JOURNAL OF
MATHEMATICAL
PHYSICS

VOLUME 8, NUMBER 5 MAY 1967

New Approach to Unified Field Theory

JEROME BLACKMAN
Department of Mathematics, Syracuse University, Syracuse, New York

(Received 26 July 1966)

A new approach to a unified field theory combining Einstein's gravitational equations and the Maxwell equations is developed using a geometry in which $GL(4, C)$ replaces $GL(4, R)$ as the group of the principal bundle. The resulting equations are the same as Einstein's in the case of empty space but differ from those he proposed for the combined gravitational and electromagnetic fields. Charge is shown to appear only at singularities, and a symmetric solution corresponding to the classical Schwarzschild metric is presented.

1. INTRODUCTION

SOON after the introduction of general relativity by Einstein in 1915, Weyl[1,2] attempted to develop a geometry which would provide not only a theory of gravitation but also of electromagnetic phenomena. The basis of his idea was to allow the length scale to vary from point to point in space; this could be described in terms of a differential form $\phi_i\, dx^i$ and the ϕ_i would then be the electromagnetic potentials. Although this theory had certain attractive features,[3] it had to be abandoned because of several deficiencies. Other attempts at a unified theory followed, notably the "already unified" theory of Rainich, Misner, and Wheeler, and Einstein's later theories based on a nonsymmetric metric tensor. The lack of any experimental results to indicate an effect on gravitation caused by electromagnetic events has left the question wide open.

The advent of quantum theory pushed these speculations out of the main stream of physics for several reasons. One was the fact that the physical scale of the phenomena was completely different. For all practical purposes gravitation could be ignored in questions of quantum theory. Seemingly, the Lorentz–Minkowski geometry of special relativity was adequate to describe the geometry, at least at a certain level of abstraction. Another reason, of course, was again the sparse nature of experimental results in general relativity and the complete lack of any results connecting gravitation and quantum phenomena.

However, the problem of the geometry of the physical world "in the small" may be at the heart of even quantum theory. For one thing, it is recognized that the uncertainty principle calls into question the concept of manifold itself since it raises doubts as to the physical meaning of a point in space–time. For another, it can be argued that behind the skepticism as to the validity of quantum theory, prevalent since the inception of the theory, there lurks the rejection of a physical theory that is so nongeometric.

It would seem then that the geometry of the physical world is still very much an open question and that there is at least the possibility of shedding light not only on the relation of gravitation and electromagnetism but even on the foundations of quantum theory. This paper does not solve these problems. What has been done is to introduce a new geometry which is

[1] H. Weyl, *Gravitation und Elektrizität* (Sitzber. Preuss. Akad. Wiss., Berlin, 1918), pp. 465–480.
[2] H. Weyl, *Space, Time, Matter* (Dover Publications, Inc., New York, 1950).
[3] R. Adler, M. Bazin, and M. Schiffer, *Introduction to General Relativity* (McGraw-Hill Book Company, Inc., New York, 1965).

clearly a variation on Weyl's ideas and which does yield an equation relating gravitation to electromagnetism which is different from Einstein's. The geometry has features reminiscent of the formalism of quantum theory, although we have not as yet been able to show a direct link. In particular, the geometry has the same type of gauge invariance as quantum electrodynamics, and it is $GL(4, C)$ rather than $GL(4, R)$ that is the Lie group of the fiber bundle over the manifold.

2. GEOMETRY

We first describe the formalism of the geometry in heuristic terms and from this derive an expression for the connection coefficients and demonstrate the gauge invariance properties. This is followed by a brief indication of how the geometry can be fitted into the modern theory of differential geometry via the theory of fiber bundles.

Instead of a change in the metric scale from point to point of a four-dimensional manifold, as introduced by Weyl, we think of length measurement as having a phase associated with it that varies from point to point. Assume that there is a nondegenerate tensor field g_{ij} which plays, in a modified way, the role of the metric tensor. The components of a vector ξ^j are allowed to be complex numbers. For the moment we do not specify the real or complex nature of g_{ij}. Using the Einstein summation convention, the square of the "length" of a vector $l^2 = g_{jk}\xi^j\xi^k$ need no longer be a real number but we can write $g_{jk}\xi^j\xi^k = |l|^2 e^{2i\phi}$, where ϕ is then the phase associated with the length of the vector ξ^j. To develop an appropriate connection assume that under parallel displacement the length $|l|$ of ξ^j does not change but the phase changes by $d\phi = \phi_j\, dx^j$, where ϕ_j is a real vector and dx^j represents a small displacement of position. If $l = |l|\, e^{i\phi}$ then under the displacement dx^j, l changes by $dl = i\,|l|\, e^{i\phi}\, d\phi = il\phi_j\, dx^j$. Defining Γ^i_{kj} by $d\xi^i = \Gamma^i_{jk}\, dx^j\xi^k$, where $d\xi^k$ is the change in the component ξ^k under the displacement dx^j, we get

$$\begin{aligned} d(l^2) &= d(g_{jk}\xi^j\xi^k) = 2l\, dl = 2il^2\phi_m\, dx^m \\ &= 2ig_{jk}\xi^j\xi^k\phi_m\, dx^m \\ &= g_{jk|m}\xi^j\xi^k\, dx^m + g_{jk}\Gamma^j_{nm}\, dx^m\xi^n\xi^k \\ &\quad + g_{jk}\Gamma^k_{nm}\xi^j\xi^n\, dx^m. \end{aligned} \tag{1}$$

Here and below $g_{jk|m} = \partial g_{jk}/\partial x^m$.

Equations (1) can be solved as in Ref. 3, p. 49, to yield

$$\Gamma^i_{jk} = -\left\{ \begin{matrix} i \\ jk \end{matrix} \right\} + ig^{mi}[g_{mj}\phi_k + g_{mk}\phi_j - g_{jk}\phi_m], \tag{2}$$

where $\left\{ \begin{matrix} i \\ jk \end{matrix} \right\}$ is the usual Christoffel symbol of the second kind defined as a function of g_{ij} by

$$\left\{ \begin{matrix} i \\ jk \end{matrix} \right\} = \tfrac{1}{2}g^{il}(g_{jl|k} + g_{kl|j} - g_{jk|l}).$$

For convenience write

$$\Gamma^i_{jk} = \Gamma^i_{jk} + i\bar{\bar{\Gamma}}^i_{jk}, \tag{3}$$

where

$$\Gamma^i_{jk} = -\left\{ \begin{matrix} i \\ jk \end{matrix} \right\}$$

and

$$\bar{\bar{\Gamma}}^i_{jk} = g^{mi}(g_{mj}\phi_k + g_{mk}\phi_j - g_{jk}\phi_m).$$

It should be clearly noted, however, that this will be the decomposition of Γ^i_{jk} into real and imaginary parts if all g_{ij} are real but otherwise need not be. Observe also that Γ^i_{jk} is invariant under the simultaneous gauge transformations

$$g_{jk} \to g_{jk}e^{2i\psi} \quad \phi_k \to \phi_k + \psi_{|k}, \tag{4}$$

where ψ is an arbitrary differentiable function on the manifold. Since this type of invariance is of basic interest, it is clear that we cannot assume that g_{jk} is always real. We can either allow g_{jk} to be complex or restrict the solutions of the equations developed below by requiring that there always be a choice of gauge function ψ such that $g_{ij}e^{2i\psi}$ is real. Some special results are developed using this assumption. At present there is no interpretation available for solutions for which there is no such gauge function.

In order to put the above discussion into the framework of current mathematics we turn to the theory of fiber bundles. The notation and terminology is mainly that of Ref. 4, particularly pp. 50ff. and pp. 140ff. The entire base manifold, and therefore the bundle, cannot be described since the topology in the large is unknown. What can be done is to give the bundle, connection, and metric in some coordinate neighborhood and then take the whole bundle to be some extension. In a later section a solution in the large is given and this can be taken as an existence theorem.

The base manifold is then U, assumed diffeomorphic to an open subset of R^4; U may also be assumed to be a coordinate neighborhood. The Lie group is $GL(4, C)$, the nonsingular 4×4 matrices with complex entries. Let X_1, X_2, X_3, and X_4 be a basis for $T_x(U)$, the tangent space at $x \in U$. The collection of all quadruples $(c^i_1X_i, \cdots, c^i_4X_i)$ where $(c^i_j) \in GL(4, C)$ is called the space of complex frames at x. $L^c(U)$ is the set of all complex frames at all points of U. $GL(4, C)$ acts on $L^c(U)$ on the right in an obvious manner: If $(x; Y_1, \cdots, Y_4) = u \in L^c(U)$ and $(c^i_j) = c \in GL(4, C)$,

[4] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry* (John Wiley & Sons, Inc., New York, 1963).

then uc is $(x; c_1^i Y_i, \cdots, c_4^i Y_i)$. Define $\pi(u) = x$ if u is a complex frame at x. It is clear that $GL(4, C)$ acts freely on $L^c(U)$. If x^1, x^2, x^3, x^4, are coordinates in U then every complex frame is of the form

$$\left(c_1^i \frac{\partial}{\partial x^i}, \cdots, c_4^i \frac{\partial}{\partial x^i}\right)$$

$\pi^{-1}(x)$ is then in one-to-one correspondence with $GL(4, C)$ and coordinates are introduced in the obvious way so that $L^c(U)(U, GL(4, C))$ becomes a principal fiber bundle.

The complex tangent bundle $E = T^c(U)$ can also be defined to be the bundle associated with $L^c(U)$ with standard fiber C^4. Then $\pi_E^{-1}(x)$ is isomorphic with C^4 and in terms of coordinates x^1, x^2, x^3, x^4, in U, is represented as the set of all $c^i \partial/\partial x^i$, where $\{c^i\}$, $i = 1, \cdots, 4$, are complex numbers. These are the "complex vectors" referred to above. If $u \in L^c(U)$ then u can be considered in the usual way as an isomorphism of C^4 on $\pi_E^{-1}(x)$ if $x = \pi(u)$. In our case if $u = (x; c_1^i X_i, \cdots, c_4^i X_i)$ and $\xi = (\xi^1, \cdots, \xi^4) \in C^4$ then $u\xi = \xi^j c_j^i X_i$.

A connection form ω can now be described in $L^c(U)$. If E_j^i is the matrix with 1 in the (i, j) position and zero elsewhere then $\{E_j^i\}_{i,j=1,\cdots,4}$ form a basis for $gl(4, C)$, the Lie algebra of $GL(4, C)$. If we set $\omega = \omega_j^i E_i^j$ and

$$\omega_j^i = Y_k^i (dX_j^k - \Gamma_{lm}^k X_j^l \, dx^m)$$

we have a connection form in $L^c(U)$. Here Γ_{lm}^k is given by (2), (x^i, X_j^k) are the coordinates in $L^c(U)$, and Y_k^i is the matrix inverse to X_j^k. The complete discussion for the real case is found in Ref. 4, pp. 140ff.; only formal changes are necessary in the complex case. Our choice of Γ_{lm}^k differs from that in Ref. 4 by a negative sign, corresponding (in the case $\phi_i \equiv 0$) to the notation in Ref. 3.

3. CURVATURE TENSOR AND FIELD EQUATIONS

Having obtained the Γ_{jk}^i, the curvature tensor R_{jkl}^i could be computed according to the formula

$$R_{jkl}^i = -\Gamma_{jk|l}^i + \Gamma_{jl|k}^i + \Gamma_{ml}^i \Gamma_{jk}^m - \Gamma_{mk}^i \Gamma_{jl}^m. \quad (5)$$

Writing $R_{jkl}^i = \bar{R}_{jkl}^i + i\bar{\bar{R}}_{jkl}^i$, and putting in (3) one obtains

$$\bar{R}_{jkl}^i = -\Gamma_{jk|l}^i + \Gamma_{jl|k}^i + \Gamma_{ml}^i \Gamma_{jk}^m - \Gamma_{mk}^i \Gamma_{jl}^m - \bar{\bar{\Gamma}}_{ml}^i \bar{\bar{\Gamma}}_{jk}^m + \bar{\bar{\Gamma}}_{mk}^i \bar{\bar{\Gamma}}_{jl}^m,$$

$$\bar{\bar{R}}_{jkl}^i = -\bar{\bar{\Gamma}}_{jl|k}^i + \bar{\bar{\Gamma}}_{jl|k}^i + \Gamma_{ml}^i \bar{\bar{\Gamma}}_{jk}^m + \bar{\bar{\Gamma}}_{ml}^i \Gamma_{jk}^m - \Gamma_{mk}^i \bar{\bar{\Gamma}}_{jl}^m - \bar{\bar{\Gamma}}_{mk}^i \Gamma_{jl}^m.$$

Since $R_{jl} = R_{jil}^i = \bar{R}_{jl} + i\bar{\bar{R}}_{jl}$ is desired, contraction gives

$$\bar{R}_{jl} = \mathring{R}_{jl} - \bar{\bar{\Gamma}}_{mi}^i \bar{\bar{\Gamma}}_{jl}^m + \bar{\bar{\Gamma}}_{mi}^i \bar{\bar{\Gamma}}_{jl}^m, \quad (6)$$

$$\bar{\bar{R}}_{jl} = -\bar{\bar{\Gamma}}_{jl|i}^i + \bar{\bar{\Gamma}}_{jl|i}^i + \Gamma_{mi}^i \bar{\bar{\Gamma}}_{jl}^m + \bar{\bar{\Gamma}}_{mi}^i \Gamma_{jl}^m - \Gamma_{mi}^i \bar{\bar{\Gamma}}_{jl}^m - \bar{\bar{\Gamma}}_{mi}^i \Gamma_{jl}^m, \quad (7)$$

where $\mathring{R}_{jl}$ is the contracted curvature tensor obtained by using only the classical metric connection corresponding to $\Gamma_{jk}^i = -\left\{ {i \atop jk} \right\}$. That is, $\mathring{R}_{jl} = R_{jl}$ in the special case $\phi_i \equiv 0$.

A simple computation shows that

$$\bar{R}_{jl} = \mathring{R}_{jl} - 2g_{jl}\phi^i\phi_i + 2\phi_j\phi_l. \quad (8)$$

The computation of $\bar{\bar{R}}_{jl}$ is longer so an outline of the results is given for the convenience of the reader.

$$\bar{\bar{\Gamma}}_{jl|l}^i = 4\phi_{j|l}, \quad (9a)$$

$$\bar{\bar{\Gamma}}_{jl|i}^i = \phi_{l|j} + \phi_{j|l} - (g_{jl}\phi^i)_{|i}, \quad (9b)$$

$$\Gamma_{ml}^i \bar{\bar{\Gamma}}_{ji}^m = -g_{lm|j}\phi^m + g_{jl|m}\phi^m - \tfrac{1}{2}g^{km}g_{km|l}\phi_j, \quad (9c)$$

$$\bar{\bar{\Gamma}}_{ml}^i \Gamma_{ji}^m = -g_{jm|l}\phi^m + g_{jl|m}\phi^m - \tfrac{1}{2}g^{km}g_{km|j}\phi_l, \quad (9d)$$

$$\Gamma_{mi}^i \bar{\bar{\Gamma}}_{jl}^m = -\tfrac{1}{2}(\phi_l g^{km}g_{km|j} + \phi_j g^{km}g_{km|l} - g_{jl}\phi^i g^{km}g_{km|i}), \quad (9e)$$

$$\bar{\bar{\Gamma}}_{mi}^i \Gamma_{jl}^m = -2\phi^m g_{jm|l} - 2\phi^m g_{lm|j} + 2\phi^m g_{jl|m}. \quad (9f)$$

Putting these in (7) yields

$$\bar{\bar{R}}_{jl} = -3\phi_{j|l} + \phi_{l|j} - (g_{jl}\phi^m)_{|m} + \phi^m g_{lm|j} + \phi^m g_{jm|l} - \tfrac{1}{2}\phi^i g_{jl}g^{km}g_{km|i}. \quad (10)$$

This can be further simplified. Note that we have two connections Γ_{jk}^i and $\Gamma_{jk}^i = -\left\{ {i \atop jk} \right\}$. Let $|||$ designate the covariant derivative based on Γ_{jk}^i and $\|$ designate that based on Γ_{jk}^i. With this notation

$$\bar{\bar{R}}_{jl} = \phi_{l\|j} - 3\phi_{j\|l} - g_{jl}\phi_{\|k}^k. \quad (11)$$

Observe that since $\Gamma_{jk}^i = \Gamma_{kj}^i$,

$$\phi_{l|j} - \phi_{j|l} = \phi_{l\|j} - \phi_{j\|l} = \phi_{l|||j} - \phi_{j|||l}.$$

From (8) and (11) it follows that

$$R_{jl} = \bar{R}_{jl} + i\bar{\bar{R}}_{jl} = [R_{jl} - 2g_{jl}\phi^i\phi_i + 2\phi_j\phi_l] + i[\phi_{l\|j} - 3\phi_{j\|l} - g_{jl}\phi_{\|i}^i]. \quad (12)$$

Since Γ_{jk}^i is invariant under the gauge transformation (4), so is R_{jkl}^i and R_{jl}. Of course, $\bar{R}_{jl}$ and $\bar{\bar{R}}_{jl}$ are not separately invariant.

Once (12) is obtained it is tempting to take (by analogy with the classical case) $R_{jl} = 0$ for the field equations. However, it is easily seen that in this case R_{jl} is not a symmetric tensor so that $R_{jl} = 0$ would

imply that both the symmetric and antisymmetric parts are zero. The antisymmetric part is

$$\tfrac{1}{2}(R_{jl} - R_{lj}) = 2i\{\phi_{l\|j} - \phi_{j\|l}\}.$$

On the other hand, we wish to take ϕ_i as the vector potential of the electromagnetic field, in which case the electromagnetic 4-tensor is $F_{ij} = \phi_{i\|j} - \phi_{j\|i}$ except for a constant multiplier. Therefore, demanding that the antisymmetric part of R_{jl} vanish would also require having zero electromagnetic field. This problem can be avoided if we observe that $\phi_{l\|j} - \phi_{j\|l} = \phi_{l|j} - \phi_{j|l}$ is gauge invariant so that

$$R_{jl} = 2i\{\phi_{l\|j} - \phi_{j\|l}\} \tag{13}$$

is a set of gauge invariant field equations. This is equivalent to requiring that only the symmetric part of R_{jl} be zero. Another gauge invariant quantity which could possibly be added to the right side of (13) is $\lambda g_{jl}R$, where λ is a constant and $R = g^{ij}R_{ij}$. Exactly as in the classical case, however, the classes of solutions of (13) and

$$R_{jl} = 2i\{\phi_{l\|j} - \phi_{j\|l}\} + \lambda g_{jl}R \tag{14}$$

are exactly the same except if $\lambda = \frac{1}{4}$, in which case the class of solutions of (14) contains those of (13) as a subset.

R can be computed from (12). Since $\phi_{l\|j}$ is the covariant derivative of ϕ_l with respect to the metric connection Γ^i_{jk} it follows as usual that $\phi_{j\|l}g^{jk} = \phi^k_{\|l}$ since $g^{jk}_{\|l} = 0$. Therefore

$$R = \mathring{R} - 6\phi^k\phi_k - 6i\phi^k_{\|k} = 0, \tag{15}$$

the last equation being a consequence of (13). From (13), (14), and (15), with $\lambda = \frac{1}{2}$ in (14), it follows that

$$\mathring{R}_{jl} - \tfrac{1}{2}g_{jl}\mathring{R} + g_{jl}\phi^i\phi_i + 2\phi_j\phi_l - i(\phi_{j\|l} + \phi_{l\|j} - 2g_{jl}\phi)^k_{\|k} = 0. \tag{16}$$

Suppose for the moment there is a real gauge, i.e., a gauge in which all g_{ij} are real. Then (15) implies $\mathring{R} = 6\phi^k\phi_k$ and $\phi^k_{\|k} = 0$. Equation (16) then becomes the pair of equations

$$\mathring{R}_{jl} - \tfrac{1}{2}g_{jl}\mathring{R} + g_{jl}\phi^i\phi_i + 2\phi_j\phi_l = 0, \tag{17a}$$

$$\phi_{j\|l} + \phi_{l\|j} = 0. \tag{17b}$$

The last is nothing but the Killing equation and is equivalent to the statement that if a coordinate system is chosen so that $\phi^j = (1, 0, 0, 0)$ then g_{ij} is independent of the first coordinate. In this gauge ϕ_j has the dual role of being both the vector potential of the electromagnetic field and the tangent to lines of symmetry. We return to this below.

4. MAXWELL EQUATIONS

Maxwell's equations can be written as

$$F_{ij|k} + F_{jk|i} + F_{ki|j} = 0, \tag{18a}$$

$$F^{ij}_{\|j} = s^i, \tag{18b}$$

where s^i is the current vector. Equation (18a) is equivalent to $F_{ij} = \phi_{i\|j} - \phi_{j\|i} = \phi_{i|j} - \phi_{j|i}$. We show that if there is a real gauge then $s^i = 0$. This statement can be interpreted as meaning that charge can occur only at singularities of the manifold. In fact, in a real gauge, Eqs. (17a, b) hold so that

$$F_{ij} = \phi_{i\|j} - \phi_{j\|i} = 2\phi_{i\|j},$$

$$F^i_{j\|k} = 2\phi^i_{\|j\|k}.$$

In the real gauge $\phi^i_{\|i} = 0$ and $\mathring{R} = 6\phi^k\phi_k$. Using these and the equation,

$$\mathring{R}^i_{njk}\phi^n = \phi^i_{\|j\|k} - \phi^i_{\|k\|j},$$

we see that

$$\begin{aligned} F^i_{j\|i} &= 2\{\phi^i_{\|j\|i} - \phi^i_{\|i\|j}\} = -2\mathring{R}^i_{nij}\phi^n = -2\mathring{R}_{nj}\phi_n \\ &= 2(g_{nj}\phi^i\phi_i - \phi_n\phi_j)\phi^n \\ &= 2(\phi_j\phi^i\phi_i - \phi_j\phi^n\phi_n) = 0. \end{aligned}$$

Therefore $s^i = F^{ij}_{\|i} = 0$. It is not obvious that this is a gauge invariant statement, but this fact can be demonstrated by showing that

$$s^i = g^{in}g^{mk}f_{nm\|k} = g^{in}g^{mk}f_{nm\|\|k}, \tag{19}$$

since the last expression has weight -4 so that if it vanishes in one gauge it vanishes in all gauges. To show this it is sufficient to observe that the two terms in (19) differ by

$$g^{in}g^{mk}\{\bar{\bar{\Gamma}}^j_{nk}f_{jm} + \bar{\bar{\Gamma}}^j_{km}f_{nj}\}$$

and to then directly compute this; it is zero.

5. A SOLUTION

It is possible to find a solution to the field equations which corresponds to the classical Schwarzschild solution of the free-space equation. This demonstrates, as mentioned above, that the field equations do have solutions.

If the equations are treated in a real gauge, (17a) is

$$\mathring{R}_{jl} - \tfrac{1}{2}g_{jl}\mathring{R} = -g_{jl}\phi^i\phi_i - 2\phi_j\phi_l. \tag{20}$$

Equation (17b) is satisfied if we take $\phi^i = (1, 0, 0, 0)$ and obtain a solution in which g_{ij} is independent of

x^1. As in Ref. 3, pp. 280ff, assume

$$g_{ij} = \begin{pmatrix} e^{\nu} & 0 & 0 & 0 \\ 0 & -e^{\lambda} & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2 \sin^2\theta \end{pmatrix} \tag{21}$$

where ν and λ are functions of r alone. Then the right side of (20) is easily computed, and the left side appears in Ref. 3, Eq. (9.118a)–(9.118d). All the equations corresponding to $j \neq l$ are identically zero. The (4, 4) equation is the same as the (3, 3) equation, and the three equations corresponding to (1, 1), (2, 2) and (3, 3) are, after minor rearrangement,

$$e^{-\lambda}(-\lambda'/r + r^{-2}) - r^{-2} = -3e^{\nu}, \tag{22a}$$

$$-e^{-\lambda}(\lambda'/r + r^{-2}) + r^{-2} = e^{\nu}, \tag{22b}$$

and

$$e^{-\lambda}\left[\frac{\nu'\lambda'}{4} - \frac{(\lambda')^2}{4} - \frac{\nu''}{2} - \frac{(\nu' - \lambda')}{2r}\right] = e^{\nu}. \tag{22c}$$

Equations (22a) and (22b) yield

$$e^{\nu+\lambda} = (\lambda' + \nu')/2r, \tag{23}$$

which is readily integrated to

$$e^{\nu} = e^{-\lambda}/(c - r^2), \tag{24}$$

where c is a constant of integration.

By differentiating (22b) and using (23) to eliminate e^{ν} from the resulting expression it is straightforward to show that (22c) follows, so that (22c) is a consequence of (22a) and (22b). Putting (24) in (22a) leads to

$$1 + [3r^2/(c - r^2)] = \lambda' r + e^{\lambda}, \tag{25}$$

which can be integrated to

$$e^{-\lambda} = (c - r^2)[(1/c) - (k/r)(c - r^2)^{\frac{1}{2}}], \tag{26}$$

where k is arbitrary. It is, in fact, a routine exercise to show that (26) and

$$e^{\nu} = (1/c) - (k/r)(c - r^2)^{\frac{1}{2}} \tag{27}$$

satisfy the equations. The result is

$$(ds)^2 = \left\{\frac{1}{c} - \frac{k}{r}(c - r^2)^{\frac{1}{2}}\right\}(dx^1)^2 - \frac{(dr)^2}{(c - r^2)[(1/c) - (k/r)(c - r^2)^{\frac{1}{2}}]} - r^2[d\theta^2 + \sin^2\theta(d\phi)^2]. \tag{28}$$

By changing the x^1 variable by $x^1 = (1/\epsilon)(x^1)'$ and changing the constants c and k, the solution can be brought to the form

$$(ds)^2 = \left\{\frac{1}{c} - k\left(\frac{c}{r^2} - \epsilon^2\right)^{\frac{1}{2}}\right\}(dx^1)^2 - \frac{(dr)^2}{(c - \epsilon^2 r^2)\{(1/c) - k[(c/r^2) - \epsilon^2]^{\frac{1}{2}}\}} - r^2[d\theta^2 + \sin^2\theta(d\phi)^2] \tag{29}$$

with $\phi^i = (\epsilon, 0, 0, 0)$. In this form it is apparent that formally letting $\epsilon \to 0$ yields the Schwarzschild metric of general relativity. Also, $\phi_1 = g_{11}\phi^1 = \epsilon/c - k\epsilon \times (c/r^2 - \epsilon^2)^{\frac{1}{2}}$. If $\epsilon^2 \ll c/r^2$, ϕ_1 is approximately $\epsilon/c - k\epsilon c^{\frac{1}{2}}/r$. Note that (29) is a real gauge solution only if $r < c^{\frac{1}{2}}/\epsilon$.

6. LINES OF SYMMETRY

It was shown above that in the real gauge the integral curves of the ϕ^i field are lines of symmetry. In the solution presented in Sec. 5, for example, the integral curves were just the x^1 lines, and translation along these lines left the metric tensor invariant. The ϕ^i vector in this case is the 4-velocity of the singularity except for a constant multiplier. We are led to the conjecture that for any singularity the ϕ^i vector near it, if nonzero, will have some interpretation as a velocity or momentum. Presumably, if a solution could be found representing two singularities, the ϕ^i field near each would represent a velocity and one could find the interaction between the two. While this is a possibility for the future we can now only derive a rather suggestive equation. We work only in the real gauge and let $z^i(p, q)$ represent the integral curve of ϕ^i through the point q, i.e., $z^i(0, q) = q^i$ and $(\partial z^i/\partial p)(0, q) = \phi^i(q)$. Assume also that in the region considered ϕ^i is timelike. Then

$$1 = \frac{dz^i}{ds}\frac{dz^j}{ds}g_{ij} = \frac{dz^i}{dp}\frac{dz^j}{dp}\left(\frac{dp}{ds}\right)g_{ij} = \phi^i\phi^j g_{ij}\left(\frac{dp}{ds}\right)^2 = \frac{\mathring{R}}{6}\left(\frac{dp}{ds}\right)^2$$

Therefore $dp/ds = \pm(6/\mathring{R})^{\frac{1}{2}}$ and $dz^k/ds = \pm\phi^k(6/\mathring{R})^{\frac{1}{2}}$.

The absolute derivative of ϕ^k along the z^k curves (with respect to the Γ^i_{jk} connection) is given by

$$\frac{D\phi^i}{Ds} = \phi^i_{\|k}\frac{dz^k}{ds} = \phi^i_{\|k}\frac{dz^k}{dp}\frac{dp}{ds} = \pm\phi^i_{\|k}\phi^k(6/\mathring{R})^{\frac{1}{2}}.$$

(See Ref. 3, p. 91.)

Remembering that $F_{ik} = 2\phi_{i\|k}$ in the real gauge this can be written

$$\frac{D\phi^i}{Ds} = \pm\frac{1}{2}\left(\frac{6}{\mathring{R}}\right)^{\frac{1}{2}}F^i_{\|k}\phi^k.$$

This is formally quite like the Lorentz equation but the interpretation is quite different.

Clebsch–Gordan Series for Symmetrized Tensor Products*

C. M. ANDERSEN
Department of Physics, Purdue University, Lafayette, Indiana

(Received 13 October 1966)

An explicit method for obtaining the Clebsch–Gordan series for the symmetrized n-fold direct products of a finite-dimensional representation of a semisimple Lie group is discussed and illustrated. The method is based on the use of weight diagrams.

1. INTRODUCTION

WE discuss in this paper an explicit method for obtaining the Clebsch–Gordan series for the direct product of a finite-dimensional representation of a semisimple Lie group $\mathfrak{L}$ taken with itself n times and symmetrized according to a definite representation of the symmetric group $\mathfrak{S}_n$.

Such symmetrized representations arise in a natural way in elementary particle physics if states describing the orbital angular momentum, spin, and internal symmetry of a collection of n similar particles are required to satisfy a generalized Pauli principle. For example, this requirement has as a consequence the result that a three-pion state arising from the decay of an ω must have a completely antisymmetric spatial wavefunction. This is due to the fact that a state consisting of bosons must be over-all totally symmetric and to the fact that an $I = 0$ representation of $SU(2)$ formed from three $I = 1$ representations must be completely antisymmetric.

Our approach is to consider first the n-fold direct product representation $[\mathfrak{D}]^n = \mathfrak{D} \otimes \mathfrak{D} \otimes \cdots \otimes \mathfrak{D}$ and reduce it according to symmetry type with one representation (in general, reducible with respect to $\mathfrak{L}$) for each standard Young tableau with n boxes. A particular case of this is the writing of a second-rank ($n = 2$) tensor as a sum of symmetric and antisymmetric tensors. Secondly, we determine a character formula for the "symmetrized" representation corresponding to each Young tableau (Sec. 2). This formula is valid for any finite dimensional representation of any semisimple Lie group. We then relate this character formula to weight diagrams and by manipulation of the weight diagrams determine the Clebsch–Gordan series for each of the symmetrized representations (Sec. 3).

In Sec. 4, we demonstrate the method by working out some examples diagrammatically for simple Lie groups of rank $\ell = 1, 2$. For groups of rank $\ell \geq 3$, the weight diagrams are in a higher dimensional space and consequently it becomes easier to work directly with the components of the weight vectors in the appropriate coordinate system. In Sec. 5, we give some examples for the higher-rank simple Lie groups.

In Sec. 6, we discuss rules to achieve the same results for $SU(\ell + 1)$ groups by manipulating Young diagrams.

2. CHARACTER FORMULA

Let $\mathfrak{D}: a \to D_{ij}(a)$ $(i, j = 1, 2, \cdots, m)$ be a finite dimensional representation of the semisimple Lie group $\mathfrak{L}$, where the $D_{ij}(a)$ are $m \times m$ matrices representing elements $a \in \mathfrak{L}$. Let us suppose that $\mathfrak{L}$ acts on the m-dimensional vectors ψ such that the vector transformation law is given by $\psi \to a\psi$ with

$$(a\psi)_i = D_{ij}(a)\psi_j, \tag{2.1}$$

where the ψ_i are the components of ψ with respect to a fixed set of basis vectors. A general unsymmetrized nth-rank tensor $\Psi_{i_1 i_2 \cdots i_n}$ belongs to the reducible representation $[\mathfrak{D}]^n \equiv \mathfrak{D} \otimes \mathfrak{D} \otimes \cdots \otimes \mathfrak{D}: a \to D_{i_1, \cdots, i_n; j_1, \cdots, j_n}(a) \equiv D_{i_1 j_1}(a) D_{i_2 j_2}(a) \cdots D_{i_n j_n}(a)$. We abbreviate the tensor transformation law to $\Psi \to a\Psi$ or

$$(a\Psi)_{(i)} = D_{(i),(j)}(a)\Psi_{(j)}; \; (i) = i_1 i_2 \cdots i_n; \tag{2.2}$$

summation over $(j) = j_1 \cdots j_n$ is understood. Let

$$\pi = \begin{pmatrix} 1 & 2 & \cdots & n \\ 1' & 2' & \cdots & n' \end{pmatrix}$$

be a permutation of the subindices and define $(\pi\Psi)_{(i)} = \Psi_{\pi(i)}$, i.e.,

$$(\pi\Psi)_{i_1 i_2 \cdots i_n} = \Psi_{i_{1'} i_{2'} \cdots i_{n'}}. \tag{2.3}$$

Then it is well known that[1] $\pi a\Psi = a\pi\Psi$. Similarly, if

$$X = \sum_k^{n!} \lambda_k \pi_k$$

* This research was supported in part by the National Science Foundation and by the Air Force Office of Scientific Research, Office of Aerospace, U.S. Air Force, under AFOSR grant number 274-66.

[1] M. Hamermesh, *Group Theory and Its Application to Physical Problems* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1964), p. 379.

