gamma'_p s\rangle$ *has the same orientation with respect to* F_0 *as it does in the same frame when in the state* $|Rs\rangle$.

Now the orientation in the state $|Rs\rangle$ is defined in (b) in terms of the projection of the spin on $R\mathbf{e}_3$. The interpretation of s in $|\Gamma'_p s\rangle$ is then complete. As $R = R_p$ the quantity s *in* $|\Gamma'_p s\rangle$ *amounts to the projection of the spin on the direction of motion* **p**. If s is maximal the spin of the state $|\Gamma'_p s\rangle$ is parallel to the direction of motion.

An important remark is here in order: s *can also be interpreted as the projection on* **p** *of the total angular momentum in the state* $|\Gamma'_p s\rangle$. This result is crucial as far as the comparison with the zero-mass case is concerned.

Obviously the passive interpretation is in terms of the frame $F = \Gamma'_p F_0$ the third axis of which is parallel to **p**.

The states $|\Gamma'_p s\rangle$ where the rotation component of Γ'_p is R_p are the so-called momentum–helicity eigenstates.[18] If the rotation component is different from the latter one, the corresponding state will be a superposition of helicity eigenstates.

(e) *Conclusion*

The spin of a system ($M \neq 0$) as actively defined above is a relative concept referring to some particular fundamental frame F_0 through a basic 4-momentum $(E_0, 0)$ (little group dependence). The spin gives the orientation with respect to F_0 of the system at rest in F_0. Studying the F_0 -spin is equivalent to studying in F_0 the angular momentum of the system at rest. And indeed we have reduced any state $|\Lambda s\rangle$ to states of the type $|1s\rangle$ or to superpositions of such states in order to interpret their s -dependence. The reduction has been performed by admitting as an axiom the identity of the "s" in two states $|Rs\rangle$, $|\Lambda s\rangle$ related by a pure Galilei transformation. In the helicity states, s may also be interpreted as the projection on **p** of the total angular momentum.

A passive interpretation involving the frames $F = \Gamma F_0$ can be developed in a manner parallel to the active one we gave above.

With the conventions introduced we know now how to read the physical content of the notation $|\Gamma_p s\rangle$. The following are meaningful: on the one hand the character p giving the momentum and the energy and on the other hand the number s as well as the rotation component of Γ_p giving the orientation of the particle.

V. TWO PARTICULAR REPRESENTATIONS OF $G_M (M \neq 0)$.

We are going to give the explicit expression of the representation (23) of G_M in two particular basis of the Hilbert space $\mathcal{H}(G_M)$. We take $(q_0 p_0) = (1, E_0, 0)$ and the representation j for the little group SU_2.

1. Preliminary Remarks

First of all we want to point out that the matrix D appearing in (23) may be written as

$$W(\Gamma; p)$$

to express that in a concrete representation of G_M the only variables are Γ and p. Naturally the func-

[18] The concept of helicity state has been especially well defined by M. Jacob and C. G. Wick, Ann. Phys. 7, 404 (1953), in particular with respect to the phase determination. The presentation of the helicity representation in relativistic theory is also very clear in the article by V. I. Ritus, Zh. Eksperim. i Teor. Fiz. 40, 352 (1961) [English transl.: Soviet Phys.—JETP 13, 240 (1961)]. The main points of our discussion are to work in nonrelativistic theory and to make more explicit both the axiomatic aspects and the deductive aspects of the usual interpretation.

tion W itself depends on the little group, on the representation of this little group and on the representatives which together fix the given induced representation.

On the other hand, as in a particular representation, the representatives of the cosets are fixed, it is convenient to write $f(\Gamma_p)$ as

$$\Theta(p) = f(\Gamma_p), \tag{42}$$

where the function θ depends on our choice of the representatives according to (42).

It is also convenient here to use the usual notation $\exp i(\theta + \hbar^{-1}\mathbf{p}\cdot\mathbf{a})$ for $\langle \exp i\theta, a \,|1; p\rangle$. With these notations, (23) becomes (we take the components)

$$[U(\exp i\theta, a, \Gamma)\Theta]_u(p) = \exp i(\theta + \hbar^{-1}\mathbf{p}\cdot\mathbf{a})W(\Gamma, p)_{uu'}\Theta_{u'}(\Gamma^{-1}p), \tag{43}$$

where $u, u' = (j, j-1, \cdots, -j)$ have so far no physical interpretation. (43) can also be written as follows

$$[U(\exp i\theta, a, \Gamma)\Theta]_u(p) = \int d\Omega_{E_0}(p'') \times \exp i(\theta + \hbar^{-1}\Gamma p''\cdot a)W(\Gamma, \Gamma p'')_{uu'} \times \delta_{E_0}(p'', \Gamma^{-1}p)\Theta_{u'}(p''). \tag{44}$$

Now let us express this result in terms of the basis vectors. We have

$$\Theta_u(p) = \langle p, u \mid \Theta\rangle$$

and

$$(U\Theta)_u(p) = \langle p, u \mid U\Theta\rangle = \langle p, u \,|U|\, p', u'\rangle\langle p', u' \mid \Theta\rangle,$$

where $|\theta\rangle$ is the state described by $\Theta_u(p)$ in the basis $|p, u\rangle$.

Hence according to (44)

$$\langle p, u \,|U|\, p', u'\rangle\langle p', u' \mid \Theta\rangle = \exp i(\Theta + \hbar^{-1}\Gamma p''\cdot a)W_{uu'}(\Gamma, \Gamma p'')\langle p'', u'' \mid \Theta\rangle$$

and the orthogonality property of the basis then gives

$$\langle p, u \,|U|\, p', u'\rangle = \exp i(\theta + \hbar^{-1}\Gamma p'\cdot a) \times W_{uu'}(\Gamma; \Gamma p')\delta_{E_0}(p', \Gamma^{-1}p). \tag{45}$$

Now as

$$U\,|p, u\rangle = |p', u'\rangle\langle p', u'|\, U\,|p, u\rangle,$$

we get from (45) *the transformation law of the basis states* $|pu\rangle$ *under* G_M

$$U(\exp i\theta, a, \Gamma)\,|p, u\rangle = \exp i(\theta + \hbar^{-1}\Gamma p\cdot a)\,|\Gamma p, u'\rangle W_{u'u}(\Gamma; \Gamma p). \tag{46}$$

2. The Momentum–Spin Representation

The momentum–spin representation is defined by taking as representatives for the left cosets of G_M mod K the pure Galilei transformations Γ_p

$$\Gamma_p = (1, 0, \mathbf{p}/M, 1). \tag{47}$$

We naturally have ($q_0 = 1$)

$$\Gamma_p[E_0, 0] = (E_0 + \mathbf{p}^2/2M, \mathbf{p}) = p.$$

For the sake of definiteness, we adopt in the present case the following notations

$$W(\Gamma; p) \to Q(\Gamma; p), \quad u \to s, \quad \Theta(p) \to \varphi(p).$$

Now according to (23) and the present conventions ($p' = \Gamma^{-1}p$),

$$[Q(\Gamma; p)]_{ss'} = D^j_{ss'}(\Gamma_p^{-1}\Gamma\Gamma_{p'}). \tag{48}$$

From (47)

$$\Gamma_p^{-1}\Gamma\Gamma_{p'} = (-\mathbf{p}/M, 1)(\mathbf{v}, R)[R^{-1}(\mathbf{p} - M\mathbf{v})/M, 1].$$

Hence

$$\Gamma_p^{-1}\Gamma\Gamma_{p'} = (0, R) \tag{49}$$

and

$$[Q(\Gamma; p)]_{ss'} = D^j_{ss'}(R). \tag{50}$$

We thus have the following theorem.

Theorem. In the momentum–spin representation, the irreducible unitary representation of G_M corresponding to spin j and rest energy E_0 can be written ($q_0 = 1$)

$$[U(\exp i\theta, a, \Gamma)\varphi]_s(p) = \exp i[\theta + \hbar^{-1}(Eb - \mathbf{p}\cdot\mathbf{a})]D^j_{ss'}(R)\varphi_{s'}(\Gamma^{-1}p), \tag{51}$$

where $D^j(R)$ is the j-representation of SU_2 in the abstract basis (J^2, J_3) and p varies on the orbit of the character $(E_0, 0)$, i.e.,

$$E - \mathbf{p}^2/2M = E_0$$

Let us point out however that so far, s has not been interpreted physically. But as the basis of the present representation (51) is $|\Gamma_p s\rangle$, where Γ_p contains no rotation, the discussion of Sec. IV.3(c) immediately applies and s is to be considered as the projection of the intrinsic angular momentum on the axis $\mathbf{e}_3$ of our laboratory system F_0.

Now according to (46) the transformation law of the basis vectors is

$$U(\exp i\theta,\ a,\ \Gamma)\ |p,\ s\rangle = \exp i(\theta + \hbar^{-1}\Gamma p\cdot a)\ |\Gamma p,\ s'\rangle D^{i}_{s's}(R). \qquad (52)$$

Hence under the transformation $(\exp i\theta,\ a,\ \Gamma)$ the spin is not affected by the velocity component and is rotated by exactly the rotation component R. Such a result is not true however for the linear momentum since under $(\exp i\theta,\ a,\ \Gamma)$

$$\mathbf{p} \to \mathbf{p}' = R\mathbf{p} + M\mathbf{v}. \qquad (53)$$

Hence $\mathbf{p}' = R\mathbf{p}$ if and only if either $\mathbf{v} = 0$ or $\mathbf{v}$ is parallel to $R\mathbf{p}$.

Under the action of $(\exp i\theta,\ a,\ \Gamma)$ the spin and the momentum are generally rotated differently. If $\mathbf{v} = 0$ or if $\mathbf{v}$ is parallel to $R\mathbf{p}$ the rotation is the same both for $\mathbf{S}$ and for $\mathbf{p}$ and is equal to R. As a result a system initially in a helicity eigenstate will not be any more in a helicity state after a transformation $(\exp i\theta,\ a,\ \Gamma)$ with $\mathbf{v}$ not satisfying the conditions above.

3. The Momentum–Helicity Representation

Instead of taking the pure Galilei transformation Γ_p to represent the left cosets, let us take the elements

$$\Gamma'_p = \Gamma_p R_p = R_p \Gamma_{p_s} \qquad (54)$$

with the notations of Sec. IV.3(d).

We have from (47)

$$\Gamma'_p = (1,\ 0,\ \mathbf{p}/M,\ R_p). \qquad (55)$$

We will write here

$$W(\Gamma; p) \to S(\Gamma; p),$$
$$u \to \sigma,$$
$$\Theta(p) \to \psi(p).$$

Let us recall that σ and s have same numerical values; taking different notations for these two quantities takes into account the difference in nature which we are going to find between them.

The corresponding representation for G_M is according to (43)

$$[U\ (\exp i\theta,\ a,\ \Gamma)\psi]_\sigma(p) = \exp i(\theta + \hbar^{-1}\mathbf{p}\cdot\mathbf{a})[S(\Gamma; p)]_{\sigma\sigma'}\psi_{\sigma'}(\Gamma^{-1}p). \qquad (56)$$

The relation of (56) to (51) is easily obtained since from (23) (48) (54) and our conventions for the little group $(p' = \Gamma^{-1}p)$

$$[S(\Gamma; p)]_{\sigma\sigma'} = D^{i}_{\sigma s}(\Gamma'^{-1}_p \Gamma \Gamma'_{p'}) = D^{i}_{\sigma s}(R_p^{-1}) D^{i}_{ss'}(R) D^{i}_{s'\sigma'}(R_{p'}) \qquad (57)$$

and hence

$$[S(\Gamma; p)]_{\sigma\sigma'} = D^{i}_{\sigma s}(R_p^{-1})[Q(\Gamma; p)]_{ss'} D^{i}_{s'\sigma'}(R_{p'}), \qquad (58)$$

Where

$$\mathbf{p}' = R^{-1}(\mathbf{p} - M\mathbf{v}).$$

As to the interpretation of this new representation, the considerations of Section IV.3(d) give us immediately the following result: the basis vector $|p\sigma\rangle$ is the state of momentum p and helicity σ, i.e., the state $|\Gamma'_p\sigma\rangle$.

Given the basis $|ps\rangle$, the transformation law to the basis $|p\sigma\rangle$ is obtained from (41) and (52). We have ($\sigma = s$ *numerically*)

$$|p\sigma\rangle = V(R_p)\ |\Gamma_{p_s}s\rangle = |ps'\rangle D^{i}_{s'\sigma}(R_p).$$

Finally we write the general transformation law under G_M of the states $|p\sigma\rangle$ which we deduce from (46)

$$U(\exp i\theta,\ a,\ \Gamma)\ |\Gamma^{-1}p,\ \sigma\rangle = \exp i(\theta + \hbar^{-1}\mathbf{p}\cdot\mathbf{a})\ |p,\ \sigma'\rangle S_{\sigma'\sigma}(\Gamma; p) \qquad (59)$$

with the formula (57) for $S_{\sigma'\sigma}$.

VI. MOMENTUM–HELICITY REPRESENTATION AND MASS-ZERO CASE

It is well known in Poincaré group theory that if we take $M = 0$ in the irreducible momentum–spin representation of mass M and spin j, we get a new representation which on the one hand is no longer irreducible and on the other hand is not yet reduced. One of the most interesting features of the helicity representation is then to provide a limit representation which is already reduced into its $(2j + 1)$ components.[19]

It has been pointed out by Levy–Leblond[9] that making $M = 0$ in the momentum–spin representation (51) for the Galilei group leads also to a reducible representation. We show here that if we work in the helicity representation, we get immediately the reduced form of the limit. This was expected from the fact that one of the central operators for the zero-mass case of class II[8] is the helicity operator $\mathbf{J}\cdot\mathbf{p}$.[20] Let us proceed to the proof.

If we put $M = 0$ in (8), we see that the transformation of the 4-momentum of the corresponding zero-mass particle is

[19] E. Inönü and E. P. Wigner, Natl. Acad. Sci. USA **39**, 510 (1953); E. P. Wigner, Rev. Mod. Phys. **29**, 255 (1957); V. I. Ritus, Ref. 18; D. W. Robinson, Helv. Phys. Acta 35, 98 (1962); F. Coester, Phys. Rev. **129**, 2816 (1963). E. J. Saletan, J. Math. Phys. **2**, 1 (1964).

[20] M. Hamermesh, Ann. Phys. **9**, 518 (1960).

$$\Gamma^{-1}p \to (E - \mathbf{p}\cdot\mathbf{v}, R^{-1}\mathbf{p})$$

under the transformation $\Gamma^{-1} = (\mathbf{v}, R)^{-1}$. Now we may let the mass go to zero while keeping unchanged the helicity state $|\mathbf{p}, \sigma\rangle$. Consider, e.g., that imparting to the particle a further velocity in the direction of $\mathbf{p}$ leaves σ unchanged. Hence the helicity states are well defined for the zero mass particle (which is naturally well known directly). Let us then compute the behavior of the representation (59) in that process.

Since

$$\Gamma p \to (E + R\mathbf{p}\cdot\mathbf{v}, R\mathbf{p}),$$

We have

$$|\Gamma p, \sigma\rangle \to |(E + R\mathbf{p}\cdot\mathbf{v}, R\mathbf{p}), \sigma\rangle.$$

and the argument of $S(\Gamma; p)$ becomes

$$(R, \mathbf{p}) = R_{\mathbf{p}}^{-1}RR_{R^{-1}\mathbf{p}}$$

which is a rotation leaving $\mathbf{e}_3$ invariant. Hence we get for the Galilei universal covering group G the true representation (the central element which is now trivial is dropped)

$$U(b, \mathbf{a}, \mathbf{v}, R)\ |\Gamma^{-1}p, \sigma\rangle$$
$$= U(b, \mathbf{a}, \mathbf{v}, R)\ |(E - \mathbf{p}\cdot\mathbf{v}, R^{-1}\mathbf{p}), \sigma\rangle$$
$$= \exp i\hbar^{-1}(Eb - \mathbf{p}\cdot\mathbf{a})\ |p\sigma')S_{\sigma'\sigma}(R, \mathbf{p})$$

with the formula (57) for $S(R, \mathbf{p})$. This is the representation of G induced by the representation $L(D)$ of its subgroup $K = T \times R_{\mathbf{p}_3}$, $R_{\mathbf{p}_3}$ being isomorphic to the two-dimensional Euclidean group in the plane $(0x_1x_2)$. Now M_3 is diagonal so that

$$D^j_{\sigma\sigma'}(R, \mathbf{p}) = \begin{bmatrix} \exp ij\varphi & & \\ & \ddots & \\ & & \exp(-ij\varphi) \end{bmatrix}$$

where φ is the angle of $(R, \mathbf{p})$. Hence the representation (59) with $M = 0$ is reduced to the $(2j + 1)$ irreducible representations $[0|p = \mathbf{p}^2, s]$ for each value of $|\mathbf{p}^2|$.

$$U(b, \mathbf{a}, \mathbf{v}, R)\ |(E - \mathbf{p}\cdot\mathbf{v}, R^{-1}\mathbf{p}), s\rangle$$
$$= \exp i[\hbar^{-1}(Eb - \mathbf{p}\cdot\mathbf{a}) - s\varphi]\ |(E, \mathbf{p}), s\rangle$$
$$(s = j \cdots -j)$$

which are nothing but the zero mass representations named class II by E. Inönü and E. Wigner. This result is to be related to a discussion of D. Korff in the case of relativistic particles. The main point of Korff's discussion is that a zero mass particle with spin s and momentum $\mathbf{p}$ can be considered as the state of momentum $\mathbf{p}$, helicity s of any massive particle of spin $|s| + k$ (where $k = 0, 1, 2 \cdots$) and the mass of which cannot be experimentally found to be different from zero. The surprising feature of this result can be emphasized by the following paradox. The spin of a zero mass particle is always parallel to the momentum. The spin of a massive particle is said to be parallel to the momentum if its projection on $\mathbf{p}$ has its maximal value. Hence, a nonzero mass particle of spin s can only have the behavior of a zero mass particle of spin s if M is very small, in contradiction with Korff's result. The answer is obviously to make more precise the concept of parallelism between spin and momentum and to distinguish between mathematical maximum and physical maximum. On the other hand, that a state $|\mathbf{p}j\rangle(j < s)$ of a particle does not tend to $|\mathbf{p}s\rangle$ when M tends to zero (or $|\mathbf{p}| \to \infty$) is seen immediately from the fact that increasing the velocity in the direction of $\mathbf{p}$ does not change the polarization.

In terms of wave functions we have

$$[U(b, \mathbf{a}, \mathbf{v}, R)\psi](E, \mathbf{p}, s)$$
$$= \exp i[\hbar^{-1}(Eb - \mathbf{p}\cdot\mathbf{a}) - s\varphi]\psi(E - \mathbf{p}\cdot\mathbf{v}, R\mathbf{p}, s).$$

As shown by A. Wightman,[4] the measure to be used in the present case is

$$d\mu(E, p) = d\omega(\mathbf{p})\, dE,$$

where $d\omega(\mathbf{p})$ is the area on the sphere of radius $|\mathbf{p}|$.

Note added in proof: The mass-zero representations of class I can also be obtained from the projective representations by letting M at S^{-1} (S = spin) go to zero in such a way that MS tends to a definite limit $\equiv$ (infinite spin representations). This and further results concerning the true representations of G will be published later.

ACKNOWLEDGMENTS

We want to express our hearty thanks to Professor E. C. G. Sudarshan who is extending to us his kind hospitality at Syracuse University. Professor Sudarshan suggested the present work and we thank him for his constant help. Professor L. O'Raifeartaigh is gratefully acknowledged for a critical reading of the manuscript and for simplifications of certain proofs. We also express our hearty thanks to Professor J. Serpe and Professor A. J. Macfarlane for their constant interest and encouragements. A Fulbright grant is gratefully acknowledged.[21]

[21] R. Stora informs us that a work analogous to ours has been performed by J. C. Guillot (unpublished).

Inequalities for the Solutions of Linear Integral Equations*

A. P. BALACHANDRAN
Physics Department, Syracuse University, Syracuse, New York
(Received 1 March 1965)

Inequalities are derived for the solutions of linear integral equations of a certain class in terms of their inhomogeneous terms and kernels. The construction of these inequalities appears to be very simple in practice as it involves only quadratures. The results presented here are, therefore, expected to be useful in the investigation of physical problems.

I. INTRODUCTION AND STATEMENT OF RESULTS

In the present note, some inequalities will be derived for the solutions of linear integral equations in an L^2-space whose kernels are subject to suitable "smallness" criteria. These were discovered in the course of an investigation into the structure of partial-wave dispersion relations[1] where it was required to find conditions under which the solution of a certain integral equation never changes sign. The proofs of these results involve only elementary considerations. Nevertheless, they do not seem to be available in the literature. As linear integral equations are of frequent occurrence in theoretical physics, and as these inequalities can be constructed very simply by quadratures from the inhomogeneous term and the kernel of the integral equation, they are expected to be useful in the study of physical problems.[2] We, therefore, believe that they merit separate publication.

We shall first summarize our conventions. The integral equation reads

$$\varphi(x) = f(x) + \int_a^b dy K(x, y)\varphi(y), \qquad (1.1)$$

where $b > a$, $f(x) \in L^2(a, b)$ and a solution $\varphi(x)$ is sought in the same space. The limits of integration are allowed to be infinite. For simplicity, it will be assumed that the interval $[a, b]$ is along the real line and that the functions in Eq. (1.1) are real. If they are complex, the equation can be rewritten as a system of coupled equations by taking its real and imaginary parts and the same considerations can be applied to this coupled system too.[3] Similarly, if the integration is along an arbitrary number of curves, it can still be converted into a real integral by introducing the parametric equations of the curves so long as they are themselves the unions of any number of differentiable curves. The resulting equation will in general involve complex functions, but we can then proceed as described above. It should also be clear from these remarks that the methods outlined here can be applied to a system of coupled integral equations.

The norm of an arbitrary vector $\psi(x)$ in $L^2(a, b)$ will be denoted by $||\psi||$. Thus,

$$||\psi||^2 = \int_a^b dx \psi^2(x). \qquad (1.2)$$

The symbol $|\psi(x)|$ will be used for the modulus of $\psi(x)$. The scalar product of any two vectors $\psi_1(x)$ and $\psi_2(x)$ will be written as (ψ_1, ψ_2):

$$(\psi_1, \psi_2) = \int_a^b dx \psi_1(x)\psi_2(x). \qquad (1.3)$$

The quantities (K_x, ψ) and $k(x)$ are defined through the equations

$$(K_x, \psi) = \int_a^b dy K(x, y)\psi(y)$$

and

$$k(x) = \left|\left[\int_a^b dy K^2(x, y)\right]^{\frac{1}{2}}\right|$$

so that, for example,

$$(k, \psi) = \int_a^b dx \left|\left[\int_a^b dy K^2(x, y)\right]^{\frac{1}{2}}\right| \psi(x),$$

* Work supported by the U. S. Atomic Energy Commission.

[1] A. P. Balachandran, "Criteria for the Solubility of Partial-Wave Dispersion Relations," Syracuse University preprint (1965) and Ann. Phys. (to be published).

[2] For example, they can be used in the study of the Lippmann–Schwinger equation in potential scattering which, it is known from the work of Coester and of Scadron, Weinberg, and Wright, can be rewritten as an integral equation with an L^2-kernel for all energies if the potential is sufficiently well-behaved. See F. Coester, Phys. Rev. **133**, B1516 (1964); and M. Scadron, S. Weinberg and J. Wright, Phys. Rev. **135**, B202 (1964). This would then lead to bounds for the T-matrix or for its norm in terms of the potential. Another application is to partial-wave dispersion relations [cf. Ref. 1].

[3] This result was pointed out by Professor E. C. G. Sudarshan.

and

$$||k||^2 = (k, k) = \int_a^b dx\, dy K^2(x, y). \quad (1.4)$$

The function $k(x)$ is, therefore, nonnegative by definition. Let us also set

$$I = [(k, |f|)^2 + (1 - ||k||^2)\, ||f||^2]^{\frac{1}{2}}, \quad (1.5)$$

where the positive root of the radical is to be taken whenever $I^2 > 0$.

The quantity $||k||^2$ is just the squared norm of K in $L^2(a, b) \times L^2(a, b)$. It should be noted that if $||k||^2$ exists, $k(x)$ also exists for almost all x by Fubini's theorem.[4]

The results of this paper can now be stated:

Let $||k||^2$ exist. Then,

(i) *If*

$$||k||^2 < 1, \quad (1.6)$$

it is true that

$$f(x) - k(x)\frac{[(k, |f|) + I]}{(1 - ||k||^2)} \leq \varphi(x) \leq f(x) + k(x)\frac{[(k, |f|) + I]}{(1 - ||k||^2)} \quad (1.7)$$

for almost all x. Also,

$$||\varphi|| \leq [(k, |f|) + I]/(1 - ||k||^2). \quad (1.8)$$

(ii) *If, instead,*

$$||k||^2 > 1 \quad (1.9)$$

and

$$I^2 > 0, \quad (1.10)$$

then

$$||\varphi|| \geq [(k, |f|) - I]/(1 - ||k||^2). \quad (1.11)$$

A few remarks are in order.

If $f(x)$ and $K(x, y)$ are continuous functions of their arguments and a solution $\varphi(x)$ is sought which is also continuous, it is clear that we need not exclude sets of zero measure in x when writing the inequality (1.7).

In deriving Eq. (1.11), no assumption is made regarding the existence or otherwise of the solutions of the homogeneous problem.

$$\varphi(x) = (K_x, \varphi). \quad (1.12)$$

Of course, if this equation has a solution, Eq. (1.1) will be consistent only if $f(x)$ satisfies certain standard criteria. When $||k||^2 < 1$, it is well known that Eq. (1.12) has no solution and this appears to be the reason why we get so much stronger statements in this case.

Now, Eq. (1.1) can be rewritten by iterating it n times. Denoting the jth iterate of $K(x, y)$ by $K^{(j)}(x, y)$, we have, with an obvious notation,

$$\varphi(x) = f(x) + \sum_{j=1}^{n-1} (K_x^{(j)}, f) + (K_x^{(n)}, \varphi) \equiv f^{(n)}(x) + (K_x^{(n)}, \varphi). \quad (1.13)$$

This is also a linear integral equation for $\varphi(x)$. Therefore,[3] (iii) *If*

$$||k^{(n)}||^2 = \int_a^b dx\, dy[K^{(n)}(x, y)]^2 \quad (1.14)$$

exists, and $f^{(n)}(x) \in L^2(a, b)$, then the results (i) *and* (ii) *are also true if $f(x)$ is replaced by $f^{(n)}(x)$ and $K(x, y)$ by $K^{(n)}(x, y)$.*

The inequalities of the form (1.7) and (1.8) for $\varphi(x)$ in terms of the nth iterates become equalities in the limit $n \to \infty$ if $||k||^2 < 1$ and $||f||^2 < \infty$. This follows from an easy application of Schwarz's inequality[5] which shows that

$$||k^{(n)}||^2 \leq ||k||^{2n} \quad (1.15a)$$

or

$$||k^{(n)}||^2 \to 0 \quad \text{as} \quad n \to \infty. \quad (1.15b)$$

By Fubini's theorem,[4]

$$k^{(n)}(x) = \left|\left[\int_a^b dy\{K^{(n)}(x, y)\}^2\right]^{\frac{1}{2}}\right| \quad (1.16)$$

thus exists for all n. Further, it vanishes for almost all x when $n \to \infty$ due to Eq. (1.15b). Schwarz's inequality[5] also gives

$$||f^{(n)}|| \leq \sum_{r=0}^{n-1} ||k||^r\, ||f||, \quad (1.17)$$

so that if $||f|| < \infty$, $||f^{(n)}||$ is bounded for all n or $f^{(n)}(x)$ converges to an element in $L^2(a, b)$ when $n \to \infty$. Therefore, on taking this limit in the analog of Eq. (1.7) for example, we have

$$\lim_{n\to\infty} f^{(n)}(x) \leq \varphi(x) \leq \lim_{n\to\infty} f^{(n)}(x) \quad (1.18a)$$

[4] See, for example, F. Riesz and B. Sz.-Nagy, *Functional Analysis* (Fredrick Ungar Publishing Company, New York, 1955), p. 83.

[5] See, for example, Ref. 4, p. 40.

or

$$\varphi(x) = \lim_{n\to\infty} f^{(n)}(x) \tag{1.18b}$$

for almost all x. Perturbation theory (which converges in the mean with our hypotheses on $||k||$ and $||f||$) leads to the same result. The convergence of the bounds to the exact solution as $n \to \infty$ implies, in particular, that there exists at least a subsequence $\{n_k\}$ such that as n increases through $\{n_k\}$, the bounds get better.

II. PROOFS

The integral equation

$$\varphi(x) = f(x) + (K_x, \varphi) \tag{1.1}$$

implies that

$$|\varphi(x)| \leq |f(x)| + |(K_x, \varphi)| \leq |f(x)| + k(x)\,||\varphi|| \tag{2.1}$$

for almost all x, where the last line follows from Schwarz's inequality.[5] Therefore, if $||k||^2$ is finite,

$$\begin{aligned} ||\varphi||^2 &= (\varphi, \varphi) \\ &= (|\varphi|, |\varphi|) \leq (|f| + k\,||\varphi||, |f| + k\,||\varphi||) \\ &= ||f||^2 + 2(k, |f|)\,||\varphi|| + ||k||^2\,||\varphi||^2 \end{aligned}$$

or

$$(1 - ||k||^2)\,||\varphi||^2 - 2(k, |f|)\,||\varphi|| - ||f||^2 \leq 0. \tag{2.2}$$

The roots of the left-hand side of this equation when regarded as a function of $||\varphi||$ are at $||\varphi||_\pm$, where

$$||\varphi||_\pm = [(k, |f|) \pm I]/[1 - ||k||^2]. \tag{2.3}$$

The expression for I is

$$I = [(k, |f|)^2 + (1 - ||k||^2)\,||f||^2]^{\frac{1}{2}}. \tag{1.5}$$

The plus and minus signs in Eq. (2.3) go together. Whenever I is real, we shall mean by I the positive root of the radical.

We can now prove the various assertions of the previous section.

Proof of (i)

As $||k||^2 < 1$, the form on the left-hand side of Eq. (2.2) is greater than zero for all sufficiently large $||\varphi||$. As it is also quadratic in $||\varphi||$, the inequality (2.2) will be satisfied only if (α) its roots $||\varphi||_\pm$ are real, and (β) $||\varphi||_- \leq ||\varphi|| \leq ||\varphi||_+$. But ($\alpha$) is no restriction as I in Eq. (1.5) is necessarily real when $||k||^2 < 1$. Also, from the same equation, we find that $I \geq (k, |f|)$, so that $||\varphi||_- \leq 0$. Therefore, since $||\varphi||$ is always nonnegative, the statement $||\varphi|| \geq ||\varphi||_-$ in (β) is also devoid of content. What is left in (β) now informs us that

$$||\varphi|| \leq ||\varphi||_+. \tag{2.4}$$

But then,

$$\begin{aligned} \varphi(x) &= f(x) + (K_x, \varphi) \qquad (1.1) \\ &\leq f(x) + |(K_x, \varphi)| \\ &\leq f(x) + k(x)\,||\varphi|| \\ &\leq f(x) + k(x)\,||\varphi||_+ \quad \text{by Eq. (2.4)} \\ &= f(x) + k(x)[(k, |f|) + I]/(1 - ||k||^2) \qquad (2.5) \end{aligned}$$

by Eq. (2.3).

Similarly,

$$\begin{aligned} \varphi(x) &\geq f(x) - |(K_x, \varphi)| \\ &\geq f(x) - k(x)\,||\varphi|| \\ &\geq f(x) - k(x)\,||\varphi||_+ \\ &= f(x) - k(x)[(k, |f|) + I]/(1 - ||k||^2), \qquad (2.6) \end{aligned}$$

where both these inequalities hold up to sets of zero measure in x.

As assertion (i) merely summarizes Eqs. (2.4), (2.5) and (2.6), it is completely proved.

Proof of (ii)

The coefficient of $||\varphi||^2$ in Eq. (2.2) is now negative so that the inequality (2.2) is satisfied for all sufficiently large or small $||\varphi||$. By assumption, I is real and positive so that the roots are also real and distinct. Thus, $||\varphi||$ must be either greater than the larger of the roots (which is $||\varphi||_-$) or less than the smaller of the roots (which is $||\varphi||_+$). That is, either

$$||\varphi|| \geq ||\varphi||_- = [(k, |f|) - I]/(1 - ||k||^2) \tag{1.11}$$

or

$$||\varphi|| \leq ||\varphi||_+ = [(k, |f|) + I]/(1 - ||k||^2). \tag{2.7}$$

But $(1 - ||k||^2)$ is negative while $(k, |f|)$ and I are positive. Therefore, the right-hand side of Eq. (2.7) is negative. As $||\varphi|| \geq 0$, this inequality is impossible and (ii) is also proved.

In conclusion, it may be observed that these results are readily modified to apply to integral equations of the type

$$\varphi(x) = f(x) + \int_a^b d\psi(y) K(x, y)\varphi(y), \tag{2.8}$$

where $\psi(y)$ is a measure. The integrations which occur in the expressions in Eqs. (1.6) to (1.11) should then be carried out with respect to $|d\psi|$. For example,

$$||\varphi||^2 = \int_a^b |d\psi(x)|\, \varphi^2(x)$$

and

$$||k||^2 = \int_a^b |d\psi(x)|\, |d\psi(y)|\, K^2(x, y). \qquad (2.9)$$

A further generalization is also possible where Schwarz's inequality in Eq. (2.1) is replaced by Hölder's inequality[5] after suitable assumptions about $\varphi(x)$, $f(x)$ and $K(x, y)$ are made. This point will not be pursued further here, however.

ACKNOWLEDGMENTS

I wish to thank Professor L. O'Raifeartaigh and Professor E. C. G. Sudarshan for many useful conversations and for suggestions on the manuscript.

Branching Rules for Simple Lie Groups*

M. L. WHIPPMAN
Physics Department, University of Pennsylvania, Philadelphia, Pennsylvania
(Received 7 December 1964)

If Γ is an irreducible representation of a group $\mathcal{G}$, and $\mathcal{H}$ is a subgroup of $\mathcal{G}$, then Γ furnishes a representation of $\mathcal{H}$ which is, in general, reducible, and the branching rules specify which irreducible representations of $\mathcal{H}$ occur in the decomposition of this representation. Branching rules are derived for various choices of $\mathcal{G}$ and $\mathcal{H}$, including most possibilities that have been discussed as higher symmetry groups.

I. INTRODUCTION

SEVERAL authors have recently discussed the possibility that the strong interactions are approximately invariant under a group $\mathcal{G}$ larger than $SU(3)$, and hence that several $SU(3)$ multiplets are contained in a single irreducible representation of $\mathcal{G}$. A problem that naturally arises in this context is that of the branching rules; that is, the determination of which $SU(3)$ multiplets are contained in a given supermultiplet of $\mathcal{G}$. More generally, we may formulate the problem in the following way. Suppose Γ is an irreducible representation of a given group $\mathcal{G}$, and that we restrict our attention to those elements of $\mathcal{G}$ which lie in a given subgroup $\mathcal{H}$. Then Γ furnishes a representation of $\mathcal{H}$ which we may denote by $\Gamma \downarrow \mathcal{H}$ and call the representation of $\mathcal{H}$ subduced by Γ.[1] In general, $\Gamma \downarrow \mathcal{H}$ will be reducible and we may write

$$\Gamma \downarrow \mathcal{H} = \sum_\alpha \nu_\alpha \Lambda_\alpha, \tag{1}$$

where the Λ_α are irreducible representations of $\mathcal{H}$. The branching rules problem consists of finding the multiplicities ν_α. For convenience we shall write $\mathcal{G} \downarrow \mathcal{H}$ for the operation of restricting $\mathcal{G}$ to $\mathcal{H}$, and when there is no possibility of confusion we shall write Eq. (1) as

$$\Gamma \to \sum \nu_\alpha \Lambda_\alpha. \tag{2}$$

This problem can always be solved by writing down the branching rule for the lowest-dimensional representations of $\mathcal{G}$ (which can generally be done by inspection) and by constructing all other representations out of direct products of the elementary representations, but this procedure is extremely tedious and it is useful to have a simple analytic method that may be applied directly to any given representation of $\mathcal{G}$ without having to solve the problem for several other representations of lower dimensionality first. Certain special cases of branching rules have been given by various authors, notably Weyl[2] and Boerner[3]. In this paper we have tried to summarize all the cases known already, as well as to discuss several others. In general, for arbitrary $\mathcal{G}$ and $\mathcal{H}$, the branching rules will be extremely complicated and in fact, if the rank of $\mathcal{G}$ is large, the rules become so complicated to state that they are of little use. For this reason, a lot of our discussion will be confined to groups of low rank where we may exploit the special local isomorphisms that obtain for some of the low-order Lie groups. Though these results do not generalize, it is, of course, the groups of low rank that are of physical interest.

We have tried to include all cases that have been suggested as possible symmetry schemes. These include $SU(4) \downarrow SU(3)$ as suggested by several authors,[4] $Sp(6) \downarrow SU(3)$[5,6] and $SO(7) \downarrow SU(3)$,[6] $SU(6) \downarrow [SU(3) \otimes SU(2)]$ as suggested by Gürsey and Radicati[7] and $SU(6) \downarrow [SU(3) \otimes SU(3)]$ as suggested by Gell-Mann.[8] It is worth emphasizing that the branching rule problem is of some interest even when $\mathcal{G}$ and $\mathcal{H}$ are not both physical symmetry groups, since certain parts of the Lagrangian may be invariant under a larger group than the full Lagrangian. For example, in the case of $SU(3)$ symmetry, the kinetic part of the Lagrangian is, in fact, invariant under $O(8)$, and though the interaction terms break this symmetry, certain cal-

* Supported by the Atomic Energy Commission.

[1] A. J. Coleman, lecture notes, Uppsala University, July 1963 (unpublished).

[2] H. Weyl, *The Theory of Groups and Quantum Mechanics* (Dover Publications, Inc., New York, 1931), 2nd ed.

[3] H. Boerner, *Representations of Groups* (North-Holland Publishing Company, Amsterdam, 1963), Chap. VIII.

[4] P. Tarjanne and V. Teplitz, Phys. Rev. Letters **11**, 447 (1963); I. S. Gerstein and M. L. Whippman; Phys. Rev. **136**, B829 (1964); B. J. Bjorken and S. L. Glashow; Phys. Letters **11**, 255 (1956); D. Amati, H. Bacry, J. Nuyts, and J. Prentki CERN preprint.

[5] H. Bacry, J. Nuyts, and L. Van Hove, CERN preprint.

[6] I. S. Gerstein and M. L. Whippman, University of Pennsylvania preprint.

[7] F. Gürsey and L. Radicati, Phys. Rev. Letters **5**, 173 (1964).

[8] M. Gell-Mann, Physics **1**, 63 (1964).

culations may be simplified by using this fact. The problem of labeling the states of a given irreducible representation of $\mathcal{G}$ may also require us to consider several different subgroups $\mathcal{H}$ in order to obtain sufficient quantum numbers. Finally, we may embed a group in a larger group as a purely mathematical device to facilitate the calculation of Clebsch–Gordan coefficients, exploiting the fact that the Clebsch–Gordan coefficients of a given group are proportional to the Clebsch–Gordan coefficients of its subgroups.[9] Thus we may derive certain relationships between $SU(3)$ and $SU(2)$ Clebsch–Gordan coefficients by subducing $SU(4)$ to both $SU(3)$ and to $SU(2) \otimes SU(2)$ or $SU(6)$ to both $SU(3) \otimes SU(3)$ and $SU(2) \otimes SU(2) \otimes SU(2)$.

In most cases we shall discuss, the results are of considerably more interest than the methods of proof, particularly as the proofs tend to be tedious and unilluminating. Hence we have summarized the main results in an appendix, and have tried to reduce the details of the proofs to the bare minimum consistent with clarity. Unfortunately, as is often true in group-theory calculations, in many cases it is comparatively easy to describe a simple procedure for obtaining the branching rules for a given representation but extremely complicated to state the rules in closed analytic form. In these cases we have tried to illustrate the method by examples.

It is worth emphasizing that a given group $\mathcal{H}$ can, in general, be embedded in $\mathcal{G}$ in several distinct ways, and the branching rules will be different for each embedding. The particular embedding we are using in each case may be established by applying the given branching rules to the representation of lowest dimension.

In the next section, we discuss the relevant properties of simple groups; our notation is summarized in Table I. In Sec. III we derive various branching rules; we summarize several important cases of these in Table II.

II. PROPERTIES OF SIMPLE LIE GROUPS

We shall only consider the nonexceptional Lie groups[10] $SU(n)$, $O(n)$ and $Sp(n)$. The irreducible single valued representations of any of these groups may be specified[11] by a partition $(\lambda) \equiv (\lambda_1\lambda_2 \cdots \lambda_k)$ of k parts with

$$\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \cdots \geq \lambda_k, \tag{3}$$

where the λ's are positive or zero integers. Strictly speaking, such a partition specifies a representation of $U(n)$ rather than of $SU(n)$, and several different partitions specify the same representation of $SU(n)$. [To be exact, the partitions $(\lambda_1 + m, \lambda_2 + m, \cdots, \lambda_n + m)$ and $(\lambda_1\lambda_2, \cdots, \lambda_n)$ specify the same representation of $SU(n)$ for all integers m.] This point is not of much importance in the following work and almost all our results apply equally well to $SU(n)$ and to $U(n)$. For $SU(n)$, $n = k$; for $Sp(n)$, $n = 2k$; and for $O(n)$, $n = 2k$ or $2k + 1$ according as n is even or odd. In the case of $O(n)$, the double-valued representations may also be specified by k numbers satisfying Eq. (3), with the λ's all integral or all half odd integral.[3]

Table I. Summary of notation.

Group	Notation for Character
$SU(n)$	$\{\lambda\}$ or $\{\lambda\}_n$
$O(n)$	$[\lambda]$ or $[\lambda]_n$
$Sp(n)$	$\langle\lambda\rangle$ or $\langle\lambda\rangle_n$

Following Murnaghan,[11] we shall denote the character of the representation (λ) of $SU(n)$ by $\{\lambda\}$ or $\{\lambda\}_n$ when we wish to emphasize the order of the group, and shall use $[\lambda]$ or $[\lambda]_n$ for the character of a representation of $O(n)$, and $\langle\lambda\rangle$ or $\langle\lambda\rangle_n$ for $Sp(n)$. It is convenient to introduce a set of auxiliary variables l_i defined by

$$l_i = \lambda_i + r_i, \tag{4}$$

where

$$\begin{aligned} r_i &= k - j && \text{for } SU(k) \\ &= k - j && \text{for } O(2k) \\ &= k - j + \tfrac{1}{2} && \text{for } O(2k+1) \\ &= k - j + 1 && \text{for } Sp(2k). \end{aligned} \tag{5}$$

The dimension of the representation (λ) is given by

$$N(\lambda) = \xi(\lambda_1\lambda_2 \cdots \lambda_k)/\xi(0, 0 \cdots, 0), \tag{6}$$

where for $SU(k)$

$$\xi(\lambda) = \prod_{i<j} (l_i - l_j),$$

for $O(2k)$

$$\xi(\lambda) = \prod_{i<j} (l_i^2 - l_j^2),$$

[9] G. Racah, lecture notes, Institute for Advanced Study, Princeton (1951).

[10] H. Weyl, *The Classical Groups* (Princeton University Press, Princeton, New Jersey, 1946).

[11] F. D. Murnaghan, *The Unitary and Rotation Groups* (Spartan Books, Washington, D. C., 1962); *The Orthogonal and Symplectic Groups* (Dublin Institute for Advanced Studies, Dublin, 1958).

TABLE II. Special cases of the branching rules. The symbols D, A, C, and I stand for defining, adjoint, coregular and identity representation respectively [C is defined by $D \otimes D = I \oplus A \oplus C$ in $Sp(n)$ and $O(n)$.] E and $\bar{F}$ are the representations {2} and {11} of $SU(n)$.

Representation of $\mathcal{G}$	D	A	C
$\mathcal{G} \downarrow \mathcal{H}$			
$SU(n) \downarrow SU(n-1)$	$D + I$	$A + D + \bar{D} + I$	
$SU(n+m) \downarrow SU(n) \otimes SU(m)$	$(D_n, I_m) + (I_n, D_m)$	$(A_n, I_m) + (I_n, A_m) + (D_n, \bar{D}_m) + (\bar{D}_n, D_m) + (I_n, I_m)$	
$SU(nm) \downarrow SU(n) \otimes SU(m)$	(D_n, D_m)	$(An, A_m) + (An, I_m) + (I_n, A_m)$	
$O(2k+1) \downarrow O(2k)$	$D + I$	$A + D$	$C + D + I$
$O(2k) \downarrow O(2k-1)$	$D + I$	$A + D$	$C + D + I$
$Sp(2k) \downarrow Sp(2k-2)$	$D + 2I$	$A + 2D + 3I$	$C + 2D + I$
$Sp(2k) \downarrow SU(k)$	$D + \bar{D}$	$A + E + \bar{E} + I$	$A + F + \bar{F}$
$O(2k) \downarrow SU(k)$	$D + \bar{D}$	$A + F + \bar{F} + I$	$A + E + \bar{E}$
$O(2k+1) \downarrow SU(k)$	$D + \bar{D} + I$	$A + F + \bar{F} + D + \bar{D} + I$	$A + E + \bar{E} + D + \bar{D} + I$

for $O(2k+1)$

$$\xi(\lambda) = \prod_{i<j} (l_i^2 - l_j^2) \prod_i l_i, \tag{7}$$

for $Sp(2k)$

$$\xi(\lambda) = \prod_{i<j} (l_i^2 - l_j^2) \prod_i l_i.$$

All characters of $SU(n)$, $O(2k+1)$ and $Sp(2k)$ are simple. The character $[\lambda]_{2k}$ of $O(2k)$ is simple if $\lambda_k = 0$; otherwise the corresponding representation of $O(2k)$ reduces to two irreducible representations of the same dimension.

III. BRANCHING RULES

A. $SU(t) \downarrow [SU(n) \otimes SU(m)]$

The simplest rule of this sort has been given by Weyl[2] for the case $SU(n) \downarrow SU(n-1)$. The rule is

$$\{\lambda\}_n \rightarrow \sum \{\lambda'\}_{n-1} \tag{8}$$

when the sum is over all (λ') satisfying

$$\lambda_1 \geq \lambda_1' \geq \lambda_2 \geq \lambda_2' \geq \cdots \geq \lambda_{n-1} \geq \lambda_{n-1}' \geq \lambda_n \tag{9}$$

There are at least two ways in which $SU(n) \otimes SU(m)$ may be embedded in some $SU(t)$, corresponding to the two types of product of representations of $SU(t)$. If $\{\lambda\}$ and $\{\mu\}$ are representations of $SU(t)$, the inner or Kronecker product of $\{\lambda\}$ and $\{\mu\}$ may be reduced in the usual way, using Young diagrams for example, to give

$$\{\lambda\}\{\mu\} = \sum m_{\lambda\mu\nu}\{\nu\}, \tag{10}$$

where

$$\sum \lambda_i + \sum \mu_i = \sum \nu_i. \tag{11}$$

Coleman[1] and Robinson[12] have shown that the $m_{\lambda\mu\nu}$ furnish the solution to the branching problem $SU(m+n) \downarrow [SU(m) \otimes SU(n)]$, where the embedding is done by considering the set of all $(m+n) \times (m+n)$ unimodular unitary matrices of the form

$$\begin{bmatrix} U_1 & O \\ O & U_2 \end{bmatrix},$$

where U_1 and U_2 are respectively $m \times m$ and $n \times n$ unimodular unitary matrices. The branching rule is then

$$\{\nu\}_{m+n} \rightarrow \sum m_{\lambda\mu\nu}\{\lambda\}_m\{\nu\}_n, \tag{12}$$

where the sum is over all partitions satisfying equation (21), and where $\{\lambda\}$ and $\{\mu\}$ are now regarded as specifying representations of $SU(m)$ and $SU(n)$ respectively. This type of embedding is considered by Gell-Mann[8] for the case $SU(6) \downarrow [SU(3) \otimes SU(3)]$.

We can, however, form a different product of representations. Let $\{\lambda\}_m$ and $\{\mu\}_n$ be representations of $SU(m)$ and $SU(n)$, respectively, subject to

$$\sum \lambda_i = \sum \mu_i = N. \tag{13}$$

We can then reduce the outer product[12] $\{\mu\} \cdot \{\lambda\}$, which is a representation of $SU(mn)$ by

$$\{\mu\}_m \cdot \{\lambda\}_n = \sum C_{\mu\lambda\nu}\{\nu\}_{mn}, \tag{14}$$

where

$$\sum \nu_i = N.$$

[12] G. de B. Robinson: Representation Theory of the Symmetric Group (The University of Toronto Press, Toronto, 1961).

The coefficients $C_{\mu\lambda\nu}$ also specify the reduction of the inner product of representations of the N-dimensional permutatiom group, and may be calculated by a method given by Robinson.[12] Coleman[1] gives a recurrence relation for the $C_{\mu\lambda\nu}$ and has tabulated them up to $N = 5$. By Frobenius' reciprocity theorem,[12] the branching rules become

$$\{\nu\}_{mn} \downarrow [SU(m) \otimes SU(n)] = \sum C_{\mu\lambda\nu}\{\mu\}_m\{\nu\}_n. \quad (15)$$

The embedding of $SU(m) \otimes SU(n)$ used here is obtained by forming the $mn \times mn$ unimodular unitary matrices which are direct products of $m \times m$ and $n \times n$ unimodular unitary matrices, and is the embedding of $SU(2) \otimes SU(3)$ in $SU(6)$ used by Gürsey and Radicati.[7]

B. $O(n) \downarrow O(n-1)$

The branching rules for this case quoted in the appendix are proved by Boerner.[3]

C. $Sp(2k) \downarrow Sp(2k-2)$

The results here are rather more complicated than for either $SU(n) \downarrow SU(n-1)$ or $O(n) \downarrow O(n-1)$, since we are going from a group of $2k \times 2k$ matrices to a group of $(2k-2) \times (2k-2)$ matrices rather than one of $(2k-1) \times (2k-1)$ matrices. The branching rules are rather involved to write down in general, and we shall only treat the cases $k = 2$ and 3. Other cases may be treated in the same way.

For the case $k = 3$, the character may be written as a function of three real angles φ_r in the form[10]

$$\langle\lambda\rangle_6 = \Xi(l_1, l_2, l_3)/\Xi(3, 2, 1), \quad (16)$$

where $\Xi(l)$ is the determinant with rth row given by

$$[\sin l_r\varphi_1 \ \sin l_r\varphi_2 \ \sin l_r\varphi_3]. \quad (17)$$

The restriction $Sp(6) \downarrow Sp(4)$ is obtained by letting $\phi_3 \to 0$. After some manipulation with determinants we find that in this limit

$$\begin{aligned}\langle\lambda\rangle_6 \to l_2^{-1}\Bigg\{&\sum_{r=1}^{l_1-1}\Bigg[\sum_{s=1}^{l_2-1} l_2(l_1-r)l_3(l_2-s)\Lambda(r,s) \\ &- \sum_{s=1}^{l_3-1} l_2(l_1-r)l_2(l_3-s)\Lambda(r,s)\Bigg] \\ &+ \sum_{r=1}^{l_2-1}\Bigg[\sum_{s=1}^{l_1-1} l_1(l_2-r)l_2(l_3-s)\Lambda(r,s) \\ &- \sum_{s=1}^{l_3-1} l_1(l_2-r)l_3(l_2-s)\Lambda(r,s)\Bigg]\Bigg\}, \end{aligned} \quad (18)$$

where

$$\Lambda(r,s) = \begin{vmatrix}\sin r\varphi_1 & \sin s\varphi_1\\ \sin r\varphi_2 & \sin s\varphi_2\end{vmatrix} \div \begin{vmatrix}\sin 2\varphi_1 & \sin\varphi_1\\ \sin 2\varphi_2 & \sin\varphi_2\end{vmatrix}. \quad (19)$$

Comparing Eqs. (17) and (19), we see that

$$\begin{aligned}\Lambda(r,s) &= \langle r-2, s-1\rangle_4 && (r > s \geq 1)\\ &= 0 && (r = s)\\ &= -\Lambda(s,r) && (r < s).\end{aligned}$$

Collecting terms we find

$$\langle\lambda\rangle_6 \to \sum \nu(\lambda')\langle\lambda'\rangle_4, \quad (20)$$

where

$$\begin{aligned}\nu(\lambda') &= (\lambda_1 - \lambda_1' + 1)(\lambda_2 - \lambda_2' + 1)(\lambda_3 + 1)\\ &\qquad (\lambda_3 \leq \lambda_2' \leq \lambda_2 \leq \lambda_1' \leq \lambda_1)\\ &= (\lambda_1 - \lambda_1' + 1)(\lambda_2 - \lambda_3 + 1)(\lambda_2' + 1)\\ &\qquad (0 \leq \lambda_2' \leq \lambda_3, \lambda_2 \leq \lambda_1' \leq \lambda_1)\\ &= (\lambda_1' - \lambda_2' + 1)(\lambda_1 - \lambda_2 + 1)(\lambda_3 + 1)\\ &\qquad (\lambda_3 \leq \lambda_2' \leq \lambda_1' < \lambda_2)\\ &= (\lambda_1' - \lambda_3 + 1)(\lambda_1' - \lambda_2 + 1)(\lambda_2' + 1)\\ &\qquad (0 \leq \lambda_2' < \lambda_3 \leq \lambda_1' < \lambda_2).\end{aligned}$$

A similar calculation shows for $Sp(4) \downarrow Sp(2)$

$$\langle\lambda\rangle_4 \to \sum \nu(\lambda')\langle\lambda'\rangle_2 \quad (21)$$

with

$$\begin{aligned}\nu(\lambda') &= (\lambda_2 + 1)(\lambda_1 - \lambda_1' + 1) \quad (\lambda_2 \leq \lambda_1' \leq \lambda_1)\\ &= (\lambda_1' + 1)(\lambda_1 - \lambda_2 + 1) \quad (0 \leq \lambda_1' < \lambda_2).\end{aligned}$$

D. $SU(n) \downarrow Sp(n)$, $SU(n) \downarrow O(n)$

These cases have been discussed by Murnaghan.[11] The results may be expressed most simply by introducing the operators ξ_p, where ξ_p is defined to be the operator which reduces the pth element of a partition by unity. That is

$$\begin{aligned}&\xi_p\{\lambda_1, \cdots, \lambda_p, \cdots, \lambda_n\}\\ &\qquad = \{\lambda_1, \cdots, \lambda_p - 1, \cdots, \lambda_n\}.\end{aligned} \quad (22)$$

Then Murnaghan has shown[11]

$$\begin{aligned}[\lambda]_n &= \prod_{\substack{p\leq q\\ 1}}^{k} (1 - \xi_q\xi_p)\{\lambda\}_n,\\ \langle\lambda\rangle_n &= \prod_{\substack{p<q\\ 1}}^{k} (1 - \xi_q\xi_p)\{\lambda\}_n,\end{aligned} \quad (23)$$

where $n = 2k$ or $2k+1$ according as it is even or odd. The inversion of these equations is complicated, and is discussed fully in Ref. (11).