(the λ_k are real numbers) is an element of the group algebra A_n of the symmetric group $\mathfrak{S}_n$, we have

$$X\Psi = \sum_k^{n!} \lambda_k \pi_k \Psi \tag{2.4}$$

and

$$Xa\Psi = aX\Psi. \tag{2.5}$$

In particular, we can let X be any one of the projection operators[2]

$$T^\alpha \equiv \frac{f^\alpha}{n!} \sum_\beta \varphi^\alpha_\beta \sum_{\pi \in \mathfrak{C}_\beta} \pi, \tag{2.6}$$

where α is a partition $[\ell^{u_\ell} \cdots 2^{u_2} 1^{u_1}]$ of n letters, u_j is the number of rows of length j in the Young tableau of shape or partition α, φ^α_β is the character component of the group $\mathfrak{S}_n$ for a representation of shape α and for the conjugation class $\mathfrak{C}_\beta$, $\beta = [\ell^{v_\ell} \cdots 2^{v_2} 1^{v_1}]$, and f^α is the dimension of the irreducible representation of $\mathfrak{S}_n$ associated with shape α, or, equivalently, is the number of standard Young tableaux of shape α. The numbers f^α satisfy the relation $\sum_\alpha (f^\alpha)^2 = n!$ and are given by[3]

$$f^\alpha = n! \prod_{\substack{r,s \\ r \le s \le k}} (\lambda_r - \lambda_s - r + s)! \Big/ \prod_{r \le k} (\lambda_r + k - r)!, \tag{2.7}$$

where λ_r is the length of the rth row and k is the number of rows. The T^α have the following properties[2]:

$$\begin{aligned} T^\alpha T^\beta &= \delta^{\alpha\beta} T^\beta, \\ \sum_\alpha T^\alpha &= 1 \quad \text{(the identity element of } A_n\text{)}. \end{aligned} \tag{2.8}$$

Thus we have

$$\begin{aligned} T^\beta(aT^\alpha\Psi) &= aT^\beta T^\alpha \Psi = \delta^{\alpha\beta}(aT^\alpha\Psi), \\ T^\beta(\pi T^\alpha \Psi) &= \delta^{\alpha\beta}(\pi T^\alpha \Psi), \end{aligned} \tag{2.9}$$

where $a \in \mathfrak{L}$ and $\pi \in \mathfrak{S}_n$. The relations (2.8) and (2.9) show that the T^α partition the space of nth-rank tensors into subspaces which are invariant under the direct product group $\mathfrak{L} \otimes \mathfrak{S}_n$.

It should be noted that the T^α are not primitive idempotents; i.e., we can write[4]

$$T^\alpha = \sum_{r=1}^{f^\alpha} e^\alpha_{rr}, \tag{2.10}$$

where the e^α_{rr} are primitive idempotents and satisfy

$$e^\alpha_{rr} e^\beta_{ss} = \delta^{\alpha\beta} \delta_{rs} e^\alpha_{rr}. \tag{2.11}$$

There is one e^α_{rr} for each standard Young tableau with n boxes. In fact one can "complete" the set of e's by defining e^α_{rs}, where α ranges over all Young tableau shapes with n boxes and r and s each range from one to f^α. The $n!$ e's are chosen to be linearly independent and to satisfy the relation[5]

$$e^\alpha_{rs} e^\beta_{uv} = \delta^{\alpha\beta} \delta_{su} e^\alpha_{rv}. \tag{2.12}$$

Thus, any $\pi \in \mathfrak{S}_n$ can be written in the form

$$\pi = \sum_{\alpha,r,s} \lambda^\alpha_{rs}(\pi) e^\alpha_{rs}, \tag{2.13}$$

so that

$$\begin{aligned} &\sum_{\alpha,r} e^\alpha_{rr} \Psi = \Psi, \\ &\pi e^\beta_{nn} \Psi = \sum_{r=1}^{f^\beta} \lambda^\beta_{rn} e^\beta_{rn} \Psi, \\ &e^\alpha_{rr}(a e^\beta_{ss} \Psi) = \delta^{\alpha\beta} \delta_{rs} (a e^\beta_{ss} \Psi), \\ &e^\alpha_{rr}(\pi e^\beta_{ss} \Psi) = \delta^{\alpha\beta}(e^\beta_{rr} \pi e^\beta_{ss}) \Psi \neq \delta^{\alpha\beta} \delta_{rs} (\pi e^\beta_{ss} \Psi). \end{aligned} \tag{2.14}$$

We see that the space of nth-rank tensors is partitioned by the e^α_{rr} into subspaces which are still invariant under $\mathfrak{L}$ but no longer invariant under $\mathfrak{S}_n$. The effect of the elements of $\mathfrak{S}_n$ is to carry a subspace of tensors of the form $e^\beta_{ss}\Psi$ (invariant under $\mathfrak{L}$) into a direct sum of the subspaces of tensors of the form $e^\beta_{rr}\Psi$, $r = 1, \cdots, f^\beta$ (each invariant under $\mathfrak{L}$).

Let us denote the representations of $\mathfrak{L}$ associated with the spaces of tensors of the form $T^\alpha\Psi$ and $e^\alpha_{rr}\Psi$ by $[\mathfrak{D}]^\alpha$ and $[\mathfrak{D}]^{(\alpha,r)}$, respectively, and the characters of the representations $\mathfrak{D}$ and $[\mathfrak{D}]^{(\alpha,r)}$ by χ and χ^α, respectively. Then by Eq. (2.10) and since $[\mathfrak{D}]^{(\alpha,r)}$ and $[\mathfrak{D}]^{(\alpha,s)}$ are equivalent representations of $\mathfrak{L}$, the character of $[\mathfrak{D}]^\alpha$ is simply $f^\alpha \chi^\alpha$.

Let the Clebsch–Gordan series for $[\mathfrak{D}]^{(\alpha,r)}$ with respect to $\mathfrak{L}$ be given by

$$[\mathfrak{D}]^{(\alpha,r)} = m^\alpha_1 \mathfrak{D}^{(\Lambda^{(1)})} + m^\alpha_2 \mathfrak{D}^{(\Lambda^{(2)})} + \cdots + m^\alpha_k \mathfrak{D}^{(\Lambda^{(k)})}, \tag{2.15}$$

then

$$\chi^\alpha = \sum_j m^\alpha_j \chi^{(\Lambda^{(j)})}. \tag{2.16}$$

(Here we use the highest weights Λ to specify the irreducible representations $\mathfrak{D}^{(\Lambda)}$.) From the definition of T^α we have

$$f^\alpha \chi^\alpha = \sum_i \frac{f^\alpha}{n!} \sum_\beta \varphi^\alpha_\beta \sum_{\pi \in \mathfrak{C}_\beta} D_{\pi(i),(i)}. \tag{2.17}$$

Finally, by the definition of conjugate classes for the group $\mathfrak{L}$ and by the use of such relations as

$$\sum_{i,j,k} D_{ij}(a) D_{jk}(a) D_{ki}(a) = \sum_i D_{ii}(a^3) = \chi(a^3) \tag{2.18}$$

[2] D. E. Rutherford, *Substitutional Analysis* (Edinburgh University Press, Edinburgh, 1948), p. 66.
[3] Reference 2, p. 26.
[4] Reference 2, p. 24.
[5] Reference 2, pp. 32, 50, and 53.

we have[6]

$$\chi^{\alpha} = \frac{1}{n!} \sum_{\beta} \varphi^{\alpha}_{\beta} h_{\beta} \chi_{\ell}^{v_{\ell}} \cdots \chi_2^{v_2} \chi_1^{v_1}, \tag{2.19}$$

where $\chi_j^{v_j}(a) \equiv [\chi(a^j)]^{v_j}$, φ^{α} is a character for $\mathfrak{S}_n$, and $h = n!/\ell^{v_\ell} \cdots 2^{v_2} 1^{v_1} v_\ell! \cdots v_2! v_1!$ is the number of permutations in the class $\mathfrak{C}_\beta$, $\beta = [\ell^{v_\ell} \cdots 2^{v_2} 1^{v_1}]$.

Evaluation of (2.19) for a given representation of a given group enables us to determine the constants m_j in Eq. (2.16) and hence the Clebsch–Gordan series of Eq. (2.15). We now turn to a diagrammatic method for evaluation of Eq. (2.19).

3. WEIGHT DIAGRAMS

At this point we consider some of the properties of weight diagrams, since these determine the functional dependence of the character. To each semisimple Lie group $\mathfrak{L}$ there is associated a unique root diagram, a set of points (roots) in an ℓ-dimensional space, where ℓ is the rank of $\mathfrak{L}$. The group of rotations and reflections in this space generated by reflections through hyperplanes normal to the roots is known as the Weyl group $\mathfrak{W}$.

To each irreducible representation (IR) of $\mathfrak{L}$ there is associated a weight diagram, a collection of points (weights) in the same dimensional space as above which have the following properties[7,8]:

(1) To each weight M there is assigned a positive integer multiplicity γ_{M}.

(2) The weight diagram is invariant under transformation by any element S of $\mathfrak{W}$, i.e., $\gamma_{\mathrm{M}} = \gamma_{S\mathrm{M}}$.

(3) If Λ designates the highest weight of the representation, then $\gamma_{\Lambda} = 1$.

(4) For arbitrary x^i $(i = 1, \cdots, \ell)$

$$\sum_{S\in\mathfrak{W}} \sum_{\mathrm{M}} \gamma_{\mathrm{M}} \xi_S \exp\left[\sum_{i=1}^{\ell} (S(\mathrm{M}+\delta))_i x^i\right] = \sum_{S\in\mathfrak{W}} \xi_S \exp\left[\sum_{i=1}^{\ell} (S(\Lambda+\delta))_i x^i\right], \tag{3.1}$$

$$\xi_S = \begin{cases} +1 \text{ if } S \text{ is a rotation,} \\ -1 \text{ if } S \text{ is a rotation-reflection;} \end{cases}$$

$$\delta = \tfrac{1}{2} \sum_{\alpha>0} \alpha,$$

i.e., δ is equal to one-half the (vector) sum of the positive roots. These four properties serve to specify all the γ_{M} given Λ.

The character χ of a representation of $\mathfrak{L}$ is a function of the conjugation classes of $\mathfrak{L}$, i.e., $\chi(a) = \chi(b)$ if $a = cbc^{-1}$, $a, b, c \in \mathfrak{L}$. The conjugation class is labeled by a suitable set of real parameters ϕ^i $(i = 1, \cdots, \ell)$ with $\phi^i = 0$ for the identity element. In terms of the ϕ^i the character for the element $a \in \mathfrak{L}$ in a IR of highest weight Λ may be written as[7,8]

$$\chi(a) = \chi(\phi) = \sum_{\mathrm{M}} \gamma_{\mathrm{M}} \exp\left[i \sum_{j=1}^{l} \mathrm{M}_j \phi^j\right] = \frac{\sum_{S\in\mathfrak{W}} \xi_S \exp\left[i \sum_{j=1}^{l} (S(\Lambda+\delta))_j \phi^j\right]}{\sum_{S\in\mathfrak{W}} \xi_S \exp\left[i \sum_{j=1}^{l} (S\delta)_j \phi^j\right]}. \tag{3.2}$$

The character of the product representation $\mathfrak{D}^{(1)} \otimes \mathfrak{D}^{(2)}$ is given by

$$\begin{aligned} \chi(a) &= \chi^{(1)}(a)\chi^{(2)}(a) \\ &= \sum_{\mathrm{M}'} \gamma^{(1)}_{\mathrm{M}'} \exp\left[i \sum_{j=1}^{l} \mathrm{M}'_j \phi_j\right] \sum_{\mathrm{M}''} \gamma^{(2)}_{\mathrm{M}''} \exp\left[i \sum_{j=1}^{l} \mathrm{M}''_j \phi^j\right] \\ &= \sum_{\mathrm{M}'\mathrm{M}''} \gamma^{(1)}_{\mathrm{M}'} \gamma^{(2)}_{\mathrm{M}''} \exp\left[i \sum_{j=1}^{l} (\mathrm{M}' + \mathrm{M}'')_j \phi^j\right] \\ &= \sum_{\mathrm{M}} \gamma_{\mathrm{M}} \exp\left[i \sum_{j=1}^{l} \mathrm{M}_j \phi^j\right], \\ \gamma_{\mathrm{M}} &= \sum_{\mathrm{M}'} \gamma^{(1)}_{\mathrm{M}'} \gamma^{(2)}_{(\mathrm{M}-\mathrm{M}')}, \end{aligned} \tag{3.3}$$

where, in fact, the $\mathfrak{D}^{(1)}$ and $\mathfrak{D}^{(2)}$ need not be irreducible. In general, the weight diagram associated with Eq. (3.3), i.e., the set of weights with multiplicities γ_{M}, does not correspond to an IR, but is the "sum" over weight diagrams of IR's of $\mathfrak{L}$.

Given a reducible representation we can determine its decomposition into IR's by "contracting" its weight diagram into a collection of highest weights.[9] By contraction we mean that each weight M with multiplicity γ_{M} is replaced by a weight $\mathrm{N} = S(\mathrm{M}+\delta) - \delta$ with multiplicity $\gamma_{\mathrm{N}} = \xi_S \gamma_{\mathrm{M}}$, where S is chosen so that N is a dominant weight. If $\mathrm{M} = \mathrm{N}$ for any S, then that point is simply discarded. Of course, several points M may contract to the same point N in which case γ_{N} is the sum over the various $\xi_S(\mathrm{M})\gamma_{\mathrm{M}}$. Contracting a weight diagram for an IR yields a diagram in which the only point with nonzero multiplicity is the highest weight Λ with $\gamma_{\Lambda} = 1$ (cf. properties 3 and 4 above); the other multiplicities have all summed to zero. Contracting the weight diagram corresponding to a reducible representation yields a set of points, each the highest weight of an IR occurring in the CG series. Moreover, the multiplicity of each point is the coefficient m_j of the corresponding IR

[6] J. S. Lomont, *Applications of Finite Groups* (Academic Press Inc., New York, 1959), p. 267.

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

[8] N. Jacobson, *Lie Algebras* (Interscience Publishers, Inc., New York, 1962), Chap. VIII.

[9] Our contracted diagrams are the same as the girdle diagrams of Behrends *et al.*[7] except that we use one point for each set of w points in a girdle diagram, where w is the order of the Weyl group $\mathfrak{W}$.

in the CG series. For example, in

$$\mathfrak{D}^{(1)} \otimes \mathfrak{D}^{(2)} = \sum_j m_j \mathfrak{D}^{(\Lambda^{(j)})}, \tag{3.4}$$

the m_j are equal to the multiplicities of the points of the weight diagram corresponding to the character given by Eq. (3.3). Note that for arbitrary x the contraction process leaves the value of

$$\sum\sum \gamma_{\mathrm{M}} \xi_S \exp\left[S(\mathrm{M} + \delta) \cdot x\right]$$

unchanged. No information is lost in the contraction process in the sense that the uncontracted weight diagram with weights M may be reconstructed from any contracted weight diagram with weights N by requiring $\gamma_{\mathrm{M}} = \gamma_{S\mathrm{M}}$ for all $S \in \mathfrak{W}$ and

$$\sum_S \sum_{\mathrm{M}} \gamma_{\mathrm{M}} \xi_S \exp\left[S(\mathrm{M} + \delta) \cdot x\right] = \sum_S \sum_{\mathrm{N}} \gamma_{\mathrm{N}} \xi_S \exp\left[S(\mathrm{N} + \delta) \cdot x\right]. \tag{3.5}$$

For the purpose of determining Clebsch–Gordan series, an explicit construction of the ϕ^i is unnecessary since all that is needed is the functional dependence on the ϕ^i and this uniquely specified by weight diagrams.

A simplification is introduced into the determinations of the γ_{N} in Eq. (3.3) if one uses the contracted multiplicities of either $\mathfrak{D}^{(1)}$ or $\mathfrak{D}^{(2)}$ (but not both) rather than the uncontracted multiplicities. If $\mathfrak{D}^{(1)}$ is irreducible, then as an intermediate step we have a diagram with multiplicities $\gamma_{\Lambda+\mathrm{M}} = \gamma_{\Lambda}^{(1)} \gamma_{\mathrm{M}}^{(2)} = \gamma_{\mathrm{M}}^{(2)}$. This diagram is neither contracted nor uncontracted, i.e., invariant under $\mathfrak{W}$, but yields the same contracted diagram as is obtained by the contraction of the diagram with[10]

$$\gamma_{\mathrm{M}} = \sum_{\mathrm{M}'} \gamma_{\mathrm{M}'}^{(1)} \gamma_{(\mathrm{M}-\mathrm{M}')}^{(2)}.$$

In order to obtain the weight diagram of χ^a of Eq. (2.19) it is necessary to discuss the "weight diagram" to be associated with $\chi_k(a) \equiv \chi(a^k)$. This is obtained by considering

$$\chi_k(a) = \chi_k(\phi) = \chi(k\phi) = \sum_{\mathrm{M}} \gamma_{\mathrm{M}} \exp\left[i \sum_{j=1}^{l} (k\mathrm{M})_j \phi^j\right], \tag{3.6}$$

where γ_{M} are multiplicities of the representation $\mathfrak{D}$. This new set of points is similar to the weight diagram of $\mathfrak{D}$ but its linear dimensions have been multiplied by a factor k. When this set of points is contracted, the resulting γ_{N} will be both positive and negative integers. It may not be a true weight diagram, but there is a one-to-one correspondence between such diagrams and the functional dependence of $\chi_k(a)$. In the construction of the contracted weight diagrams associated with the various terms of Eq. (2.19) one multiplies together in the manner described above one contracted and one uncontracted diagram, then contracts this product before multiplying again by an uncontracted diagram. Finally, the contracted diagrams for each β are added together (in the sense of adding multiplicities). The resulting multiplicities will turn out to be positive integers even though negative nonintegers were involved in the intermediate steps. These multiplicities are the coefficients m_j in Eq. (2.15) which we have set out to find.

[10] A. J. Macfarlane, L. O'Raifeartaigh, and P. S. Rao, Syracuse University preprint.

4. GROUPS OF RANKS ONE AND TWO

Let us now illustrate the technique by considering tensor products of definite symmetry type for the compact rank-two groups $SU(3)$, G_2, and $R(5)$ or $USp(4)$. Afterwards we consider the rank-one groups $SU(2)$ and $R(3)$.

In the case of rank-two groups it is convenient to consider two parallel sets of Weyl planes: one set (the usual set) having a common intersection at the origin and the second having a common intersection at $-\delta$ [cf. Eq. (3.1)]. The points M and $\mathrm{N} = S(\mathrm{M} + \delta) - \delta$, $S \in \mathfrak{W}$ are related by successive reflections through the second set of Weyl planes.

Let us consider first the symmetric and antisymmetric products of two $SU(3)$ octets. For $n = 2$ Eq. (2.19) becomes simply

$$\begin{aligned} \chi^{[2]} &= \tfrac{1}{2}(\chi^2 + \chi_2), \\ \chi^{[1^2]} &= \tfrac{1}{2}(\chi^2 - \chi_2), \end{aligned} \tag{4.1}$$

where χ is the character of the octet representation, $\chi^{[2]}$ is the character of the symmetric product, $\chi^{[1^2]}$ is the character of the antisymmetric product, and χ^2, χ_2 are the same as in Eq. (2.19). In Fig. 1 we give the uncontracted and contracted weight diagrams for χ, χ^2, χ_2, $\chi^{[2]}$, and $\chi^{[1^2]}$. Here, as in the succeeding figures, the displaced Weyl planes are indicated by dashed lines and the outlines of the root diagrams by dotted lines. The dashed and dotted lines thus serve to indicate the origin and the scale for each diagram. Figure 1 shows that the $\{1\}$ and $\{27\}$ representations appear only in the symmetric product, that the $\{10\}$ and $\{\overline{10}\}$ appear only in the antisymmetric product, and that there is one symmetric $\{8\}$ and one antisymmetric $\{8\}$, all of which is well known.

As an application, consider an exact $SU(3)$ theory in which a resonance decays strongly into two spin-zero mesons belonging to the same octet. If this resonance is somehow known to be a unitary singlet or 27-plet, then we see that the generalized Pauli

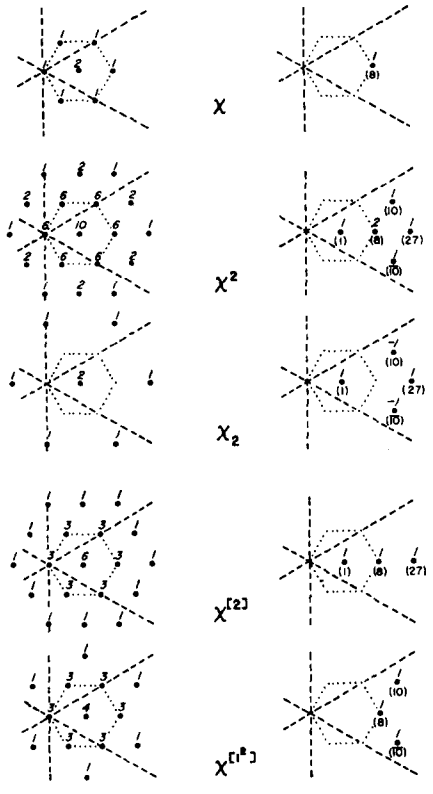


FIG. 1. Weight diagrams for χ, χ^2, χ_2, $\chi^{[2]}$, and $\chi^{[1^2]}$ for the octet (adjoint) representation of $SU(3)$. The uncontracted diagrams are shown on the left, the contracted diagrams on the right. The symmetric and antisymmetric twofold products correspond to $\chi^{[2]}$ and $\chi^{[1^2]}$, respectively.

principle requires that it must have $J^P = \text{even}^+$, while if it is known to be a unitary decuplet, then it must have $J^P = \text{odd}^-$.

The CG series for $\{8 \otimes 8 \otimes 8\}$ can be partitioned into four sets of IR's corresponding to the representations [3], [2, 1], [2, 1], and $[1^3]$ of $\mathfrak{S}_n$. In Fig. 2(a) we

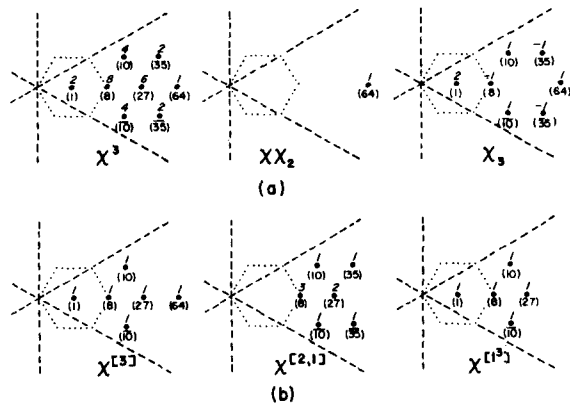


FIG. 2. (a) Contracted weight diagrams for χ^3, $\chi\chi_2$, and χ_3 for the octet representation of $SU(3)$. (b) Contracted weight diagrams giving the Clebsch–Gordan series for the α-symmetrized threefold products of octets (α = [3], [2, 1], $[1^3]$).

show the contracted weight diagrams (cwd) for χ^3, $\chi\chi_2$, χ_3. Note that the cwd for $\chi\chi_2$ consists only of a single point. Use of the character formulas [Eq. (2.19)] for $n = 3$:

$$\begin{aligned} \chi^{[3]} &= \tfrac{1}{6}(\chi^3 + 2\chi_3 + 3\chi\chi_2), \\ \chi^{[2,1]} &= \tfrac{1}{3}(\chi^3 - \chi_3), \\ \chi^{[1^3]} &= \tfrac{1}{6}(\chi^3 + 2\chi_3 - 3\chi\chi_2), \end{aligned} \tag{4.2}$$

and of the diagrams of Fig. 2(a) yields the contracted weight diagrams of Fig. 2(b).

For the four- and five-fold direct products of $SU(3)$ octets we obtain the partitioning shown in Tables I and II. Two of the diagrams used in this computation are shown in Fig. 3. In the first columns of Tables I and II we list the representations occurring in the CG series (with respect to $\mathfrak{L}$) of the four- and five-fold direct products, respectively. These representations

TABLE I. Coefficients in the Clebsch–Gordan series for the α-symmetrized and the unsymmetrized fourfold direct products of $SU(3)$ octets.

Representation		Symmetry type $\binom{\alpha}{f_\alpha}$					
(p_1, p_2)	Dimension	[4] 1	[3, 1] 3	[2, 2] 2	[2, 1²] 3	[1⁴] 1	Unsymmetrized
(0, 0)	{1}	1	...	2	1	...	8
(1, 1)	{8}	2	4	2	4	2	32
(2, 2)	{27}	2	4	4	3	2	33
(3, 3)	{64}	1	2	1	1	...	12
(4, 4)	{125}	1	...	...	...	...	1
(3, 0)	{10}	...	3	1	3	...	20
(0, 3)	$\{\overline{10}\}$	...	3	1	3	...	20
(4, 1)	{35}	1	2	1	2	...	15
(1, 4)	$\{\overline{35}\}$	1	2	1	2	...	15
(5, 2)	{81}	...	1	...	...	...	3
(2, 5)	$\{\overline{81}\}$	...	1	...	...	...	3
(6, 0)	{28}	...	...	1	...	...	2
(0, 6)	{28}	...	...	1	...	...	2

TABLE II. Coefficients in the Clebsch–Gordan series for the α-symmetrized and the unsymmetrized five-fold direct products of $SU(3)$ octets.

Representation		Symmetry type $\binom{\alpha}{f^\alpha}$							
(p_1, p_2)	Dimension	[5] 1	[4, 1] 4	[3, 2] 5	[3, 1^2] 6	[2^2, 1] 5	[2, 1^3] 4	[1^5] 1	Unsymmetrized
(0, 0)	{1}	1	1	1	2	1	1	1	32
(1, 1)	{8}	2	5	6	7	6	5	1	145
(2, 2)	{27}	2	7	8	9	7	7	1	180
(3, 3)	{64}	2	5	5	4	3	2	...	94
(4, 4)	{125}	1	2	1	4	...	...	...	20
(5, 5)	{216}	1	...	...	...	...	...	...	1
(3, 0)	{10}	1	1	4	5	4	...	1	100
(0, 3)	$\{\overline{10}\}$	1	1	4	5	4	...	1	100
(4, 1)	{35}	1	4	5	5	4	2	...	100
(1, 4)	$\{\overline{35}\}$	1	4	5	5	4	2	...	100
(5, 2)	{81}	1	2	2	2	1	...	...	36
(2, 5)	$\{\overline{81}\}$	1	2	2	2	1	...	...	36
(6, 3)	{154}	...	1	...	...	...	...	...	4
(3, 6)	$\{\overline{154}\}$	...	1	...	...	...	...	...	4
(6, 0)	{28}	...	1	1	1	1	...	...	20
(0, 6)	$\{\overline{28}\}$	...	1	1	1	1	...	...	20
(7, 1)	{80}	...	...	1	...	...	...	...	5
(1, 7)	$\{\overline{80}\}$	...	...	1	...	...	...	...	5

are specified in two different ways. The first way is by giving (p_1, p_2), where p_1 is the number of boxes in the first row of the corresponding Young shape minus the number in the second row, and p_2 is the difference in the lengths of the second and third rows. The second way is giving the dimension of the representation. For example, the octet representation is specified by $(p_1, p_2) = (1, 1)$ and by {8}.

Consider now the group G_2. The symmetric and antisymmetric direct products of the 14-dimensional

TABLE III. Coefficients in the Clebsch–Gordan series for the α-symmetrized and the unsymmetrized three-fold direct products of the 14-dimensional adjoint representation of G_2.

Representation	Symmetry type $\binom{\alpha}{f^\alpha}$			
	[3] 1	[2, 1] 2	[1^3] 1	Unsymmetrized
{1}	...	...	1	1
{7}	1	...	...	1
{27}	...	1	1	3
$\{77\}_1$	1	1	1	4
{182}	...	...	1	1
{14}	1	2	...	5
{64}	...	1	...	2
{189}	1	1	...	3
{448}	...	1	...	2
$\{77\}_2$	...	1	1	3
{273}	1	...	...	1

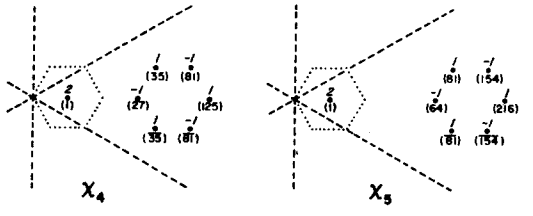


FIG. 3. Contracted weight diagrams used in the calculation of the Clebsch–Gordan series of the α-symmetrized four- and five-fold products of $SU(3)$ octets.

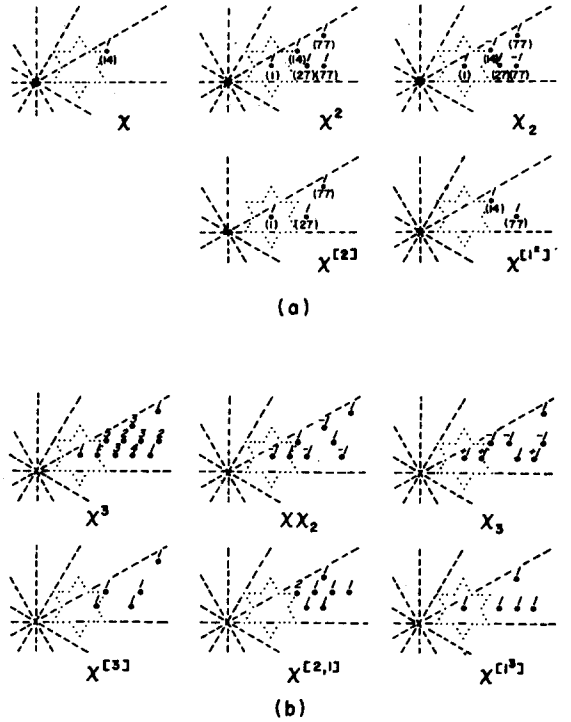


FIG. 4. Contracted weight diagrams which show the calculation of the Clebsch–Gordan series for the α-symmetrized (a) two-fold, (b) three-fold products of the 14-dimensional (adjoint) representation of G_2.

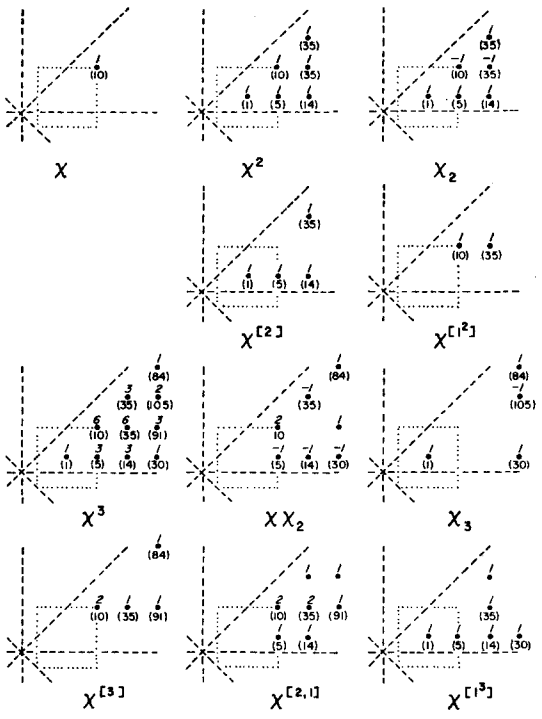


FIG. 5. Contracted weight diagrams which show the calculation of the Clebsch–Gordan series for the α-symmetrized two- and three-fold products of the four-dimensional spinor representation of $R(5)$.

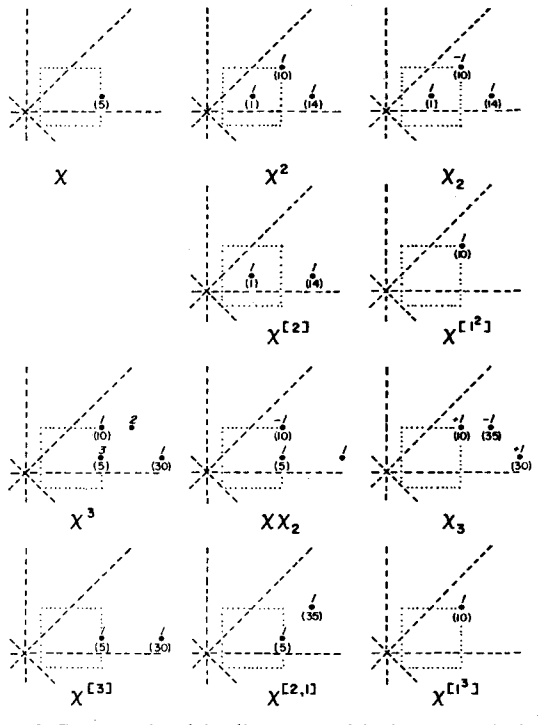


FIG. 6. Contracted weight diagrams which show the calculation of the Clebsch–Gordan series for the α-symmetrized two- and three-fold products of the five-dimensional representation of $R(5)$.

(adjoint) representation with itself contain the representations $\{1\} + \{27\} + \{77\}_2$ and $\{14\} + \{77\}_1$, respectively. The contracted weight diagrams for χ, χ^2, χ_2, $\chi^{[2]}$, and $\chi^{[1^2]}$ are shown in Fig. 4(a). For the three-fold product of the adjoint representation we obtain Table III from Fig. 4(b).

We turn now to the groups $R(5)$ and $USp(4)$. In Figs. 5, 6, and 7 we have the diagrams for the symmetrized two- and three-fold products of the four-, five-, and ten-dimensional representations, respectively.

Having treated the rank-two groups, it is a simple matter to treat the rank-one groups, $SU(2)$ and $R(3)$. Since these groups have one-dimensional weight diagrams, the CG series of the symmetrized n-fold products of any low-dimensional representation can be obtained very quickly by this method. In Fig. 8 we illustrate this for the two-, three-, and four-fold products of the spin $\frac{3}{2}$ representation.