E. $O(2k) \downarrow SU(k)$, $O(2k+1) \downarrow SU(k)$

The case $k = 1$ is trivial. For $k = 2$, $O(4) \downarrow SU(2)$ and $O(5) \downarrow SU(2)$ may be found by using the branching rules for $O(n) \downarrow O(n-1)$ since $SU(2)$ and $O(3)$ are locally isomorphic. When $k = 3$, $O(6) \downarrow SU(3)$ may be found by using the fact that $O(6)$ and $SU(4)$ are locally isomorphic. The problem is not completely trivial, however, since we must find the connection between the specification of a representation in terms of a four part partition $\{\mu_1\mu_2\mu_3\mu_4\}_4$ and that in terms of a three part partition $[\lambda_1\lambda_2\lambda_3]_6$. This may be done by expressing both in terms of the components of the highest weight of the representation along the simple weights. Let α_1, α_2, α_3 be the simple weights of $SU(4)$ and let M be the highest weight of the representation. Define

$$p_r = 2(M,\alpha_r)/(\alpha_r, \alpha_r). \tag{24}$$

Then Dynkin[13] shows

$$\begin{aligned} p_r &= \mu_r - \mu_{r+1} \qquad (r = 1, 2, 3), \\ p_1 &= \lambda_2 + \lambda_3, \\ p_2 &= \lambda_1 - \lambda_2, \\ p_3 &= \lambda_2 - \lambda_3. \end{aligned}$$

Hence

$$\begin{aligned} \mu_1 &= \lambda_1 + \lambda_2, \\ \mu_2 &= \lambda_1 - \lambda_3, \\ \mu_3 &= \lambda_2 - \lambda_3, \end{aligned} \tag{25}$$

and without loss of generality we can take $\mu_4 = 0$. It is important to remember when using these formulas, that $[\lambda]_6$ is not a simple character if $\lambda_3 \neq 0$. It is easy to prove that in this case we should write

$$\begin{aligned} [\lambda_1\lambda_2\lambda_3]_6 = {} & \{\lambda_1 + \lambda_2, \lambda_1 - \lambda_3, \lambda_2 - \lambda_3, 0\}_4 \\ & + \{\lambda_1 + \lambda_2, \lambda_1 + \lambda_3, \lambda_2 + \lambda_3, 0\}_4. \end{aligned} \tag{26}$$

The branching rules can now be obtained from Eq. (9). For example, consider the representation [1, 1, 1] of $SO(6)$,

$$\begin{aligned} [1,1,1]_6 = {} & \{2,0,0,0\}_4 + \{2,2,2,0\}_4 \\ & \to \{2,0,0\}_3 + \{1,0,0\}_3 + \{0,0,0\}_3 \\ & + \{2,2,0\}_3 + \{2,2,1\}_3 + \{2,2,2\}_3. \end{aligned} \tag{27}$$

The case $O(7) \downarrow SU(3)$ can now be solved via the chain $O(7) \to O(6) \to SU(3)$. Finally, we can find the branching rules for $O(8) \downarrow SU(4)$ and $O(9) \downarrow SU(4)$ by the chains $O(8) \to O(7) \to O(6)$ and $O(9) \to O(8) \to O(7) \to O(6)$ and by Eq. (25).

F. $Sp(2k) \downarrow SU(k)$

As before, the case $k = 1$ is trivial, and $k = 2$ may be solved using the isomorphism of $Sp(4)$ and $O(5)$. The case $k = 3$ may be treated by introducing an intermediate stage of $O(6)$. Of course, $O(6)$ is not a subgroup of $Sp(6)$, but it is possible to express the characters of one in terms of the characters of the other, and since the branching rules are merely algebraic relations between characters, this procedure is justified.

From Eq. (23) we may write

$$\langle\lambda\rangle_{2k} = \left[\prod_{p=1}^{k} (1 - \xi_p^2)\right]^{-1} [\lambda]_{2k} \tag{28}$$

and hence

$$\langle\lambda_1\lambda_2\lambda_3\rangle_6 = \sum [\lambda_1'\lambda_2'\lambda_3']_6, \tag{29}$$

where the sum is over all (λ') such that

$$\lambda_r - \lambda_r' \equiv 0 \pmod 2 \qquad (r = 1, 2, 3). \tag{30}$$

Note that some of the terms in this sum will not satisfy

$$\lambda_1' \geq \lambda_2' \geq \lambda_3'.$$

These terms must be brought into standard form using[12]

$$\begin{aligned} [\lambda_1'\lambda_2'\lambda_3'] &= -[\lambda_2' - 1, \lambda_1' + 1, \lambda_3'] \\ &= -[\lambda_1', \lambda_3' - 1, \lambda_2' + 1]. \end{aligned} \tag{31}$$

From this we see that $[\lambda']$ vanishes if $\lambda_2' = \lambda_1' + 1$ or $\lambda_3' = \lambda_2' + 1$. It also vanishes if $\lambda_1 \geq \lambda_2' \geq \lambda_3'$, and $\lambda_3' < 0$. This procedure is best illustrated by an example. Consider the representation $\langle 110\rangle$ of $Sp(6)$. Then

$$\begin{aligned} &\langle 110\rangle_6 \\ &= [110]_6 + [-110]_6 + [1\ -10]_6 + \cdots \\ &= [110]_6 - [000]_6 \\ &= \{2110\}_4 - \{0000\}_4 \\ &\to \{211\}_3 + \{111\}_3 + \{210\}_3 + \{110\}_3 - \{000\}_3 \\ &= \{100\}_3 + \{210\}_3 + \{110\}_3. \end{aligned}$$

That is, the 14-dimensional representation of $Sp(6)$ decomposes into representations of $SU(3)$ according to

$$14 \to 3 + 8 + \bar{3}.$$

[13] E. B. Dynkin: Amer. Math. Soc. Transl. II **6**, 245 (1957).

The case $Sp(8) \downarrow SU(4)$ may be treated similarly, by writing $\langle\lambda\rangle_8$ in terms of $[\lambda]_8$.

G. $Sp(m) \downarrow O(n)$; $O(m) \downarrow Sp(n)$

These cases can be handled analogously to the last one, using Eq. (28). That is, we write $\langle\lambda\rangle_n$ in terms of $[\lambda]_m$, or vice versa and then use the results of Sec. B or C. This method will work for any m and n (provided, of course, that these are such that the one group is a subgroup of the other) and does not depend on any isomorphisms between different groups.

APPENDIX: SUMMARY OF BRANCHING RULES

(1) $SU(n) \downarrow SU(n-1)^2$

$$\{\lambda\}_n \to \sum \{\lambda'\}_{n-1}$$

with

$$\lambda_1 \geq \lambda_1' \geq \lambda_2 \geq \lambda_2' \cdots \geq \lambda_{n-1} \geq \lambda_{n-1}' \geq \lambda_n.$$

(2) $SU(n+m) \downarrow [SU(n) \otimes SU(m)]^1$

$$\{\nu\}_{n+m} \to \sum m_{\lambda\mu\nu}\{\lambda\}_n\{\mu\}_n$$

where

$$\sum \lambda_i + \sum \mu_i = \sum \nu_i = N$$

and

$$\{\lambda\}_{n+m}\{\mu\}_{n+m} = \sum m_{\lambda\mu\nu}\{\nu\}_{n+m}.$$

The $m_{\lambda\mu\nu}$ have been tabulated by Murnaghan[14] for $N \leq 10$.

(3) $SU(nm) \downarrow [SU(n) \otimes SU(m)]$

$$\{\nu\}_{nm} \to \sum C_{\lambda\mu\nu}\{\lambda\}_n\{\mu\}_m,$$

where

$$\sum \lambda_i = \sum \mu_i = \sum \nu_i = N$$

and

$$\{\lambda\}\cdot\{\mu\} = \sum C_{\lambda\mu\nu}\{\nu\}.$$

The $C_{\lambda\mu\nu}$ have been tabulated by Coleman[1] for $N \leq 5$.

(4) $O(2k) \downarrow O(2k-1)^3$

$$[\lambda]_{2k} \to \sum [\lambda']_{2k-1}$$

where

$$\lambda_1 \geq \lambda_1' \geq \lambda_2 \geq \lambda_2' \cdots \geq \lambda_{k-1}' \geq \lambda_k,$$

and the λ_i' are integral or half-integral according to what the λ_i are.

[14] F. D. Murnaghan: Amer. J. Math. 59, 437 (1937); 60, 44 (1938).

(5) $O(2k+1) \downarrow O(2k)^3$

$$[\lambda]_{2k+1} \to \sum [\lambda']_{2k},$$

where

$$\lambda_1 \geq \lambda_1' \geq \lambda_2 \geq \cdots \geq \lambda_k \geq \lambda_k' \geq 0$$

and the λ_i' are integral or half-integral according to what the λ_i are.

(6) $Sp(6) \downarrow Sp(4)$

$$\langle\lambda\rangle_6 \to \sum \nu(\lambda')\langle\lambda'\rangle_4,$$

with

$$\begin{aligned}\nu(\lambda') &= (\lambda_1 - \lambda_1' + 1)(\lambda_2 - \lambda_2' + 1)(\lambda_3 + 1), \\ &\qquad (\lambda_3 \leq \lambda_2' \leq \lambda_2 \leq \lambda_1' \leq \lambda_1) \\ &= (\lambda_1 - \lambda_1' + 1)(\lambda_2 - \lambda_3 + 1)(\lambda_2' + 1) \\ &\qquad (0 \leq \lambda_2' \leq \lambda_3 \leq \lambda_2 \leq \lambda_1' \leq \lambda_1) \\ &= (\lambda_1' - \lambda_2' + 1)(\lambda_1 - \lambda_2 + 1)(\lambda_3 + 1) \\ &\qquad (\lambda_3 \leq \lambda_2' \leq \lambda_1' < \lambda_2) \\ &= (\lambda_1' - \lambda_3 + 1)(\lambda_1 - \lambda_2 + 1)(\lambda_2' + 1) \\ &\qquad (0 \leq \lambda_2' < \lambda_3 \leq \lambda_1' < \lambda_2).\end{aligned}$$

(7) $Sp(4) \downarrow Sp(2)$

$$\langle\lambda\rangle_4 \to \sum \nu(\lambda')\langle\lambda'\rangle_2,$$

with

$$\begin{aligned}\nu(\lambda') &= (\lambda_2 + 1)(\lambda_1 - \lambda_1' + 1) \qquad (\lambda_2 \leq \lambda_1' \leq \lambda_1) \\ &= (\lambda_1' + 1)(\lambda_1 - \lambda_2 + 1) \qquad (0 \leq \lambda_1' < \lambda_2).\end{aligned}$$

(8) $O(6) \downarrow SU(3)$

$$[\lambda]_6 \to \sum \{\lambda'\}_3 + \sum \{\lambda''\}_3,$$

where (a) if $\lambda_3 = 0$,

$$\lambda_1 + \lambda_2 \geq \lambda_1'' \geq \lambda_1 \geq \lambda_2' \geq \lambda_2 \geq \lambda_3' \geq 0$$

and second term does not appear,

(b) if $\lambda_3 \neq 0$,

$$\lambda_1 + \lambda_2 \geq \lambda_1' \geq \lambda_1 - \lambda_3 \geq \lambda_2' \geq \lambda_2 - \lambda_3 \geq \lambda_3' \geq 0,$$

$$\lambda_1 + \lambda_2 \geq \lambda_1'' \geq \lambda_1 + \lambda_3 \geq \lambda_2'' \geq \lambda_2 + \lambda_3 \geq \lambda_3'' \geq 0.$$

(9) $Sp(6) \downarrow SU(3)$

Use

$$\langle\lambda\rangle_6 = \sum [\lambda']_6,$$

where

$$\lambda_r \equiv \lambda_r' \pmod 2 \qquad (r = 1, 2, 3)$$

and above result.

Recursion Relations for the Wigner Coefficients of Unitary Groups

T. A. Brody, M. Moshinsky, and I. Renero
Instituto de Física, Universidad de México, México, D. F.
(Received 3 February 1965)

The polynomials in the components of a set of n-dimensional vectors that form a basis for an irreducible representation of $S\mathfrak{U}_n$ are shown to be part of the basis of the group $\mathbf{U}_{nr}$, in which the subgroup $\mathfrak{U}_n \times \mathbf{U}_r$ is explicitly reduced and $r \geq n-1$. Using this result, the concept of auxiliary Wigner coefficient is introduced, for which the problem of multiplicity does not arise and the phase convention is related to that of Gel'fand and Zetlin; recursion relations for this auxiliary coefficient are obtained in a straightforward way, and the connection between it and the ordinary Wigner coefficient is shown to be simple. The recursion relations are being programmed for an electronic computer to allow the systematic evaluation of the Wigner coefficients of $S\mathfrak{U}_3$ and $S\mathfrak{U}_4$.

1. INTRODUCTION

In a recent publication[1], to be referred to here as I, one of the authors (M.M.) gave a complete discussion of the Wigner coefficients of the unitary groups. These coefficients were derived as scalar products of polynomials in creation operators, polynomials which formed bases for the irreducible representations (BIR's) of the unitary groups. The technique for determining these polynomials, including the resolution of the multiplicity problem, was given in I.

The general algebraic or numerical evaluation of these scalar products is laborious. The recursion relations derived from them in the present paper will permit the systematic calculation of numerical values for the coefficients.

Specific coefficients of the unitary groups[2] and of $S\mathfrak{U}_3$ in particular[3-6] have been extensively discussed in the literature, and tables of those used in strong-interaction applications are available[7]. However, elementary-particle physics and nuclear-structure problems[8-10] require a much wider set of Wigner coefficients. For this reason, two of the authors (T.A.B. and I.R.) are programming the recursion relations obtained in the present paper and plan to make the program or partial tables available to those interested.

Besides this pragmatic purpose, it will be seen that the concept of auxiliary Wigner coefficient introduced in this paper allows one to formulate a characterization of the multiplicity problem which differs from some that have been proposed recently[1,11] and seems to be applicable not only to the unitary groups but also to other semisimple compact Lie groups such as R_n and Sp_{2n}.

2. THE BIR OF $S\mathfrak{U}_n$ AS PART OF THE BIR OF $\mathbf{U}_{nr} \supset \mathfrak{U}_n \times \mathbf{U}_r$

In I, the BIR of $S\mathfrak{U}_n$ was given in terms of polynomials in the components of the n-dimensional vectors $a^+_{\mu s}$ with $\mu = 1 \cdots n$ the component index and $s = 1 \cdots r$ the vector index. The general IR of $S\mathfrak{U}_n$ requires only $n-1$ vectors; to build up Kronecker products of the BIR's, however, $r > n-1$ is convenient. The $a^+_{\mu s}$ are Bose creation operators whose properties were described in I. From them the operators

$$\mathbf{C}^{\nu t}_{\mu s} = a^+_{\mu s} a^{\nu t}, \quad \mathbf{C}^{\nu}_{\mu} = \sum_s \mathbf{C}^{\nu s}_{\mu s}, \quad \mathbf{C}^{t}_{s} = \sum_\mu \mathbf{C}^{\mu s}_{\mu t} \tag{2.1}$$

are constructed,[9] which are seen to be, respectively, the generators of the unitary groups $\mathbf{U}_{nr}$, $\mathfrak{U}_n$, $\mathbf{U}_r$. As was seen in I, $a^{\mu s}$ may be defined as $\partial/\partial a^+_{\mu s}$.

With respect to $\mathbf{U}_{nr}$, $a^+_{\mu s}$ corresponds to a single vector of dimension nr and so the set of all linearly independent homogeneous polynomials of degree N in the $a^+_{\mu s}$ forms the BIR of the completely symmetric representation of $\mathbf{U}_{nr}$, characterized by $[N]$. This set contains the BIR of $\mathfrak{U}_n \times \mathbf{U}_r \subset \mathbf{U}_{nr}$ to which belong the homogeneous polynomials P of degree N that

* Work supported by the Comisión Nacional de Energía Nuclear, México.

[1] M. Moshinsky, J. Math. Phys. 4, 1128 (1963).
[2] G. E. Baird and L. C. Biedenharn, J. Math. Phys. 4, 1449 (1963).
[3] M. Moshinsky, Rev. Mod. Phys. 34, 813 (1962).
[4] D. Lurié and A. J. Macfarlane, J. Math. Phys. 5, 565 (1964).
[5] A. R. Edmonds, Proc. Roy. Soc. (London) A268, 567 (1962). The problem of multiplicity is not discussed in this article.
[6] K. T. Hecht has used recursion relations of the Wigner coefficients of $S\mathfrak{U}_3$ to derive particular coefficients. We are indebted for a copy of his paper prior to publication.
[7] Y. Dothan and H. Harari, Report IA-777, Rehovoth, Israel, 1962; H. Goldberg, Report IA-834, Rehovoth, Israel, 1963; P. McNamee and F. Chilton, Rev. Mod. Phys. 36, 1005 (1964).
[8] J. P. Elliott, Proc. Roy. Soc. (London) A245, 128 (1958).
[9] M. Moshinsky, "Group Theory and the Many-Body Problem," to appear in *Physics of Many-Particle Systems*, edited by E. Meeron (Gordon and Breach, Inc., New York, to be published).
[10] J. Flores, E. Chacón, P. A. Mello, and M. de Llano, Nuclear Phys. (to be published).
[11] G. E. Baird and L. C. Biedenharn, J. Math. Phys. 5, 682 (1965).

are of highest weight in $\mathfrak{U}_n$ and U_r and therefore satisfy

$$\mathfrak{C}^{\mu}_{\mu}P = h_{\mu}P, \quad \mathfrak{C}^{\nu}_{\mu}P = 0, \quad \mu < \nu;$$
$$\mathsf{C}^{s}_{s}P = k_{s}P, \quad \mathsf{C}^{t}_{s}P = 0, \quad s < t. \tag{2.2}$$

It was shown in I that there are solutions of (2.2) for a partition $[h_1 \cdots h_n]$ of N if and only if $k_s = h_s$, $1 \leq s \leq n$; $k_s = 0$, $n < s$; the solutions are unique and are given by

$$P = A(h_1 \cdots h_n)(\Delta^{1}_{1})^{h_1-h_2}(\Delta^{12}_{12})^{h_2-h_3} \cdots (\Delta^{12\cdots n}_{12\cdots n})^{h_n}, \tag{2.3}$$

where

$$A(h_1 \cdots h_n) = \left[\prod_{j=2}^{n}\prod_{i=1}^{j-1}(h_i - h_j + j - i)\right]^{\frac{1}{2}} \times \left[\prod_{i=1}^{n}(h_i + n - i)!\right]^{-\frac{1}{2}} \tag{2.4}$$

and

$$\Delta^{s_1\cdots s_j}_{\mu_1\cdots\mu_j} = \sum_{\mathfrak{p}}(-1)^{\mathfrak{p}}\mathfrak{p}a^{+}_{\mu_1 s_1}\cdots a^{+}_{\mu_j s_j} \tag{2.5}$$

with $\mathfrak{p}$ standing for all permutations of $s_1, s_2, \cdots s_j$.

The remainder of the BIR of U_{nr} classified according to the canonical chains

$$\mathsf{U}_{nr} \supset \mathfrak{U}_n \times \mathsf{U}_r, \qquad \begin{matrix}\mathfrak{U}_n \supset \mathfrak{U}_{n-1} \supset \cdots \mathfrak{U}_1 \\ \mathsf{U}_r \supset \mathsf{U}_{r-1} \supset \cdots \mathsf{U}_1\end{matrix} \tag{2.6}$$

is obtained by applying the lowering operators

$$\mathfrak{L}^{\mu}_{\nu},\ \nu > \mu = 1 \cdots n-1;\ \mathsf{L}^{s}_{t},\ t > s = 1 \cdots r-1,$$

introduced by Nagel and Moshinsky[12]. Using the notation of Gel'fand and Zetlin[13], the resulting polynomials will be

$$P(h_{ij}, k_{lm}) \equiv P\begin{pmatrix} h_{1n}\ h_{2n} \cdots h_{nn} & h_{1n} \cdots h_{nn}\ 0\ 0 \cdots 0 \\ h_{1n-1} \cdots h_{n-1n-1}; & k_{1r-1} \cdots k_{nr-1} 0 \cdots 0 \\ \cdots & \cdots \\ h_{11} & k_{11} \end{pmatrix} \tag{2.7}$$

where the h_{ij}, k_{lm} satisfy the usual inequalities[14]

$$h_{i+1j} \geq h_{i+1j-1} \geq h_{ij}, \qquad 1 \leq i \leq j \leq n,$$
$$k_{l+1m} \geq k_{l+1m-1} \geq k_{lm}, \qquad 1 \leq l \leq m \leq r. \tag{2.8}$$

In the notation of (2.3), $h_{pn} = h_p$.

Put $r = 2n - 2$ and consider the ploynomials (2.7) which are a BIR for $S\mathfrak{U}_n$, i.e., with $h_{nn} = 0$; those which are of highest weight in U_{2n-2} must satisfy

$$\mathsf{C}^{s}_{s}P = k_{s2n-2}P; \qquad \mathsf{C}^{t}_{s}P = 0, \quad t > s = 1 \cdots 2n-3. \tag{2.9}$$

It was seen in I that (2.9) has solutions only for $k_{s,2n-2} = h_{sn}$ for $1 \leq s \leq n-1$, $k_{s2n-2} = 0$ for $n \leq s \leq 2n-2$. These solutions therefore depend only on the first $n-1$ vectors and are characterized by (2.9) in the same way as the first BIR of $S\mathfrak{U}_n$ discussed in I. Since these polynomials are of highest weight in U_{2n-2}, $k_{lm} = k_{l2n-2}$ for any m, and the BIR of $S\mathfrak{U}_n$ may be written

$$P_{\max}(h_{ij}) \equiv P(h_{ij}, k_{lm} = k_{l2n-2} = h_l). \tag{2.10}$$

Another set of polynomials that form a BIR of $S\mathfrak{U}_n$ in terms of the components of an independent set of vectors $s = n \cdots 2n-2$ may be obtained by taking the polynomials (2.7), again with $h_{nn} = 0$, which are of lowest weight in U_{2n-2} and thus satisfy

$$\mathsf{C}^{s}_{s}P = k_{s2n-2}P; \qquad \mathsf{C}^{t}_{t}P = 0, \quad t > s = 1 \cdots 2n-3. \tag{2.11}$$

The equations (2.11) result from (2.9) by means of the permutation of the indices

$$\begin{pmatrix} 1 & 2 & \cdots & 2n-3 & 2n-2 \\ 2n-2 & 2n-3 & & 2 & 1 \end{pmatrix} \tag{2.12}$$

and thus they have solutions only when

$$k_{s,2n-2} = 0, \qquad 1 \leq s \leq n-1;$$
$$k_{s,2n-2} = h_{2n-1-s,n}, \qquad n \leq s \leq 2n-2. \tag{2.13}$$

These solutions therefore depend only on the components of the last $n-1$ vectors and correspond to the second BIR of $S\mathfrak{U}_n$ discussed in I. They may be written, in analogy to (2.10),

$$P_{\min}(h_{ij}) \equiv P(h_{ij}, k_{m-l+1,m} = k_{2n-1-l,2n-2} = h_{ln}). \tag{2.14}$$

Because of (2.11), $P_{\min}$ is of minimum weight in U_{2n-2}, so that the k's give the representations of U_{2n-2} and its subgroups in reverse order; thus $[k_{2n-2,2n-1} \cdots k_{n,2n}0 \cdots 0]$ is that of U_{2n-2}.

As was seen in I, the Wigner coefficients of $S\mathfrak{U}_n$ may be expressed as scalar products of $P_{\max}(h'_{ij})P_{\min}(h''_{ij})$ with one of the polynomials of the BIR of $\mathfrak{U}_n \times \mathsf{U}_{2n-2}$ if $\mathfrak{U}_n$ is characterised by the canonical chain (2.6), U_{2n-2}, however, by $\mathsf{U}_{2n-2} \supset \mathsf{U}_{n-1} + \mathsf{U}_{n-1}$. Such a polynomial will be denoted by

[12] J. Nagel and M. Moshinsky, J. Math. Phys. **6**, 682 (1965) **14**, 29 (1965); Rev. Mexicana Fís.

[13] I. M. Gel'fand and M. L. Zetlin, Doklady Akad. Nauk SSSR **71**, 825 (1950).

[14] H. Weyl, *The Theory of Groups and Quantum Mechanics* (Dover Publications, Inc., New York, 1931), p. 391.

$$\mathcal{P}\begin{bmatrix} h_{ij} & h_{in}\ 0 & \chi \\ & \bar{h}'_{ij}\ \bar{h}''_{ij} & \end{bmatrix} \equiv \mathcal{P}\left[\begin{array}{ll} h_{1n}\cdots h_{nn} & h_{1n}\cdots h_{nn}\ 0\cdots 0 \\ h_{1n-1}\cdots h_{n-1n-1} & \bar{h}'_{1n-1}\cdots \bar{h}'_{n-1n-1}\quad \bar{h}''_{1n-1}\cdots \bar{h}''_{n-1n-1} \\ \cdots\ ; & \bar{h}'_{1n-2}\cdots \qquad\qquad \bar{h}''_{1n-2}\cdots \quad \chi \\ h_{11} & \cdots \qquad\qquad \cdots \\ & \bar{h}'_{11} \qquad\qquad \bar{h}''_{11} \end{array}\right], \tag{2.15}$$

where the IR of U_{2n-2} is given by $[h_{1n}\cdots h_{nn}\,0\cdots 0]$ and those of its subgroups U_{n-1} and their canonical chain of subgroups are characterized by $\bar{h}'_{ij}$, $\bar{h}''_{ij}$, $1 \le i \le j \le n-1$. The rows of the BIR of U_{2n-2} are not completely defined by these subgroups, so that (see I) the eigenvalues χ of another $\frac{1}{2}n(n-3)+1$ operators X are required to complete the definition.

Both (2.7) and (2.15) are BIR of $\mathfrak{U}_n \times \mathsf{U}_{2n-2}$, and so (2.15) may be expanded in terms of (2.7). The coefficients are evidently independent of h_{ij}, $1 \le i \le j \le n-1$, and since the first subgroup U_{n-1} coincides in both chains, they will be different from zero only if $k_{ij} = \bar{h}'_{ij}$, $1 \le i \le j \le n-1$, and moreover will be independent of the $\bar{h}'_{ij}$ with $j < n-1$. The polynomials (2.15) are completely characterized by the eigenvalues of the Casimir operators[15-17] G_l of U_{n-1} and of the operators[1] X, all of which are functions only of the C^s_s and so commute with the $\mathcal{C}^r_\mu$; in consequence the matrix elements of the G_l and X with respect to the states formed by applying the polynomials (2.7) to the ground state $|0\rangle$ (defined by $a^{\mu s}|0\rangle = 0$ for all μ, s) will not depend on the h_{ij} for $j < n$. They need thus only be found for the case $h_{ij} = h_{in}$ for all j, i.e., for the Gel'fand states[2,12,13] of U_{2n-2}; they may be derived from the matrix elements of C^t_s, calculated by Gel'fand and Zetlin. The matrices so obtained may be diagonalised simultaneously, since the G_l and X operators commute; the components of the eigenvectors so found are the coefficients of the expansion of the $\mathcal{P}$ of (2.15) in terms of the P of (2.7).

Once these expansion coefficients are obtained, the full Wigner coefficient of $S\mathfrak{U}_n$, written as

$$\langle 0|\, P^\dagger_{\max}(h'_{ij})P^\dagger_{\min}(h''_{ij}) \times \mathcal{P}\begin{bmatrix} h_{ij}; & h_{in}\ 0 & \chi \\ & \bar{h}'_{ij} = h'_{in}\bar{h}''_{ij} = h''_{n-i+j-1,n} & \end{bmatrix} |0\rangle \tag{2.16}$$

may be found from the scalar product

$$\langle 0|\, P^\dagger_{\max}(h'_{in})P^\dagger_{\min}(h''_{n-i+j-1,n})P(h_{ij}, k_{lm})\,|0\rangle \tag{2.17}$$

in which

$$k_{l2n-2} = \begin{cases} h_{ln}, & 1 \le l \le n-1 \\ 0 & n \le l \le 2n-2 \end{cases};$$

$$k_{ij} = h'_{in}, \qquad 1 \le i \le j \le n-1. \tag{2.18}$$

Here $h'_{nn} = h''_{nn} = 0$, the symbol $\dagger$ indicates the Hermitian conjugate, and the representation $[h_{pn}]$ is derived from $[h'_{pn}]$ and $[h''_{pn}]$ by the usual Littlewood rules.[18]

The scalar product (2.17) will be called the *auxiliary Wigner coefficient of* $S\mathfrak{U}_n$ (AW_n); it differs from (2.17) in that the canonical chain of subgroups (2.6) characterizes the polynomial $P(h_{ij}, k_{lm})$, whereas the chain $\mathsf{U}_{2n-2} \supset \mathsf{U}_{n-1} \dotplus \mathsf{U}_{n-1}$ was used for $\mathcal{P}$, so that the χ are needed to resolve the multiplicity problem. The canonical structure of (2.18) makes it possible to find recursion relations for it, as will be done in the following sections. The case of $S\mathfrak{U}_3$ is treated first, for the sake of simplicity and because of its intrinsic importance.

3. RECURSION RELATIONS FOR THE AW_3

In this section the notation to be used is

$$\begin{aligned} k_{i4} &= h_{i3} = h_i, & k_{44} &= 0 \\ h_{i2} &= q_i & k_{i3} &= u_i \end{aligned} \tag{3.1}$$

with a similar one for the primed and double-primed h's.

The three polynomials in (2.17), written for the case $n = 3$, each form a BIR for the subgroup $S\mathfrak{U}_2$ of $S\mathfrak{U}_3$, and therefore the Wigner–Eckart theorem shows that the scalar product will depend on h'_{11} and h''_{11} through the ordinary Wigner coefficient

$$\langle t't''\tau'\tau'' \mid t\tau\rangle:$$

$$\begin{aligned} t' &= \tfrac{1}{2}(q'_1 - q'_2), & t'' &= \tfrac{1}{2}(q''_1 - q''_2), \\ \tau' &= h'_{11} - \tfrac{1}{2}(q'_1 + q'_2), & \tau'' &= h''_{11} - \tfrac{1}{2}(q''_1 + q''_2), \\ t &= \tfrac{1}{2}(q_1 - q_2), & \tau &= h_{11} - \tfrac{1}{2}(q_1 + q_2). \end{aligned} \tag{3.2}$$

[15] V. Bargmann and M. Moshinsky, Nucl. Phys. 18, 697 (1960).

[16] G. Racah, "Group Theory and Spectroscopy," Lecture notes, Institute of Advanced Study, Princeton, New Jersey (1951).

[17] The operators G_l are defined by

$$G_l = \sum_{s_1\cdots s_l=n}^{2n-2} \mathsf{C}^{s_1}_{s_1}\mathsf{C}^{s_2}_{s_2}\cdots\mathsf{C}^{s_l}_{s_l}.$$

[18] E. E. Littlewood, *The Theory of Group Characters* (Oxford University Press, Oxford, England, 1940), p. 40.

It is thus sufficient to determine the value of (2.17) for appropriately chosen values of h'_{11} and h''_{11}; other values are then derived by using (3.2). It is easily seen that the choice

$$h'_{11} = q'_1 \qquad h''_{11} = q'_2 \tag{3.3}$$

satisfies in all cases the inequalities (2.8). It will also be seen that the Littlewood rules force the values

$$k_{i2} = h'_i \qquad k_{11} = h'_1 \tag{3.4}$$

but that the values of the u_i are not so determined. The AW_3 to be evaluated are thus

$$\langle 0|\, P^{\dagger}_{\max} \begin{bmatrix} h'_1 h'_2 0 \\ q'_1 q'_2 \\ q'_1 \end{bmatrix} P^{\dagger}_{\min} \begin{bmatrix} h''_1 h''_2 0 \\ q''_1 q''_2 \\ q''_2 \end{bmatrix} P \begin{bmatrix} h_1 h_2 h_3 & & h_1 h_2 h_3 0 \\ q_1 q_2 & ; & u_1 u_2 u_3 \\ q'_1 + q''_2 & & h'_1 h'_2 \\ & & h'_1 \end{bmatrix} |0\rangle$$

$$\equiv \left\langle \begin{matrix} h'_1 h'_2 \\ q'_1 q'_2 \end{matrix}, \begin{matrix} h''_1 h''_2 \\ q''_1 q''_2 \end{matrix} \middle| \begin{matrix} h_1 h_2 h_3 \\ q_1 q_2 \end{matrix}; \begin{matrix} u_1 u_2 u_3 \end{matrix} \right\rangle . \tag{3.5}$$

In the abbreviated notation only nonrepeated indices are retained; thus in the ket, the first row is common to both groups, except for a zero, and is written only once.

In (3.5), $P_{\max}$ and $P_{\min}$ are now not only of highest and lowest weight in $\mathbf{U}_4$, but also in the subgroup $\mathfrak{U}_2$ of $S\mathfrak{U}_3$. $P_{\max}$ was obtained in Eq. (A18) of Ref. 3. $P_{\min}$ may then be obtained by using the permutation (2.12) for both $\mathbf{U}_4$ and $\mathfrak{U}_2$, i.e., $\left(\begin{smallmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{smallmatrix}\right)$ for the s indices and $\left(\begin{smallmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{smallmatrix}\right)$ for the μ's, so that

$$P_{\min} = A \begin{bmatrix} h''_1 h''_2 0 \\ q''_1 q''_2 \end{bmatrix}$$

$$\times (\Delta_2^4)^{q_1'' - h_2''} (\Delta_3^4)^{h_1'' - q_1''} (\Delta_{12}^{34})^{q_2'' - h_2''} (\Delta_{23}^{34})^{h_2'' - q_2''} . \tag{3.6}$$

To obtain the correct normalisation, however, it is necessary to lower the weight from $P_{\max}$ by means of the lowering operators[12] mentioned above; it turns out that A in (3.6) is positive if the indices on the Δ's are arranged in the natural order and that for all polynomials it takes the form[19]

$$A \begin{bmatrix} h_1 & h_2 & h_3 \\ & q_1 & q_2 \end{bmatrix}$$

$$= \left[\frac{(h_1 - h_3 + 2)!\,(h_2 - h_3 + 1)!\,(h_1 - h_2 + 1)!\,(q_1 - q_2 + 1)!}{(h_1 + 2)!\,(h_2 + 1)!\,h_3!\,(q_2 - h_3)!\,(q_1 - h_3 + 1)!\,(h_2 - q_2)!\,(h_1 - q_2 + 1)!\,(q_1 - h_2)!\,(h_1 - q_1)!} \right]^{\frac{1}{2}} . \tag{3.7}$$

From (3.6) it is clear that one can write

$$A^{-1} \begin{bmatrix} h''_1 & h''_2 & 0 \\ & q''_1 & q''_2 \end{bmatrix} P_{\min} \begin{bmatrix} h''_1 & h''_2 & 0 \\ & q''_1 & q''_2 \end{bmatrix} = \Delta_{23}^{34} A^{-1} \begin{bmatrix} h''_1 - 1 & h''_2 - 1 & 0 \\ & q''_1 - 1 & q''_2 \end{bmatrix} P_{\min} \begin{bmatrix} h''_1 - 1 & h''_2 - 1 & 0 \\ & q''_1 - 1 & q''_2 \end{bmatrix}$$

$$= \Delta_{12}^{34} A^{-1} \begin{bmatrix} h''_1 - 1 & h''_2 - 1 & 0 \\ & q''_1 - 1 & q''_2 - 1 \end{bmatrix} P_{\min} \begin{bmatrix} h''_1 - 1 & h''_2 - 1 & 0 \\ & q''_1 - 1 & q''_2 - 1 \end{bmatrix} = \Delta_3^4 A^{-1} \begin{bmatrix} h''_1 - 1 & h''_2 & 0 \\ & q''_1 & q''_2 \end{bmatrix} P_{\min} \begin{bmatrix} h''_1 - 1 & h''_2 & 0 \\ & q''_1 & q''_2 \end{bmatrix}$$

$$= \Delta_2^4 A^{-1} \begin{bmatrix} h''_1 - 1 & h''_2 & 0 \\ & q''_1 - 1 & q''_2 \end{bmatrix} P_{\min} \begin{bmatrix} h''_1 - 1 & h''_2 & 0 \\ & q''_1 - 1 & q''_2 \end{bmatrix} . \tag{3.8}$$

The four forms of (3.8) yield the four possible recursion relations for the AW_3. Substituting, for instance, the first form in (3.5) and making use of the completeness of the P's,

$$\left\langle \begin{matrix} h'_1 & h'_2 \\ q'_1 & q'_2 \end{matrix}, \begin{matrix} h''_1 & h''_2 \\ q''_1 & q''_2 \end{matrix} \middle| \begin{matrix} h_1 & h_2 & h_3 \\ & q_1 & q_2 \end{matrix}; \begin{matrix} u_1 & u_2 & u_3 \end{matrix} \right\rangle = \sum_{\substack{\bar{h}_i \bar{q}_i \bar{r}_1 \\ \bar{u}_i}} \left\langle \begin{matrix} h'_1 & h'_2 & h''_1 - 1 & h''_2 - 1 \\ q'_1 & q'_2 & q''_1 - 1 & q''_2 \end{matrix} \middle| \begin{matrix} \bar{h}_1 & \bar{h}_2 & \bar{h}_3 \\ & \bar{q}_1 & \bar{q}_2 \end{matrix}; \begin{matrix} \bar{u}_1 & \bar{u}_2 & \bar{u}_3 \end{matrix} \right\rangle$$

$$\times \frac{A \begin{bmatrix} h''_1 & h''_2 & 0 \\ & q''_1 & q''_2 \end{bmatrix}}{A \begin{bmatrix} h''_1 - 1 & h''_2 - 1 & 0 \\ & q''_1 - 1 & q''_2 \end{bmatrix}} \left\langle \begin{matrix} \bar{h}_1 & \bar{h}_2 & \bar{h}_3 & \bar{u}_1 & \bar{u}_2 & \bar{u}_3 \\ & \bar{q}_1 & \bar{q}_2 & h'_1 & h'_2 & \\ & & \bar{r}_1 & & h'_1 & \end{matrix} \middle| (\Delta_{23}^{34})^{\dagger} \middle| \begin{matrix} h_1 & h_2 & h_3 & u_1 & u_2 & u_3 \\ & q_1 & q_2 & h'_1 & h'_2 & \\ & q'_1 + q''_2 & & h'_1 & & \end{matrix} \right\rangle \tag{3.9}$$

[19] M. Moshinsky, T. A. Brody, and I. Renero, Rev. Mexicana Fís. (to be published).

where

$$\left\langle \begin{vmatrix} \bar{h}_i & \bar{u}_i \\ \bar{q}_i & h_i' \\ \bar{r}_1 & h_1' \end{vmatrix} (\Delta_{23}^{34})^\dagger \begin{vmatrix} h_i & u_i \\ q_i & h_i' \\ q_1'+q_2'' & h_1' \end{vmatrix} \right\rangle$$

$$\equiv \langle 0|\, P^\dagger \begin{bmatrix} h_i & h_i & 0 \\ q_i & u_i & \\ q_1'+q_2'' & h_i' & \\ & h_1' & \end{bmatrix} \Delta_{23}^{34} P \begin{bmatrix} \bar{h}_i & \bar{h}_i & 0 \\ \bar{q}_i & \bar{u}_i & \\ \bar{r}_1 & h_i' & \\ & h_1' & \end{bmatrix} |0\rangle$$

$$= \left\langle \begin{vmatrix} h_i & u_i \\ q_i & h_i' \\ q_1'+q_2'' & h_1' \end{vmatrix} \Delta_{23}^{34} \begin{vmatrix} \bar{h}_i & \bar{u}_i \\ \bar{q}_i & h_i' \\ \bar{r}_1 & h_1' \end{vmatrix} \right\rangle. \qquad (3.10)$$

The last relation holds because the matrices are real.[12]

The calculation for the matrix elements of Δ_2^4, Δ_3^4, Δ_{23}^{34} and Δ_{12}^{34} is straightforward if laborious; it will be described elsewhere.[19] The results will be found in the Appendix; on substituting them into (3.9) and similar equations, the required recursion relations for the AW_3 are obtained.

If the recursion relations using (A5) or (A7) are used h_2'' times, one arrives at an AW_3 with a single row, namely $[h_1'' - h_2'' 0]$; for this a closed expression is available[3] which has been programmed for computer evaluation by one of us (T.A.B.). It can also be obtained by now applying (A1) or (A2) $h_1'' - h_2''$ times. By the time this paper appears it is hoped that a computer program exists for carrying out the recursions.

In order to find from the AW_3 the Wigner coefficients of $S\mathfrak{U}_3$, it is merely necessary to carry out the procedure described above; for the present case, this implies diagonalizing the matrix elements[1]

$$\langle \bar{k}_{lm}|\, G_2\, |k_{lm}\rangle, \qquad \langle \bar{k}_{lm}|\, \mathrm{X}\, |k_{lm}\rangle, \qquad (3.11)$$

where

$$G_2 = \sum_{s,t=3}^{4} \mathsf{C}_s^t \mathsf{C}_t^s, \qquad \mathrm{X} = \sum_{s,t,w=1}^{2} \mathsf{C}_{s+2}^t \mathsf{C}_t^w \mathsf{C}_w^{s+2}$$

and $1 \leq l \leq m \leq 4$. These are derived trivially from the matrix elements of the C_s^t obtained by Gel'fand and Zetlin.[13]

It should be noted that the Wigner coefficients obtained in this manner are not identical to the isoscalar factors defined by Edmonds;[5] these would have been obtained if instead of (3.3) the choice $h_{11}' = q_1'$, $h_{11}'' = q_1''$ had been made and a factor

$$\langle t'\ t'\ t''\ -t''\,|\,t\ t'\ -t''\rangle = \left[\frac{(q_1-q_2+1)(q_1'-q_2')!\,(q_1''-q_2')!}{(q_1'+q_1''-q_1)!\,(q_1'+q_1''-q_2+1)!}\right]^{\frac{1}{2}}$$

introduced.

The form (3.6) of $P_{\min}$ is obtained by using the lowering operators of Nagel and Moshinsky[12] both in U_4 and $\mathfrak{U}_3$; since the normalization for these was chosen in such a way that the Gel'fand states obtained by means of them give matrix elements of the generators of U_4 and $\mathfrak{U}_3$ which coincide with those given by Gel'fand and Zetlin,[13] the phase convention used here for the AW_3 is the same as that for Wigner coefficients derived from these generator matrix elements.[2] The ordinary Wigner coefficient will, however, still contain an undefined phase which arises in the diagonalization of the matrices (3.11).

From the orthonormality property of the ordinary Wigner coefficient it may be deduced, using the fact that $P_{\max}$, $P_{\min}$ and P in (3.5) form complete bases for polynomials in the first two, last two, and all four vectors, respectively, that the AW_3 obey the rule

$$\sum_{h_i'' q_i' q_i''} \left\langle \begin{matrix} h_1' & h_2', & h_1'' & h_2'' \\ q_1' & q_2' & q_1'' & q_2'' \end{matrix} \middle| \begin{matrix} h_1 & h_2 & h_3 \\ q_1 & q_2 \end{matrix} ; u_1\ u_2\ u_3 \right\rangle$$

$$\times \left\langle \begin{matrix} h_1' & h_2', & h_1'' & h_2'' \\ q_1' & q_2' & q_1'' & q_2'' \end{matrix} \middle| \begin{matrix} \bar{h}_1 & \bar{h}_2 & \bar{h}_3 \\ \bar{q}_1 & \bar{q}_2 \end{matrix} ; \bar{u}_1\ \bar{u}_2\ \bar{u}_3 \right\rangle$$

$$= \prod_{i=1}^{3} \delta_{h_i \bar{h}_i} \delta_{u_i \bar{u}_i} \prod_{j=1}^{2} \delta_{q_j \bar{q}_j}, \qquad (3.12)$$

which differs from the usual orthonormality relation in that a summation over h_i'' appears; because of (2.18), there is none over h_i'. In the orthonormality relation for the Wigner coefficients obtained by the procedure just described, no summation over h_i' or h_i'' appears, of course; they are orthonormal in the h_i, q_i and χ.

4. RECURSION RELATIONS FOR THE AW_n

These may be derived by a procedure which simply generalizes that of the preceding section and will, therefore, only be outlined.

Assuming that the Wigner coefficients of $S\mathfrak{U}_{n-1}$ are available, $P_{\max}$ and $P_{\min}$ may be chosen to be of highest and lowest weight, respectively, both in U_{2n-2} and $\mathfrak{U}_{n-1}$, so that

$$h_{ij}' = h_{in-1}',\ h_{ij}'' = h_{n+i-j-1,n-1}'' \qquad \text{for } 1 \leq i \leq j \leq n-2. \qquad (4.1)$$

$P_{\min}$ must now be a solution of (2.11) and of

$$\mathcal{C}^{\mu}_{\mu}P_{\min} = q''_{n-\mu}P_{\min},\ \mathcal{C}^{\nu}_{\mu}P_{\min} = 0,$$
$$1 \leq \mu < \nu \leq n-1 \qquad (4.2)$$

(the notation $h_{in} = h_i$, $h_{in-1} = q_i$, etc., will be used when convenient). Using the permutations (2.12) and

$$\begin{pmatrix} 1 & 2 & \cdots & n-1 & n \\ n-1 & n-2 & \cdots & 1 & n \end{pmatrix}$$

for the s and μ, the solution is seen to be

$$P_{\min}(h''_i,\ q''_i) = A\begin{pmatrix} h''_1 & \cdots & h''_{n-1} & 0 \\ q''_1 & \cdots & q''_{n-1} & \end{pmatrix}(\Delta^{2n-2}_{n-1})^{q_1''-h_2''}$$
$$\times (\Delta^{2n-2}_{n})^{h_1''-q_1''}(\Delta^{2n-3\ 2n-2}_{n-2\ n-1})^{q_2''-h_3''}$$
$$\times (\Delta^{2n-3\ 2n-2}_{n-1\ n})^{h_2''-q_2''}\cdots(\Delta^{n\ n+1\ \cdots\ 2n-2}_{2\ 3\ \cdots\ n})^{h_{n-1}''-q_{n-1}''}, \qquad (4.3)$$

where[19]

$$A\begin{pmatrix} h''_i \\ q''_i \end{pmatrix} = \left[\prod_{j=1}^{n}\prod_{i=1}^{j}\frac{(h_i-h_j+j-i)!\,(q_i-q_j+j-i)!}{(h_i-q_j+j-i)!\,(q_i-h_{j+1}+j-i)!}\right]^{\frac{1}{2}}. \qquad (4.4)$$

Now (4.3) may be written in alternative forms analogous to (3.8), and these lead to recursion relations in which the coefficients are, except for ratios of the normalization constants (4.4), the matrix elements

$$\langle h_{ij},\ k_{lm}|\ \Delta^{2n-2-q\ \cdots\ 2n-3\ 2n-2}_{n-1-q\ \cdots\ n-2\ n-1}\ |\bar{h}_{ij},\ \bar{k}_{lm}\rangle,$$
$$\langle h_{ij},\ k_{lm}|\ \Delta^{2n-2-q\ \cdots\ 2n-3\ 2n-2}_{n-q\ \cdots\ n-1\ n}\ |\bar{h}_{ij},\ \bar{k}_{lm}\rangle,$$
$$q = 0, 1, \cdots, n-2. \qquad (4.5)$$

Their derivation is described elsewhere.[19]

By diagonalizing simultaneously the matrices of the Casimir and X operators, one obtains the coefficients necessary to convert the AW_n gotten from (4.5) into particular Wigner coefficients of $S\mathfrak{U}_n$, i.e., those satisfying (4.1). There are now however, $n-2$ Casimir operators and $\frac{1}{2}n(n-3)+1$ operators of the type X. Care must be taken in using these particular Wigner coefficients of $S\mathfrak{U}_n$, since to obtain the general Wigner coefficients the product with the general coefficients of $S\mathfrak{U}_{n-1}$ must be summed over the eigenvalues χ of the X of $S\mathfrak{U}_{n-1}$.

It is planned to carry out the procedure described here for $S\mathfrak{U}_4$, because of its importance in supermultiplet theory.

APPENDIX

The matrix elements of Δ^{s}_{μ}, $\Delta^{st}_{\mu\nu}$ [see Eq. (3.10)] are different from zero only if certain of the following conditions hold:

$$\bar{h}_i = h_i + \delta_{ia} + \delta_{ib}, \qquad \bar{q}_i = q_i + \delta_{ia'} + \delta_{ib'},$$
$$\bar{r}_1 = r_1 + 1, \qquad \bar{u}_i = u_i + \delta_{id}.$$

Which of these must hold may easily be seen from the particular matrix element; thus Δ^{4}_{3} can only affect the h_i and only increases one of them by 1, so that the first condition, with $b = 0$, is the only one applicable; for Δ^{4}_{2}, the first two conditions hold, with $b = 0$ and $b' = 0$; and so on. All quantities not governed by these relations have, of course, the same value in bra and ket.

In order to simplify the notation, the convention is adopted in this Appendix of indicating in the bra merely the indices which correspond to places where a quantity has been increased by one; the first row, which, except for a zero, is common to both chains of groups, is written only once; the triplets (a, b, c) and (d, e, f) are taken to be arbitrary permutations of (1, 2, 3), and (a', b') is a permutation of (1, 2); and

$$S(x) = \begin{cases} +1 & \text{for } x \geq 0 \\ -1 & \text{for } x < 0. \end{cases}$$

Thus for instance,

$$\left\langle \begin{vmatrix} h_i+\delta_{ia}+\delta_{ib} & h_i+\delta_{ia}+\delta_{ib} & 0 \\ q_i+\delta_{ia'} & ; \quad u_i+\delta_{id} & \\ r_1 & v_i & \\ & w_1 & \end{vmatrix} \Delta^{34}_{23} \begin{vmatrix} h_i & h_i & 0 \\ q_i & ; \quad u_i & \\ r_1 & v_i & \\ & w_1 & \end{vmatrix} \right\rangle \equiv \left\langle \begin{matrix} a & ; & b \\ a' & & d \end{matrix} \middle| \Delta^{34}_{23} \right\rangle .$$

With these conventions, the matrix elements of interest for $S\mathfrak{U}_3$ are given below. That for Δ^{34}_{13}, though not needed for the computation of Wigner coefficients, is given for completeness' sake.

$$\left\langle \begin{matrix} a \\ ; \end{matrix} \middle| \Delta^{4}_{3} \right\rangle = [(h_a - q_1 + 2 - a)(h_a - q_2 + 3 - a)B]^{\frac{1}{2}}, \qquad (A1)$$

$$\left\langle \begin{matrix} a \\ a \end{matrix}; \left| \Delta_2^4 \right| \right\rangle$$

$$= S(a'-a)\left[\frac{(h_a - q_{b'} + b' - a + 1)(h_b - q_{a'} + a' - b)(h_c - q_{a'} + a' - c)(q_{a'} - r_1 + 2 - a')}{(q_{a'} - q_{b'} + b' - a')(q_{a'} - q_{b'} + b' - a' + 1)} B\right]^{\frac{1}{2}}, \tag{A2}$$

$$\left\langle \begin{matrix} a \\ a' \\ 1 \end{matrix}; \left| \Delta_1^4 \right| \right\rangle = (-1)^{a'+1} S(a'-a)$$

$$\times \left[-\frac{(h_a - q_{b'} + b' - a + 1)(h_b - q_{a'} + a' - b)(h_c - q_{a'} + a' - c)(q_{b'} - r_1 + 1 - b')}{(q_{a'} - q_{b'} + b' - a')(q_{a'} - q_{b'} + b' - a' + 1)} B\right]^{\frac{1}{2}}. \tag{A3}$$

In these formulas,

$$B = \frac{(h_a - u_1 + 2 - a)(h_a - u_2 + 3 - a)(h_a - u_3 + 4 - a)}{(h_a - h_b + b - a)(h_a - h_b + b - a + 1)(h_a - h_c + c - a)(h_a - h_c + c - a + 1)}, \tag{A4}$$

$$\left\langle \begin{matrix} a\ b \\ 1\ 2; d \\ 1 \end{matrix} \left| \Delta_{12}^{34} \right| \right\rangle = (-1)^{c+d} S(d-c)[-(h_c - q_1 + 1 - c)(h_c - q_2 + 2 - c)K]^{\frac{1}{2}}, \tag{A5}$$

$$\left\langle \begin{matrix} a\ b \\ a'; d \\ 1 \end{matrix} \left| \Delta_{13}^{34} \right| \right\rangle = (-1)^d S(a'-c)S(d-c)$$

$$\times \left[-\frac{(h_a - q_{b'} + b' - a + 1)(h_b - q_{b'} + b' - b + 1)(h_c - q_{a'} + a' - c)(q_{b'} - r_1 + 1 - b')}{(q_{a'} - q_{b'} + b' - a')(q_{a'} - q_{b'} + b' - a' + 1)} K\right]^{\frac{1}{2}}, \tag{A6}$$

$$\left\langle \begin{matrix} a\ b \\ a'\ d \end{matrix} \left| \Delta_{23}^{34} \right| \right\rangle = (-1)^{b'+d} S(a'-c)S(d-c)$$

$$\times \left[\frac{(h_a - q_{b'} + b' - a + 1)(h_b - q_{b'} + b' - b + 1)(h_c - q_{a'} + a' - c)(q_{a'} - r_1 + 2 - a')}{(q_{a'} - q_{b'} + b' - a')(q_{a'} - q_{b'} + b' - a' + 1)} K\right]^{\frac{1}{2}}. \tag{A7}$$

In these formulas,

$$K = \frac{(h_a - u_e + e - a + 1)(h_b - u_e + e - b + 1)(h_a - u_f + f - a + 1)(h_b - u_f + f - b + 1)}{(h_a - h_c + c - a)(h_a - h_c + c - a + 1)(h_b - h_c + c - b)(h_b - h_c + c - b + 1)}$$

$$\times \frac{(h_c - u_d + d - c)(u_d - v_1 + 2 - d)(u_d - v_2 + 3 - d)(u_d + 4 - d)}{(u_d - u_e + e - d)(u_d - u_e + e - d + 1)(u_d - u_f + f - d)(u_d - u_f + f - d + 1)}. \tag{A8}$$

Algebraic Tabulation of the Clebsch–Gordan Coefficients for Reduction of the Product $(\lambda, \mu) \oplus (3, 0)$ of Irreducible Representations of $SU(3)$*

L. K. PANDIT† AND N. MUKUNDA‡
Department of Physics and Astronomy, University of Rochester, Rochester, New York
(Received 25 January 1965)

Tables of Clebsch–Gordan coefficients of SU (3) for the reduction of the product $(\lambda, \mu) \otimes (3, 0)$ of representations of $SU(3)$ are constructed by use of the tensor method. Derivations of some crossing and symmetry relations for the SU_3 Clebsch–Gordan coefficients are given, and the Clebsch–Gordan coefficients for the product $(\mu, \lambda) \otimes (0, 3)$ are related to those mentioned above. The phase convention used in compiling the tables is stated and explained.

INTRODUCTION

ON account of the great importance recently attained by the group $SU(3)$ in the field of high-energy physics,[1] the job of constructing tables of "Clebsch–Gordan (CG) coefficients" for this group has become of practical interest. Various authors[2] have made numerical tables for the reduction of products of some low-dimensional representations. Since no general expression is known yet for any arbitrary CG coefficient of $SU(3)$, the next thing one may attempt is to obtain general algebraic expressions for the CG-coefficients for reduction of a product of an arbitrary representation (λ, μ) with a low-dimensional representation of practical interest. It is now well-known that unitary irreducible representations (UIR's) of dimensionality 8 and 10 play a special role in the "eightfold-way" of $SU(3)$ symmetry[1b,c]—being the representations definitely "filled" by experimentally well-established meson and baryon states. An algebraic tabulation of the CG-coefficients for reduction of the product of the UIR (λ, μ) and the UIR (1,1) of dimensionality 8 has very recently been made by Kuriyan, Lurié, and Macfarlane.[3] In the present paper we give a tabulation for the reduction of the product of (λ, μ) with the UIR's (3, 0) and (0, 3) of dimensionality 10.