5. GROUPS OF RANK $\ell \geq 3$

In the case of rank $\ell \geq 3$, the weight diagrams have three or more dimensions and consequently the graphic procedures just illustrated are not suitable. However, if one chooses to work in the appropriate coordinate systems, namely the systems discussed by Racah,[11] the corresponding operations on the weight components M_i are not difficult. For the groups $SU(\ell + 1)$ we imbed the weight diagram in an $(\ell + 1)$-dimensional space and impose the condition that

$$\sum_{i=1}^{\ell+1} M_i = 0 \tag{5.1}$$

so that there are only ℓ independent coordinates. The Weyl group $\mathfrak{W}$ is the group of permutations of the components of M. We have $\xi_S = +1$ (-1) for even (odd) permutations. The M_i are fractions with denominator $\ell + 1$ and the differences $(M_i - M_j)$ are all integers. The δ_j are given by $\delta_j = \frac{1}{2}\ell + 1 - j$. Such a treatment of $SU(3)$ would label the $SU(3)$ root diagram as in Fig. 9.

For the groups $R(2\ell + 1)$ and $USp(2\ell)$, $\mathfrak{W}$ is generated by the permutations of the components M_i $(i = 1, \cdots \ell)$ and by the changes of sign of any number of the M_i. We have $\xi_S = +1$ (-1) for an even (odd) permutation and any number of changes of sign. Here all ℓ components of M are independent. Finally for the groups $R(2\ell)$, $\mathfrak{W}$ is the same as for $R(2\ell + 1)$ except that only an even number of changes of sign are allowed.

By this procedure we find for $SU(4)$ that

$$\begin{aligned} \{15 \otimes 15\}_S &= \{1\} + \{15\} + \{20\} + \{84\}, \\ \{15 \otimes 15\}_A &= \{15\} + \{45\} + \{\overline{45}\}. \end{aligned} \tag{5.2}$$

[11] G. Racah, Ergeb. Exakt. Naturwiss. 37, 28 (1965).

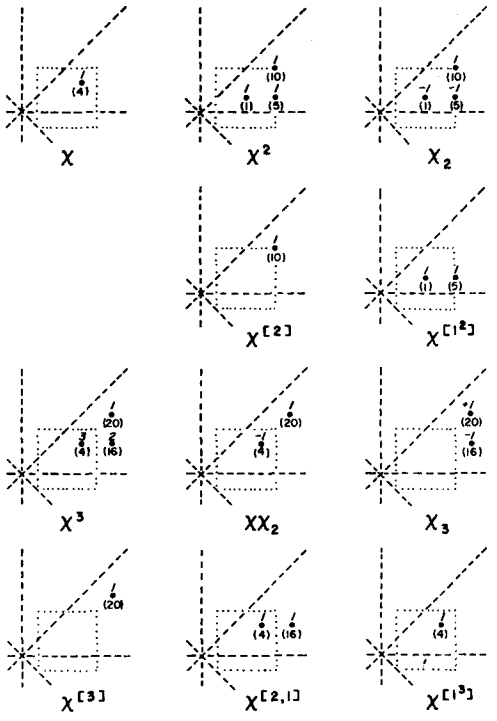


FIG. 7. Contracted weight diagrams which show the calculation of the Clebsch–Gordan series for the α-symmetrized two- and three-fold products of the ten-dimensional (adjoint) representation of $R(5)$.

For the three-fold product of the adjoint representation $[(p_1, p_2, p_3) = (1, 0, 1)]$ the coefficients in the CG series are given in Table IV. Again the representations in the CG series are specified by both their Young shape and their dimensionality.

For $SU(6)$ we find, for example,

$$\{35 \otimes 35\}_S = \{1\} + \{35\} + \{189\} + \{405\},$$
$$\{35 \otimes 35\}_A = \{35\} + \{280\} + \{\overline{280}\}. \tag{5.3}$$

For the three-fold direct product of {35}'s we have

TABLE IV. Coefficients in the Clebsch–Gordan series for the α-symmetrized and the unsymmetrized three-fold direct products of the adjoint representation of $SU(4)$.

Representation		Symmetry type $\binom{\alpha}{f^\alpha}$			
(p_1, p_2, p_3)	Dimension	[3] 1	[2, 1] 2	[1³] 1	Unsymmetrized
(0, 0, 0)	{1}	1	...	1	2
(1, 0, 1)	{15}	2	3	1	9
(0, 2, 0)	{20}	...	2	1	5
(4, 0, 0)	{35}	...	...	1	1
(0, 0, 4)	$\{\overline{35}\}$	...	...	1	1
(2, 1, 0)	{45}	1	2	1	6
(0, 1, 2)	$\{\overline{45}\}$	1	2	1	6
(2, 0, 2)	{84}	1	2	1	6
(1, 2, 1)	{175}	1	1	1	4
(3, 1, 1)	{256}	...	1	...	2
(1, 1, 3)	$\{\overline{256}\}$	...	1	...	2
(3, 0, 3)	{300}	1	...	...	1

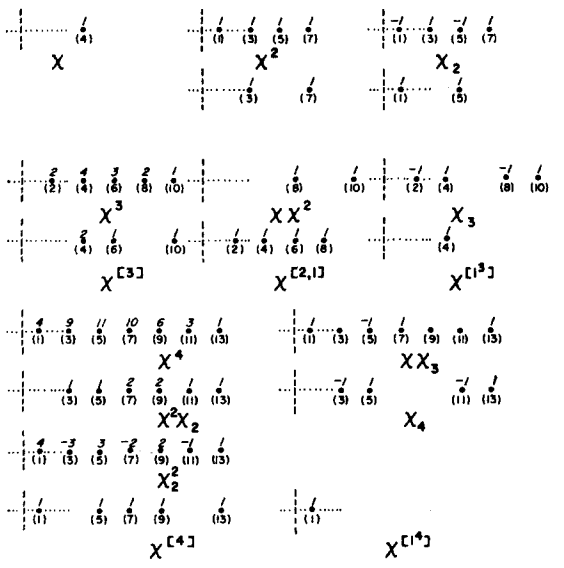


FIG. 8. Contracted weight diagrams which show the calculation of the Clebsch–Gordan series for the α-symmetrized two-, three-, and four-fold products of the spin $\frac{3}{2}$ representation of $SU(2)$.

Table V. Finally, in Table VI we give the coefficients in the CG series of the α-symmetrized two-, three-, and four-fold direct products of the 56-dimensional representation of $SU(6)$. The representations appearing in the CG series are labeled by $(p_1, p_2, p_3, \cdots)$,

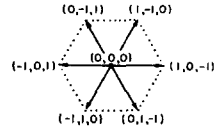


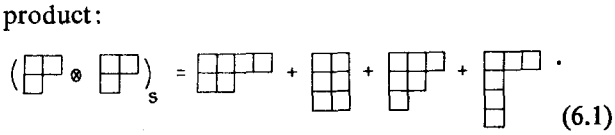
FIG. 9. Root diagram for $SU(3)$ imbedded in a three-dimensional space showing components of the root vectors.

where p_i is the number of boxes in the ith row of the corresponding Young shape minus the number in the $(i + 1)$th row. Zeros on the right have been dropped.

6. YOUNG SHAPES

For the nonexceptional Lie groups other methods for obtaining the CG series of an α-symmetrized n-fold direct product representation have been discussed.[12] These methods are based on the fact that the decomposition of the representation $[\mathfrak{D}]^{(\alpha,r)}$ of $SU(\ell + 1)$ as expressed in terms of Young shapes (partitions) is independent of ℓ for sufficiently large ℓ, and for small ℓ the only difference which arises is that one must delete those shapes which have more than $\ell + 1$ rows.

For example, we have the following symmetrized product:



(6.1)

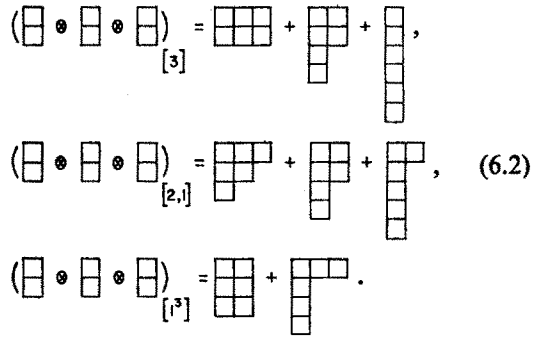
[12] D. E. Littlewood, *The Theory of Group Characters* (Oxford University Press, Oxford, England, 1950), 2nd ed., Appendix A and references cited therein, in particular, J. A. Todd, Proc. Cambridge Phil. Soc. 45, 328 (1949).

TABLE V. Coefficients in the Clebsch–Gordan series for the α-symmetrized and the unsymmetrized three-fold direct products of the adjoint representation of $SU(6)$.

Representation		Symmetry type $\binom{\alpha}{f^\alpha}$			
$(p_1, p_2, p_3, p_4, p_5)$	Dimension	[3] 1	[2, 1] 2	[1³] 1	Unsymmetrized
(0, 0, 0, 0, 0)	{1}	1	...	1	2
(1, 0, 0, 0, 1)	{35}	2	3	1	9
(0, 0, 2, 0, 0)	{175}	1	...	...	1
(0, 1, 0, 1, 0)	{189}	1	2	1	6
(2, 0, 0, 1, 0)	{280}	1	2	1	6
(0, 1, 0, 0, 2)	$\overline{\{280\}}$	1	2	1	6
(2, 0, 0, 0, 2)	{405}	1	2	1	6
(3, 0, 1, 0, 0)	{840}	...	...	1	1
(0, 0, 1, 0, 3)	$\overline{\{840\}}$	...	...	1	1
(1, 1, 1, 0, 0)	{896}	...	1	...	2
(0, 0, 1, 1, 1)	$\overline{\{896\}}$	...	1	...	2
(3, 0, 0, 0, 3)	{2695}	1	...	...	1
(3, 0, 0, 1, 1)	{3200}	...	1	...	2
(1, 1, 0, 0, 3)	{3200}	...	1	...	2
(1, 1, 0, 1, 1)	{3675}	1	1	1	4

For $SU(4)$, all four terms contribute, but for $SU(3)$, the last term must be deleted to give $\{8 \otimes 8\}_S = \{27\} + \{1\} + \{8\}$. Similarly for $SU(2)$, only the first term on the right-hand side has few enough rows to contribute. In this case, we have simply $\{2 \otimes 2\}_S = \{3\}$.

An analogous observation pertains for three- and higher-fold products. As an example of this, consider the $SU(\ell + 1)$ representations which have the same Young shape as the $\{\bar{3}\}$ representation of $SU(3)$. We have



(6.2)

These diagrams are valid for any $SU(\ell + 1)$ provided one deletes those diagrams which do not apply to the particular unitary group under consideration.

For the particular case of finding the CG series for the symmetric and antisymmetric twofold direct products of a $SU(3)$ representation $\mathfrak{D}$, the rules can be quite simply stated in terms of the Young shapes which occur in the CG series. This is because for $SU(3)$ the coefficients which appear in the CG series corresponding to χ_2 take on only the values 1, 0, and

TABLE VI. Coefficients in the Clebsch–Gordan series for the α-symmetrized and the unsymmetrized two-, three-, and four-fold direct products of the 56-dimensional representation of $SU(6)$.

Representation		Symmetry type $\binom{\alpha}{f^\alpha}$		
(p_1, p_2)	Dimension	[2] 1	[1²] 1	Unsymmetrized
(6, 0)	{462}	1	...	1
(4, 1)	{1050}	...	1	1
(2, 2)	{1134}	1	...	1
(0, 3)	{4903}	...	1	1

Representation		Symmetry type $\binom{\alpha}{f^\alpha}$			
(p_1, p_2, p_3)	Dimension	[3] 1	[2, 1] 2	[1³] 1	Unsymmetrized
(9, 0, 0)	{980}	1	...	...	1
(7, 1, 0)	{5720}	...	1	...	2
(5, 2, 0)	{8910}	1	1	...	3
(3, 3, 0)	{9240}	1	1	1	4
(1, 4, 0)	{5880}	...	1	...	2
(6, 0, 1)	{6160}	...	...	1	1
(4, 1, 1)	{11550}	...	1	...	2
(2, 2, 1)	{11340}	...	1	1	3
(0, 3, 1)	{4704}	1	...	...	1
(3, 0, 2)	{6000}	1	...	...	1
(1, 1, 2)	{5880}	...	1	...	2
(0, 0, 3)	{980}	...	...	1	1

Representation		Symmetry type $\binom{\alpha}{f^\alpha}$				
(p_1, p_2, p_3, p_4)	[4] 1	[3, 1] 3	[2²] 2	[2, 1²] 3	[1⁴] 1	Unsymmetrized
(12, 0, 0, 0)	1	...	...	...	...	1
(10, 1, 0, 0)	...	1	...	...	...	3
(8, 2, 0, 0)	1	1	1	...	...	6
(6, 3, 0, 0)	1	2	...	1	...	10
(4, 4, 0, 0)	1	1	2	1	...	11
(2, 5, 0, 0)	...	2	...	1	...	9
(0, 6, 0, 0)	1	...	1	...	1	4
(9, 0, 1, 0)	...	...	...	1	...	3
(7, 1, 1, 0)	...	1	1	1	...	8
(5, 2, 1, 0)	...	2	1	2	1	15
(3, 3, 1, 0)	1	2	1	2	1	16
(1, 4, 1, 0)	...	1	1	2	...	11
(6, 0, 2, 0)	1	1	1	...	...	6
(4, 1, 2, 0)	1	2	1	2	...	15
(2, 2, 2, 0)	1	2	2	1	1	15
(0, 3, 2, 0)	...	1	...	1	...	6
(3, 0, 3, 0)	...	1	...	2	1	10
(1, 1, 3, 0)	...	1	1	1	...	8
(0, 0, 4, 0)	1	...	...	...	...	1
(8, 0, 0, 1)	...	...	...	...	1	1
(6, 1, 0, 1)	...	...	...	1	...	3
(4, 2, 0, 1)	...	...	1	1	1	6
(2, 3, 0, 1)	...	1	...	1	...	6
(0, 4, 0, 1)	...	...	1	...	1	3
(5, 0, 1, 1)	...	1	...	...	...	3
(3, 1, 1, 1)	...	1	1	1	...	8
(1, 2, 1, 1)	1	1	...	1	...	7
(2, 0, 2, 1)	...	...	1	1	1	6
(0, 1, 2, 1)	...	1	...	...	...	3
(4, 0, 0, 2)	1	...	...	...	...	1
(2, 1, 0, 2)	...	1	...	...	...	3
(0, 2, 0, 2)	...	...	1	...	...	2
(1, 0, 1, 2)	...	...	...	1	...	3
(0, 0, 0, 3)	...	...	...	...	1	1

-1. By Eq. (4.1), the coefficient of a given representation (p_1, p_2) in the CG series of χ_2 must equal zero if (p_1, p_2) occurs an even number of times in χ^2, and this coefficient must be ± 1 if (p_1, p_2) occurs an odd number of times in χ^2. Since a calculation of the CG series of $[\mathfrak{D}]^2 = \mathfrak{D} \otimes \mathfrak{D}$ by manipulating Young shapes is well known,[13] the only problem remaining is to determine the sign of each nonzero term in χ_2. This sign may be determined by the following rule.

Let k be the total number of boxes occurring in the odd-numbered rows of a Young shape which corresponds to a nonzero term in χ^2, i.e., in the CG series of $\mathfrak{D} \otimes \mathfrak{D}$. Then for any $SU(\ell + 1)$, the corresponding coefficient in χ_2 can only be positive (negative) if k is even (odd).[14] The number k, in turn, is even or odd according as $p_1 + p_2 + p_5 + p_6 + \cdots$ is even or odd, i.e.,

$$k \equiv \sum_{i=0}^{\infty} (p_{4i+1} + p_{4i+2}) \pmod 2. \qquad (6.3)$$

Thus for $SU(3)$, the numbers $(p_1 + p_2)$ together with χ^2 determine χ_2 and hence, by Eq. (4.1), determine the CG series of $[\mathfrak{D}]^{[2]}$ and $[\mathfrak{D}]^{[1^2]}$. For example, in Eq. (6.1) the first, second, and fourth terms on the right-hand side each appear once in χ^2, and thus they have coefficients either $+1$ or -1 in χ_2. In each of these terms the number k of boxes in the first and third rows is four, and hence all three have coefficient plus one in χ_2.

[13] Littlewood, Ref. 12, pp. 94–98.

[14] This is implicit in the article by Todd cited in Ref. 12.

The above rule can be extended as follows. For any $SU(\ell + 1)$ the coefficient of $(p_1, p_2, p_3, \cdots)$ in the CG series corresponding to χ_2 is zero unless the number of odd-numbered rows which contain an odd number of boxes is equal to the number of even-numbered rows which contain an odd number of boxes.[14] For example, in the third term on the right-hand side of Eq. (6.1), there are two odd-numbered rows (rows one and three) which each contain an odd number of boxes and only one even-numbered row (row two) which contains an odd number of boxes. Thus, by this rule its coefficient in χ_2 is zero and it appears an equal number of times in $[\mathfrak{D}]^{[2]}$ and $[\mathfrak{D}]^{[1^2]}$. This is consistent with the fact that this term appears an even number of times (twice) in χ^2.

For any $SU(\ell + 1)$ with initial representation $(p_1, p_2, 0, 0, \cdots)$ and for $SU(4)$ and $SU(5)$, if in the set of p_i which specify the initial representation $\mathfrak{D}$ there are only one or two p_i which are nonzero, then again the coefficients in χ_2 can only be $+1$, 0, and -1, and the above rules uniquely determine the CG series for $[\mathfrak{D}]^{[2]}$ and $[\mathfrak{D}]^{[1^2]}$. In the general case the procedure is considerably more complicated.

ACKNOWLEDGMENTS

The author wishes to thank Professor S. P. Rosen for suggesting this problem. He is indebted to him as well as to L. O'Raifeartaigh, J. Belinfante, S. M. Harris, and J. Nagel for useful discussions.

Complete Excitation Spectra*

J. T. ANDERSON†
Physics Department, University of California, Riverside, California

(Received 2 August 1966)

A type of nonhomogeneous integral equation frequently encountered in interaction problems is solved by means of the Heaviside expansion theorem. The energy spectrum is obtained in the usual way from the solutions of the dispersion equation for the excitations, $D(\mathbf{k}, s) = 0$. The momentum spectrum is obtained as the solutions of $\partial D/\partial \mathbf{k} = 0$, which is shown to be also a form of the Cauchy–Riemann equations. It is proved that the existence of a solution of $\partial D/\partial \mathbf{k} = 0$ is a necessary condition for the existence of an analytic region on the momentum plane. It is also proved that the existence of a solution of $dD/d\mathbf{k} = 0$ and $\partial D/\partial s = 0$ is a necessary and sufficient condition for the existence of analytic regions in both variables, provided the group velocity is finite and continuous. As an example, two linearized self-consistent equations with arbitrary coupling are solved.

INTRODUCTION

MANY problems in statistical mechanics and field theory have formal solutions in the form of integral equations in which the kernel is a function of a space–time interval, and often such equations may be solved explicitly by convolution integrals which eliminate one set of variables; that is, one can express the solutions in terms of the initial and boundary values.[1] The result, however, is another integral equation involving a difference kernel in the form of a propagator or dispersion relation appearing in the denominator of the integrand. Examples are found in the calculations of the correlation energy of the electron gas.[2] Similar equations also result from self-consistent field solutions of an equation of motion.[3,4]

An equation of this type yields the energy spectrum of the excitations as the solutions of the dispersion equation but in general the momentum spectrum can be found only in certain regions. In the case of particle excitations in the Fermi gas, the number distribution of particles in various momentum states has been found in the high-density region by Kulik[5] and in the low-density region by Belyakov[6]; but the momentum spectrum itself has not been found. In this paper we show that in the case of complex energy and momentum the complete spectra, including the collective excitations, may be obtained provided the derivative of the dispersion equation with respect to either the energy or the momentum has nontrivial solutions.

1. FORMAL SOLUTION

Consider the nonhomogeneous equation

$$\psi(\mathbf{x}, t) = \phi(\mathbf{x}, t) + \lambda \iint d\mathbf{x}_0\, dt_0 K(\mathbf{x}_0, t_0; \mathbf{x}, t)\psi(\mathbf{x}_0, t_0). \quad (1)$$

In general the kernel is a functional of the potential and a Green's function obtained from a homogeneous differential equation representing the response of the four-dimensional system to a disturbance at $\mathbf{x}_0$, t_0 and $\psi(\mathbf{x}_0, t_0)$ includes the initial and boundary conditions. A well-known example is, of course, the S-matrix equation, whose analytic properties for complex k have been discussed by Barut and Ruei.[7] Provided $K(\mathbf{x}_0, t_0; \mathbf{x}, t)$ is a difference kernel in both variables, the Fourier transforms in space and time yield

$$\Psi(\mathbf{k}, s) = \Phi(\mathbf{k}, s) + \lambda V(\mathbf{k}, s)\Psi(\mathbf{k}, s), \quad (2)$$

where we have absorbed a factor of $(2\pi)^2$ into V. Thus we find

$$\psi(\mathbf{x}, t) = \frac{1}{(2\pi)^2} \int_{C_2} ds e^{-ist} \int_{C_1} d\mathbf{k} e^{-i\mathbf{k}\cdot\mathbf{x}} \frac{\Phi(\mathbf{k}, s)}{1 - \lambda V(\mathbf{k}, s)}, \quad (3)$$

where s and k are complex variables. In order that (3) converge, the amplitude must be continuable into the s^* and k^* half-planes and we confine our discussion to integrands which satisfy these conditions, and are real on the real axes.

* Work begun in the Department of Applied Mathematics and Theoretical Physics, Cambridge University, Cambridge, England.

† Permanent address: Box 94, La Jolla, California.

[1] Solutions of integral equations by Fourier transformation is discussed in most standard works. A thorough treatment may be found in Chap. 8 of P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill Book Company, Inc., New York, 1953).

[2] See, for example, the calculations of Hubbard and of Goldstone in *The Many-Body Problem*, D. Pines, Ed. (W. A. Benjamin, Inc., New York, 1962).

[3] L. D. Landau, J. Phys. (USSR) **10**, 25 (1946).

[4] H. Ehrenreich and M. Cohen, Phys. Rev. **115**, 786 (1959).

[5] I. O. Kulik, Zh. Eksperim. i Teor. Fiz. **40**, 1243 (1961) [English transl.: Soviet Phys.—JETP **13**, 946 (1961)].

[6] V. A. Belyakov, Zh. Eksperim. i Teor. Fiz. **40**, 1210 (1961) [English transl.: Soviet Phys.—JETP **13**, 850 (1961)].

[7] A. O. Barut and K. H. Ruei, J. Math. Phys. **2**, 181 (1961).

In interaction problems λ is a variable coupling factor and if there exists a constant λ' such that

$$1 - \lambda' V(\mathbf{k}^*, s^*) \equiv 0 \tag{4}$$

for every $\mathbf{k}^*$ and s^*, then λ' are the eigenvalues of (1) and $\mathbf{k}^*$ and s^* are continuous. With variable coupling, however, (3) may still have singularities for particular values of momentum and energy, $\mathbf{k}^{*\prime}$ and $s^{*\prime}$ (the excitation eigenvalues). The solutions of (4) in the continuum yields some function, $s^* = f(\lambda', k^*)$ whereas the particular solutions of (3) for variable coupling represent the discrete part of the spectrum; but in general the momenta are arbitrary. The complete spectra may thus be found by solving the eigenvalue equation for these two conditions.

Formal solutions, including the momentum eigenvalues, may be obtained by use of the Heaviside expansion theorem for each of the four complex integrations.[8] The Heaviside theorem permits the evaluation of an integral over a meromorphic rational function by replacement of the integral operator by an inverse differential operator acting on the denominator alone. The theorem holds without factorization of the integrand provided the integrand is analytic. This principle may be used to delineate the analytic regions on the energy and momentum planes in the following way. Following Landau contours in each complex plane, the four integrations may be written in the usual way as the sum of pole terms in the analytic regions and principal value terms in the nonanalytic regions. Thus formal solutions of (3) read

$$\psi(\mathbf{x}, t) = \sum_{k^*} \lim_{k^* \to k^{*\prime}} \sum_{s^*} \lim_{s^* \to s^{*\prime}} \frac{\Phi(\mathbf{k}^*, s^*) e^{-i\mathbf{k}\cdot\mathbf{x}} e^{-ist}}{(\partial/\partial s)(\partial/\partial \mathbf{k})[1 - \lambda V(k^*, s^*)]} + P\int_{-\infty}^{+\infty} ds e^{-ist} \int_{-\infty}^{+\infty} d\mathbf{k} e^{-i\mathbf{k}\cdot\mathbf{x}} \frac{\Phi(\mathbf{k}^*, s^*)}{[1 - \lambda V(k^*, s^*)]}, \tag{5}$$

where $s^{*\prime}$ and $\mathbf{k}^{*\prime}$ satisfy

$$D(\mathbf{k}^*, s^{*\prime}) = 1 - \lambda V(\mathbf{k}^*, s^{*\prime}) \equiv 0 \tag{6}$$

and

$$\partial D(\mathbf{k}^{*\prime}, s^*)/\partial \mathbf{k} \equiv 0 \tag{7}$$

and $P\int$ is the principal value integral. The operations over $dk_1\, dk_2\, dk_3$ may be carried out in succession, writing $k_j = \pi_j + i\kappa_j$ and summing over j so that (5) holds for a momentum distribution of arbitrary shape. Of course, (5) is of little interest unless the analytic regions of $D(k^*, s^*)$ are known. In the following section we show that the solutions of (7) exactly define the analytic regions in the energy plane. The proof, which is given for a single coordinate, holds for cylindrical symmetry of the momentum distribution and may be extended to more complicated cases by iteration.

2. ANALYTIC PROPERTIES

Let $D(k^*)$ be a function of $k^* = \pi - i\kappa$. Then, if D is real on the real axis, $D(k^*) = D^*(k)$ and

$$\frac{\partial D^*}{\partial k^*} = \frac{1}{2}\left(\frac{\partial D^*}{\partial \pi} + i\frac{\partial D^*}{\partial \kappa}\right), \quad \frac{\partial D^*}{\partial k} = \frac{1}{2}\left(\frac{\partial D^*}{\partial \pi} - i\frac{\partial D^*}{\partial \kappa}\right) \tag{8}$$

The factor $\frac{1}{2}$ comes from solving k and k^* for π and κ before obtaining the partial derivatives. Now putting $D^*(k) = u(\pi, \kappa) - iv(\pi, \kappa)$, we find

$$\frac{\partial D^*}{\partial k} = \frac{1}{2}\left[\left(\frac{\partial u}{\partial \pi} - \frac{\partial v}{\partial \kappa}\right) - i\left(\frac{\partial u}{\partial \kappa} + \frac{\partial v}{\partial \pi}\right)\right] \tag{9}$$

If the Cauchy–Riemann equations are satisfied on the upper half-plane,

$$\frac{\partial D(k^*)}{\partial k^*} = \frac{1}{2}\frac{\partial D(k^*)}{\partial k} - i\frac{\partial D(k^*)}{\partial \kappa} \tag{10}$$

and

$$\partial D(k^*)/\partial k = 0, \tag{11}$$

and there are similar equations for the lower half-plane. Hence if D^* is analytic in k, it is independent of k but not of k^* except in the limit $\partial D^*/\partial \kappa = 0$. This provides a necessary but not a sufficient condition that D^* be analytic on the upper half-plane. The real and imaginary parts of (11) are necessary conditions that the derivative be unique as $\delta\pi \to 0$ and $\delta\kappa \to 0$ independently. A sufficient condition requires that D^* be differentiable[8] or independent of the angle in both limits.

Thus if the ordinary Cauchy–Riemann equations are satisfied, one obtains a guarantee that the derivative is unique as $\delta\pi = h\cos\theta \to 0$ and $\delta\kappa = h\sin\theta \to 0$ independently, where h is an infinitesimal; but this condition is independent of the angle in both limits and tells us nothing of the behavior as $\delta\kappa/\delta\pi \to 0$. The sufficient condition, that the function be differentiable, is simply that the derivative be unique also as $\tan\theta \to 0$; that this condition is independent of the $\delta\pi \to 0$, $\delta\kappa \to 0$ limits is evident from the observation that both limits cannot be taken independently except as $h \to 0$ at constant angle. In brief, the necessary condition is obtained from a variation of the length of the radius vector at each point on D^*, whereas the sufficient condition is obtained from the

[8] H. Jeffreys and B. S. Jeffreys, *Methods of Mathematical Physics* (Cambridge University Press, New York, 1962), pp. 237–238.

variation of the normal at each point. The necessary and sufficient condition that the derivative be unique is therefore that the derivative be independent of both variations in the limits $h \to 0$ and $\tan\theta \to 0$. Noting that $d\kappa/d\pi = \tan\theta$, the sufficient condition is included in

$$\frac{\partial}{\partial(d\kappa/d\pi)} \frac{\partial D^*}{\partial k} = 0, \tag{12}$$

where D^* is understood to be a function of angle as well. Equation (11), which is the Cauchy–Riemann equation for the lower half of the k plane, is a necessary condition that D^* be an analytic function in that region; thus (12) is a necessary and sufficient condition provided the partial derivatives are all continuous. Unless D^* is an explicit function of $d\kappa/d\pi$, (12) is satisfied identically.

We extend these results to two complex variables as follows; a necessary condition that $D(k^*, s^*)$ be analytic in k is

$$\frac{dD(k^*, s^*)}{dk} = \frac{\partial D(k^*, s^*)}{\partial s}\frac{ds}{dk} + \frac{\partial D(k^*, s^*)}{\partial k} = 0. \tag{13}$$

The first equality has been shown to hold quite generally for several complex variables.[9] With continuous partial derivatives of all variables, an analytic region on the momentum plane is guaranteed by the solutions of $\partial D^*/\partial k = 0$. Similarly an analytic region on the energy plane is guaranteed by the existence of solutions of $\partial D^*/\partial s = 0$, provided D^* is differentiable on the s plane. Thus (13) also provides a necessary condition either that D^* be analytic in s (if $\partial D^*/\partial s = 0$) or, if $\partial D^*/\partial s \neq 0$, that there exist an analytic region on the energy plane as a conformal mapping of an analytic region of the momentum plane (if $ds/dk = 0$). Both conditions require that $\partial D^*/\partial k = 0$; but (13) is also satisfied by $\partial D^*/\partial k = -(\partial D^*/\partial s)\, ds/dk$.

A sufficient condition is that dD^*/dk be differentiable or that all derivatives be finite and continuous. Thus a sufficient condition of analyticity in both variables is

$$\frac{\partial}{\partial(ds/dk)} \frac{dD(k^*, s^*)}{dk} = \frac{\partial D(k^*, s^*)}{\partial s} = 0 \tag{14}$$

provided s is not analytic in k. Hence s must not be a function of k. In this case the complex group velocity can be written in the usual way as the sum of real and imaginary parts and (14) expressed as the sum of two equations similar to (12); the second term in (14) is then obtained after taking the derivatives with respect to the real and imaginary velocities.

Now we note that the Cauchy–Riemann equation, (11), is identical with (7) obtained from the Heaviside expansions so that the solutions of (7) are certainly single valued and analytic provided D^* is differentiable. In general the numerator of (5) is a function of the initial and boundary conditions so that the singularities of Φ contribute terms which are independent of the interactions. Thus, excepting the trivial case, $1 - \lambda V = \Phi$, the existence of solutions of (13) and (14) proves the existence of an analytic region of $\Psi(k, s)$ in the k and s planes and the boundaries of the analytic region are given by the solutions of $\mathrm{Re}\,(D^*) = 0$ and $\mathrm{Im}\,(D^*) = 0$.

Noting again that the Cauchy–Riemann equation is formally equivalent to (7) obtained from the Heaviside expansions, and that solutions of (14) are analytic in both variables, we obtain Theorems 1 and 2.

Theorem 1: The existence of solutions of $\partial D^*/\partial k = 0$ and $\partial D^*/\partial s = 0$ is a necessary condition for the existence of analytic regions of $\Psi(k, s)$ in both variables, and the solutions define the boundaries of the analytic regions, and

Theorem 2: The existence of solutions of $dD^*/dk = 0$ and $\partial D^*/\partial s = 0$ is a necessary and sufficient condition for the existence of analytic regions of $\Psi(k, s)$ in both variables provided the group velocity is finite and continuous.

Thus if (14) has solutions, then $D(k^{*\prime}, s^{*\prime})$ is an analytic function of both variables and $k^{*\prime}$ and $s^{*\prime}$ are the correct excitation eigenvalues, irrespective of the analytic properties of Ψ. Thus we obtain the corollary.

Corollary: Provided ds/dk is finite and continuous, the existence of solutions of $\partial D^*/\partial k = 0$ and $\partial D^*/\partial s = 0$ is a necessary and sufficient condition that the Heaviside expansions are a solution of (5).

3. SOLUTION OF THE LINEARIZED SELF-CONSISTENT FIELD EQUATIONS

As an example of the method developed in Secs. 1 and 2 we consider the linearized self-consistent field equations with arbitrary coupling. A self-consistent formulation of an interaction problem requires that the solutions of an equation of motion for the particles be compatible with a coupled field equation representing the interactions. Self-consistent methods are of interest because it is not necessary to specify the

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

exact nature of the source, that is, as a sum of two-body, three-body potentials, continuous charge distribution, etc., but specific sources may be introduced without changing the procedure.

In principle, self-consistent formulations may be written for any particle interactions via a conservative field. Such a scheme is made possible by noting that a divergence equation may be written for any conservative field, that is, for any field analytic in all variables, and related to the change in density in the manner familiar in continuum mechanics.