The present work is motivated both by practical as well as methodological interest. The former aspect has already been touched upon in the foregoing. The methodological interest derives from the fact that our results have been obtained using purely tensor methods, which turn out to be particularly simple for the $SU(3)$ group. In a recent paper[4] (hereafter referred to as I) we presented a construction of orthonormal basic states of an arbitrary UIR (λ, μ) in terms of general irreducible tensors. This construction allows one to derive several algebraic results for the group in a simple manner, some examples of which were given in I. In the derivation of the results of the present work we have employed these tensor methods. Some time ago Moshinsky[5] had obtained an expression in the form of a series for the CG coefficients for the product $(\lambda, \mu) \otimes (\nu, 0)$, a case general enough to cover the product $(\lambda, \mu) \otimes (3, 0)$ to be discussed here. However, since the methods employed by Moshinsky are more abstract, we feel that our derivations of the closed algebraic expressions for the CG coefficients tabulated here are of immediate interest; and this interest is the more, since we have taken care to incorporate a phase convention which is becoming rapidly standardized.[6] We also take this occasion

* Research supported in part by the U. S. Atomic Energy Commission.

† On leave of absence from Tata Institute of Fundamental Research, Bombay, India.

‡ Present Address: Palmer Physical Laboratory, Princeton University, Princeton, New Jersey. On leave of absence from the Atomic Energy Establishment, Bombay, India.

[1] (a) M. Ikeda, S. Ogawa, and Y. Ohnuki, Progr. Theoret. Phys. (Kyoto), **22**, 715 (1960); **23**, 1073 (1960); O. Klein, Arkiv Fysik **16**, 191 (1959); J. E. Wess, Nuovo Cimento **15**, 52 (1960); Y. Yamaguchi, Progr. Theoret. Phys. (Kyoto) Suppl. **11**, 1, 37 (1959). (b) Y. Ne'eman, Nucl. Phys. **26**, 222 (1961). (c) M. Gell-Mann, Phys. Rev. **125**, 1067 (1962). (d) S. Okubo, Progr. Theoret. Phys. (Kyoto) **27**, 949 (1962).

[2] (a) S. Sawada and M. Yonezawa, Progr. Theoret. Phys. (Kyoto) **23**, 662 (1960); M. A. Rashid, Nuovo Cimento **26**, 118 (1962). (b) A. R. Edmonds, Proc. Roy. Soc. (London) **A268**, 567 (1962). (c) Y. Dothan and H. Harari, Israel Atomic Energy Commission, Report IA-777 (1963); P. Tarjanne, Carnegie Institute of Technology Report NYO 9290A (1963); also Ann. Acad. Sci. Fenn., AVI Physica 105 (1962). (d) J. J. deSwart, Rev. Mod. Phys. **35**, 916 (1963). (e) P. McNamee and F. Chilton, Rev. Mod. Phys. **36**, 1005 (1964).

[3] J. G. Kuriyan, D. Lurié, and A. J. Macfarlane, J. Math. Phys. **6**, 722 (1965).

[4] N. Mukunda and L. K. Pandit, J. Math. Phys. **6**, 746 (1965). This paper is referred to as I in the present work. We refer the reader to this paper for details of notation and results used here, as also for more references on our subject.

[5] M. Moshinsky, Rev. Mod. Phys. **34**, 813 (1962).

[6] For a detailed discussion of the question of phases, see J. G. Kuriyan *et al.*, Ref. 3.

to derive by the tensor method some useful crossing and symmetry relations for the CG coefficients.[7]

Section I contains a discussion of the UIR's (3, 0) and (0, 3) of $SU(3)$, of the direct products (3, 0) $\otimes$ (λ, μ) and (0, 3) $\otimes$ (μ, λ), and a definition of the corresponding CG coefficients. Section II contains a discussion of the behavior of the basic states of a UIR under complex conjugation. This leads to an important "crossing relation" for CG coefficients. It is also shown how the CG coefficients for the product (0, 3) $\otimes$ (μ, λ) may be directly obtained from those for the product (3, 0) $\otimes$ (λ, μ). Finally, a clear statement of the phase convention to be used in defining CG coefficients, and a complete list of relations between the different CG coefficients, are given. In Sec. III we outline the method of derivation of the CG coefficients. The results of the computations are given in Tables V–XIV. (All tables cited herein are to be found, in numerical order, at the end of the article.)

SECTION I

The *decuplet representation* of $SU(3)$ is the ten-dimensional unitary irreducible representation (UIR) denoted, in the highest weight notation, by (3, 0). It is provided by a symmetric third-rank tensor S^{abc} with three upper indices. [We use here the notation and terminology of I, see Ref. (4)]. The isospin and hypercharge content of this UIR may be conveniently displayed as in Table I, which gives the orthonormal basic states of (3, 0) in terms of S^{abc}. (See Appendix B of I).

An *irreducible tensor operator* of the decuplet type is a set of ten operators labeled and transforming *exactly* like the standard orthonormal basic states of the UIR (3, 0), under transformations of $SU(3)$. (By "standard," we refer to the Biedenharn phase convention for the matrix elements of the generators of $SU(3)$, in any UIR.)[8] This definition of a decuplet tensor operator can be converted in the usual manner into a standard set of commutation relations between these operators and the generators of $SU(3)$. We refer to the quartet of operators with $I = \frac{3}{2}$ as $\mathfrak{Q}_M$, $M = \pm\frac{3}{2}, \pm\frac{1}{2}$; the triplet with $I = 1$ as $\mathfrak{Z}_M$, $M = \pm 1, 0$; the doublet with $I = \frac{1}{2}$ as $\mathfrak{D}_M$, $M = \pm\frac{1}{2}$; and the singlet with $I = 0$ as $\mathfrak{S}$.

We go on now to consider the direct product of the UIR (3, 0) and an arbitrary UIR (λ, μ) of $SU(3)$. Such a product can be reduced into a direct sum of UIR's of $SU(3)$, according to

$$(\lambda, \mu) \otimes (3, 0) = \sum\nolimits_{\oplus} (\lambda', \mu'). \qquad \text{(I.1)}$$

The contents of the rhs of (I.1) may be found by several methods, such as the use of Young's diagrams,[9] the technique of Speiser,[10] or the method of tensors.[11] We find that, in general, the rhs of (I.1) consists of ten *distinct* UIR's of $SU(3)$. They are listed in Table II, along with the construction of the corresponding irreducible tensors. As before, the UIR (3, 0) is described by the tensor S^{abc}, while the UIR (λ, μ) is given by an irreducible tensor $T^{m\cdots}{}_{n\cdots}$ with λ upper and μ lower indices. The operations indicated in Table II must be followed by symmetrization and the removal of traces, when necessary. The tensor ϵ_{abc} is the usual completely antisymmetric invariant tensor of $SU(3)$; note that the numerical values of its components are unaffected by any particular phase convention for upper and lower tensor indices, such as that adopted in I. However, as in I, the process of contraction of an upper and a lower tensor index must, by definition, always be done with the tensor g^a_b of I, even when an index being contracted appears in ϵ_{abc}.

Now let (λ', μ') denote any one of the ten UIR's of Table II. The orthonormal basic states of (λ', μ') are then obtained from direct products of basic states of (λ, μ) and (3, 0), by the use of the corresponding Clebsch–Gordan (CG) coefficients of $SU(3)$. We use the notation of Kuriyan *et al.*[3] and write this CG coefficient in the following form:

$$C(\lambda\mu\ 30\ \lambda'\mu'\ ;\ IMY\ JNZ\ I'M'Y'). \qquad \text{(I.2)}$$

The labels I, M, Y are, respectively, the isospin, Z-component of isospin, and hypercharge labels for states of (λ, μ); the labels JNZ and $I'M'Y'$ refer similarly to (3, 0) and (λ', μ'), respectively. Note that since no UIR (λ', μ') appears more than once in Table II, the labels appearing in (I.2) are sufficient and specify unambiguously the corresponding CG coefficient.

The factorization theorem of Racah[12] allows us to write (I.2) as the product of a CG coefficient for isospin and an "*isoscalar factor*" as follows:

$$C(\lambda\mu\ 30\ \lambda'\mu';\ IMY\ JNZ\ I'M'Y')$$
$$= C(IJI';\ MNM')U(\lambda\mu\ 30\ \lambda'\mu';\ IY\ JZ\ I'Y'). \qquad \text{(I.3)}$$

The matrix elements of a decuplet tensor operator S, between states belonging to (λ, μ) and (λ', μ'), can be expressed in terms of the CG coefficient (I.2) and a single reduced matrix element. Thus

[7] See also J. J. deSwart, Ref. 2(d).
[8] L. C. Biedenharn, Phys. Letters **3**, 69 (1962).
[9] See, for example, A. R. Edmonds, Ref. 2(b).
[10] (a) D. R. Speiser, Proceedings of the Istanbul Summer School, Istanbul (1962). (b) Also J. J. deSwart, Ref. 2(d).
[11] N. Mukunda and L. K. Pandit, Progr. Theoret. Phys. (Kyoto), **34**, 46 (1965).
[12] G. Racah, Phys. Rev. **76**, 1352 (1949).

$$\begin{aligned}
\langle \lambda'\mu'; I'M'Y' \,|\mathfrak{Q}_N|\, \lambda\mu; IMY\rangle &= C(\lambda\mu\ 30\ \lambda'\mu'\ ; IMY\ \tfrac{3}{2}N1\ I'M'Y')\langle \lambda'\mu'\ ||S||\ \lambda\mu\rangle,\\
\langle \lambda'\mu'; I'M'Y' \,|\mathfrak{I}_N|\, \lambda\mu; IMY\rangle &= C(\lambda\mu\ 30\ \lambda'\mu'\ ; IMY\ 1N0\ I'M'Y')\langle \lambda'\mu'\ ||S||\ \lambda\mu\rangle,\\
\langle \lambda'\mu'; I'M'Y' \,|\mathfrak{D}_N|\, \lambda\mu; IMY\rangle &= C(\lambda\mu\ 30\ \lambda'\mu'\ ; IMY\ \tfrac{1}{2}N,-1\ I'M'Y')\langle \lambda'\mu'\ ||S||\ \lambda\mu\rangle,\\
\langle \lambda'\mu'; I'M'Y' \,|\mathfrak{S}|\, \lambda\mu; IMY\rangle &= C(\lambda\mu\ 30\ \lambda'\mu'\ ; IMY\ 00,-2\ I'M'Y')\langle \lambda'\mu'\ ||S||\ \lambda\mu\rangle.
\end{aligned} \tag{I.4}$$

We present the *isoscalar factors* U appearing in (I.3), in a series of ten Tables (in numerical order at the end of this paper), corresponding to the ten cases of Table II. In Sec. III, we describe briefly the method of evaluation of the isoscalar factors. In Sec. II, among other things, we relate the CG coefficient (I.2) required for the reduction of $(\lambda, \mu) \otimes (3, 0)$, to the following one:

$$C(30\ \lambda\mu\ \lambda'\mu'\ ; JNZ\ IMY\ I'M'Y'), \tag{I.5}$$

required for the reduction of $(3, 0) \otimes (\lambda, \mu)$.

We close the present section with a few comments on the UIR (0, 3). The ten-dimensional UIR (0, 3) of $SU(3)$ is equivalent to the complex conjugate of the UIR (3, 0), and is provided by a symmetric tensor S_{abc} with three lower indices. The IMY content of (0, 3) is obtained from that of (3, 0), by changing the sign of Y.

Let (μ, λ) be any UIR of $SU(3)$. [The UIR (μ, λ) is the complex conjugate of the UIR (λ, μ) that appears in (I.1). For the sake of symmetry, we consider here (μ, λ) rather than (λ, μ).] Then analogously to (I.1), we have the reduction

$$(\mu, \lambda) \otimes (0, 3) = \sum_{\oplus} (\mu', \lambda'). \tag{I.6}$$

The UIR's (μ', λ') that appear on the rhs of (I.6) are just the complex conjugates of those that appear on the rhs of (I.1); and the allowed values of μ' and λ' in (I.6) are the same as those given by Table II. The CG coefficient for the reduction (I.6) is written

$$C(\mu\lambda\ 03\ \mu'\lambda'\ ; IMY\ JNZ\ I'M'Y'), \tag{I.7}$$

and will be related, in Sec. II, both to the CG coefficient (I.2) and to one of type

$$C(03\ \mu\lambda\ \mu'\lambda'\ ; JNZ\ IMY\ I'M'Y'), \tag{I.8}$$

required for reducing the product $(0, 3) \otimes (\mu, \lambda)$.

SECTION II

In this Section, a derivation is given of a "crossing relation" for the CG coefficient (I.2),[13] whereby four of the cases in Table II may be directly related to four others. In addition, the phase convention to be used in order to completely define the CG coefficients (I.2) and (I.7), and the relationship of these to one another and to (I.5, 8) will be made clear.

Consider first the construction of an invariant from the direct product of two UIR's that are complex conjugates of one another, namely (λ, μ) and (μ, λ). Denote the corresponding irreducible tensors by $T^{m\cdots}_{n\cdots}$ and $S^{n\cdots}_{m\cdots}$, respectively, and the corresponding auxiliary quantities (in the sense of I) by

$$T^{j_1m_1}_{j_2m_2},\ \psi^{j_1m_1}_{j_2m_2},\ \psi^{I(j_1j_2)}_{M},\ \Psi^{IY}_{M};$$

and

$$S^{j_1m_1}_{j_2m_2},\ \phi^{j_1m_1}_{j_2m_2},\ \phi^{I(j_1j_2)}_{M},\ \Phi^{IY}_{M};$$

respectively. [Here both Ψ^{IY}_M and Φ^{IY}_M *include* the final choices of phases made in I, Appendix B, Eq. (B1).] The invariant $\mathcal{I}$ that can be formed from T and S is obtained by contracting all upper indices in T with all lower indices in S, and vice versa, that is:

$$\mathcal{I} = T^{ab\cdots}_{pq\cdots} S^{pq\cdots}_{ab\cdots}. \tag{II.1}$$

We may now go through the same sequence of steps as employed in Sec. III of I, to successively rewrite (II.1) in terms of $\psi^{j_1m_1}_{j_2m_2}$, $\phi^{j_1m_1}_{j_2m_2}$ then $\psi^{I(j_1j_2)}_M$, $\phi^{I(j_1j_2)}_M$ and finally Ψ^{IY}_M, Φ^{IY}_M. In doing so, we must take account of the signs introduced by the tensor g^a_b in the contractions in (II.1). In this manner we find

$$\begin{aligned}
\mathcal{I} &= \sum_{j_1m_1} \sum_{j_2m_2} [N_1(\lambda\mu; j_1m_1j_2m_2)]^2 \\
&\qquad \times (-1)^{\mu+m_2-j_2-\lambda+m_1+j_1} T^{j_1m_1}_{j_2m_2} S^{j_2,-m_2}_{j_1,-m_1},\\
&= \sum_{j_1j_2} \sum_{IM} (-1)^{\mu-\lambda+M+j_1-j_2} \psi^{I(j_1j_2)}_M \phi^{I(j_2j_1)}_{-M} \qquad \text{(II.2)}\\
&= \sum_{IMY} (-1)^{M-\frac{1}{2}Y+\frac{1}{3}(\mu-\lambda)} \Psi^{IY}_M \Phi^{I,-Y}_{-M}.
\end{aligned}$$

At this stage, we conveniently add a (λ, μ)-dependent phase factor to (II.2), normalize it, and finally redefine

$$\begin{aligned}
\mathcal{I} &= \left[(\lambda+1)(\mu+1)\left(\frac{\lambda+\mu}{2}+1\right)\right]^{-\frac{1}{2}} \\
&\qquad \times \sum_{IMY} (-1)^{M-\frac{1}{2}Y-\frac{1}{3}(\lambda+2\mu)} \Psi^{IY}_M \Phi^{I,-Y}_{-M}.
\end{aligned} \tag{II.3}$$

The phase chosen in (II.3) is such that it is $+1$ when I, M, Y correspond to the *highest weight* in the UIR (λ, μ): $I = M = \frac{1}{2}(\lambda + \mu)$, $\lambda = \frac{1}{3}(\lambda - \mu)$. The normalizing factor is given by the dimensionality of the UIR (λ, μ):

$$D(\lambda, \mu) = (\lambda + 1)(\mu + 1)[\tfrac{1}{2}(\lambda + \mu) + 1].$$

[13] See also: (a) J. J. deSwart, Ref. 2 (d); (b) A. J. Macfarlane, N. Mukunda, and E. C. G. Sudarshan, J. Math. Phys. 5, 576 (1964).

The expression (II.3) for $\mathcal{I}$ shows explicitly how one is to construct a *normalized invariant state*, given states Ψ and Φ transforming in the standard manner according to the UIR's (λ, μ) and (μ, λ), respectively. It may be put to several uses. We deduce that if Ψ^{IY}_M transforms according to the UIR (λ, μ) and if we *define* Φ^{IY}_M by

$$\Phi^{IY}_M = (-1)^{-M+\frac{1}{2}Y-\frac{1}{3}(\lambda+2\mu)}[\Psi^{I,-Y}_{-M}]^*, \qquad \text{(II.4)}$$

then the Φ^{IY}_M will transform in the standard manner according to (μ, λ).[14] The asterisk in (II.4) means complex conjugation if we deal with ordinary complex numbers, and Hermitian conjugation if we are dealing with tensor operators.

We next use (II.3) to derive the "crossing relations" for CG coefficients, following the general method of Macfarlane[13(b)] *et al.* Let (λ, μ), (λ'', μ''), and (λ', μ') be three UIR's such that the third is contained only *once* in the direct product of the first two. It then follows that (μ, λ) appears only once in the product of (μ', λ') and (λ'', μ'').[13(b)] Suppose next that Ψ, Φ, Ξ are quantities transforming, respectively, according to (λ, μ), (λ'', μ'') and (μ', λ'). (Note that the last UIR is (μ', λ'), *not* (λ', μ')!) Then there exists a unique (up to a phase) normalized state in the triple product space of Ψ, Φ, Ξ, that is invariant under SU_3. On comparing two alternative ways of constructing this state, namely, first coupling the UIR's (λ, μ) and (λ'', μ'') to (λ', μ'), and then using (II.3) to make an invariant, or first coupling (μ', λ') and (λ'', μ'') to (μ, λ) and then using (II.3), we obtain the relationship

$$\begin{aligned} &C(\lambda\mu\ \lambda''\mu''\ \lambda'\mu';\ IMY\ JNZ\ I'M'Y') \\ &= \eta(-1)^{\frac{1}{3}(\lambda'+2\mu'+2\lambda+\mu)+\frac{1}{2}Z-N}[D(\lambda', \mu')/D(\lambda, \mu)]^{\frac{1}{2}} \\ &\quad \times C(\mu'\lambda'\ \lambda''\mu''\ \mu\lambda;\ I'-M'-Y'\ JNZ\ I-M-Y). \end{aligned} \qquad \text{(II.5)}$$

In this equation, η is a phase factor dependent only on variables of the type λ, μ. [For the explicit and general derivation of (II.5), including the generalization to the case that (λ', μ') appears more than once in $(\lambda, \mu) \otimes (\lambda'', \mu'')$, see Ref. 13(b).] Using (II.5) for the case $(\lambda'', \mu'') = (3, 0)$, we find that the CG coefficients for cases A, B, E, H, of Table II may be simply expressed in terms of those for cases D, C, G, J, respectively, while cases F, K are not related to any other cases. [A similar situation clearly occurs for $(\lambda'', \mu'') = (0, 3)$.] However, for ease in applications, we have given separate and full tables for all the ten cases of Table II. The relationship for isoscalar factors that follows from (II.5) is

$$\begin{aligned} &U(\lambda\mu\ 30\ \lambda'\mu';\ IY\ JZ\ I'Y') \\ &= \eta(-1)^{\frac{1}{3}(\lambda'+2\mu'+2\lambda+\mu)+J+\frac{1}{2}Z} \\ &\quad \times [\{D(\lambda', \mu')\cdot(2I+1)\}/\{D(\lambda, \mu)(2I'+1)\}]^{\frac{1}{2}} \\ &\quad \times U(\mu'\lambda'\ 30\ \mu\lambda;\ I'-Y'\ JZ\ I-Y). \end{aligned} \qquad \text{(II.6)}$$

We deal next with CG coefficients of type (I.7). Since we can always choose phases of states so that all CG coefficients are real, straightforward application of (II.4) leads to

$$\begin{aligned} &C(\lambda\mu\ 30\ \lambda'\mu';\ IMY\ JNZ\ I'M'Y') \\ &= \xi C(\mu\lambda\ 03\ \mu'\lambda';\ I-M-Y\ J-N-Z\ I'-M'-Y'), \end{aligned} \qquad \text{(II.7)}$$

whereby CG coefficients for the product $(\mu, \lambda) \otimes (0, 3)$ are given in terms of those for $(\lambda, \mu) \otimes (3, 0)$.[2(d)] Here ξ is a phase again depending only on the UIR's involved. In deriving (II.7), one uses the fact that the CG coefficients appearing therein vanish unless $M' = M + N$, $Y' = Y + Z$. The crossing relationship for the (0, 3) cases can be read off directly from (II.5), or obtained on combining (II.7) with the crossing relationship for the (3, 0) case. We write it in the form

$$\begin{aligned} &C(\mu\lambda\ 03\ \mu'\lambda';\ IMY\ JNZ\ I'M'Y') \\ &= \bar{\eta}(-1)^{\frac{1}{3}(\lambda'+2\mu'+2\lambda+\mu)-\frac{1}{2}Z+N}[D(\mu', \lambda')/D(\mu, \lambda)]^{\frac{1}{2}} \\ &\quad \times C(\lambda'\mu'\ 03\ \lambda\mu;\ I'-M'-Y'\ JNZ\ I-M-Y). \end{aligned} \qquad \text{(II.8)}$$

As in (II.5), so in (II.8) we have a phase factor $\bar{\eta}$ to be evaluated.

We can relate the CG coefficient (I.2) to (I.5). Either from their properties as invariant tensors, or by the method of Kuriyan *et al.*,[3] we deduce that the two may differ at most by a phase ζ, and write

$$\begin{aligned} &C(\lambda\mu\ 30\ \lambda'\mu';\ IMY\ JNZ\ I'M'Y') \\ &= \zeta C(30\ \lambda\mu\ \lambda'\mu';\ JNZ\ IMY\ I'M'Y'). \end{aligned} \qquad \text{(II.9)}$$

Similarly, for (I.7) and (I.8), one has

$$\begin{aligned} &C(\mu\lambda\ 03\ \mu'\lambda';\ IMY\ JNZ\ I'M'Y') \\ &= \bar{\zeta} C(03\ \mu\lambda\ \mu'\lambda';\ JNZ\ IMY\ I'M'Y'). \end{aligned} \qquad \text{(II.10)}$$

The last relationship between CG coefficients that we consider is the following:

$$\begin{aligned} &C(30\ \lambda\mu\ \lambda'\mu';\ JNZ\ IMY\ I'M'Y') \\ &= \bar{\xi} C(03\ \mu\lambda\ \mu'\lambda';\ J-N-Z\ I-M-Y\ I'-M'-Y'). \end{aligned} \qquad \text{(II.11)}$$

[14] See also D. Lurié and A. J. Macfarlane, J. Math. Phys. **5**, 565 (1964); J. J. deSwart, Ref. 2(d).

The determination of the six phase factors appearing in (II.5–II.11), as well as the complete definition of the CG coefficients themselves, depends on the choice of a *phase convention* for CG coefficients of $SU(3)$. We follow the convention of Kuriyan *et al.*,[3] in whose work an excellent discussion of this question may be found. Before stating this convention for our case, a brief discussion is desirable. For fixed values of (λ, μ) and (λ', μ'), the *relative* signs of the CG coefficients (or isoscalar factors) of (I.2) are completely fixed by the standard choice of the matrix elements of the $SU(3)$ generators in any UIR, namely the Biedenharn convention[8]; similarly for the CG coefficients of type (I.5), (I.7) or (I.8). For given (λ, μ) and (λ', μ'), the problem remains of choosing in some definite way the sign of any one of the CG coefficients (I.2) for a definite choice of the variables of type I, M and Y (making sure, of course, that the corresponding CG coefficient does not vanish). A similar choice must be made in each of the cases (I.5), (I.7), (I.8). The prescription of Kuriyan *et al.*,[3] adapted to the present case, is the following: Let $\bar{I}$, $\bar{M}$, $\bar{Y}$ and $\bar{I}'$, $\bar{M}'$, $\bar{Y}'$ denote the labels of the *highest weight* states of the UIR's (λ, μ) and (λ', μ'), respectively. Then for each (λ', μ') there is always one (and only one) allowed set of values for JNZ such that (I.2) does not vanish. We demand

$$C(\lambda\mu\ 30\ \lambda'\mu'\ ;\ \bar{I}\bar{M}\bar{Y}\ JNZ\ \bar{I}'\bar{M}'\bar{Y}') > 0. \quad \text{(II.12)}$$

A similar requirement is made for the CG coefficients (I.7):

$$C(\mu\lambda\ 03\ \mu'\lambda'\ ;\ \bar{I}\bar{M}\bar{Y}\ JNZ\ \bar{I}'\bar{M}'\bar{Y}') > 0. \quad \text{(II.13)}$$

The requirements (II.12) and (II.13) suffice for an unambiguous determination of the corresponding CG coefficients, for two reasons: Firstly, because the weights that occur in both (3, 0) and (0, 3) are simple, and so the labels JNZ that appear in (II.12) and (II.13) are unique, if they exist at all. Secondly, we have explicitly verified from our tables that the particular CG coefficients (II.12) and (II.13) never accidentally vanish.[15,16]

For the CG coefficients (I.5) [and similarly for (I.8)] in general the weights of the UIR "in the middle" (λ, μ) are nonsimple. If we fix $\bar{J}\bar{N}\bar{Z}$ and $\bar{I}'\bar{M}'\bar{Y}'$ to refer to the *highest* weights of (3, 0) and (λ', μ'), respectively, this determines M, Y in (I.5), but there are many values of I such that (I.5) does not vanish. We pick the *highest* value of I that occurs in (λ, μ) for the given Y and call it $\bar{I}$. (Note that $\bar{I}$ is *not*, in general, the highest isospin in (λ, μ).) We then demand that

$$C(30\ \lambda\mu\ \lambda'\mu'\ ;\ \bar{J}\bar{N}\bar{Z}\ \bar{I}MY\ \bar{I}'\bar{M}'\bar{Y}') > 0. \quad \text{(II.14)}$$

Similarly for (I.8), we demand

$$C(03\ \mu\lambda\ \mu'\lambda'\ ;\ \bar{J}\bar{N}\bar{Z}\ \bar{I}MY\ \bar{I}'\bar{M}'\bar{Y}') > 0. \quad \text{(II.15)}$$

We have verified explicitly with the help of our tables that, with the choice of $\bar{I}$ prescribed above, the CG coefficients appearing in (II.14) and (II.15) never vanish accidentally.[15] Equations (II.12)–(II.15) define the phase conventions for the CG coefficients dealt with here.

All the phase factors appearing in (II.5)–(II.11) can now be determined, with occasional use of the tables. After explicit evaluation, it turns out to be possible to write each of the phase factors analytically in terms of λ, μ, λ', μ'. For conciseness, we give below all six of the equations in their final forms:

$$\begin{aligned}
&C(\lambda\mu\ 30\ \lambda'\mu'\ ;\ IMY\ JNZ\ I'M'Y') \\
&\qquad = (-1)^{\frac{1}{2}(\lambda'-\mu'+\mu-\lambda)+\frac{1}{2}Z-N}\cdot\left[\frac{D(\lambda',\mu')}{D(\lambda,\mu)}\right]^{\frac{1}{2}}\cdot C(\mu'\lambda'\ 30\ \mu\lambda\ ;\ I'-M'-Y'\ JNZ\ I-M-Y), \\
&C(\lambda\mu\ 30\ \lambda'\mu'\ ;\ IMY\ JNZ\ I'M'Y') = (-1)^{\mu'-\mu}\cdot C(30\ \lambda\mu\ \lambda'\mu'\ ;\ JNZ\ IMY\ I'M'Y'), \\
&C(\lambda\mu\ 30\ \lambda'\mu'\ ;\ IMY\ JNZ\ I'M'Y') = (-1)^{\lambda'-\lambda+1}\cdot C(\mu\lambda\ 03\ \mu'\lambda'\ ;\ I-M-Y\ J-N-Z\ I'-M'-Y'), \\
&C(\mu\lambda\ 03\ \mu'\lambda'\ ;\ IMY\ JNZ\ I'M'Y') \\
&\qquad = (-1)^{\frac{1}{2}Z-N+1}\cdot[D(\lambda',\mu')/D(\lambda,\mu)]^{\frac{1}{2}}\cdot C(\lambda'\mu'\ 03\ \lambda\mu\ ;\ I'-M'-Y'\ JNZ\ I-M-Y), \\
&C(\mu\lambda\ 03\ \mu'\lambda'\ ;\ IMY\ JNZ\ I'M'Y') = (-1)^{\mu'-\mu}C(03\ \mu\lambda\ \mu'\lambda'\ ;\ JNZ\ IMY\ I'M'Y'), \\
&C(30\ \lambda\mu\ \lambda'\mu'\ ;\ JNZ\ IMY\ I'M'Y') = (-1)^{\lambda'-\lambda+1}\cdot C(03\ \mu\lambda\ \mu'\lambda'\ ;\ J-N-Z\ I-M-Y\ I'-M'-Y').
\end{aligned} \quad \text{(II.16)}$$

[15] As stated by Kuriyan *et al.* (Ref. 3), "a CG coefficient is said to vanish accidentally if the vanishing is not the result of a selection rule."

[16] The following point is worth noting. As the numbers (λ', μ') range over the ten cases listed in Table II, the labels J, N, Z, that appear in (II.12) range over the ten "weights" of the UIR (3,0). Thus a one-to-one correspondence is set up between the irreducible constituents of $(\lambda, \mu) \otimes (3, 0)$ and the "weights" of the UIR (3, 0). This correlation has been stated, for the general case, by L. C. Biedenharn, Phys. Letters **3**, 254 (1963). It has been proved in Ref. 11 by the present authors, and also independently by B. Preziosi, A. Simoni and B. Vitale, preprint, Naples (1964). See Ref. 3 for the analogous situation in the case of the product $(\lambda, \mu) \otimes (1, 1)$. It may be noted that the graphical methods of Speiser [see ref. 10(a)] also implies this correlation.

These relations can easily be written in terms of isoscalar factors.

SECTION III

In this Section we outline the method of computing the isoscalar factors corresponding to the CG coefficients (I.2), namely

$$U(\lambda\mu\ 30\ \lambda'\mu'\ ;\ IY\ JZ\ I'Y'). \tag{III.1}$$

We pick just one case, case (J) of Table II, for illustration, as all the other cases follow the same pattern. The method is the same as that used, and explained in detail, in I for the derivation of the matrix elements of the octet operators. Since, for any particular case, we calculate as a first step the isoscalar factors U only up to an overall normalization factor and an overall phase characteristic of the particular case, the method of calculation to be now described leads us first to the evaluation of a set of auxiliary quantities V which may be defined as:

$$V(\lambda\mu\ 30\ \lambda'\mu'\ ;\ IY\ JZ\ I'Y') = \epsilon[N(\lambda\mu;\ \lambda'\mu')]^{-1}U(\lambda\mu\ 30\ \lambda'\mu'\ ;\ IY\ JZ\ I'Y'). \tag{III.2}$$

Here, $N(\lambda\mu;\ \lambda'\mu')$ is a positive normalizing factor independent of the IY-type variables, and ϵ is a sign factor, ± 1, also independent of the IY-type variables. For each $(\lambda',\ \mu')$, i.e., each case of Table II, we first obtain the quantities V, then determine ϵ by means of the phase convention of Sec. II, namely Eq. (II.12), and finally determine $N(\lambda\mu;\ \lambda'\mu')$ by the requirement that U be normalized. This gives

$$[N(\lambda\mu;\ \lambda'\mu')]^{-2} = \sum_{IY}\sum_{JZ}[V(\lambda\mu\ 30\ \lambda'\mu'\ ;\ IY\ JZ\ I'Y')]^2. \tag{III.3}$$

The sum in (III.3) is, of course, to be carried out keeping I', Y' fixed: as a matter of convenience it turns out best to choose $I' = 0$.

Case J: $(\lambda',\ \mu') = (\lambda - 2,\ \mu + 1)$.

The first step is to construct an irreducible tensor of type $(\lambda - 2,\ \mu + 1)$, in the product space of the irreducible tensors S [of type (3, 0)] and T [of type $(\lambda,\ \mu)$]. Following Table II, we write down

$$R^{m_1\cdots m_{\lambda-s}}_{n_1\cdots n_{\mu-1};\alpha\beta} = \epsilon_{\alpha br}\epsilon_{\beta cs}T^{m_1\cdots m_{\lambda-s}rs}_{n_1\cdots n_{\mu-1}a}S^{abc}. \tag{III.4}$$

By construction, R is a tensor of type $(\lambda - 2,\ \mu + 1)$, completely symmetric in its upper indices, and completely traceless. It is symmetric in the indices $n_1\cdots n_{\mu-1}$ among themselves, and also with respect to α, β; it will become *irreducible* if we symmetrize it with respect to interchanges of any two lower indices, one appearing before and one after the semicolon. This is straightforward. Denote the resulting irreducible tensor by P, and the corresponding (orthogonal and relatively normalized) basic states by Ψ'^{IY}_M. To compute the functions V of (III.2), it suffices to consider Ψ'^{IY}_I, which, according to Eq. (B.2), of Appendix B of I, is given in terms of the tensor P by

$$\Psi'^{IY}_I = N'_3(I,\ Y)N'_1(j_1j_1j_2j_2)P^{(2j_1,0,\lambda'-2j_1)}_{(0,2j_2,\mu'-2j_2)}, \tag{III.5}$$

where

$$j_1 = \tfrac{1}{2}I + \tfrac{1}{4}Y + \tfrac{1}{6}(\lambda' - \mu'),\qquad j_2 = \tfrac{1}{2}I - \tfrac{1}{4}Y - \tfrac{1}{6}(\lambda' - \mu'); \tag{III.6}$$

and where the N'-factors are the same as the N-factors of I with λ, μ replaced by λ', μ'. Note that (III.5) does not yet contain the final choice of phases made in Eq. (B1) of I, but corresponds rather to Eq. (B2) of I. The final phases will be incorporated later. As explained in I, in (III.5) we have that component of P which has 1 as $2j_1$ upper and 0 lower indices, 2 as 0 upper and $2j_2$ lower indices, and 3 as $(\lambda' - 2j_1)$ upper and $(\mu' - 2j_2)$ lower indices. Now symmetrization of R to obtain P involves only a proper combinatorial counting. Suppressing for the moment the upper indices for which there is no complication in the present case, we have the result

$$P_{(0,2j_2,\mu'-2j_2)} = (2j_2)(2j_2 - 1)R_{(0,2j_2-2,\mu'-2j_2);22} + 2(\mu' - 2j_2)(2j_2)R_{(0,2j_2-1,\mu'-2j_2-1);23} + (\mu' - 2j_2)(\mu' - 2j_2 - 1)R_{(0,2j_2,\mu'-2j_2-2);33}. \tag{III.7}$$

We have introduced here the notation $R_{(x,y,z);\alpha\beta}$ for the component of R with x indices equal to 1, y indices equal to 2, and z indices equal to 3, *before* the semicolon. The components of R occurring in (III.7) can now be expressed in terms of S and T, and we have to look for the terms involving S^{111}, $\sqrt{3}S^{113}$, $\sqrt{3}S^{133}$ and S^{333} to extract all the functions V. For example, the terms in Ψ'^{IY}_I containing S^{111} are found to be

$$S^{111}N'_3(I,\ Y)N'_1(j_1j_1j_2j_2)[2j_2(2j_2 - 1)T^{(2j_1,0,\lambda-2j_1)}_{(1,2j_2-2,\mu-2j_2+1)} - 2(2j_2)(\mu - 2j_2 + 1)T^{(2j_1,1,\lambda-2j_1-1)}_{(1,2j_2-1,\mu-2j_2)} + (\mu - 2j_2 + 1)(\mu - 2j_2)T^{(2j_1,2,\lambda-2j_1-2)}_{(1,2j_2,\mu-2j_2-1)}]. \tag{III.8}$$

Further, in terms of the orthonormalized states Ψ^{IY}_M in the space of the tensor T [cf. Eq. (B2) of I], we have

$$T^{(2j_1,0,\lambda-2j_1)}_{(1,2j_2-2,\mu-2j_2+1)} = [N_1(j_1 j_1,\ j_2-\tfrac{1}{2},\ j_2-\tfrac{3}{2})]^{-1}\sum_{I'} C(j_1,\ j_2-\tfrac{1}{2},\ I';\ j_1,\ j_2-\tfrac{3}{2},\ I-\tfrac{3}{2})$$
$$\times N_2(j_1,\ j_2-\tfrac{1}{2},\ I')[N_3(I',\ Y-1)]^{-1}\Psi^{I',Y-1}_{I-\frac{3}{2}},\qquad I'=I-\tfrac{1}{2},\ I-\tfrac{3}{2}. \tag{III.9}$$

$$T^{(2j_1,1,\lambda-2j_1-1)}_{(1,2j_2-1,\mu-2j_2)} = [N_1(j_1+\tfrac{1}{2},\ j_1-\tfrac{1}{2},\ j_2,\ j_2-1)]^{-1}\sum_{I'} C(j_1+\tfrac{1}{2},\ j_2,\ I';\ j_1-\tfrac{1}{2},\ j_2-1,\ I-\tfrac{3}{2})$$
$$\times N_2(j_1+\tfrac{1}{2},\ j_2,\ I')[N_3(I',\ Y-1)]^{-1}\Psi^{I',Y-1}_{I-\frac{3}{2}},\qquad I'=I+\tfrac{1}{2},\ I-\tfrac{1}{2},\ I-\tfrac{3}{2}. \tag{III.10}$$

$$T^{(2j_1,2,\lambda-2j_1-2)}_{(1,2j_2,\mu-2j_2-1)} = [N_1(j_1+1,\ j_1-1,\ j_2+\tfrac{1}{2},\ j_2-\tfrac{1}{2})]^{-1}\sum_{I'} C(j_1+1,\ j_2+\tfrac{1}{2},\ I';\ j_1-1,\ j_2-\tfrac{1}{2},\ I-\tfrac{3}{2})$$
$$\times N_2(j_1+1,\ j_2+\tfrac{1}{2},\ I')[N_3(I',\ Y-1)]^{-1}\Psi^{I',Y-1}_{I-\frac{3}{2}},\qquad I'=I\pm\tfrac{1}{2},\ I\pm\tfrac{3}{2}. \tag{III.11}$$

All the information required for calculating the four functions $V(\lambda\mu\ 30\ \lambda-2,\mu+1;\ IY\ \frac{3}{2}1\ I'Y')$ is now contained in Eqs. (III.8)–(III.11). The N-factors are given in I and the CG coefficients of isospin are well known. For instance, on combining and simplifying all the factors, we find the term in Ψ^{IY}_I containing $S^{111}\Psi^{I+\frac{3}{2},Y-1}_{I-\frac{3}{2}}$ to be

$$S^{111}\Psi^{I+\frac{3}{2},Y-1}_{I-\frac{3}{2}}[(2I+2)(2I+3)]^{-1}$$
$$\times\left\{\frac{(\mu+1)}{\lambda(\lambda-1)(\lambda+\mu+1)}\,\frac{6g(g-1)(e+1)e(d+2)(a+2)(b+3)(b+2)(b+1)}{(2I+4)(2I+1)}\right\}^{\frac{1}{2}}. \tag{III.12}$$

The symbols a, b, d, e, g are defined in Table III. At this point we incorporate the final phase convention of I [Eq. (B1)], by setting

$$\Psi^{I,Y}_M = (-1)^{\,I-\frac{1}{2}Y-\frac{1}{2}(\lambda+2\mu)}|\lambda\mu;\ I,\ M,\ Y\rangle,\ \Psi^{I,Y}_M = (-1)^{I-\frac{1}{2}Y-\frac{1}{2}(\lambda'+2\mu')}\ |\lambda'\mu';\ I,\ M,\ Y\rangle. \tag{III.13}$$

Dropping a *common* $(\lambda,\ \mu)$-dependent factor, since we are still calculating the result up to an overall normalization factor, we identify from (III.12):

$$C(I+\tfrac{3}{2},\ \tfrac{3}{2},\ I;\ I-\tfrac{3}{2},\ \tfrac{3}{2},\ I)V(\lambda\mu\ 30\ \lambda-2,\mu+1\ ;\ I+\tfrac{3}{2},Y-1\ \tfrac{3}{2}1\ IY)$$
$$=[(2I+3)(2I+2)]^{-1}\left\{\frac{6g(g-1)(e+1)e(d+2)(a+2)(b+3)(b+2)(b+1)}{(2I+4)(2I+1)}\right\}^{\frac{1}{2}}. \tag{III.14}$$

After making the substitutions $I\to I-\frac{3}{2}$, $Y\to Y+1$, we finally get

$$V(\lambda\mu\ 30\ \lambda-2,\mu+1\ ;\ IY\ \tfrac{3}{2}1\ I-\tfrac{3}{2},Y+1) = (-)\left\{\frac{(a+1)(b+1)b(b-1)de(e-1)(g+2)(g+1)}{(2I)(2I-1)(2I-2)}\right\}^{\frac{1}{2}}. \tag{III.15}$$

In this manner all the V-functions have been evaluated. As explained earlier, once all the V-functions are known for a particular case, one finds the phase ϵ and normalizing factor $N(\lambda\mu;\ \lambda'\mu')$, which then lead to the isoscalar factors of the $SU(3)$ CG coefficients. The functions $U'=\epsilon V$ and N are presented in Tables V–XIV.

Structure of Tables V–XIV

We present in Tables V–XIV, the *isoscalar factors* of the $SU(3)$ CG coefficients for $(\lambda,\ \mu)\otimes(3,\ 0)$ corresponding, respectively, to the ten cases A to K of Table II. The plan of these tables is explained by Table IV, which shows that a common overall positive normalization factor $N(\lambda\mu;\ \lambda'\mu')$ for each case is given at the bottom of the relevant table. The algebraic expressions entered in each table, when multiplied by the corresponding normalization factor give the isoscalar factors U. Thus, calling these algebraic expressions as U', we have

$$U(\lambda\mu\ 30\ \lambda'\mu'\ ;\ IY\ JZ\ I'Y') = N(\lambda\mu;\ \lambda'\mu')U'(\lambda\mu\ 30\ \lambda'\mu'\ ;\ IY\ JZ\ I'Y').$$

The symbols used in writing the results are listed in Table III.

To obtain the CG coefficients, or the isoscalar factors, for the products $(3,\ 0)\otimes(\lambda,\ \mu)$, $(\mu,\ \lambda)\otimes(0,\ 3)$ and $(0,\ 3)\otimes(\mu,\ \lambda)$, one has simply to make use of Eqs. (II.16) along with the tables given here for $(\lambda,\ \mu)\otimes(3,\ 0)$.

ACKNOWLEDGMENTS

We would like to express our sincere thanks to Professor S. Okubo, Professor E. C. G. Sudarshan, and Professor A. J. Macfarlane for much encouragement. One of us (LKP) wishes to thank the U. S. Educational Foundation in India for the award of a Fulbright Travel Grant.

TABLE I. The UIR (3, 0).

	$I_z = M$							
Y	$-\frac{3}{2}$	-1	$-\frac{1}{2}$	0	$+\frac{1}{2}$	$+1$	$+\frac{3}{2}$	
1	S^{222}		$\sqrt{3}S^{122}$		$\sqrt{3}S^{112}$		S^{111}	$I = \frac{3}{2}$: $\mathfrak{Q}_M$
0		$\sqrt{3}S^{223}$		$\sqrt{6}S^{123}$		$\sqrt{3}S^{113}$		$I = 1$: $\mathfrak{T}_M$
-1			$\sqrt{3}S^{233}$		$\sqrt{3}S^{133}$			$I = \frac{1}{2}$: $\mathfrak{D}_M$
-2				S^{333}				$I = 0$: $\mathfrak{s}$

TABLE II. Reduction of $(\lambda, \mu) \otimes (3, 0)$.

	(λ', μ')	Remarks	Construction
A.	$(\lambda + 3, \mu)$	—	$S^{abc}T^{m\cdots}_{n\cdots}$
B.	$(\lambda + 2, \mu - 1)$	$\mu \neq 0$	$S^{abc}T^{m\cdots}_{an\cdots}$
C.	$(\lambda + 1, \mu - 2)$	$\mu \neq 0, 1$	$S^{abc}T^{m\cdots}_{abn\cdots}$
D.	$(\lambda, \mu - 3)$	$\mu \neq 0, 1, 2$	$S^{abc}T^{m\cdots}_{abcn\cdots}$
E.	$(\lambda + 1, \mu + 1)$	$\lambda \neq 0$	$\epsilon_{apn}S^{abc}T^{pm\cdots}_{n'\cdots}$
F.	(λ, μ)	$\lambda \neq 0; \mu \neq 0$	$\epsilon_{apn}S^{abc}T^{pm\cdots}_{bn'\cdots}$
G.	$(\lambda - 1, \mu - 1)$	$\lambda \neq 0; \mu \neq 0, 1$	$\epsilon_{apn}S^{abc}T^{pm\cdots}_{bcn'\cdots}$
H.	$(\lambda - 1, \mu + 2)$	$\lambda \neq 0, 1$	$\epsilon_{apn}\epsilon_{bqn'}S^{abc}T^{pqm\cdots}_{n''\cdots}$
J.	$(\lambda - 2, \mu + 1)$	$\lambda \neq 0, 1; \mu \neq 0$	$\epsilon_{apn}\epsilon_{bqn'}S^{abc}T^{pqm\cdots}_{cn''\cdots}$
K.	$(\lambda - 3, \mu + 3)$	$\lambda \neq 0, 1, 2$	$\epsilon_{apn}\epsilon_{bqn'}\epsilon_{crn''}S^{abc}T^{pqrm\cdots}_{n'''\cdots}$

TABLE III. Notation employed in tables of isoscalar factors.

$$
\begin{aligned}
a &= \tfrac{1}{3}(\lambda + 2\mu) + (I + \tfrac{1}{2}Y) \\
b &= \tfrac{1}{3}(2\lambda + \mu) + (I - \tfrac{1}{2}Y) \\
d &= \tfrac{1}{3}(\mu - \lambda) + (I - \tfrac{1}{2}Y) \\
e &= \tfrac{1}{3}(\lambda - \mu) + (I + \tfrac{1}{2}Y) \\
f &= \tfrac{1}{3}(2\lambda + \mu) - (I + \tfrac{1}{2}Y) \\
g &= \tfrac{1}{3}(\lambda + 2\mu) - (I - \tfrac{1}{2}Y)
\end{aligned}
$$

TABLE IV. Plan of the tables of isoscalar factors.

ΔY	ΔI	$(\lambda, \mu) \to (\lambda', \mu')$				
	$+\frac{3}{2}$	$U'(\lambda\mu$	30	$\lambda'\mu'; IY)$	$\frac{3}{2}1$	$I + \frac{3}{2}, Y + 1)$
	$+\frac{1}{2}$	$U'(\lambda\mu$	30	$\lambda'\mu'; IY)$	$\frac{3}{2}1$	$I + \frac{1}{2}, Y + 1)$
$+1$	$-\frac{1}{2}$	$U'(\lambda\mu$	30	$\lambda'\mu'; IY)$	$\frac{3}{2}1$	$I - \frac{1}{2}, Y + 1)$
	$-\frac{3}{2}$	$U'(\lambda\mu$	30	$\lambda'\mu'; IY)$	$\frac{3}{2}1$	$I - \frac{3}{2}, Y + 1)$
	$+1$	$U'(\lambda\mu$	30	$\lambda'\mu'; IY)$	10	$I + 1, Y)$
0	0	$U'(\lambda\mu$	30	$\lambda'\mu'; IY)$	10	$IY)$
	-1	$U'(\lambda\mu$	30	$\lambda'\mu'; IY)$	10	$I - 1, Y)$
	$+\frac{1}{2}$	$U'(\lambda\mu$	30	$\lambda'\mu'; IY)$	$\frac{1}{2}, -1$	$I + \frac{1}{2}, Y - 1)$
-1	$-\frac{1}{2}$	$U'(\lambda\mu$	30	$\lambda'\mu'; IY)$	$\frac{1}{2}, -1$	$I - \frac{1}{2}, Y - 1)$
-2	0	$U'(\lambda\mu$	30	$\lambda'\mu'; IY)$	$0, -2$	$I, Y - 2)$
				$N(\lambda\mu; \lambda'\mu')$		

TABLE V. Isoscalar factors for $(\lambda, \mu) \otimes (3, 0) \to (\lambda + 3, \mu)$.

ΔY	ΔI	$(\lambda, \mu) \to (\lambda + 3, \mu)$
	$+\frac{3}{2}$	$[(2I + 4)(2I + 3)(2I + 2)]^{-\frac{1}{2}}[(a + 4)(a + 3)(a + 2)(b + 4)(b + 3)(b + 2)(e + 3)(e + 2)(e + 1)]^{\frac{1}{2}}$
	$+\frac{1}{2}$	$[(2I + 3)(2I + 2)(2I)]^{-\frac{1}{2}}[3(a + 3)(a + 2)(b + 3)(b + 2)d(e + 2)(e + 1)(f + 1)(g + 1)]^{\frac{1}{2}}$
$+1$	$-\frac{1}{2}$	$[(2I + 2)(2I)(2I - 1)]^{-\frac{1}{2}}[3(a + 2)(b + 2)d(d - 1)(e + 1)(f + 2)(f + 1)(g + 2)(g + 1)]^{\frac{1}{2}}$
	$-\frac{3}{2}$	$[(2I)(2I - 1)(2I - 2)]^{-\frac{1}{2}}[d(d - 1)(d - 2)(f + 3)(f + 2)(f + 1)(g + 3)(g + 2)(g + 1)]^{\frac{1}{2}}$
	$+1$	$[(2I + 3)(2I + 2)]^{-\frac{1}{2}}[3(a + 3)(a + 2)(b + 4)(b + 3)(b + 2)(e + 2)(e + 1)(f + 1)]^{\frac{1}{2}}$
0	0	$[2I(I + 1)]^{-\frac{1}{2}}[3(a + 2)(b + 3)(b + 2)d(e + 1)(f + 2)(f + 1)(g + 1)]^{\frac{1}{2}}$
	-1	$[2I(2I - 1)]^{-\frac{1}{2}}[3(b + 2)d(d - 1)(f + 3)(f + 2)(f + 1)(g + 2)(g + 1)]^{\frac{1}{2}}$
	$+\frac{1}{2}$	$[2I + 2)^{-\frac{1}{2}}[3(a + 2)(b + 4)(b + 3)(b + 2)(e + 1)(f + 2)(f + 1)]^{\frac{1}{2}}$
-1	$-\frac{1}{2}$	$[2I]^{-\frac{1}{2}}[3(b + 3)(b + 2)d(f + 3)(f + 2)(f + 1)(g + 1)]^{\frac{1}{2}}$
-2	0	$[(b + 4)(b + 3)(b + 2)(f + 3)(f + 2)(f + 1)]^{\frac{1}{2}}$

$$N(\lambda\mu; \lambda + 3, \mu) = [(\lambda + 3)(\lambda + 2)(\lambda + 1)(\lambda + \mu + 4)(\lambda + \mu + 3)(\lambda + \mu + 2)]^{-\frac{1}{2}}$$

TABLE VI. Isoscalar Factors for $(\lambda, \mu) \otimes (3, 0) \rightarrow (\lambda + 2, \mu - 1)$.

ΔY	ΔI	$(\lambda, \mu) \rightarrow (\lambda + 2, \mu - 1)$
	$+\frac{3}{2}$	$[(2I+4)(2I+3)(2I+2)]^{-\frac{1}{2}}[3(a+3)(a+2)(b+3)(b+2)(e+3)(e+2)(e+1)fg]^{\frac{1}{2}}$
	$+\frac{1}{2}$	$[(2I+3)(2I+2)(2I)]^{-\frac{1}{2}}[(a+2)(b+2)d(e+2)(e+1)]^{\frac{1}{2}}[(a+1)(b+1)+2(\lambda+\mu+1+fg)]$
$+1$	$-\frac{1}{2}$	$[(2I+2)(2I)(2I-1)]^{-\frac{1}{2}}[d(d-1)(e+1)(f+1)(g+1)]^{\frac{1}{2}}[2(a+2)(b+1)+g(f+2)]$
	$-\frac{3}{2}$	$[(2I)(2I-1)(2I-2)]^{-\frac{1}{2}}[3(a+1)(b+1)d(d-1)(d-2)(f+2)(f+1)(g+2)(g+1)]^{\frac{1}{2}}$
	$+1$	$(-)[(2I+3)(2I+2)]^{-\frac{1}{2}}[(a+2)(b+3)(b+2)(e+2)(e+1)g]^{\frac{1}{2}}[a-2f+1]$
0	0	$[(2I+2)(2I)]^{-\frac{1}{2}}[2(b+2)d(e+1)(f+1)]^{\frac{1}{2}}[(a+2)(b+1)+g(f-a)]$
	-1	$[2I(2I-1)]^{-\frac{1}{2}}[(a+1)d(d-1)(f+2)(f+1)(g+1)]^{\frac{1}{2}}[2b-g+2]$
-1	$+\frac{1}{2}$	$(-)[2I+2]^{-\frac{1}{2}}[(b+3)(b+2)(e+1)(f+1)g]^{\frac{1}{2}}[2a-f+2]$
	$-\frac{1}{2}$	$[2I]^{-\frac{1}{2}}[(a+1)(b+2)d(f+2)(f+1)]^{\frac{1}{2}}[b-2g+1]$
-2	0	$(-)[3(a+1)(b+3)(b+2)(f+2)(f+1)g]^{\frac{1}{2}}$

$$N(\lambda\mu; \lambda+2, \mu-1) = [(\lambda+2)(\lambda+1)(\mu+1)(\lambda+\mu+4)(\lambda+\mu+2)(\lambda+\mu+1)]^{-\frac{1}{2}}$$

TABLE VII. Isoscalar factors for $(\lambda, \mu) \otimes (3, 0) \rightarrow (\lambda + 1, \mu - 2)$.

ΔY	ΔI	$(\lambda, \mu) \rightarrow (\lambda + 1, \mu - 2)$
	$+\frac{3}{2}$	$[(2I+4)(2I+3)(2I+2)]^{-\frac{1}{2}}[3(a+2)(b+2)(e+3)(e+2)(e+1)f(f-1)g(g-1)]^{\frac{1}{2}}$
	$+\frac{1}{2}$	$[(2I+3)(2I)(2I+2)]^{-\frac{1}{2}}[d(e+2)(e+1)fg]^{\frac{1}{2}}[3(a+b+1)+2ab+fg]$
$+1$	$-\frac{1}{2}$	$[(2I+2)(2I)(2I-1)]^{-\frac{1}{2}}[(a+1)(b+1)d(d-1)(e+1)]^{\frac{1}{2}}[ab+2(\lambda+\mu+fg)]$
	$-\frac{3}{2}$	$[2I(2I-1)(2I-2)]^{-\frac{1}{2}}[3(a+1)a(b+1)bd(d-1)(d-2)(f+1)(g+1)]^{\frac{1}{2}}$
	$+1$	$[(2I+3)(2I+2)]^{-\frac{1}{2}}[(b+2)(e+2)(e+1)fg(g-1)]^{\frac{1}{2}}[f-2a-3]$
0	0	$(-)[(2I+2)(2I)]^{-\frac{1}{2}}[2(a+1)d(e+1)g]^{\frac{1}{2}}[2a+ab+f(g-b-1)]$
	-1	$[2I(2I-1)]^{-\frac{1}{2}}[(a+1)a(b+1)d(d-1)(f+1)]^{\frac{1}{2}}[b-2g]$
-1	$+\frac{1}{2}$	$[2I+2]^{-\frac{1}{2}}[(a+1)(b+2)(e+1)g(g-1)]^{\frac{1}{2}}[a-2f]$
	$-\frac{1}{2}$	$(-)[2I]^{-\frac{1}{2}}[(a+1)ad(f+1)g]^{\frac{1}{2}}[2b-g+3]$
-2	0	$[3(a+1)a(b+2)(f+1)g(g-1)]^{\frac{1}{2}}$

$$N(\lambda\mu; \lambda+1, \mu-2) = [(\lambda+1)(\mu+1)\mu(\lambda+\mu+3)(\lambda+\mu+2)(\lambda+\mu)]^{-\frac{1}{2}}$$

TABLE VIII. Isoscalar factors for $(\lambda, \mu) \otimes (3, 0) \rightarrow (\lambda, \mu - 3)$.

ΔY	ΔI	$(\lambda, \mu) \rightarrow (\lambda, \mu - 3)$
	$+\frac{3}{2}$	$[(2I+4)(2I+3)(2I+2)]^{-\frac{1}{2}}[(e+3)(e+2)(e+1)f(f-1)(f-2)g(g-1)(g-2)]^{\frac{1}{2}}$
	$+\frac{1}{2}$	$[(2I+3)(2I+2)(2I)]^{-\frac{1}{2}}[3(a+1)(b+1)d(e+2)(e+1)f(f-1)g(g-1)]^{\frac{1}{2}}$
$+1$	$-\frac{1}{2}$	$[(2I+2)(2I)(2I-1)]^{-\frac{1}{2}}[3(a+1)a(b+1)bd(d-1)(e+1)fg]^{\frac{1}{2}}$
	$-\frac{3}{2}$	$[2I(2I-1)(2I-2)]^{-\frac{1}{2}}[(a+1)a(a-1)(b+1)b(b-1)d(d-1)(d-2)]^{\frac{1}{2}}$
	$+1$	$(-)[(2I+3)(2I+2)]^{-\frac{1}{2}}[3(a+1)(e+2)(e+1)f(f-1)g(g-1)(g-2)]^{\frac{1}{2}}$
0	0	$(-)[2I(I+1)]^{-\frac{1}{2}}[3(a+1)a(b+1)d(e+1)fg(g-1)]^{\frac{1}{2}}$
	-1	$(-)[2I(2I-1)]^{-\frac{1}{2}}[3(a+1)a(a-1)(b+1)bd(d-1)g]^{\frac{1}{2}}$
-1	$+\frac{1}{2}$	$[2I+2]^{-\frac{1}{2}}[3(a+1)a(e+1)fg(g-1)(g-2)]^{\frac{1}{2}}$
	$-\frac{1}{2}$	$[2I]^{-\frac{1}{2}}[3(a+1)a(a-1)(b+1)dg(g-1)]^{\frac{1}{2}}$
-2	0	$(-)[(a+1)a(a-1)g(g-1)(g-2)]^{\frac{1}{2}}$

$$N(\lambda\mu; \lambda, \mu-3) = [(\mu+1)\mu(\mu-1)(\lambda+\mu+2)(\lambda+\mu+1)(\lambda+\mu)]^{-\frac{1}{2}}$$

TABLE IX. Isoscalar factors for $(\lambda, \mu) \otimes (3, 0) \rightarrow (\lambda + 1, \mu + 1)$.

ΔY	ΔI	$(\lambda, \mu) \rightarrow (\lambda + 1, \mu + 1)$
	$+\frac{3}{2}$	$(-)[(2I+4)(2I+3)(2I+2)]^{-\frac{1}{2}}[3(a+4)(a+3)(a+2)(b+3)(b+2)(d+1)(e+2)(e+1)f]^{\frac{1}{2}}$
	$+\frac{1}{2}$	$[(2I+3)(2I+2)(2I)]^{-\frac{1}{2}}[(a+3)(a+2)(b+2)(e+1)(g+1)]^{\frac{1}{2}}[(b+3)e-2df]$
$+1$	$-\frac{1}{2}$	$[(2I+2)(2I)(2I-1)]^{-\frac{1}{2}}[(a+2)d(f+1)(g+2)(g+1)]^{\frac{1}{2}}[2(b+2)e-(d-1)f]$
	$-\frac{3}{2}$	$[2I(2I-1)(2I-2)]^{-\frac{1}{2}}[3(b+1)d(d-1)e(f+2)(f+1)(g+3)(g+2)(g+1)]^{\frac{1}{2}}$
	$+1$	$[(2I+3)(2I+2)]^{-\frac{1}{2}}[(a+3)(a+2)(b+3)(b+2)(d+1)(e+1)]^{\frac{1}{2}}[e-2f]$
0	0	$[(2I+2)(2I)]^{-\frac{1}{2}}[2(a+2)(b+2)(f+1)(g+1)]^{\frac{1}{2}}[e(b+d+3)-fd]$
	-1	$[2I(2I-1)]^{-\frac{1}{2}}[de(f+2)(f+1)(g+2)(g+1)]^{\frac{1}{2}}[2b+d+3]$
-1	$+\frac{1}{2}$	$[2I+2]^{-\frac{1}{2}}[(a+2)(b+3)(b+2)(d+1)(f+1)]^{\frac{1}{2}}[2e-f]$
	$-\frac{1}{2}$	$[2I]^{-\frac{1}{2}}[(b+2)e(f+2)(f+1)(g+1)]^{\frac{1}{2}}[b+2d+3]$
-2	0	$[3(b+3)(b+2)(d+1)e(f+2)(f+1)]^{\frac{1}{2}}$

$$N(\lambda\mu; \lambda+1, \mu+1) = [(\lambda+3)(\lambda+1)\lambda(\mu+1)(\lambda+\mu+3)(\lambda+\mu+2)]^{-\frac{1}{2}}$$

TABLE X. Isoscalar factors for $(\lambda, \mu) \otimes (3, 0) \rightarrow (\lambda, \mu)$.

ΔY	ΔI	$(\lambda, \mu) \rightarrow (\lambda, \mu)$
	$+\frac{3}{2}$	$(-)[(2I+4)(2I+3)(2I+2)]^{-\frac{1}{2}}[6(a+3)(a+2)(b+2)(d+1)(e+2)(e+1)f(f-1)g]^{\frac{1}{2}}$
	$+\frac{1}{2}$	$(-)[(2I+3)(2I+2)(2I)]^{-\frac{1}{2}}[2(a+2)(e+1)f]^{\frac{1}{2}}[(a+3)(b+1)d+dfg-(b+2)ge]$
$+1$	$-\frac{1}{2}$	$[(2I+2)(2I)(2I-1)]^{-\frac{1}{2}}[2(b+1)d(g+1)]^{\frac{1}{2}}[(a+2)(\lambda+be-df)+e(f+1)g]$
	$-\frac{3}{2}$	$[(2I)(2I-1)(2I-2)]^{-\frac{1}{2}}[6(a+1)(b+1)bd(d-1)e(f+1)(g+2)(g+1)]^{\frac{1}{2}}$
	$+1$	$[(2I+3)(2I+2)]^{-\frac{1}{2}}[2(a+2)(b+2)(d+1)(e+1)fg]^{\frac{1}{2}}[a+e-f+3]$
0	0	$(-)[(2I+2)(2I)]^{-\frac{1}{2}}[efg(2d+g-b-3)-2adfg+(a+2)df(\lambda+\mu+1)+e(g-d)$ $\times(\lambda+\mu+1)(2I+2)]$
	-1	$[2I(2I-1)]^{-\frac{1}{2}}[2(a+1)(b+1)de(f+1)(g+1)]^{\frac{1}{2}}[b+d-g]$
-1	$+\frac{1}{2}$	$(-)[2I+2]^{-\frac{1}{2}}[2(b+2)(d+1)g]^{\frac{1}{2}}[(a+1)(e-f)-(e+1)f]$
	$-\frac{1}{2}$	$(-)[2I]^{-\frac{1}{2}}[2(a+1)e(f+1)]^{\frac{1}{2}}[(b+2)(g-d)+d(g+1)]$
-2	0	$(-)[6(a+1)(b+2)(d+1)e(f+1)g]^{\frac{1}{2}}$

$$N(\lambda\mu; \lambda\mu) = [(\lambda+2)\lambda(\mu+2)\mu(\lambda+\mu+3)(\lambda+\mu+1)]^{-\frac{1}{2}}$$

TABLE XI. Isoscalar factors for $(\lambda, \mu) \otimes (3, 0) \rightarrow (\lambda - 1, \mu - 1)$.

ΔY	ΔI	$(\lambda, \mu) \rightarrow (\lambda - 1, \mu - 1)$
	$+\frac{3}{2}$	$(-)[(2I+4)(2I+3)(2I+2)]^{-\frac{1}{2}}[3(a+2)(d+1)(e+2)(e+1)f(f-1)(f-2)g(g-1)]^{\frac{1}{2}}$
	$+\frac{1}{2}$	$(-)[(2I+3)(2I+2)(2I)]^{-\frac{1}{2}}[(b+1)(e+1)f(f-1)g]^{\frac{1}{2}}[2(a+2)d-e(g-1)]$
$+1$	$-\frac{1}{2}$	$(-)[(2I+2)(2I)(2I-1)]^{-\frac{1}{2}}[(a+1)(b+1)bdf]^{\frac{1}{2}}[(a+2)(d-1)-2eg]$
	$-\frac{3}{2}$	$[2I(2I-1)(2I-2)]^{-\frac{1}{2}}[3(a+1)a(b+1)b(b-1)d(d-1)e(g+1)]^{\frac{1}{2}}$
	$+1$	$[(2I+3)(2I+2)]^{-\frac{1}{2}}[(d+1)(e+1)f(f-1)g(g-1)]^{\frac{1}{2}}[2a+e+4]$
0	0	$[(2I+2)(2I)]^{-\frac{1}{2}}[2(a+1)(b+1)fg]^{\frac{1}{2}}[d(a+e+2)-e(g-1)]$
	-1	$[2I(2I-1)]^{-\frac{1}{2}}[(a+1)a(b+1)bde]^{\frac{1}{2}}[d-2g-1]$
-1	$+\frac{1}{2}$	$(-)[2I+2]^{-\frac{1}{2}}[(a+1)(d+1)fg(g-1)]^{\frac{1}{2}}[a+2e+2]$
	$-\frac{1}{2}$	$(-)[2I]^{-\frac{1}{2}}[(a+1)a(b+1)eg]^{\frac{1}{2}}[2d-g+1]$
-2	0	$[3(a+1)a(d+1)eg(g-1)]^{\frac{1}{2}}$

$$N(\lambda\mu; \lambda-1, \mu-1) = [(\lambda+1)(\mu+2)(\mu+1)(\mu-1)(\lambda+\mu+2)(\lambda+\mu+1)]^{-\frac{1}{2}}$$

TABLE XII. Isoscalar factors for $(\lambda, \mu) \otimes (3, 0) \to (\lambda - 1, \mu + 2)$.

ΔY	ΔI	$(\lambda, \mu) \to (\lambda - 1, \mu + 2)$
	$+\frac{3}{2}$	$[(2I+4)(2I+3)(2I+2)]^{-\frac{1}{2}}[3(a+4)(a+3)(a+2)(b+2)(d+2)(d+1)(e+1)f(f-1)]^{\frac{1}{2}}$
	$+\frac{1}{2}$	$(-)[(2I+3)(2I+2)(2I)]^{-\frac{1}{2}}[(a+3)(a+2)(d+1)f(g+1)]^{\frac{1}{2}}[2(b+2)e - d(f-1)]$
$+1$	$-\frac{1}{2}$	$[(2I+2)(2I)(2I-1)]^{-\frac{1}{2}}[(a+2)(b+1)e(g+2)(g+1)]^{\frac{1}{2}}[(b+2)(e-1) - 2df]$
	$-\frac{3}{2}$	$[2I(2I-1)(2I-2)]^{-\frac{1}{2}}[3(b+1)bde(e-1)(f+1)(g+3)(g+2)(g+1)]^{\frac{1}{2}}$
	$+1$	$[(2I+3)(2I+2)]^{-\frac{1}{2}}[(a+3)(a+2)(b+2)(d+2)(d+1)f]^{\frac{1}{2}}[f-2e-1]$
0	0	$[(2I+2)(2I)]^{-\frac{1}{2}}[2(a+2)e(d+1)(g+1)]^{\frac{1}{2}}[(b+2)(e-f-1) - df]$
	-1	$[2I(2I-1)]^{-\frac{1}{2}}[(b+1)e(e-1)(f+1)(g+2)(g+1)]^{\frac{1}{2}}[b+2d+2]$
-1	$+\frac{1}{2}$	$(-)[2I+2]^{-\frac{1}{2}}[(a+2)(b+2)(d+2)(d+1)e]^{\frac{1}{2}}[2f-e+1]$
	$-\frac{1}{2}$	$[2I]^{-\frac{1}{2}}[(d+1)e(e-1)(f+1)(g+1)]^{\frac{1}{2}}[2b+d+4]$
-2	0	$[3(b+2)(d+2)(d+1)e(e-1)(f+1)]^{\frac{1}{2}}$

$$N(\lambda\mu; \lambda - 1, \mu + 2) = [(\lambda+2)(\lambda+1)(\lambda-1)(\mu+2)(\mu+1)(\lambda+\mu+2)]^{-\frac{1}{2}}$$

TABLE XIII. Isoscalar factors for $(\lambda, \mu) \otimes (3, 0) \to (\lambda - 2, \mu + 1)$.

ΔY	ΔI	$(\lambda, \mu) \to (\lambda - 2, \mu + 1)$
	$+\frac{3}{2}$	$[(2I+4)(2I+3)(2I+2)]^{-\frac{1}{2}}[3(a+3)(a+2)(d+2)(d+1)(e+1)f(f-1)(f-2)g]^{\frac{1}{2}}$
	$+\frac{1}{2}$	$[(2I+3)(2I+2)(2I)]^{-\frac{1}{2}}[(a+2)(b+1)(d+1)f(f-1)]^{\frac{1}{2}}[(a+3)d - 2eg]$
$+1$	$-\frac{1}{2}$	$(-)[(2I+2)(2I)(2I-1)]^{-\frac{1}{2}}[(b+1)bef(g+1)]^{\frac{1}{2}}[2(a+2)d - (e-1)g]$
	$-\frac{3}{2}$	$[2I(2I-1)(2I-2)]^{-\frac{1}{2}}[3(a+1)(b+1)b(b-1)de(e-1)(g+2)(g+1)]^{\frac{1}{2}}$
	$+1$	$(-)[(2I+3)(2I+2)]^{-\frac{1}{2}}[(a+2)(d+2)(d+1)f(f-1)g]^{\frac{1}{2}}[a+2e+3]$
0	0	$(-)[(2I+2)(2I)]^{-\frac{1}{2}}[2(b+1)(d+1)ef]^{\frac{1}{2}}[(a+2)d - g(a+e+1)]$
	-1	$(-)[2I(2I-1)]^{-\frac{1}{2}}[(a+1)(b+1)be(e-1)(g+1)]^{\frac{1}{2}}[g-2d]$
-1	$+\frac{1}{2}$	$[(2I+2)]^{-\frac{1}{2}}[(d+2)(d+1)efg]^{\frac{1}{2}}[2a+e+3]$
	$-\frac{1}{2}$	$(-)[2I]^{-\frac{1}{2}}[(a+1)(b+1)(d+1)e(e-1)]^{\frac{1}{2}}[2g-d]$
-2	0	$(-)[3(a+1)(d+2)(d+1)e(e-1)g]^{\frac{1}{2}}$

$$N(\lambda\mu; \lambda - 2, \mu + 1) = [(\lambda+1)\lambda(\mu+3)(\mu+1)\mu(\lambda+\mu+2)]^{-\frac{1}{2}}$$

TABLE XIV. Isoscalar factors for $(\lambda, \mu) \otimes (3, 0) \to (\lambda - 3, \mu + 3)$.

ΔY	ΔI	$(\lambda, \mu) \to (\lambda - 3, \mu + 3)$
	$+\frac{3}{2}$	$(-)[(2I+4)(2I+3)(2I+2)]^{-\frac{1}{2}}[(a+4)(a+3)(a+2)(d+3)(d+2)(d+1)f(f-1)(f-2)]^{\frac{1}{2}}$
	$+\frac{1}{2}$	$[(2I+3)(2I+2)(2I)]^{-\frac{1}{2}}[3(a+3)(a+2)(b+1)(d+2)(d+1)ef(f-1)(g+1)]^{\frac{1}{2}}$
$+1$	$-\frac{1}{2}$	$(-)[(2I+2)(2I)(2I-1)]^{-\frac{1}{2}}[3(a+2)(b+1)b(d+1)e(e-1)f(g+2)(g+1)]^{\frac{1}{2}}$
	$-\frac{3}{2}$	$[2I(2I-1)(2I-2)]^{-\frac{1}{2}}[(b+1)b(b-1)e(e-1)(e-2)(g+3)(g+2)(g+1)]^{\frac{1}{2}}$
	$+1$	$[(2I+3)(2I+2)]^{-\frac{1}{2}}[3(a+3)(a+2)(d+3)(d+2)ef(f-1)]^{\frac{1}{2}}$
0	0	$(-)[(2I+2)(2I)]^{-\frac{1}{2}}[6(a+2)(b+1)(d+2)(d+1)e(e-1)f(g+1)]^{\frac{1}{2}}$
	-1	$[2I(2I-1)]^{-\frac{1}{2}}[3(b+1)b(d+1)(g+2)(g+1)e(e-1)(e-2)]^{\frac{1}{2}}$
-1	$+\frac{1}{2}$	$(-)[2I+2]^{-\frac{1}{2}}[3(a+2)(d+3)(d+2)(d+1)e(e-1)f]^{\frac{1}{2}}$
	$-\frac{1}{2}$	$[2I]^{-\frac{1}{2}}[3(b+1)(d+2)(d+1)e(e-1)(e-2)(g+1)]^{\frac{1}{2}}$
-2	0	$[(d+3)(d+2)(d+1)e(e-1)(e-2)]^{\frac{1}{2}}$

$$N(\lambda\mu; \lambda - 3, \mu + 3) = [(\lambda+1)\lambda(\lambda-1)(\mu+3)(\mu+2)(\mu+1)]^{-\frac{1}{2}}$$

Formalization of the Lagrangian, the Hamiltonian, and Related Concepts

M. C. PEASE
Stanford Research Institute, Menlo Park, California
(Received 7 December 1964)

We consider linear system of finite dimensionality in which the system matrix is a well-behaved function of the independent variable—z or t. The system is assumed to obey a conservation law with zero signature, but is otherwise unrestricted. Using what we call the "vectorial derivative," it is shown that we can obtain Lagrangian and Hamiltonian functions in terms of which the system behavior is described in the usual way. Thus, the methods of classical physics are made available for the study of such systems, without the physical connotations usually carried by such functions.

In carrying out this formalization, we define a functional—e.g., the Hamiltonian—and, at the same time, define the operator that generates the function. It is shown that the interrelation of two functionals involves the appropriate commutator of the operators involved. For example, the Poisson bracket of a functional, F, and the Hamiltonian, H, is the functional obtained from the operator $\mathbf{\Gamma}[\mathbf{\Gamma}H, \mathbf{\Gamma}F]$, where

$$\mathbf{\Gamma} = \begin{pmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{pmatrix}$$

and $\mathbf{F}$ and $\mathbf{H}$ are the operators involved in F and H. The analogy to quantum mechanics is striking.

Finally, it is shown that the contact transformations that preserve the Hamiltonian operator form a rotational system in the Lie algebra, L, in which $\mathbf{\Gamma H}$ is embedded. That is, they obey equations of the form $\dot{\mathbf{M}} = [\mathbf{\Gamma H}, \mathbf{M}]$. It follows, then, that all $\mathbf{H}$-preserving transformations can be split into components in the intersection of the $\mathbf{\Gamma}$-unitary group and either L or a subspace derived from L.

IT is our purpose, here, to apply the methods of classical mechanics[1,2] to the analysis of linear systems describable by a vector differential equation of the form

$$d\mathbf{x}/dt = \mathbf{S}\mathbf{x}, \tag{1}$$

where $\mathbf{x}$ is a $2n$-dimensional vector represented as a column vector.[3] $\mathbf{S}$ is, then, a $2n \times 2n$ matrix whose coefficients may be functions of t. We shall at the start, assume only that the coefficients of $\mathbf{x}$ and $\mathbf{S}$ are bounded complex functions of t that obey the Lipschitz condition in the range of interest. We call $\mathbf{S}(t)$ the *system matrix*. We can, of course, use z in place of t as the independent variable, and consider Eq. (1) as the basic equation of coupled-mode theory.

Such systems are of tremendous importance in physics and engineering. When $\mathbf{S}$ is constant, its solution is, in principle, straightforward and directly calculable. When $\mathbf{S}$ is not constant, however, Eq. (1) may be quite intractable, even when the dimensionality is small, and the functional dependence of $\mathbf{S}$ on t apparently simple. Indeed, there are only a few cases where a solution in closed form is known. As illustration, we may observe that the conditions stated permit us to represent in the form of Eq. (1) any scalar linear differential equation of order $2n$ in any range excluding the singularities of its coefficients. It is of interest, therefore, that the methods of classical physics can be applied in this formal context where the concepts of energy and action may have little meaning.

The analysis given here may also be helpful in illuminating the connection between classical and quantum physics. We find, for example, that the Poisson bracket of an arbitrary functional and the Hamiltonian goes over fairly automatically into a kind of commutator of the associated operators. It therefore appears that this interrelation is a formal one that is implied by the mathematical structure. The interpretation of this interrelation in terms of possible measurements is, of course, another matter entirely.

Finally, when we investigate the contact transformations that preserve the Hamiltonian operator, and so describe essential symmetries of the system, we shall find that they have some rather remarkable properties. Usually, these symmetries are studied

[1] J. L. Synge, "Classical Dynamics," in *Handbuch der Physik*, edited by S. Flügge (Springer-Verlag, Berlin, 1960), Vol. III/1.

[2] A. Mercier, *Analytical and Canonical Formalism in Physics* (North-Holland Publishing Company, Amsterdam, 1959).

[3] We shall indicate vectors, when expressed as column matrices, by lower-case boldfaced type. Matrices will be indicated by boldfaced capitals. We will use a dagger (†) to indicate the Hermitian conjugate, or complex-conjugate transpose. The asterisk (*) will indicate the complex conjugate of a scalar.

through the infinitesimal contact transformations, which identify the Lie algebra[4] involved, and so at least partially identify the group. Quite surprisingly, we find it unnecessary to do this. We can consider an arbitrary linear transformation, without restriction to infinitesimal ones, and find that it is confined to the minimum covering Lie algebra of $\Gamma\mathbf{H}$, where Γ is a constant operator, and $\mathbf{H}$ the Hamiltonian operator. The operator $\Gamma\mathbf{H}$ is, at worst, in the Lie algebra of Γ-skew-Hermitian operators. Contact transformations in general are in the Lie group of Γ-unitary operators. Hence, we obtain the rather remarkable result that the contact transformations preserving $\mathbf{H}$ are, at most, curves in both the algebra and the group.

In developing this analysis, we initially assume nothing more than has been stated. We shall shortly, however, assume the existence of what we call the *formal Lagrangian* of a particular form. We find that this implies the existence of a conservation law of the system. In particular we find that there exists a $\mathbf{K}$ that is Hermitian and of zero signature such that, if $\mathbf{x}(t)$ is a solution of Eq. (1), then the scalar

$$K = \mathbf{x}^\dagger \mathbf{K} \mathbf{x} \tag{2}$$

is independent of t, and so describes a quantity that is conserved by the system.

The restriction of the system to one with a conservation law does not appear to be serious. Most systems of interest, at least in the idealized form that is generally the subject of analysis, do have a conservation law. As examples, we may cite systems conserving energy, number of quanta, reduced energy in the sense of either the linearized Manley–Rowe equations[5] or Chu's kinetic power law,[6] and the like.[7]

The restriction to conservation laws with zero signature is more serious. However, it does appear that the linearized approximations to most systems of physical interest do involve such laws. In coupled-mode problems, for example, where z replaces t, we find that the modes do occur in pairs with opposite parities. While this is a restriction we would prefer to avoid, it does not appear to limit unduly the practical application of these methods.

VECTORIAL DERIVATIVE

As a preliminary to our analysis, we define what we call the *vectorial derivative* with respect to a vector.

We are concerned, in general, with functionals that are quadratic forms:

$$A = \mathbf{x}^\dagger \mathbf{A} \mathbf{x} = \sum_{ij} x_i^* A_{ij} x_j. \tag{3}$$

If we consider x_i and x_i^* to be independent variables, A is a function of the $2n$ variables, x_i and x_i^*, $1 \leq i \leq n$. Hence, we can obtain

$$\frac{\partial A}{\partial x_j} = \sum_i x_i^* A_{ij} = (\mathbf{x}^\dagger \mathbf{A})_j,$$

$$\frac{\partial A}{\partial x_i^*} = \sum_j A_{ij} x_j = (\mathbf{A}\mathbf{x})_i.$$

This suggests that we construct vectors of these derivatives by defining the vectorial derivatives of Eq. (3) as

$$\partial A/\partial \mathbf{x} = \mathbf{x}^\dagger \mathbf{A}, \qquad \partial A/\partial \mathbf{x}^\dagger = \mathbf{A}\mathbf{x}. \tag{4}$$

It is these operators that we call the vectorial derivative. It is not difficult to verify that, acting on a suitable operand, they obey the usual rules of differentiation.[8]

Of more immediate interest, we can show the validity of the Euler–Lagrange equations by an adaptation of the usual development. In particular, let $\mathfrak{L}(\mathbf{u}, \mathbf{u}^\dagger, \dot{\mathbf{u}}, \dot{\mathbf{u}}^\dagger, t)$ be a functional of the indicated variables where the dot indicates the derivative with respect to time. We now require that $\mathbf{u}$, $\mathbf{u}^\dagger$, $\dot{\mathbf{u}}$, $\dot{\mathbf{u}}^\dagger$ be functions of t, so that they describe a trajectory through phase space, such that the integral $L =$

[4] An *algebra* is, briefly, a linear vector space that is also a ring. As a linear vector space, it is closed under scalar multiplication and addition. As a ring, there is also defined a product relation between any two elements that is linear in both elements, and gives an element of the ring, i.e., the ring is closed under the defined product. Division need not be possible in a ring, nor must the ring contain a unit element. Algebras are either associative or not, depending on whether the product relation is. A *Lie algebra* is a nonassociative algebra using a *Lie product*. A Lie product is one which is skew-symmetric, so that $[A, B] = -[B, A]$, and obeys the Jacobi identity: $[[A, B], C] + [[B, C], A] + [[C, A], B] = 0$. The Lie product that interests us, here, is the commutator of two matrices:

$$[\mathbf{A}, \mathbf{B}] = \mathbf{AB} - \mathbf{BA}$$

The set of all K-skew-Hermitian matrices, as defined in Eq. (13), can be easily shown to be a Lie algebra over the field of real numbers. It is both a linear vector space over the reals, and closed with respect to commutation.

As an introduction to the concept of a Lie algebra and its relation to a continuous group, see M. Hamermesh, *Group Theory* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1962). For a much deeper study, see N. Jacobson, *Lie Algebras* (Interscience Publishers, a Division of John Wiley & Sons, Inc., New York, 1962).

[5] J. M. Manley and H. E. Rowe, Proc. IRE 44, 904 (1956).

[6] H. A. Haus, MIT Rept. No. 316 (April 1958).

[7] M. C. Pease, J. Appl. Phys., 31, 1988, 2028 (1960).

[8] The vectorial derivative operators are written in the same form as ordinary partial derivatives, except for the boldface symbol indicating that the differentiation is with respect to a vector. The distinction is important, however. Unrestrained use of the formalism can lead to trouble. In particular, it is necessary to pay careful attention to the order of factors in an equation involving the vectorial derivative of a scalar function.

$\int_{t_1}^{t_2} \mathfrak{L}\, dt$ shall be stationary for any variation that preserves the endpoints. The condition for this can readily be found to be the Euler–Lagrange equations

$$\frac{\partial \mathfrak{L}}{\partial \mathbf{u}} = \frac{d}{dt}\left(\frac{\partial \mathfrak{L}}{\partial \dot{\mathbf{u}}}\right) \qquad \frac{\partial \mathfrak{L}}{\partial \mathbf{u}^\dagger} = \frac{d}{dt}\left(\frac{\partial \mathfrak{L}}{\partial \dot{\mathbf{u}}^\dagger}\right). \tag{5}$$

It is with these equations that we shall seek to describe the system.

The Lagrangian Function

We are now prepared to seek a functional $\mathfrak{L}$ such that any stationary trajectory it defines will be a solution of Eq. (1).

The Euler–Lagrange equations, Eq. (5), are second order, while Eq. (1) is first order. We can always trade order for dimensionality, however. Suppose, then, we partition $\mathbf{x}$ and $\mathbf{S}$, in Eq. (1), into

$$\mathbf{x} = \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} \qquad \mathbf{S} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{bmatrix}. \tag{6}$$

We specify that this partition must be symmetric, so that $\mathbf{u}$ and $\mathbf{v}$ are n-dimensional vectors, and the submatrices of $\mathbf{S}$ are $n \times n$ square. This does not appear to be entirely necessary, but there seems little reason to study the more general case.

Equation (1) is thus split into two equations:

$$\dot{\mathbf{u}} = \mathbf{S}_{11}\mathbf{u} + \mathbf{S}_{12}\mathbf{v}, \tag{7}$$

$$\dot{\mathbf{v}} = \mathbf{S}_{21}\mathbf{u} + \mathbf{S}_{22}\mathbf{v}. \tag{8}$$

We solve Eq. (7) for $\mathbf{v}$:

$$\mathbf{v} = \mathbf{S}_{12}^{-1}(\dot{\mathbf{u}} - \mathbf{S}_{11}\mathbf{u}). \tag{9}$$

This assumes that the partition of Eq. (6) has been made in such a way that $\mathbf{S}_{12}$ is nonsingular.

For the sake of flexibility, we introduce a constant nonsingular matrix, $\mathbf{P}$, into Eq. (8) and write it as

$$(i\mathbf{P}\mathbf{v})^{\cdot} = i\mathbf{P}\mathbf{S}_{21}\mathbf{u} + i\mathbf{P}\mathbf{S}_{22}\mathbf{v}. \tag{10}$$

We may now identify the left side of Eq. (10) as $(d/dt)(\partial \mathfrak{L}/\partial \dot{\mathbf{u}}^\dagger)$, and the right as $(\partial \mathfrak{L}/\partial \mathbf{u}^\dagger)$ in Eq. (5). We can, then, write

$$\begin{aligned} \mathfrak{L} &= i\dot{\mathbf{u}}^\dagger\mathbf{P}\mathbf{v} + i\mathbf{u}^\dagger(\mathbf{P}\mathbf{S}_{21}\mathbf{u} + \mathbf{P}\mathbf{S}_{22}\mathbf{v}) \\ &= \dot{\mathbf{u}}^\dagger(i\mathbf{P}\mathbf{S}_{12}^{-1})\dot{\mathbf{u}} + \dot{\mathbf{u}}^\dagger(-i\mathbf{P}\mathbf{S}_{12}^{-1}\mathbf{S}_{11})\mathbf{u} \\ &\quad + \mathbf{u}^\dagger(i\mathbf{P}\mathbf{S}_{22}\mathbf{S}_{12}^{-1})\dot{\mathbf{u}} + \mathbf{u}^\dagger(i\mathbf{P}\mathbf{S}_{21} - i\mathbf{P}\mathbf{S}_{22}\mathbf{S}_{12}^{-1}\mathbf{S}_{11})\mathbf{u} \end{aligned} \tag{11}$$

using Eq. (9).

The $\mathfrak{L}$ of Eq. (11) will, now regenerate Eq. (1) if we use the second form of Eq. (4). However, we also require that the first form of Eq. (4) shall generate Eq. (1). The simplest way to do this is to require that $\mathfrak{L}$ be Hermitian. If we impose this condition on Eq. (11), we find that we must have

$$\begin{aligned} \mathbf{S}_{12}^\dagger\mathbf{P} &= -\mathbf{P}^\dagger\mathbf{S}_{12}, \\ \mathbf{S}_{21}^\dagger\mathbf{P}^\dagger &= -\mathbf{P}\mathbf{S}_{21}, \\ \mathbf{S}_{11}^\dagger\mathbf{P} &= -\mathbf{P}\mathbf{S}_{22}. \end{aligned} \tag{12}$$

We can combine these constraints by saying that $\mathbf{S}$ must be $\mathbf{K}$-skew-Hermitian, as defined by

$$\mathbf{K}\mathbf{S} + \mathbf{S}^\dagger\mathbf{K} = \mathbf{0}, \tag{13}$$

for the nonsingular, constant, Hermitian matrix $\mathbf{K}$ given by

$$\mathbf{K} = \begin{bmatrix} \mathbf{O} & \mathbf{P} \\ \mathbf{P}^\dagger & \mathbf{O} \end{bmatrix}. \tag{14}$$

That $\mathbf{S}$ should be K-skew-Hermitian is the necessary and sufficient condition for the conservation of the functional given in Eq. (2). The $\mathbf{K}$ of Eq. (14) is distinguished by the fact that it has zero signature, as is most easily seen by noting that the traces of all odd powers of $\mathbf{K}$ are zero.

It will be noted that there was a good deal of arbitrariness in the development of the $\mathfrak{L}$ of Eq. (11). The partition of $\mathbf{x}$ and $\mathbf{S}$ is arbitrary as long as $\mathbf{S}_{12}$ is nonsingular. The division of Eq. (8) given in Eq. (10) is largely arbitrary. We could also have used a $\mathbf{P}$ that is not constant. The factor i in Eq. (10) was introduced so as to obtain a $\mathbf{K}$ in Eq. (14) that is Hermitian, but there are other ways of doing this. Finally, the requirement that $\mathfrak{L}$ be Hermitian is stronger than we need. All that we require is that the two forms of Eq. (4) shall both lead to Eq. (1).

The $\mathfrak{L}$ of Eq. (11) is, thus, to be regarded as a convenient Lagrangian, but not as one having any other reason for preference. The particular $\mathbf{K}$ of Eq. (14) is incidental to these arbitrary choices, which involve the choice of basis for the system. That a conservation law of the form of Eq. (2) exists at all, and that it has zero signature, are more fundamental requirements, and are probably independent of the particular way the arbitrary choices are made.

In what follows, then, we shall use the Lagrangian of Eq. (11).

THE HAMILTONIAN AND THE CANONICAL EQUATIONS

The next step is to form the Hamiltonian. For this, we define in a formal manner the appropriate conjugate momentum vector $\mathbf{p}$:

$$\mathbf{p} = \partial \mathfrak{L}/\partial \dot{\mathbf{u}}^\dagger = i\mathbf{P}\mathbf{S}_{12}^{-1}\dot{\mathbf{u}} - i\mathbf{P}\mathbf{S}_{12}^{-1}\mathbf{S}_{11}\mathbf{u}, \tag{15}$$

and similarly for $\mathbf{p}^\dagger$. If we solve this for $\dot{\mathbf{u}}$ and substitute into Eq. (11), we find that

$$\mathfrak{L} = -i\mathbf{p}^\dagger \mathbf{S}_{12}\mathbf{P}^{-1}\mathbf{p} + i\mathbf{u}^\dagger \mathbf{P}\mathbf{S}_{21}\mathbf{u}, \tag{16}$$

using Eq. (12).

We define the Hamiltonian as

$$H = \mathbf{p}^\dagger \dot{\mathbf{u}} + \dot{\mathbf{u}}^\dagger \mathbf{p} - \mathfrak{L}, \tag{17}$$

when expressed as a function of $\mathbf{u}$, $\mathbf{u}^\dagger$, $\mathbf{p}$, $\mathbf{p}^\dagger$, and t. We find, then, that

$$H = (\mathbf{u}^\dagger \mathbf{p}^\dagger)\mathbf{H}\begin{pmatrix}\mathbf{u}\\ \mathbf{p}\end{pmatrix}, \tag{18}$$

where

$$\mathbf{H} = \begin{bmatrix} -i\mathbf{P}\mathbf{S}_{21} & \mathbf{S}_{11}^\dagger \\ \mathbf{S}_{11} & -i\mathbf{S}_{12}\mathbf{P}^{-1}\end{bmatrix}. \tag{19}$$

The canonical equations that we expect H to satisfy are

$$\dot{\mathbf{u}} = \partial H/\partial \mathbf{p}^\dagger, \qquad \dot{\mathbf{u}}^\dagger = \partial H/\partial \mathbf{p}, \tag{20}$$

$$\dot{\mathbf{p}} = -\partial H/\partial \mathbf{u}^\dagger, \qquad \dot{\mathbf{p}}^\dagger = -\partial H/\partial \mathbf{u}. \tag{21}$$

We find that Eq. (20) reformulates Eq. (15) for $\mathbf{p}$. If we develop Eq. (21) and use Eqs. (15) and (9), we obtain Eq. (10), the Euler–Lagrange equation. Hence, Eqs. (20) and (21) do describe the system.

It is convenient to write the canonical equations in the form

$$\frac{d}{dt}\begin{bmatrix}\mathbf{u}\\ \mathbf{p}\end{bmatrix} = \mathbf{\Gamma}\begin{bmatrix}\partial/\partial \mathbf{u}^\dagger \\ \partial/\partial \mathbf{p}^\dagger\end{bmatrix} H(\mathbf{u}, \mathbf{p}), \tag{22}$$

where $\mathbf{\Gamma}$ is defined, as in Synge,[1] by

$$\mathbf{\Gamma} = \begin{bmatrix}\mathbf{O} & \mathbf{I}\\ -\mathbf{I} & \mathbf{O}\end{bmatrix}. \tag{23}$$

If we substitute Eq. (18) in Eq. (22) and observe that

$$\begin{bmatrix}\partial/\partial \mathbf{u}^\dagger \\ \partial/\partial \mathbf{p}^\dagger\end{bmatrix}(\mathbf{u}^\dagger \mathbf{p}^\dagger) = \mathbf{I},$$

then Eq. (22) can be written simply as

$$\frac{d}{dt}\begin{bmatrix}\mathbf{u}\\ \mathbf{p}\end{bmatrix} = \mathbf{\Gamma}\mathbf{H}\begin{bmatrix}\mathbf{u}\\ \mathbf{p}\end{bmatrix}. \tag{24}$$

Given the functional Hamiltonian H, or the matrix-valued function $\mathbf{H}$, then Eq. (22) or (24) describes the system.

POISSON BRACKET

We consider, now, an arbitrary functional $F(\mathbf{u}, \mathbf{u}^\dagger, \mathbf{p}, \mathbf{p}^\dagger, t)$. Its total derivative with time is

$$\frac{dF}{dt} = \frac{\partial F}{\partial t} + \dot{\mathbf{u}}^\dagger \frac{\partial F}{\partial \mathbf{u}^\dagger} + \frac{\partial F}{\partial \mathbf{u}}\dot{\mathbf{u}} + \dot{\mathbf{p}}^\dagger \frac{\partial F}{\partial \mathbf{p}^\dagger} + \frac{\partial F}{\partial \mathbf{p}}\dot{\mathbf{p}}. \tag{25}$$

Using Eqs. (20) and (21), we are lead to define the Poisson bracket of two functionals, F and G, as

$$\{F, G\} = \frac{\partial F}{\partial \mathbf{u}}\frac{\partial G}{\partial \mathbf{p}^\dagger} - \frac{\partial F}{\partial \mathbf{p}}\frac{\partial G}{\partial \mathbf{u}^\dagger} + \frac{\partial G}{\partial \mathbf{p}}\frac{\partial F}{\partial \mathbf{u}^\dagger} - \frac{\partial G}{\partial \mathbf{u}}\frac{\partial F}{\partial \mathbf{p}^\dagger}, \tag{26}$$

so that Eq. (25) becomes

$$dF/dt = \partial F/\partial t + \{F, H\}. \tag{27}$$

If F does not depend explicitly on t, then a necessary and sufficient condition for it to be a dynamical invariant of the system is that its Poisson bracket with H shall vanish.

If, now, F is a quadratic form,

$$F = (\mathbf{u}^\dagger \mathbf{p}^\dagger)\mathbf{F}\begin{bmatrix}\mathbf{u}\\ \mathbf{p}\end{bmatrix} = (\mathbf{u}^\dagger \mathbf{p}^\dagger)\begin{bmatrix}\mathbf{F}_{11} & \mathbf{F}_{12}\\ \mathbf{F}_{21} & \mathbf{F}_{22}\end{bmatrix}\begin{bmatrix}\mathbf{u}\\ \mathbf{p}\end{bmatrix},$$

then we can find by straightforward calculation that

$$\{F, H\} = (\mathbf{u}^\dagger \mathbf{p}^\dagger)\{\mathbf{F}\mathbf{\Gamma}\mathbf{H} - \mathbf{H}\mathbf{\Gamma}\mathbf{F}\}\begin{bmatrix}\mathbf{u}\\ \mathbf{p}\end{bmatrix}. \tag{28}$$

The term in brackets is a kind of a commutation of the operators $\mathbf{F}$ and $\mathbf{H}$. We can make it explicitly a commutator by noting that $\mathbf{\Gamma}^2 = -\mathbf{I}$, so that Eq. (28) can be written as

$$\{F, H\} = (\mathbf{u}^\dagger \mathbf{p}^\dagger)\mathbf{\Gamma}[\mathbf{\Gamma}\mathbf{H}, \mathbf{\Gamma}\mathbf{F}]\begin{bmatrix}\mathbf{u}\\ \mathbf{p}\end{bmatrix}. \tag{29}$$

Thus, we find the commutator of operators emerging as the equivalent of the formation of the Poisson bracket of the functionals. From Eq. (27) we find also that the total derivative of a functional involves the commutation of the appropriate operator and $\mathbf{\Gamma}\mathbf{H}$.

CONTACT TRANSFORMATIONS

In seeking transformations that will preserve the form of the canonical equations, we follow the usual procedure of defining a contact transformation in terms of a generating functional $T(\mathbf{u}, \mathbf{u}^\dagger, \mathbf{q}, \mathbf{q}^\dagger, t)$, where $\mathbf{q}$ and $\mathbf{q}^\dagger$ are to be the new momentum vectors. We require that T be such that

$$\mathbf{p} = \partial T/\partial \mathbf{u}^\dagger, \qquad \mathbf{p}^\dagger = \partial T/\partial \mathbf{u}, \tag{30}$$

and define the position vectors by

$$\mathbf{w} = \partial T/\partial \mathbf{q}^\dagger, \qquad \mathbf{w}^\dagger = \partial T/\partial \mathbf{q}. \tag{31}$$

Again we will confine our attention to functionals that are Hermitian quadratic forms, and write

$$T = \mathbf{u}^\dagger \mathbf{U}\mathbf{u} + \mathbf{q}^\dagger \mathbf{V}\mathbf{q} + \mathbf{u}^\dagger \mathbf{X}\mathbf{q} + \mathbf{q}^\dagger \mathbf{X}^\dagger \mathbf{u}, \tag{32}$$

where $\mathbf{U}$ and $\mathbf{V}$ are Hermitian and $\mathbf{X}$ is nonsingular. We find that

$$\begin{aligned} \mathbf{p} &= \mathbf{U}\mathbf{u} + \mathbf{X}\mathbf{q}, \\ \mathbf{w} &= \mathbf{V}\mathbf{q} + \mathbf{X}^\dagger\mathbf{u}. \end{aligned} \tag{33}$$

From these equations we obtain

$$\begin{pmatrix} \mathbf{u} \\ \mathbf{p} \end{pmatrix} = \mathbf{M}\begin{pmatrix} \mathbf{w} \\ \mathbf{q} \end{pmatrix}, \qquad \begin{pmatrix} \mathbf{w} \\ \mathbf{q} \end{pmatrix} = \mathbf{M}^{-1}\begin{pmatrix} \mathbf{u} \\ \mathbf{p} \end{pmatrix}, \tag{34}$$

where

$$\mathbf{M} = \begin{pmatrix} \mathbf{X}^{\dagger-1} & -\mathbf{X}^{\dagger-1}\mathbf{V} \\ \mathbf{U}\mathbf{X}^{\dagger-1} & \mathbf{X} - \mathbf{U}\mathbf{X}^{\dagger-1}\mathbf{V} \end{pmatrix}, \tag{35}$$

$$\mathbf{M}^{-1} = \begin{pmatrix} \mathbf{X}^\dagger - \mathbf{V}\mathbf{X}^{-1}\mathbf{U} & \mathbf{V}\mathbf{X}^{-1} \\ -\mathbf{X}^{-1}\mathbf{U} & \mathbf{X}^{-1} \end{pmatrix}. \tag{36}$$

We observe, first, that $\mathbf{M}$ and $\mathbf{M}^\dagger$ are Γ-unitary,[9] i.e.,

$$\begin{aligned} \mathbf{M}^\dagger\Gamma\mathbf{M} &= \Gamma, \\ \mathbf{M}\Gamma\mathbf{M}^\dagger &= \Gamma, \end{aligned} \tag{37}$$

as may be verified directly. The second equation can be obtained from the first by taking its inverse and noting that $\Gamma^{-1} = -\Gamma$.

Now, substituting Eq. (34) in Eq. (24), we obtain

$$\begin{aligned} \frac{d}{dt}\begin{pmatrix} \mathbf{w} \\ \mathbf{q} \end{pmatrix} &= -\mathbf{M}^{-1}\dot{\mathbf{M}}\begin{pmatrix} \mathbf{w} \\ \mathbf{q} \end{pmatrix} + \mathbf{M}^{-1}\Gamma\mathbf{H}\mathbf{M}\begin{pmatrix} \mathbf{w} \\ \mathbf{q} \end{pmatrix} \\ &= -\mathbf{M}^{-1}\dot{\mathbf{M}}\begin{pmatrix} \mathbf{w} \\ \mathbf{q} \end{pmatrix} + (\mathbf{M}^{-1}\Gamma\mathbf{M}^{\dagger-1})(\mathbf{M}^\dagger\mathbf{H}\mathbf{M})\begin{pmatrix} \mathbf{w} \\ \mathbf{q} \end{pmatrix} \\ &= \Gamma\{\Gamma\mathbf{M}^{-1}\dot{\mathbf{M}} + \mathbf{M}^\dagger\mathbf{H}\mathbf{M}\}\begin{pmatrix} \mathbf{w} \\ \mathbf{q} \end{pmatrix}, \end{aligned} \tag{38}$$

where we have used Eq. (37) and the fact that $\Gamma^2 = -\mathbf{I}$.

Hence, if we define

$$\mathfrak{H} = \Gamma\mathbf{M}^{-1}\dot{\mathbf{M}} + \mathbf{M}^\dagger\mathbf{H}\mathbf{M}, \tag{39}$$

Eq. (38) will have the same form as Eq. (24), i.e., will be canonical. The $\mathfrak{H}$ of Eq. (39) is the transformed Hamiltonian operator.

To put Eq. (39) in more familiar form, we can observe, first, that the transformation of $\mathbf{H}$ to $\mathbf{M}^\dagger\mathbf{H}\mathbf{M}$ is the transformation that changes $H(\mathbf{u}, \mathbf{p})$ to $H(\mathbf{w}, \mathbf{q})$ —i.e., a change of basis. We can also verify by direct (if tedious) calculation, that

$$\Gamma\mathbf{M}^{-1}\dot{\mathbf{M}} = \begin{pmatrix} \mathbf{X}^{-1} & \mathbf{O} \\ -\mathbf{V}\mathbf{X}^{-1} & \mathbf{I} \end{pmatrix}\begin{pmatrix} \dot{\mathbf{U}} & \dot{\mathbf{X}} \\ \dot{\mathbf{X}}^\dagger & \dot{\mathbf{V}} \end{pmatrix}\begin{pmatrix} \mathbf{X}^{\dagger-1} & -\mathbf{X}^{\dagger-1}\mathbf{V} \\ \mathbf{O} & \mathbf{I} \end{pmatrix}. \tag{40}$$

The central matrix is $\partial\mathbf{T}/\partial t$, where $\mathbf{T}$ is the operator of the generating functional, Eq. (32). The matrices on each side are those appropriate to changing the basis of $\mathbf{T}$ from $(\mathbf{u}, \mathbf{q})$ to $(\mathbf{w}, \mathbf{q})$. Hence Eq. (39) becomes, in functional form,

$$\mathfrak{H} = \partial T/\partial t + H(\mathbf{w}, \mathbf{q}). \tag{41}$$

Thus, the Hamiltonian function transforms in the usual way under a contact transformation of the form given.

INVARIANT TRANSFORMATIONS

Of particular interest are the contact transformations that leave the Hamiltonian operator unchanged. These transformations tell us the symmetry properties of the system.

This problem is usually attacked by considering the infinitesimal transformations, i.e., those such that $\mathbf{M}$ is near the identity. However, rather surprisingly, this turns out not to be necessary in the present type of problem. We suspect that this conclusion is a consequence of the linearity of Eq. (1).

We consider Eq. (39) and require that $\mathfrak{H} = \mathbf{H}$. Recalling that $\Gamma^2 = -\mathbf{I}$ and using Eq. (37), we find that we require that

$$\dot{\mathbf{M}} = \Gamma\mathbf{H}\mathbf{M} - \mathbf{M}\Gamma\mathbf{H} = [\Gamma\mathbf{H}, \mathbf{M}]. \tag{42}$$

This is a rather strange equation to find in this context. It implies that $\mathbf{M}(t)$ is involved with a Lie algebra,[4] rather than the group.

Consider $\Gamma\mathbf{H}$, as computed from Eq. (19). Using Eq. (12), we find that it is Γ-skew-Hermitian, i.e., that

$$\Gamma(\Gamma\mathbf{H}) + (\Gamma\mathbf{H})^\dagger\Gamma = \mathbf{0}. \tag{43}$$

The linear envelope of all such matrices is the Lie algebra of the infinitesimal transformations of the Lie group of Γ-unitary matrices, defined by Eq. (37). We shall call this algebra L. $\Gamma\mathbf{H}$ is, then, a curve in L.

If $\mathbf{M}(0)$ is in L, then it is easy to see that Eq. (42) implies that $\mathbf{M}(t)$ remains in L for all t. $\mathbf{M}(t)$ is, then, a curve that is in the intersection of the algebra and the group. [We have already seen, in Eq. (37), that $\mathbf{M}$ is in the group.]

If $\mathbf{M}(0)$ is not in L, then we have shown elsewhere[10]

[9] The set of K-unitary matrices, i.e., matrices for which $\mathbf{M}\dagger\mathbf{K}\mathbf{M} = \mathbf{K}$, form a group providing $\mathbf{K}$ is nonsingular.

[10] M. C. Pease, J. Math. Phys. 6, 111 (1965).

that we can split $\mathbf{M}(0)$ into two components, one of which is in L and the other is orthogonal to all of L under the Killing form. The first component generates a curve in L, independent of the second component. And the second generates a curve in a linear subspace that is orthogonal to all of L, and independent of the first component. Thus $\mathbf{M}(t)$ is split into two noninteracting curves.

It may also happen that $\Gamma\mathbf{H}$ is a curve in a proper subalgebra of L, say L_1. We can then split $\mathbf{M}$ further. In the first place L_1 may not be simple. If it is not, we can split it into simple subalgebras, and split $\mathbf{M}(t)$ accordingly. We can then split $\mathbf{M}(t)$ into noninteracting curves in each of these subalgebras.

Thus it is possible to discuss in considerable detail the possible behavior of the transformations leaving $\mathbf{H}$ invariant, knowing only the structure of the Lie algebra in which $\Gamma\mathbf{H}$ is embedded.

CONCLUSIONS

We have been able, here, to remove all physical connotations from the notions of the Lagrangian, the Hamiltonian, etc., when applied to linear systems, and so have broadened the range of their possible application and value. In so doing, we have made these methods available for the study of nonuniform or time dependent systems, which are generally intractable to the more direct forms of analysis. Presumably, we can now apply to such systems the entire body of technique that has been developed for classical mechanics.

Of immediate consequence, we have obtained means for investigating directly the existence of conservation laws of the system, or of specifying systems that exhibit specified conservation laws. Our analysis depends on the existence of one such law, given by Eq. (2) or (14). But it is of evident utility to investigate if the system obeys others as well. If so, these other laws will be found as constant solutions of Eq. (42), or as Γ-unitary matrices that commute with all of L, or L_1, the minimum covering Lie algebra of $\Gamma\mathbf{H}$. They are, in other words, Casimir operators of L_1, or of the subgroup of the Γ-unitary group that generates L_1 as the corresponding algebra.[11]

As a byproduct, we have seen the emergence of the commutator as the significant type of relation between operators. In the context of this investigation, it has been natural for us to consider, more or less simultaneously, the quadratic forms that we used to define functionals and the operators generating these forms. We have found that, where we wish to consider the interrelation of two functionals, e.g., in the Poisson bracket, the corresponding interrelation of the operators is by some appropriately modified commutator.

As a second byproduct, we have discovered that the transformations that leave the Hamiltonian operator invariant obey Eq. (42). This is an interesting result in that it shows that these transformations may be resolved into curves in the Lie algebra in which $\Gamma\mathbf{H}$ is embedded, and the linear vector space orthogonal to it. In other words, we can infer many of the consequences of the symmetry of the system from knowledge of the structure of this algebra.

ACKNOWLEDGMENT

We wish to acknowledge the support of the U.S. Air Force for part of the work reported here.

[11] G. Racah, Rend. Acad. Lincei 8, 138 (1950). See also M. Hamermesh, *Group Theory* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1962), p. 117 *et seq.*

N-Particle Kinematics and Group-Theoretical Treatment of Phase Space I. Nonrelativistic

JEAN-MARC LÉVY-LEBLOND* AND FRANÇOIS LURÇAT†
Laboratoire de Physique Théorique et Hautes Energies, Orsay, Seine et Oise, France
(Received 11 February 1965)

This paper is a group-theoretical study of the kinematics of n nonrelativistic particles. A systematic method is given to construct a new complete set of commuting observables. The method is based on the existence of a group (the "great group") which acts transitively on the phase-space manifold and preserves the phase-space volume element; the observables are then Casimir operators of the great group and of some of its subgroups, including the usual three-dimensional rotation group.

Among these collective observables, in addition to the total angular momentum, the most interesting is the "togetherness operator" which describes the simultaneous localization of the n particles. This operator is a generalization to $n > 2$ of the square of orbital angular momentum; its use allows to generalize to n particles the familiar centrifugal barrier arguments.

1. INTRODUCTION

AFTER the essential work by Wigner,[1] the relevance of the Poincaré group and its representations for particle physics has been gradually recognized. However, there remains a great number of unsolved problems, mainly concerning the application of this type of group-theoretical idea to the analysis of concrete physical situations. This is the case when one looks for a convenient description of several-particle states; a description which should clearly exhibit their essential kinematical properties. In our opinion this is *not* an academic type of problem. On the contrary we think such questions to be of great physical interest; the answers to them might give new and useful tools for the study of dynamical problems. This view has strong historical support: think of the phase-shift analysis and of the great importance this method has gained in scattering studies. As a matter of fact, this phase-shift method at once results from a kinematical study of the nonrelativistic two-body problem (though this is not the historical point of view). It is significant that only recently has such a study been consistently achieved in the relativistic case,[2,3] which, of course, presents much more difficulty.

A certain amount of work has been done concerning the three-particle case.[3,4] However, until very recently, the only methods available for three or more particles relied upon the usual trick of coupling each particle, one after the other, to the cluster of all the preceding ones. The solutions thus obtained present great disadvantages in many cases, due to to the awkwardness and asymmetry of the coupling scheme. In particular when one was to treat all the three particles on the same footing (as far as possible!), considerable difficulties were met. A good example of this type of problem is afforded by the analysis of the Dalitz plot distributions arising from three identical particles. Usually this had been done by constructing "matrix elements" with the hypothetical symmetries of the decaying particle.[5] But in this approach, one is faced with a good deal of arbitrariness, and ambiguities arise. This has not been a great hindrance up to now, since one was only interested in some rather crude properties of the analyzed state (spin and parity, essentially). But, if we want to look for finer details (say, the degree of spatial correlation or localization), we obviously need more-refined methods. What is perhaps more important, the quantum numbers used up to now to characterize three-particle states were not always physically useful, in that they did not reflect the properties of the three-particle state as a whole, but only its properties as a two-plus-one-particle state. This hindered, for instance, the generalization in a simple way of the centrifugal-barrier argument. In another area, let us recall the difficulties encountered when trying to extend the concept of Regge trajectory in the three-particle case. These might be due to a bad choice of the quantum numbers to be complexified.