We consider the coupled equations of motion

$$\{[\mathcal{H}_0 + \mathfrak{U}(\chi)], N(k, \chi)\} = i(\partial/\partial t)N(k, \chi) \quad (15)$$

and

$$\Box\mathfrak{U}(\chi) = \lambda\int d^3k\, N(k, \chi), \quad (16)$$

where $N(k, \chi)$ is the distribution function for particles with momenta k at $\chi = (\mathbf{x}, t)$ interacting via the internal field $\mathfrak{U}(\chi)$ and with coupling strength λ. In writing the inhomogeneous scalar wave equation in the form (16) we obviously do not consider radiative interactions or external fields. As we do include longitudinal collective modes, the integrations are over all final particle states including excited states. Now writing $N(k, \chi) = n_0(k) + n(\chi)$, we assume the field developed from $n_0(k)$ adds a constant term to the initial Hamiltonian and $n(\chi)$ then represents the density fluctuations in space and time. The corresponding field, obtained from (16) is the sum of phonons or plasmons, etc., the field modes being determined by the type of coupling.

Linearizing and taking matrix elements of (15) for transitions between initial and final states p and k, then Fourier transforming with respect to the momentum-energy transfer, $\tau = q, E$, we obtain

$$\mathfrak{U}(\tau) = U_0(\tau_0) + \lambda\Gamma(\tau)\Big[\tilde{U} + \int d^3k\, n(k, \tau)\Big] \quad (17)$$

and

$$N(k, \tau) = n_0(k_0, \tau) + G(p, \tau)[n_0(k)U_0 + n(k, \tau)U_0 + n_0(k)U(\tau) + \tilde{n}], \quad (18)$$

where $\tilde{n} = \tilde{n}(k, \bar{\tau})$ and $\tilde{U} = \tilde{U}(k, \bar{\tau})$ are functions of the initial and boundary values of $n(k, \tau)$ and $U(\tau)$ and their derivatives, $G(p, \tau)$ is

$$G(p, k, E) = (E_p - E_k + E)^{-1}, \quad (19)$$

the propagator for noninteracting particles in the initial state, and $\Gamma(\tau)$ is

$$\Gamma(q, E) = (q^2 - E^2/c^2)^{-1}, \quad (20)$$

the propagator for noninteracting fields in the absence of particles. Setting $U_0 = 0$ (since U_0 is included in the unperturbed Hamiltonian), after cancelling the zero-order terms, (17) and (18) become

$$U(\tau) = \frac{\lambda\Gamma(\tau)\int d^3kG(p, \tau)\tilde{n}}{1 - \lambda\Gamma(\tau)\int d^3kG(p, \tau)n_0(k)} \quad (21)$$

and

$$n(k, \tau) = G(p, \tau)\Big\{\tilde{n} + n_0(k)\frac{\lambda\Gamma(\tau)[\tilde{U} + \int d^3kG(p, \tau)\tilde{n}]}{1 - \lambda\Gamma(\tau)\int d^3kG(p, \tau)n_0(k)}\Big\}. \quad (22)$$

In obtaining (21) we have carried out a unitary transformation of the form $O(k, E)U(k, E)O^\dagger(k^\dagger, E^\dagger)$, with $k^\dagger = -k$, etc., which in effect continues the dispersion relation into the third quadrant for each integration. The integral operator can be shown to be unitary directly by carrying through the substitutions $k^\dagger = -k$, $E^\dagger = -E$ in $D^\dagger(k^\dagger, E^\dagger)$, whereupon one finds $D^\dagger(k^\dagger, E^\dagger) = D^{-1}(k, E)$ so that one recovers the initial dispersion relation valid in the first quadrant. The result, (21), can be shown to be valid quite independently by substituting $\rho = \int d^3kn(k, \tau)$ into (15) and (16) and carrying through the same calculation. As our example is essentially illustrative, however, and the results are not sensitive to the formal details, we omit the latter.

With (5), the solutions of (21) may be written

$$U(x, t) = \sum_E \lim_{E\to E'} \sum_q \lim_{q\to q'} \times \frac{\lambda[\tilde{U} + \int d^3kG(p, q, E)\tilde{n}]e^{-iq\cdot x - iEt}}{(\partial/\partial E)(\partial^3/\partial q^3)[q^2 - E^2/c^2 - \lambda\int d^3kG(p, q, E)n_0(k)]} + \text{principal value integrals}, \quad (23)$$

where E' and q' are solutions of

$$D(p, q, E') = q^2 - E'^2/c^2 - \lambda\int d^3kG(p, q, E')n_0(p, q) \equiv 0 \quad (24)$$

and, after utilizing $\partial k_i/\partial q_i = 1$, and taking the average of the first term,

$$\lim(q \to q')\frac{\partial^3 D(p, q, E)}{\partial q^3} = \lim(q \to q')\Big[+\frac{8q_1q_2q_3}{(q_1^2 + q_2^2 + q_3^2)^2} - \lambda G(p, q, E)n_0(p, q)\Big] \equiv 0. \quad (25)$$

$D^{-1}(p, q, E)$ has the usual form of a field propagator modified by a vertex part representing the interactions

with particles. The solutions of (24) thus give the energy eigenvalues of the excitations in the interacting system and (25) is an analogous equation for the momentum eigenvalues. Writing $a_0 = 1/2m\lambda$ for the virtual scattering length, (25) becomes, for a spherical distribution,

$$q'^2 + 2[p + a_0^{-1} n_0(p, q')]q' - 2mE = 0. \quad (26)$$

Now writing $q' = \pi' + i\kappa'$, $E = \omega + i\gamma$, $n_0(p, q') = \mathrm{Re}\, n_0(p, q') + i\, \mathrm{Im}\, n_0(p, q')$, and resolving, one obtains

$$2\kappa'[\pi' + p + a_0^{-1} \mathrm{Re}\, n_0(p, q')] + 2a_0^{-1}\pi' \mathrm{Im}\, n_0(p, q') - 2m\gamma = 0 \quad (27)$$

and

$$\pi'[\pi' + 2p + 2a_0^{-1} \mathrm{Re}\, n_0(p, q')] - \kappa'[\kappa' + 2a_0^{-1} \mathrm{Im}\, n_0(p, q')] - 2m\omega = 0. \quad (28)$$

The solutions of (27) and (28) for π' and κ' define the boundaries of the analytic region in the momentum plane, and similar equations may be written for the energy plane requiring that ω' and γ' also satisfy either (14) or its conjugate; that is, if $dE/dq = 0$ then $dE^*/dq \neq 0$ or the converse, so that the analytic momentum region can always be mapped on to one half of the E plane. A straightforward calculation gives

$$-m\frac{\partial\omega}{\partial\kappa'} = m\frac{\partial\gamma}{\partial\pi'} = \kappa' + a_0^{-1} \mathrm{Im}\, n_0(p, q') \quad (29)$$

and

$$m\frac{\partial\gamma}{\partial\kappa'} = m\frac{\partial\omega}{\partial\pi'} = \pi' + p + a_0^{-1} \mathrm{Re}\, n_0(p, q') \quad (30)$$

so that

$$\frac{dE}{dk'} = \frac{\partial\omega}{\partial\pi'} - i\frac{\partial\omega}{\partial\kappa'}. \quad (31)$$

Comparing (27)–(30) we find that the analytic domain defined by (27) and (28) may be divided into regions defined by certain properties of the group velocity.

In region I, $\partial\omega/\partial\pi' = \partial\omega/\partial\kappa' = 0$ so that in equilibrium

$$\pi_1' = -p - a_0^{-1} \mathrm{Re}\, n_0(p, q'),$$
$$\omega_1 = -\pi_1'^2/2m, \quad \gamma_1' = \mathrm{Im}\, n_0(p, q') = 0. \quad (32)$$

The excitations are undamped but have no oscillations. In this region of stationary particle excitations the interactions simply add the momentum of the virtual field $-a_0^{-1} \mathrm{Re}\, n_0(p, q')$ to the initial particle momentum $-p$ and the energy shift is obtained from a straightforward renormalization of the kinetic mass. In region I the energy region is a conformal mapping, with intercepts $\gamma_\pm' = 0$,

$$\omega_\pm' = [\mp p_{\max} - a_0^{-1} \mathrm{Re}\, n_0(p_{\max}, q'_{\max})]^2/2m$$

of the analytic momentum region with intercepts

$$\pi_\pm' = \pm p_{\max} - a_0^{-1} \mathrm{Re}\, n_0(p_{\max}, q'_{\max}), \quad \kappa_\pm' = 0.$$

In region II, $\partial\omega/\partial\pi' = 0$, $\partial\omega/\partial\kappa' \neq 0$ and to the particle excitations are added stationary collective oscillations defined by

$$\pi_2' = -p - a_0^{-1} \mathrm{Re}\, n_0(p, q'),$$
$$\kappa_2' = -\pi_2'^2/2m - \kappa_2'[\kappa_2' + 2a_0^{-1} \mathrm{Im}\, n_0(p, q')]/2m,$$
$$\gamma_2' = m^{-1}a_0^{-1}\pi_2' \mathrm{Im}\, n_0(p, q'). \quad (33)$$

In region II the interactions among particle excitations are screened by bound collective modes for energies $\omega > \pi'^2/2m$ and the total energy is analytic only if $\partial D/\partial\omega = 0$ in region II.

In region III, $\partial\omega/\partial\kappa' = 0$, $\partial\omega/\partial\pi' \neq 0$ so that

$$\kappa_3' = -a_0^{-1} \mathrm{Im}\, n_0(p, q'),$$
$$\omega_3' = \pi[\pi + 2p + 2a_0^{-1} \mathrm{Re}\, n_0(p, q')]/2m - \kappa_3'^2/2m, \quad (34)$$
$$\gamma_3' = \kappa_3'[p + a_0^{-1} \mathrm{Re}\, n_0(p, q')]/m$$

representing the excitation of field particles with definite limits on their energies imposed by the convergence requirements of $n_0(p_{\max}, q'_{\max})$ and which have a diffusion length proportional to κ'^{-1}. In region III the analytic energy region is once more a conformal mapping with intercepts

$$\omega_\mp' = -\pi_{\max}^2/2m \mp 2\pi_{\max} \times [p_{\max} + 2a_0^{-1} \mathrm{Re}\, n_0(p_{\max}, q'_{\max})]/2m - \kappa_\pm'^2/2m$$

of the analytic momentum region with intercepts

$$\kappa_\pm' = -a_0^{-1} \mathrm{Im}\, n_0(p_{\max}, q'_{\max}) \quad \text{and} \quad \pi_\pm = \pm\pi_{\max}$$

determined from the convergence requirements of the distribution function. This requirement defines the threshold for decay of field particle excitations into particle pairs.

In region IV, $\partial\omega/\partial\kappa' \neq 0$, $\partial\omega/\partial\pi' \neq 0$ and the field excitations propagate as free particle excitations in the continuum with

$$\partial^3 D(p, q_3', E_3')/\partial q_3'^3 = \partial D(p, q_3', E_3)/\partial E_3' = 0. \quad (35)$$

For undamped excitations in region I, $\pi_1' = -p - a_0^{-1} \mathrm{Re}\, n_0(p, q')$ and $\omega_1' = -\pi_1'^2/2m$ reducing to the initial state in the zero coupling limit. The second term is thus the momentum of the "polarization" field due to the dressing of particles by their interactions and the particle excitations have undamped stable states in the interacting system. In region II if $\gamma_2' = 0$, π_2' and κ_2' remain invariant. In region III if $\gamma_3' = 0$, $\omega_3' = \pi'^2/2m$ and κ_3' is invariant. Hence undamped excitations may occur in all three regions.

The renormalized phase velocity, as well as all other physical observables in the interacting system, may be obtained from the statistical averages over the new distribution function, which we obtain from the solution of (22). The result is

$$N(k, x, t) = n_0(k) + \sum_E \lim_{E \to E_{s'}} \sum_q \lim_{q \to q_{s'}} \frac{\tilde{n} e^{-q\cdot x} e^{-iEt}}{(\partial/\partial E)(\partial^3/\partial q^3) G^{-1}(p, q, E)} + \sum_E \lim_{E \to E''} \sum_q \lim_{q \to q''} \frac{\lambda[\tilde{U} + \int d^3k G(p, q, E)\tilde{n}] e^{-q\cdot x} e^{-iEt}}{(\partial/\partial E)(\partial^3/\partial q^3) G^{-1}(p, q, E) D(p, q, E)} + \text{principal value integrals}, \quad (36)$$

where E_s' and q_s' are solutions of the single particle dispersion equations,

$$G^{-1}(p, q, E_s') = \lim (E \to E_s')(E_p - E_k + E) \equiv 0 \quad (37)$$

and

$$\lim (q \to q_s')(\partial^3/\partial q^3) G^{-1}(p, q, E) \equiv 0 \quad (38)$$

and q'' and E'' are solutions of the appropriate eigenvalue equations in the second term; there are, however, only two sets of values corresponding to the solutions of Eqs. (24), (25), (37), and (38), as the free field equations, $\Gamma^{-1} = 0$ and $\partial^3\Gamma^{-1}/\partial q^3 = 0$ have only the trivial solutions, $q_{f'} = E_f' = 0$. As we have already found q' and E' which appear also in (23), there remain only the single-particle eigenvalues obtained from (37) and (38). They are

$$\pi_s' = -p, \quad \kappa_s' = 0, \quad E_s' = -p^2/2m, \quad (39)$$

the familiar results for noninteracting particles. Now carrying out the operations in the denominator of the second term we find that $G^{-1}D = 0$ and

$$(\partial^3/\partial q^3)(\partial/\partial E) G^{-1} D = (\partial^4/\partial q^4) G^{-1} D$$

yield either $G^{-1} = 0$ or $D = 0$, and therefore either

$$G^{-1}\partial^3 D/\partial q^3 = 0 \; (G^{-1} \neq 0)$$

or $D\partial^3 G^{-1}/\partial q^3 = 0$ $(D \neq 0)$, which give just the two sets of eigenvalue equations we have already found for the single particle and field excitations. Hence the distribution function separates quite naturally into two parts, a short-range contribution

$$n_s(k, x, t) = \lambda \sum_E \lim_{E \to -p/2m} \sum_q \lim_{q \to -p} \times \frac{A_1[\tilde{U} + \int d^3k G(p, q, E)\tilde{n}] e^{-q\cdot x} e^{-iEt}}{D(p, q, E)(\partial/\partial E)(\partial^3/\partial q^3) G^{-1}(p, q, E)} \quad (40)$$

plus principal value terms, and a long-range contribution

$$n_l(k, x, t) = \lambda \sum_E \lim_{E \to E'} \sum_q \lim_{q \to q'} \times \frac{A_1[\tilde{U} + \int d^3k G(p, q, E)\tilde{n}] e^{-q\cdot x} e^{-iEt}}{G^{-1}(p, q, E)(\partial/\partial E)(\partial^3/\partial q^3) D(p, q, E)}, \quad (41)$$

where the eigenvalues of n_s are the single-particle values and those of n_l are those for the field. A_1 and A_2 are constants resulting from the partial fraction expansions. The remaining terms in the denominator, $(\partial G^{-1}/\partial E)(\partial^3 D/\partial q^3)$, $(\partial D/\partial E)(\partial^3 G^{-1}/\partial q^3)$, $D\partial^3 G^{-1}/\partial q^3$, $G^{-1}\partial^3 D/\partial q^3$, contribute only to the principal part in the limits $\partial G^{-1}/\partial E = \infty$, $\partial D/\partial E = \infty$, $D = \infty$, and so on.

The distribution (36) now consists of a noninteracting single-particle contribution independent of the coupling strength, plus two interaction terms, both λ dependent and representing the long-range and short-range parts of the correlations, plus principal value terms. The noninteracting single-particle contribution diverges however due to the vanishing of $(\partial^4/\partial q^4) G^{-1}$ as $q \to q_s'$ unless $\tilde{n} = 0$ in the limit; this means there can be no noninteracting single particle contributions to the interacting part of the distribution. The divergence in n_s due to vanishing of $\partial^4/\partial q^4 G^{-1}$ in the limit is canceled by a divergence of the same order due to $D(q, E)$ in the denominator. Thus

$$\lim (q \to -p) \lim (E \to -p^2/2m) \times \left[q^2 - E^2/c^2 - a_0^{-1} \int d^3k \frac{n_0(p, q)}{-q^2 - 2pq + mE}\right] = p^2 - p^4/c^2 - a_0^{-1} \lim (q \to -p) \frac{qpn_0(p, q)}{-2q - 2p} \quad (42)$$

diverges in the same manner as $(\partial^4/\partial q^4) G^{-1}$ approaches zero.

The principal value contributions to both $U(x, t)$ and $n(k, x, t)$ can be separated into contributions due to the integrations along the boundaries of the analytic regions plus terms from the nonanalytic regions. The boundary contributions, represented by the Cauchy principal values, are just the arithmetic averages of the contributions inside and outside the analytic domain and of course represent the collective excitations of the system. The nonanalytic parts then represent the uncorrelated background fluctuations.

ACKNOWLEDGMENTS

The author thanks Dr. A. Rajagopal for reading the manuscript and for several helpful suggestions.

Analytic Continuation of Laplace Transforms by Means of Asymptotic Series

W. A. BEYER AND L. HELLER
University of California, Los Alamos Scientific Laboratory, Los Alamos, New Mexico

(Received 14 July 1966)

Conditions under which a Laplace transform $L\{F(t)\} = f(s)$ may be analytically continued, by means of an asymptotic expansion of F, outside the half plane of convergence of the Laplace transform integral are investigated. For $t > k$ define R_N by $F(t) = t^{-\beta}\{\sum_{n=0}^{N} a_n t^{-n} + R_N(t)\}$ for some fixed β with $\mathrm{Re}\,\beta < 1$ and suppose that F is integrable on $[0, k)$ for some $k \geq 0$. *First*, it is shown that if $R_N(t) = O\{N!(\sigma/t)^{N+1}\}$ uniformly in N and $t > k$ for some $\sigma > 0$, then the singular part of f at $s = 0$ can be determined in terms of a_i. If β is an integer, then in some neighborhood of $s = 0$ it is shown that

$$L\{F(t)\} = s^{\beta-1} \sum_{i=0}^{-\beta} a_i \Gamma(1 - i - \beta) s^i + (\log s) g(s) + h(s),$$

where g and h are analytic at $s = 0$ and $g(s) = \sum_{i=1}^{\infty} (-1)^i a_{i-\beta} s^{i-1}/(i-1)!$. If β is not an integer, in some neighborhood of $s = 0$ it is shown that $L\{F(t)\} = s^{\beta-1} g(s) + h(s)$, where $g(s) = \sum_{i=0}^{\infty} a_i \Gamma(1 - i - \beta) s^i$, with g and h analytic at $s = 0$. *Second*, if the estimate on $R_N(t)$ holds uniformly in N and in the complex t plane in the region $(|t| > k) \cap (|\arg t| < (\frac{1}{2}\pi) + \lambda)$ for some $\lambda > 0$, then the analytic continuation of f can be determined in terms of the a_i. For any $k' > k$ and for $|\arg s| < \pi$ we have $L\{F(t)\} = \int_0^{k'} e^{-st} F(t)\, dt + \int_0^{\infty} a(t) \Gamma(2 - \beta, k'(s + t))(s + t)^{\beta-2}\, dt$, where Γ is the incomplete gamma function and $a(t)$ is the analytic continuation of $\sum_{i=0}^{\infty} a_i t^i/i!$. If $k = 0$ in the hypotheses, then with a slight further restriction on $F(t)$ one has $L\{F(t)\} = \Gamma(2 - \beta) \int_0^{\infty} a(t)(s + t)^{\beta-2}\, dt$. A generalization and application to a problem in nonrelativistic dispersion theory which includes a Coulomb potential are discussed.

I. INTRODUCTION

A. Background

TO solve a nonrelativistic potential scattering problem by the method of partial wave dispersion relations it is necessary to know the location and nature of all the singularities of the scattering amplitude $A(q)$ in the upper half q (momentum) plane. Allowing for the possible presence of a Coulomb potential, an integral equation for the partial wave-function $\mathfrak{U}(q, r)$ is

$$\mathfrak{U}(q, r) = \mathfrak{f}(q, r) - M \int_0^{\infty} \mathfrak{f}(q, r_<) \mathfrak{h}^{(+)}(q, r_>) V(r') \mathfrak{U}(q, r')\, dr',$$

where M is twice the reduced mass, V is the nuclear potential, and $\mathfrak{f}$ and $\mathfrak{h}^{(+)}$ are regular and irregular Coulomb functions, respectively, defined by

$$\mathfrak{f}(q, r) = r^{L+1} e^{iqr} \Phi(L + 1 + i\eta, 2L + 2, -2iqr),$$

$$\mathfrak{h}^{(+)}(q, r) = D(q) r^{L+1} e^{iqr} \Psi(L + 1 + i\eta, 2L + 2, -2iqr),$$

where Φ and Ψ are confluent hypergeometric functions,[1] $\eta = Me^2/2q$ with e the common charge of the scattering particles, and $D(q)$ is chosen to make the Wronskian $\mathfrak{f}_r \mathfrak{h}^{(+)} - \mathfrak{f} \mathfrak{h}_r^{(+)} = 1$,

$$D(q) = -ie^{-i\pi L} (2q)^{2L+1} \Gamma(L + 1 + i\eta)/\Gamma(2L + 2).$$

The dependence upon the orbital angular momentum quantum number L has been suppressed in the notation. $r_<$ ($r_>$) denotes the smaller (greater) of r and r'. $\mathfrak{f}$ is an entire function of q for all finite r. $\mathfrak{h}^{(+)}$ is analytic in the q plane except for $q = 0$ and some points on the negative imaginary q axis which accumulate at $q = 0$, but approaches a unique limit as $q \to 0$ along any path which avoids a small sector containing the negative imaginary axis.[2]

The integral equation is equivalent to the differential equation for $r > 0$:

$$\left[\frac{d^2}{dr^2} + q^2 - \frac{Me^2}{r} - \frac{L(L+1)}{r^2}\right] \mathfrak{U}(q, r) = MV(r) \mathfrak{U}(q, r),$$

and the two conditions: $\mathfrak{U}(q, 0) = 0$ and $\mathfrak{U}(q, r) \to \mathfrak{f}(q, r) + A(q) \mathfrak{h}^{(+)}(q, r)$ as $r \to \infty$. $\mathfrak{f}$ and $\mathfrak{h}^{(+)}$ are solutions of the differential equation with $V(r) = 0$. $A(q)$ is the scattering amplitude. From the integral equation for $\mathfrak{U}$, $A(q)$ is seen to be

$$A(q) = -M \int_0^{\infty} \mathfrak{f}(q, r) V(r) \mathfrak{U}(q, r)\, dr.$$

The units are such that $\hbar = c = 1$.

[1] Bateman Manuscript Project, *Higher Transcendental Functions* (McGraw-Hill Book Company, Inc., New York, 1953), Vol. 1, Chap. VI.

[2] L. Heller and M. Rich. Phys. Rev. **144**, 1324 (1966).

For a potential $V(r)$ which is a superposition of exponential potentials,

$$V(r) = \int_{\mu}^{\infty} \sigma(\alpha) e^{-\alpha r}\, d\alpha,$$

and such that $V(r)$ is less singular than r^{-2} at $r = 0$, it is known[3-5] that the only singularities of $A(q)$ in the upper half q plane are: branch points at $q = \frac{1}{2}ni\mu$, $n = 1, 2, 3 \cdots$; possible bound state poles; and an essential singularity at $q = 0$ which, however, is of no consequence[6] for contours which avoid a small sector containing the negative imaginary q axis.

Furthermore the discontinuity in $A(q)$ across the upper imaginary q axis,

$$\lim_{\epsilon \to 0} [A(q + \epsilon) - A(q - \epsilon)],$$

which is the needed input for the dispersion relation, is correctly given[3] in the region $\frac{1}{2}\mu \leq -iq < \frac{1}{2}(n + 1)\mu$, except at bound state poles,[7] by the sum of the discontinuities of the first n Born approximations to $A(q)$, whether or not the Born series converges. The nth Born approximation is obtained by putting the $(n - 1)$ iterate for $\mathfrak{U}$ from the integral equation into the above expression for $A(q)$, the zeroth iterate being $\mathfrak{f}$.

The problem arises, therefore, of how to determine the location and nature of the singularities of a given Born term. If one can find an analytic expression for the desired integral (which is n dimensional in nth Born approximation) then the problem is essentially solved. In general, however, one cannot find analytic expressions for these integrals[8] and this leads to the attempt to find a general way of extracting their singular parts. Looking at the first Born approximation for a Yukawa potential, the integral under consideration is

$$-M \int_0^{\infty} \mathfrak{f}^2(q, r) \frac{e^{-\mu r}}{r}\, dr.$$

Since the integrand is an entire function of q for each fixed r, the existence of singularities of the integral is associated with a failure of the integral to converge at its upper limit. There are two distinct series in the asymptotic behavior of $\mathfrak{f}$, one of which has a common factor of $\exp(iqr)$, and the other of $\exp(-iqr)$. Therefore $\mathfrak{f}^2$ has three distinct series with common factors of $\exp(2iqr)$, 1, and $\exp(-2iqr)$. When these three kinds of terms are put into the integral, it is reasonable to expect that only $\exp(-2iqr)$ can produce singularities in the upper half q plane since the integrals involving terms of the other two types will converge if $\operatorname{Im} q \geq 0$. Defining

$$B(s, r) \equiv r^{-1} e^{(s-\mu)r} \mathfrak{f}^2((s - \mu)/2i, r),$$

where $s = \mu + 2iq$, puts the integral in the form

$$-M \int_0^{\infty} e^{-sr} B(s, r)\, dr.$$

$B(s, r)$ does not have any increasing exponential behavior at $r = \infty$. If B were independent of s, this integral would be a Laplace integral. If B were weakly dependent on s, the analytic properties of the integral should differ little from those of a Laplace integral. In Sec. VI this generalization of a Laplace integral (or transform) is discussed.

B. Outline

For $F(t)$ integrable on each finite subinterval of $[0, \infty)$, as is well known, the one-sided Laplace transform

$$f(s) = \lim_{w \to \infty} \int_0^{w} e^{-st} F(t)\, dt \equiv L\{F(t)\}$$

defines, in general, a function of a complex variable s which is analytic in a half-plane $\operatorname{Re}(s) > \alpha$. Unlike power series, there may not be singularities of $f(s)$ on the axis of convergence, $\operatorname{Re}(s) = \alpha$, of $L\{F(t)\}$. In the literature there seem to be few criteria for determining singularities of $f(s)$.[9]

Our purpose is to show how term-by-term application of the Laplace transform to the asymptotic expansion of $F(t)$ sometimes gives information on singularities of $f(s)$ on the axis of convergence.

Any singularity of $L\{F(t)\}$ results from the behavior of $F(t)$ as $t \to \infty$. This should be reflected in an asymptotic expansion of $F(t)$ as $t \to \infty$, say $\sum_{i=0}^{\infty} a_i t^{-i}$. But

$$\int_1^{\infty} e^{-st} t^{-i}\, dt = [(-1)^i/(i - 1)!\, s^{i-1}] \log s + (\text{entire function of } s)$$

[3] H. Cornille and A. Martin, Nuovo Cimento **26**, 298 (1962).
[4] D. Y. Wong and H. P. Noyes, Phys. Rev. **126**, 1866 (1962).
[5] J. R. Rix, thesis, Harvard University (1965).
[6] See Appendix A of Ref. 2, and Appendix B of A. Scotti and D. Y. Wong, Phys. Rev. **138**, B145 (1965).
[7] If the rest of the discontinuity is known exactly, then omitting the discontinuity at the bound states has no effect on the solution of the dispersion relation.
[8] Some problems which have been solved analytically are the first Born approximation for a Yukawa potential without charge and with charge (see Ref. 4 for the latter). The second Born approximation for a single Yukawa without charge has been treated in P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill Book Company, Inc., New York, 1953), Vol. II, p. 1082, and in M. Luming, Phys. Rev. **136**, B1120 (1964).
[9] G. Doetsch, *Handbuch der Laplace Transform* (Birkhäuser, Basel, 1950), Vol. I; (1955), Vol. II; (1956), Vol. III.

for i a positive integer.[10] If F is integrable on (0, 1) we have, proceeding formally:

$$L\{F(t)\} = \int_1^\infty e^{-st}\left(\sum_{i=0}^{\infty} a_i t^{-i}\right) dt + \int_0^1 e^{-st}F(t)\,dt$$
$$= \frac{a_0}{s} + \log s\left(\sum_{i=1}^{\infty} \frac{(-1)^i a_i}{(i-1)!}\, s^{i-1}\right) + h(s),$$

where perhaps $h(s)$ is analytic at $s = 0$. For reasons discussed in part A of this Introduction, one is interested only in $f(s) = L\{F(t)\}$ modulo a function analytic at a singularity of $f(s)$. The above expression for $f(s)$ provides a convenient decomposition of $f(s)$ at $s = 0$; $f(s)$ is expressed as the sum of a function analytic at $s = 0$, a function a_0/s, and the product of $\log s$ and a function analytic at $s = 0$. The purpose of this paper is to find justification for this formal procedure.

Doetsch[11] states a theorem that if $F(t)$ is asymptotic to

$$\sum_{\nu=0}^{N} c_\nu t^{\alpha_\nu} (\operatorname{Re}\alpha_0 > \operatorname{Re}\alpha_1 > \cdots > \operatorname{Re}\alpha_N > -1)$$

for $t \to \infty$ then near $s = 0$, $f(s)$ is asymptotic to

$$\sum_{\nu=0}^{N} c_\nu \Gamma(1 + \alpha_\nu) s^{-1-\alpha_\nu}$$

in the right half-plane. Only those powers of t are included which give *infinite* singularities at $s = 0$.

Section II gives a stronger theorem under which a weak Watson condition on the asymptotic expansion of F permits the determination of the complete singular part of $L\{F\}$ at $s = 0$. An application of this theorem is given in Sec. III. Section IV gives another theorem under which a strong Watson condition on F ensures that $s = 0$ is the only singularity of $L\{F\}$ in $|\arg s| < \pi + \lambda$ for some $\lambda > 0$, and gives a formula for the analytic continuation of $L\{F\}$ to that portion of the complex plane. Section V gives an example of the application of Theorem 2. Section VI discusses the problem mentioned in part A of this Introduction and outlines a generalization of Theorem 1. Section VII reviews the literature on estimates on the remainders of asymptotic series which are necessary to apply these theorems.

The term "Watson condition" arises from G. N. Watson's investigation of the characteristics of certain asymptotic expansions.[12]

The first theorem of this paper can be regarded as an example of a more general phenomenon where a linear transformation T is given and the term-by-term application of T to a general expansion $F \sim \sum F_n$ yields, in general, an expansion $T(F) \sim \sum T(F_n) \sim \sum f_n$. The expansions $F(t) \sim \sum F_n(t)$ are usually associated with a certain region of t space and have an "error term" $R(N, t)$ which satisfies various conditions. Other examples of this phenomenon are discussed under titles such as "Watson's lemma." See Erdélyi[13,14] and Erdélyi and Wyman.[15]

Throughout, the term "continue" is short for "analytically continue."

II. FIRST THEOREM AND PROOF

Theorem 1: Suppose $F(t)$ is a function, real or complex valued, defined and continuous on $(0, \infty)$ and for $t > k$:

$$F(t) = t^{-\beta}\left[\sum_{i=0}^{N} a_i t^{-i} + R_N(t)\right],$$

where

$$R_N(t) = O\{N!(\sigma/t)^{N+1}\}$$

uniformly in N and $t > k$. Without loss of generality, restrict $\operatorname{Re}\beta < 1$. k and σ are fixed positive numbers. Suppose $F(t)$ is integrable on $[0, k]$. Then $L\{F(t)\}$ is analytic for $|\arg s| < \frac{1}{2}\pi$, $s \neq 0$, and in some deleted neighborhood of $s = 0$, whose radius is proportional to σ^{-1}, the continuation of $L\{F(t)\}$ has the form, in case β is an integer,

$$s^{\beta-1}\sum_{i=0}^{-\beta} a_i \Gamma(1 - i - \beta) s^i + (\log s) g(s) + h(s),$$

where g and h are analytic at $s = 0$ and

$$g(s) = \sum_{i=1}^{\infty} \frac{(-1)^i}{(i-1)!}\, a_{i-\beta} s^{i-1}.$$

For β not an integer, the continuation of $L\{F(t)\}$ has the form

$$s^{\beta-1} g(s) + h(s),$$

where g and h are analytic at $s = 0$ and

$$g(s) = \sum_{i=0}^{\infty} a_i \Gamma(1 - i - \beta) s^i.$$

Remark: The condition on $R_N(t)$ is called the weak Watson condition.

Since the proof is somewhat involved, we first outline it. An estimate of a_n shows that $a_n = O(n!\,\sigma^n)$. We then construct the series $\sum_{n=0}^{\infty} a_n u^n/n! = a(u)$ which is convergent for $|u| < 1/\sigma$. [$a(u)$ is called the Borel transform of the series $\sum a_n t^{-n}$.] Using $a(u)$, we construct a function $\tilde{F}(t)$ and show that $\tilde{f}(s) = L\{\tilde{F}(t)\}$

[10] See Ref. 9, Vol. I, p. 468.
[11] See Ref. 9, Vol. II, p. 97.
[12] G. N. Watson, Phil. Trans. Roy. Soc. London (A) **211**, 279 (1912).
[13] A. Erdélyi, *Asymptotic Expansions* (Dover Publications, Inc., New York, 1956), p. 34.
[14] A. Erdélyi, Arch. Ratl. Mech. Anal. **7**, 1 (1961).
[15] A. Erdélyi and M. Wyman, Arch. Ratl. Mech. Anal. **14**, 239 (1963).

satisfies the conclusion of the theorem, considering separately the cases β an integer and β not an integer. We then show that $F(t) - \tilde{F}(t) = O(e^{-bt})$ for some $b > 0$ and thus that $L\{F(t) - \tilde{F}(t)\}$ is analytic at $s = 0$.