In our opinion, a decisive step towards a solution

*,† Postal address: Laboratoire de Physique Théorique et Hautes Energies, Bâtiment 211, Faculté des Sciences, Orsay (S.-et-O.), France.

[1] E. P. Wigner, Ann. Math. **40**, 149 (1939).

[2] A. J. Macfarlane, J. Math. Phys. **4**, 490 (1963) and other references there; P. Moussa and R. Stora, Symposium on the Lorentz Group, Boulder, Colorado, (1964).

[3] H. Joos, Fortschr. Phys. **10**, 65 (1962).

[4] G. C. Wick, Ann. Phys. **18**, 65 (1962).

[5] L. W. Alvarez, B. C. Maglič, A. H. Rosenfeld, and M. L. Stevenson, Phys. Rev. **125**, 687 (1962); B. Barsella and E. Fabri, *ibid.* **126**, 1561 (1962); C. Zemach, *ibid.* **133**, B1201 (1964).

of many of these questions has been taken by Dragt,[6] who fully exploited preceding ideas of Smith.[7] Dragt tells us how to obtain a new basis for describing three-particle states. In his classification, the basis states are labeled, as usual, by the total energy–momentum 4-vector, the total angular momentum, and one of its components. The novelty consists in the introduction as the additional quantum numbers, instead of the energy and angular momentum of any pair of particles for instance, of two *discrete* quantum numbers with the following properties:

(a) they can be considered as labels of an irreducible representation of the SU_3 group.

(b) one of them at least is of fundamental physical significance in that it is related to the eigenvalues of the "togetherness operator." This is a dynamical variable introduced by Smith.[7] In the classical case, this is a function of the three-particle trajectories which decreases as the corresponding world lines come closer to each other. In other words, low values for this quantum number mean a highly localized (at some time) three-particle state.

We see at once all the benefit to be gained from such a concept. For instance, in many physical situations, one knows beforehand that the considered state is essentially localized within a small volume, so that only the lowest eigenvalues of the "togetherness operator" will have to be considered; indeed Dragt has shown[6] that in the $K \to 3\pi$ decay, one practically deals with the first eigenvalue only. This is to be contrasted with the standard "two-plus-one-particle" method used by Dalitz[8] in his fundamental work on this decay.

Further, Dragt's states are completely "democratic" with respect to the three particles, in that they belong to irreducible subspaces with respect to the permutation group on the three particles. In other words, they enjoy simple and explicit symmetry properties, which makes it very easy to enforce upon them the Bose–Einstein or Fermi–Dirac statistics.

It seems to us, however, that an elucidation of the meaning of Dragt's results may be useful, not only for a clarification of the three-particle case, but also for the generalization to more complicated situations ($n > 3$; relativistic case). Chacòn and Moshinsky[9] have given another derivation of Dragt's results, which is somewhat clearer, but being based upon the Schrödinger equation, does not either allow for a relativistic generalization.

The present work is devoted to a clarification of the situation, together with the formulation of a method valid for any number of particles. Here we stick to the nonrelativistic case. The problem of multiparticle states classification will be entirely formulated from a group-theoretical point of view. In the following section, we show how this problem reduces to the decomposition into irreducible components of a tensor product of irreducible representations of the Galilei group. The purely translational properties of the n-particle states are easily dealt with, being of course related to the total energy–momentum of the state. When the energy–momentum has a fixed value, the various possible configurations of the n-particles 4-momenta generate what we will call the "phase space" of the system. We show the study of this phase space, equipped with the usual "volume element," to be of fundamental importance for classifying the "internal" (nontranslational) properties of the n-particle system. In particular the angular momentum properties of the system are related to the rotation group SO_3 which acts in a natural way on phase space and keeps the measure invariant. But the nonrelativistic phase space has an obvious sphere structure and one can imbed the SO_3 group into a larger rotation group which may be called the "great group" acting now in a transitive way in phase space. We show how the n-particle states with fixed momentum–energy and angular momentum can be further classified with the help of quantum numbers corresponding to irreducible representations of the "great group" and of subgroups of it. This method is illustrated by a simple example in the Appendix. Furthermore, this great group can be defined in such a way (in fact, it is here the most natural one) that one of the quantum numbers thus introduced will always correspond to the "togetherness operator" alluded to before.[7] This means we introduce a new complete set of commuting observables for an n-particle system. This set comprises, in addition to usual observables such as the total angular momentum, new observables which describe collective properties of the system. Here is in fact the physical justification of the method.

We indicate then that any group transitive on phase-space could do the job as a "great group." One could think of determining such a group by "democracy" arguments, trying to keep all the particles on the same footing. Indeed, this is what

[6] A. J. Dragt, J. Math Phys. **6**, 533 (1965).
[7] F. T. Smith, Phys. Rev. **120**, 1058 (1960).
[8] Cf. for instance R. H. Dalitz, Repts. Progr. Phys. **20**, 163 (1957).
[9] E. Chacòn and M. Moshinsky, "The Three-Body Problem and the SU_4 Group" (University of Mexico preprint).

Dragt does[6] when he chooses to work with SU_3 instead of SO_6 (they are both transitive on the three-particle phase space). Unfortunately, this is not always possible; as a matter of fact, in many cases there exists no other group transitive on a sphere than the corresponding orthogonal group.

In a companion paper,[10] it will be shown how the general method enables us to recover Dragt's results in the three-particle case.

Finally, let us say that the method, as exposed here, can be *in principle* easily generalized to the relativistic case. However some new problems arise, of interpretation on the one hand, of computation on the other. We defer a treatment of these to a forthcoming publication.

2. KINEMATICS AND PHASE-SPACE

Quantum Kinematics and the Galilei Group

Kinematical properties of nonrelativistic particles are described by the transformation law of their states under the Galilei group. We call a particle "elementary" when its states span a representation space for an *irreducible* representation of the Galilei group. The solutions of the free-particle Schrödinger equation form such a vector space. Let

$$g = (b, \mathbf{a}, \mathbf{v}, R) \tag{1}$$

be any element of the Galilei group, acting upon space–time coordinates $(\mathbf{x}, t)$ according to

$$\begin{aligned} \mathbf{x}' &= R\mathbf{x} + \mathbf{v}t + \mathbf{a}, \\ t' &= t + b. \end{aligned} \tag{2}$$

An irreducible unitary representation of the Galilei group is labeled by a positive number m (the mass) and an integer or half-integer s (the spin).[11] It can be cast in the form

$$[U(g)\psi]_\alpha(\mathbf{p}) = \exp(i\mathbf{p}\cdot\mathbf{a} - iEb) \times \sum_\beta \mathfrak{D}^s_{\alpha\beta}(R)\psi_\beta[R^{-1}(\mathbf{p} - m\mathbf{v})] \tag{3}$$

where

$$E = \mathbf{p}^2/2m + \mathcal{V} \qquad (\mathcal{V} = \text{constant potential energy})$$

and $\mathfrak{D}^s(R)$ is the $(2s+1)$-dimensional representation of the rotation group. This representation of the Galilei group we denote by $[m \mid \mathcal{V}, s]$.

The representation is unitary with the scalar product given by

$$(\phi, \psi) = \sum_\alpha \int \bar{\phi}_\alpha(\mathbf{p})\psi_\alpha(\mathbf{p})\, d\mathbf{p}. \tag{4}$$

A convenient but improper basis for these one-particle states is given by the "vectors" $|\mathbf{p}, \alpha\rangle$ with

$$\langle \mathbf{p}_0, \alpha_0 \mid \mathbf{p}, \alpha\rangle = \delta^{(3)}(\mathbf{p} - \mathbf{p}_0)\delta_{\alpha\alpha_0}.$$

Consider now n-particle states. These obviously belong to the tensor product of the n one-particle-state spaces. A basis for n-particle wavefunctions is a basis in this tensor product space. Of course, the physicist is interested in a choice which exhibits as clearly as possible the kinematical properties of these n-particle states, i.e., their total mass, momentum, angular momentum, etc. That is to say, we want basis states belonging to Hilbert spaces irreducible with respect to the Galilei group representation. We have thus to decompose the "total" Hilbert space into a direct sum of Hilbert spaces, each of them carrying an irreducible representation of the Galilei group. Or, in other words, we must decompose the tensor product of the n individual Galilei group representations into a sum of irreducible ones.

N-Particle Systems

Let us state this problem explicitly. Consider for simplicity n particles without spin and with the same mass m. The Galilei group representation to be considered acts in the Hilbert space of functions $\psi(\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_n)$ with the scalar product:

$$(\phi, \psi) = \int \bar{\phi}(\mathbf{p}_1, \cdots, \mathbf{p}_n)\psi(\mathbf{p}_1, \cdots, \mathbf{p}_n)\, d\mathbf{p}_1 \cdots d\mathbf{p}_n \tag{5}$$

according to

$$\begin{aligned} &[U(g)\psi](\mathbf{p}_1, \cdots, \mathbf{p}_n) \\ &= \exp[i\mathbf{a}\cdot(\mathbf{p}_1 + \cdots + \mathbf{p}_n) - ib(E_1 + \cdots + E_n)] \\ &\times \psi[R^{-1}(\mathbf{p}_1 - m\mathbf{v}), R^{-1}(\mathbf{p}_2 - m\mathbf{v}), \cdots, R^{-1}(\mathbf{p}_n - m\mathbf{v})]. \end{aligned} \tag{6}$$

We now introduce the following new variables:

$$\begin{aligned} \mathbf{P} &= \mathbf{p}_1 + \mathbf{p}_2 + \cdots + \mathbf{p}_n, \\ \mathbf{q}_1 &= 2^{-\frac{1}{2}}(\mathbf{p}_2 - \mathbf{p}_1), \\ \mathbf{q}_2 &= 6^{-\frac{1}{2}}(2\mathbf{p}_3 - \mathbf{p}_2 - \mathbf{p}_1), \\ &\vdots \\ \mathbf{q}_{n-1} &= [n(n-1)]^{-\frac{1}{2}}[(n-1)\mathbf{p}_n - \mathbf{p}_{n-1} - \cdots - \mathbf{p}_1]. \end{aligned} \tag{7}$$

Then (incorrectly keeping the same symbol for the function of the new variables)

$$[U(g)\psi](\mathbf{P}, \mathbf{q}_1, \cdots, \mathbf{q}_n) = \exp(i\mathbf{a}\cdot\mathbf{P} - ibE) \times \psi[R^{-1}(\mathbf{P} - M\mathbf{v}), R^{-1}\mathbf{q}_1, \cdots, R^{-1}\mathbf{q}_{n-1}], \tag{8}$$

[10] J. M. Lévy-Leblond and M. Lévy-Nahas, J. Math. Phys. 6, 1571 (1965) (following paper).
[11] J. M. Lévy-Leblond, J. Math. Phys. 4, 776 (1963).

where

$$E = E_1 + \cdots + E_n = (1/2m)(\mathbf{p}_1^2 + \cdots + \mathbf{p}_n^2)$$

and $M = nm$. The scalar product is now

$$(\phi, \psi) = n^{-\frac{1}{2}} \int \bar{\phi}(\mathbf{P}, \mathbf{q}_1, \cdots, \mathbf{q}_{n-1}) \times \psi(\mathbf{P}, \mathbf{q}_1, \cdots, \mathbf{q}_{n-1})\, d\mathbf{P}\, d\mathbf{q}_1 \cdots d\mathbf{q}_{n-1}. \quad (9)$$

Let us introduce still another notation. Put

$$E = \mathbf{P}^2/2M + \mathcal{U} \quad (10)$$

where

$$2m\mathcal{U} = \mathbf{q}_1^2 + \cdots + \mathbf{q}_{n-1}^2 = Q^2. \quad (11)$$

But note now that the $\mathbf{q}_i$'s such that (11) is verified can be thought of as a convenient parametrization of the n-particle phase space defined by

$$\sum_i \mathbf{p}_i = \mathbf{P}, \qquad \sum_i \frac{\mathbf{p}_i^2}{2m} = E, \quad (12)$$

corresponding to total momentum–energy $(\mathbf{P}, E)$. (Note that $\mathcal{U}$ is the total energy in the center-of-momentum frame.) In other words we can choose as independent variables the total momentum–energy $(\mathbf{P}, E)$, or else the momentum $\mathbf{P}$ and the internal energy $\mathcal{U}$, and a point on the corresponding n-particle phase space. This phase space (with dimension $3n - 4$) and its generic element will be noted respectively Π_n and π. What we have now is the Galilei group representation:

$$[U(g)\psi](\mathbf{P}, \mathcal{U}, \pi) = \exp(i\mathbf{a}\cdot\mathbf{P} - ibE) \times \psi[R^{-1}(\mathbf{P} - M\mathbf{v}), \mathcal{U}, R^{-1}\pi]. \quad (13)$$

Obviously the representation of the space–time translation subgroup is thus already in reduced form. All that remains to be done, in order to decompose the representation into a sum of irreducible ones [in the form of (3)], is to reduce the representation of the rotation group [which is nothing but the "little group" of the total momentum–energy vector[11]]

$$\phi(\pi) \to \phi(R^{-1}\pi) \quad (14)$$

for each value of $\mathcal{U}$ and $\mathbf{P}$.

Angular Momentum and the Usual Coupling Scheme

Let us see first how this works in the two-particle case.[11] Phase space is then parametrized by the vector $\mathbf{q}_1$ subject to the constraint

$$\mathbf{q}_1^2 = 2m\mathcal{U} = \text{constant}. \quad (15)$$

The rotation group representation is the so-called "quasiregular" representation,[12] in the functions over the two-dimensional sphere. It is readily reduced by expanding the function in spherical harmonics:

$$\Psi(\mathbf{P}, \mathcal{U}, \mathbf{q}_1) = \sum_{l\mu} \psi_{l\mu}^{\mathcal{U}}(\mathbf{P}) Y_l^{\mu}\left(\frac{\mathbf{q}_1}{|\mathbf{q}_1|}\right). \quad (16)$$

The functions $\psi_{l\mu}^{\mathcal{U}}$ then belong to the irreducible representation $[2m|\mathcal{U}, l]$ of the Galilei group. Of course, l is the intrinsic (orbital) angular momentum of the two-particle system. We remark that each irreducible representation enters once only in the above decomposition. Therefore, the tensor product initial representation is multiplicity-free.[13]

This is the attractive feature of the two-particle case which we lose as soon as we go over to the three- (and more) particle case. The rotation group representation in the functions over phase–space is no longer multiplicity-free and it is not sufficient to give the total angular momentum in order to specify the irreducible representation of the Galilei group, or the corresponding states. Degeneracy parameters must be introduced if we are to distinguish the various equivalent representations (in fact for each angular momentum, there are an infinity of these!). The usual way of reducing the little group representation in the functions over phase–space is to expand these in spherical harmonics with respect to each of the vector $\mathbf{q}_i$'s

$$\Psi(\mathbf{P}, \mathcal{U}, \mathbf{q}_1, \cdots, \mathbf{q}_{n-1}) = \sum_{\substack{l_1\mu_1|\mathbf{q}_1| \\ \cdots\cdots \\ l_{n-1}\mu_{n-1}|\mathbf{q}_{n-1}|}} \psi^{\mathcal{U}}_{l_1\mu_1, \cdots, l_{n-1}\mu_{n-1}|\mathbf{q}_1| \cdots |\mathbf{q}_{n-1}|}(\mathbf{P}) \times Y_{l_1}^{\mu_1}(\hat{q}_1) \cdots Y_{l_{n-1}}^{\mu_{n-1}}(\hat{q}_{n-1}), \quad (17)$$

where

$$\hat{q}_i = \mathbf{q}_i/|\mathbf{q}_i|.$$

The rotation group representation is then

$$\mathfrak{D}^{l_1} \otimes \mathfrak{D}^{l_2} \cdots \otimes \mathfrak{D}^{l_{n-1}}. \quad (18)$$

By successively recoupling, with standard methods, these $(n - 1)$ angular momenta, one can completely reduce this representation. Finally, we obtain as degeneracy parameters the internal energy and angular momentum of the first k particles along with the relative ones (with respect to this cluster) of the $(k + 1)$ particle, for $k = 1, 2, \cdots, n - 1$. One immediately sees the awkwardness of such a coupling scheme in cases where it would be convenient to

[12] N. Ia. Vilenkin, Trudy Mosk. Mat. Obsc. **12**, 185 (1963).

[13] For a rigorous definition of multiplicity-free representations, see G. W. Mackey, Bull. Amer. Math. Soc. **69**, 628 (1963).

treat all the particles on the same footing, as far as it is possible!.

New Reduction Scheme

Anyway, there are other methods available for reducing the rotation group representation. We now present one of these, which in fact gives rise in the three-particle case to the Dragt[6] classification of three-particle states.

Consider anew the SO_3 representation:

$$\phi(\pi) \to \phi(R^{-1}\pi), \qquad R \in SO_3, \tag{14}$$

where π belongs to the phase space Π_n. The functions ϕ defined over phase space span a Hilbert space where the scalar product is defined with the aid of the measure on phase-space induced by the ordinary measure [see (5)] in the surrounding momentum space. But the phase space Π_n is clearly a $(3n-4)$-dimensional sphere S_{3n-4} imbedded in this momentum space and the measure on it is seen to be the uniform measure on the sphere. Let us write it for instance as

$$d\mu(\pi) = \delta(\mathbf{q}_1^2 + \cdots + \mathbf{q}_n^2 - 2mU)\, d\mathbf{q}_1 \cdots d\mathbf{q}_n. \tag{19}$$

In order for the SO_3 representation in this Hilbert space to be unitary, it has to preserve the measure (one says to be ' isobaric"), which is readily verified. Now, one quickly sees there is a much larger group of isobaric transformations. Let SO_{3n-3} be the orthogonal group preserving the sphere S_{3n-4} and $\mathcal{R}$ be its generic element. All the transformations

$$\phi(\pi) \to \phi(\mathcal{R}^{-1}\pi), \qquad \mathcal{R} \in SO_{3n-3} \tag{20}$$

are isobaric, so that the Hilbert space considered is a unitary representation space for the group SO_{3n-3}.

The SO_{p+1} group is obviously transitive[14] onto the sphere S_p. One says also S_p is an homogeneous space of SO_{p+1}, and the considered representation of SO_{p+1} in the functions over S_p is called the "quasiregular" one. Many properties of it are known;[12] in particular, this quasiregular representation is multiplicity-free: in its decomposition into irreducible components, each irreducible representation shows up once at most.

We now reduce the SO_3 representation in two steps:

(a) Firstly, we embed SO_3 in SO_{3n-3} and reduce the SO_{3n-3} representation. In virtue of its mutiplicity-free character, no degeneracy parameter is needed here.

(b) Secondly, we restrict ourselves to the SO_3 group. The SO_{3n-3} irreducible representations are no longer irreducible representations of the SO_3 subgroup. Then we reduce them. They are not, in general, multiplicity-free representations (for SO_3) and it can be necessary to introduce some degeneracy parameter.

The situation is then the following: the initial SO_3 representation is completely reduced. Each irreducible representation is labeled (in addition to its proper label as an SO_3 representation) by the label of the SO_{3n-3} representation where it comes from plus, if necessary, "inner" degeneracy parameters. This is then a new method of introducing degeneracy parameters, which, up to now, has treated alike the n particles. This method is illustrated in the Appendix by a simple but significative example.

Physical Significance

This is interesting, but could be only formal, were it not for a most important *physical* fact. One could ask the question: do the labels thus introduced, and especially the quantum numbers given by the SO_{3n-3} representation, have any physical significance? The answer is *yes*. Consider the quadratic Casimir operator for SO_{3n-3} of which one of these quantum numbers is an eigenvalue. It can be written as[15]

$$\Lambda = \sum_{\substack{i,j \\ \alpha,\beta}} (L_{i,\alpha;j,\beta})^2 \qquad \begin{cases} \alpha, \beta = 1, 2, 3 \\ i, j = 1, 2, \cdots, n-1 \end{cases} \tag{21}$$

where

$$L_{i,\alpha;j,\beta} = q_{i,\alpha}\, d_{j,\beta} - q_{j,\beta}\, d_{i,\alpha}$$

and

$$d_{k,\alpha} = i\partial/\partial q_{k,\alpha}.$$

The $L_{i,\alpha;j,\beta}$ operators constitute the Lie algebra of the SO_{3n-3} group. $\mathbf{d}_k$ is the quantum operator corresponding to the distance between the $(k+1)$th particle and the center of mass of the first k ones. The $\mathbf{d}_i$'s are given in terms of the $\mathbf{r}_i$'s, position vector for the jth particle, as are the $\mathbf{q}$'s given in terms of the $\mathbf{p}$'s [see Eq. (7)].

But consider now the classical dynamical variable corresponding to Λ. As shown by Smith,[7] this dynamical variable is a constant of the motion which vanishes if and only if the straight-line trajectories

[14] A group G is said to act transitively on a space S if, given any two points of S, there is always an element in G taking one into the other. The space S is then said to be a homogeneous space of the group G.

[15] G. Racah, "Group Theory and Spectroscopy," Princeton lectures edited as CERN report, CERN 61-8.

of the n particles cross each other in some point *at the same time*. More generally, this quantity is the smaller as the particle world lines come closer to each other.

It can thus be considered as measuring the minimal spatial extension of the n-particle system, or else as a generalized impact parameter. This is why one calls Λ the "togetherness operator." Due to this physical interpretation of the Λ operator, one expects its eigenfunctions to form a particularly useful basis when one has to analyze n-particle states with a certain degree of simultaneous spatial localization. In fact, one can generalize the usual two-particle centrifugal barrier argument, so that we know in advance that, when dealing with short-range forces, the states with the lowest eigenvalues of Λ will be dominating.

Summarizing what has been said up to now, we can state the method for constructing a basis of n-particle states according to the above scheme, as follows:

(a) look for generalized spherical harmonics, i.e., functions over the sphere S_{3n-4} (the phase–space manifold) carrying irreducible representations of the group SO_{3n-3} (transitive on the sphere S_{3n-4}),

(b) for the spherical harmonics belonging to one and the same representation of SO_{3n-3}, remove the degeneracy by classifying them with the aid of the representations of a chain of subgroups terminated by $SO_3 \supset SO_2$, i.e., the ordinary three- and two-dimensional rotation groups.

These functions will then automatically be eigenfunctions of the total angular momentum $\mathbf{J}^2$ and its component J_0, distinguished from each other by the SO_{3n-3} representation parameters and perhaps other degeneracy labels.

The Democracy Concept

It is time now to recall that, in the three-particle case, Dragt does more.[6] Indeed, his three-particle wavefunctions are labeled by the SU_3 representation parameters, and not by SO_6 ones. The reason is that, as is well known, the sphere S_{2k-1} is a homogeneous space, not only for the orthogonal group SO_{2k} but also for its subgroup SU_k. Of course the above described scheme is valid for any transitive group[16] on the phase–space manifold. In the case of SU_3, it is known[17] that its representation in the functions over the sphere S_5 is multiplicity-free and, by decomposing it, one obtains all the irreducible representations of SU_3, once for each of them. The advantage of using a transitive group smaller than SO_{3n-3}, is that in each of its irreducible representations, the multiplicity of a given SO_3 representation will be lower, so that a smaller number of additional degeneracy parameters will be needed.

Now, still in the three-particle case, the SU_3 subgroup of SO_6 can be physically characterized. In fact, the permutation group $\mathcal{S}_3$ on the three particles is a subgroup of the SO_6 group and as such acts naturally in the phase–space S_5 and induces inner automorphisms in the SO_6 group. The group elements which are stable under the action of the alternating group $\mathcal{A}_3 \subset \mathcal{S}_3$ generate a subgroup of SO_6 which is nothing else than SU_3.[6,9] As a consequence of this relationship between SU_3 and $\mathcal{S}_3$, the Dragt wavefunctions, classified according to the SU_3 representations, automatically belong to subspaces which are irreducible with respect to the operations of the $\mathcal{S}_3$ group in the space of three-particle states. In other words, any basis vector among Dragt's is also a basis vector for an $\mathcal{S}_3$ representation. This makes it very easy to construct three-particle states obeying Bose (or Fermi) statistics.

Obviously, it would be quite desirable to carry on this feature in the n-particle case. One easily shows that, except for $n = 2$, the SO_{3n-3} group itself is never "democratic" (i.e., generated by elements invariant with respect to the alternating group $\mathcal{A}_n$). One would then like to consider a "democratic" and transitive subgroup of SO_{3n-3}. But this is a hopeless program; indeed, it can be shown[18] that, on an even-dimensional sphere S_{2k}, there exists no transitive compact connected Lie group other than SO_{2k+1}. Thus, for all even-number-particle cases (beginning with 4), there exists no democratic subgroup. Even for n odd greater than 3, it seems that the democracy requirements are too strong to permit the existence of a democratic subgroup. We are then led to conclude that "democracy" is a concept with, alas, a far too short a range of applicability.

4. CONCLUSIONS

(a) We have shown how the problem of finding convenient bases for n-particle wavefunctions can be dealt with by group-theoretical methods. Indeed, it amounts to the problem of reducing a tensor product of n irreducible representations of the Galilei group (in the nonrelativistic case).

[16] The method may be generalized, with some modifications, to any group acting on the phase-space manifold (possibly in a nontransitive way) and having the usual SO_3 group as a subgroup.

[17] M. A. B. Bég and H. Ruegg, "A Set of Harmonic Functions for the Group SU_3" (Princeton preprint).

[18] A. L. Onichtchik, Dokl. Akad. Nauk. SSSR **135**, 531 (1961), and other references therein.

(b) The purely translational properties are easily treated. What remains to be done then, is to consider the rotational properties. It has been shown how at this stage, phase space naturally occurs in these group-theoretical considerations.

(c) By embedding the ordinary rotation group in a larger group acting in phase space—we have considered here a group which acts transitively—a new reduction scheme has been devised, which supersedes the usual step-by-step coupling of n particles. The n-particle states so defined are characterized by new quantum numbers, one of which describes the collective localization of the system, and which are thus of great physical interest.

(d) The "democracy" concept has been studied and it has been indicated that although our method does not at first introduce any distinction between the n particles, it is not, in general, possible to keep them on the same footing until the end of the treatment.

(e) A relativistic extension of this work will soon be published.

APPENDIX

We can illustrate the above described reduction process by the following simple example:

Consider the Hilbert space of functions on the two-dimensional sphere S_2 with the scalar product

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

$d\Omega$ being the usual invariant measure on S_2. Consider then the orthogonal group SO_2 of rotations around some fixed axis. The above Hilbert space is a representation space for this group, the representation D being defined by

$$[U(\phi)f](\Omega) = f(\Omega_\phi),$$

where $\phi \in SO_2$ and Ω_ϕ is the point on S_2 deduced from Ω by the rotation with angle $(-\phi)$ around the chosen axis. Suppose one is to reduce this representation. We follow the method indicated and embed SO_2 into the group SO_3. The considered SO_2 representation can then be viewed as the restriction to SO_2 of the SO_3 representation:

$$[U(R)f](\Omega) = f(R^{-1}\Omega), \qquad \text{where } R \in SO_3.$$

This is the quasiregular representation of SO_3. Let us call it $\mathfrak{D}$. Its reduction into irreducible representations $\mathfrak{D}^l$ of SO_3 is well known

$$\mathfrak{D} = \bigoplus_{l=0}^{\infty} \mathfrak{D}^l.$$

Let us call now $\mathfrak{D}^l_{\text{res}}$ the restriction to SO_2 of the $\mathfrak{D}^l$ representation of SO_3. This is a reducible representation of SO_2 whose reduction into irreducible (one-dimensional) representations d^m of SO_2 is naturally

$$\mathfrak{D}^l_{\text{res}} = \bigoplus_{m=-l}^{+l} d^m.$$

The contents of irreducible components of the initial SO_2 representation D is then given by

$$D = \bigoplus_{l=0}^{\infty} \mathfrak{D}^l_{\text{res}} = \bigoplus_{l=0}^{\infty} \bigoplus_{m=-l}^{+l} d^m_{(l)} = \bigoplus_{m=-\infty}^{+\infty} \bigoplus_{l=|m|}^{+\infty} d^m_{(l)}.$$

Of course each d^m representation shows up with an infinite multiplicity, i.e., once for each $l \geq m$. But the associate degeneracy is naturally removed by labeling each d^m with the index of the $\mathfrak{D}^l_{\text{res}}$ representation from which it originates. This is the meaning of the notation $d^m_{(l)}$.

As concerns a basis for the Hilbert space which exhibits this two-step reduction process, one is obviously led to consider the spherical harmonics $Y^m_l(\Omega)$—the z axis, relatively to which m is quantized, being of course the axis which defines the SO_2 group. Then the Y^m_l with l fixed span an irreducible representation of SO_3. Restricting ourselves now to SO_2, we see that Y^m_l is precisely the basis vector for the (one-dimensional) representation $d^m_{(l)}$. We have thus obtained a basis for the initial SO_2 representation which clearly exhibits its contents in irreducible components and such that degeneracy is removed by the SO_3 representation label.

Three-Particle Nonrelativistic Kinematics and Phase Space

JEAN-MARC LÉVY-LEBLOND* AND MONIQUE LÉVY-NAHAS*

Laboratoire de Physique Théorique et Hautes Energies, Orsay (Seine-et-Oise), France
(Received 11 February 1965)

The kinematics of a nonrelativistic three-particle system is studied with the help of the general method devised by Lévy-Leblond and Lurçat. Basis states are constructed which are eigenstates, in addition to the total momentum–energy, angular momentum, etc., of new observables; among these, the "togetherness tensor" describes the simultaneous localization of the three particles and therefore is of great physical interest. All of these observables arise as Casimir operators of a "great group" acting on the three-particle phase-space manifold in a transitive way, and of some of its subgroups. In the present case, by trying to keep all the particles on the same footing ("democracy" arguments), we are led to choose the SU_3 group as a particularly convenient "great group". We thus recover completely the Dragt classification of non-relativistic three-particle states. The explicit calculation of the basis functions is done in a new way, by analytical methods, solving partial derivative equations. This enables us to establish the most general form of these basis functions.

INTRODUCTION

IN the preceding paper,[1] it has been shown how group-theoretical considerations on the phase-space manifold of N particles enables one to construct a new complete set of commuting observables, thus giving rise to an interesting classification of N-particle states. We wish to apply here this method to the case $N = 3$ thus recovering Dragt's results.[2] For an account of the general theory, and of its physical meaning, the reader is referred to.[1] Let us state briefly that our goal is to construct a basis for three-particle wavefunctions carrying the irreducible representations of the Galilei group arising in the decomposition of the tensor product of the three irreducible representations corresponding to the three particles to be coupled. The translational properties are easily dealt with by fixing the total momentum–energy 4-vector. The wavefunctions of the system can now be considered as functions on the phase-space of this system, i.e., on the manifold defined by

$$\begin{aligned} \mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3 &= \mathbf{P}, \\ \mathbf{p}_1^2 + \mathbf{p}_2^2 + \mathbf{p}_3^2 &= 2mE, \end{aligned} \tag{1}$$

where $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$ are the momenta of the three particles, m is their common mass, and $(\mathbf{P}, E)$ the total momentum-energy 4-vector.

PHASE SPACE AND THE ROTATION GROUP

In order to classify the "rotational" properties of the system, we have to study the representation of the rotation group in the functions on the phase-space.

With $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$ the momenta of the three particles we consider, we define

$$\begin{aligned} \mathbf{P} &= \mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3, \\ \mathbf{q} &= (1/\sqrt{2})(\mathbf{p}_2 - \mathbf{p}_1), \\ \mathbf{q}' &= 6^{-\frac{1}{2}}(2\mathbf{p}_3 - \mathbf{p}_2 - \mathbf{p}_1). \end{aligned}$$

The phase-space is then parametrized by $(\mathbf{q}, \mathbf{q}')$ with the constraint

$$2m\mathcal{V} = \mathbf{q}^2 + \mathbf{q}'^2 = Q^2. \tag{3}$$

We deal thus with a five-dimensional sphere S_5. We have to reduce the following representation of the rotation group:

$$\psi(\mathbf{q}, \mathbf{q}') \to \psi(R^{-1}\mathbf{q}, R^{-1}\mathbf{q}'), \qquad R \in SO_3, \tag{4}$$

where ψ is any vector of the Hilbert space of square-integrable functions on the S_5 sphere with the scalar product

$$\begin{aligned} (\phi, \psi) &= \int \bar{\phi}(\mathbf{q}, \mathbf{q}')\psi(\mathbf{q}, \mathbf{q}')\delta(\mathbf{q}^2 + \mathbf{q}'^2 - Q^2)\, d\mathbf{q}\, d\mathbf{q}' \\ &= \int \phi(\pi)\psi(\pi)\, d\pi; \end{aligned} \tag{5}$$

π is any point on S_5 [i.e., a six-component vector $(\mathbf{q}, \mathbf{q}')$ with length fixed] and $d\pi$ is the invariant measure on S_5, suitably normalized. This Hilbert space is also a representation space for the SO_6 group according to

$$\psi(\pi) \to \psi(\mathcal{R}^{-1}\pi), \ \mathcal{R} \in SO_6 \tag{6}$$

THE DEMOCRATIC SUBGROUP

We consider now the group S_3 of permutations among the three particles. This is a subgroup of the

* Postal Address: Laboratoire de Physique Théorique et Hautes Energies, Bâtiment 211, Faculté des Sciences, Orsay (S.-et-O.) France.

[1] J. M. Lévy-Leblond and F. Lurçat, J. Math. Phys. 6, 1564 (1965).

[2] A. J. Dragt, J. Math. Phys. 6, 533 (1965).

SO_6 group. Indeed, with each permutation

$$\begin{pmatrix} 1 & 2 & 3 \\ i_1 & i_2 & i_3 \end{pmatrix}$$

is then naturally associated a mapping of phase-space onto itself, defined by the transformation

$$(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3) \to (\mathbf{p}_{i_1}, \mathbf{p}_{i_2}, \mathbf{p}_{i_3}).$$

Looking for instance at the two generators $\mathcal{P}_{12}$ [the (1 2) interchange permutation] and $\mathcal{C}$ [the (1 2 3) cyclic permutation] of $\mathcal{S}_3$, we obtain

$$\mathcal{P}_{12}\begin{bmatrix} \mathbf{q} \\ \mathbf{q}' \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \mathbf{q} \\ \mathbf{q}' \end{bmatrix}, \tag{7}$$

$$\mathcal{C}\begin{bmatrix} \mathbf{q} \\ \mathbf{q}' \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & -\frac{1}{2} \end{bmatrix}\begin{bmatrix} \mathbf{q} \\ \mathbf{q}' \end{bmatrix}, \tag{7'}$$

where the 6×6 orthogonal matrices $\mathcal{P}_{12}$ and $\mathcal{C}$ are written as four 3×3 blocks. Consider now the cyclic permutation $\mathcal{C}$. It generates an automorphism of SO_6, in fact an inner one since $\mathcal{C} \in SO_6$. This is the transformation

$$\mathcal{R}' = \mathcal{C}\mathcal{R}\mathcal{C}^{-1} \quad \text{for each} \quad \mathcal{R} \in SO_6.$$

We search for the SO_6 elements which are stable with respect to this automorphism—and thus for all the group of automorphisms induced by the alternating subgroug $\mathcal{A}_3 \subset \mathcal{S}_3$. These form a subgroup $D \in SO_6$, called "democratic subgroup" (we shall explain this name a few lines below) and defined by

$$d = \mathcal{C}d\mathcal{C}^{-1} \quad \text{for each} \quad d \in D. \tag{8}$$

Let us write SO_6 matrices acting on the six-dimensional vector $\pi = \begin{pmatrix} \mathbf{q} \\ \mathbf{q}' \end{pmatrix}$, as four real 3×3 matrices:

$$g = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \text{ with } gg^{\mathrm{T}} = \mathbb{1}, \text{ i.e., } \begin{cases} \alpha\alpha^{\mathrm{T}} + \beta\beta^{\mathrm{T}} = \mathbb{1} \\ \gamma\gamma^{\mathrm{T}} + \delta\delta^{\mathrm{T}} = \mathbb{1} \\ \alpha\gamma^{\mathrm{T}} + \beta\delta^{\mathrm{T}} = 0. \end{cases} \tag{9}$$

The democracy condition (8) then requires

$$\beta + \gamma = 0, \qquad \alpha - \delta = 0. \tag{10}$$

The D subgroup is thus generated by the matrices

$$d = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} \text{ with } dd^{\mathrm{T}} = \mathbb{1}, \text{ i.e., } \begin{cases} \alpha\alpha^{\mathrm{T}} + \beta\beta^{\mathrm{T}} = \mathbb{1} \\ \alpha\beta^{\mathrm{T}} - \beta\alpha^{\mathrm{T}} = 0. \end{cases} \tag{11}$$

Associate now with each $d \in D$, the 3×3 matrix

$$u(d) = \alpha + i\beta. \tag{12}$$

After (11) these matrices obey $uu^{\dagger} = \mathbb{1}$, and generate an U_3 group. But the one-to-one correspondence $u \leftrightarrow d$ is easily seen to be a group isomorphism so that in fact D is isomorphic to U_3. The way D acts on phase-space (the S_5 sphere) can now be restated by introducing the complex three-dimensional vector $\mathbf{z}(\pi) = \mathbf{q}' + i\mathbf{q}$ which spans a complex sphere. We have the correspondence

$$\pi \to d\pi \text{ implies } \mathbf{z}(\pi) \to u(d)\mathbf{z}(\pi).$$

Note that, as the subgroup SU_3 of U_3 still acts transitively on S_5, it will be sufficient, from now on, to restrict to it our attention. The advantage of the SU_3 (or U_3) group over the SO_6 group is that, due to its "democratic" nature, the functions belonging to irreducible representations of it will automatically enjoy nice symmetry properties with respect to the permutation group on the three particles. Indeed, since, by definition, each element of the "democratic" group commutes with the cyclic permutation $\mathcal{C}$, the same property will hold for the various operators constructed from its Lie algebra, operators of which the basis functions belonging to irreducible representations will have to be eigenfunctions. As a consequence of this commutativity property, these basis functions automatically will be also eigenfunctions of the cyclic permutation operator.

Finally, let us remark that the ordinary SO_3 rotation group is the subgroup of SO_6 characterized by

$$\beta = \gamma = 0, \qquad \alpha = \delta, \qquad \alpha\alpha^{\mathrm{T}} = \mathbb{1}. \tag{13}$$

Obviously it is still a subgroup of SU_3 with simply

$$u = \alpha, \qquad \alpha\alpha^{\mathrm{T}} = \mathbb{1}. \tag{13'}$$

CHOICE OF COORDINATES

What we have to do is to reduce the SU_3 representation in the functions on the sphere S_5. The result of this reduction is already known: each irreducible representation of SU_3 appears exactly once in the decomposition. Bég and Ruegg[3] have accomplished this reduction and introduced a set of harmonic functions on the sphere S_5 which carry these irreducible representations. However, their explicit results are of no use to us. Indeed, in order to obtain states of definite angular momentum, what we need are basis states which belong to irreducible subspaces with respect to the SO_3 group, the representation of which we, finally, have to reduce. In other terms, we define the basis functions we are

[3] M. A. B. Bég and H. Ruegg, J. Math. Phys. **6**, 677 (1965).

looking for, as eigenfunctions of a complete set of commuting observables which has to comprise invariant operators of SU_3 and the usual total angular momentum observables (say, $\mathbf{J}^2$ and J_0) associated to the chain $SO_3 \supset SO_2$ of SU_3 subroups, plus others, if necessary. The chain $SU_3 \supset SO_3 \supset SO_2$ is *not* the one discussed by Bég and Ruegg. Their harmonic functions are in fact classified with the help of the more familiar chain $SU_3 \supset SU_2 \supset U_1$. The first thing to do is to introduce a convenient coordinate system on the sphere S_5. This has been discussed by Dragt.[2] A given configuration $(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3)$ is parametrized[4]:

(1) by its projection onto the Dalitz plot, i.e., by the two polar coordinates (ρ, φ) such that

$$\mathbf{p}_k^2 = \tfrac{1}{3}Q^2(1 + \rho\xi_k), \qquad \xi_k = \cos(\varphi + \tfrac{2}{3}k\pi), \qquad k = 1, 2, 3. \tag{14}$$

(2) by the rotation necessary to transform the momentum triangle from some "reference" orientation to its actual one. The most elegant way of doing this is to consider the momentum triangle as a solid by fixing unit masses to each vertex. Its principal axes of inertia define three orthogonal unit vectors $(\mathbf{u}, \mathbf{v}, \mathbf{w})$ [$\mathbf{u}$, $\mathbf{v}$ in the momentum triangle plane $\mathbf{w}$ orthogonal to it]. Let $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ be three orthonormal reference vectors. We choose as a second set of parameters the Euler angles (α, β, γ) of the rotation needed to carry $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ on $(\mathbf{u}, \mathbf{v}, \mathbf{w})$. Defining

$$\rho = \cos 2\psi, \qquad 0 \leq \psi \leq \tfrac{1}{4}\pi, \tag{15}$$

it can be shown that

$$\begin{aligned} \mathbf{q} &= Q(\cos\psi \sin\tfrac{1}{2}\varphi\mathbf{u} - \sin\psi \cos\tfrac{1}{2}\varphi\mathbf{v}), \\ \mathbf{q}' &= Q(\cos\psi \cos\tfrac{1}{2}\varphi\mathbf{u} + \sin\psi \sin\tfrac{1}{2}\varphi\mathbf{v}), \end{aligned} \tag{16}$$

or else

$$\mathbf{z} = Qe^{\frac{1}{2}i\varphi}(\cos\psi\mathbf{u} - i\sin\psi\mathbf{v}), \tag{17}$$

with

$$\begin{bmatrix} \mathbf{u} \\ \mathbf{v} \\ \mathbf{w} \end{bmatrix} = \begin{bmatrix} \cos\gamma\cos\beta\cos\alpha - \sin\gamma\sin\alpha & -\sin\gamma\cos\beta\cos\alpha - \cos\gamma\sin\alpha & \sin\beta\cos\alpha \\ \cos\gamma\cos\beta\sin\alpha + \sin\gamma\cos\alpha & -\sin\gamma\cos\beta\sin\alpha + \cos\gamma\cos\alpha & \sin\beta\sin\alpha \\ -\cos\gamma\sin\beta & \sin\gamma\sin\beta & \cos\beta \end{bmatrix} \begin{bmatrix} \mathbf{i} \\ \mathbf{j} \\ \mathbf{k} \end{bmatrix}. \tag{18}$$

This set of coordinates has the great advantage that, acting on phase-space with any SO_3 element, R modifies only the (α, β, γ) coordinates. Indeed, call R_0 the (α, β, γ) rotation. We see at once that:

$$\mathbf{z} \to R\mathbf{z} \quad \text{implies} \quad R_0 \to RR_0.$$

In other terms, the orbits of the SO_3 group in phase-space are isomorphic to SO_3 itself. The (φ, ψ) parameters label the different orbits, while the (α, β, γ) parameters label the different points of a given orbit. We have yet to describe the measure on phase space with the help of these coordinates. A simple Jacobian calculation gives

$$d\mathbf{p}_1\, d\mathbf{p}_2\, d\mathbf{p}_3 \approx d\mathbf{P}Q^5\, dQ\rho\, d\rho\, d\varphi\, dR, \tag{19}$$

where dR is the invariant volume element for SO_3,

$$dR = d\alpha\, |d(\cos\beta)|\, d\gamma. \tag{20}$$

On the S_5 sphere itself the measure is then defined by

$$d\pi = \rho\, d\rho\, d\varphi\, dR. \tag{21}$$

It is clear that by projecting onto the Dalitz diagram, one gets the usual uniform measure in the plane, $\rho d\rho d\varphi$.

[4] In what follows, we stick to the center-of-mass frame, i.e., the condition $\mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3 = \mathbf{P} = 0$ is always fulfilled. Since pure Galilei transformations leave phase-space invariant, this is no restriction.

THE DIFFERENTIAL OPERATORS

We are now in position to write as differential operators the various operators of which the basis functions we look for are to be eigenfunctions.

The SU_3 representation is characterized by the eigenvalues of:

(a) The Laplace–Beltrami operator[5] on the sphere S_5 (quadratic Casimir operator for SU_3)—which is also the "togetherness operator"[6]— is obtained by a tedious calculation:

$$\begin{aligned} \Delta = {} & \frac{\partial^2}{\partial\psi^2} + 4\cot 4\psi \frac{\partial}{\partial\psi} \\ & + \frac{4}{\cos^2 2\psi}\frac{\partial^2}{\partial\varphi^2} + 4\frac{\sin 2\psi}{\cos^2 2\psi}\frac{\partial^2}{\partial\varphi\,\partial\alpha} \\ & + \left(\frac{1}{\cos^2 2\psi} + 2\frac{1 + \cos 2\psi\cos 2\alpha}{\sin^2 2\psi}\frac{\cos^2\beta}{\sin^2\beta}\right)\frac{\partial^2}{\partial\alpha^2} \\ & - 2\frac{\cos 2\psi \sin 2\alpha}{\sin^2 2\psi}\frac{1 + \cos^2\beta}{\sin^2\beta}\frac{\partial}{\partial\alpha} \end{aligned}$$

[5] See, for instance, T. Y. Thomas, *Concepts from Tensor-Analysis and Differential Geometry* (Academic Press Inc., New York, 1961).

[6] What we call the "togetherness operator" is the squared magnitude of the "grand angular momentum tensor" introduced by Smith.[7] See Refs. 1 and 2 for its relation to the Casimir operator of SU_3.

[7] F. T. Smith, Phys. Rev. **120**, 1058 (1960).

$$+ 2\frac{1 - \cos 2\psi \cos 2\alpha}{\sin^2 2\psi}\frac{\partial^2}{\partial\beta^2}$$

$$+ 2\frac{1 + \cos 2\psi \cos 2\alpha}{\sin^2 2\psi}\frac{\cos\beta}{\sin\beta}\frac{\partial}{\partial\beta}$$

$$+ 2\frac{1 + \cos 2\psi \cos 2\alpha}{\sin^2 2\psi}\frac{1}{\sin^2\beta}\frac{\partial^2}{\partial\gamma^2}$$

$$+ 4\frac{\cos 2\psi \cos 2\alpha}{\sin^2 2\psi}\frac{\cos\beta}{\sin^2\beta}\frac{\partial}{\partial\gamma}$$

$$+ 4\frac{\cos 2\psi \sin 2\alpha}{\sin^2 2\psi}\frac{\cos\beta}{\sin\beta}\frac{\partial^2}{\partial\alpha\,\partial\beta}$$

$$- 4\frac{1 + \cos 2\psi \cos 2\alpha}{\sin^2 2\psi}\frac{\cos\beta}{\sin^2\beta}\frac{\partial^2}{\partial\alpha\,\partial\gamma}$$

$$- 4\frac{\cos 2\psi \sin 2\alpha}{\sin^2 2\psi}\frac{1}{\sin\beta}\frac{\partial^2}{\partial\beta\,\partial\gamma}. \qquad (22)$$

It looks quite a formidable task to diagonalize this monster, but in fact, we shall find many simplifications.[8]

(b) A linear operator[3]

$$C = \mathbf{z}\cdot\frac{\partial}{\partial\mathbf{z}} - \mathbf{z}^*\cdot\frac{\partial}{\partial\mathbf{z}^*} = \frac{2}{i}\frac{\partial}{\partial\varphi}. \qquad (23)$$

If the harmonic function F belongs to the (p, q) representation of SU_3, then

$$\Delta F = -\lambda(\lambda + 4)F, \qquad \lambda = p + q, \qquad (24)$$

$$CF = \mu F, \qquad \mu = p - q. \qquad (24')$$

Now, in order to distinguish the various harmonic functions belonging to a given SU_3 representation we have chosen to diagonalize

$$\mathbf{J}^2 = -\frac{1}{\sin\beta}\frac{\partial}{\partial\beta}\left(\sin\beta\frac{\partial}{\partial\beta}\right) - \frac{1}{\sin^2\beta}\left(\frac{\partial^2}{\partial\alpha^2} - 2\cos\beta\frac{\partial^2}{\partial\alpha\,\partial\gamma} + \frac{\partial^2}{\partial\gamma^2}\right), \qquad (25)$$

$$J_0 = i\partial/\partial\gamma, \qquad (26)$$

corresponding to the subgroup chain $SU_3 \supset SO_3 \supset SO_2$. As discussed by Dragt,[2] the restriction to SO_3 of a given SU_3 irreducible representation is not multiplicity-free and a further degeneracy parameter is needed. A known cubic operator Ω can play this role. However the low-dimension SO_3 irreducible representations ($J = 0, 1$) show up only once[9] in the restriction of a SU_3 representation to SO_3. For the time being we will limit ourselves to these J values, so that we will not have to consider Ω.

SOLVING THE EIGENVALUE PROBLEM

Consider now the eigenfunctions of $\mathbf{J}^2$ and J_0. After the Peter–Weyl theorem,[10] any square-integrable functions (with respect to the invariant measure) on a compact group can be expanded on the orthonormal basis of the group constituted by the matrix elements of the irreducible representations. Thus, for *any* function on S_5, we can write

$$F(\rho, \varphi, R) = \sum_{jmm'} \mathcal{G}^j_{mm'}(\rho, \varphi)\mathfrak{D}^j_{mm'}(R), \qquad (27)$$

where R stands for the SO_3 element (α, β, γ). Now if F is to be an eigenfunction of $\mathbf{J}^2$ and J_0 [with eigenvalues $L(L + 1)$ and M], it is necessarily of the form

$$F_{LM}(\rho, \varphi, R) = \sum_m \mathcal{G}^L_{mM}(\rho, \varphi)\mathfrak{D}^L_{mM}(R). \qquad (28)$$

We now deal immediately with the φ dependence. Indeed, in order for F to be an eigenfunction of C (with eigenvalue μ), it must be written

$$F^\mu_{LM}(\rho, \varphi, R) = e^{\frac{1}{2}i\mu\varphi}\sum_m g^L_{mM}(\rho)\mathfrak{D}^L_{mM}(R). \qquad (29)$$

We have now determined the general form of the functions we are looking for.

All that remains to be done is to determine the $g^L_{mM}(\rho)$ functions. In order to do it, one substitutes (29) into (22). After having applied the differential operator Δ, one expands the rhs on the SO_3 representation matrix elements anew and identifies it term by term with the lhs expansion. One thus gets a coupled system of second-order differential linear equations for the g's. The task is further reduced by the following remarks:

The lowering and raising operators $J_\pm$ commute with Δ, so that the $g^L_{mM}(\rho)$ are in fact independent of M. It suffices then to do the job for $M = 0$ (no more γ dependence).

Here parity is, of course, a good quantum number, commuting with all the operators we already considered. As seen from (18) the parity operation corresponds to the substitution

$$P: \alpha \to \pi + \alpha \ (\rho, \varphi, \beta, \gamma \text{ unchanged}). \qquad (30)$$

For a given L, the g's then split into two subsets corresponding to $P = +1$ (m even) and $P = -1$ (m odd), whose differential systems are uncoupled.

[8] It must be said that there exist several papers dealing with the eigenvalues and eigenvectors of this operator. However, the coordinate system used are very inconvenient for our purposes—although they lead to an expression easier to handle. See, for instance: J. D. Louck, J. Mol. Spectr. 4, 298 (1960); K. D. Granzow, J. Math. Phys. 4, 897 (1963); 5, 1474 (1964).

[9] G. Racah, Rev. Mod. Phys. 21, 494 (1949).

[10] See for instance: L. Pontrjagin, *Topological Groups* (Princeton University Press, Princeton, New Jersey, 1939), Sec. 29.

This process can be illustrated in the two simplest cases where only one g function is to be considered.

EXAMPLE A: $L^P = 0^+$

The functions have the following form:

$$F^{\lambda;\mu}_{0^+;0}(\rho, \varphi, R) = e^{\frac{1}{2}i\mu\varphi} g(\rho). \tag{31}$$

The g function has to obey the differential equation

$$(1 - \rho^2)g'' + \left(\frac{1}{\rho} - 3\rho\right)g' + \tfrac{1}{4}[\lambda(\lambda + 4) - \mu^2/\rho^2]g = 0. \tag{32}$$

This is easily solved by putting

$$g = \rho^{\frac{1}{2}|\mu|} h(\rho^2). \tag{33}$$

We obtain for the h function, the equation

$$x(1 - x)h'' + [1 + \tfrac{1}{2}|\mu| - (2 + \tfrac{1}{2}|\mu|)x]h' + \tfrac{1}{16}(\lambda - |\mu|)(\lambda + |\mu| + 4)h = 0, \tag{34}$$

whose solutions are Jacobi polynomials[11]:

$$h(x) = \mathcal{P}^{(\lambda-|\mu|)}_{\frac{1}{2}(\lambda-|\mu|)}(1 + \tfrac{1}{2}|\mu|; 1 + \tfrac{1}{2}|\mu|; x). \tag{35}$$

Finally, the complete solution reads

$$F^{\lambda;\mu}_{0^+;0}(\rho, \varphi, R) = e^{\frac{1}{2}i\mu\varphi}\rho^{\frac{1}{2}|\mu|}\mathcal{P}^{(\lambda-|\mu|)}_{\frac{1}{2}(\lambda-|\mu|)} \times (1 + \tfrac{1}{2}|\mu|; 1 + \tfrac{1}{2}|\mu|; x^2). \tag{36}$$

It is readily verified that for different (λ, μ) values, the corresponding functions, due to the properties of Jacobi polynomials, are othogonal with respect to the scalar product induced by the measure (21).

EXAMPLE B: $L^P = 1^+$

The basis functions have the following form:

$$F^{\lambda;\mu}_{1^+;M}(\rho, \varphi, R) = e^{\frac{1}{2}i|\mu|\varphi} g(\rho)\mathfrak{D}^1_{0M}(R). \tag{37}$$

As a function of ψ, g obeys

$$\frac{d^2g}{d\psi^2} + 4\cot 4\psi \frac{dg}{d\psi} + \left[\lambda(\lambda + 4) - \frac{\mu^2}{\cos^2 2\psi} - \frac{4}{\sin^2 2\psi}\right]g = 0. \tag{38}$$

Singular terms in the last coefficient are dealt with by putting

$$g = \sin 2\psi \rho^{\frac{1}{2}|\mu|} k(\rho^2). \tag{39}$$

We then obtain for $k(x)$ the equation

$$x(1 - x)k'' + [1 + \tfrac{1}{2}|\mu| - (3 + \tfrac{1}{2}|\mu|)x]k' + \tfrac{1}{4}[\lambda(\lambda + 4) - |\mu|(|\mu| + 8) - 12]k = 0, \tag{40}$$

whose solutions are still Jacobi polynomials[11]:

$$k(x) = \mathcal{P}^{(\lambda-|\mu|-2)}_{\frac{1}{2}(\lambda-|\mu|-2)}(2 + \tfrac{1}{2}|\mu|; 1 + \tfrac{1}{2}|\mu|; x). \tag{41}$$

The complete solution now appears as

$$F^{\lambda;\mu}_{1^+;N}(\rho, \varphi, R) = e^{\frac{1}{2}i\mu\varphi} \sin 2\psi(\cos 2\psi)^{\frac{1}{2}|\mu|} \times \mathcal{P}^{(\lambda-|\mu|-2)}_{\frac{1}{2}(\lambda-|\mu|-2)}(2 + \tfrac{1}{2}|\mu|; 1 + \tfrac{1}{2}|\mu|; \cos^2 2\psi)\mathfrak{D}^1_{0M}(R). \tag{42}$$

These solutions are the ones found by Dragt[2] (with a slightly different definition of Euler angles).