Proof. From the hypothesis there exists $\hat{K}$ such that for $t > k$:

$$|R_N| < \hat{K}N!\,(\sigma/t)^{N+1}.$$

Let $N > 0$.

$$\begin{aligned} |a_N| &\le (|R_{N-1}| + |R_N|)t^N \le \hat{K}N!\,\sigma^N[(1/N) + (\sigma/t)] \\ &\le \hat{K}N!\,\sigma^N[(1/N) + (\sigma/k)] \le KN!\,\sigma^N, \end{aligned} \tag{1}$$

where $K = \hat{K}(1 + \sigma/k)$. Thus the series

$$a(u) = \sum_{n=0}^{\infty} (a_n/n!)u^n$$

is convergent for $|u| < 1/\sigma$. Put $\tau = 1/(4\sigma)$ and define

$$\tilde{F}(t) \equiv t^{1-\beta}\int_0^{\tau} e^{-tu}a(u)\,du. \tag{1a}$$

For $\operatorname{Re} s > 0$ and $u \ge 0$ we have

$$\Gamma(2-\beta)(u+s)^{-(2-\beta)} = \int_0^{\infty} e^{-(s+u)t}t^{1-\beta}\,dt. \tag{2}$$

For $\operatorname{Re} s > 0$ define

$$\tilde{f}(s) = \int_0^{\infty} e^{-st}\tilde{F}(t)\,dt, \tag{3}$$

and using (1a) and (2) obtain for $\operatorname{Re} s > 0$:

$$\tilde{f}(s) = \int_0^{\infty} e^{-st}t^{1-\beta}\left\{\int_0^{\tau} e^{-tu}a(u)\,du\right\}dt,$$

$$\tilde{f}(s) = \int_0^{\tau} a(u)\left\{\int_0^{\infty} e^{-(s+u)t}t^{1-\beta}\,dt\right\}du, \tag{4}$$

$$\tilde{f}(s) = \Gamma(2-\beta)\int_0^{\tau} a(u)(u+s)^{-(2-\beta)}\,du. \tag{5}$$

Since the iterated integral in (4) is absolutely convergent for $\operatorname{Re} s > 0$, the interchange of order of integration is justified. The integral (5) defines an analytic function of s in the s plane cut along $[-\tau, 0]$.[16] We write from (5):

$$\tilde{f}(s) = \Gamma(2-\beta)\sum_{n=0}^{\infty}\frac{a_n}{n!}\int_0^{\tau}\frac{u^n}{(u+s)^{2-\beta}}\,du \tag{6}$$

since the series $\sum (a_n/n!)u^n$ is uniformly convergent on $[0, \tau]$.

[16] Another way than that below to discuss the analytic properties of $\tilde{f}(s)$ is to integrate (5) by parts until $u + s$ is raised to a positive power and then to break up the remaining integral into two pieces: $0 \to -s$ and $-s \to \tau$. The first integral together with the integrated terms gives the desired answer and the second is analytic at $s = 0$.

Suppose β is not an integer. Then

$$\tilde{f}(s) = \Gamma(2-\beta)\sum_{n=0}^{\infty}\frac{a_n}{n!} \times \left\{\sum_{\nu=0}^{n}\binom{n}{\nu}\frac{(-s)^{n-\nu}}{\beta+\nu-1}\left[(s+\tau)^{\nu-1+\beta} - s^{\nu-1+\beta}\right]\right\}. \tag{7}$$

Using (1), we have

$$\begin{aligned} &\left|\sum_{n=0}^{\infty}\frac{a_n}{n!}\sum_{\nu=0}^{n}\binom{n}{\nu}\frac{(-s)^{n-\nu}}{\beta+\nu-1}(s+\tau)^{\nu}\right| \\ &\quad\le K\sum_{n=0}^{\infty}|\sigma s|^n\sum_{\nu=0}^{n}\binom{n}{\nu}\frac{|s|^{-\nu}}{|\beta+\nu-1|}|s+\tau|^{\nu} \\ &\quad\le K\sum_{n=0}^{\infty}|\sigma s|^n\sum_{\nu=0}^{n}\binom{n}{\nu}\frac{1}{|\beta+\nu-1|}\left|1+\frac{\tau}{s}\right|^{\nu} \\ &\quad\le \frac{K}{\min\limits_{\nu\ge -1}|\beta+\nu|}\sum_{n=0}^{\infty}|\sigma s|^n\sum_{\nu=0}^{n}\binom{n}{\nu}\left|1+\frac{\tau}{s}\right|^{\nu} \\ &\quad= \frac{K}{\min\limits_{\nu\ge -1}|\beta+\nu|}\sum_{n=0}^{\infty}|\sigma s|^n\left(1+\left|1+\frac{\tau}{s}\right|\right)^n \\ &\quad\le \frac{K}{\min\limits_{\nu\ge -1}|\beta+\nu|}\sum_{n=0}^{\infty}(2\,|\sigma s| + \tfrac{1}{4})^n. \end{aligned} \tag{7a}$$

In the circle $|s| < 1/(4\sigma)$

$$\sum_{n=0}^{\infty}\frac{a_n}{n!}\sum_{\nu=0}^{n}\binom{n}{\nu}\frac{(-s)^{n-\nu}}{\beta+\nu-1}(s+\tau)^{\nu} = \sum_{n=0}^{\infty}G_n(s)$$

is, from (7a), a uniformly convergent polynomial series and consequently represents an analytic function of s in that domain. In particular, it is analytic at $s = 0$. For the remaining part of (7) we have, using problem 9 on page 260 of Whittaker and Watson[17]:

$$\begin{aligned} &-\sum_{n=0}^{\infty}\frac{a_n}{n!}\sum_{\nu=0}^{n}\binom{n}{\nu}\frac{(-s)^{n-\nu}}{\beta+\nu-1}s^{\nu-1+\beta} \\ &\quad= -s^{\beta-1}\sum_{n=0}^{\infty}a_n(-s)^n\frac{\Gamma(\beta-1)}{\Gamma(n+\beta)} \\ &\quad= \frac{s^{\beta-1}}{\Gamma(2-\beta)}\sum_{n=0}^{\infty}a_n\Gamma(1-n-\beta)s^n. \end{aligned}$$

Thus, for β not an integer, we have

$$\tilde{f}(s) = s^{\beta-1}\left(\sum_{n=0}^{\infty}a_n s^n\Gamma(1-n-\beta)\right) + h(s), \tag{8}$$

where $h(s)$ is analytic at $s = 0$ and the series is convergent for $|s| < 1/\sigma$.

[17] E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis* (Cambridge University Press, Cambridge, England, 1950), 4th ed.

If $\beta = 0$, then from (6),

$$\tilde{f}(s) = \sum_{n=0}^{\infty} \frac{a_n}{n!} \left\{ \sum_{\substack{\nu=0 \\ \nu \neq 1}}^{n} \binom{n}{\nu} \frac{(-s)^{n-\nu}}{\nu - 1} [(s+\tau)^{\nu-1} - s^{\nu-1}] + n(-s)^{n-1} \log \frac{s+\tau}{s} \right\}$$

$$= \left(\log \frac{s+\tau}{s} \right) \sum_{n=1}^{\infty} \frac{a_n (-1)^{n-1}}{(n-1)!} s^{n-1} + \sum_{n=0}^{\infty} \frac{a_n}{n!} \sum_{\substack{\nu=0 \\ \nu \neq 1}}^{n} \binom{n}{\nu} \frac{(-s)^{n-\nu}}{\nu - 1} [(s+t)^{\nu-1} - s^{\nu-1}]. \quad (9)$$

As before, the series

$$\sum_{n=0}^{\infty} \frac{a_n}{n!} \sum_{\substack{\nu=0 \\ \nu \neq 1}}^{n} \binom{n}{\nu} \frac{(-s)^{n-\nu}}{\nu - 1} (s+\tau)^{\nu-1}$$

is uniformly convergent to a function analytic for $|s| < 1/(4\sigma)$. The omitted term corresponding to $\nu = 1$ is not essential. From the other contribution to the double sum in (9) the $n = 0$ term is explicitly evaluated as a_0/s and we now examine the remainder. Again using (1), we have

$$\left| \sum_{n=1}^{\infty} \frac{a_n}{n!} \sum_{\substack{\nu=0 \\ \nu \neq 1}}^{n} \binom{n}{\nu} \frac{(-s)^{n-\nu}}{\nu - 1} s^{\nu-1} \right|$$

$$\leq |s|^{-1} K \sum_{n=1}^{\infty} |\sigma s|^n \sum_{\substack{\nu=0 \\ \nu \neq 1}}^{n} \binom{n}{\nu} \frac{1}{|\nu - 1|}$$

$$\leq |s|^{-1} K \sum_{n=1}^{\infty} |\sigma s|^n \sum_{\nu=0}^{n} \binom{n}{\nu} = |s|^{-1} K \sum_{n=1}^{\infty} |\sigma s|^n 2^n$$

$$= 2K\sigma \sum_{n=0}^{\infty} |2\sigma s|^n.$$

Thus if $|s| < 1/(2\sigma)$,

$$\sum_{n=1}^{\infty} \frac{a_n}{n!} \sum_{\substack{\nu=0 \\ \nu \neq 1}}^{n} \binom{n}{\nu} \frac{(-s)^{n-\nu}}{\nu - 1} s^{\nu-1}$$

is uniformly convergent to a function analytic for $|s| < 1/(2\sigma)$. Thus, from (9), we have for $\beta = 0$,

$$\tilde{f}(s) = \log s \sum_{n=1}^{\infty} \frac{a_n (-1)^n}{(n-1)!} s^{n-1} + \frac{a_0}{s} + (\text{function analytic at } s = 0), \quad (10)$$

where the series is convergent for $|s| < 1/\sigma$.

If β is an integer, one writes for $t > k$:

$$F(t) = \sum_{i=0}^{-\beta-1} a_i t^{-\beta-i} + \sum_{i=-\beta}^{N} a_i t^{-\beta-i} + t^{-\beta} R_N(t).$$

The first summand on the right side is treated separately. The theory developed above for the case $\beta = 0$ can be applied to the second summand and remainder by defining $b_i \equiv a_{i-\beta}$, except that the estimates on the b_i now include an extra factor $(\sigma')^i$ which can be absorbed into the σ of the hypothesis.

To complete the proof, we show that the Laplace transform of $F(t) - \tilde{F}(t)$ is analytic at $s = 0$. Integration by parts $N+1$ times of the formula (1a) gives

$$\tilde{F}(t) = t^{-\beta} \left[\sum_{n=0}^{N} \frac{a_n}{t^n} - e^{-t\tau} \sum_{n=0}^{N} \frac{a^{(n)}(\tau)}{t^n} + t^{-N} \int_0^{\tau} e^{-tu} a^{(N+1)}(u)\, du \right],$$

and hence

$$F(t) - \tilde{F}(t) = \left[O\left(N! \left(\frac{\sigma}{t} \right)^{N+1} \right) + e^{-t\tau} \sum_{n=0}^{N} \frac{a^{(n)}(\tau)}{t^n} - t^{-N} \int_0^{\tau} e^{-tu} a^{(N+1)}(u)\, du \right] t^{-\beta}. \quad (11)$$

We need to estimate

$$a^{(n)}(u) = \sum_{\nu=n}^{\infty} \frac{a_\nu u^{\nu-n}}{(\nu - n)!}$$

We have

$$|a^{(n)}(u)| \leq K \sum_{\nu=n}^{\infty} \frac{\nu!\, \sigma^\nu}{(\nu - n)!} |u|^{\nu-n}$$

$$= Kn! \sum_{\nu=n}^{\infty} \binom{\nu}{n} \sigma^\nu |u|^{\nu-n}$$

$$= Kn! |u|^{-n} \sum_{p=0}^{\infty} \binom{p+n}{n} (\sigma |u|)^{p+n}$$

$$= Kn!\, \sigma^n (1 - \sigma|u|)^{-(n+1)}. \quad (12)$$

Using (12), we have for $t > k$ that

$$e^{-t\tau} \sum_{n=0}^{N} \frac{a^{(n)}(\tau)}{t^n} = O\left(e^{-t\tau} \sum_{n=0}^{N} \frac{n!}{t^n} \left(\tfrac{4}{3}\sigma \right)^n \right)$$

and

$$t^{-N} \int_0^{\tau} e^{-tu} a^{(N+1)}(u)\, du$$

$$= O\left(t^{-N} \sigma^{N+1} (N+1)! \int_0^{\tau} e^{-tu} (1 - \sigma u)^{-N-2}\, du \right)$$

$$= O\left(t^{-N} \sigma^{N+1} (N+1)! \int_0^{\tau} (1 - \sigma u)^{-N-2}\, du \right)$$

$$= O\left(t^{-N} \sigma^{N+1} (N+1)! \frac{(1 - \sigma\tau)^{-N-1} - 1}{\sigma(N+1)} \right)$$

$$= O\left(t^{-N} \sigma^{N+1} (N+1)! \frac{(1 - \sigma\tau)^{-(N+1)}}{\sigma(N+1)} \right)$$

$$= O(N!(4\sigma/3t)^N).$$

Thus, from (11),

$$F(t) - \tilde{F}(t) = t^{-\operatorname{Re}\beta} \left\{ O\left(N! \left(\frac{4\sigma}{3t} \right)^N \right) + O\left(e^{-t\tau} \sum_{n=0}^{N} n! \left(\frac{4\sigma}{3t} \right)^n \right) \right\} \quad (13)$$

uniformly in N and $t > k$.

We want to show that $F(t) - \tilde{F}(t)$ falls off sufficiently rapidly (exponentially) as $t \to \infty$. Since the two functions on the right side of (13) are functions of t and N, with the bound being uniform in N, we may choose any value of N giving the desired result. The choice is made by finding, for each value of t, the N value which (approximately) minimizes the first term on the right side of (13). Put $\tilde{t} = 3t/4\sigma$ and

$$g(N, \tilde{t}) = \Gamma(N + 1)\tilde{t}^{-N}. \tag{14}$$

Contour lines of g are shown in Fig. 1 where we now regard N as a continuous variable. We have that

$$\begin{aligned}\partial g/\partial N &= g[-\log \tilde{t} + (d/dN) \log \Gamma(N + 1)] \\ &= g[-\log \tilde{t} + \psi(N + 1)].\end{aligned}$$

Along the curve in the $(\tilde{t}, N)$ plane defined by $\partial g/\partial N = 0$ or

$$\tilde{t} = \exp [\psi(N + 1)] \tag{15}$$

g has a minimum value for fixed $\tilde{t}$. A formula in Whittaker and Watson[17] (page 248) gives

$$\psi(N + 1) - \log N = \log \left(1 + \frac{1}{N}\right) + \frac{1}{2(N + 1)} - \int_0^\infty \left[\frac{1}{2} - \frac{1}{t} + \frac{1}{e^t - 1}\right] e^{-t(N+1)}\, dt. \tag{16}$$

As $N \to \infty$ the right side of (16) $\to 0$. Thus from (15)

$$\tilde{t} \to N$$

along the line $\partial g/\partial N = 0$ as $N \to \infty$. Hence choose N to be the largest integer in $\tilde{t} = 3t/4\sigma$. Using

$$\Gamma(N + 1) = (2\pi)^{\frac{1}{2}} N^{N+\frac{1}{2}} e^{-N} e^{\theta/12N}$$

for $N > 0$, $0 < \theta < 1$,

$$N^{\frac{1}{2}} = O(e^{\delta N})$$

for any $\delta > 0$, uniformly in $N > 0$, and

$$e^{\theta/12N} = O(1)$$

gives

$$\Gamma(N + 1) = O[N^N e^{-(1-\delta)N}]$$

and therefore

$$g(\tilde{t}, t) = O\{e^{-(1-\delta)3t/4\sigma}\}. \tag{17}$$

We now examine the second term on the right side of (13). For $1 \leq n \leq N = [3t/4\sigma]$ we have

$$n!\,(4\sigma/3t)^n \leq (n - 1)!\,(4\sigma/3t)^{n-1}.$$

Hence

$$\begin{aligned}e^{-tr} \sum_{n=0}^{N} n! \left(\frac{4\sigma}{3t}\right)^n &\leq e^{-tr}(N + 1) \leq e^{-t/4\sigma}\left(\frac{3t}{4\sigma} + 1\right) \\ &= O(e^{-(1-\delta)t/4\sigma}).\end{aligned} \tag{18}$$

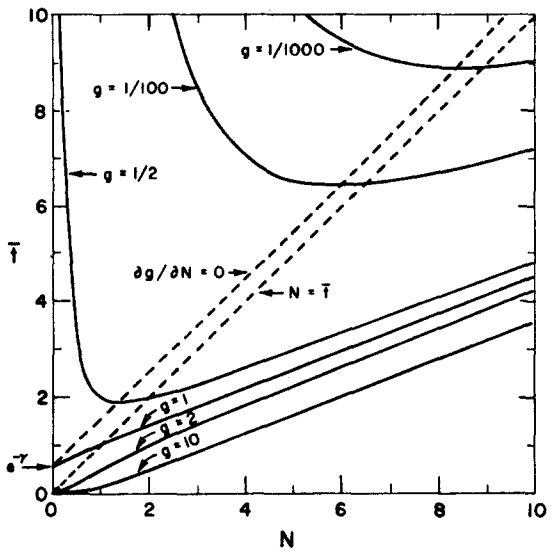


FIG. 1. Contour lines of $g = \Gamma(N + 1)\tilde{t}^{-N}$.

The combination of (13), (17), and (18) together with the entirety of $\int_0^k e^{-st}F(t)\, dt$ shows that $L\{F(t) - \tilde{F}(t)\}$ is analytic for $\operatorname{Re} s > -\frac{1}{4}\sigma$. Thus

$$L\{F(t)\} = L\{\tilde{F}(t)\} + \tilde{h}(s),$$

where $\tilde{h}(s)$ is analytic for $\operatorname{Re} s > -1/4\sigma$. Use of (8) and (10) completes the proof of the theorem.

We remark that some of the work involved in proving Theorem 1 involves questions of convergence of polynomial series near $s = 0$. Such questions arise in a very similar context in the excellent book of Kline and Kay,[18] where it would seem they are not adequately dealt with. See in particular page 309, lines 11 and 12 of Ref. 18.

III. EXAMPLE AND DISCUSSION

It is known[19] that

$$L\{e^t \operatorname{Ei}(-t)\} \equiv L\left\{-e^t \int_t^\infty e^{-u}u^{-1}\, du\right\} = \frac{(\log s)}{(1 - s)},$$

which is analytic in the sector $|\arg s| < 2\pi$ except for the logarithmic branch point at $s = 0$. We have, by integration by parts of the integral N times, that

$$e^t \operatorname{Ei}(-t) = \sum_{i=1}^{N} (-1)^i \frac{(i - 1)!}{t^i} + (-1)^{N+1} N!\, e^t \int_t^\infty \frac{e^{-u}}{u^{N+1}}\, du$$

and hence $|R_N(t)| \leq N!\, t^{-(N+1)}$ for $t > 0$. Since

[18] M. Kline and I. W. Kay, *Electromagnetic Theory and Geometric Optics* (Interscience Publishers, Inc., New York, 1965).
[19] See Ref. 9, Vol. II, p. 166.

$e^t \operatorname{Ei}(-t)$ is integrable on $(0, k)$ for any $k > 0$ the theorem applies with $\beta = 0$, $a_0 = 0$, and $a_i = (-1)^i(i-1)!$, $i \geq 1$, and implies that

$$L\{e^t \operatorname{Ei}(-t)\} = \log s\left(\sum_{i=1}^{\infty} s^{i-1}\right) + h(s)$$

$$= \frac{(\log s)}{(1-s)} + h(s),$$

where $h(s)$ is analytic at $s = 0$. In this case $h(s) \equiv 0$.

To show the requirement of uniformity in the estimate of R_N in the theorem cannot be completely dispensed with, consider $f(s) = L\{\exp(-t^{\frac{1}{2}})\}$. Re $s = 0$ is obviously the axis of convergence of $f(s)$. Also, since $\exp(-t^{\frac{1}{2}}) > 0$, $s = 0$ must be a singular point of $f(s)$.[20] But $\exp(-t^{\frac{1}{2}}) \sim \sum_{n=0}^{\infty} 0/t^n$ and the theorem would say (without the uniformity requirement) that $f(s)$ is analytic at $s = 0$. That

$$\exp(-t^{\frac{1}{2}}) \neq O(N!(\sigma/t)^{N+1})$$

for any $\sigma > 0$ uniformly in N and $t >$ some $T > 0$ is clear from Sec. II, since choosing $N = t/\sigma$ makes the right side decrease as $\exp(-t/\sigma)$.

IV. STRONG WATSON CONDITION AND UNIQUENESS OF SINGULAR POINT

In this section we state a condition on F (to be called strong Watson condition) under which $s = 0$ is the only singular point of $L\{F(t)\}$ in some sector $-\pi - \lambda < \arg s < \pi + \lambda$ $(\lambda > 0)$. And we obtain the continuation of $L\{F(t)\}$ through this sector.

We say that $\hat{F}(z)$ satisfies the strong Watson condition if $\hat{F}$ is regular in the set (see Fig. 2):

$$D(\lambda, k) = \{z \mid |z| > k, \quad |\arg z| \leq \tfrac{1}{2}\pi + \lambda\},$$

where k is a non-negative number and λ is a positive number, and

$$\hat{F}(z) = \sum_{i=0}^{N} a_i z^{-i} + R_N(z), \tag{19}$$

where

$$R_N(z) = O\{N!\,(\sigma/|z|)^{N+1}\} \tag{20}$$

uniformly in N and z in $D(\lambda, k)$ for some fixed $\sigma > 0$.

Watson[12,21] shows that if $\hat{F}(z)$ satisfies the strong Watson condition, then the Borel transform of the series (19):

$$a(t) = \sum_{n=0}^{\infty} \frac{a_n}{n!} t^n \tag{21}$$

can be continued in the t plane so that $a(t)$ is regular

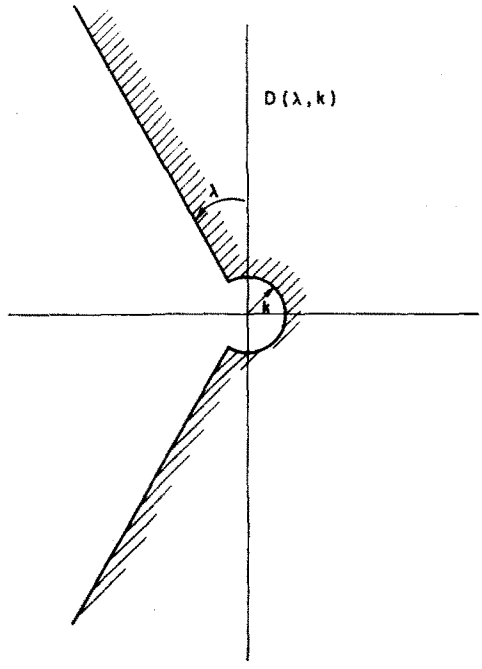


FIG .2. The region $D(\lambda, k)$: $|z| > k$ and $|\arg z| \leq \frac{1}{2}\pi + \lambda$.

in the angle $|\arg t| < \lambda$; and

$$\hat{F}(z) = \int_0^{\infty} e^{-w} a\left(\frac{w}{z}\right) dw$$

for $|z| > k$ and $|\arg z| < \lambda$. In addition Watson (page 309 of Ref. 12) shows that

$$\hat{F}(z) = z\int_0^{\infty} e^{-zt} a(t)\, dt \tag{22}$$

for Re $z > k$, the integral being absolutely convergent. Further, for $|\arg t| < \lambda$:

$$a(t) = \frac{1}{2\pi i}\int_L \hat{F}\left(\frac{u}{t}\right)\frac{e^u}{u}\, du, \tag{23}$$

where the contour L is the boundary of the domain $D(\nu, l)$ (see Fig. 2) with $0 < \nu < \lambda$ and $l > k\,|t|$. The contour is described from below.

For example, if $\hat{F}(t) = e^t \operatorname{Ei}(-t)$, then

$$a(t) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n} t^n = -\log(1+t).$$

We need the incomplete gamma function:

$$\Gamma(\alpha, X) = \int_X^{\infty} e^{-y} y^{\alpha-1}\, dy \quad (|X| > 0) \tag{24}$$

$X = 0$ is a branch point of Γ. When $\arg X = 0$ the path of integration is a straight line on the positive y axis. The path is to be deformed continuously in the region $|y| > 0$ as $\arg X$ varies.

[20] See Ref. 9, Vol. I, p. 153.

[21] G. H. Hardy, *Divergent Series* (Oxford University Press, Oxford, England, 1949), pp. 191–195.

We now make the following assertion.

Theorem 2: Suppose for some β with $\operatorname{Re}\beta < 1$, $z^\beta F(z)$ satisfies the strong Watson condition and $F(z)$ is integrable on $[0, k]$ and not identically zero. (a) In the sector $|\arg s| < \pi + \lambda$, the continuation of $L\{F(t)\}$ has precisely one singularity, $s = 0$, and the continuation of $L\{F(t)\}$ to all points s with $|\arg s| < \pi$ is given by

$$\int_0^{k'} e^{-st}F(t)\,dt + \int_0^\infty a(t)\frac{\Gamma(2-\beta, k'(s+t))}{(s+t)^{2-\beta}}\,dt, \quad (25)$$

where $k' > k$ and $a(t)$ is the auxiliary function defined by (19) and (21). (b) Further, suppose $k = 0$ in the strong Watson condition and $F(t) = O(|t|^\alpha)$ with $\alpha > -1$ as $t \to 0$ in $D(\lambda, 0)$. Then the continuation of $L\{F(t)\}$ to all values s in the sector $|\arg s| < \pi$ is given by

$$\Gamma(2-\beta)\int_0^\infty \frac{a(t)}{(s+t)^{2-\beta}}\,dt. \quad (26)$$

By contour change (25) and (26) can be used to continue $L\{F\}$ into the sector $|\arg s| < \pi + \lambda$.

Proof. We first give an outline of the proof, We choose any $k' > k$ and write for $L\{F(t)\}$ [with $\hat{F}(z) = z^\beta F(z)$]:

$$\begin{aligned} L\{F(t)\} &= \int_0^\infty e^{-st}F(t)\,dt \\ &= \int_0^{k'} e^{-st}F(t)\,dt + \int_{k'}^\infty e^{-st}t^{-\beta}\hat{F}(t)\,dt. \end{aligned} \quad (27)$$

We then substitute in the second integral of (27) Watson's expression (22) for $\hat{F}(t)$, justify and interchange order of integration for $\operatorname{Re} s > 0$ and finally obtain (25) above. The expression (25) provides the continuation of $L\{F(t)\}$. We then, by deforming a contour slightly, continue across the negative real axis. Formula (26) is obtained from (25) from right continuity (under suitable restrictions on F) in the variable k' at $k' = 0$.

The first integral in (27) is an entire function of s since F is integrable. Now consider the second integral in (27). Substituting for $\hat{F}$ from Watson's result (22) we have

$$\int_{k'}^\infty e^{-st}t^{-\beta}\hat{F}(t)\,dt = \int_{k'}^\infty e^{-st}t^{1-\beta}\left\{\int_0^\infty e^{-t\tilde{t}}a(\tilde{t})\,d\tilde{t}\right\}dt. \quad (28)$$

Now suppose $\operatorname{Re} s > 0$. Then

$$\begin{aligned} &\int_{k'}^\infty |e^{-st}|\,|t^{1-\beta}|\left\{\int_0^\infty |e^{-t\tilde{t}}|\,|a(\tilde{t})|\,d\tilde{t}\right\}dt \\ &\quad = \int_{k'}^\infty e^{-(\operatorname{Re} s)t}t^{1-\operatorname{Re}\beta}\left\{\int_0^\infty e^{-t\tilde{t}}|a(\tilde{t})|\,d\tilde{t}\right\}dt \\ &\quad = \int_0^\infty |a(\tilde{t})|\left\{\int_{k'}^\infty e^{-((\operatorname{Re} s)+\tilde{t})t}t^{1-\operatorname{Re}\beta}\,dt\right\}d\tilde{t} \\ &\quad = \int_0^\infty |a(\tilde{t})|\frac{\Gamma(2-\operatorname{Re}\beta, k'(\operatorname{Re} s+\tilde{t}))}{(\operatorname{Re} s+\tilde{t})^{2-\operatorname{Re}\beta}}\,d\tilde{t}. \end{aligned} \quad (29)$$

We show that the last integral in (29) is convergent for $\operatorname{Re} s > 0$ and thus, by Fubini's theorem, that the interchange of order of integration in (28) is permitted and obtain for $\operatorname{Re} s > 0$ and $|\arg s| < \frac{1}{2}\pi$:

$$\begin{aligned} \int_{k'}^\infty e^{-st}t^{-\beta}\hat{F}(t)\,dt &= \int_0^\infty a(\tilde{t})\left\{\int_{k'}^\infty e^{-(s+\tilde{t})t}t^{1-\beta}\,dt\right\}d\tilde{t} \\ &= \int_0^\infty a(\tilde{t})\frac{\Gamma(2-\beta, k'(s+\tilde{t}))}{(s+\tilde{t})^{2-\beta}}\,d\tilde{t}. \end{aligned} \quad (30)$$

We now prove the following lemma.

Lemma: If $a(t)$ is such that $\int_0^\infty e^{-k'u}a(u)\,du$ is absolutely convergent, then the last integral in (29) is convergent for $\operatorname{Re} s > 0$; the last integral in (30) is uniformly convergent in any closed bounded region S where $|\arg s| < \pi$ and $s \neq 0$. Hence the last integral in (30) is analytic for $|\arg s| < \pi$, $s \neq 0$.

Proof. We record an estimate of[22] $\Gamma(\alpha, X)$: for $|\arg X| < \frac{3}{2}\pi$, one has

$$\Gamma(\alpha, X) = X^{\alpha-1}e^{-X}[1 + O(X^{-1})]$$

as $|X| \to \infty$.

To prove the assertion about uniform convergence, write (using the above estimate of Γ)

$$\begin{aligned} &\int_R^\infty dt\,|a(t)|\frac{|\Gamma(2-\beta, k'(s+t)|}{|(s+t)^{2-\beta}|} \\ &\le |k'^{1-\beta}|\,|e^{-k's}|\int_R^\infty \left|\left[1 + O\left(\frac{1}{k'(s+t)}\right)\right]\right|\frac{|a(t)|\,e^{-k't}}{|(s+t)|}\,dt \\ &\le |k'^{1-\beta}|\,\frac{e^{-k'M_1}}{M_2}\left[1 + O\left(\frac{1}{k'M_2}\right)\right]\int_R^\infty |a(t)|\,e^{-k't}\,dt, \end{aligned} \quad (31)$$

where

$$M_1 = \min_{s\in S} \operatorname{Re}(s)$$

and

$$M_2 = \min_{\substack{s\in S \\ \operatorname{Re} s < 0}} |\operatorname{Im} s|.$$

To prove the last integral in (29) is convergent for $\operatorname{Re} s > 0$, replace in (31) β by $\operatorname{Re}\beta$, s by $\operatorname{Re} s$, M_1 by $\operatorname{Re} s$, and M_2 by $|s|$. This completes the proof of the lemma.

From the absolute convergence of the integral on the right side of (22) for $\operatorname{Re} z > k$ follows the absolute convergence of

$$\frac{\hat{F}(k')}{k'} = \int_0^\infty e^{-k't}a(t)\,dt.$$

[22] M. Abramowitz and I. A. Stegun, in *Handbook of Mathematical Functions* (U.S. Department of Commerce, Washington, 1964), p. 263, formula 6.5.32.

It follows from the convergence of the last integral in (29) that the interchange of order of integration which occurs in (30) is justified. It follows from the uniform convergence of (30) (see Titchmarsh[23]) that it represents the continuation of

$$\int_{k'}^{\infty} e^{-st} t^{-\beta} \hat{F}(t)\, dt$$

to the s plane cut along the negative axis.

Since $a(t)$ is analytic in a sector containing the positive axis (Watson's result) we continue (30) over the negative s axis by writing (30) in the form

$$\left\{\int_0^{-s-\xi} + \oint_{-s-\xi}^{-s+\xi} + \int_{-s+\xi}^{\infty}\right\} \frac{a(t)\Gamma(2-\beta, k'(s+t))}{(s+t)^{2-\beta}}\, dt,$$

where ξ is a small positive number and $\oint$ denotes contour integration over a small semicircle about $-s$ contained in the sector of analyticity of $a(t)$ ($|\arg t| < \lambda$). (This device is used in Kline and Kay.[24]) This completes the proof of part (a) of the theorem.

We now consider part (b) of the theorem. We need first to show that under the hypothesis on F we have that

$$a(t) = O(t^{-\alpha-\beta}) \tag{32}$$

for real t as $t \to +\infty$. To this end we express the integral in (23) as $a(t) = \sum_{i=1}^{3} A_i(t)$, where $A_1(t)$ is the part of the integral taken along the circular part of the contour L, $A_2(t)$ is the part taken along the upper straight portion of the contour L, and $A_3(t)$ is that part taken along the lower straight portion. This contour is shown in Fig. 2 if λ is replaced by ν and k by l. We have on change of variable to $u = le^{i\theta}$ that

$$A_1(t) = \frac{1}{2\pi} \int_{\theta=-\frac{1}{2}\pi-\nu}^{\frac{1}{2}\pi+\nu} \hat{F}\left(\frac{le^{i\theta}}{t}\right) \exp(le^{i\theta})\, d\theta$$

and hence

$$|A_1(t)| \le \frac{e^l(\nu + \frac{1}{2}\pi)}{\pi} \max_{-\frac{1}{2}\pi-\nu<\theta<\nu+\frac{1}{2}\pi} \left|\hat{F}\left(\frac{le^{i\theta}}{t}\right)\right|.$$

Because of the analyticity hypothesis on F, l can be any positive number. Hence we can write

$$|A_1(t)| = O((l/t)^{\alpha+\beta}) = O(t^{-\alpha-\beta})$$

as $t \to +\infty$. For $A_2(t)$ we have on putting $u = yt \times \exp[i(\frac{1}{2}\pi + \nu)]$:

$$A_2(t) = \frac{1}{2\pi i} \int_{y=l/t}^{\infty} \hat{F}(ye^{i(\frac{1}{2}\pi+\nu)}) e^{yt\exp[i(\frac{1}{2}\pi+\nu)]} \frac{dy}{y}$$

and hence

$$|A_2(t)| \le \frac{1}{2\pi} \max_{l/t\le y<\infty} |\hat{F}(ye^{i(\frac{1}{2}\pi+\nu)})|\, |\tilde{\mathrm{E}}\mathrm{i}\,[l\cos(\tfrac{1}{2}\pi+\nu)]|$$
$$= O(t^{-\alpha-\beta}).$$

Similarly, $A_3(t) = O(t^{-\alpha-\beta})$ as $t \to +\infty$. From this we have the estimate (32).