CONCLUSIONS

It has been shown how the general considerations of Ref. 1 can be readily applied to the three particle case, leading to a classification of three-particle states according to the SU_3 representations, which is the one introduced by Dragt.[2]

The explicit calculation of the basis functions in this scheme has been undertook by analytical methods, the eigenvalue problem of our complete set of commuting observables being reduced to the solution of partial derivative equations. We have thus been able to establish the general form of these basis functions [see (29)], what had not been done before.

However, except in some simple cases, the complete solution of the problem is quite complicated and algebraic methods seem to be more powerful.[2,12]

For a discussion of the symmetry properties of the basis functions thus obtained (which is a most important topic), the reader is referred to Dragt's original work.[2] Dragt indicates also how to generalize the method used here to the case of three particles with unequal masses.

As a last remark, the possibility exists to still generalize it to the case of three particles with spin, while retaining the nice symmetry properties of the basis functions. It suffices to consider an L–S coupling between the total angular momentum (introduced as before) and the total spin, this one being obtained by a symmetrical coupling of the three individual spins.[13]

ACKNOWLEDGMENTS

It is a pleasure to thank F. Lurçat for many helpful comments and for his continued interest in this work.

[11] H. Hochstadt, *Special Functions of Mathematical Physics*, (Hott, Rinehart and Winston, New York, 1961).

[12] E. Chacòn and M. Moshinsky, "The Three-Body Problem and the SU_4 Group" (University of Mexico preprint, 1964).

[13] A. Chakrabarti, Ann. Inst. Henri Poincaré 1, 301 (1964). See also Ref. 1.

S-Matrix Theory and Double Scattering

DANIEL IAGOLNITZER
Centre d'Etudes Nucléaires de Saclay Service de Physique Théorique, Gif-sur-Yvette (S.-et-O.) France,
and
Ecole Nationale des Ponts et Chaussées, Paris, France
(Received 23 November 1964)

In the double scattering case, we show that the introduction of a pole in the S-matrix elements, is consistent with a description in terms of successive interactions.

We then show that among Olive's class of possible "propagators" $c(k^2 - m^2 + i\epsilon)^{-1} - (1 - c) \cdot (k^2 - m^2 - i\epsilon)^{-1}$, only the usual value $c = 1$ gives a consistent result.

I. INTRODUCTION

We consider double scattering and study the relationship between the propagation of an intermediate real particle (phenomenological description) and the propagator pole corresponding to it in the framework of S-matrix theory.

What we want to prove is that the existence of the conventional pole in S-matrix theory is consistent with a phenomenological description in terms of successive scatterings. This consistency is a physical requirement which, as far as we know, has only been verified in theoreies where it is possible to follow the evolution of the physical process in space–time.[1]

The first thing we need is a correspondence between the ways of describing things in S-matrix and phenomenological formalisms, in order to compare S-matrix and phenomenological results. In S-matrix theory we use wave packets in momentum space, in order to avoid divergence problems associated with the presence of the pole. On the other hand, we use currents (probability density and current density) in the phenomenological description of the motion of particles. We adopt the usual correspondence between wave packets and currents.

We now note that the quantities we have to compare are not scattering amplitudes, but transition probabilities, since these are the quantities that appear in the phenomenological description.

In Sec. II, we study simple scattering. This is fairly trivial, but we use it to exhibit the formalism. Using the conventional relation between S-matrix elements and cross section, we verify that the phenomenological formula, which links the transition probability to the currents and the cross section, gives the same result as the S-matrix formula. This is shown under some reasonable conditions, usual in quantum mechanics.[2]

In Sec. III, we study double scattering. On the one hand, we obtain the S-matrix formula with the usual pole and by using the factorization of the residue into the product of the amplitudes for the individual scatterings. On the other hand, we get the phenomenological formula by writing the transition probability for the first scattering, the current of the intermediate particle and the transition probability for the second scattering. The two formulas, which are quite different, disagree for finite separation between the two targets. However, we find that in the limit of infinite separation, the two formulas give the same result. This is sufficient to satisfy the physical consistency requirement.

We finally show on a simple example that this result is not true for an arbitrary propagator satisfying unitarity plus cluster decomposition. In fact, among Olive's class[3]

$$\frac{c}{k^2 - m^2 + i\epsilon} - \frac{1 - c}{k^2 - m^2 - i\epsilon},$$

where c is an arbitrary real number, only the value 1 of c agrees with the phenomenological result (or zero if we had used opposite sign convention in the Fourier transform). Eden and Landshoff[4] also recently tried to show that the value 1 of c was fixed by a causality requirement. But we find their argument unconvincing not only because it lacks mathematical rigor, as the authors themselves recognize but because it makes ad hoc use of quantities such as negative energies, off-mass-shell elements, etc., ···

[1] See A. Messiah, *Quantum Mechanics* (North-Holland Publishing Company, Amsterdam, and Interscience Publishers, Inc., New York, 1961), Vol. 2, Chap. 19 for the study of the Schrödinger case.

[2] The problem was first studied by Chew and Low. See Ref. 1, Vol. 1, Chap. 10.

[3] D. I. Olive, "The Exploration of S-Matrix Theory", (Cambridge University and Carnegie Institute of Technology Preprints).

[4] R. J. Eden and P. V. Landshoff, "The Problem of Causality in S-Matrix Theory." (Cambridge University preprint).

which are foreign to S-matrix theory, and which we do not find necessary to use.

The particles are supposed spinless and with masses different from zero.

II. THE SIMPLE SCATTERING PROBLEM

A. S-matrix result

We first calculate the transition probability using the S-matrix formalism. A state of one spinless particle is characterized by a wavefunction $\varphi(\mathbf{p})$ where $\mathbf{p}$ is the 3-momentum of the particle and is written as

$$|a\rangle = \int \frac{\varphi_a(\mathbf{p})}{2p_0} |\mathbf{p}\rangle \, d\mathbf{p} \quad \text{with} \quad p_0 = (\mathbf{p}^2 + m^2)^{\frac{1}{2}}.$$

The normalization is

$$\langle \mathbf{p}' \mid \mathbf{p} \rangle = 2p_0 \delta(\mathbf{p} - \mathbf{p}'), \qquad \int \frac{|\varphi(\mathbf{p})|^2}{2p_0} d\mathbf{p} = 1 \tag{1}$$

so that $\langle a|a\rangle = 1$.

In what follows, $\varphi_a(\mathbf{p})(\varphi_b(\mathbf{q}), \varphi_c(\mathbf{r}))$ will be taken infinitely differentiable and vanishes outside a sphere with center at $\mathbf{P}$, $(\mathbf{Q}, \mathbf{R})$ and diameter d. We shall ultimately have to take limits when d goes to zero, and then define our wavefunctions as

$$\begin{aligned} \varphi_a(\mathbf{p}) &= c_a \Phi_a[(\mathbf{p} - \mathbf{P})/d + \mathbf{P}], \\ \varphi_b(\mathbf{q}) &= c_b \Phi_b[(\mathbf{q} - \mathbf{Q})/d + \mathbf{Q}], \\ \varphi_c(\mathbf{r}) &= c_c \Phi_c[(\mathbf{r} - \mathbf{R})/d + \mathbf{R}], \end{aligned} \tag{1'}$$

where c_a, c_b, c_c are normalization coefficients and Φ_a, Φ_b, Φ_c are fixed functions, vanishing outside a sphere with diameter 1.

One integrates over the final states with the measure $E(\mathbf{p}_1\mathbf{q}_1)\, d\mathbf{p}_1/2p_{10}\, d\mathbf{q}_1/2q_{10}$, where $\mathbf{p}_1$ and $\mathbf{q}_1$ are the momenta of the final particles and E characterizes the efficiency of the measurement apparatus. We finally define T through $S = 1 + iT$ and t through

$$\langle i| \, T \, |f\rangle = \langle i| \, t \, |f\rangle \delta(P_f - P_i). \tag{2}$$

P_f and P_i being the momenta in the f and i states.

We then write

$$W = \int E(\mathbf{p}_1\mathbf{q}_1) \frac{d\mathbf{p}_1}{2p_{10}} \frac{d\mathbf{q}_1}{2q_{10}} |\langle ab| \, T \, |p_1q_1\rangle|^2 ,$$

$$\begin{aligned} W = \int \frac{E(\mathbf{p}_1\mathbf{q}_1)}{2p_{10}2q_{10}} d\mathbf{p}_1 \, d\mathbf{q}_1 \langle p'q'| \, t' \, |p_1q_1\rangle\langle p_1q_1| \, t \, |pq\rangle \\ \times \delta^{(4)}(p + q - p_1 - q_1)\varphi_a^*(\mathbf{p}')\varphi_a(\mathbf{p})\varphi_b^*(\mathbf{q}')\varphi_b(\mathbf{q}) \\ \times \delta^{(4)}(p' + q' - p - q) \frac{d\mathbf{p}}{2p_0} \frac{d\mathbf{p}'}{2p_0'} \frac{d\mathbf{q}}{2q_0} \frac{d\mathbf{q}'}{2q_0'}. \end{aligned} \tag{3}$$

We use $\delta^{(4)}(p + q - p_1 - q_1)$ to integrate over the new variables $\mathbf{p}_1 + \mathbf{q}_1$ and $|\mathbf{p}_1|$ and are then left with integrations over

$\hat{p}_1 = \mathbf{p}_1/|\mathbf{p}_1|$, $\mathbf{p}$, $\mathbf{q}$, $\mathbf{p}'$, $\mathbf{q}'$:

$$\begin{aligned} W = \int G(\hat{p}_1, \mathbf{p}, \mathbf{q}, \mathbf{p}', \mathbf{q}') \, d\Omega_{\hat{p}_1} \varphi_a^*(\mathbf{p}')\varphi_a(\mathbf{p})\varphi_b^*(\mathbf{q}')\varphi_b(\mathbf{q}) \\ \times \delta(p + q - p' - q') \, d\mathbf{p} \, d\mathbf{p}' \, d\mathbf{q} \, d\mathbf{q}', \end{aligned} \tag{4}$$

where G is a regular function, since the t-matrix elements are supposed to be. Then for small enough d, we have

$$G^{-1} \, \partial G/\partial p_i \, d \ll 1. \tag{5}$$

p_i is any of the variables $\mathbf{p}\,\mathbf{q}\,\mathbf{p}'\mathbf{q}'$.

We then approximate W by putting $\mathbf{P}$, $\mathbf{Q}$ in place of $\mathbf{p}$, $\mathbf{p}'$; $\mathbf{q}$, $\mathbf{q}'$ in G. It is legitimate as one verifies that due to (1) and (1')

$$W \propto d^2(1 + O(d \, G^{-1}\partial G/\partial p_i)). \tag{6}$$

This gives the formula

$$\begin{aligned} W \approx \frac{1}{16P_0^2Q_0^2} \int \frac{E(\mathbf{p}_1\mathbf{q}_1)}{2p_{10}2q_{10}} d\mathbf{p}_1 \, d\mathbf{q}_1 \\ \times \delta(P + Q - p_1 - q_1) \, |\langle PQ| \, t \, |p_1q_1\rangle|^2 \\ \times \int \varphi_a^*(\mathbf{p}')\varphi_a(\mathbf{p})\varphi_b^*(\mathbf{q}')\varphi_b(\mathbf{q}) \\ \times \delta^4(p + q - p' - q') \, d\mathbf{p} \, d\mathbf{p}' \, d\mathbf{q} \, d\mathbf{q}'. \end{aligned} \tag{7}$$

Equation (5) is, in fact, satisfied if

$$(1/E)(\partial E/\partial p_i)d \ll 1;$$

p_i is any of the variables $\mathbf{p}_1\mathbf{q}_1$; (5a)

$m^{-1}d \ll 1$, where m is any of the masses of the particles involved; (5b)

$$\left| \frac{1}{t(pqp_1q_1)} \frac{\partial t}{\partial p_i} d \right| \ll 1, \text{ where } t(pqp_1q_1) = \langle pq| \, t \, |p_1q_1\rangle. \tag{5c}$$

Equation (5a) means that the measuring apparatus cannot discriminate between the momentum components of the wave packets. (5b) ensures that $p_0^{-1}(\partial p_0/\partial p)d \ll 1$ and means that $1/d$ is large compared to the Compton wavelength of the particles. $1/d$ will be a little later interpreted as the extension in space of the wave packet and that condition is well known from quantum mechanics.[2] The condition (5c) is known too. Let us only note that Froissart, Goldberger, and Watson[5] need it to define such physical notions as impact parameters.

[5] M. Froissart, M. L. Goldberger, and K. M. Watson, Phys. Rev. **131**, 2820 (1963).

The condition then means that the impact parameter is small compared to the extension in space of the wave packet.

B. Phenomenological Result

We now write a phenomenological formula for W. We use currents $j_\mu(x)$ to describe the motion of the particles. x is a point in time and space, $j_0(x)$ the probability density and $\mathbf{j}(x)$ the current density of the particle.

The usual formula which linked the transition probability to the number n_1 of incident particles with velocity $v_1 = P_1/P_0$, the number n_2 of scattering centers at rest and the cross section is $W_E = n_1 n_2 v_1 \sigma_E$.

In the limit where the velocity is supposed well defined, we write it, by using the probability densities $j_0^a(x)$ for the incident particle and $j_0^b(x)$ for the scattering center:

$$W_E = \int w_E(x)\, dx$$

with

$$W_E(x) = (P_1/P_0) j_0^a(x) j_0^b(x) \sigma_E. \tag{8}$$

Since $P_1/P_0 = P_1Q_0/P_0Q_0$, and P_1Q_0 is $((PQ)^2 - (m_a m_b)^2)^{\frac{1}{2}}$ in the particular frame where the second particle is at rest, we get the invariant formula

$$W_E = [(PQ)^2 - (m_a m_b)^2]^{\frac{1}{2}} \sigma_E \frac{1}{P_0 Q_0} \int j_0^a(x) j_0^b(x)\, dx, \tag{8'}$$

where it is also possible to write $j_0^a/P_0 = j_\mu^a/P_\mu$ $(\mu = 1, 2, 3)$ and $j_0^b/Q_0 = j_\nu^b/Q_\nu$.

C. Correspondence between Secs. A and B

We now take

$$j_\mu(x) = f^*(x) \overleftrightarrow{\partial}_\mu f(x), \tag{9}$$

$$f(x) = (2\pi)^{-\frac{3}{2}} \int \theta(p_0)\, \delta(p^2 - m^2) \varphi(\mathbf{p}) e^{-ipx}\, d^4p. \tag{9'}$$

This gives

$$\frac{1}{P_\mu Q_\nu} \int j_\mu^a(x) j_\nu^b(x)\, dx \approx \frac{1}{4P_0^2 Q_0^2} \frac{1}{(2\pi)^2} \int \varphi_a^*(\mathbf{p}') \varphi_a(\mathbf{p}) \varphi_b^*(\mathbf{q}') \varphi_b(\mathbf{q}) \times \delta(p + q - p' - q')\, d\mathbf{p}\, d\mathbf{p}'\, d\mathbf{q}\, d\mathbf{q}'. \tag{10}$$

Equation (8') is then identical to Eq. (7) with

$$\sigma_E = \frac{(2\pi)^2}{4((PQ)^2 - (m_a m_b)^2)^{\frac{1}{2}}} \int \frac{E(\mathbf{p}_1 \mathbf{q}_1)}{2p_{10} 2q_{10}}\, d\mathbf{p}_1\, d\mathbf{q}_1 \times |\langle PQ|\, t\, |p_1 q_1\rangle|^2\, \delta^{(4)}(P + Q - p_1 - q_1). \tag{10'}$$

Thus we make a correspondence between the S-matrix and the phenomenological formalisms by interpreting the abstract variable x in (9') as time and space and $j_\mu(x)$ defined by (9) as the current.

It is satisfying because the ordinary properties of a current are realized:

$$\partial^\mu j_\mu = 0, \qquad \int j_0(x)\, d\mathbf{x} = 1, \tag{11}$$

and for small-enough d, $j_0(x) > 0$, due to (1) and (1'),

$$1/d^3 j_0(x/d) = |\psi_d(x)|^2 + O(d^2), \tag{12}$$

where

$$\psi_d(x) = \int \Phi(\mathbf{p}) e^{ipx} \{2[p_0 d + P_0(1 - d)]\}^{-\frac{1}{2}}.$$

III. THE DOUBLE SCATTERING PROBLEM

We now study the double scattering problem. The initial state is defined by the three wavefunctions $\varphi_a(\mathbf{p})$, $\varphi_b(\mathbf{q})$, $\varphi_c(\mathbf{r})$ and we have here to consider an efficiency $E(\mathbf{p}_1 \mathbf{q}_1 \mathbf{r}_1)$. If there were no singularity in the t-matrix elements, one would obtain in the same way as in the simple scattering case (for d small enough so that the S-matrix elements are slowly varying)

$$W \approx \int \frac{E(\mathbf{p}_1 \mathbf{q}_1 \mathbf{r}_1)}{2p_{10} 2q_{10} 2r_{10}}\, d\mathbf{p}_1\, d\mathbf{q}_1\, d\mathbf{r}_1 \times |\langle PQR|\, t\, |p_1 q_1 r_1\rangle|^2 \delta(\textstyle\sum P_\lambda - \sum p_\lambda) \times \frac{(2\pi)^4}{8P_\mu Q_\nu R_\lambda} \int j_\mu^a(x) j_\nu^b(x) j_\lambda^c(x)\, dx, \tag{13}$$

with

$$\textstyle\sum P = P + Q + R, \qquad \sum p_\lambda = p_1 + q + r_1.$$

Here we suppose the kinematics as determined by the vanishing of E and the φ's, such that $(p+q-p_1)^2$ can get equal to m^2 where m is the mass of some possible "intermediate" particle. The theories then introduce a singularity in the t-matrix element proportional to the "propagator" $D(k)^2 = -1/2\pi/(k^2 - m^2 + i\epsilon)$, where $k = p + q - p_1$ and $1/2\pi$ comes from our normalizations. We suppose there is no other combination of initial and final momenta giving a singularity in the integrand.

We then make the usual hypothesis in a S-matrix theory:

$$\langle pqr|\, t\, |p_1 q_1 r_1\rangle = \mathcal{R}(pqrp_1 q_1 r_1) \times D(k^2), \tag{14}$$

where

$$\mathcal{R}(pqr, p_1q_1r_1) = \langle pq|\ t\ |p_1k\rangle\langle kr|\ t\ |q_1r_1\rangle \quad \text{for} \quad k^2 = m^2$$

and is a regular function.

We never use t-matrix elements off the mass shell.

With the correspondence made above between currents and wavefunctions, the wavefunction of the third particle translated in space through a vector ϱ is $\varphi_c(\mathbf{r})e^{-i\varrho\mathbf{r}}$.

We first study the formula for transition probability obtained through the S-matrix formalism with the usual propagator pole, then derive the phenomenological formula, and compare them for large ρ.

A. Transition Probability with the Usual Propagator

We are free to take $\mathbf{R}=0$ in the following. We have

$$W(\varrho) = \int \frac{d\mathbf{p}_1}{2p_{10}}\frac{d\mathbf{q}_1}{2p_{10}}\frac{d\mathbf{r}_1}{2r_{10}} E(\mathbf{p}_1\mathbf{q}_1\mathbf{r}_1)$$
$$\times \mathcal{R}(pqr, p_1q_1r_1)\mathcal{R}^*(p'q'r'p_1q_1r_1)D(k^2)D^*(k'^2)$$
$$\times \delta^{(4)}(\sum p - \sum p_1)\ \delta^{(4)}(\sum p - \sum p')e^{-i\varrho(\mathbf{r}-\mathbf{r}')}$$
$$\times \frac{\varphi_a^*(\mathbf{p}')}{2p_0'}\frac{\varphi_a(\mathbf{p})}{2p_0}\frac{\varphi_b^*(\mathbf{q}')}{2q_0'}\frac{\varphi_b(\mathbf{q})}{2q_0}$$
$$\times \frac{\varphi_c^*(\mathbf{r}')}{2r_0'}\frac{\varphi_c(\mathbf{r})}{2r_0}\, d\mathbf{p}\, d\mathbf{p}'\, d\mathbf{q}\, d\mathbf{q}'\, d\mathbf{r}\, d\mathbf{r}'. \tag{14}$$

Then one has (and this is one of the main points to achieve our proof!):

$$4\pi^2 D(k^2)D^*(k'^2) = \frac{1}{k^2 - m^2 + i\epsilon} \times \frac{1}{k'^2 - m^2 - i\epsilon}$$
$$= \frac{1/2}{k'^2 - k^2 - i\epsilon}\left[\frac{1}{k^2 - m^2 + i\epsilon} - \frac{1}{k^2 - m^2 - i\epsilon} + \frac{1}{k'^2 - m^2 + i\epsilon} - \frac{1}{k^2 - m^2 - i\epsilon}\right] + \frac{1/2}{k'^2 - k^2 - i\epsilon}$$
$$\times \left[\frac{1}{k^2 - m^2 - i\epsilon} + \frac{1}{k^2 - m^2 + i\epsilon} - \frac{1}{k'^2 - m^2 - i\epsilon} - \frac{1}{k'^2 - m^2 + i\epsilon}\right], \tag{15}$$

$$4\pi^2 D(k^2)D^*(k'^2)$$
$$= \frac{i\pi}{k'^2 - k^2 - i\epsilon}\left[\delta(k^2 - m^2) + \delta(k'^2 - m^2)\right] + \frac{1}{k'^2 - k^2 - i\epsilon}\left[\frac{P}{k^2 - m^2} - \frac{P}{k'^2 - m^2}\right], \tag{15'}$$

with

$$k'^2 - k^2 - i\epsilon = -(2\beta\mathbf{K}_1 + i\epsilon) + O_2(\alpha, \beta, \mathbf{p} - \mathbf{P}, \cdots), \tag{16}$$

where

$$\mathbf{K}_1 = \mathbf{P} + \mathbf{Q} - \mathbf{p}_1, \qquad \beta = \mathbf{r} - \mathbf{r}', \qquad \alpha = \mathbf{r} + \mathbf{r}',$$

and O_2 is a second-order quantity.

For the sake of simplicity, we now suppose the kinematics such that the equation

$$(p + q - p_1)^2 - m^2 = 0$$

has at most one solution for every fixed direction of $\mathbf{k} = \mathbf{p} + \mathbf{q} - \mathbf{p}_1$, and that the derivative of $(p + q - p_1)^2$ relative to $|\mathbf{k}|$ does not vanish. In Eq. (14), we use $\delta^{(4)}(\sum p - \sum p_1)$ to integrate over the new variables $\sum \mathbf{p}_1$ and $|\mathbf{q}_1|$ and then use the new variables $\hat{k}_1 = \mathbf{k}_1/|\mathbf{k}_1|$ and $s = (p + q - q_1)^2 - m^2$.

Due to (15') W separates into three parts, W_{I}, W_{II}, W_{III};

$$W_i = \int G(\hat{q}_1, \hat{K}_1, s, \mathbf{p}, \mathbf{q}, \mathbf{p}', \mathbf{q}')\frac{H_i}{k^2 - k'^2 - i\epsilon}$$
$$\times e^{-i\varrho\beta}\ \delta^{(4)}(\sum p' - \sum p)\varphi_a^*(\mathbf{p}')\varphi_a(\mathbf{p})\varphi_b^*(\mathbf{q}')\varphi_b(\mathbf{q})$$
$$\times \varphi_c^*(\mathbf{r}')\varphi_c(\mathbf{r})\, d\mathbf{p}\, d\mathbf{p}'\, d\mathbf{q}\, d\mathbf{q}'\, d\mathbf{r}\, d\mathbf{r}'\, d\Omega_{\hat{q}_1}\, d\Omega_{\hat{K}_1}\, ds \tag{15''}$$

with

$$H_1 = i\pi[\delta(k^2 - m^2) + \delta(k'^2 - m^2)],$$
$$H_2 = P/(k^2 - m^2),$$
$$H_3 = -P/(k'^2 - m^2).$$

The function G is regular and satisfies conditions (5a, b, c) for small-enough d, as the $\mathcal{R}$-matrix elements are supposed regular.

We now define

$$\varphi(\beta) = \int \varphi_a^*(\mathbf{p}')\varphi_a(\mathbf{p})\varphi_b^*(\mathbf{q}')\varphi_b(\mathbf{q})\varphi_c^*[\tfrac{1}{2}(\alpha - \beta)]$$
$$\times \varphi_c[\tfrac{1}{2}(\alpha + \beta)]\delta(\sum p' - \sum p)\, d\mathbf{p}\, d\mathbf{p}'\, d\mathbf{q}\, d\mathbf{q}'\, d\alpha. \tag{17}$$

For small enough d and with the hypothesis that the velocities of the first two particles are different, i.e., $|\mathbf{P}/P_0 - \mathbf{Q}/Q_0|$ is supposed large compared to d, $\varphi(\beta)$ is infinitely differentiable. In fact,

$$\delta^4(\sum p' - \sum p) = \delta^3(\sum \mathbf{p}' - \sum \mathbf{p}) \times \delta[(\mathbf{P}/P_0 - \mathbf{Q}/Q_0)(\mathbf{p} - \mathbf{p}') + O(\mathbf{p} + \mathbf{p}' - 2\mathbf{p}, \mathbf{p} - \mathbf{p}', \alpha, \beta)].$$

We verify that the derivative of the function inside the last δ-function relative to the variable $(\mathbf{P}/P_0 -$

$\mathbf{Q}/Q_0)(\mathbf{p} - \mathbf{p}')$ never vanishes and is almost unity if $\mathbf{p}$ and $\mathbf{p}'$ lie on support of $\varphi_a(\mathbf{p})$. We then integrate with the new variables $\sum \mathbf{p}'$, $\sum \mathbf{p}$, $(\mathbf{P}/P_0 - \mathbf{Q}/Q_0) \times (\mathbf{p} - \mathbf{p}')$, $\cdots$ and $\varphi(\beta)$ is infinitely differentiable as are all the functions involved. Furthermore, $\varphi(\beta)$ vanishes outside a sphere with center at $\mathbf{0}$ and diameter $2d$, owing to the vanishing of φ_c.

With the same kind of approximation as in the simple scattering case, one then gets by putting $P(Q)$ in place of $pp'(qq')$ in G, $\delta(k^2 - m^2)$, $\delta(k'^2 - m^2)$:

$$\frac{dW_I}{d\Omega_{q_1}} \approx \int F_1(\hat{K}_1)\, d\Omega_{\hat{K}_1} \frac{1}{\beta \mathbf{K}_1 + i\epsilon} \varphi(\beta) e^{-i\varrho\beta}\, d\beta, \quad (18)$$

where we do not write explicitly the dependence of F_1 upon $\hat{q}_1$.

We now write

$$\frac{-1}{\beta \mathbf{K}_1 + i\epsilon} = \frac{1}{|\mathbf{K}_1|} \int \theta(\hat{K}_1 \cdot \varrho') \times \delta^{(2)}(\mathbf{K}_1 \times \varrho') e^{i\varrho'\beta}\, d\varrho', \quad (19)$$

$$\int F_1(\hat{K}_1)\, d\Omega_{\hat{K}_1} \theta(\hat{K}_1 \cdot \varrho')\, \delta^{(2)}(\hat{K}_1 \times \varrho') = \frac{F_1(\hat{\rho}')}{\varrho'^2}. \quad (20)$$

We have defined $\delta^{(2)}(\hat{K}_1 \times \varrho)$ as follows. Take a frame where $\hat{K}_1$ is the unit vector of the first axis:

$$\delta^{(2)}(\hat{K}_1 \times \varrho) = \delta(\rho_2)\delta(\rho_3).$$

We then get

$$\frac{dW_1}{d\Omega_{\hat{q}_1}} = \frac{1}{|\mathbf{K}_1|} \int F_1(\hat{\rho}') \frac{d\varrho'}{\rho'^2} \int e^{i(\varrho' - \varrho)\beta} \varphi(\beta)\, d\beta = \frac{1}{|\mathbf{K}_1|} \int F_1(\hat{\rho}') \frac{d\varrho'}{\rho'^2} g(\varrho' - \varrho)\, d\varrho' \quad (21)$$

with

$$g(\varrho) = \int e^{i\varrho\beta} \varphi(\beta)\, d\beta. \quad (22)$$

As φ is infinitely differentiable, g is rapidly decreasing with ρ and one has

$$|g(\varrho)| < 1/|\rho|^n \quad \text{for} \quad |\rho| > b_n \quad \text{and } n \text{ is arbitrary.}$$

We take ρ large enough so that $\rho^{\frac{1}{2}} > b_n$ and write

$$\int \frac{F_1(\hat{\rho}')}{\rho'^2} g(\varrho' - \varrho)\, d\varrho' = \int_{|\varrho' - \varrho| < \rho^{\frac{1}{2}}} + \int_{|\varrho' - \varrho| > \rho^{\frac{1}{2}}}.$$

As ρ goes to $+\infty$, the second term goes to zero as $1/|\rho|^{n-1}$. As $F(\hat{\rho}')/\rho'^2$ is then nearly constant, the first one gives in the lowest order in $1/\rho$ (n is taken large enough)

$$dW_1/d\Omega_{\hat{q}_1} \sim F(\hat{\rho})\varphi(0)/\rho^2, \quad (23)$$

where

$$F(\hat{\rho}) = \left[\frac{(2\pi)^3}{K_1} \frac{1}{4\pi} \frac{1}{64 P_0^2 Q_0^2 R_0^2} \times \int \frac{E(\mathbf{p}_1\mathbf{q}_1\mathbf{r}_1)}{2p_{10}2q_{10}2r_{10}} |\Re(PQRp_1q_1r_1)|^2 \times \delta^4(\sum P - \sum p_1)\, \delta((P + Q - p_1)^2 - m^2) \times d(\sum \mathbf{p}_1)\mathbf{q}_1^2\, d\,|\mathbf{q}_1|\,|\mathbf{K}_1|^2\, d\,|\mathbf{K}_1|\right]_{\hat{K}_1 = \hat{\rho}} \quad (24)$$

and

$$|\Re(PQRp_1q_1r_1)|^2 = |\langle PQ|\ t\ |p_1\rangle|^2\ |\langle R|\ t\ |q_1r_1\rangle|^2$$

for $(P + Q - g_1) = m^2$. These are the only elements to consider due to $\delta(P + Q - p_1)^2 - m^2)$.

We now consider W_{II} and W_{III} of Eq. (15''), and use the variables $t = s$ for W_{II}, $t = s + k'^2 - k^2$ for W_{III}.

For small $k'^2 - k^2$: $G(t - k'^2 + k^2) \sim G(t) - (k'^2 - k^2)\partial G/\partial t$. This gives

$$W_{II} + W_{III} \approx \int \left[G(t) - \left(G(t) - \frac{\partial G}{\partial t}(k'^2 - k^2)\right)\right] \times \frac{dt}{k'^2 - k^2 - i\epsilon} \frac{P}{t} \varphi(\beta) e^{-i\varrho\beta}\, d\beta \approx \int \frac{\partial G}{\partial t} \frac{P}{t}\, dt \varphi(\beta) e^{-i\varrho\beta}\, d\beta.$$

One gets a well-defined quantity for

$$\int \frac{\partial G}{\partial t} \frac{P}{t}\, dt,$$

and $W_{II} + W_{III}$ decreases rapidly with ρ. We then neglect it, and have the result

$$dW/d\Omega_{\hat{q}_1} \sim F(\hat{\rho})[\varphi(0)/\rho^2], \quad (25)$$

where $F(\hat{\rho})$ is given by formula (24).

B. Phenomenological Formula for Two Successive Interactions

We consider now we are dealing with two successive scatterings and an intermediate real particle of mass m and momentum $\mathbf{K}$. We first write its current for a given $\mathbf{p}_1$. We call $W_1(W_2)$ and $\sigma_1(\sigma_2)$ the transition probability and cross section of the first (second) interaction.

As in the calculation of the simple scattering transition probability, we calculate the density at point y for an interaction at point x and integrate over x. The density at point $y(y_0, y)$ depends on the transition probability in the direction $(y - x)\hat{}$, the speed of the intermediate particle and the de-

crease in $(\mathbf{x}-\mathbf{y})^{-2}$ of the current ($\int jr^2\,d\Omega = \text{const}$). This gives

$$\frac{dj_0^{\text{int}}(y)}{d\mathbf{p}_1} = [(P,Q)^2 - (m_a m_b)^2]^{\frac{1}{2}} \times \int dx\, \frac{j_\mu^a(x) j_\nu^b(x)}{P_\mu Q_\nu} \frac{d}{d\mathbf{p}_1}\left[\frac{d\sigma_1}{d\Omega_K}\right]_{\hat{K}=(y-x)\hat{}} \frac{1}{(\mathbf{x}-\mathbf{y})^2} \times \delta[|\mathbf{x}-\mathbf{y}| - (|\mathbf{K}|/K_0)\,|y_0 - x_0|]\theta(y_0 - x_0), \quad (26)$$

where $\theta(y_0 - x_0)$ means the intermediate particle only exists for $y_0 > x_0$.

One easily verifies that

$$\int \frac{dj_0^{\text{int}}(y)}{d\mathbf{p}_1}\,d\mathbf{y} = \frac{dW_1}{d\mathbf{p}_1}$$

by taking the new variable $\mathbf{u} = \mathbf{y} - \mathbf{x}$. Equation (10′) gives σ_1 and we may write

$$\frac{dj_0^{\text{int}}(y)}{d\mathbf{p}_1} = \frac{(2\pi)^2}{4P_\mu Q_\nu} \int j_\mu^a(x) j_\nu^b(x)\,dx\, \frac{1}{(\mathbf{x}-\mathbf{y})^2} \times \delta\left(|\mathbf{x}-\mathbf{y}| - \frac{|\mathbf{K}|}{K_0}(y_0 - x_0)\right) \times \left[\int \frac{E(\mathbf{p}_1\mathbf{K})}{2p_{10}2K_0}\,|\mathbf{K}|^2\,d\,|\mathbf{K}| \times |\langle PQ|\,t\,|p_1K\rangle|^2\,\delta(P+Q-p_1-K)\right]_{\hat{K}=(y-x)\hat{}} \quad (26')$$

We now use Eqs. (7) and (8′) once more for the second interaction

$$\frac{dW(\rho)}{d\mathbf{p}_1} = \frac{|\mathbf{K}|}{K_0}\,\sigma_2 \int \frac{dj_0^{\text{int}}(y')}{d\mathbf{p}_1}\, j_0^c(y_0', \mathbf{y}' - \boldsymbol{\rho})\,dy'. \quad (27)$$

By using the new variable $y(y_0', \mathbf{y}' - \boldsymbol{\rho})$ and Eq. (10′), this gives

$$\frac{dW(\boldsymbol{\rho})}{d\mathbf{p}_1} = \frac{(2\pi)^4}{16P_\mu Q_\nu R_0 K_0} \int j_\mu^a(x) j_\nu^b(x) \times \frac{1}{(-\mathbf{x}+\mathbf{y}+\boldsymbol{\rho})^2}\, j_0^c(y)\,\delta(|-\mathbf{x}+\mathbf{y}+\boldsymbol{\rho}| - (|\mathbf{K}|/K_0)(y_0 - x_0))\,dx\,dy \left[\int \frac{E(\mathbf{p}_1\mathbf{K})}{2p_{10}2K_0}\,|\mathbf{K}|^2\,d\,|\mathbf{K}| \times |\langle PQ|\,t\,|p_1K\rangle|^2\,\delta(P+Q-p_1-K) \times [E(\mathbf{q}_1\mathbf{r}_1)/2q_{10}2r_{10}]\,\delta(K+R-q_1-r_1) \times |\langle KR|\,t\,|q_1r_1\rangle|^2\,d\mathbf{q}_1\,d\mathbf{r}_1\right]_{\hat{K}=(y-x)\hat{}+\hat{\rho}}. \quad (27')$$

In fact, we do not measure anything about $\mathbf{K}$ and some correlations may exist between the measurements of $\mathbf{p}_1\mathbf{q}_1\mathbf{r}_1$. We then take $E(\mathbf{p}_1\mathbf{q}_1\mathbf{r}_1)$ instead of $E(\mathbf{p}_1\mathbf{K})E(\mathbf{q}_1\mathbf{r}_1)$ and obtain $W(\boldsymbol{\rho})$ by integrating over $\mathbf{p}_1$.

We now state the following result whose proof is fairly long, but not too difficult:

$$|j_0(\mathbf{x} - (\mathbf{P}/P_0)x_0, x_0)| < 1/|\mathbf{x} - \mathbf{v}x_0|^r \quad \text{for} \quad |\mathbf{x} - \mathbf{v}x_0| > A_r + x_0 d,$$

for any value of r, with $\mathbf{v} = \mathbf{P}/P_0$ and

$$|j_0(\mathbf{x}, 0)| < 1/|\mathbf{x}|^r \quad \text{for} \quad |\mathbf{x}| > A_r$$

[$j_0(\mathbf{x}, 0)$ being rapidly decreasing because the wave-functions φ are infinitely differentiable].

It is clear that $j_0^a(x) j_0^b(x)$ is rapidly decreasing with $|\mathbf{x}|$ and x_0, for $|v_a - v_b| > 2d$. By using the same kind of approximation as for formulas (21) and (23), we get

$$\frac{dW(\boldsymbol{\rho})}{d\Omega_{\hat{q}_1}} = \frac{(2\pi)^4}{16P_\mu Q_\nu R_0 K_0}\,\frac{1}{\rho^2}\left[\int \frac{E(\mathbf{p}_1\mathbf{q}_1\mathbf{r}_1)}{2p_{10}2q_{10}2r_{10}} \times |\mathbf{K}|^2\,d\,|\mathbf{K}|\,|\mathbf{q}_1|^2\,d\,|\mathbf{q}_1|\,d(\textstyle\sum \mathbf{p}_1) \times \delta(\sum P - \sum p_1)\,\delta((P+Q-p_1)^2 - m^2) \times |\langle PQ|\,t\,|p_1K\rangle|^2\,|\langle KR|\,t\,|q_1r_1\rangle|^2\right]_{\hat{K}=\hat{\rho}} \times \int j_\mu^a(x) j_\nu^b(x) j_0^c(y)\,dx\,dy \times \delta(|\mathbf{x}-\mathbf{y}| - (|\mathbf{K}|/K_0)(y_0 - x_0)). \quad (28)$$

We have also used

$$d\mathbf{K}/2K_0 = \delta(K^2 - m^2)\,d^4K, \qquad \int \delta(P+Q-p_1-K)\,d^4K = 1,$$

and

$$d\mathbf{p}_1 = -d\mathbf{K} \quad \text{with} \quad \mathbf{K} = \mathbf{P} + \mathbf{Q} - \mathbf{p}_1.$$

Then, one has

$$\int j_\mu^a(x) j_\nu^b(x)\,dx\,dy\,\delta(|x-y| - (K/K_0)(y_0 - x_0)) \times j_0^c(y) \approx 8P_\mu Q_\nu R_0 \frac{1}{(2\pi)^9} \times \int \frac{\varphi_a^*(\mathbf{p}')}{2p_0'}\frac{\varphi_a(\mathbf{p})}{2p_0}\frac{\varphi_b^*(\mathbf{q}')}{2q_0'}\frac{\varphi_b(\mathbf{q})}{2q_0}\frac{\varphi_c^*(\mathbf{r}')}{2r_0'}\frac{\varphi_c(\mathbf{r})}{2r_0} \times d\mathbf{p}\,d\mathbf{p}'\,d\mathbf{q}\,d\mathbf{q}'\,d\mathbf{r}\,d\mathbf{r}'\,e^{-i(p+q-p'-q')x} \times e^{i(r-r')y}\,dx\,dy\,\delta\left(|\mathbf{x}-\mathbf{y}| - \frac{K}{K_0}(y_0 - x_0)\right). \quad (29)$$

We now write

$$\int e^{i(p+q-p'-q')x} e^{i(r-r')y}\,dx\,dy \times \delta(|\mathbf{x}-\mathbf{y}| - (K/K_0)(y_0 - x_0)) = \int e^{i(p+q+r-p'-q'-r')x}\,dx \times \int e^{i(r-r')u}\,\delta\left(|\mathbf{u}| - \frac{|\mathbf{K}|}{K_0}u_0\right)du = (2\pi)^4\,\delta(\sum p' - \sum p) \times \int e^{i\lambda|u|}e^{-i(r-r')u}\,du\,\frac{K_0}{|\mathbf{K}|}, \qquad (29')$$

where $\lambda = (|\mathbf{K}|/K_0)(r_0 - r_0')$ is a second-order quantity. We find

$$\int e^{i\lambda|u|}e^{-i\beta u}\,du \approx \delta(\beta)(2\pi)^3$$

at the limit of small support of test functions $\varphi(\beta)$.

One then gets using (28) and (29)

$$dW/d\Omega_{\hat{q}_1} \sim F(\hat{\rho})[\varphi(0)/\rho^2], \qquad (30)$$

which is exactly formula (25).

C. Choice between Other Possible Propagators

Unitarity gives $D(k^2) - D^*(k^2) = i\delta(k^2 - m^2)$, i.e.

$$D(k^2) = -(1/2\pi)(k^2 - m^2 + i\epsilon)^{-1} + R(k^2), \qquad (31)$$

where R is some real function. In particular, one may make Olive's choice[3]:

$$-2\pi D(k^2) = \frac{c}{k^2 - m^2 + i\epsilon} - \frac{1-c}{k^2 - m^2 - i\epsilon} = \frac{1}{k^2 - m^2 + i\epsilon} + 2(c-1)\frac{P}{k^2 - m^2}, \qquad (32)$$

where c is an arbitrary constant, although the reasons for restricting R in this way are not very clear.

We now show that we need $c = 1$ in (32) to obtain the phenomenological result. Let us note first that if we take $c = 0$, i.e.,

$$D = \frac{-1}{2\pi}\frac{1}{k^2 - m^2 - i\epsilon},$$

the calculation is similar but we have

$$\frac{-1}{\beta\mathbf{K}_1 - i\epsilon} = \frac{1}{K_1}\int \theta(-\hat{K}_1\varrho')\,\delta(\hat{K}_1 \times \varrho')e^{i\varrho'\beta}\,d\varrho'$$

is place of formula (19). The end of the calculation is similar, but instead of the expected $F(\hat{\rho})$, we get $F(-\hat{\rho})$, i.e., the value of F for the direction opposite to the phenomenological one.

We now use formula (32) and have

$$4\pi^2 D(k^2)D^*(k'^2) = \frac{c^2}{(k^2 - m^2 + i\epsilon)(k'^2 - m^2 - i\epsilon)} + \frac{(1-c)^2}{(k^2 - m^2 - i\epsilon)(k'^2 - m^2 + i\epsilon)} - c(1-c)\left[\frac{1}{k^2 - m^2 + i\epsilon}\frac{1}{k'^2 - m^2 + i\epsilon} + \frac{1}{k^2 - m^2 - i\epsilon}\frac{1}{k'^2 - m^2 - i\epsilon}\right].$$

The first term gives $c^2F(\hat{\rho})[\varphi(0)/\rho^2]$. The second one $(1-c)^2F(-\hat{\rho})\varphi(0)/\rho^2$. For the third one, we write

$$c(1-c)\left[\frac{1}{k^2 - m^2 + i\epsilon}\frac{1}{k'^2 - m^2 + i\epsilon} + \frac{1}{k^2 - m^2 - i\epsilon}\frac{1}{k'^2 - m^2 - i\epsilon}\right] = c(1-c)\left[2\pi^2\,\delta(k^2 - m^2)\,\delta(k'^2 - m^2) + \frac{2P}{k^2 - m^2}\frac{P}{k'^2 - m^2}\right].$$

Since

$$\frac{P}{k^2 - m^2}\frac{P}{k'^2 - m^2} + \pi^2\,\delta(k^2 - m^2)\,\delta(k'^2 - m^2) = \frac{P}{k'^2 - k^2}\left[\frac{P}{k^2 - m^2} - \frac{P}{k'^2 - m^2}\right],$$

the third term is to be neglected for the same reason as $W_{II} + W_{III}$ in the end of Sec. IIID.

In conclusion, one does not find the phenomenological result but rather

$$[\varphi(0)/\rho^2][c^2F(\hat{\rho}) + (1-c)^2F(-\hat{\rho})].$$

IV. CONCLUSION

We have thus shown that the formula obtained in the S-matrix formalism gives the same result as the phenomenological one.

This was shown using the usual correspondence between wavefunctions and currents and under the following assumptions: the distance between the two targets is large compared to the extension in space of the wave packets of the particles, which in turn is large compared to the Compton wavelength and to the impact parameters. Furthermore the measurement apparatus cannot discriminate between the momentum components of the wave packets.

Let us now point out that the only thing we have done was to obtain consistent results. Maybe one could try to find another propagator, another correspondence between wavefunctions and currents,

other "reasonable conditions" and show it is possible to describe a double scattering in terms of two successive scatterings. It would then also yield consistent results. Although we have no absolute reason to say it is hopeless, it does not seem very easy. The consistency of our results may be expressed in various ways. For instance one may say it proves the usual propagator is a good one. But one might also admit the existence of the usual pole and say it proves that the correspondence between wave-functions and currents is good. This could be interpreted as a definition of space–time and currents in S-matrix theory, in an asymptotic sense.

However, we only studied double scattering. On the one hand, to say that the usual propagator is good would need to generalize our results to any multiple scattering and to any number of initial and final particles. The generalization to any multiple scattering does not seem difficult. The generalization for more than three final particles of the double scattering case is easy. Another point is to see what are the restrictions on the function $R(k^2)$ of formula (31).

On the other hand, to say that the definition of space–time and currents is good would need not only the generalizations we have quoted above, but also the study of all possible forms of interactions (including for instance the propagation of unstable particles and such things). One should make a correspondence between properties (singularities for instance) of the S-matrix and corresponding phenomenological processes, including all possible processes.

It might then yield, and this is the more ambitious program one could try to achieve, the possibility of a measurement theory in the framework of S-matrix theory, a measurement being merely a special kind of process taking place in space–time.

ACKNOWLEDGMENTS

I chiefly want to thank Marcel Froissart, who has directed this work and has considerably helped me in many discussions. I also thank Professor Messiah for his kind hospitality at the Service de Physique Théorique and for study help and encouragement.

Racah Algebra for an Arbitrary Group*

JEAN-ROBERT DEROME† AND W. T. SHARP
Department of Mathematics, University of Toronto, Toronto, Ontario, Canada
(Received 19 May 1964)

$3j$- and $6j$-symbols are studied for a general group without assuming that the group is ambivalent or multiplicity-free. The choice of multiplicity label r that distinguishes the equivalent irreducible representations that may arise in a Kronecker product of irreducible representations is left arbitrary and no special choices of phase are made. It is found that the $3j$- and $6j$-symbols obtained have essentially the same properties as the familiar $3j$- and $6j$-symbols for the rotation group in three dimensions.

1. INTRODUCTION

IN this paper the main features of Racah algebra (often called the quantum theory of angular momentum) are developed for any finite or compact group. It can be shown[1] that there is no essential difference between the Racah algebra for R_3, the rotation group in three dimensions,[2,3] and that for any other simply reducible (SR) group. A group G is said to be SR if it is both ambivalent and multiplicity free (mf).[4] The statement that G is ambivalent means that every element of the group is in the same class as its inverse. The group G is mf if in the reduction of the Kronecker product of any two irreducible representations of G no irreducible constituent appears more than once. Sharp[5] has relaxed the requirement that G be ambivalent, thus obtaining Racah algebra for "quasi SR" groups.

Hamermesh[6] has shown how to obtain the properties of the Clebsch–Gordan coefficients for S_n, the symmetric group on n objects. These groups are not mf for $n > 4$, but they are ambivalent.

In this paper we remove completely the assumption that G is ambivalent and multiplicity free. This generalization is interesting physically as well as mathematically: for example the group $SU(3)$ which is neither ambivalent nor mf has been used in nuclear structure[7] as well as in the classification of elementary particles.[8]

The $3j$-symbols are defined, in the usual way, in terms of the elements of the unitary matrix which reduces the Kronecker product of two irreducible representations of the group G into its irreducible constituents. However the set of $3j$-symbols thus defined is not unique so that there remains a certain freedom in the choice of $3j$-symbols. This freedom can be exploited to impose some simple symmetry relations among the various $3j$-symbols.

We define the $1j$-symbol as a special case of a $3j$-symbol and the $6j$-symbol as a sum of products of $3j$-symbols. These generalized $6j$-symbols have the familiar properties of symmetry and unitarity, and satisfy a Biedenharn identity.[3,9] The added complexity of these generalized relations is twofold: there are summations over multiplicity indices and the "phases" are more complicated.

We label the equivalence classes of the irreducible representations of G by the letter j with appropriate subscripts. By $j_1 = j_2$ we shall mean that j_1 is equivalent to j_2. We denote the representation matrices by $j(R)^m{}_{m'}$, where $R \in G$ and $[j]$ will stand for the dimension of the representation space of j. The matrices $j(R)^m{}_{m'}$ depend on the choice of basis in representation space. For instance $j_1 = j_2$ does not imply that $j_1(R)^{m_1}{}_{m_1'} = j_2(R)^{m_2}{}_{m_2'}$, but only that these two sets of matrices are equivalent. If G is not ambivalent, G has irreducible representations that are neither integer nor half-integer; such j will be called complex. The representations will be taken in unitary form and we will write $\{j(R)^m{}_{m'}\}^* = j(R)_m{}^{m'}$. We denote by j^* the representation which is complex conjugate to j. When $j \neq j^*$, i.e., when j

* A less general version of this work was submitted by one of us (J.-R. D.) in October 1962 in partial fulfillment of the requirements for the M. Sc. degree in physics at the University of Alberta.

† Holder of a Studentship from the National Research Council of Canada.

[1] E. P. Wigner, "On the Matrices Which Reduce the Kronecker Product of Representations of S. R. Groups." Privately circulated manuscript, Princeton, 1951; see also E. P. Wigner, *Group Theory* (Academic Press Inc., New York, 1959).

[2] G. Racah, Phys. Rev. **62**, 438 (1942).

[3] U. Fano and G. Racah, *Irreducible Tensorial Sets* (Academic Press Inc., New York, 1959).

[4] E. P. Wigner, Am. J. Math. **63**, 57 (1941).

[5] W. T. Sharp, "Racah Algebra and the Contraction of Groups." Report AECL-1098 (Atomic Energy of Canada Ltd., Chalk River, Ontario, 1960).

[6] M. Hamermesh, *Group Theory* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1962).

[7] J. P. Elliot, Proc. Roy. Soc. (London) **A245**, 128 (1958); M. Kretzschmar, Z. Physik **157**, 433 (1960); V. Bargmann and M. Moshinsky, Nucl. Phys. **18**, 697 (1960).

[8] M. Gell-Mann, Phys. Rev. **125**, 1067 (1962); Y. Ne'eman, Nucl. Phys. **26**, 222 (1961).

[9] L. C. Biedenharn, J. Math. and Phys. **31**, 287 (1953).

is complex, we choose the basis in the representation space of j^* such that

$$j^*(R)^m{}_{m'} = \{j(R)^m{}_{m'}\}^* = j(R)_m{}^{m'}. \tag{1.1}$$

2. THE 3j-SYMBOL

We define the $3j$-symbol $(j_1j_2j_3)_{r,m_1m_2m_3}$ by the equation

$$j_1(R)^{m_1}{}_{m_1'} j_2(R)^{m_2}{}_{m_2'} = \sum_{j_3} [j_3]\{(j_1j_2j_3)_{r,m_1m_2m_3}\}^* \times j_3(R)_{m_3}{}^{m_3'}(j_1j_2j_3)_{r,m_1'm_2'm_3'}, \tag{2.1}$$

where a sum is implied over repeated m and r indices (once covariant and once contravariant). If a covariant m or r index appears in a complex conjugated symbol it is to be considered for purposes of summation convention as contravariant and vice versa. The index r is a multiplicity index and it takes on as many values (e.g., 1, 2, 3, $\cdots$) as the multiplicity of j_3^* in the Kronecker product $j_1 \times j_2$. We say that j_1, j_2, and j_3 form a triad if j_3^* is contained at least once in the product $j_1 \times j_2$. The $3j$-symbols thus form a square matrix having as row indices (m_1m_2) and as column indices (j_3m_3r). It is sometimes useful to think of the $3j$-symbols with j's and m's fixed as the components of a vector. The multiplicity index serves as a label for these components. It follows from Eq. (2.1) that the $3j$-symbols satisfy the unitarity conditions

$$[j_3]\{(j_1j_2j_3)_{r,m_1m_2m_3}\}^*(j_1j_2j_3')_{r',m_1m_2m_3'} = \delta_{j_3j_3'}\delta^{m_3}{}_{m_3'}\delta^r{}_{r'} \tag{2.2}$$

and

$$\sum_{j_3} [j_3]\{(j_1j_2j_3)_{r,m_1m_2m_3}\}^*(j_1j_2j_3)_{r,m_1'm_2'm_3} = \delta^{m_1}{}_{m_1'}\delta^{m_2}{}_{m_2'}, \tag{2.3}$$

where the indicated sum is over all irreducible representations of G since by convention the $3j$-symbol is zero whenever $j_1j_2j_3$ do not form a triad.

Clearly the phases of the $3j$-symbols are not determined by the defining equation. More generally let

$$\{(j_1j_2j_3)_{r,m_1m_2m_3}\}' = U(123)_r{}^{r'}(j_1j_2j_3)_{r',m_1m_2m_3}, \tag{2.4}$$

where $U(123)$ is a unitary matrix which may be a function of j_1, j_2, and j_3 but is independent of m_1, m_2, and m_3. The transformed set of $3j$-symbols satisfies the same defining relation as the original set. This matrix U may also depend on the order in which j_1, j_2, and j_3 appear in the $3j$-symbol. For instance $U(123)$ and $U(132)$ may be different. Thus the $3j$-symbols are defined only up to a unitary transformation and this freedom can be exploited to impose "simple" symmetry properties on the $3j$-symbols. By an appropriate choice of the matrices U one might hope for example to make the $3j$-symbol symmetric or antisymmetric under permutations of the j's and of the corresponding m's. However, the most convenient choice of the matrices U (i.e., of an r labeling scheme) may depend both on the group and the specific application envisaged. Hence in this paper we shall not make a definite choice of these matrices U; the $3j$-symbols are required only to satisfy Eq. (2.1) and the theorems proved are independent of special choices of basis. They apply to any set of $3j$-symbols (consistent with the defining relation) of any compact or finite group.

We now wish to see what happens to the $3j$-symbols when the j's and m's are permuted. Using the unitarity of the $3j$-symbols and of the representation matrices in Eq. (2.1) one obtains

$$\frac{1}{A}\int j_1(R)^{m_1}{}_{m_1'} j_2(R)^{m_2}{}_{m_2'} j_3(R)^{m_3}{}_{m_3'}\, dR = \{(j_1j_2j_3)_{r,m_1m_2m_3}\}^*(j_1j_2j_3)_{r,m_1'm_2'm_3'}, \tag{2.5}$$

where $A = \int dR$ and the integration or sum is to be performed over the group G. It follows that

$$\sum_r |(j_1j_2j_3)_{r,m_1m_2m_3}|^2 = \sum_r |(j_1j_3j_2)_{r,m_1m_3m_2}|^2 = \sum_r |(j_2j_3j_1)_{r,m_2m_3m_1}|^2. \tag{2.6}$$

Hence there is a unitary matrix M such that

$$(j_2j_1j_3)_{r,m_2m_1m_3} = M(12,3)_r{}^{r'}(j_1j_2j_3)_{r',m_1m_2m_3}, \tag{2.7}$$

and using the unitarity of the $3j$-symbols, one obtains

$$M(12,3)_r{}^{r'} = (j_2j_1j_3)_{r,m_2m_1m_3}\{(j_1j_2j_3)_{r',m_1m_2m_3}\}^*. \tag{2.8}$$

This matrix $M(12, 3)$ does not depend on m_1, m_2, or m_3 as can be seen from Eq. (2.5). However it may depend on the ordered set $j_1j_2j_3$ and also on what transposition is being considered. The comma in $M(12, 3)$ indicates that j_1 and j_2 are to be permuted. The argument (1, 23) would mean that j_2 and j_3 are to be permuted and finally when no comma appears, i.e., (123) we mean that j_1 and j_3 are being permuted. Clearly any cyclic permutation of the j's in a $3j$-symbol can be obtained with a product of two matrices M and we need consider only transpositions. Given a set of $3j$-symbols one can obtain a set of matrices M which express a certain symmetry of the $3j$-symbols under permutations of the j's and m's. As an example consider a simply reducible group. The matrices M are just

complex numbers of modulus one and the usual choice of phases consists in taking

$$M(123) = M(1, 23) = M(12, 3) = (-1)^{j_1+j_2+j_3},$$

where $(-1)^j$ is a certain phase factor associated with j.[1,5] It does not seem possible to define $(-1)^j$ with the desired properties in the general case.

3. THE 1j-SYMBOL

The 1j-symbol $(j)_{mm'}$ and its complex conjugate $(j)^{mm'}$ are defined in the usual way:

$$(j)_{mm'} = [j]^{\frac{1}{2}}(j0j^*)_{1,m0m'} = \{(j)^{mm'}\}^*. \tag{3.1}$$

The representation 0 is the identity representation which is of course one-dimensional so that the corresponding m takes on exactly one value denoted by $_0$. From the properties of the 3j-symbols it follows that

$$j(R)^{m_1}{}_{m_1'} = (j)^{m_1m_2}j^*(R)_{m_2}{}^{m_2'}(j)_{m_1'm_2'}, \tag{3.2}$$

$$(j)^{m_1m_2}(j)_{m_1'm_2} = \delta^{m_1}{}_{m_1'}, \tag{3.3}$$

$$(j)^{m_1m_2}(j)_{m_1m_2'} = \delta^{m_2}{}_{m_2'}. \tag{3.4}$$

Also,

$$(j^*)_{mm'} = M(j0j^*)(j)_{m'm}, \tag{3.5}$$

where the matrix $M(j0j^*)$ is just a phase factor since $0 = 0^*$ appears exactly once in the product $j \times j$. From the choice made in Eq. (1.1) and from Eq. (3.2) it follows that

$$(j)_{mm'} = \beta_j(j)_{m'm}, \tag{3.6}$$

$$(j^*)_{mm'} = \gamma_j(j)_{mm'}, \tag{3.7}$$

where

$$\beta_j = +1 \quad \text{if } j \text{ is integer or complex}$$
$$= -1 \quad \text{if } j \text{ is half-integer,}$$

and

$$\gamma_j = +1 \quad \text{if} \quad j = j^*$$
$$= M(j0j^*) \quad \text{if} \quad j \neq j^*.$$

Note that γ_j could be chosen to be $+1$ in all cases but in accord with the point of view outlined previously we shall not make this particular choice. We will write $\lambda_j = \beta_j\gamma_j$ and also $\mu(123) = \lambda_1\lambda_2\lambda_3$ if $j_1j_2j_3$ form a triad. Clearly $\gamma_{j^*} = (\gamma_j)^*$ and therefore $\lambda_{j^*} = (\lambda_j)^*$. If the group G is SR, we can take $\lambda_j = (-1)^{2j}$ and it follows that all the phase factors μ are $+1$.

4. RAISING AND LOWERING OF INDICES

Before proving the basic lemma of the theory we define a unitary matrix A which will be used to raise and lower multiplicity indices. Let

$$A(123)_{rr'} = \{A(123)^{rr'}\}^*$$
$$= (j_1j_2j_3)_{r,m_1m_2m_3}(j_1^*j_2^*j_3^*)_{r',n_1n_2n_3}$$
$$\times (j_1)^{m_1n_1}(j_2)^{m_2n_2}(j_3)^{m_3n_3}. \tag{4.1}$$

The unitarity of A is easily shown to follow from the unitarity of the 1j- and 3j-symbols. The proofs of the basic lemma and of the theorems of the next section are similar to those given by Sharp[5] for quasi SR groups. Most of these proofs are straightforward but long. For this reason we will give detailed proofs only for the lemma and for some of the theorems.

Lemma: $\{(j_1j_2j_3)_{r,m_1m_2m_3}\}^*$

$$= A(123)^{rr'}(j_1)^{m_1n_1}(j_2)^{m_2n_2}(j_3)^{m_3n_3}(j_1^*j_2^*j_3^*)_{r',n_1n_2n_3}.$$

Proof: Because of the unitarity of the 1j-symbol, Eq. (3.2) can be written

$$j_1(R)^{m_1}{}_{m_1'}(j_1)^{m_1'n_1'} = j_1^*(R)_{n_1}{}^{n_1'}(j_1)^{m_1n_1}.$$

Using similar equations for j_2 and j_3, one obtains

$$j_1(R)^{m_1}{}_{m_1'}j_2(R)^{m_2}{}_{m_2'}j_3(R)^{m_3}{}_{m_3'}(j_1)^{m_1'n_1'}(j_2)^{m_2'n_2'}(j_3)^{m_3'n_3'}$$
$$= j_1^*(R)_{n_1}{}^{n_1'}j_2^*(R)_{n_2}{}^{n_2'}j_3^*(R)_{n_3}{}^{n_3'}(j_1)^{m_1n_1}(j_2)^{m_2n_2}(j_3)^{m_3n_3}.$$

Integrating each side over the group and using Eq. (2.5), one obtains

$$\{(j_1j_2j_3)_{r,m_1m_2m_3}\}^*(j_1j_2j_3)_{r,m_1'm_2'm_3'}$$
$$\times (j_1)^{m_1'n_1'}(j_2)^{m_2'n_2'}(j_3)^{m_3'n_3'}$$
$$= (j_1)^{m_1n_1}(j_2)^{m_2n_2}(j_3)^{m_3n_3}(j_1^*j_2^*j_3^*)_{r',n_1n_2n_3}$$
$$\times \{(j_1^*j_2^*j_3^*)_{r',n_1'n_2'n_3'}\}^*.$$

It then follows, from the unitarity of the 1j- and 3j-symbols, that

$$\{(j_1j_2j_3)_{r,m_1m_2m_3}\}^*$$
$$= \{(j_1j_2j_3)_{r,m_1'm_2'm_3'}(j_1^*j_2^*j_3^*)_{r',n_1'n_2'n_3'}$$
$$\times (j_1)^{m_1'n_1'}(j_2)^{m_2'n_2'}(j_3)^{m_3'n_3'}\}^*$$
$$\times (j_1)^{m_1n_1}(j_2)^{m_2n_2}(j_3)^{m_3n_3}(j_1^*j_2^*j_3^*)_{r',n_1n_2n_3}.$$

This is easily recognized to be the desired result.

The above lemma makes it useful to define systematically contravariant (raised) m and r indices as in the SR case.[1,5] We use the 1j-symbol to raise a single m index according to the definition

$$(j_1j_2j_3^*)_{r,m_1m_2}{}^{m_3} = (j_3^*)^{m_3m_3'}(j_1j_2j_3)_{r,m_1m_2m_3'}, \tag{4.2}$$

and the A matrix to raise a single multiplicity index according to

$$(j_1j_2j_3)^r{}'_{m_1m_2m_3} = A(1^*2^*3^*)^{rr'}(j_1j_2j_3)_{r',m_1m_2m_3}. \tag{4.3}$$

It follows that

$$(j_1j_2j_3^*)_{r,m_1m_2m_3} = (j_1j_2j_3)_{r,m_1m_2}{}^{m_3'}(j_3)_{m_3'm_3} \tag{4.4}$$

and

$$(j_1j_2j_3)_{r,m_1m_2m_3} = (j_1j_2j_3)^{r',}{}_{m_1m_2m_3}A(1^*2^*3^*)_{r',r}. \tag{4.5}$$

Our lemma states then that

$$\{(j_1j_2j_3)_{r,m_1m_2m_3}\}^* = (j_1j_2j_3)^{r,m_1m_2m_3}. \tag{4.6}$$

The matrix $(j)_{mm'}$ is symmetric when j is integer or complex but it is skew-symmetric when j is half-integer. Similarly the matrix A need not be symmetric. It is easy to show that

$$A(1^*2^*3^*)_{r',r} = \mu(123)^*A(123)_{rr'}. \tag{4.7}$$

If the group is ambivalent and if $\mu(123) = 1$ then $A(123)$ is symmetric.

We shall often drop the m indices and indicate the fact that a particular m index is contravariant by making the corresponding j boldface. Then our summation convention for m is that we must sum over all possible values of m when the corresponding j occurs twice (once in boldface type and once not). The $1j$-symbols can be used to raise or lower m's in any order. However the matrix A depends on three j's (in a certain order) and it raises the r index of the $3j$-symbol with all m's covariant. If some of the m's in the $3j$-symbol are contravariant a different matrix A must be used to raise the multiplicity index. It is easy to see that for each contravariant m the corresponding j in the argument of A must be replaced by its complex conjugate. Thus we write

$$(\mathbf{j}_1j_2j_3)_r = (j_1j_2\mathbf{j}_3)^*A(12^*3)_{*r},$$

where we use the fact that $j^{**} = j$. We are now in a position to define the $6j$-symbol.

5. THE $6j$-SYMBOL

Our definition of the $6j$-symbol is similar to the usual one. However the $6j$-symbol now depends on four multiplicity indices, one for each triad. We define

$$\begin{Bmatrix} j_{12}j_{13}j_{23} \\ j_{34}j_{24}j_{14} \end{Bmatrix}_{r_1r_2r_3r_4}$$
$$= (\mathbf{j}_{34}j_{24}j_{23})_{r_1}(j_{34}\mathbf{j}_{13}j_{14})_{r_2}(j_{12}j_{24}\mathbf{j}_{14})_{r_3}(j_{12}j_{13}\mathbf{j}_{23})_{r_4}. \tag{5.1}$$

This definition is to be extended to $6j$-symbols with contravariant multiplicity indices in an obvious way: any contravariant multiplicity index on the right of Eq. (5.1) must appear contravariantly on the $6j$-symbol. The matrices A can then be used to raise and lower multiplicity indices on the $6j$-symbols. Our first theorem gives an expression for the complex conjugated $6j$-symbol.

Theorem 1:

$$\begin{Bmatrix} j_{12}j_{13}j_{23} \\ j_{34}j_{24}j_{14} \end{Bmatrix}^*_{r_1r_2r_3r_4} = \begin{Bmatrix} j_{12}{}^*j_{13}{}^*j_{23}{}^* \\ j_{34}{}^*j_{24}{}^*j_{14}{}^* \end{Bmatrix}^{r_1r_2r_3r_4}.$$

Proof: To obtain the complex conjugate of a partly contravariant $3j$-symbol the properties of the $1j$-symbol must be used. For instance

$$\begin{aligned}\{(j_1j_2j_3)_r,{}^{m_1m_2}{}_{m_3}\}^* &= \{(j_1)^{m_1n_1}(j_2)^{m_2n_2}(j_1{}^*j_2{}^*j_3)_{r,n_1n_2m_3}\}^* \\ &= (j_1)_{m_1n_1}(j_2)_{m_2n_2}(j_1{}^*j_2{}^*j_3)^{r,n_1n_2m_3} \\ &= (j_1{}^*j_2{}^*j_3)^{r,n_1n_2m_3}(j_1{}^*)_{n_1m_1}(j_2{}^*)_{n_2m_2}(\lambda_1\lambda_2)^* \\ &= (\lambda_1\lambda_2)^*(j_1j_2j_3)^{r'}{}_{,m_1m_2}{}^{m_3}.\end{aligned}$$

For the $6j$-symbol one has

$$\begin{Bmatrix} j_{12}j_{13}j_{23} \\ j_{34}j_{24}j_{14} \end{Bmatrix}^*_{r_1r_2r_3r_4}$$
$$= (\lambda_{12}\lambda_{13}\lambda_{23}\lambda_{34}\lambda_{24}\lambda_{14})^*(\mathbf{j}_{34}\mathbf{j}_{24}\mathbf{j}_{23})^{r_1}$$
$$\times (\mathbf{j}_{34}\mathbf{j}_{13}\mathbf{j}_{14})^{r_2}(\mathbf{j}_{12}\mathbf{j}_{24}\mathbf{j}_{14})^{r_3}(\mathbf{j}_{12}\mathbf{j}_{13}\mathbf{j}_{23})^{r_4}.$$

If one now uses relations of the type

$$(j_1j_2j_3)_r(j_1'j_2'\mathbf{j}_3')_* = \lambda_3(j_1j_2\mathbf{j}_3{}^*)_r(j_1'j_2'j_3{}^*)_*$$

one obtains

$$\begin{Bmatrix} j_{12}j_{13}j_{23} \\ j_{34}j_{24}j_{14} \end{Bmatrix}^*_{r_1r_2r_3r_4}$$
$$= (\mathbf{j}_{34}{}^*j_{24}{}^*j_{23}{}^*)^{r_1}(j_{34}{}^*\mathbf{j}_{13}{}^*\mathbf{j}_{14}{}^*)^{r_2}$$
$$\times (\mathbf{j}_{12}{}^*\mathbf{j}_{24}{}^*\mathbf{j}_{14}{}^*)^{r_3}(\mathbf{j}_{12}{}^*\mathbf{j}_{13}{}^*\mathbf{j}_{23}{}^*)^{r_4},$$

which is the required result.

In the following theorem we give the symmetry properties of the $6j$-symbols under permutations of the j's that preserve the triads.

Theorem 2:

(i) $$\begin{Bmatrix} j_{12}j_{13}j_{23} \\ j_{34}j_{24}j_{14} \end{Bmatrix}_{r_1r_2r_3r_4} = \begin{Bmatrix} j_{12}{}^*j_{24}j_{14}{}^* \\ j_{34}{}^*j_{13}j_{23}{}^* \end{Bmatrix}_{r_3r_1r_4r_2} (\lambda_{12}\lambda_{23}\lambda_{14}\lambda_{34}),$$

(ii) $$\begin{Bmatrix} j_{12}j_{13}j_{23} \\ j_{34}j_{24}j_{14} \end{Bmatrix}_{r_1r_2}{}^{r_3r_4} = \begin{Bmatrix} j_{13} & j_{12} & j_{23} \\ j_{24}{}^* & j_{34}{}^* & j_{14}{}^* \end{Bmatrix}_{r_1' \quad r_2'}{}^{r_3' \quad r_4'} (\lambda_{24}\lambda_{14}\lambda_{34})$$

$$\times M(12\ 13,\ 23)_{r_4',}{}^{r_4}M(12^*24,\ 14^*)_{r_3',}{}^{r_3}\ M(34\ 13,\ 14^*)^{r_2'}{}_{,r_2}M(34^*24,\ 23)^{r_1'}{}_{,r_1},$$

(iii) $$\begin{Bmatrix} j_{12} j_{13} j_{23} \\ j_{34} j_{24} j_{14} \end{Bmatrix}_{r_1 r_2}^{\quad r_3 r_4} = \begin{Bmatrix} j_{13} j_{23} j_{12} \\ j_{24} j_{14} j_{34} \end{Bmatrix}_{r_1' r_3'}^{\quad r_3' \quad r_4'} \{M(13, 12\ 23)M(12\ 13, 23)\}_{r_4', r_4}$$
$$\times \{M(24, 12{*}14{*})M(12{*}24, 14{*})\}_{r_3', r_3} \{M(13, 34\ 14{*})M(34\ 13, 14{*})\}^{r_3', r_2}$$
$$\times \{M(24, 34{*}23)M(34{*}24, 23)\}^{r_1', r_1},$$

where $M(1\ 2\ 3)^{r}{}_{s} = \{M(1\ 2\ 3)_{r}{}^{s}\}^{*}$ and where $\{M(2, 1\ 3)M(1\ 2, 3)\}_{r,}{}^{r} = M(2, 1\ 3)_{r,}{}^{s}M(1\ 2, 3)_{s}{}^{r}$.

Any other permutation of the j's which preserves the triads can be obtained from the three basic permutations given above. The multiplicity indices on the $6j$-symbols can be raised and lowered with the unitary matrices A with the appropriate arguments.

In the quantum theory of angular momentum the $6j$-symbols are often thought of as recoupling coefficients since they arise from the associativity of the Kronecker product. The next theorem shows that the general $6j$-symbol can also be considered as a recoupling coefficient.

Theorem 3:

$$(\mathbf{j}_{12}\mathbf{j}_{13}\mathbf{j}_{23})^{r_4}(j_{34}j_{24}j_{23})_{r_1} = \sum_{j_{14}} [j_{14}]\lambda_{24}{*}(j_{34}\mathbf{j}_{13}\mathbf{j}_{14})^{r_2}(\mathbf{j}_{12}j_{24}j_{14})^{r_3}\begin{Bmatrix} j_{12} & j_{13}j_{23} \\ j_{34}{*} & j_{24}j_{14}{*} \end{Bmatrix}_{r_1 r_2 r_3}^{\quad r_4}.$$

Proof: Let T be defined by

$$(\mathbf{j}_{12}\mathbf{j}_{13}\mathbf{j}_{23})^{r_4}(j_{34}j_{24}j_{23})_{r_1} = \sum_{j_{14}} [j_{14}]\lambda_{24}{*}(j_{34}\mathbf{j}_{13}\mathbf{j}_{14})^{r_2}(\mathbf{j}_{12}j_{24}j_{14})^{r_3}T_{r_1 r_2 r_3}{}^{r_4}. \tag{5.2}$$

We wish to show that

$$T_{r_1 r_2 r_3}{}^{r_4} = \begin{Bmatrix} j_{12} & j_{13}\ j_{23} \\ j_{34}{*} & j_{24}\ j_{14}{*} \end{Bmatrix}_{r_1 r_2 r_3}^{\quad r_4}.$$

The index m_{34} can be raised on each side of Eq. (5.2) and after using the unitarity of the $3j$-symbols one obtains

$$(\mathbf{j}_{12}\mathbf{j}_{13}\mathbf{j}_{23})^{r_4}(j_{34}{*}j_{24}\ j_{23})_{r_1}(j_{34}{*}j_{13}j_{14})_{r_2} = \lambda_{24}{*}(j_{12}j_{24}j_{14})^{r_3}T_{r_1 r_2 r_3}{}^{r_4}.$$

If m_{14} is raised, then it follows from the unitarity of the $3j$-symbols that

$$T_{r_1 r_2 r_3}{}^{r_4} = (\mathbf{j}_{12}\mathbf{j}_{13}\mathbf{j}_{23})^{r_4}(\mathbf{j}_{34}{*}j_{24}j_{23})_{r_1}(j_{34}{*}j_{13}\mathbf{j}_{14}{*})_{r_2}(j_{12}\mathbf{j}_{24}j_{14}{*})_{r_3},$$

and by definition this is the $6j$-symbol which appears in the statement of the theorem.

Theorem 4: The $6j$-symbols satisfy the unitarity condition

$$\sum_{j_{14}} [j_{14}][j_{23}] \begin{Bmatrix} j_{12} j_{13} j_{23} \\ j_{34} j_{24} j_{14} \end{Bmatrix}^{*}_{r_1 r_2 r_3 r_4} \begin{Bmatrix} j_{12} j_{13} j_{23}' \\ j_{34} j_{24} j_{14} \end{Bmatrix}_{r_1' r_2' r_3' r_4'} = \delta_{j_{23} j_{23}'} \delta_{r_1, r_1'} \delta_{r_4, r_4'},$$

except in the trivial cases (i.e., $j_{23}{*}$ does not appear in $j_{12} \times j_{13}$ or in $j_{34}{*} \times j_{24}$) when the left-hand side is zero.

Proof: The preceding Theorem 3 can be used to expand the expression

$$\alpha = (j_{12}j_{13}j_{23})^{r_4}(\mathbf{j}_{34}{*}\mathbf{j}_{24}\mathbf{j}_{23})^{r_1}(\mathbf{j}_{12}\mathbf{j}_{13}\mathbf{j}_{23}')_{r_4'}(j_{34}{*}j_{24}j_{23}')_{r_1'}$$
$$= \sum_{j_{14}, j_{14}'} [j_{14}][j_{14}']\mu(12\ 13\ 23)\lambda_{24}\lambda_{24}{*}$$
$$\times \{(j_{34}{*}\mathbf{j}_{13}\mathbf{j}_{14})^{r_2}(\mathbf{j}_{12}j_{24}j_{14})^{r_3}\}^{*}(j_{34}{*}\mathbf{j}_{13}\mathbf{j}_{14}')^{r_2'}(\mathbf{j}_{12}j_{24}j_{14}')^{r_3'}$$
$$\times \begin{Bmatrix} j_{12} j_{13} j_{23} \\ j_{34} j_{24} j_{14}{*} \end{Bmatrix}^{*}_{r_1 r_2 r_3 r_4} \begin{Bmatrix} j_{12} j_{13} j_{23}' \\ j_{34} j_{24} j_{14}'{*} \end{Bmatrix}_{r_1' r_2' r_3' r_4'}.$$

It follows from the unitarity of the $3j$-symbols that

$$\alpha = \sum_{j_{14}} [j_{14}]\mu(12\ 13\ 23) \begin{Bmatrix} j_{12} j_{13} j_{23} \\ j_{34} j_{24} j_{14}{*} \end{Bmatrix}^{*}_{r_1 r_2 r_3 r_4} \begin{Bmatrix} j_{12} j_{13} j_{23}' \\ j_{34} j_{24} j_{14}{*} \end{Bmatrix}_{r_1' r_2' r_3' r_4'}.$$

But if one uses the unitarity of the $3j$-symbols directly in the expression for α one obtains

$$\alpha = (1/[j_{23}])\,\delta_{j_{23}j_{23}'}\delta_{r_1r_1'}\delta_{r_4r_4'}\mu(12\ 13\ 23),$$

which completes the proof.

The proofs of the following two theorems are long and uninspiring. The procedure is the same as for quasi-SR groups[5] and will not be given here.

Theorem 5: (Racah back-coupling rule)

$$\begin{Bmatrix} j_{12} & j_{13} & j_{23} \\ j_{34} & j_{24} & j_{14} \end{Bmatrix}^{r_3r_4}_{r_1r_2} = \sum_j [j]\lambda_j(\lambda_{34}\lambda_{12}\lambda_{24})^* \begin{Bmatrix} j_{13} & j_{12} & j_{23} \\ j_{34} & j_{24} & j \end{Bmatrix}^{r' \quad r_4'}_{r_1 \quad r''} \begin{Bmatrix} j_{12} & j_{34} & j^* \\ j_{13} & j_{24} & j_{14} \end{Bmatrix}^{r'' \quad r_3}_{r_2' \quad r}$$

$$\times M(13\ 34,\ 14^*)_{r_2}{}^{r_2'}M(12\ 13,\ 23)_{r_4'}{}^{r_4}M(12^*34^*,\ j)_{r''}{}^{r'}.$$

Theorem 6: (Biedenharn identity)

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ j_{23} & j_{13} & j_{12} \end{Bmatrix}^{r_1r}_{r_2r_2} \begin{Bmatrix} j_1 & j_2 & j_3 \\ j_{23}' & j_{13}' & j_{12}' \end{Bmatrix}^{*\ r_1'r}_{r_2'r_2'}$$

$$= \sum_j [j] \begin{Bmatrix} j_{13}' & j_{13}^* & j \\ j_{12} & j_{12}'j_1^* \end{Bmatrix}^{r_1'' \quad r''}_{s \quad s_1} \begin{Bmatrix} j_{12}' & j_{12}^* & j \\ j_{23} & j_{23}'j_2^* \end{Bmatrix}^{s_2r'}_{s'r_2''} \begin{Bmatrix} j_{23}' & j_{23}^* & j \\ j_{13} & j_{13}'j_3^* \end{Bmatrix}^{r_3's''}_{rr_2''}$$

$$\times \lambda_2\lambda_3(\lambda_{12'}{}^2\lambda_{23'})^*\, M(13^*13',\ j)_{r''}{}^{r}M(12^*12',\ j)_{r'}{}^{s}M(23^*23',\ j)_{s''}{}^{s'}$$

$$\times M(12^*13\ 1^*)^{r_1}{}_{r_1'}M(23,\ 12^*2)_{r_2}{}^{r_2''}M(13\ 23^*,\ 3)_{r_3}{}^{r_3''}$$

$$\times \{M(12'^*13'1^*)M(13'12'^*,\ 1^*)\}_{r_1'}{}^{s_1}$$

$$\times \{M(2\ 23'12'^*)M(23'2,\ 12'^*)\}_{s_2}{}^{r_2'}.$$

We now give a generalization of the not very well-known group integral formulas.

Theorem 7:

(i) $$\begin{Bmatrix} j_1 & j_2j' \\ j_1^* & j_2j \end{Bmatrix}^{r\ r'}_{r'\ r} = \lambda_1^*\lambda_2(1/A)^2 \iint \chi^{i'}(R)\chi^{i}(S)\chi^{i_1}(RS)\chi^{i_2}(RS^{-1})\,dR\,dS,$$

where $\chi^i(R)$ is the character of R in the representation j,

(ii) $$\begin{Bmatrix} j_{12}j_{13}j_{23} \\ j_{34}j_{24}j_{14} \end{Bmatrix}_{r_1r_2r_3r_4} \begin{Bmatrix} j_{12}j_{13}j_{23} \\ j_{34}j_{24}j_{14} \end{Bmatrix}^*_{r_1r_2r_3r_4} = (1/A)^3 \iiint \chi^{i_{12}}(R_1)\chi^{i_{13}}(R_2)\chi^{i_{23}}(R_3)\chi^{i_{24}}(R_1^{-1}R_3)$$

$$\times \chi^{i_{14}}(R_3^{-1}R_2)\chi^{i_{14}}(R_1R_2^{-1})\,dR_1\,dR_2\,dR_3,$$

(iii) $$\begin{Bmatrix} j'j_1j_2 \\ j\ j_3j_2^* \end{Bmatrix}^{r_{23}\quad\ r_{12}'}_{r_{12}r_{23}'} \begin{Bmatrix} j'\ j_1j_2 \\ j''j_3j_2^* \end{Bmatrix}^{*\ r_{12}'}_{r_{23}''r_{12}''r_{23}'} \begin{Bmatrix} j\ \ j_1j_2 \\ j''\ j_3j_2^* \end{Bmatrix}^{r_{12}}_{r_{23}''r_{12}''r_{23}}$$

$$= (1/A)^6(\lambda_3) \int \cdots \int \chi^{i}(TV^{-1})\chi^{i'}(QS^{-1})\chi^{i''}(UW^{-1})$$

$$\times \chi^{i_1}(QUT)\chi^{i_2}(QWTSUV)\chi^{i_3}(SWV)\,dQ\,dS\,dT\,dU\,dV\,dW$$

$$= (1/A)^4(\lambda_3) \iiiint \chi^{i}(R)\chi^{i'}(R')\chi^{i''}(R'')\chi^{i_1}(PR')$$

$$\times \chi^{i_2}(R'R''^{-1}P^2R^{-1})\chi^{i_3}(R''^{-1}PR^{-1})\,dR\,dR'\,dR''\,dP,$$

(iv) $$\begin{Bmatrix} j_{12}j_{13}^*\ j_{23} \\ j_{34}j_{24}\ \ j_{14} \end{Bmatrix}^{r_1r_2r_3r_4} \begin{Bmatrix} j_{12}j_{13}^*\ j_{23} \\ j_{12}j_{24}\ \ j_{14} \end{Bmatrix}^*_{r_1'r_2'} \begin{Bmatrix} j_{12}\ \ j_{13}j_{23} \\ j_{34}^*\ j_{13}j_{14}^* \end{Bmatrix}^{r_4'}_{r_3'r_2'r_2'} \left[\begin{Bmatrix} j_{12}j_{13}j_{23} \\ j_{34}j_{24}j_{23}^* \end{Bmatrix}^{r_1\quad r_1'r_4'}_{r_3'} \right]^*$$

$$= (1/A)^8(\lambda_{23}\lambda_{12}^*\lambda_{14}^*) \int \cdots \int \chi^{i_{12}}(P_1P_2)\chi^{i_{13}}(P_5P_3P_7^{-1})$$

$$\times \chi^{i_{13}}(P_1^{-1}P_4^{-1}P_3P_8P_5^{-1})\chi^{i_{24}}(P_2^{-1}P_7P_6)\chi^{i_{23}}(P_1P_7P_8P_3P_6)$$

$$\times \chi^{j_{14}}(P_2P_4^{-1}P_5^{-1})\chi^{j_{56}}(P_4P_6^{-1}P_8)\,dP_1\,dP_2\,dP_3\,dP_4\,dP_5\,dP_6\,dP_7\,dP_8$$

$$= (1/A)^5(\lambda_{23}\lambda_{12}{}^*\lambda_{14}{}^*)\int\cdots\int \chi^{j_{12}}(R)\chi^{j_{13}}(R_2^{-1}R_1^{-1}S)$$

$$\times \chi^{j_{13}}(R^{-1}SR_3R_1)\chi^{j_{24}}(R_2)\chi^{j_{35}}(SRR_2R_3)\chi^{j_{14}}(R_1)\,\chi^{j_{56}}(R_3)\,dR\,dS\,dR_1\,dR_2\,dR_3.$$

Proof: As an example we prove the first group integral formula. By definition,

$$\begin{Bmatrix} j_1 & j_2 & j' \\ j_1^* & j_2 & j \end{Bmatrix}^{r\,r'}_{r'\,r} = (j_1j_2j')^{r',m_1m_2'm'}(j_1j_2j)_{r,m_1}{}^{m_2}{}_m\,(j_1{}^*j_2j)^{r',}{}_{m_1'm_2'}{}^m(j_1{}^*j_2j')_{r',}{}^{m_1'}{}_{m_2m'}$$

$$= \lambda_1{}^*(j_1j_2j')^{r',m_1m_2'm'}(j_1j_2j')_{r',m_1'm_2m'}\,(j_2)^{m_2m_2'''}(j_1j_2{}^*j)_{r,m_1m_2'''m}(j_1j_2{}^*j)^{r,m_1'm_2''m}(j_2{}^*)_{m_2''m_2'},$$

where we have used the properties of the $1j$-symbols. Using Eq. (2.5) we obtain

$$\begin{Bmatrix} j_1 & j_2 & j' \\ j_1^* & j_2 & j \end{Bmatrix}^{r\,r'}_{r'\,r} = \lambda_1{}^*\lambda_2(1/A)^2\iint j_1(R)^{m_1}{}_{m_1'}j_2(R)^{m_2'}{}_{m_2}j'(R)^{m'}{}_{m'}$$

$$\times j_1(S)^{m_1'}{}_{m_1}j_2{}^*(S)^{m_2''}{}_{m_2'''}j(S)^m{}_m\,dR\,dS\,(j_2)^{m_2m_2'''}(j_2)_{m_2'm_2''}$$

$$= \lambda_1{}^*\lambda_2(1/A)^2\iint \chi^{j'}(R)\chi^{j}(S)\chi^{j_1}(RS)j_2(R)^{m_2'}{}_{m_2}j_2(S)_{m_2'}{}^{m_2}\,dR\,dS,$$

where we use Eq. (3.2). Now since the representation matrices are unitary,

$$j_2(R)^{m_2'}{}_{m_2}j_2(S)_{m_2'}{}^{m_2} = j_2(RS^{-1})^{m_2'}{}_{m_2'} = \chi^{j_2}(RS^{-1}),$$

and the result follows.

We have seen how the properties of the $6j$-symbols follow from the properties of the $3j$-symbols in a rather straightforward manner. Similarly there does not seem to be any difficulty in extending the theory to the $9j$-symbols. The $9j$-symbol could be defined by

$$\begin{bmatrix} j_{11} & j_{12} & j_{13} \\ j_{21} & j_{22} & j_{23} \\ j_{31} & j_{32} & j_{33} \end{bmatrix}^{uvw}_{rst} = (j_{11}j_{12}j_{13})_r(j_{21}j_{22}j_{23})_s(j_{31}j_{32}j_{33})_t(j_{11}j_{21}j_{31})^u(j_{12}j_{22}j_{32})^v(j_{13}j_{23}j_{33})^w.$$

The reader can verify for example that

$$\begin{bmatrix} j_{11} & j_{12} & j_{13} \\ j_{12} & j_{22} & j_{23} \\ j_{13} & j_{23} & j_{33} \end{bmatrix}^{rst}_{rst} = (1/A)^3\iiint \chi^{j_{11}}(R_1)\chi^{j_{22}}(R_2)\chi^{j_{33}}(R_3)\chi^{j_{12}}(R_1R_2)\chi^{j_{13}}(R_1R_3)\chi^{j_{23}}(R_2R_3)\,dR_1\,dR_2\,dR_3.$$

6. CONCLUSION

$3j$-symbols for $SU(3)$ have been studied by many authors[10] in connection with elementary-particle and nuclear theory and these symbols have been investigated also for some other groups[11] that are not multiplicity free. In each case it is necessary to make some explicit choice of labeling scheme for the multiplicity index. It is plausible that the most advantageous choice for a given group depends on the specific application to be made. The present paper provides a framework in which to compare different choices and to see what properties are independent of this arbitrary choice of basis. We have seen that $3j$- and $6j$-symbols can indeed be defined generally with essentially all the usual properties. In a subsequent paper we hope to discuss what simplifying choices of phase can be made for an arbitrary group. (Can M for example be taken in diagonal form?) and some interesting special properties of the $3j$- and $6j$-symbols for particular physically interesting groups like $SU(3)$.

[10] J. J. de Swart, Rev. Mod. Phys. **35**, 916 (1963); J. J. de Swart, Nuovo Cimento **31**, 420 (1964); M. Moshinsky, Rev. Mod. Phys. **34**, 813 (1962); G. E. Baird and L. C. Biedenharn, J. Math. Phys. **4**, 1449 (1963); V. Vanagas and A. Jucys, Sb. Litov. Fiz. **2**, 199 (1962).

[11] L. C. Biedenharn, Phys. Letters **3**, 254 (1963); see also G. E. Baird and L. C. Biedenharn, J. Math Phys. **5**, 1730 (1964); G. F. Koster, J. O. Dimmock, R. G. Wheeler, and H. Statz, *Properties of The Thirty-Two Point Groups* (MIT Press, Cambridge, Massachusetts, 1963).

Time-Dependent Perturbation Theory*

C. L. HAMMER AND T. A. WEBER
Institute for Atomic Research and Department of Physics, Iowa State University, Ames, Iowa
(Received 11 September 1964; final manuscript received 12 April 1965)

A time-dependent perturbation theory which is based upon the U-matrix approach is presented using the interaction $e^{-\alpha t}(1 + e^{-\beta t})^{-1}V(\mathbf{x})$ with β real, α complex and $\beta > \text{Re}\,\alpha > 0$. Thus the adiabatic or time-independent approximation ($\beta \to 0$) and the sudden approximation ($\beta \to \infty$) can be obtained using just one formalism. The usual series for the U matrix is derived and shown to converge for a semi infinite range if the interaction is of the form $|t|^{-\delta}$, $\delta > 1$, for large times t. Two interesting results are (1) The derivation of the "golden rule" from the discrete-state time-dependent perturbation theory presented in most text books of quantum mechanics leads to erroneous physical interpretations because the energy-level shift is ignored. (2) When considering scattering between states in the continuum, it is found that a characteristic feature of time-dependent interactions is a discrete momentum spectrum of final states which, in the relativistic case, leads to a mass spectrum. These spectra cannot be obtained using the S matrix.

INTRODUCTION

IN most books on quantum mechanics[1] the "golden rule" is found from first-order perturbation theory to be

$$w = (2\pi/\hbar)\ |\langle k|\ A\ |0\rangle|^2\ \delta(E_k - E_0 - \hbar\omega), \tag{1}$$

where w is the rate at which transitions take place from the E_0 to the E_k energy state, and $A(\mathbf{x})$ arises from a step-function perturbation of the form

$$\begin{aligned} V(\mathbf{x}, t) &= A(\mathbf{x})e^{-i\omega t}; \qquad t \geq 0, \\ V(\mathbf{x}, t) &= 0; \qquad\qquad\ t < 0. \end{aligned} \tag{2}$$

Difficulties are found to occur when transitions of second, or higher, order are considered because terms arise which do not conserve energy (usually these discussions are made with $\omega = 0$). The nonconservation of energy is blamed upon the sudden turning on of the perturbation and the explanation is made that the rapid turn on is an artifice used to simplify the calculation. These terms are usually dropped[2,3] or subtracted out by altering the initial conditions.[4]

On the other hand when the "sudden approximation,"[5] based on a step-function perturbation, such as given by Eq. (2) with $\omega = 0$, is considered, it is found that the probability for being in an excited state is independent of time. Consequently for this approximation for $t > 0$ the transition probability w is zero. It is apparent that a paradox exists between the two different approaches.

One wonders whether this paradox arises because of improper limiting processes. That this is not the case is argued below where it is shown that the "non-energy-conserving" terms are needed to account for the shift in energy of the energy levels of the unperturbed system due to the perturbation. When this energy shift is accounted for, the transition rate vanishes in agreement with the "sudden approximation."

To examine this point, time-dependent perturbation theory is reformulated in terms of the U matrix[6]. The expansion of this matrix in powers of the perturbation is shown to converge absolutely and uniformly providing that the perturbation $H'(\mathbf{x}, t)$ is bounded and approaches zero faster than $|t|^{-\delta}$, $\delta > 1$ for $|t| \to \infty$. By considering the perturbation to be of the form

* Work was performed in the Ames Laboratory of the U. S. Atomic Energy Commission, Contribution No. 1444.
[1] (a) L. I. Schiff, *Quantum Mechanics*, (McGraw-Hill Book Company, Inc., New York, 1955), 2nd ed., p. 195; (b) L. D. Landau and E. W. Lifshitz, *Quantum Mechanics* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1958), p. 140.
[2] Reference 1(a), p. 202; Ref. 1(b), p. 145.
[3] W. Heitler, *Quantum Theory of Radiation* (Oxford University Press, London, 1954), 3rd ed., p. 140.
[4] Reference 3, p. 164.
[5] Reference 1(a), p. 217.
[6] S. Schweber, *An Introduction to Relativistic Quantum Field Theory* (Row, Peterson and Company, New York, 1961), pp. 7, 317, 331, 334, 335.

$$H'(\mathbf{x}, t) = e^{-\alpha t}[1 + e^{-\beta t}]^{-1} V(\mathbf{x}) \tag{3}$$

with β real, α complex, and $\beta > \text{Re } \alpha > 0$, both the "adiabatic" and "sudden" approximations can be investigated by allowing $\beta \to 0$ or $\beta \to \infty$ at the conclusion of the calculations so that limiting questions are no longer pertinent.

In addition to the discrete case, scattering from one continuum state to another due to time-dependent interactions is also considered. The nth term of the perturbation expansion for the U matrix is obtained for the familiar "adiabatic" potential $V(x)e^{-\alpha|t|}$, with α real, which can be considered as a limiting case for $\beta \to \infty$ of the interaction $H'(\mathbf{x}, t) + H'(\mathbf{x}, -t)$. Identical results to the usual ones are obtained for the S matrix[6] in the limit $\alpha \to 0$.

In addition, it is found that a characteristic feature of time-dependent interactions is a discrete momentum spectrum of final states, which in the relativistic case, can be interpreted as leading to a mass spectrum. These spectra cannot be obtained using the S matrix.

U-Matrix Formalism

One considers the time-dependent Schrödinger equation

$$[H_0(\mathbf{x}) + H'(\mathbf{x}, t)]\psi_S(\mathbf{x}, t) = i\hbar\, \partial\psi_S(\mathbf{x}, t)/\partial t, \tag{4}$$

where $H'(\mathbf{x}, t)$ is Hermitian and is to be regarded as a time-dependent perturbation and H_0 is the unperturbed Hamiltonian, time-independent, but not necessarily representing a free particle. The perturbation expansion is most easily derived by utilizing the interaction or Dirac picture. The relationship between the Schrödinger wavefunction $\psi_S(\mathbf{x}, t)$ and the interaction wavefunction $\psi_I(\mathbf{x}, t)$ is the unitary transformation $\exp[(i\hbar)^{-1}H_0(\mathbf{x})t]$, where

$$\psi_S(\mathbf{x}, t) = \exp[(i\hbar)^{-1}H_0(x)t]\psi_I(\mathbf{x}, t). \tag{5}$$

The equation for the time development of ψ_I can be obtained from Eq. (4) by direct substitution. Thus

$$H_I'(\mathbf{x}, t)\psi_I(\mathbf{x}, t) = i\hbar\, \partial\psi_I(\mathbf{x}, t)/\partial t, \tag{6}$$

where

$$H_I'(\mathbf{x}, t) = \exp[-(i\hbar)^{-1}H_0(\mathbf{x})t]H'(\mathbf{x}, t) \exp[(i\hbar)^{-1}H_0(\mathbf{x})t]. \tag{7}$$

A formal solution to Eq. (6) in terms of the U matrix is (6)

$$\psi_I(t) = U(t, -t_0)\psi_I(-t_0), \tag{8}$$

where

$$U(t, -t_0) = 1 + (i\hbar)^{-1}\int_{-t_0}^{t} d\xi_1\, H_I'(\xi_1) + \cdots + (i\hbar)^{-n}\int_{-t_0}^{t} d\xi_1 \int_{-t_0}^{\xi_1} d\xi_2 \cdots \int_{-t_0}^{\xi_{n-1}} d\xi_n\, H_I'(\xi_2)\cdots H_I'(\xi_n) + \cdots . \tag{9}$$

Time differentiation of Eq. (9) shows that

$$i\hbar[\partial U(t, -t_0)/\partial t] = H_I'(t)U(t, -t_0), \tag{10}$$

so that $\psi_I(t)$ as defined by Eqs. (8) and (9) is a solution to Eq. (6).

The following properties of the U matrix are readily demonstrated[6]:

$$U^\dagger(t, -t_0)U(t, -t_0) = 1, \tag{11a}$$

$$U(t, t')U(t', -t_0) = U(t, -t_0), \tag{11b}$$

$$U(t, -t_0) = U^{-1}(-t_0, t) = U^\dagger(-t_0, t), \tag{11c}$$

$$U(-t_0, -t_0) = 1. \tag{11d}$$

The series given in Eq. (9) converges if $H_I'\psi_I(-t_0)$ is bounded, that is, provided that

$$H_I'(\mathbf{x}, t)|\ \psi_I(-t_0) \leq |H'(\mathbf{x}, t)|\ \psi_I(-t_0) \leq M\psi_I(-t_0), \tag{12}$$

where M is finite and independent of t. Thus for the nth term of the series,

$$|U^{(n)}(t, -t_0)| \leq \hbar^{-n}M^n \int_{-t_0}^{t} d\xi_1 \int_{-t_0}^{\xi_1} d\xi_2 \cdots \int_{-t_0}^{\xi_{n-1}} d\xi_n = \hbar^{-n}M^n(t + t_0)^n(n!)^{-1}. \tag{13}$$

The right-hand side of Eq. (13) is the general term of the exponential series. Since this series converges absolutely and uniformly for the finite interval $t + t_0$, the series for $U(t, -t_0)$ converges absolutely and uniformly in this interval.

It can also be shown that the series converges for semi-infinite and infinite intervals. For example, for $t < -t_1$, where t_1 is finite but as large as desired, let

$$|H_I'(\mathbf{x}, t)| \leq N\ |t|^{-\delta}; \qquad \delta > 1, \tag{14}$$

where N can be any large but finite function of $\mathbf{x}$, and consider Eq. (9) with $t_0 = \infty$. Then,

$$\begin{aligned} U^{(n)}(t, -\infty) = (i\hbar)^{-n}\Big\{ &\int_{-\infty}^{-t_1} d\xi_1 \int_{-\infty}^{\xi_1} d\xi_2 \cdots \int_{-\infty}^{\xi_{n-1}} d\xi_n [H_I'(\xi_1) \cdots H_I'(\xi_n)] \\ &+ \int_{-t_1}^{t} d\xi_1 \int_{-\infty}^{-t_1} d\xi_2 \int_{-\infty}^{\xi_2} d\xi_3 \cdots \int_{-\infty}^{\xi_{n-1}} d\xi_n\ [H_I'(\xi_1) \cdots H_I'(\xi_n)] \\ &+ \int_{-t_1}^{t} d\xi_1 \int_{-t_1}^{\xi_1} d\xi_2 \int_{-\infty}^{-t_1} d\xi_3 \int_{-\infty}^{\xi_3} d\xi_4 \cdots \int_{-\infty}^{\xi_{n-1}} d\xi_n\ [H_I'(\xi_1) \cdots H_I'(\xi_n)] + \cdots \\ &+ \int_{-t_1}^{t} d\xi_1 \int_{-t_1}^{\xi_1} d\xi_2 \cdots \int_{-t_1}^{\xi_{n-2}} d\xi_{n-1} \int_{-\infty}^{t_1} d\xi_n\ [H_I'(\xi_1) \cdots H_I'(\xi_n)] \\ &+ \int_{-t_1}^{t} d\xi_1 \int_{-t_1}^{\xi_1} d\xi_2 \cdots \int_{-t_1}^{\xi_{n-1}} d\xi_n\ [H_I'(\xi_1) \cdots H_I'(\xi_n)]\Big\}. \end{aligned} \tag{15}$$

For the finite interval, $-t_1 \leq \xi_n \leq t$, Eq. (12) applies and results are obtained which are similar to those of Eq. (13). For the semi-infinite interval one observes that a typical term of Eq. (15) contains the integral

$$\begin{aligned} &\left| \int_{-\infty}^{-t_1} d\xi_1 \int_{-\infty}^{\xi_1} d\xi_2 \cdots \int_{-\infty}^{\xi_{l-1}} d\xi_l\ [H_I'(\xi_1) \cdots H_I'(\xi_l)] \right| \\ &\quad \leq N^l \int_{-\infty}^{-t_1} d\xi_1 \int_{-\infty}^{\xi_1} d\xi_2 \cdots \int_{-\infty}^{\xi_{l-1}} d\xi_l\ [(-\xi_1)^{-\delta}(-\xi_2)^{-\delta} \cdots (-\xi_l)^{-\delta}] = N^l(t_1)^{l(1-\delta)}/(\delta - 1)^l(l)!. \end{aligned} \tag{16}$$

Therefore,

$$\begin{aligned} |U^{(n)}(t, -\infty)| &\leq \hbar^{-n}\left\{\frac{N^n(t_1)^{n(1-\delta)}}{(\delta - 1)^n n!} + \cdots + \frac{N^{n-m}(t_1)^{(n-m)(1-\delta)}}{(\delta - 1)^{n-m}(n - m)!} \frac{M^m(t + t_1)^m}{m!} + \cdots + \frac{M^n(t + t_1)^n}{n!}\right\}, \\ &\leq \frac{\hbar^{-n}}{n!}\left[\frac{Nt_1^{(1-\delta)}}{\delta - 1} + M(t + t_1)\right]^n. \end{aligned} \tag{17}$$

Since the right-hand side of Eq. (17) is the general term of an exponential series, the series for $U(t, -\infty)$ converges absolutely and uniformly. A parallel proof can be made for $U(\infty, t)$, and, for the S matrix, $U(\infty, -\infty)$, the proof can be made if for $t > -t_1$,

$$|H_I'(\mathbf{x}, t)| \leq L(t_2 + t)^{-\delta}; \qquad t_2 > t_1; \qquad \delta > 1, \tag{18}$$

where L is any large but finite function of $\mathbf{x}$.

TIME-DEPENDENT PERTURBATION THEORY

The function $\psi_I(t)$ can be expanded in terms of a complete, orthonormal set of eigenfunctions $u_n(x)$,

$$\psi_I(t) = \sum_n a_n(t)u_n(\mathbf{x}) \tag{19}$$

where

$$H_0(\mathbf{x})u_n(\mathbf{x}) = E_n u_n(\mathbf{x}).$$

As an initial condition, consider the system to be in the stationary state $u_0(\mathbf{x})$ at $t = -t_0$ so that

$$\psi_I(-t_0) = u_0(\mathbf{x}) \tag{20}$$

Comparison of Eq. (19) to Eq.(8) shows that

$$\sum a_n(t)u_n(\mathbf{x}) = U(t, -t_0)\psi_I(-t_0) = U(t, -t_0)u_0(\mathbf{x}). \tag{21}$$

Therefore,

$$a_k(t) = \langle u_k| U(t, -t_0) |u_0\rangle \equiv U_{k0}(t, -t_0). \tag{22}$$

The isometric condition for $U(t, -t_0)$, given by Eq. (11a), guarantees that

$$\sum_k |a_k(t)|^2 = 1,$$

so that the probability for being in the kth state at the time t is

$$|a_k(t)|^2 = |U_{k0}(t, -t_0)|^2. \tag{23}$$

The rate of transition, $w(t)$, from the 0th to the kth state is then given by

$$w(t) = d\,|a_k(t)|^2/dt = d\,|U_{k0}(t, -t_0)|^2/dt. \tag{24}$$

The expansion for the matrix element $U_{k0}(t, -t_0)$ in powers of the interaction can be directly obtained from Eq. (9) as

$$\begin{aligned} U_{k0}(t, -t_0) = a_k(t) = \delta_{k0} &+ (i\hbar)^{-1} \int_{-t_0}^{t} d\xi_1 \langle u_k| H_I'(\xi_1) |u_0\rangle \\ &+ (i\hbar)^{-2} \int_{-t_0}^{t} d\xi_1 \int_{-t_0}^{\xi_1} d\xi_2 \langle u_k| H_I'(\xi_1) |u_l\rangle\langle u_l| H_I'(\xi_2) |u_0\rangle \\ &+ (i\hbar)^{-3} \int_{-t_0}^{t} d\xi_1 \int_{-t_0}^{\xi_1} d\xi_2 \int_{-t_0}^{\xi_2} d\xi_3 \langle u_k| H_I'(\xi_1) |u_l\rangle\langle u_l| H_I'(\xi_2) |u_m\rangle\langle u_m| H_I'(\xi_3) |u_0\rangle + \cdots . \end{aligned} \tag{25}$$

The repeated indices indicate the summation over discrete states and the integration over the continuous states. Substitution for $H_I'(\xi_n)$ from Eq. (7) gives

$$\begin{aligned} U_{k0}(t, -t_0) = \delta_{k0} &+ (i\hbar)^{-1} \int_{-t_0}^{t} d\xi_1 \langle u_k| H'(\xi_1) |u_0\rangle \exp(i\omega_{k0}\xi_1) \\ &+ (i\hbar)^{-2} \int_{-t_0}^{t} d\xi_1 \langle u_k| H'(\xi_1) |u_l\rangle \exp(i\omega_{kl}\xi_1) \int_{-t_0}^{\xi_1} d\xi_2 \langle u_l| H'(\xi_2) |u_0\rangle \exp(i\omega_{l0}\xi_2) + \cdots , \end{aligned} \tag{26}$$

where

$$\omega_{kn} = \hbar^{-1}(E_k - E_n).$$

A suitable way to treat transitions between discrete states and scattering between continuous states in a parallel manner is to use the Fourier transform

$$H'(\mathbf{x}, t) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} d\sigma\, g(\sigma) \exp(i\sigma t), \tag{27}$$

and to consider only interactions with the initial condition

$$\lim_{t_0 \to \infty} H'(\mathbf{x}, -t_0) = 0 \tag{28}$$

in accordance with the conditions established in Eq. (14). If

$$\lim_{\sigma \to \pm\infty} g(\sigma) = 0, \tag{29}$$

then by Jordan's lemma,[7] Eq. (27) can be replaced by a contour integration along the real axis, closing

[7] E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis* (Cambridge University Press, Cambridge, England, 1952), 4th ed., p. 115.

the contour in the upper or lower half-plane depending upon whether t is greater or less then zero. Thus

$$\begin{aligned} H'(\mathbf{x}, t) &= (2\pi)^{-\frac{1}{2}} \int_{C_l} d\sigma\, g(\sigma) \exp(i\sigma t), \qquad t < 0; \\ &= (2\pi)^{-\frac{1}{2}} \int_{C_u} d\sigma\, g(\sigma) \exp(i\sigma t), \qquad t > 0. \end{aligned} \tag{30}$$

If it is assumed that $g(\sigma)$ is analytic[8] in the neighborhood of the contour, then the condition given in Eq. (28) is satisfied since any singularities within the contour lead to terms which damp exponentially for large negative values of t.

With the substitutions defined by Eq. (27) and Eq. (30), all the time integrations in Eq. (26) can be done. For example, consider the second-order term

$$\begin{aligned} U_{k0}^{(2)}(t, -\infty) = \lim_{t_0\to\infty} (i\hbar)^{-2} \int_{-t_0}^{t} d\xi_1 \langle u_k|\, H'(\xi_1)\, |u_l\rangle \exp(i\omega_{kl}\xi_1) \\ \times \int_{-t_0}^{\xi_1} d\xi_2 \langle u_l|\, H'(\xi_2)\, |u_0\rangle \exp(i\omega_{l0}\xi_2) = \lim_{t_0\to\infty} (i\hbar)^{-2}(2\pi)^{-1} \int_{-t_0}^{t} d\xi_1 \int_{-t_0}^{\xi_1} d\xi_2 \\ \times \int_{-\infty}^{\infty} d\sigma_1 \langle u_k|\, g(\sigma_1)\, |u_l\rangle \exp[i(\sigma_1 + \omega_{kl})\xi_1] \int_{-\infty}^{\infty} d\sigma_2 \langle u_l|\, g(\sigma_2)\, |u_0\rangle \exp[i(\sigma_2 + \omega_{l0})\xi_2]. \end{aligned} \tag{31}$$

Performing the integration over ξ_2 gives rise to the integral

$$\begin{aligned} I &= \lim_{t_0\to\infty} \int_{-\infty}^{\infty} d\sigma_2 \langle u_l|\, g(\sigma_2)\, |u_0\rangle \int_{-t_0}^{\xi_1} d\xi_2 \exp[i(\sigma_2 + \omega_{l0})\xi_2] \\ &= \lim_{t_0\to\infty} \int_{-\infty}^{\infty} d\sigma_2 \frac{\langle u_l|\, g(\sigma_2)\, |u_0\rangle}{i(\sigma_2 + \omega_{l0})} \{\exp[i(\sigma_2 + \omega_{l0})\xi_1] - \exp[-i(\sigma_2 + \omega_{l0})t_0]\}. \end{aligned} \tag{32}$$

Since the integrand is analytic for $\sigma_2 = -\omega_{l0}$, σ_2 can be replaced by $\sigma_2 - i\epsilon_2$, where ϵ_2 is a small, positive displacement. Therefore

$$\begin{aligned} I = \lim_{\epsilon_2\to 0} \int_{-\infty}^{\infty} d\sigma_2 \frac{\langle u_l|\, g(\sigma_2 - i\epsilon_2)\, |u_0\rangle}{i(\sigma_2 + \omega_{l0} - i\epsilon_2)} \exp[i(\sigma_2 + \omega_{l0} - i\epsilon_2)\xi_1] \\ - \lim_{\epsilon_2\to 0} \lim_{t_0\to\infty} \int_{C_l} \frac{d\sigma_2 \langle u_l|\, g(\sigma_2 - i\epsilon_2)\, |u_0\rangle}{i(\sigma_2 + \omega_{l0} - i\epsilon_2)} \exp[-i(\sigma_2 - i\epsilon_2 + \omega_{l0})t_0]. \end{aligned} \tag{33}$$

Since the integrand is analytic in the neighborhood of the contour, all singularities inside the contour give rise to terms which damp exponentially for large values of t_0. Therefore,

$$I = \lim_{\epsilon_2\to 0} \int_{-\infty}^{\infty} d\sigma_2 \frac{\langle u_l|\, g(\sigma_2 - i\epsilon_2)\, |u_0\rangle}{i(\sigma_2 + \omega_{l0} - i\epsilon_2)} \exp[i(\sigma_2 + \omega_{l0} - i\epsilon_2)\xi_1]. \tag{34}$$

The integration over ξ_1 can now be performed. This gives

$$\begin{aligned} U_{k0}^{(2)} = (-\hbar)^{-2}(2\pi)^{-1} \lim_{t_0\to\infty} \int_{-\infty}^{\infty} d\sigma_1 \langle u_k|\, g(\sigma_1)\, |u_l\rangle \lim_{\epsilon_2\to 0} \int_{-\infty}^{\infty} d\sigma_2 \frac{\langle u_l|\, g(\sigma_2 - i\epsilon_2)\, |u_0\rangle}{(\sigma_2 + \omega_{l0} - i\epsilon_2)} \\ \times \{\exp[i(\sigma_1 + \sigma_2 + \omega_{k0} - i\epsilon_2)t] - \exp[-i(\sigma_1 + \sigma_2 + \omega_{k0} - i\epsilon_2)t_0]\}(\sigma_1 + \sigma_2 + \omega_{k0} - i\epsilon_2)^{-1}, \end{aligned} \tag{35}$$

since

$$\omega_{k0} = \omega_{kl} + \omega_{l0}.$$

The integrand has no singularities at $\sigma_1 = -\sigma_2 + i\epsilon_2 - \omega_{k0}$. Therefore, replacing σ_1 by $\sigma_1 - i\epsilon_1$, one arrives, in a manner parallel to that used for Eq. (34), at the result

$$\begin{aligned} U_{k0}^{(2)} = (-\hbar)^{-2}(2\pi)^{-1} \lim_{\epsilon_1\to 0} \int_{-\infty}^{\infty} d\sigma_1 \lim_{\epsilon_2\to 0} \int_{-\infty}^{\infty} d\sigma_2 \\ \times \frac{\langle u_k|\, g(\sigma_1 - i\epsilon_1)\, |u_l\rangle\langle u_l|\, g(\sigma_2 - i\epsilon_2)\, |u_2\rangle}{(\sigma_1 + \sigma_2 + \omega_{k0} - i\epsilon_1 - i\epsilon_2)(\sigma_2 + \omega_{l0} - i\epsilon_2)} \exp[i(\sigma_1 + \sigma_2 + \omega_{k0} - i\epsilon_1 - i\epsilon_2)t]. \end{aligned} \tag{36}$$

[8] Although this assumption is reasonable for Fourier transforms of physical quantities, it is not necessary since it can be shown that for $|t| \to \infty$, $H'(\mathbf{x}, t) \to t^{\delta-1}$; $\delta < 1$, for singularities on the real axis of the form $\omega^{-\delta}$, $\omega^{-\delta}\log|\omega|$, $|\omega|^{-\delta}$ and $|\omega|^{-\delta}\log|\omega|$.