For $k' \ge 0$ we have for large R:

$$\left|\int_R^{\infty} a(t) \frac{\Gamma(2-\beta, k'(s+t))}{(s+t)^{2-\beta}}\, dt\right|$$
$$\le \int_R^{\infty} \frac{|a(t)|\,|\Gamma(2-\beta, k'(s+t))|}{|(s+t)^{2-\beta}|}\, dt$$
$$= O\left\{\int_R^{\infty} |t^{-\alpha-\beta}| \left|\int_{k'(s+t)}^{\infty} e^{-y} y^{1-\beta}\, dy\right| \frac{1}{|(s+t)^{2-\beta}|}\, dt\right\}. \tag{33}$$

In the y integral in (33) put $y = u + k'(s+t)$. The last expression in (33) becomes

$$O\left\{\int_R^{\infty} \frac{|t^{-\alpha-\beta}|}{|(s+t)^{2-\beta}|} \times \left|\int_0^{\infty} e^{-[u+k'(s+t)]}[u+k'(s+t)]^{1-\beta}\, du\right| dt\right\}$$
$$= O\left\{\int_R^{\infty} \frac{|t^{-\alpha-\beta}|}{|(s+t)^{2-\beta}|} e^{-k't} \times \int_0^{\infty} e^{-u}\,|[u+k'(s+t)]^{1-\beta}|\, du\, dt\right\}$$
$$= O\left\{\int_R^{\infty} \frac{|t^{-\alpha-\beta}|}{|(s+t)^{2-\beta}|} e^{-k't} \times \int_0^{\infty} e^{-u}[|u+k'(s+t)|]^{1-\operatorname{Re}\beta}\, du\, dt\right\}$$
$$= O\left\{\int_R^{\infty} t^{-2-\alpha}\, dt\right\} = O\left\{\frac{1}{R^{1+\alpha}}\right\}$$

since $\alpha > -1$ and $\operatorname{Re}\beta < 1$. The third equality is not trivial. From the above follows the uniform convergence in $k' \ge 0$ of the second integral in (25).

From the uniform convergence it follows that the second integral in (25) is right-continuous at $k' = 0$. Hence (25) holds for $k' = 0$ under the hypothesis on $\hat{F}$ given in part (b) of Theorem 2. This gives (26).

Alternatively, one could use the estimate (32) of $a(t)$ and apply Theorem 4a on page 334 of Widder's book[25] to obtain (26). This completes the proof of the theorem.

To show there exist F for which (25) is appropriate, but (26) is not, consider $F(z) = 1/(z-b)$ where $\operatorname{Re} b > 0$. Choose $k = 2|b|$. Then $a(t) = (\exp(bt) - 1)/b$ and (26) fails to converge for this $a(t)$.

[23] E. C. Titchmarsh, *The Theory of Functions* (Oxford University Press, Oxford, England, 1939) 2nd ed., p. 100.

[24] Reference 18, pp. 288–291.

[25] D. V. Widder, *The Laplace Transform* (Princeton University Press, Princeton, N.J., 1946).

More generally, if $F(z)$ has any singularity in the right half of the z plane, (26) fails to converge. Now suppose $F(z)$ is analytic in the right half of the z plane and has only a finite number of singularities on the imaginary axis, say at z_k ($k = 1, 2, \cdots, n$). Suppose $F(z) = O(|z_k - z|^{\alpha_k})$ as $z \to z_k$ ($\mathrm{Re}\, z \geq 0$). If $z_k = 0$ as above, we require $\alpha_k > -1$. If $z_k \neq 0$, then if we require $\alpha > -3$, (26) will still converge; i.e., (26) will converge under hypotheses weaker than stated in Theorem 2.

We remark that Theorem 2 could perhaps be generalized by using Nevanlinna's generalization of Watson's result.[26] For example, perhaps λ could be zero. For a review of Nevanlinna's paper, see Ref. 27.

V. EXAMPLES OF THEOREM II

In Sec. III, $F(t) = e^t \,\mathrm{Ei}\,(-t)$ was shown to satisfy Theorem 1. It is now shown[28] that it also satisfies Theorem 2 with k equal to zero and $\lambda < \pi$, i.e., $|t| > 0$, $|\arg t| < \frac{3}{2}\pi$. One way to do this is to express the exponential integral in terms of the confluent hypergeometric function as[1]

$$-e^t \,\mathrm{Ei}\,(-t) = \Psi(1, 1; t),$$

and then the result is a special case of the same property for $\Psi(a, c; t)$, at least for $\mathrm{Re}\, a > 0$. Instead we demonstrate the result directly, using the form of the remainder from Sec. III,

$$R_N(t) = (-1)^{N+1} N!\, e^t \int_t^\infty \frac{e^{-u}}{u^{N+1}}\, du.$$

If $\frac{1}{2}\pi < \arg t < \frac{3}{2}\pi$, then choose the contour shown in Fig. 3(a), which is a straight line starting at $u = t$, and passing through $u = i\,|t|$. (Closing the contour at infinity contributes nothing.) This is equivalent to the change of variable $u = t + ye^{i\theta}$, with $\theta = \frac{1}{2}\arg t - \frac{1}{4}\pi$, and the real variable y ranging from 0 to $+\infty$. Examination of Fig. 3(a) shows that the minimum value of $|u|$ on the contour is $|t| \cos\theta$, and therefore

$$|R_N(t)| \leq \frac{N!\, e^{\mathrm{Re}\, t}}{|t|^{N+1}(\cos\theta)^{N+1}} \int_0^\infty e^{-\mathrm{Re}\, t} e^{-y\cos\theta}\, dy$$

or

$$|R_N(t)| \leq \frac{N!}{|t|^{N+1}[\cos(\frac{1}{2}\arg t - \frac{1}{4}\pi)]^{N+2}}. \qquad (34)$$

If $0 \leq \arg t \leq \frac{1}{2}\pi$, then choose $\theta = 0$ [see Fig. 3(b)], and the cosine factor in Eq. (34) is replaced by unity. On this horizontal contour $|u| \geq |t|$. For $\arg t < 0$ all contours are reflected in the real u axis. The result, therefore, is that

$$|R_N(t)| \leq AN!\, \sigma^{N+1}/|t|^{N+1},$$

where $A = \sigma = \sec\frac{1}{2}\lambda$, for $|t| > 0$ and $|\arg t| \leq \frac{1}{2}\pi + \lambda$. If $|\arg t| \leq \frac{1}{2}\pi$, $A = \sigma = 1$. Evidently, λ must be less than π.

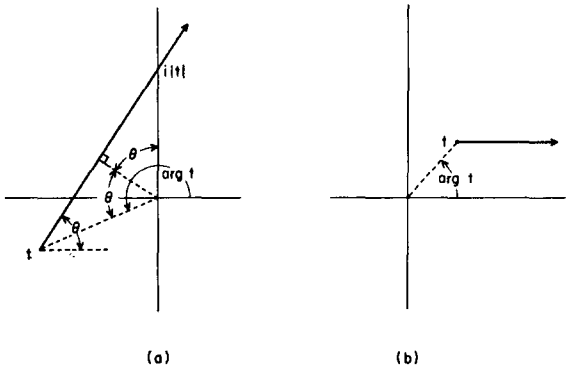


FIG. 3. Contours of integration in the u plane which establish that $e^t \,\mathrm{Ei}\,(-t) = -e^t \int_t^\infty e^{-u}u^{-1}\, du$ satisfies strong Watson conditions for $|t| > 0$, $|\arg t| < \frac{3}{2}\pi$. (a) $\frac{1}{2}\pi \leq \arg t < \frac{3}{2}\pi$. θ is chosen to be $(\frac{1}{2}\arg t - \frac{1}{4}\pi)$. The minimum value of $|u|$ on the contour is $|t|\cos\theta$. (b) $0 \leq \arg t \leq \frac{1}{2}\pi$. θ is chosen to be zero, i.e., the contour is horizontal. The contours are reflected in the real u axis if $\arg t < 0$.

According to the first part of Theorem 2, the Laplace transform of $e^t \,\mathrm{Ei}\,(-t)$ has only one singularity, $s = 0$, for $|\arg s| < \pi + \lambda < 2\pi$. Indeed $\log s/(1 - s)$, which is the Laplace transform in question, does have this property. According to the second part of Theorem 2, the Laplace transform just given is equal to

$$\int_0^\infty \frac{a(t)}{(s + t)^2}\, dt$$

for $|\arg s| < \pi$, where $a(t)$ was shown in Sec. IV for this example to equal $-\log(1 + t)$, and the principal value is to be used. Integration by parts gives

$$\begin{aligned} -\int_0^\infty \frac{\log(1 + t)}{(s + t)^2}\, dt &= \frac{\log(1 + t)}{s + t}\bigg|_0^\infty - \int_0^\infty \frac{1}{(s + t)(1 + t)}\, dt \\ &= 0 + \frac{1}{1 - s}\int_0^\infty \left[\frac{1}{1 + t} - \frac{1}{s + t}\right] dt \\ &= \frac{\ln s}{1 - s}, \end{aligned}$$

confirming the result.

For the confluent hypergeometric function

$$\Psi(a, c; x) = \frac{1}{\Gamma(a)} \int_0^{\infty e^{i\phi}} e^{-xt} t^{a-1}(1 + t)^{c-a-1}\, dt \qquad (35)$$

with $\mathrm{Re}\, a > 0$, $|\phi| < \pi$, and $|\phi + \arg x| < \frac{1}{2}\pi$, the

[26] F. Nevanlinna, Ann. Acad. Sci. Fenn. (A) **12**, 1 (1916).
[27] L. Bieberbach, *Fortschritte der Mathematik* **46**, 1463 (1916).
[28] In Ref. 12 it was shown that $e^t \,\mathrm{Ei}\,(-t)$ satisfies the strong Watson condition for $|\arg t| < \pi$. We extend this result to $\frac{3}{2}\pi$.

validity of the strong Watson condition can be demonstrated for $|x| > 0$ and $|\arg x| \leq \frac{1}{2}\pi + \lambda$, with $\lambda < \pi$, by expanding

$$g(t) \equiv (1+t)^{c-a-1} = \sum_{n=0}^{N} \frac{g^{(n)}(0)}{n!} t^n + \frac{t^{N+1}}{N!} \int_0^1 dy (1-y)^N g^{(N+1)}(ty).$$

The result (which we do not prove here) is

$$x^a \Psi(a, c; x) = \frac{1}{\Gamma(a)} \sum_{M=0}^{N} \frac{\Gamma(M+1+a-c)}{M!\,\Gamma(1+a-c)} \times \frac{\Gamma(M+a)}{(-x)^M} + R_N(a, c, x), \quad (36)$$

where it can be shown that there exist constants A and σ independent of N and x such that

$$|R_N| < \frac{AN!\,\sigma^{N+1}}{|x|^{N+1}}, \quad (37)$$

and just as in the example of the exponential integral A and σ become infinite as λ approaches π. A and σ depend upon a and c, but can be chosen to be independent of them so long as $|a|$ and $|c|$ are bounded.

VI. GENERALIZATION AND EXAMPLE FROM SCATTERING THEORY

We now return to the problem mentioned in the Introduction, which is not a Laplace transform, but is only a slight generalization from the standpoint of the theorems in this paper. Changing from the variable s to $k = \frac{1}{2}(\mu - s)$, the function under consideration is

$$A_1(k) = \int_0^\infty e^{2kr} e^{-\mu r} H(k, r)\, dr \quad (38)$$

where

$$H(k, r) = r^{-1} e^{-2kr} [r^{L+1} e^{-kr} \Phi(a, c; 2kr)]^2, \quad (39)$$

with

$$a = L + 1 + Me^2/2k \quad \text{and} \quad c = 2L + 2.$$

As mentioned in the Introduction, $H(k, r)$ is an entire function of k. The aim is to show that $A_1(k)$ has only one singularity for $-\frac{1}{2}\pi \leq \arg k \leq \frac{1}{2}\pi$, namely at $k = \frac{1}{2}\mu$, provided one cuts the k plane from there to $+\infty$, and further to find the singular part of the function in a neighborhood of that point. $k = 0$ requires special consideration which we defer until later.

The portion of the integral (38) from $r = 0$ to any fixed R is an entire function of k, $E(k)$, and therefore we are led to consider

$$B(k) = A_1(k) - E(k) = \int_R^\infty e^{2kr} e^{-\mu r} H(k, r)\, dr. \quad (40)$$

For $k \neq 0$, use the relation[29]

$$\Phi(a, c; 2kr) = \frac{\Gamma(c)}{\Gamma(c-a)} e^{i\epsilon a\pi} \Psi(a, c; 2kr) + \frac{\Gamma(c)}{\Gamma(a)} e^{i\epsilon(a-c)\pi} e^{2kr} \Psi(c-a, c; -2kr), \quad (41)$$

where the single-valued function Φ has been expressed in terms of the multivalued functions Ψ, and one must choose

$$\arg(-2kr) = \arg(2kr) - \epsilon\pi \quad (42)$$

with $\epsilon = +1$ everywhere, or $\epsilon = -1$ everywhere. In Ref. 1 ϵ is taken as $+1$ if $0 < \arg(2kr) < \pi$, and $\epsilon = -1$ if $-\pi < \arg(2kr) < 0$, but either one of those relations can be continued in $\arg(2kr)$ indefinitely, with ϵ held fixed. We understand $\arg(2kr)$ to be $\arg k + \sigma$, with $\sigma = \arg r$.

From Eqs. (39), (40), and (41),

$$B(k) = B_1(k) + B_2(k) + B_3(k), \quad (43)$$

where

$$B_1(k) = \frac{\Gamma^2(c)}{\Gamma^2(c-a)} e^{2i\epsilon a\pi} \times \int_R^\infty e^{-\mu r} e^{-2kr} r^{2L+1} \Psi^2(a, c; 2kr)\, dr,$$

$$B_2(k) = \frac{2\Gamma^2(c)}{\Gamma(c-a)\Gamma(a)} e^{i\epsilon(2a-c)\pi} \times \int_R^\infty e^{-\mu r} r^{2L+1} \Psi(a, c; 2kr) \Psi(c-a, c; -2kr)\, dr, \quad (44)$$

and

$$B_3(k) = \frac{\Gamma^2(c)}{\Gamma^2(a)} e^{2i\epsilon(a-c)\pi} \times \int_R^\infty e^{-\mu r} e^{2kr} r^{2L+1} \Psi^2(c-a, c; -2kr)\, dr.$$

Since all three integrands as well as the multiplying functions outside the integrals are analytic functions of k for all finite k (except $k = 0$), the functions $B_i(k)$ will be analytic in any region of k for which the integrals converge uniformly. If an arbitrarily small region containing $k = 0$ is avoided, then one can show directly from (35) that the Ψ functions can be bounded uniformly in k by a constant multiplied by a fixed power of $|r|$, for all r on the path of integration, provided: $|\arg k + \sigma| \leq \frac{1}{2}\pi + \lambda < \frac{3}{2}\pi$ for $\Psi(a, c; 2kr)$; and

$$|\arg k + \sigma - \epsilon\pi| \leq \tfrac{1}{2}\pi + \lambda < \tfrac{3}{2}\pi$$

for $\Psi(c-a, c; -2kr)$.

To begin, consider region ① of the k plane:

[29] Reference 1, p. 259, formula (7).

$0 \leq \operatorname{Re} k \leq \frac{1}{2}\mu - \delta$, $\delta > 0$, $\operatorname{Im} k \geq 0$, but $k \neq 0$. See Fig. 4(a). Choose $\epsilon = +1$. With $\sigma = 0$ in all three integrals of (44), one has

$$|\arg k + \sigma| \quad \text{and} \quad |\arg k + \sigma - \epsilon\pi|$$

both less than $\frac{3}{2}\pi$, and the exponentials in the integrands are all decreasing. All three integrals converge uniformly in k, and therefore represent analytic functions of k in this region.

We want to continue $B_1 + B_2 + B_3$ beyond region ①. For B_1 and B_2 the expressions in (44) show that these two functions have no singularities for $-\frac{1}{2}\pi \leq \arg k \leq \frac{1}{2}\pi$. It is also clear from (44) that $B_3(k)$ is not analytic at $k = \frac{1}{2}\mu$ since the exponential is equal to unity and differentiating with respect to k enough times must lead to a divergent integral. To continue B_3 around $k = \frac{1}{2}\mu$, first rotate the path of integration [Fig. 4(b)] in the r plane. See Fig. 4(c). In region ① of the k plane any value of σ may be used such that $0 \leq \sigma < \frac{1}{2}\pi$. By taking σ almost $\frac{1}{2}\pi$, one effects the continuation of B_3 clockwise around $k = \frac{1}{2}\mu$ into region ② (almost to the real k axis). A further rotation of the path of integration continues B_3 clockwise across the real k axis and this is sufficient for the dispersion relation application, but this process can be repeated for more than one sheet around $k = \frac{1}{2}\mu$, finally being stopped by the requirement on $|\arg k + \sigma - \epsilon\pi|$.

To continue $B_1 + B_2 + B_3$ counterclockwise around $k = \frac{1}{2}\mu$, start over in region ①, this time with $\epsilon = -1$. The reasoning goes just as before with the sign of all arguments changed.

One proves that $A_1(k)$ is analytic at $k = 0$ by showing from its powers series that $\Phi(a, c; 2kr)$ is bounded uniformly in k for $|k| < k_0$, by a constant multiplying $\exp(\sigma k_0 r)$, with σ another constant. For sufficiently small k_0 the integral (38) is uniformly convergent.

To find the singular part of $A_1(k)$ in a neighborhood of $k = \frac{1}{2}\mu$ requires a generalization of Theorem 1 with a similar proof.

Theorem 3: Suppose there exists a neighborhood of $s = 0$ in which: (i) $F(s, t)$ is an analytic function of s for all $t > k$; (ii)

$$F(s, t) = t^{-\beta(s)}\left[\sum_{i=0}^{N} \frac{a_i(s)}{t^i} + R_N(s, t)\right], \quad t > k,$$

where $a_i(s)$ and $\beta(s)$ are analytic functions of s, with $\operatorname{Re} \beta(s) < 1$; and (iii)

$$R_N(s, t) = O[N!\, \sigma^{N+1}/t^{N+1}]$$

uniformly in s (in the neighborhood), N, and $t > k$. Then if $\beta(s)$ is not an integer constant, one has for $\int_T^\infty e^{-st}F(s, t)\, dt$:

$$s^{\beta(s)-1}\left[\sum_{m=0}^{\infty} a_m(s)\Gamma(1 - m - \beta(s))s^m\right] + h(s).$$

If $\beta(s)$ is an integer constant, then

$$\int_T^\infty e^{-st}F(s, t)\, dt$$

is given by

$$s^{\beta-1}\sum_{i=0}^{-\beta} a_i(s)\Gamma(1 - i - \beta)s^i + (\log s)\left[\sum_{i=1}^{\infty} \frac{(-1)^i}{(i-1)!} a_{i-\beta}(s)s^{i-1}\right] + h(s).$$

In both cases $h(s)$ is analytic at $s = 0$. T is a fixed positive constant.

Proof. One constructs

$$a(s, u) = \sum_{m=0}^{\infty} a_m(s)u^m/m!$$

which is uniformly convergent in s for $|u| < 1/\sigma$. Define

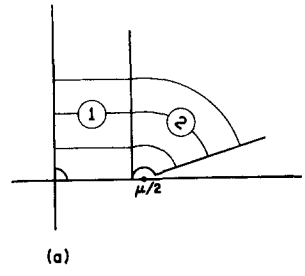
$$\tilde{F}(s, t) \equiv t^{1-\beta(s)}\int_0^\tau e^{-tu}a(s, u)\, du$$

which is an analytic function of s, and for $\operatorname{Re} s > 0$:

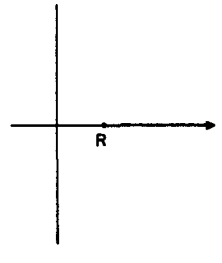
$$\tilde{f}(s) \equiv \int_0^\infty e^{-st}\tilde{F}(s, t)\, dt = \Gamma(2 - \beta(s))\int_0^\tau a(s, u)(u + s)^{\beta(s)-2}\, du.$$

The remainder of the proof proceeds as in Theorem 1.

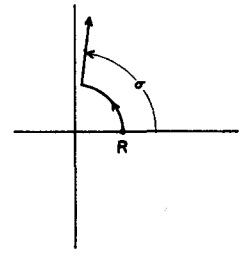
The hypotheses of Theorem 3 are met by the integral



(a)



(b)



(c)

FIG. 4. (a) The k plane showing some of the regions of analyticity of the functions $B_i(k)$ which are discussed in Sec. VI. (b) The original path of integration in the r plane for Eqs. (44). (c) A rotated path of integration in the r plane which effects the analytic continuation of $B_3(k)$ clockwise around $k = \frac{1}{2}u$ into region ② of the k plane.

for B_3 in (44), if one puts $s = u - 2k$, in a neighborhood of $k = \frac{1}{2}\mu$ (or $s = 0$) and therefore the asymptotic series for Ψ, (36), can be used to determine the singular part of $B_3(k)$ and consequently of $A_1(k)$.[2]

VII. REMAINDER ESTIMATES

Estimates on the remainder term in terms of z and N (complete bounds) of an asymptotic expansion are required to determine if the expansion satisfies a Watson condition. We are not aware of any tables of such estimates. There are investigations made by Watson,[12,30] Olver[31,32,33] and Olver and Stenger.[34] In this section we survey these estimates.

Watson[30] has analyzed an asymptotic expansion of $\log \Gamma(z + a)$. If $0 < a \leq 1$, $|z| > 1$, $|\arg z| < \pi - 2\theta$, then it is shown that

$$\log \Gamma(z + a) - (z + a - \tfrac{1}{2}) \log z + z = \sum_{n=0}^{N} \frac{a_n}{z^n} + R_N(z),$$

where

$$|R_N(z)| < K\Gamma(N + 1)[(1 + \delta) \csc (\tfrac{1}{2}\theta)/2\pi]^N |z|^{-N-1}$$

for some K and for fixed $\delta > 0$. The case of arbitrary a is also discussed and is slightly more complicated. It is not reproduced here.

For the G function of Barnes,

$$G(x + 1) = [\exp \{\tfrac{1}{2}x \log \tfrac{1}{2}\pi - \tfrac{1}{2}x(x + 1) - \tfrac{1}{2}\gamma x^2\}] \times \prod_{n=1}^{\infty} \left(1 + \frac{x}{n}\right)^n \exp (-x + \tfrac{1}{2}n^{-1}x^2)$$

and for $|z| > 1$, $|\arg z| < \pi - 2\theta$, it is shown[30] that

$$\log G(z + a) - \tfrac{1}{12} - \tfrac{1}{2}(z + a - 1) \log 2\pi + \log A - \tfrac{1}{2}\{(z + a - 1)^2 - \tfrac{1}{12}\} \log z + \tfrac{3}{4}z^2 + (a - 1)z$$

possesses an asymptotic expansion $\sum_{n=0}^{N} a_n/z^n + R_N(z)$ having the same condition as given above for the gamma function. Here A is certain definite constant having numerical value 1.28242713 $\cdots$ and γ is Euler's number. The interest in Barnes' G function arises from the fact that it satisfies the relation $G(z + 1) = \Gamma(z)G(z)$ with $G(1) = 1$.

Watson[30] also considered the Mittag-Leffler function $E_\alpha(x)$ defined by

$$E_\alpha(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(\alpha n + 1)},$$

where $\alpha > 0$. (This is an entire function satisfying a Watson condition.)

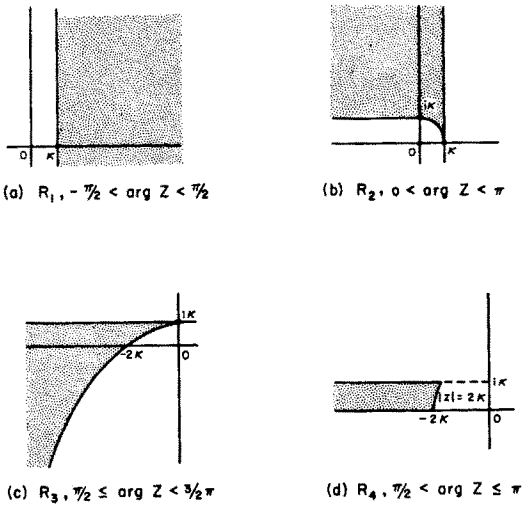


FIG. 5. Regions for bounds on remainder terms for the Whittaker functions $W_{k,m}(z)$.

Olver[32] has considered the Whittaker functions $W_{k,m}$ (or, equally, the confluent hypergeometric functions). The Whittaker functions have important specializations: Bessel functions, parabolic cylinder functions, the exponential integral, sine and cosine integrals, Fresnel integrals, error function, and incomplete gamma functions.

Put $K = 2k$ and let R_1, R_2, R_3, and R_4 be the regions shown in Fig. 5. The bounding curve in R_3 is the parabola $x^2 + 4Ky \geq 4K^2$. $\bar{R}_i$ $(i = 2, 3, 4)$ is the set of complex values conjugate to R_i. From Olver's results it follows that there exist $\bar{K}$ and $N(k, m)$ so that for $n > N(k, m)$:

$$|\epsilon_n(z)| < \bar{K}n! \, |z|^{-n}$$

for z in $R \cup R_2 \cup \bar{R}_2$ and (with $\theta = \arg z$ and φ a certain angle dependent on z):

$$|\epsilon_n(z)| < \bar{K}n! \csc^n (\theta - \varphi) \, |z|^{-n}$$

for z in $R_3 \cup \bar{R}_3$. This result overlaps that given in Sec. V concerning the confluent hypergeometric functions.

Additional results on complete error bounds on many asymptotic expansions are derived by Olver[31–33] and Olver and Stenger.[34] Generally the bounds are not directly in the form required for applications discussed in this paper but no doubt can easily be put in the Watson form we use. Older results due on

[30] G. N. Watson, Quart. J. Pure Appl. Math. **43**, 63 (1912).
[31] F. W. J. Olver, *Asymptotic Solutions of Differential Equations*, C. H. Wilcox, Ed. (John Wiley & Sons, Inc., New York, 1964), p. 163.
[32] F. W. J. Olver, J. Soc. Ind. Appl. Math. **B2**, 225 (1965).
[33] F. W. J. Olver, *Error in Digital Computation*, L. B. Rall, Ed. (John Wiley & Sons, Inc., New York, 1965), p. 65.
[34] F. W. J. Olver and F. Stenger, J. Soc. Ind. Appl. Math. **B2**, 244 (1965).

Hankel, Stieltjes, Weber, and Schläfli are given in Watson.[35]

Some general theorems relating to the calculation of complete bounds on asymptotic expansions are given by Watson.[12] In particular, the asymptotic expansion of the product of two functions is investigated. In addition, results are given on complete bounds on remainder terms of an asymptotic expansion of a function $\Phi\{f(z)\}$ where certain things are known about Φ and where complete bounds are known on remainder terms of an asymptotic expansion of $f(z)$.

Watson[12] investigates the asymptotic expansion in negative powers of z of Laplace transforms of functions $\varphi(t)$ defined by series $\varphi(t) = \sum_{i=0}^{\infty} a_i t^i/i!$ where $|a_i| < Ai!\,\sigma^i$. In addition, it is assumed that $\varphi(t)$ has no singularities in the sector $|\arg t| \leq \lambda$ and in this sector $|\varphi(t)| < K \exp\{\gamma |t|\}$ where K and γ are constants. Watson then shows that if

$$F(z) = z\int_0^{\infty} e^{-zt}\varphi(t)\,dt,$$

then $F(z)$ or its analytic continuation has the form

$$\sum_{n=0}^{N} \frac{a_n}{z^n} + R_N(z)$$

for $|\arg z| < \frac{1}{2}\pi + \lambda - \theta$ $(\theta > 0)$ and

$$\operatorname{Re}\{z \exp(-iv)\} > \gamma + 1$$

where $|v| < \lambda - \theta$; $R_N(z)$ satisfies:

$$|R_N(z)| < B(\sigma_1 \csc\theta)^N \Gamma(N+1)\,|z|^{-N-1},$$

where σ_1 is any number $> \sigma$. B does not depend on N or z.

It is possible, but apparently not completely trivial, to carry out an analysis similar to Watson's analysis for Laplace transforms of the form

$$\int_0^{\infty} e^{-zt} t^{\beta}\varphi(t)\,dt,$$

where β is an arbitrary complex number and $\varphi(t)$ satisfies the above hypotheses. This case includes the confluent hypergeometric functions discussed in Sec. V and Sec. VI.

This result by Watson together with Watson's work discussed at the beginning of Sec. IV yield the following, which shows the equivalence between functions with asymptotic expansions satisfying a strong Watson condition and functions which are Laplace transforms of certain analytic functions. More precisely, let $\lambda > 0$ and $k > 0$ be given. Then $f(z)$ is regular in $D(\lambda, k) = \{z \mid |z| > k, |\arg z| \leq \frac{1}{2}\pi + \lambda\}$ and in $D(\lambda, k)$,

$$f(z) = \sum_{i=0}^{n} a_i z^{-i} + R_n(z),$$

with $R_n(z) = O\{\Gamma(n)(\sigma/z)^{n+1}\}$ for some $\sigma > 0$ uniformly in n, and $z \in D(\lambda, k)$ if and only if there exists $\varphi(t)$ regular in $\{|t| < \sigma_1\} \cup \{|\arg t| < \lambda\}$, and $O(\exp k|t|)$ for some $\sigma_1 > 0$, such that for $\operatorname{Re} z > k$:

$$f(z) = z\int_0^{\infty} e^{-zt}\varphi(t)\,dt.$$

Perhaps a similar correspondence could be made between functions with asymptotic expansions satisfying a weak Watson condition and a different class of Laplace transforms.

VIII. PROPOSAL FOR FURTHER RESEARCH

Nevanlinna[25,26] obtains the following result (a generalization of Watson's result): Suppose $F(z)$ is an analytic function, which for $\operatorname{Re} z^k > \gamma^k$, $\gamma > 0$, $k > 0$ is regular and can be represented there by a series $\sum_1^{\infty} a_\nu/z^\nu$ and for arbitrarily small positive ξ and each $\rho' > \rho$ satisfies in $\operatorname{Re} z^k > \gamma^k + \xi$ the inequality

$$\left|F(z) - \sum_1^{n-1} a_\nu/z^\nu\right| < \Gamma\left(\frac{n}{k} + 1\right)\rho'^n\,|z|^{-n}$$

for sufficiently large values of n. Then the function $a(z)$ defined by

$$a(z) = \sum_1^{\infty} \frac{a_\nu}{\Gamma(\nu/k + 1)}\, z^{\nu/k}$$

is analytic in $|z| < 1/\rho^k$, can be continued along the positive real axis, and for $\operatorname{Re} z^k > \gamma^k$:

$$F(z) = z^k\int_0^{\infty} a(t)e^{-z^k t}\,dt.$$

Consider now the Laplace transform of $F(z)$. Proceeding formally we have

$$\begin{aligned} f(s) &= \int_0^{\infty} e^{-st}F(t)\,dt \\ &= \int_0^{\infty} e^{-st}t^k\left\{\int_0^{\infty} a(u)e^{-t^k u}\,du\right\}dt \\ &= \int_0^{\infty} a(u)\,du\int_0^{\infty} t^k e^{-(st+ut^k)}\,dt \\ &= \int_0^{\infty} a(u)\,\frac{1}{ku^{1+1/k}}\,g_k'(-su^{-1/k})\,du, \end{aligned} \tag{45}$$

[35] G. N. Watson, *Theory of Bessel Functions* (Cambridge University Press, Cambridge, England, 1950), 4th ed., pp. 205–220.

where

$$g_k(z) = \sum_{n=0}^{\infty} z^n \frac{\Gamma[1 + (1/k)n]}{\Gamma(1 + n)} \quad \left(0 \leq \frac{1}{k} \leq 1\right)$$

is a function discussed by Barnes:[36,37] For $k = 2$, $g_k(z)$ is expressible as a confluent hypergeometric function.[38]

[36] E. W. Barnes, Phil. Trans. Roy. Soc. London A206, 283 (1906).
[37] E. W. Barnes, Cambridge Phil. Trans. 20, 215–232 (1906).
[38] Bateman Manuscript Project, *Tables of Integral Transforms*, A. Erdélyi, Ed. (McGraw-Hill Book Company, Inc., New York, 1954), Vol. 1, p. 142, formula (24).

What information about the singularities of $f(s)$ can be determined from (45)?

ACKNOWLEDGMENTS

The authors thank Dr. D. Liberman for many valuable discussions on this topic.

This work was performed under the auspices of the U.S. Atomic Energy Commission.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 5 MAY 1967

Cluster Expansion of an Inverse Overlap Matrix for Solids*

TADASHI ARAI AND JOHN R. GABRIEL
Argonne National Laboratory, Argonne, Illinois

(Received 18 April 1966)

In this paper, a method is developed to compute an inverse overlap matrix based on a linked cluster expansion of a determinant. The inverse is expanded in terms of cluster integrals represented by diagrams and a recurrence relation for generation of all diagrams required is found. Thus the computation of an exact inverse overlap matrix is reduced to solving the recurrence equations self-consistently. An approach for solving the equations is suggested and bounds for errors accompanying this procedure are calculated. The method is applied to the hydrogen lattice.

I. INTRODUCTION

WHENEVER atomic functions are used as bases in calculating the electronic structure of solids, one has to deal with the difficult problem of calculating the inverse overlap matrices.[1–3]

The principle of the calculation is simple; that is, the overlap matrix S is transformed to diagonal form S', inverted and transformed back to the original representation. The inverse is then given by a Fourier transform of $(S')^{-1}$, an integral, because the transformation used is just a Fourier transformation.