The introduction of ϵ_1 into the integrand of Eq. (35) is not necessary but it permits one to carry out the σ_1 and σ_2 integrations without regard to order. The general term as inferred from Eq. (36) is

$$U_{k0}^{(n)}(t, -\infty) = (-1)^n(2\pi)^{-\frac{1}{2}n}\hbar^{-n} \lim_{\epsilon_1\to 0}\int_{-\infty}^{\infty} d\sigma_1 \lim_{\epsilon_2\to 0}\int_{-\infty}^{\infty} d\sigma_2 \cdots \lim_{\epsilon_n\to 0}\int_{-\infty}^{\infty} d\sigma_n \exp\{it[{}^{1n}\Omega_{k0}]\}$$
$$\times\left[\frac{\langle u_k|\, g(\sigma_1 - i\epsilon_1)\,|u_l\rangle\langle u_l|\, g(\sigma_2 - i\epsilon_2)\,|u_m\rangle \cdots \langle u_p|\, g(\sigma_n - i\epsilon_n)\,|u_0\rangle}{{}^{1n}\Omega_{k0}\,{}^{2n}\Omega_{l0}\cdots {}^{nn}\Omega_{p_0}}\right], \tag{37}$$

where

$${}^{in}\Omega_{k0} = \omega_{k0} + \sum_{m=i}^{n}(\sigma_m - i\epsilon_m).$$

APPLICATIONS

The adiabatic and sudden approximations can be treated by one formalism using

$$H'(\mathbf{x}, t) = V(x)e^{-\alpha t}/(1 + e^{-\beta t}), \text{ with } \beta \text{ real, } \alpha \text{ complex and } \beta > \text{real } \alpha > 0. \tag{38}$$

When α is complex it is understood that $H^{\dagger}(\mathbf{x}, t)$ is to be added to Eq. (38) so that the total interaction is Hermitian. The Fourier transform of $H'(\mathbf{x}, t)$ given by Eq. (38) is

$$g(\sigma) = \left(\frac{1}{2\pi}\right)^{\frac{1}{2}}\frac{\pi}{\beta}\frac{V(\mathbf{x})}{\sin[(\pi/\beta)(\alpha + i\sigma)]}. \tag{39}$$

The function $g(\sigma)$ has singularities for $\alpha + i\sigma = m\beta$ where $m = 0, \pm 1, \pm 2, \cdots$. Equation (39) can be be rewritten as

$$g(\sigma) = (-1)^m(2\pi)^{-\frac{1}{2}}(\pi/\beta)\frac{V(\mathbf{x})}{\sin[(\pi/\beta)(\alpha + i\sigma + \beta m)]}. \tag{40}$$

Thus, in the neighborhood of the singularities, $g(\sigma)$ has the power series expansion

$$g(\sigma) = \frac{(-1)^m(2\pi)^{-\frac{1}{2}}V(\mathbf{x})}{(\alpha + i\sigma + \beta m)}\{1 + O[(\alpha + i\sigma + \beta m)^2]\}; \qquad m = 0, \pm 1, \pm 2, \cdots \tag{41}$$

so that $g(\sigma)$ has only simple poles.

The first three terms of the U matrix for $t > 0$ are given by Eq. (37) as

$$U_{k0}^{(0)}(t, -\infty) = \delta_{k0};$$

$$U_{k0}^{(1)}(t, -\infty) = -\hbar^{-1}(2\pi)^{-\frac{1}{2}} \lim_{\epsilon_1\to 0}\int_{C_u} d\sigma_1 \frac{\langle u_k|\, g(\sigma_1 - i\epsilon_1)\,|u_0\rangle}{(\sigma_1 + \omega_{k0} - i\epsilon_1)} \exp[i(\sigma_1 + \omega_{k0} - i\epsilon_1)t],$$

$$U_{k0}^{(2)}(t, -\infty) = \hbar^{-2}(2\pi)^{-1} \lim_{\epsilon_1\to 0}\int_{C_u} d\sigma_1 \lim_{\epsilon_2\to 0}\int_{C_u} d\sigma_2$$
$$\times \frac{\langle u_k|\, g(\sigma_1 - i\epsilon_1)\,|u_l\rangle\langle u_l|\, g(\sigma_2 - i\epsilon_2)\,|u_0\rangle}{(\sigma_1 + \sigma_2 + \omega_{k0} - i\epsilon_1 - i\epsilon_2)(\sigma_2 + \omega_{l0} - i\epsilon_2)} \exp[i(\sigma_1 + \sigma_2 + \omega_{k0} - i\epsilon_1 - i\epsilon_2)t]. \tag{42}$$

With the substitution of $g(\sigma)$ from Eq. (39) into Eqs. (42), one obtains, using the residue theorem,

$$U_{k0}^{(1)}(t, -\infty) = (i\hbar)^{-1}V_{k0}\left\{(\pi/\beta)\frac{1}{\sin[(\pi/\beta)(\alpha - i\omega_{k0})]} - \sum_{m=0}^{\infty}(-1)^m\frac{\exp[-(\alpha + \beta m - i\omega_{k0})t]}{\alpha + \beta m - i\omega_{k0}}\right\};$$

$$U_{k0}^{(2)}(t, -\infty) = (i\hbar)^{-2}V_{kl}V_{l0}\left\{\frac{\pi/\beta}{\sin[(\pi/\beta)(2\alpha - i\omega_{k0})]}\left[\sum_{m=0}^{\infty}\frac{1}{(\alpha + \beta m - i\omega_{kl})} + \sum_{m=1}^{\infty}\frac{1}{(\alpha - \beta m - i\omega_{l0})}\right]\right.$$
$$- \frac{(\pi/\beta)}{\sin[(\pi/\beta)(\alpha - i\omega_{l0})]}\sum_{m=0}^{\infty}\frac{(-1)^m\exp[-(\alpha + \beta m - i\omega_{kl})t]}{(\alpha + \beta m - i\omega_{kl})}$$
$$\left. + \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{(-1)^m(-1)^n\exp[-(2\alpha + \beta m + \beta n - i\omega_{k0})t]}{(\alpha + \beta n - i\omega_{l0})(2\alpha + \beta m + \beta n - i\omega_{k0})}\right\}. \tag{43}$$

Sudden Approximation

To obtain the sudden approximation one takes $\alpha = 0$ and $\beta = \infty$ so that Eq. (38) can be written as

$$H'(\mathbf{x}, t) = \begin{cases} 0; & t < 0 \\ V(\mathbf{x}); & t > 0, \end{cases} \tag{44}$$

which is a step function at the origin. Note that this is the interaction commonly used in deriving the "golden rule" from time-dependent perturbation theory.[2] In this derivation one considers $H'(\mathbf{x}, t)$ to be independent of time except for being turned on at $t = 0$. Setting $\alpha = 0$ and taking the limit $\beta \to \infty$, one obtains from Eq. (43)

$$U_{k0}^{(1)}(t, -\infty) = -(i\hbar)^{-1} V_{k0} \frac{1 - e^{i\omega_{k0}t}}{(i\omega_{k0})}$$

$$U_{k0}^{(2)}(t, -\infty) = (i\hbar)^{-2} V_{kl} V_{l0} \left\{ \frac{1}{(i\omega_{k0})(i\omega_{kl})} + \frac{e^{i\omega_{k0}t}}{(i\omega_{l0})(i\omega_{k0})} - \frac{e^{i\omega_{kl}t}}{(i\omega_{l0})(i\omega_{kl})} \right\}; \qquad t > 0, \tag{45}$$

where $V_{kl} = \langle u_k | V(\mathbf{x}) | u_l \rangle$. Rewriting Eqs. (45), taking into account the cases where the ω's are zero, one obtains

$$U_{00}^{(1)} = (i\hbar)^{-1} V_{00} t,$$

$$U_{00}^{(2)} = (i\hbar)^{-2} \left\{ (V_{00})^2 (t^2/2) + \sum_{l \neq 0} V_{0l} V_{l0} \left[\frac{t}{i\omega_{l0}} - \frac{1}{(i\omega_{l0})^2} (1 - e^{-i\omega_{l0}t}) \right] \right\}, \tag{46}$$

and for $k \neq 0$

$$U_{k0}^{(1)} = -(i\hbar)^{-1} V_{k0} \frac{1 - e^{i\omega_{k0}t}}{(i\omega_{k0})},$$

$$U_{k0}^{(2)} = (i\hbar)^{-2} \left\{ V_{k0} V_{00} \left[\frac{1}{(i\omega_{k0})^2} (1 - e^{i\omega_{k0}t}) + \frac{t e^{i\omega_{k0}t}}{(i\omega_{k0})} \right] + V_{kk} V_{k0} \left[\frac{1}{(i\omega_{k0})^2} (e^{i\omega_{k0}t} - 1) - \frac{t}{i\omega_{k0}} \right] \right.$$

$$\left. + \sum_{l \neq k, 0} V_{kl} V_{l0} \left[\frac{1}{(i\omega_{k0})(i\omega_{kl})} + \frac{e^{i\omega_{k0}t}}{(i\omega_{l0})(i\omega_{k0})} - \frac{e^{i\omega_{kl}t}}{(i\omega_{l0})(i\omega_{kl})} \right] \right\}. \tag{47}$$

In the Schrödinger picture, Eq. (19) is

$$\psi_S(\mathbf{x}, t) = \sum_k a_k(t) u_k(\mathbf{x}) \exp(-i\hbar^{-1} E_k t) = \sum_k u_k(\mathbf{x}) U_{k0}(t, -\infty) \exp(-i\hbar^{-1} E_k t). \tag{48}$$

Substituting for the first three terms of $U_{k0}(t, -\infty)$, using Eqs. (46) and (47), gives through second order

$$\psi_S(\mathbf{x}, t) = v_0(\mathbf{x}) \exp(-i\hbar^{-1} W_0 t) \left(1 - \frac{1}{2} \sum_{l \neq 0} \frac{V_{0l} V_{l0}}{(E_0 - E_l)^2} \right)$$

$$+ \sum_{k \neq 0} v_k(\mathbf{x}) \exp(-i\hbar^{-1} W_k t) \left\{ \frac{V_{k0}}{E_k - E_0} - \frac{V_{kk} V_{k0}}{(E_k - E_0)^2} + \sum_{l \neq k} \frac{V_{kl} V_{l0}}{(E_k - E_0)(E_k - E_l)} \right\}; \qquad t > 0, \tag{49}$$

where

$$v_0(\mathbf{x}) = u_0(\mathbf{x}) + \sum_{k \neq 0} \frac{u_k(\mathbf{x}) V_{k0}}{E_0 - E_k}$$

$$+ \sum_{k \neq 0} u_k(\mathbf{x}) \left\{ \sum_{l \neq 0} \frac{V_{kl} V_{l0}}{(E_l - E_0)(E_k - E_0)} - \frac{V_{k0} V_{00}}{(E_k - E_0)^2} \right\} - \tfrac{1}{2} u_0(\mathbf{x}) \sum_{l \neq 0} \frac{V_{0l} V_{l0}}{(E_0 - E_l)^2}, \tag{50a}$$

$$v_k(\mathbf{x}) = u_k(\mathbf{x}) + \sum_{l \neq k} \frac{u_l(\mathbf{x}) V_{lk}}{E_k - E_l}, \tag{50b}$$

$$W_0 = E_0 + V_{00} + \sum_{l \neq 0} \frac{V_{0l} V_{l0}}{E_0 - E_l}, \tag{50c}$$

and

$$W_k = E_k + V_{kk}. \tag{50d}$$

These results are precisely what one expects from the "sudden" approximation (5). Here the functions $v_k(\mathbf{x})$ are solutions of the stationary-state problem

$$[H_0 + V(\mathbf{x})]v_k(\mathbf{x}) = W_k v_k(\mathbf{x}). \tag{51}$$

Since $V(\mathbf{x})$ is presumed small $v_k(\mathbf{x})$ can be expressed in terms of $u_k(\mathbf{x})$ using stationary state perturbation theory[9]. This procedure gives Eq. (50). The coefficients of $v_k(\mathbf{x})$ expressed in Eq. (49) are given by $\langle v_k(\mathbf{x}) \mid u_0(\mathbf{x})\rangle$. These are the same as those obtained from the "sudden" approximation[5] when the initial condition $a_n = \delta_{n0}$ is used.

As a result of Eqs. (49) and (50), it is seen that the states $v_k(\mathbf{x})$ are populated according to a time-independent distribution for $t > 0$. Thus all transitions from the state $u_0(\mathbf{x})$ to the states $v_k(\mathbf{x})$ have taken place at the time $t = 0$. The transition rate vanishes for $t > 0$.

If one calculates the probability for a transition from the state $u_0(\mathbf{x})$ to the state $u_k(\mathbf{x})$ ignoring the fact that the physical states are given by $v_k(\mathbf{x})$ for $t > 0$, a non-vanishing transition probability $w(t)$ is obtained. For this case the transition probability is given by Eq. (24). Considering only the first order term $U_{k0}^{(1)}$ given in Eq. (47) one obtains

$$w(t) = \frac{|V_{k0}|^2}{\hbar^2}\frac{d}{dt}\frac{[\sin(\omega_{k0}t/2)]^2}{(\omega_{k0}/2)^2} = \frac{2\pi}{\hbar^2}|V_{k0}|^2\frac{\sin\omega_{k0}t}{\pi\omega_{k0}}. \tag{52}$$

For t large compared to $(\omega_{k0})^{-1}$, $(\pi\omega_{k0})^{-1}\sin\omega_{k0}t$ is replaced by the delta function $\delta(\omega_{k0})$. Hence one obtains the "golden rule"

$$w = (2\pi/\hbar)\,|V_{k0}|^2\,\delta(E_k - E_0). \tag{53}$$

It is therefore seen that this derivation of the "golden rule" is not physically meaningful since the states $u_k(\mathbf{x})$ do not exist for $t > 0$.

Adiabatic Approximation

For the adiabatic approximation one takes the limit as β approaches zero and sets $\alpha = 0$ so that Eq. (38) can be written as

$$H'(\mathbf{x}, t) = V(\mathbf{x})/(1 + e^{-\beta t}).$$

Thus one obtains from Eq. (43)

$$\begin{aligned} U_{00}^{(1)}(t, -\infty) &= (i\hbar)^{-1}V_{00}\left\{t - \sum_{m=1}^{\infty}(-1)^m\frac{e^{-\beta m t}}{\beta m}\right\}, \\ &= (i\hbar)^{-1}V_{00}\{t + (1/\beta)\log(1 + e^{-\beta t})\}, \\ &= (i\hbar)^{-1}V_{00}\int_{-\infty}^{t}\frac{d\xi}{1 + e^{-\beta\xi}}; \end{aligned} \tag{54a}$$

$$\begin{aligned} U_{00}^{(2)}(t, -\infty) = (i\hbar)^{-2}(V_{00})^2&\left\{\frac{t^2}{2} - t\sum_{m=1}^{\infty}\frac{(-1)^m e^{-\beta m t}}{\beta m} + \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\frac{(-1)^n(-1)^m e^{-\beta(m+n)t}}{\beta n(\beta m + \beta n)}\right\} \\ &+ (i\hbar)^{-2}\sum_{l\neq 0}V_{0l}V_{l0}\left\{\frac{t}{i\omega_{l0}} + \frac{1}{(\omega_{l0})^2} - \sum_{m=1}^{\infty}\frac{1}{(\beta m + i\omega_{l0})^2}\right. \\ &\left. - \frac{1}{i\omega_{l0}}\sum_{m=1}^{\infty}\frac{(-1)^m e^{-\beta m t}}{\beta m} + \sum_{n=1}^{\infty}\frac{(-1)^n e^{-\beta n t}}{(\beta n)(\beta n - i\omega_{l0})} + \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{(-1)^{m+n}e^{-\beta(m+n)t}}{\beta(m+n)(\beta n - i\omega_{l0})}\right\}. \end{aligned} \tag{54b}$$

The sums appearing in Eq. (54b) can be done, giving

$$\begin{aligned} U_{00}^{(2)}(t, -\infty) = (i\hbar)^{-2}(V_{00})^2&\left\{\frac{t^2}{2} + \frac{t}{\beta}\log(1 + e^{-\beta t}) + \frac{1}{2\beta^2}\log^2(1 + e^{-\beta t})\right\} \\ &+ (i\hbar)^{-2}\sum_{l\neq 0}V_{0l}V_{l0}\left\{\frac{t}{i\omega_{l0}} + \frac{1}{i\omega_{l0}\beta}\log(1 + e^{-\beta t})\right. \\ &\left. - \frac{1}{i\omega_{l0}\beta}\frac{1}{(1 + e^{-\beta t})} + \frac{1}{2(\omega_{l0})^2}\left[\left(\frac{e^{-\beta t}}{1 + e^{-\beta t}}\right)^2 - \frac{2e^{-\beta t}}{1 + e^{-\beta t}} + 1\right]\right\} \end{aligned}$$

[9] Reference 1(a), p. 151.

$$= (i\hbar)^{-2}\frac{(V_{00})^2}{2}\left[\int_{-\infty}^{t}\frac{d\xi}{(1+e^{-\beta\xi})}\right]^2 + (i\hbar)^{-2}\sum_{l\neq 0}\frac{V_{0l}V_{l0}}{i\omega_{l0}}\int_{-\infty}^{t}\frac{d\xi}{(1+e^{-\beta\xi})^2}$$

$$+\frac{(i\hbar)^{-2}}{2}\sum_{l\neq 0}\frac{V_{0l}V_{l0}}{(\omega_{l0})^2}\frac{1}{(1+e^{-\beta t})^2}. \tag{54c}$$

For $k \neq 0$,

$$U_{k0}^{(1)} = (i\hbar)^{-1}\frac{V_{k0}}{i\omega_{k0}}\frac{e^{i\omega_{k0}t}}{1+e^{-\beta t}}, \tag{55a}$$

$$U_{k0}^{(2)} = -(i\hbar)^{-2}\sum_{l\neq 0}\frac{V_{kl}V_{l0}}{\omega_{k0}\omega_{l0}}\frac{e^{i\omega_{k0}t}}{(1+e^{-\beta t})^2}$$

$$+\frac{(i\hbar)^{-2}}{i\omega_{k0}}\frac{V_{k0}V_{00}}{(1+e^{-\beta t})}e^{i\omega_{k0}t}\int_{-\infty}^{t}\frac{d\xi}{(1+e^{-\beta\xi})}+\frac{(i\hbar)^{-2}}{(\omega_{k0})^2}\frac{V_{k0}V_{00}e^{i\omega_{k0}t}}{(1+e^{-\beta t})^2}\left(1+e^{-\beta t}\int_{-\infty}^{\beta t}\frac{dy}{1+e^{-y}}\right). \tag{55b}$$

Substitution into Eq. (48) for $U_{k0}(t, -\infty)$ from Eqs. (54) and (55) gives through second order

$$\psi_S(\mathbf{x}, t) = \chi(\mathbf{x}, t)\exp\left(-\frac{i}{\hbar}\left\{E_0 t + \int_{-\infty}^{t} d\xi\,[W(\xi) - E_0]\right\}\right) \tag{56}$$

where

$$\chi(\mathbf{x}, t) = \left[1 - \tfrac{1}{2}\sum_{l\neq 0}\frac{|V_{l0}(t)|^2}{(E_0 - E_l)^2}\right]u_0 + \sum_{k\neq 0}\frac{u_k V_{k0}(t)}{E_0 - E_k}$$

$$+\sum_{k\neq 0}u_k\left\{\sum_{l\neq 0}\frac{V_{kl}(t)V_{l0}(t)}{(E_0 - E_k)(E_0 - E_l)} - \frac{V_{k0}(t)V_{00}(t)}{(E_0 - E_k)^2} + \frac{1}{(E_0 - E_k)^2}\frac{\partial V_{k0}(t)}{\partial t}\int_{-\infty}^{t} d\xi\, V_{00}(\xi)\right\}, \tag{57a}$$

$$W(\xi) = E_0 + V_{00}(\xi) + \sum_{l\neq 0}\frac{|V_{0l}(\xi)|^2}{E_0 - E_l}, \tag{57b}$$

and

$$V_{kl}(t) = V_{kl}(1+e^{-\beta t})^{-1}, \tag{57c}$$

$$\frac{\partial V_{k0}(t)}{\partial t}\int_{-\infty}^{t} d\xi\, V_{00}(\xi) = -V_{k0}(t)V_{00}(t)e^{-\beta t}\int_{-\infty}^{\beta t} dy\,(1+e^{-y})^{-1}. \tag{57d}$$

In the limit as β approaches zero, the integral in the exponent of Eq. (56) becomes infinite. It is clear that this infinity must always occur as a phase factor for all orders of the perturbation expansion since it was shown, even for arbitrarily small but finite β, that the series for the U matrix converges and that $U^\dagger U = 1$.

Except for the term involving $(\partial V_{k0}(t)/\partial t)$ in Eq. (57a), which approaches zero exponentially [see Eq. (57d)] for large βt, the results expressed by Eqs. (56) and (57) are identical to the results obtained with stationary-state perturbation theory.[9] If βt is not large the form of $\psi_S(\mathbf{x}, t)$ differs in second order from the stationary-state result.

Equations (56) and (57) apply equally well for an arbitrary potential of the form $V(t) = V'(\mathbf{x}, \beta t)$, which approaches zero for large negative times and where β is less than any of the natural frequencies of the system, with the exception, as shown below, of the second term of Eq. (57a) which arises from the third term of Eq. (54c). This can be shown by considering, as an example, the derivation of the $U_{00}^{(2)}$ term directly from Eq. (26) replacing $H'(t)$ by $V'(\mathbf{x}, \beta t)$. One has

$$U_{00}^{(2)} = (i\hbar)^{-2}\int_{-\infty}^{t} d\xi_1 \exp(i\omega_{0l}\xi_1)V'_{0l}(\beta\xi_1)\int_{-\infty}^{\xi_1} d\xi_2 \exp(i\omega_{l0}\xi_2)V'_{l0}(\beta\xi_2),$$

where

$$\langle k|\, V'(\mathbf{x}, \beta t)\,|l\rangle = V'_{kl}(\beta t) = V_{kl}(t).$$

The substitution $\beta\xi_2 = u + \beta\xi_1$ allows $U_{00}^{(2)}$ to be written as

$$U_{00}^{(2)} = (i\hbar)^{-2}(1/\beta)\int_{-\infty}^{t} d\xi_1\, V'_{0l}(\beta\xi_1)\int_{-\infty}^{0} du \exp(i\omega_{l0}\beta^{-1}u)V'_{l0}(u + \beta\xi_1).$$

If $l \neq 0$, the integration path for the u integration can be rotated into the upper half complex u-plane, encircling poles and indenting around branch points where necessary, until it lies along the imaginary axis. Since these singularities lie in the upper half of the complex u-plane, they will contribute terms which damp exponentially for large (ω_{l0}/β). Thus the u integral is converted to a Laplace integral. An asymptotic approximation can be obtained for large (ω_{l0}/β) (small β) by expanding $V'_{l0}(u + \beta\xi_1)$ about $u = 0$ and integrating term by term. If two orders are kept, one has, to order β,

$$U_{00}^{(2)} = (i\hbar)^{-2}\left\{\frac{1}{i\omega_{l0}}\int_{-\infty}^{t} d\xi_1\ V'_{0l}(\beta\xi_1)V'_{l0}(\beta\xi_1) + \frac{\beta}{(\omega_{l0})^2}\int_{-\infty}^{t} d\xi_1\ V'_{0l}(\beta\xi_1)\frac{\partial V'_{l0}(\beta\xi_1)}{\partial(\beta\xi_1)} + O(\beta)\right\}, \qquad l \neq 0.$$

Provided the time dependence of V'_{0l} is the same as V'_{l0} ,which is the case for all potentials of the form $V(\mathbf{x})f(\beta t)$, the last term can be integrated giving the result

$$U_{00}^{(2)} = (i\hbar)^{-2}\left\{\frac{1}{i\omega_{l0}}\int_{-\infty}^{t} d\xi_1\ V_{0l}(\xi_1)V_{l0}(\xi_1) + \frac{1}{2}\frac{1}{(\omega_{l0})^2}\ V_{0l}(t)V_{l0}(t) + O(\beta)\right\}, \qquad l \neq 0,$$

in agreement with the second and third terms of Eq. (54c). All others terms leading to Eqs. (56) and (57) can be calculated in a similar fashion.

Transient Perturbation

For the transient perturbation one assumes that $H'(\mathbf{x}, \pm\infty)$ is zero. Thus the probability for finding the system in the kth state is determined by the S-matrix $U(\infty, -\infty)$ for which the nth term is from Eq. (37)

$$U_{k0}^{(n)}(\infty, -\infty) = \lim_{t\to\infty} U_{k0}^{(n)}. \tag{58}$$

If one closes the contour integration for σ_1 in the upper half complex plane and assumes that $g(\sigma_1 - i\epsilon_1)$ has no singularities on the real axis, only the pole due to ${}^{1n}\Omega_{k0}$ contributes. Contributions from the singularities in the upper half complex plane have positive imaginary parts which give rise to an exponential damping factor which goes to zero as t approaches infinity.

Thus one obtains

$$U_{k0}^{(n)}(\infty, -\infty) = (-1)^n(2\pi)^{1-n/2}i\hbar^{-n}\lim_{\epsilon_2\to 0}\int d\sigma_2 \cdots$$

$$\times \lim_{\epsilon_n\to 0}\int d\sigma_n \frac{\langle u_k|\ g(\omega_{1k} - {}^{2n}\Omega_{10})\ |u_l\rangle\langle u_l|\ g(\sigma_2 - i\epsilon_2)\ |u_m\rangle \cdots \langle u_p|\ g(\sigma_n - i\epsilon_n)\ |u_0\rangle}{{}^{2n}\Omega_{l0} \cdots {}^{nn}\Omega_{p0}}. \tag{59}$$

The transition probability through second order is

$$|a_k(\infty)|^2 = \frac{2\pi}{\hbar^2}\left|\langle u_k|\ g(\omega_{0k})\ |u_0\rangle + \lim_{\epsilon_2\to 0}\frac{(2\pi)^{-\frac{1}{2}}}{\hbar}\int_{-\infty}^{\infty} d\sigma_2 \frac{\langle u_k|\ g(\omega_{0k} - \sigma_2 + i\epsilon_2)\ |u_l\rangle\langle u_l|\ g(\sigma_2 - i\epsilon_2)\ |u_0\rangle}{\omega_{0l} - \sigma_2 + i\epsilon_2}\right|^2. \tag{60}$$

If one assumes that $E_0 > E_k$, the above equation corresponds, in first order, to the emission of a quantum of energy ω_{0k}. The second-order term for $\sigma_2 > 0$ corresponds to the emission of a quantum of energy σ_2, leaving the system in the intermediate state u_l, followed by emission $(\omega_{0k} - \sigma_2 > 0)$ or absorption $(\omega_{0k} - \sigma_2 < 0)$ of a quantum leaving the system in the final state u_k.

For this case the number of transitions per unit time is

$$w = N\ |a_k(\infty)|^2, \tag{61}$$

where N is the number of transients per unit time.

Continuous Case

The continuous case differs from those previously discussed in that the asymptotic scattering boundary condition must be imposed on the wavefunction. That is, in the Schrödinger picture, one must have (for scalar wavefunctions)

$$\lim_{|x|\to\infty} \psi_S(\mathbf{x}, t) \to \exp\left[i\hbar^{-1}(\mathbf{p}_0\cdot\mathbf{x} - E_0t)\right] + \frac{f(\theta, \varphi)}{|\mathbf{x}|}\exp\left[i\hbar^{-1}(p\ |\mathbf{x}| - Et)\right], \tag{62}$$

where $\mathbf{p}_0$, E_0, and $\mathbf{p}$, E are, respectively, the momentum and energy of the incident plane wave and outgoing spherical wave. Note that the polar angle θ is defined by

$$\hat{p}_0 \cdot \hat{p} = \cos\theta, \tag{63}$$

and the azimuthal angle φ is measured in the plane perpendicular to $\mathbf{p}_0$. It has also been assumed for simplicity that the scattering center is fixed and the scattering takes place in the vicinity of $\mathbf{x} = 0$. The scattering cross section per unit solid angle is related to the amplitude $f(\theta, \varphi)$ simply by[10]

$$d\sigma(\theta, \varphi)/d\Omega = (v/v_0)\ |f(\theta, \varphi)|^2, \tag{64}$$

where v and v_0 are the velocities of the outgoing and incident waves.

Since only the continuum states are being discussed, the simplest way to impose the asymptotic condition given in Eq. (62) is to choose

$$u_\nu(\mathbf{x}) = (2\pi\hbar)^{-\frac{3}{2}} \exp(i\hbar^{-1}\mathbf{p}\cdot\mathbf{x}), \tag{65}$$

so that Eq. (48) becomes

$$\psi_S(\mathbf{x}, t) = (2\pi\hbar)^{-\frac{3}{2}} \int d\mathbf{q} \exp\{i\hbar^{-1}[\mathbf{q}\cdot\mathbf{x} - E(q)t]\} U_{a\nu_0}(t, -\infty) = (2\pi\hbar)^{-\frac{3}{2}} \exp[i\hbar^{-1}(\mathbf{p}_0\cdot\mathbf{x} - E_0 t)]$$

$$+ (2\pi\hbar)^{-\frac{3}{2}} \int d\mathbf{q} \int d\mathbf{y} \exp\{i\hbar^{-1}[\mathbf{q}\cdot(\mathbf{x} - \mathbf{y}) - Et]\}[U(t, -\infty) - 1] \exp(i\hbar^{-1}\mathbf{p}_0\cdot\mathbf{y}). \tag{66}$$

The integration over the momentum directions can be done by choosing

$$\mathbf{q}\cdot(\mathbf{x} - \mathbf{y}) = q\ |\mathbf{x} - \mathbf{y}| \cos\theta'. \tag{67}$$

Thus, Eq. (66) becomes

$$\psi_S(\mathbf{x}, t) = (2\pi\hbar)^{-\frac{3}{2}} \exp[i\hbar^{-1}(\mathbf{p}_0\cdot\mathbf{x} - E_0 t)]$$

$$- i(2\pi\hbar)^{-\frac{1}{2}} \int_{-\infty}^{\infty} q\, dq \int dy \frac{\exp(i\hbar^{-1}[q\ |\mathbf{x} - \mathbf{y}| - Et])}{|\mathbf{x} - \mathbf{y}|} [U(t, -\infty) - 1] \exp(i\hbar^{-1}\mathbf{p}_0\cdot\mathbf{y}). \tag{68}$$

If one assumes that $V(\mathbf{y})$ falls off with sufficient rapidity so that negligible error is made by integrating over some finite region rather than over all space, then in the usual way

$$\lim_{|\mathbf{x}|\to\infty} \psi_S(\mathbf{x}, t) = (2\pi\hbar)^{-\frac{3}{2}} \exp[i\hbar^{-1}(\mathbf{p}_0\cdot\mathbf{x} - E_0 t)] - i(2\pi\hbar)^{-\frac{1}{2}}\ |\mathbf{x}|^{-1}\ I_C, \tag{69}$$

where

$$I_C = \int_{-\infty}^{\infty} q\, dq \exp[i\hbar^{-1}(q\ |\mathbf{x}| - Et)]\langle u_a|\ U(t, -\infty) - 1\ |u_{\nu_0}\rangle \tag{70}$$

and

$$\mathbf{q} = q\mathbf{x}\ |\mathbf{x}|^{-1}. \tag{71}$$

The above integral cannot contain incoming spherical waves, since if it did, these waves would also be present at $t = -\infty$ for large $|\mathbf{x}|$. One sees from Eq. (69) that this would imply that $U(-\infty, -\infty) \neq 1$, contrary to the definition for $U(t, -t_0)$ given in Eq. (8). It follows, therefore, that no contributions to the integral I_C can come from singularities that may lie in the left-hand complex q-plane; that is, from singularities with Re $q < 0$. Thus the definition of $U(t, -t_0)$ coupled with the proper choice of basis eigenfunctions fixes the asymptotic boundary conditions.

It is simpler at this juncture to restrict the discussion to the nonrelativistic limit in which case $E = q^2/2m$. The integral I_C can be evaluated asymptotically for large t considering $|\mathbf{x}|/t$, the position of the detectors divided by the time of observation, to be finite. Thus, consider the contour shown in Fig. 1. The contributions from the contours C_1 and C_3 at $|q| = \infty$ are zero because of exponential damping. The contour C_2, which is the straight line

$$q_I = -q_R + iq', \tag{72}$$

[10] Reference 1(a), pp. 100–102.

where

$$q' = m\,|\mathbf{x}|/t, \qquad q = q_R + iq_I, \tag{73}$$

is uniquely determined by the requirement that the integral I_{C_2} be a Laplace transform so that it can be evaluated asymptotically using the Laplace method without further distortion of the contour. The integral along C_2 is shown in the Appendix to be of order $|\mathbf{x}|^{-\frac{1}{2}}$ and therefore does not contribute to the scattered wave. Thus, by the residue theorem

$$I_C = 2\pi i \sum_n (R_{U_n} - R_{L_n}). \tag{74}$$

This procedure is easily generalized to include branch points. These are not discussed here since the functions given by Eqs. (43) have only simple poles. Examination of Eqs. (69) and (74) shows that from each pole one obtains a scattered outgoing spherical wave of momentum q_R of the form

$$|\mathbf{x}|^{-1} \exp\{i\hbar^{-1}[q_R\,|\mathbf{x}| - q_R^2(t/2m)]\} \exp[(q_I q_R t'/\hbar m)] f_a(\theta, \varphi) \exp[i\hbar^{-1}(q_I^2 t/2m)], \tag{75}$$

where

$$t' = t - (m\,|\mathbf{x}|/q_R), \tag{76}$$

and

$$q = q_R + iq_I,$$

is the location of the pole.

The quantity t' can be rewritten in terms of q' as

$$(t'/t) = 1 - (q'/q_R). \tag{77}$$

For the poles within the contour in the lower half-plane $q_R > q'$, $q_I < 0$ and $t' > 0$, whereas for the poles within the contour in the upper half-plane, $q_R < q'$, $q_I > 0$ and $t' < 0$. Therefore, for both cases the scattered waves damp exponentially to zero unless t' remains finite as $|\mathbf{x}|$ and t become large. The time t' is to be interpreted as the retarded time, or the time at which the scattering occurred, since then q_R is the physical momentum,

$$p = q_R = m\,|\mathbf{x}|/(t - t'), \tag{78}$$

defined from Eqs. (62) and (76). Since, in general, each pole has a different q_R, a discrete momentum spectrum is obtained for the scattered wave. Only those poles that are within the contour can give rise to scattered waves that can be observed for a given $|\mathbf{x}|$ and t. Those poles outside the contour in the upper half-plane have been scattered at times $t' < 0$ with momentum p so large that at the time t they have passed the point of observation, whereas those poles in the lower half-plane not included in the contour have been scattered at times $t' > 0$ but with insufficient momentum to reach the point of observation at the time t.

Note from Eq. (77) that for $t = \infty$, if a scattered wave is to exist, q' must equal q_R. In this case at most only one pole can contribute to the scattering since for poles such that $q_R = \epsilon q'\epsilon \neq 1$, Eq. (77) becomes

$$t' = (1 - \epsilon)t. \tag{79}$$

Thus $|t'| = \infty$ for $t = \infty$ and the outgoing waves damp exponentially to zero. To obtain the complete momentum spectrum, t must therefore be taken large but not infinite. Since the S matrix is $U(t, -\infty)$ evaluated at $t = \infty$, it is clear that a formalism based upon the S matrix alone cannot take account of such a momentum spectrum.

As a particular example, Eq. (69) can be evaluated by substituting for $U(t, -\infty)$, the first-order term from Eqs. (43). This gives rise to the integral

$$I_C = 2m \int_{-\infty}^{\infty} q\, dq\, V_{qp_0} \exp(i\hbar^{-1} q\,|\mathbf{x}|)\Big\{\frac{(\pi/\beta)(-1)^l \exp[-i\hbar^{-1}(q^2t/2m)](2mi\hbar)^{-1}}{\sin[(-i\pi/2m\beta\hbar)(q^2 - 2mE_0 + 2mi\hbar(\alpha + \beta l)]} - \sum_{n=0}^{\infty}(-1)^n \frac{\exp\{-i\hbar^{-1}[E_0 - i\hbar(\alpha + \beta n)]t\}}{q^2 - 2mE_0 + 2mi\hbar(\alpha + \beta n)}\Big\}; \tag{80}$$

where l is any integer.

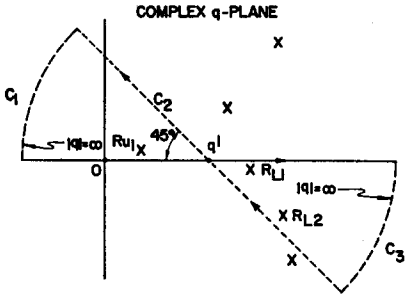


FIG. 1. Contour integration for $\psi(\mathbf{x}, t)$.

The first terms can be evaluated using the contour shown in Fig. 1 and the second term can be evaluated by completing the contour in the upper half-plane alone. Contributions from the second term therefore come from the poles

$$q = -(2m)^{\frac{1}{2}}[E_0 - i\hbar(\alpha + \beta n)]^{\frac{1}{2}} \tag{81}$$

which are in the second quadrant. This contribution is precisely canceled, as expected, by the contribution from the poles of the first term which are in the second quadrant,

$$q = -(2m)^{\frac{1}{2}}[E_0 - i\hbar(\alpha + \beta l)]^{\frac{1}{2}}, \qquad l = 0, 1, 2, \cdots. \tag{82}$$

The evaluation of I_C can thus be made by watching only those poles of the first term which lie in the first and fourth quadrant. These are

$$\begin{aligned} q_{lL} &= (2m)^{\frac{1}{2}}[E_0 - i\hbar(\beta l + \alpha)]^{\frac{1}{2}}, \qquad l = 0, 1, 2, \cdots, \\ q_{lu} &= (2m)^{\frac{1}{2}}[E_0 + i\hbar(\beta l - \alpha)]^{\frac{1}{2}}, \qquad l = 1, 2, \cdots. \end{aligned} \tag{83}$$

It should be noted that the magnitude of the phase for the lth pole is always less than 45°. The result for all q' such that all the poles in the fourth quadrant are included within the contour is

$$I_C = -2\pi i m \sum_{l=0}^{\infty} V_{qp_0}(-1)^l \exp\{i\hbar^{-1}[q_R\,|\mathbf{x}| - (q_R^2/2m)t]\} \times \exp\{-(\beta l + \alpha)[t - (m\,|\mathbf{x}|/q_R)]\} \exp[i\hbar^{-1}(q_I^2/2m)t)], \tag{84}$$

where

$$\begin{aligned} q_R &= (m)^{\frac{1}{2}}\{[E_0^2 + \hbar^2(\beta l + \alpha)^2]^{\frac{1}{2}} + E_0\}^{\frac{1}{2}}, \\ q_I &= -(m)^{\frac{1}{2}}\{[E_0^2 + \hbar^2(\beta l + \alpha)^2]^{\frac{1}{2}} - E_0\}^{\frac{1}{2}}. \end{aligned} \tag{85}$$

It has been assumed here that α is real. For α complex, the above equations apply if α is replaced by Re α and E_0 is replaced by $E_0 + \hbar$ Im α.

It should be noted here that energy is not conserved. If α is real then $(q_R^2 - q_I^2)/(2m) = E_0$, the incident energy. However the outgoing velocity and momentum are interpreted from Eqs. (78) and (84) to be q_R/m and q_R respectively. Thus the final energy is given by $q_R^2/2m$. This nonconservation of energy is due to the fact that in the nonrelativistic case, the mass is not allowed to change.

For the relativistic case $E(q) = (q^2c^2 + m^2c^4)^{\frac{1}{2}}$ where m is the mass of the incident particle. Here there are branch points at $\pm im$ which make the contour of integration more complicated. However, it is easy to see, that in the relativistic case the poles will give rise to terms of the form

$$\exp[i\hbar^{-1}(q_R\,|\mathbf{x}| - E_0 t)] \exp\{-(\beta l + \alpha)[t - (E_0/c^2 q_R)\,|\mathbf{x}|]\}.$$

Hence q_R is interpreted to be the final momentum and c^2q_R/E_0 the final velocity so that E_0 must also be the final energy. Thus energy is conserved and one obtains

$$(q_0^2c^2 + m^2c^4)^{\frac{1}{2}} = (q_R^2c^2 + M^2c^4)^{\frac{1}{2}},$$

where M is the mass of the final particle. Hence, conservation of energy in the relativistic case implies that the mass M of the final particle is different from the mass m of the initial particle. It is therefore apparent that because of the discrete momentum spectrum a discrete mass spectrum will be obtained in the relativistic case.

Regardless of the form of the potential (including time independent potentials) that is used to calculate $U(t, -\infty)$, a mass spectrum can be obtained in a manner similar to the above procedure if the potential gives rise to singularities that produce a momentum spectrum. This point is investigated in greater detail by the authors[11] in conjunction with the baryon mass spectrum.

In the following only the nonrelativistic case, for the potential given in Eq. (3), will be discussed. Two limiting cases are of interest. These are:

1. The limit $\beta \to \infty$, Re $\alpha \to 0$. Because Eq. (84) is derived only for $t' > 0$, this limit gives identical results to the "adiabatic" potential $e^{-\alpha|t|}$ used in the adiabatic hypothesis.[6] In addition it is equivalent to the step function used by Heitler.[3]

2. The limit β, Re $\alpha \to 0$. This is the adiabatic hypothesis using, instead of the "adiabatic" potential, the function given in Eq. (38).

The result of the above limiting processes for complex α are

$$I_C \to -2\pi im \exp\{i\hbar^{-1}[p\,|\mathbf{x}| - (p^2/2m)t]\}V_{pp_0} \exp[-(\mathrm{Re}\,\alpha)t'] \tag{86}$$

for

$$t' > 0, \quad \mathrm{Re}\,\alpha \to 0, \quad \beta \to \infty,$$

where,

$$p \gg \hbar\,\mathrm{Re}\,\alpha,$$

and

$$I_C \to -2\pi im \exp\{i\hbar^{-1}[p\,|\mathbf{x}| - (p^2/2m)t]\}V_{pp_0} \frac{\exp[-(\mathrm{Re}\,\alpha)t']}{1 + \exp(-\beta t')}, \tag{87}$$

for

$$t' > 0, \quad \mathrm{Re}\,\alpha \to 0, \quad \beta \to 0,$$

where, for both equations,

$$p = [2m(E_0 + \hbar\,\mathrm{Im}\,\alpha)]^{\frac{1}{2}}.$$

It is assumed, in the derivation of Eq. (87), that

$$\sum_{l=0}^{n} (-1)^l e^{-\beta l} \cong (1 + e^{-\beta t})^{-1}, \tag{88}$$

and

$$p \gg \hbar(\beta n + \alpha),$$

where n is some large positive integer.

If Im $\alpha = 0$, the scattering is elastic and results from the value of the potential at the time t' of scattering. If Im $\alpha = \pm\omega$, an absorption or emission of a quantum of energy $\hbar\omega$ occurs.

In the limit $\beta \to \infty$ with $q' >$ Re $[(2m)(E_0 - i\hbar\alpha)]^{\frac{1}{2}}$, $(t' < 0)$, the pole for $q = [(2m)(E_0 - i\hbar\alpha)]^{\frac{1}{2}}$ is not included within the contour and all other contributions damp exponentially to zero giving, for any α,

$$\lim_{\beta\to\infty} I_S = 0, \qquad t' < 0. \tag{89}$$

Thus no scattering occurs, as expected, since in this limit the interaction is zero for $t' < 0$.

[11] C. L. Hammer and T. A. Weber, Nuovo Cimento **37**, 88 (1965).

If instead of Eq. (66), the time dependence

$$\exp(\mathrm{Re}\,\alpha t)/(1+e^{\beta t}) + \exp(-\mathrm{Re}\,\alpha t)/(1+e^{-\beta t}), \tag{90}$$

is assumed, then the "adiabatic" potential is obtained in the limit $\beta \to \infty$. This gives rise to an additional pole in the upper half-plane at $q = [(2m)(E_0 + i\hbar\alpha)]^{\frac{1}{2}}$. Since

$$\mathrm{Re}\,[E_0 + i\hbar\alpha]^{\frac{1}{2}} = \mathrm{Re}\,[E_0 - i\hbar\alpha]^{\frac{1}{2}},$$

a result for $t' < 0$ identical to Eq. (86) is obtained, except that α is replaced by $-\alpha$, giving for any t', in the limit $\beta \to \infty$ and real $\alpha \to 0$,

$$I_C \to -2\pi i m \exp\{i\hbar^{-1}[p\,|\mathbf{x}| - (p^2/2m)t]\}V_{pp_0}\exp(-\mathrm{Re}\,\alpha\,|t'|). \tag{91}$$

Similarily, if in the limit $\beta \to 0$, $q' > \mathrm{Re}\,\{(2m)[E_0 + i\hbar(\beta n - \alpha)]\}^{\frac{1}{2}}$, then essentially only the poles in the first quadrant are included in the contour giving for $t' < 0$ the same result as Eq. (87).

Comparison of Eqs. (62), (64), (69), (86), (87), and (91) gives the first-order cross section for any of the limiting cases

$$d\sigma(\theta, \varphi) = \frac{m}{p_0}\frac{2\pi}{\hbar}\frac{mp\,d\Omega}{(2\pi\hbar)^3}\,|\langle u_p'|\,H'(\mathbf{x}, t')\,|u_{p_0}'\rangle|^2, \tag{92}$$

where

$$u_q' = \exp(i\hbar^{-1}\mathbf{q}\cdot\mathbf{x}) \tag{93}$$

are plane waves normalized to a unit volume. Since the density of states for these plane waves is

$$\rho(E) = mp\,d\Omega/(2\pi\hbar)^3, \tag{94}$$

one sees that the transition probability for the outgoing wave is

$$w = (2\pi/\hbar)\rho(E)\,|\langle u_p'|\,H'(\mathbf{x}, t')\,|u_{p_0}'\rangle|^2 \tag{95}$$

as given by the "golden rule."

The general term of the U matrix can be easily obtained for the case of the "adiabatic" potential ($\beta \to \infty$). The result of Eq. (43) with the appropriate substitutions gives the σ_1 integration in the general term given in Eq. (37) as

$$U_{qp_0}^{(n)}(t, -\infty) = (-1)^n(2\pi)^{-n/2}\hbar^{-n}\lim_{\epsilon_2\to 0}\int_{-\infty}^{\infty}d\sigma_2\lim_{\epsilon_3\to 0}\int_{-\infty}^{\infty}d\sigma_3$$
$$\cdots\times\int_{-\infty}^{\infty}d\sigma_n\frac{[I_{ql}\,\langle u_l|\,g(\sigma_2 - i\epsilon_2)\,|u_m\rangle\cdots\langle u_p|\,g(\sigma_n - i\epsilon_n)\,|u_{p_0}\rangle]}{{}^{2n}\Omega_{lp_0}\,{}^{3n}\Omega_{mp_0}\cdots{}^{nn}\Omega_{pp_0}}, \tag{96}$$

where

$$I_{ql} = i(2\pi)^{\frac{1}{2}}V_{ql}\left\{\frac{(\pi/\beta)}{\sin[(-i\pi/\beta)(i\alpha + {}^{2n}\Omega_{qp_0})]} - i\sum_{m=0}^{\infty}\frac{(-1)^m\exp\{i[{}^{2n}\Omega_{qp_0} + i(\alpha + \beta m)]t\}}{{}^{2n}\Omega_{qp_0} + i(\alpha + m\beta)}\right\}. \tag{97}$$

Just as for the first-order result of Eq. (80) the contributions from the poles of the second term of Eq. (97) must precisely cancel the contributions from the poles in the second quadrant of the first term. Thus Eq. (97) becomes

$$I_{ql} = i(2\pi)^{\frac{1}{2}}V_{ql}\frac{(\pi/\beta)}{\sin[(-i\pi/\beta)(i\alpha + {}^{2n}\Omega_{qp_0})]}, \tag{98}$$

where only poles in the fourth quadrant of the final q integration are to be considered.
For β large this becomes

$$I_{ql} = -(2\pi)^{\frac{1}{2}}V_{ql}(i\alpha + {}^{2n}\Omega_{qp_0})^{-1}. \tag{99}$$

The σ_2 integration involves the terms

$$I_{qm} = \lim_{\epsilon_2\to 0}\int_{-\infty}^{\infty}d\sigma_2\,I_{ql}\frac{\langle u_l|\,g(\sigma_2 - i\epsilon_2)\,|u_m\rangle}{{}^{2n}\Omega_{lp_0}}. \tag{100}$$

This integral can be evaluated by closing the contour in the lower half-plane so that

$$I_{qm} = 2\pi i V_{ql} V_{lm} \frac{(\pi/\beta)}{\sin\left[(-i\pi/\beta)(2i\alpha + {}^{3n}\Omega_{qp_0})\right]} \left\{ \sum_{m=0}^{\infty} \left[\omega_{lq} - i(\alpha + \beta m)\right]^{-1} - \sum_{m=1}^{\infty} \left[{}^{3n}\Omega_{lp_0} + i(\alpha - \beta m)\right]^{-1} \right\}. \quad (101)$$

For large β,

$$I_{qm} = -2\pi V_{ql} V_{lm} [2i\alpha + {}^{3n}\Omega_{qp_0}]^{-1} [\omega_{lq} - i\alpha]^{-1}. \quad (102)$$

Since only one of the denominators of Eq. (102) contains σ_3, the σ_3 integration becomes identical to the σ_2 integration if appropriate substitutions are made. This process can be repeated for all the σ integrations giving, for any α,

$$U^{(n)}_{qp_0}(t, -\infty) = \frac{\hbar^{-n} V_{ql} V_{lm} \cdots V_{pp_0}}{(\omega_{ql} + i\alpha)(\omega_{qm} + i2\alpha) \cdots (\omega_{qp_0} + in\alpha)}, \quad (103)$$

where only the poles in the fourth quadrant of the final q integration are to be considered. In the limit $t \to \infty$, the q integration will vanish unless $q_R = q'$ [see Eq. (79)]. Since E_l, E_m, etc., are integration variables, their values can be fixed so that the poles associated with them do not contribute. This leaves only the pole at $E_q = E_{p_0} - in\alpha$. For this pole q_R becomes p_0 in the limit $\alpha \to 0$. Consequently choosing $q' = p_0$ for small α gives for the nth term of I_C, in the limit $t \to \infty$,

$$I^{(n)}_{C\ t\to\infty} = -2\pi i \hbar^{-n} \frac{\exp\left[i(p_0 |\mathbf{x}| - E_0 t)\right] V_{p_0 s} V_{sm} \cdots V_{rp} V_{pp_0}}{[\omega_{p_0 s} + i\alpha][\omega_{p_0 m} + 2i\alpha] \cdots [\omega_{p_0 r} + i(n-1)\alpha]}. \quad (104)$$

As a result, the nth term of an S matrix,

$$S^{(n)} = -2\pi i \hbar^{n-1} \delta(E_q - E_0) \frac{V_{qs} V_{sm} \cdots V_{rp} V_{pp_0}}{[\omega_{qs} + i\alpha][\omega_{qm} + 2i\alpha] \cdots [\omega_{qr} + i(n-1)\alpha]}, \quad (105)$$

can be constructed in agreement with Eq. (167a) of Schweber.[6]

ACKNOWLEDGMENTS

The authors wish to thank Professor Joseph Keller and Professor Roland H. Good, Jr. for many stimulating discussions concerning these calculations.

APPENDIX

The integration along the contour C_2 can be done using a method of asymptotic expansion for Fourier type integrals that has recently been published.[12] The integral to be evaluated is, from Eq. (70),

$$I_{C_2} = \int_{\infty e^{-i\pi/4}}^{\infty e^{3i\pi/4}} q\, dq \exp\left\{i\hbar^{-1}[q |\mathbf{x}| - (q^2/2m)t]\right\} \langle u_q | U(t, -\infty) - 1 | u_{p_0} \rangle. \quad (106)$$

By making the substitutions

$$q_I = -q_R + q' = -(\hbar m \xi)^{\frac{1}{2}}$$

where $q = q_R + iq_I$, this integral can be put in the form of the Laplace transform

$$I_{C_2} = \tfrac{1}{2}(1 - i)\hbar m \exp\left[i\hbar^{-1}(q'^2/2m)t\right] \int_{\infty}^{0+} \left[(1 - i + q'(\hbar m\xi)^{-\frac{1}{2}}\right] \exp(-\xi t) \langle u_q | U(t, -\infty) - 1 | u_{p_0} \rangle\, d\xi, \quad (107)$$

where the phase of ξ is zero at the end of the contour. To obtain the asymptotic expansion, the integrand is expanded about $\xi = 0$. The first term in this expansion is

$$I_{C_2} \sim \tfrac{1}{2}(1 - i)(\hbar m)^{\frac{1}{2}} q' \exp\left[i\hbar^{-1}(q'^2/2m)t\right] \langle u_{q'} | U(t, -\infty) - 1 | u_{p_0} \rangle \int_{\infty}^{0+} \xi^{-\frac{1}{2}} \exp(-\xi t)\, d\xi. \quad (108)$$

The Laplace transform is easily evaluated and is of order $t^{-\frac{1}{2}}$. Since t is related to $|\mathbf{x}|$ through Eq. (73), I_{C_2} is of order $|\mathbf{x}|^{-\frac{1}{2}}$ and therefore does not contribute to the scattering cross section.

[12] T. A. Weber, D. M. Fradkin, and C. L. Hammer, Ann. Phys. (N. Y.) 27, 362 (1964).