Although, for a linear chain, this integral may be converted into a contour integral and evaluated in a closed form[3,4] it is doubtful if such a technique can be extended to the two- and three-dimensional cases. Gilbert[3] has discussed difficulties associated with this approach. Calais and Appel[5] have avoided the difficulties by carrying out the integration numerically. Lack of knowledge on singularities involved in the integrand, however, makes this approach somewhat uncertain. There might be possibilities that the denominator of the integrand becomes vanishingly small.

The expansion method[2] used by Löwdin converges very slowly and is not considered very useful for practical purposes. Another approach used by Löwdin[6] is to solve directly a set of linear equations which defines the inverse. A numerical calculation can be carried out only if the infinite set of equations is replaced by a set of a small number, say Q, of equations. Unfortunately, there has been no way to estimate the error that may accompany the truncation.

In this paper, we develop a linked cluster expansion for computing an inverse overlap matrix; the method is a generalization of Cauchy's expansion of a determinant[7] and also a simplification of the linked cluster expansion[8] proposed previously for computing the

* Based on work performed under the auspices of the U.S. Atomic Energy Commission.
[1] P. O. Löwdin, J. Chem. Phys. 18, 365 (1950).
[2] P. O. Löwdin, Advan. Phys. 5, 1 (1956).
[3] T. L. Gilbert, J. Math. Phys. 3, 107 (1962).
[4] Also see the exact methods of calculating cyclic matrices by P. O. Löwdin, R. Pauncz, and J. de Heer, J. Math. Phys. 1, 461 (1960); P. B. Abraham and G. Weiss, *ibid.* 3, 340 (1962).
[5] J. L. Calais and K. Appel, J. Math. Phys. 5, 1001 (1964).
[6] P. O. Löwdin, J. Chem. Phys. 19, 1579 (1951).
[7] See, for instance, A. C. Aitken, *Determinants and Matrices* (Oliver and Boyd Ltd., Edinburgh, 1959), Chap. IV.
[8] T. Arai, Phys. Rev. 134, A824 (1964); Progr. Theoret. Phys. (Kyoto) 36, 473 (1966). Henceforth, the first paper is referred to as I and the second as II.

where

$$g_k(z) = \sum_{n=0}^{\infty} z^n \frac{\Gamma[1 + (1/k)n]}{\Gamma(1 + n)} \quad \left(0 \leq \frac{1}{k} \leq 1\right)$$

is a function discussed by Barnes:[36,37] For $k = 2$, $g_k(z)$ is expressible as a confluent hypergeometric function.[38]

[36] E. W. Barnes, Phil. Trans. Roy. Soc. London A206, 283 (1906).
[37] E. W. Barnes, Cambridge Phil. Trans. 20, 215–232 (1906).
[38] Bateman Manuscript Project, *Tables of Integral Transforms*, A. Erdélyi, Ed. (McGraw-Hill Book Company, Inc., New York, 1954), Vol. 1, p. 142, formula (24).

What information about the singularities of $f(s)$ can be determined from (45)?

ACKNOWLEDGMENTS

The authors thank Dr. D. Liberman for many valuable discussions on this topic.

This work was performed under the auspices of the U.S. Atomic Energy Commission.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 5 MAY 1967

Cluster Expansion of an Inverse Overlap Matrix for Solids*

TADASHI ARAI AND JOHN R. GABRIEL
Argonne National Laboratory, Argonne, Illinois

(Received 18 April 1966)

In this paper, a method is developed to compute an inverse overlap matrix based on a linked cluster expansion of a determinant. The inverse is expanded in terms of cluster integrals represented by diagrams and a recurrence relation for generation of all diagrams required is found. Thus the computation of an exact inverse overlap matrix is reduced to solving the recurrence equations self-consistently. An approach for solving the equations is suggested and bounds for errors accompanying this procedure are calculated. The method is applied to the hydrogen lattice.

I. INTRODUCTION

WHENEVER atomic functions are used as bases in calculating the electronic structure of solids, one has to deal with the difficult problem of calculating the inverse overlap matrices.[1–3]

The principle of the calculation is simple; that is, the overlap matrix S is transformed to diagonal form S', inverted and transformed back to the original representation. The inverse is then given by a Fourier transform of $(S')^{-1}$, an integral, because the transformation used is just a Fourier transformation.

Although, for a linear chain, this integral may be converted into a contour integral and evaluated in a closed form[3,4] it is doubtful if such a technique can be extended to the two- and three-dimensional cases. Gilbert[3] has discussed difficulties associated with this approach. Calais and Appel[5] have avoided the difficulties by carrying out the integration numerically. Lack of knowledge on singularities involved in the integrand, however, makes this approach somewhat uncertain. There might be possibilities that the denominator of the integrand becomes vanishingly small.

The expansion method[2] used by Löwdin converges very slowly and is not considered very useful for practical purposes. Another approach used by Löwdin[6] is to solve directly a set of linear equations which defines the inverse. A numerical calculation can be carried out only if the infinite set of equations is replaced by a set of a small number, say Q, of equations. Unfortunately, there has been no way to estimate the error that may accompany the truncation.

In this paper, we develop a linked cluster expansion for computing an inverse overlap matrix; the method is a generalization of Cauchy's expansion of a determinant[7] and also a simplification of the linked cluster expansion[8] proposed previously for computing the

* Based on work performed under the auspices of the U.S. Atomic Energy Commission.
[1] P. O. Löwdin, J. Chem. Phys. 18, 365 (1950).
[2] P. O. Löwdin, Advan. Phys. 5, 1 (1956).
[3] T. L. Gilbert, J. Math. Phys. 3, 107 (1962).
[4] Also see the exact methods of calculating cyclic matrices by P. O. Löwdin, R. Pauncz, and J. de Heer, J. Math. Phys. 1, 461 (1960); P. B. Abraham and G. Weiss, *ibid.* 3, 340 (1962).
[5] J. L. Calais and K. Appel, J. Math. Phys. 5, 1001 (1964).
[6] P. O. Löwdin, J. Chem. Phys. 19, 1579 (1951).
[7] See, for instance, A. C. Aitken, *Determinants and Matrices* (Oliver and Boyd Ltd., Edinburgh, 1959), Chap. IV.
[8] T. Arai, Phys. Rev. 134, A824 (1964); Progr. Theoret. Phys. (Kyoto) 36, 473 (1966). Henceforth, the first paper is referred to as I and the second as II.

Heitler–London energy. The inverse is then expanded in terms of connected diagrams and a recurrence relation for generating all possible diagrams is found. Thus the computation of the inverse is reduced to solving the recurrence relation self-consistently. Since the recurrence relation is written as the set of linear equations that defines the inverse overlap matrix, the method thus developed is formally equivalent to Löwdin's second approach. In the course of the derivation, however, it becomes clear as to how the inverse should behave and what kind of approximation is involved in the truncation used. Thus bounds for errors accompanying the approximation can be calculated and the results, in turn, ensure the accuracy of the method developed here.

II. INVERSE OVERLAP MATRIX

Let R_h be the positional vector of the hth atom in a lattice composed of a large number N of atoms and denote by $\varphi_n(r - R_h)$ the one-electron wavefunction localized at the hth atom. We further impose the periodic boundary conditions that

$$\varphi_n(r - \overline{R_h + \Lambda_i}) = \varphi_n(r - R_h), \quad \text{for} \quad i = x, y, z, \tag{1}$$

where Λ_i is the length of the lattice in the i direction.

The overlap matrix is defined as the matrix whose hl element S_{hl} is given by the overlap integral

$$S_{hl} \equiv \langle h|l\rangle \equiv \int \varphi_n(r - R_h)^* \varphi_n(r - R_l)\, d\tau, \tag{2}$$

and the inverse S^{-1} is determined by

$$\sum_{k=1}^{N} (S^{-1})_{hk} S_{kl} = \delta_{hl}. \tag{3}$$

Because of the periodic boundary condition (1), $(S^{-1})_{hk}$ is a function of the vector $R_{h-k} \equiv R_h - R_k$ connecting atom k to atom h and can be denoted by $(S^{-1})_{h-k}$. Hence the inverse S^{-1} is in principle determined by solving the N linear equations for a fixed h:

$$\sum_{k=1}^{N} \langle k|l\rangle (S^{-1})_{h-k} = \delta_{hl}, \quad \text{for} \quad l = 1, 2, \cdots, N. \tag{4}$$

Since it is not possible to solve these equations for the limit $N \to \infty$, Löwdin[6] has extended the summation up to the 12th neighbors starting from atom h and neglects higher-order terms in overlaps without estimating the possible error associated with this truncation.

III. CLUSTER EXPANSION AND THE RECURRENCE RELATION FOR THE INVERSE

The inverse S^{-1} introduced in (3) is formally written as adj S, the adjugate of the overlap matrix S, divided by the determinant $|S|$ of S:

$$S^{-1} = \text{adj } S/|S|. \tag{5}$$

Since adj S is the $N \times N$ matrix whose hl element is given by the cofactor $|S[l\,|h]|$ of $S_{lh} = \langle l\,|h\rangle$ in $|S|$, we find that

$$(S^{-1})_{hl} = |S[l\,|h]|/|S|. \tag{6}$$

The cofactor $|S[l\,|h]|$ is the determinant of the $(N-1) \times (N-1)$ matrix $S[l\,|h]$ obtained from S by suppressing the lth row and the hth column and by attaching the proper sign $(-1)^{l-h}$.

We want to calculate the quotient on the right of (6) by expanding the determinants successively in terms of cofactors, and hence it is convenient to introduce smaller submatrices $S[l_1 l_2 \,|\, h_1 h_2]$, $S[l_1 l_2 \cdots \,|\, h_1 h_2 \cdots]$, etc. obtained from $S[l_1 \,|\, h_1]$ by deleting its rows and columns one by one. For instance, the $(N-r) \times (N-r)$ matrix

$$S[l_1 \cdots l_{r-1} l_r \,|\, h_1 \cdots h_{r-1} h_r]$$

is generated from the $(N-r+1) \times (N-r+1)$ matrix $S[l_1 \cdots l_{r-1} \,|\, h_1 \cdots h_{r-1}]$ by deleting the "l_rth" row and the "h_rth" column[9] and by inserting a proper sign so that $|S[l_1 \cdots l_r \,|\, h_1 \cdots h_r]|$ is the cofactor of $S_{l_r h_r}$ in $|S[l_1 \cdots l_{r-1} \,|\, h_1 \cdots h_{r-1}]|$.

The expansion of a determinant $|S|$ in terms of its cofactors $|S[k|h]|$

$$|S| = \sum_k S_{kh}\, |S[k|h]| = |S[h|h]| + \sum_{k \neq h}{}' \langle k|h\rangle\, |S[k|h]| \tag{7}$$

can be extended to the determinant $|S[l\,|h]|$, yielding the result

$$|S[l\,|h]| = \langle h\,|l\rangle\, |S[lh|hl]| + \sum_{\substack{k \neq h \\ \neq l}}{}'' \langle k\,|l\rangle\, |S[lk|hl]|, \tag{8}$$

where $|S[lk|hl]|$ is the cofactor of S_{kl} in $|S[l|h]|$. Let us now introduce the notation:

$$F[l|h] \equiv (S^{-1})_{hl} \equiv |S[l\,|h]|/|S| \tag{9}$$

and generalize it to quotients of smaller determinants as follows:

$$F_{h_1 \cdots h_r}[l\,|h_{r+1}] \equiv |S[h_1 \cdots h_r l \,|\, h_1 \cdots h_r h_{r+1}]|/|S[h_1 \cdots h_r \,|\, h_1 \cdots h_r]|, \tag{10}$$

where the indices $h_1 \cdots h_r$ indicate that the h_1th, $\cdots$, h_rth rows and columns are suppressed in both determinants. By multiplying (7) and (8) by $|S|^{-1}$ and by using the above notation, we find that

$$F[h|h] = 1 - \sum_{k \neq h}{}' \langle k\,|h\rangle F[k|h], \tag{11}$$

$$-F[l|h] = \langle h\,|l\rangle F_l[h|h] F[l\,|l] + \sum_{\substack{k \neq h \\ \neq l}}{}'' \langle k\,|l\rangle F_l[k|h] F[l|l]. \tag{12}$$

[9] The l_rth row and the h_rth column of $S[l_1 \cdots l_{r-1} \,|\, h_1 \cdots h_{r-1}$ mean, respectively, those corresponding to the l_rth row and the h_rth column of the original matrix S.

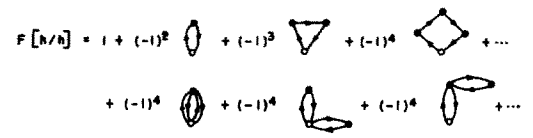


FIG. 1. The expansion of $F[h|h]$. An open dot indicates atom h.

The second term on the right of (12), for instance, is obtained as follows.

$$|S[lk|hl]|/|S| = -\{|S[kl|hl]|/|S[l\,|l]|\} \times \{|S[l|l]|/|S|\}$$
$$= -F_i[k|h]F[l|l]$$

since, by definition of cofactors,

$$|S[lk|hl]| = -|S[kl|hl]|.$$

The function $F_{i_1\cdots i_r}[l|h]$ will also satisfy identities of the same forms as (11) and (12) except that the summations do not include atoms $l_1, \cdots, l_r$. This shows that the relations (11) and (12) are recurrence relations for the expansion of $F[l\,|h]$ in powers of the overlap integrals; the expansion described here is, in fact, a generalization of Cauchy's expansion of a determinant.[7]

The manipulation is simplified by the use of diagrams. For example, we list series of diagrams appearing in the expansions of $F[h|h]$ and $F[l\,|h]$ in Figs. 1 and 2, respectively. Here a straight line starting from l and ending at h represents the overlap integral $\langle h|l\rangle$; and therefore the value of a diagram will decay exponentially when the length of the line increases. A solid dot indicates that the sum over all sites in the lattice should be taken under the provision that dots in a diagram are all distinct. An open dot is a fixed point h or l.

All diagrams necessary are constructed according to the following rules.

(a) The first path starts from l in the case of $F[l\,|h]$ and h in the case of $F[h|h]$ and ends at h in both cases. A path never goes through a point more than once.

(b) Terms in higher orders are obtained by adding more rings. A ring starts from a point on paths drawn previously in the diagram, and always comes back to the starting point.

(c) After completing a diagram, attach sign $(-1)^{t+s}$ to the diagram, where t is the number of rings in the diagram and s the number of electron permutations required in drawing the diagram.

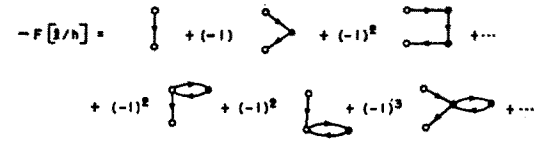


FIG. 2. The expansion of $F[l|h]$. The lines start from atom h and end at atom l.

FIG. 3. The recurrence relation for generation of diagrams.

The factor $(-1)^t$ comes from the minus sign of the second term on the right of (11) and $(-1)^s$ appears because of the minus sign on the left of (12). The number s is obtained as follows. Count the number of dots involved in each path excluding those at the origin and the end of the path. The sum of the numbers thus obtained in a diagram is s. The method (b) implies that the diagrams obtained are all connected.

Unless overlap integrals are all small, however, the convergence of the series is so slow that the direct summation of the series becomes impractical. Instead, use of the recurrence relation shown in Fig. 3 is more convenient, since the relation generates all diagrams for $F[l\,|h]$ except those which have rings starting from point h. For instance, the second, third, and fourth diagrams in Fig. 2 can be generated by inserting the expression on the right of the recurrence relation into the second term of the same expression. By multiplying diagrams thus obtained by those for $F[h|h]$, we obtain all diagrams for $F[l\,|h]$. Hence, the quantity represented by a double line on the left of the recurrence relation is $-F[l\,|h]/F[h|h]$. We denote it by $f(h \leftarrow l)$. The recurrence relation is then written as

$$f(h \leftarrow l) = \langle h|l\rangle - \sum_k{}'' f(h \leftarrow k)\langle k|l\rangle, \quad h \neq l, \quad (13)$$

where

$$f(h \leftarrow l) = -F[l\,|h]/F[h|h] = -|S[l|h]|/|S[h|h]|. \quad (14)$$

The relations (13) and (14) are nothing but a transformation of (3) and (4). However, it has become clear that $f(h \leftarrow l)$ is a well-behaved function given by the sum of connected diagrams which links two points l and h and hence the value will be bounded and decrease when the distance between l and h increases as long as the inverse exists.[10]

Since the overlap integral $\langle k|l\rangle$ decreases exponentially with increasing distance R_{k-l} between k and l, the summation on the right of (13) converges and we need to consider only k's within the range $|R_{k-l}| < |R|$, where the overlap integral $S_R = \int \varphi^*(r)\varphi(r - R)\,d\tau$ practically vanishes. This implies that the functions $f(h \leftarrow k)$ needed in the summation are in the ranges

$$||R| - |R_{h-l}|| < |R_{h-k}| < |R| + |R_{h-l}|| \quad (15)$$

because of the vector relation $R_{h-k} = R_{h-l} + R_{l-k}$.

[10] Even if the expansions of the determinants $|F[l|h]|$ described in Figs. 1 and 2 diverge, the values of the determinants will be finite and the present method is still applicable to this case because an infinite number of terms appearing in the expansion of $f(h \leftarrow l)$ is included correctly in the calculation. See the discussion in Sec. IV.

To compute $f(h \leftarrow l)$ for small distances $|R_{h-l}|$, therefore, we need to know the value of $f(h \leftarrow k)$ in the range $|R_{h-k}| \lesssim |R|$ since $|R_{h-l}| \ll |R|$. However, to estimate the value of $f(h \leftarrow k)$ for which $|R_{h-k}| \approx |R|$, we have to supply $f(h \leftarrow k')$ in the range $|R_{h-k'}| \lesssim 2|R|$ and so on. Thus the linear equations (13) are related to an infinitely large number of variables $f(h \leftarrow l)$, but the difficulty may be avoided by solving the equations iteratively. In the following section, we describe a method to do this and estimate the error.

In case $\varphi_n(r - R)$ is an S function, the Eq. (13) can be simplified as follows. Let $k_1, k_2, \cdots, k_n$ be the Kth neighbors of atom h, which have the common overlap integral $S_K = \langle h|k_i\rangle$ as well as the common function $f_K = f(h \leftarrow k_i)$. Then the summation on the right of (13) can be split into two; the first one $\sum_k^{K=\text{const}}$ extends over all sites k which are the Kth neighbors of atom h and the second one over all K; that is,

$$f_L = S_L - \sum_K \Delta_{LK} f_K, \tag{16}$$

where

$$\Delta_{LK} = \sum_k^{K=\text{const}} \langle k|l\rangle. \tag{17}$$

IV. A METHOD OF SOLVING THE LINEAR EQUATIONS AND ITS ACCURACY

In this section, we describe a method of solving the set of simultaneous equations (16) and discuss its accuracy. Under the matrix representation, (16) is written as

$$(1 + \Delta)\mathbf{f} = \mathbf{S}, \tag{18}$$

where $\mathbf{f}$ and $\mathbf{S}$ are vectors whose elements f_K and S_K are arranged in order of K while the matrix Δ is composed of Δ_{LK}.

As is discussed in Sec. III, the value of f_K decreases with increasing K. Let us assume, as the first approximation, that $f_K = 0$ for $K > Q$. This corresponds to generation of only a certain type of diagrams; but, within this limitation, an infinite number of all possible diagrams are generated and summed correctly by means of the recurrence relation in Fig. 3. The certain type of diagrams here means all possible diagrams, in which the distance between any pair of atoms that are connected by two lines is less than R_Q, where R_Q is the distance between the Qth neighbors. The infinite set of Eq. (18) is then truncated to a set of Q simultaneous equations[11]

$$(\mathbf{1}^{(Q)} + \Delta_0^{(Q)})\mathbf{f}_0^{(Q)} = \mathbf{S}_0^{(Q)}. \tag{19}$$

[11] Notations such as $\Delta_0^{(Q)}$ are used to represent $Q \times Q$ matrices or Q-dimensional vectors and, by Δ_0, the same matrix as $\Delta_0^{(Q)}$ is given in the whole space by adding zeros.

Although the solution $\mathbf{f}_0^{(Q)}$ is computed without difficulty as long as $Q \lesssim 100$, it is not possible to increase Q indefinitely. In certain cases, however, the difference between (18) and (19) can be included into the calculation correctly by using the series expansion described in Sec. IVA. Even if the expansion method does not work, however, the first few elements of $\mathbf{f}_0^{(Q)}$ obtained by solving (19) become exact anyway, as is discussed in Sec. IVB.

A. Exact Expansion for f_K (An Iteration Method)

Let Δ_1 and $\mathbf{S}_1$ be the remainders of Δ and $\mathbf{S}$; that is, $\Delta_1 = \Delta - \Delta_0$ and $\mathbf{S}_1 = \mathbf{S} - \mathbf{S}_0$, respectively. Use of the algebraic identity

$$1 + \Delta = (1 + \Delta_0)[1 + (1 + \Delta_0)^{-1}\Delta_1]$$

in the original Eq. (18) yields that

$$(1 + \mathbf{T})\mathbf{f} = (1 + \Delta_0)^{-1}\mathbf{S} \equiv \mathbf{f}_s, \tag{20}$$

where the notations

$$\mathbf{T} \equiv (1 + \Delta_0)^{-1}\Delta_1, \tag{21}$$

$$\mathbf{f}_s \equiv \mathbf{f}_0 + \mathbf{S}_1 \tag{22}$$

are used. The exact solution $\mathbf{f}$ is then expanded as

$$\begin{aligned}\mathbf{f} &= (1 + \mathbf{T})^{-1}\mathbf{f}_s \\ &= (1 - \mathbf{T} + \mathbf{T}^2 - \cdots (-1)^{n-1}\mathbf{T}^{n-1} + \mathcal{R}_n)\mathbf{f}_s,\end{aligned} \tag{23}$$

and the remainder is given by

$$\mathcal{R}_n\mathbf{f}_s = (-1)^n(1 + \mathbf{T})^{-1}\mathbf{T}^n\mathbf{f}_s. \tag{24}$$

The expansion (23) is equivalent to solving (18) self-consistently by an iteration method. Inserting (19) and (22) into (18), we find that the difference $\Delta\mathbf{f}$ between $\mathbf{f}$ and $\mathbf{f}_s$ is given by

$$\Delta\mathbf{f} = \mathbf{f} - \mathbf{f}_s = -\mathbf{T}\mathbf{f}. \tag{25}$$

If the correction term $-\mathbf{T}\mathbf{f}$ is calculated by using $\mathbf{f}_0$, the second term in the expansion (23) is obtained. Use of $\mathbf{f}_s + \Delta\mathbf{f}$ thus obtained in $-\mathbf{T}\mathbf{f}$ yields the third term and so on. This process, of course, recovers the type of diagrams which are omitted in solving (19).

If the quantity

$$\lambda_{\max} = \max \text{ of } \sum_L |T_{KL}| \tag{26}$$

is less than one, the elements of vector $\mathcal{R}_n\mathbf{f}_s$ are bounded as

$$(\mathcal{R}_n\mathbf{f}_s)_{\max} \leq (1 - \lambda_{\max})^{-1}\lambda_{\max}^n f_{\max}, \tag{27}$$

where

$$f_{\max} = \max \text{ of } |f_K|. \tag{28}$$

Hence, the remainder $(\mathcal{R}_n\mathbf{f}_s)_{\max}$ vanishes for large n,

justifying the expansion (or iteration) procedure described above.

The bound for $\lambda_{\max}$ is given by

$$\lambda_{\max} \leq \text{max of} \sum_L |[(1+\Delta_0)^{-1}]_{KL}| \times \text{max of} \sum_L |(\Delta_1)_{KL}|. \quad (29)$$

Since $(\Delta_1)_{KL}$ defined by (17) is a sum of a finite number of overlap integrals connecting K and L and decays exponentially as the distance K and L increases, the sum $\sum_L |(\Delta_1)_{KL}|$ is always finite. The quantity $\sum_L |[(1+\Delta_0)^{-1}]_{KL}|$ involves a finite number of terms and the upper bound $\mu_{\max}$ is found to be of the order of one in our examples in Sec. V. Therefore, the condition for $(\mathfrak{R}_n \mathbf{f}_s)_{\max} \to 0$ is fulfilled if

$$\text{max of} \sum_L |(\Delta_1)_{KL}| < \frac{1}{\mu_{\max}}, \quad (30)$$

where

$$\mu_{\max} = \text{max of} \sum_L |[(1+\Delta_0)^{-1}]_{KL}|. \quad (31)$$

B. Asymptotic Expansion for f_K

If overlap integrals are large and the condition (30) is not satisfied, the bound $(\mathfrak{R}_n \mathbf{f}_s)_{\max}$ for the remainder thus computed will increase indefinitely for large n. In the following, it is shown that the method can still be used as an asymptotic expansion of f_K.

From the definitions of $\mathbf{T}$ and Δ_1, it is found that the matrix $\mathbf{T}$ has the structure as shown on the left of Fig. 4 and the first $Q \times Q$ elements, T_{KL} for K, $L \leq Q$, vanish exactly. Let us now redivide $\mathbf{T}$ as on the right of Fig. 4 and compute the quantity defined by

$$\alpha_K \begin{cases} \equiv \sum_L |T_{KL}|, & \text{for } K = 1, 2, \cdots, Q' < Q, \\ \equiv \sum_{L=1}^{Q'} |T_{KL}|, & \text{for } K \geq Q'+1. \end{cases} \quad (32)$$

As $(1+\Delta_0)^{-1}$ has only a finite number of nonvanishing off-diagonal elements and is nearly equal to a unit matrix, the quantities T_{KL} may be estimated by $\mu_{\max}(\Delta_1)_{KL}$ and considered as proportional to the overlap integral connecting K and L. Since nonvanishing elements T_{KL} appearing in the summations in (32) are those having large K–L distances only, α_K's are all small and, by choosing a suitably large Q, the bound ϵ for α_K can be made arbitrarily small;

$$\text{max of } \alpha_K \leq \epsilon \ll 1, \quad (33)$$

because the K–L distance for the first nonvanishing element T_{KL} involved in α_K increases when Q increases.

Use of (32) and (33) in successive multiplications by $\mathbf{T}$ of the vector $\mathbf{f}_s$ leads to the following bounds for vector $\mathbf{T}^n\mathbf{f}_s$:

$$\text{max of the first } Q' \text{ elements of } \mathbf{T}^n\mathbf{f}_s \leq \epsilon^2 \lambda_{\max}^{n-2} f_{\max} + \epsilon \lambda_{\max}^{n-1} f_{\max}^{(-Q')}, \quad (34)$$

$$\text{max of the other elements of } \mathbf{T}^n\mathbf{f}_s \leq \epsilon \lambda_{\max}^{n-1} f_{\max} + \lambda_{\max}^{n} f_{\max}^{(-Q')}, \quad (35)$$

where $f_{\max}^{(-Q')}$ is the maximum of elements f_K for $K > Q'$ and decreases exponentially by increasing Q' since f_K decreases exponentially for increasing K as described in Sec. III. According to the discussion in the Appendix, off-diagonal elements of $(1+\mathbf{T})^{-1}$ are small and the remainder $\mathfrak{R}_n \mathbf{f}_s$ itself is bounded similar to (34) and (35). This would justify the truncation proposed in (19) and, by choosing sufficiently large Q and Q', errors in the first Q' elements f_K become negligibly small without applying the iteration procedure.

The inequality (34) also shows that the iteration (expansion) procedure would increase the accuracy asymptotically—at least up to the second time in iterations. In practice, continuation of the expansion procedure would improve the accuracy until the first Q' elements of vector $\mathbf{T}^n\mathbf{f}_s$ begin to increase.

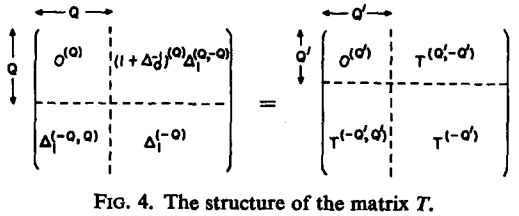


FIG. 4. The structure of the matrix T.

V. SUMMARY OF THE METHOD

The method developed in this paper can be summarized as follows.

(a) The method is formally equivalent to solving the set of linear equations (4), which defines the inverse overlap matrix.

(b) The truncation used in reducing the number of equations to Q corresponds to construction of only a certain type of diagrams; but, within this limitation, an infinite number of them are generated and summed correctly. In this sense, the method covers Löwdin's expansion method; but, even for the cases where Löwdin's expansion diverges our method gives the correct results because of the infinite summation.

(c) The error due to the truncation can be made arbitrarily small as is discussed in Sec. IVB.

TABLE I. The inverse overlap matrix for the body-centered cubic lattice of hydrogen.

K	R_K	S_K	$(S^{-1})_K$					$(S^{-1})_K$[a]	
			$Q=10$	$Q=20$	$Q=30$	$Q=50$	$Q=100$	$Q=18$	$Q=31$
0	000	1.000 00	1.6387 48	1.6393 04	1.6393 17	1.6393 18	1.6393 18	1.6392 83	1.6393 14
1	111	0.2890 11	−0.2737 96	−0.2740 39	−0.2740 43	−0.2740 44	−0.2740 44	−0.2740 37	−0.2740 43
2	200	0.2105 28	−0.0669 03	−0.0673 42	−0.0673 58	−0.0673 58	−0.0673 58	−0.0673 27	−0.0673 59
3	220	0.0716 95	0.0681 38	0.0681 02	0.0681 05	0.0681 05	0.0681 05	0.0681 07	0.0681 05
4	311	0.0361 18	0.0191 16	0.0199 00	0.0199 21	0.0199 22	0.0199 22	0.0198 82	0.0199 22
5	222	0.0291 95	0.0303 56	0.0316 22	0.0316 34	0.0316 37	0.0316 37	0.0316 19	0.0316 38
6	400	0.0132 25	−0.0018 01	−0.0016 62	−0.0015 96	−0.0015 96	−0.0015 96	−0.0016 03	−0.0015 95
7	331	0.0076 73	−0.0118 86	−0.0126 93	−0.0127 02	−0.0127 05	−0.0127 05	−0.0126 68	−0.0127 06
8	420	0.0064 49	−0.0072 06	−0.0080 48	−0.0081 00	−0.0081 02	−0.0081 03	−0.0080 94	−0.0081 03
9	422	0.0033 23		−0.0029 45	−0.0029 78	−0.0029 81	−0.0029 81	−0.0029 12	−0.0029 83
10	333	0.0020 80	−0.0025 29	−0.0026 89	−0.0027 08	−0.0027 14	−0.0027 14	−0.0027 79	−0.0027 15
11	511	0.0020 80	0.0002 13	0.0001 56	0.0000 44	0.0000 43	0.0000 43	0.0001 33	0.0000 41
12	440	0.0009 97		0.0034 88	0.0034 82	0.0034 97	0.0034 98	0.0033 87	0.0034 99
13	531	0.0006 56		0.0017 42	0.0018 86	0.0018 91	0.0018 91	0.0018 84	0.0018 94
14	442	0.0005 73		0.0020 45	0.0020 11	0.0020 22	0.0020 23	0.0019 58	0.0020 24
15	600	0.0005 73		0.0002 39	0.0001 86	0.0001 92	0.0001 92	0.0001 06	0.0001 91
16	620	0.0003 37			0.0003 85	0.0003 91	0.0003 92	0.0002 15	0.0003 93
17	533	0.0002 30		0.0001 01	0.0002 58	0.0002 59	0.0002 60	0.0001 36	0.0002 66
18	622	0.0002 03			−0.0000 78	−0.0000 76	−0.0000 76	−0.0001 57	−0.0000 74
19	444	0.0001 25		0.0001 64	0.0000 88	0.0001 00	0.0001 00		0.0000 96
20	551	0.0000 88		−0.0010 17	−0.0007 22	−0.0007 88	−0.0007 89		−0.0007 87
21	771	0.0000 88			−0.0000 79	−0.0000 93	−0.0000 94		−0.0000 90
22	640	0.0000 78			−0.0006 09	−0.0006 06	−0.0006 10		−0.0006 16
23	642	0.0000 49			−0.0003 45	−0.0003 41	−0.0003 43		−0.0003 50
24	553	0.0000 35		−0.0003 19	−0.0002 36	−0.0002 70	−0.0002 70		−0.0002 73
25	731	0.0000 35			−0.0001 06	−0.0001 24	−0.0001 24		−0.0001 26
26	800	0.0000 21				0.0000 14	0.0000 17		0.0000 02
27	733	0.0000 15			0.0000 48	0.0000 42	0.0000 42		0.0000 44
28	644	0.0000 14			0.0000 03	0.0000 08	0.0000 07		0.0000 00
29	820	0.0000'14				0.0000 07	0.0000 07		0.0000 01
30	660	0.0000 09				0.0002 33	0.0002 34		0.0002 13
31	822	0.0000 09				0.0000 34	0.0000 34		0.0000 31
32	555	0.0000 07		0.0000 60	0.0000 31	0.0000 20	0.0000 21		
33	751	0.0000 07			0.0002 02	0.0001 38	0.0001 51		
34	662	0.0000 06				0.0001 52	0.0001 52		

[a] The results obtained by Calais and Appel (see Ref. 5).

(d) For certain cases, there is an expansion method which leads to an exact solution as is described in Sec. IVA.

Finally, we note that the present method is a special version of the cluster expansion method in the Heitler–London approach,[8] and is generated from the general theory when it is applied to the ferromagnetic ground state. Since in the case of the ferromagnetic ground state the wavefunction given by (2.1) of I is a single Slater determinant, the quantity $\mathbf{S}[k|h]$ defined by (2.18) together with (2.12) of I[12] becomes equal to $|S[k|h]|$, and $\mathbf{F}[k|h]$ introduced by (3.1) of I is the same as $F[k|h]$. In fact, Eqs. (11) and (12) for $F[h|h]$ and $F[k|h]$ are equivalent to (2.11) and (2.12) of II, showing that the inverse overlap matrix S^{-1} is given by the $\mathbf{F}[k|h]$ of the general theory applied to the ferromagnetic ground state.

VI. APPLICATION TO THE HYDROGEN LATTICE

The present method was applied to the hydrogen lattice and the programming was carried out on the CDC 3600 for simple, body-centered and face-centered cubic lattices with overlap integrals between $1s$ orbitals given by

$$S_K = (1 + R_K + \tfrac{1}{3}R_K^2)e^{-R_K}. \tag{36}$$

[12] Equation (2.18) of I should read

$$\mathbf{S}[k_1 \cdots k_n \,|\, h_1 \cdots h_n] \equiv \int \Psi[k_1(h_1) \cdots k_n(h_n)]^\dagger \times \Psi[h_1(h_1) \cdots h_n(h_n)]\, d\tau.$$

The inverse $(S^{-1})_K$ is obtained from (11) and (14) and written as

$$(S^{-1})_0 = F[h|h] = \left(1 - \sum_{K'} n_{K'} S_{K'} f_{K'}\right)^{-1} \quad (37)$$

and

$$(S^{-1})_K = F[k \mid h]_{K=h-k} = -f_K \left(1 - \sum_{K'} n_{K'} S_{K'} f_{K'}\right)^{-1}, \quad (38)$$

where n_K is the number of Kth neighbors.

The results for a body-centered cubic lattice are summarized in Table I, where the lattice constant ($= R_2$) is assumed to be 3.8336 a.u. in order to compare them with those of Calais and Appel. The second column of Table I shows the components of the vector R_K arranged in the order of increasing distances, the third column the corresponding overlap integrals. The next five columns give the values of the inverse $(S^{-1})_K$ calculated for various values of Q; the Q being the dimension of the linear equations (19). Those numbers are compared with the results obtained by Calais and Appel quoted in the last two columns of Table I, where Q indicates that the calculation includes all contributions up to the Qth overlap integral. This is not the case in our method as is clear from the description in Sec. IV.

In order to save space, we list in this table the values of $(S^{-1})_K$ only up to $K = 34$ and omit others; the elements omitted are all smaller than 10^{-4} and insignificant. Higher-order terms are automatically generated by the machine and included in the calculation, but not necessarily in the order of increasing distances. This is why, for example, the term $(S^{-1})_{11}$ is included while $(S^{-1})_9$ is missing when $Q = 10$.

From Table I, it is evident that, in this example, both methods have virtually attained convergence at $Q \approx 30$. We have added two more terms, $K = 30$ and 34, to our third case where $Q = 30$ and this practically eliminates the small difference which existed between the cases $Q = 30$ and $Q = 100$.

In conclusion, we have shown that the simple method of calculating the inverse overlap matrix, given by (4), is based on sound mathematics and is not just an arbitrary approximation. From the practical point of view, the method converges reasonably quickly.

ACKNOWLEDGMENT

The authors would like to thank T. L. Gilbert for valuable discussions on this subject.

APPENDIX. CALCULATION OF $\mathcal{R}_n\mathbf{f}_s$ FOR $\lambda_{\max} > 1$

In this appendix we show that the remainder $\mathcal{R}_n\mathbf{f}_s$ is bounded similar to (34) and (35). For this purpose, we prove the following theorems.

Theorem 1: The length L of vector $\mathbf{Tf}$ is bounded as

$$L \leq \lambda_{\max}^2 L_0, \quad (A1)$$

where $\lambda_{\max}$ is the maximum of eigenvalues of the Hermitian matrix $\mathbf{T}$ and L_0 is the length of vector $\mathbf{f}$.

As $\mathbf{T}$ can be brought into a diagonal form $\boldsymbol{\lambda}$ by a similarity transformation such that $\boldsymbol{\lambda} = \mathbf{UTU}^\dagger$, the length L of $\mathbf{Tf}$ is calculated as

$$L = (\mathbf{Tf})^\dagger(\mathbf{Tf}) = (\mathbf{f}^\dagger\mathbf{U}^\dagger)(\mathbf{UT}^\dagger\mathbf{U}^\dagger)(\mathbf{UTU}^\dagger)(\mathbf{Uf}) = (\mathbf{Uf})^\dagger\boldsymbol{\lambda}^\dagger\boldsymbol{\lambda}(\mathbf{Uf}) = \sum_K |(\mathbf{Uf})_K|^2 |\lambda_K|^2. \quad (A2)$$

Since the length L_0 of f is invariant under a unitary transformation and equal to $\sum_K |(\mathbf{Uf})_K|^2$, the inequality (A1) follows from (A2).

Let $\mathbf{M}_1$ and $\mathbf{M}_2$ be Hermitian matrices and denote by $\nu_{1\max}$ and $\nu_{2\max}$ the maxima of eigenvalues of $\mathbf{M}_1$ and $\mathbf{M}_2$, respectively. Then we obtain

Theorem 2: If $\nu_{1\max}\nu_{2\max} < 1$, the matrix

$$(\mathbf{1} + \mathbf{M}_1\mathbf{M}_2)^{-1}$$

can be expanded in powers of $\mathbf{M}_1\mathbf{M}_2$.

Repeated use of Theorem 1 yields the following bound for the length $L^{(n)}$ of vector $(\mathbf{M}_1\mathbf{M}_2)^n\mathbf{f}$;

$$L^{(n)} \leq (\nu_{1\max}\nu_{2\max})^{2n} L_0, \quad (A3)$$

which leads to vanishingly small $L^{(n)}$ for $n \to \infty$. By expanding $(\mathbf{1} + \mathbf{M}_1\mathbf{M}_2)^{-1}$ in powers of $\mathbf{M}_1\mathbf{M}_2$ and by using (A3) in each term in the expansion series, we find the inequality

$$L \leq (1 - \nu_{1\max}\nu_{2\max})^{-1} L_0, \quad (A4)$$

showing the convergence of the expansion.

Theorem 3: (the theorem of Levy–Hadamard[13]) The maximum $\nu_{\max}$ of eigenvalues of a Hermitian matrix $\mathbf{M}$ is bounded as

$$\nu_{\max} \leq \max \text{ of } \sum_L |M_{KL}|. \quad (A5)$$

Let us introduce a Hermitian matrix $\mathbf{T}'$ defined by

$$\mathbf{T}' = (\mathbf{1} + \boldsymbol{\Delta}_0)^{-\frac{1}{2}} \boldsymbol{\Delta}_1 (\mathbf{1} + \boldsymbol{\Delta}_0)^{-\frac{1}{2}}. \quad (A6)$$

Since $\mathbf{T}'$ is found to have a structure similar to $\mathbf{T}$ shown in Fig. 4, $\mathbf{1} + \mathbf{T}'$ can be divided further into

[13] See, for instance, E. Bodewig, *Matrix Calculus* (North-Holland Publishing Company, Amsterdam, 1959), p. 67.

two parts X and Y as follows.[14]

$$\mathbf{X} \equiv \begin{bmatrix} \mathbf{1}^{(Q')} & \mathbf{0}^{(Q',-Q')} \\ \mathbf{0}^{(-Q',Q')} & \mathbf{1}^{(-Q')} + \mathbf{T}^{(-Q')} \end{bmatrix}, \tag{A7}$$

$$\mathbf{Y} \equiv \begin{bmatrix} \mathbf{0}^{(Q')} & \mathbf{T}^{(Q',-Q')} \\ \mathbf{T}^{(-Q',Q')} & \mathbf{0}^{(-Q')} \end{bmatrix}. \tag{A8}$$

The original matrix $(\mathbf{1} + \mathbf{T})^{-1}$ is then written in terms of the Hermitian matrices X and Y as

$$(\mathbf{1} + \mathbf{T})^{-1} = (\mathbf{1} + \mathbf{\Delta}_0)^{-\frac{1}{2}}(\mathbf{1} + \mathbf{T}')^{-1}(\mathbf{1} + \mathbf{\Delta}_0)^{\frac{1}{2}}, \tag{A9}$$

where

$$(\mathbf{1} + \mathbf{T}')^{-1} = (\mathbf{X} + \mathbf{Y})^{-1} = (\mathbf{1} + \mathbf{X}^{-1}\mathbf{Y})^{-1}\mathbf{X}^{-1}. \tag{A10}$$

It is easily shown that elements of $\mathbf{T}'$ are bounded as (32) and (33) and hence

$$\text{max of } \sum_L |Y_{KL}| \leq \epsilon \ll 1, \tag{A11}$$

while matrix X is nearly equal to $\mathbf{T}'$. By Theorem 3, therefore, eigenvalues of Hermitian matrices Y and $\mathbf{X}^{-1}$ are bounded as

$$\text{max of eigenvalues of } \mathbf{Y} \leq \epsilon \ll 1, \tag{A12}$$

$$\text{max of eigenvalues of } \mathbf{X}^{-1} \leq 1/(1 - |\lambda_{\min}|), \tag{A13}$$

where $\lambda_{\min}$ is the minimum of eigenvalues of T'.

[14] By $\mathbf{T}^{(Q',-Q')}$, we denote the $Q' \times (N - Q')$ submatrix of T consisting of the first Q' rows and the last $N - Q'$ columns; and, by $\mathbf{T}^{(-Q')}$, the $(N - Q') \times (N - Q')$ submatrix consisting of the last $N - Q'$ rows and $N - Q'$ columns.

According to Theorem 2, the matrix $(\mathbf{1} + \mathbf{X}^{-1}\mathbf{Y})^{-1}$ can be expanded in powers of $\mathbf{X}^{-1}\mathbf{Y}$, if

$$\epsilon/(1 - |\lambda_{\min}|) < 1. \tag{A14}$$

As long as the basic functions are linearly independent, eigenvalues $1 + \lambda$ of overlap matrix $1 + \Delta$ are always positive and hence $1 - |\lambda_{\min}| \geq \delta > 0$, where δ is a given number for a given overlap matrix. Since ϵ can be made arbitrarily small by choosing a large Q, the inequality (A14) should be always satisfied, justifying the expansion of $(\mathbf{1} + \mathbf{X}^{-1}\mathbf{Y})^{-1}$. If $\epsilon/(1 - |\lambda_{\min}|) \ll 1$, the order of magnitude of $(\mathbf{1} + \mathbf{X}^{-1}\mathbf{Y})^{-1}$ can be estimated by $\mathbf{1} - \mathbf{X}^{-1}\mathbf{Y}$ and hence

$$(\mathbf{1} + \mathbf{T})^{-1} \approx (\mathbf{1} - \mathbf{X}^{-1}\mathbf{Y})\mathbf{X}^{-1} \tag{A15}$$

because $(\mathbf{1} + \mathbf{\Delta}_1)^{-1}$ has a finite number of small off-diagonal elements and nearly equal to a unit matrix. Use of (A12) $\sim$ (A15) as well as (34) and (35) in $(\mathbf{1} + \mathbf{T})^{-1}\mathbf{T}^n\mathbf{f}_s$ leads to

max of the first Q' elements of $\mathcal{R}_n\mathbf{f}_s$

$$\leq \frac{\epsilon^2\lambda_{\max}^{n-1}}{(1 - |\lambda_{\min}|)^2} f_{\max} + \frac{\epsilon\lambda_{\max}^{n}}{(1 - |\lambda_{\min}|)^2} f_{\max}^{(-Q)}. \tag{A16}$$

This proves that the remainder $\mathcal{R}_n\mathbf{f}_s$ is bounded similar to (34) and (35).

Simple Force Multipoles in the Theory of Deformable Surfaces

M. M. BALABAN
University of California, Berkeley, California

A. E. GREEN
University of Newcastle upon Tyne, England

P. M. NAGHDI
University of California, Berkeley, California

(Received 4 August 1966)

This paper is concerned with a nonlinear theory of simple force multipoles for a deformable surface, embedded in a Euclidean 3-space; the surface is not necessarily elastic. The theory is developed with the use of basic thermodynamical principles, together with invariance conditions under superposed rigid body motions. For simplicity, the basic kinematical ingredients are restricted to be the (ordinary) monopolar velocity of the surface and suitable first- and second-order gradients of the velocity. The theory of an elastic surface and other special cases of the general theory which bear on the foundations of the classical theory of shells are also discussed.

1. INTRODUCTION

THIS paper is concerned with a nonlinear theory of simple force multipoles for a deformable surface, embedded in a Euclidean 3-space, and may be regarded as a counterpart of a general theory of simple force and stress multipoles for three-dimensional bodies given recently by Green and Rivlin.[1] References to related and other special developments in the theory of simple force multipoles and theories of elasticity with couple-stress may be found in the paper by Green and Rivlin[1] and in the monograph by Truesdell and Toupin.[2] The kinematic variables (in the rate of energy equation) in the work of Green and Rivlin[1] are the (classical) monopolar velocity and the gradients of velocity, up to order n, in a rectangular Cartesian coordinate system. Here, however, we limit ourselves to a theory of deformable surfaces in which the basic kinematic ingredients are the velocity of the surface, together with *suitable* first- and second-order gradients of the velocity. The further inclusion of higher-order velocity gradients is possible and may be carried out in the same manner, although at the expense of additional complications.

We find it particularly convenient in this paper to employ an invariant vector notation in the development of the basic theory; the notation used is essentially similar to that in Green and Zerna[3] or Naghdi.[4] While each symbol is defined when first introduced, a few helpful remarks about the notation are made at the end of Sec. 2, where we have collected some basic formulas concerning the geometry of a surface embedded in a Euclidean 3-space.

In Sec. 3, we consider the kinematics of a surface corresponding to the velocity of the surface (a three-dimensional vector field), and suitable first- and second-order gradients (with respect to surface coordinates) of velocity which are invariant under coordinate transformations of the surface. Related and associated kinematical results when motions of the surface differ from those of a given motion only by superposed rigid body motions are discussed in Sec. 4. The measures of deformation for the surface emerging from these kinematical considerations are equivalent to the first and the second fundamental forms of the surface (in both deformed and undeformed configurations), as well as their gradients, all of which are surface tensors.

Next, using a principle of balance of energy and a Clausius–Duhem inequality, valid for a surface embedded in a Euclidean 3-space, together with invariance conditions under superposed rigid body motion, we develop in Sec. 5 a theory of simple force multipoles for deformable surfaces. As already noted, the basic kinematic variables of this theory are the velocity and its first- and second-order gradients and, consistent with these kinematical ingredients, only monopolar, dipolar, and tripolar contact and body forces are admitted. The derivation in Sec. 5, carried out in a neat vectorial form, conceals the relative complexity of the results and hence alternative forms of the basic equations of the theory, in terms of tensor components,

[1] A. E. Green and R. S. Rivlin, Arch. Ratl. Mech. Anal. **16**, 325 (1964).
[2] C. Truesdell and R. A. Toupin, in *Encyclopedia of Physics*, S. Flügge, Ed. (Springer-Verlag, Berlin, 1960), Vol. III/1, p. 226.
[3] A. E. Green and W. Zerna, *Theoretical Elasticity* (Clarendon Press, Oxford, England, 1954).
[4] P. M. Naghdi, in *Progress in Solid Mechanics*, I. N. Sneddon and R. Hill, Eds. (North-Holland Publishing Company, Amsterdam, 1963), Vol. 4, p. 1.

are collected in Sec. 6. The theory of Sec. 5 or 6 is exact and entirely consistent with the basic dynamical and thermodynamical principles of continuum mechanics. Moreover, it is valid for nonisothermal deformations and is not necessarily limited to elastic surfaces.

The remainder of the paper deals with special cases of the theory. Nonlinear constitutive relations in terms of the Helmholtz free energy for an elastic surface are derived in Sec. 7 and special cases, including a determinate theory of an elastic surface with simple force dipoles, are discussed in Sec. 8.

Our present study is a contribution to the theory of deformable surfaces by a *direct* approach, in contrast to that which may begin with the three-dimensional equations. For other developments in the theory of deformable surfaces by a direct approach, we mention a linear isothermal theory for an elastic surface by Serbin,[5] an isothermal nonlinear theory of elastic surfaces by Cohen and DeSilva,[6] and a general theory of a Cosserat surface by Green, Naghdi, and Wainwright.[7] The latter contains other references, where the works of E. and F. Cosserat[8] and Ericksen and Truesdell[9] are particularly cited.

While we do not consider here the linearized version of the theory of Secs. 6 and 7, this may be carried out in a manner similar to that in the paper by Green, Naghdi, and Wainwright.[7] There has always existed an interest in constructing theories of elastic plates and shells from the three-dimensional equations of linear elasticity by admitting higher moments of stress (see, e.g., the papers by Tiffen and Lowe[10,11]) and, in this connection, the theory of Secs. 6 and 7 is rather illuminating. Moreover, the nature of the determinate tripolar and dipolar theories discussed in Sec. 8 sheds further light on the foundations of the classical theory of elastic shells.

2. PRELIMINARIES. GEOMETRY OF A SURFACE IN A EUCLIDEAN SPACE

Let z_i, $(i = 1, 2, 3)$, refer to a fixed right-handed rectangular Cartesian coordinate system in a Euclidean 3-space, and let x^i denote an arbitrary (real) curvilinear coordinate system defined by the transformation relations

$$z_i = z_i(x^1, x^2, x^3). \tag{2.1}$$

We assume (2.1) to be nonsingular, so that a unique inverse

$$x^i = x^i(z_1, z_2, z_3) \tag{2.2}$$

exists at each point of space with a neighborhood $\mathcal{R}$, where

$$\det [\partial z_i/\partial x^j] \neq 0. \tag{2.3}$$

Denoting by $\mathbf{p}(x^i)$ the position vector of a generic point of the Euclidean space, then the metric tensor of the coordinate system x^i is given by

$$g_{ij} = \mathbf{p}_{,i} \cdot \mathbf{p}_{,j} = (\partial z_k/\partial x^i)(\partial z_k/\partial x^j), \tag{2.4}$$

where z_k are the Cartesian components of $\mathbf{p}$ and a comma stands for partial differentiation. From (2.4) follows

$$g = \det [g_{ij}] = \left| \frac{\partial z_k}{\partial x^i} \frac{\partial z_k}{\partial x^j} \right| = \left| \frac{\partial z_k}{\partial x^i} \right|^2 > 0, \tag{2.5}$$

in view of (2.3). The conjugate tensor g^{ij}, namely the inverse of (2.4) which exists on account of (2.5), together with g_{ij} satisfy

$$g^{ik} g_{kj} = g_{jk} g^{ki} = \delta^i_j, \tag{2.6}$$

where δ^i_j is the Kronecker symbol. We also define the contravariant and covariant ϵ systems by

$$\epsilon^{ijk} = g^{-\frac{1}{2}} e^{ijk}, \quad \epsilon_{ijk} = g^{\frac{1}{2}} e_{ijk}, \tag{2.7}$$

where e^{ijk} and e_{ijk} are the standard permutation symbols in 3-space.

Consider now a surface $\mathfrak{s}$ embedded in a Euclidean 3-space and let the position vector of a point on this surface be denoted by $\mathbf{r}(x^1, x^2)$ with $\{x^1, x^2\}$ being simply parameters, as yet unrelated to the curvilinear coordinates x^i utilized between (2.1) to (2.7). Let $\mathbf{a}_\alpha$ be the base vectors along the x^α curves on $\mathfrak{s}$ and let $a_{\alpha\beta}$ denote the first fundamental form of the surface. Then

$$\mathbf{a}_\alpha = \mathbf{r}_{,\alpha}, \quad a_{\alpha\beta} = \mathbf{a}_\alpha \cdot \mathbf{a}_\beta, \tag{2.8}$$

where in (2.8) a comma denotes partial differentiation with respect to x^α and $\alpha, \beta = 1, 2$ only. The reciprocal of $(2.8)_1$ and $(2.8)_2$, denoted by $\mathbf{a}^\alpha$ and $a^{\alpha\beta}$ and defined for all points of $\mathfrak{s}$ for which

$$a = \det (a_{\alpha\gamma}) \neq 0, \tag{2.9}$$

together with $\mathbf{a}_\alpha$ and $a_{\alpha\beta}$, satisfy

$$\begin{gathered} a^{\alpha\beta} = \mathbf{a}^\alpha \cdot \mathbf{a}^\beta, \quad \mathbf{a}^\alpha \cdot \mathbf{a}_\beta = \delta^\alpha_\beta, \\ a^{\alpha\lambda} a_{\lambda\beta} = a_{\beta\lambda} a^{\lambda\alpha} = \delta^\alpha_\beta, \end{gathered} \tag{2.10}$$

where δ^α_β is the Kronecker symbol in 2-space.

[5] H. J. Serbin, J. Math. Phys. 4, 838 (1963).

[6] H. Cohen and C. N. DeSilva, J. Math. Phys. 7, 246 (1966). Although Cohen and DeSilva begin by including a director in their work, they soon abandon this. In the context of the present paper, the work of Cohen and DeSilva may be compared to a special case of our theory in Sec. 8, called the restricted theory of simple force dipoles. However, even in this case, there are some differences between their equations and those of our restricted dipolar theory.

[7] A. E. Green, P. M. Naghdi, and W. L. Wainwright, Arch. Ratl. Mech. Anal. 20, 287 (1965).

[8] E. and F. Cosserat, *Théorie des Corps Déformables* (Hermann et Cie, Paris, 1909).

[9] J. L. Ericksen and C. Truesdell, Arch. Ratl. Mech. Anal. 1, 295 (1958).

[10] R. Tiffen and P. G. Lowe, Proc. London Math. Soc. (3) 13, 653 (1963).

[11] R. Tiffen and P. G. Lowe, J. London Math. Soc. 40, 72 (1965).

We introduce the second fundamental form of the surface through

$$b_{\alpha\beta} = \mathbf{a}_3 \cdot \mathbf{a}_{\alpha,\beta} = \mathbf{a}_3 \cdot \mathbf{a}_{\beta,\alpha} = b_{\beta\alpha}, \tag{2.11}$$

where $\mathbf{a}_3$ is the unit normal to any point of $\mathfrak{s}$ and satisfies the relations

$$\mathbf{a}_3 \cdot \mathbf{a}_3 = 1, \quad \mathbf{a}_3 \cdot \mathbf{a}_\alpha = 0, \quad \mathbf{a}_3 \cdot \mathbf{a}_{3,\alpha} = 0. \tag{2.12}$$

We recall, from differential geometry, the formulas

$$\begin{aligned} \mathbf{a}_{\alpha|\beta} &= b_{\alpha\beta}\mathbf{a}_3, \\ \mathbf{a}_{3,\beta} &= -b^{\alpha}_{\beta}\mathbf{a}_\alpha, \\ b_{\alpha\beta|\gamma} &= b_{\alpha\gamma|\beta}, \end{aligned} \tag{2.13}$$

and also note, for future reference, the result

$$T_{\alpha|\beta\gamma} - T_{\alpha|\gamma\beta} = R^{\lambda}_{.\alpha\beta\gamma}T_\lambda, \tag{2.14}$$

where in (2.13) and (2.14) a single stroke designates covariant differentiation with respect to $a_{\alpha\beta}$, T_α is any surface tensor, and $R^{\lambda}_{.\alpha\beta\gamma}$ is the Riemann–Christoffel surface tensor given by

$$\begin{aligned} R^{\lambda}_{.\alpha\beta\gamma} &= -R^{\lambda}_{.\alpha\gamma\beta} \\ &= b_{\alpha\gamma}b^{\lambda}_{\beta} - b_{\alpha\beta}b^{\lambda}_{\gamma}. \end{aligned} \tag{2.15}$$

Consider again the surface $\mathfrak{s}$ and assume the existence of a neighborhood $\mathcal{R}(\mathfrak{s})$ in which points in space lie along one and only one normal to $\mathfrak{s}$. Let x^3, a parameter measured (to the scale of z_i) along the positive direction of the uniquely defined normal from $\mathfrak{s}$, denote the distance to any point in $\mathcal{R}(\mathfrak{s})$. It then follows that any point in $\mathcal{R}$ can be located by means of the relation

$$\mathbf{p}(x^1, x^2, x^3) = \mathbf{r}(x^1, x^2) + x^3\mathbf{a}_3(x^1, x^2). \tag{2.16}$$

If we now identify the parameters x^1, x^2, and x^3 with the curvilinear coordinates x^i, then $x^i = \{x^\alpha, x^3\}$ may be regarded as a system of convected normal coordinates and (2.16) represents the transformation relations between these normal coordinates and z_i, the Cartesian components of $\mathbf{p}$.

The existence of a neighborhood $\mathcal{R}(\mathfrak{s})$ in which every point is uniquely located by (2.16) may be verified in the manner discussed by Naghdi.[4] Indeed, if R_1 and R_2 are the principal radii of curvature of the surface, in view of (2.9), we need only choose

$$\mathcal{R}(\mathfrak{s}) = \{(x^\alpha, x^3): |x^3| < \min(|R_1|, |R_2|)\} \tag{2.17}$$

to ensure that (2.3) is always satisfied. With this choice of $\mathcal{R}(\mathfrak{s})$, (2.16) is nonsingular and hence the various formulas given earlier in this section remain valid.

We also note here that on the surface $x^3 = 0$, we have

$$\begin{aligned} &g_{\alpha\beta} = a_{\alpha\beta}, \quad g_{\alpha 3} = 0, \quad g_{33} = 1, \\ &g^{\alpha\beta} = a^{\alpha\beta}, \quad g^{\alpha 3} = 0, \quad g^{33} = 1, \\ &g = a. \end{aligned} \tag{2.18}$$

A few remarks concerning the notation and convention used may be helpful at this stage. Throughout the paper Latin indices (subscripts or superscripts) have the range 1, 2, 3, Greek indices have the range 1, 2, and the usual summation convention is employed. We use a single stroke (|) for covariant differentiation with respect to the first fundamental form of the surface, a semicolon (;) for components of the covariant derivative of a vector function as in Eq. (3.6), a comma for partial differentiation with respect to surface coordinates x^α, and a superposed dot for material derivative, i.e., differentiation with respect to time, holding x^α fixed. The lowering and raising of superscripts and subscripts of space tensors defined on $x^3 = 0$ (e.g., the components v_i of the velocity vector in Sec. 3) is accomplished by using the metric tensor g_{ij} and its conjugate defined in (2.18).

3. KINEMATICS OF A DEFORMABLE SURFACE

Let the motion of a surface $\mathfrak{s}$, embedded in a Euclidean 3-space, be referred to a fixed right-handed system of rectangular Cartesian axes. The position vector of a typical particle of $\mathfrak{s}$ at time t may be designated by $\mathbf{r}$, where

$$\mathbf{r} = \mathbf{r}(x^1, x^2, t), \tag{3.1}$$

and x^α together with x^3 represent a system of convected normal coordinates introduced in Sec. 2. Also, we impose the restriction

$$\det[\partial\mathbf{r}/\partial x^\alpha] > 0 \tag{3.2}$$

for physically admissible motions.

The first and the second fundamental forms of the surface, as well as various formulas of Sec. 2 involving $\mathbf{r}$, $\mathbf{a}_3$ and their derivatives, are still valid except that now these functions depend also on t. We designate the initial value of $\mathbf{r}$ at time $t = 0$ by $\mathbf{R}$ and refer to the initial (undeformed) surface by $\mathfrak{S}$ and denote its first and second fundamental forms by $A_{\alpha\beta}$ and $B_{\alpha\beta}$, respectively.

Let $\mathbf{v}$, a three-dimensional vector field, denote the velocity of $\mathfrak{s}$ at time t. Then, when referred to the base vectors $\mathbf{a}_i = \{\mathbf{a}_\alpha, \mathbf{a}_3\}$ of $\mathfrak{s}$, $\mathbf{v}$ may be written as

$$\mathbf{v} = v^i\mathbf{a}_i = v^\alpha\mathbf{a}_\alpha + v^3\mathbf{a}_3 = v_i\mathbf{a}^i. \tag{3.3}$$

Since the coordinate curves on $\mathfrak{s}$ are convected, we have[12]

$$\begin{aligned}\dot{\mathbf{a}}_\alpha &= \mathbf{v}_{|\alpha} = \mathbf{v}_{,\alpha} \\ &= (v_{\beta|\alpha} - b_{\beta\alpha}v_3)\mathbf{a}^\beta + (v_{3,\alpha} + b^\beta_\alpha v_\beta)\mathbf{a}^3 \end{aligned} \tag{3.4}$$

and then, by (2.12), we can show that

$$\dot{\mathbf{a}}_3 = -(v_{3,\alpha} + b^\beta_\alpha v_\beta)\mathbf{a}^\alpha, \tag{3.5}$$

where a superposed dot denotes the material derivative with respect to t, holding x^α fixed. We also introduce here the notation $v_{i;\alpha}$ by writing the right-hand side of $(3.4)_2$ as $v_{i;\alpha}\mathbf{a}^i$ so that

$$\mathbf{v}_{|\alpha} = \mathbf{v}_{,\alpha} = v_{i;\alpha}\mathbf{a}^i \tag{3.6}$$

and

$$\begin{aligned} v_{\lambda;\alpha} &= \mathbf{v}_{|\alpha}\cdot\mathbf{a}_\lambda = (v_{\lambda|\alpha} - b_{\alpha\lambda}v_3), \\ v_{3;\alpha} &= \mathbf{v}_{|\alpha}\cdot\mathbf{a}_3 = (v_{3,\alpha} + b^\lambda_\alpha v_\lambda). \end{aligned} \tag{3.7}$$

We now record without proofs certain kinematical results established previously.[7] Thus

$$\begin{aligned} \dot{\mathbf{a}}_\alpha &= (\eta_{k\alpha} + \psi_{k\alpha})\mathbf{a}^k, \\ \dot{\mathbf{a}}^\alpha &= a^{\alpha\lambda}(\psi_{k\lambda} - \eta_{k\lambda})\mathbf{a}^k, \\ \dot{\mathbf{a}}_3 &= \dot{\mathbf{a}}^3 = \psi_{k3}\mathbf{a}^k, \end{aligned} \tag{3.8}$$

where

$$\begin{aligned} 2\eta_{\alpha\beta} = 2\eta_{\beta\alpha} &= (v_{\alpha|\beta} + v_{\beta|\alpha} - 2b_{\alpha\beta}v_3) \\ &= (v_{\alpha;\beta} + v_{\beta;\alpha}), \end{aligned} \tag{3.9}$$

$$\eta_{3\alpha} = \eta_{\alpha 3} = 0, \quad \eta_{33} = 0,$$

and

$$\begin{aligned} 2\psi_{\alpha\beta} = -2\psi_{\beta\alpha} &= (v_{\alpha|\beta} - v_{\beta|\alpha}) \\ &= (v_{\alpha;\beta} - v_{\beta;\alpha}), \\ \psi_{\alpha 3} = -\psi_{3\alpha} &= -(v_{3,\alpha} + b^\beta_\alpha v_\beta) \\ \psi_{33} &= 0. \end{aligned} \tag{3.10}$$

In view of (3.9) and (3.10), the components of the velocity gradient $\mathbf{v}_{|\alpha}$ given by (3.7) may now be put in the form

$$v_{\lambda;\alpha} = \eta_{\lambda\alpha} + \psi_{\lambda\alpha}, \quad v_{3;\alpha} = \psi_{3\alpha}. \tag{3.11}$$

For future convenience, we define here the kinematic measures

$$e_{\alpha\beta} = \tfrac{1}{2}(a_{\alpha\beta} - A_{\alpha\beta}), \quad \kappa_{\alpha\beta} = -(b_{\alpha\beta} - B_{\alpha\beta}), \tag{3.12}$$

introduce the notation[13]

$$\boldsymbol{\eta}_\alpha = \eta_{k\alpha}\mathbf{a}^k = \tfrac{1}{2}(\mathbf{v}_{|\alpha}\cdot\mathbf{a}_\lambda + \mathbf{v}_{|\lambda}\cdot\mathbf{a}_\alpha)\mathbf{a}^\lambda \tag{3.13}$$

and note that

$$\begin{aligned} \dot{e}_{\alpha\beta} &= \eta_{\alpha\beta} = \tfrac{1}{2}\dot{a}_{\alpha\beta}, \\ \dot{\kappa}_{\alpha\beta} &= -\dot{b}_{\alpha\beta}. \end{aligned} \tag{3.14}$$

[12] See, for example, Ref. 3.

[13] $\boldsymbol{\eta}_\alpha$ as defined by (3.13) is negative of the corresponding quantity in Ref. 7.

Also, the rate of change of the determinant of $a_{\alpha\beta}$ is given by

$$\begin{aligned} \dot{a} &= \dot{\overline{\det(a_{\alpha\beta})}} \\ &= (\partial/\partial a_{\lambda\nu})[\det(a_{\alpha\beta})]\dot{a}_{\lambda\nu} \\ &= aa^{\lambda\nu}\dot{a}_{\lambda\nu} = 2a\eta^\lambda_\lambda. \end{aligned} \tag{3.15}$$

It is clear from (3.9) to (3.12) that $e_{\alpha\beta}$, $\kappa_{\alpha\beta}$, $\eta_{\alpha\beta}$, as well as $\psi_{\alpha\beta}$ (a subtensor of ψ_{ij}), are *surface tensors*, whereas ψ_{ij} is a *space tensor*. The tensors $\eta_{\alpha\beta}$ and $\psi_{\alpha\beta}$ may be referred to as the surface rate of deformation and the surface spin, respectively.

We consider now the higher-order velocity gradients. The second-order velocity gradient $\mathbf{v}_{|\alpha\beta}$, by $(3.4)_1$ and $(2.13)_1$, may be expressed as

$$\begin{aligned} \mathbf{v}_{|\alpha\beta} = (\dot{\mathbf{a}}_\alpha)_{|\beta} &= \dot{\overline{(b_{\alpha\beta}\mathbf{a}_3)}} + \dot{\Gamma}^\lambda_{\alpha\beta}\mathbf{a}_\lambda \\ &= \dot{b}_{\alpha\beta}\mathbf{a}_3 - b_{\alpha\beta}v_{3;\lambda}\mathbf{a}^\lambda + \dot{\Gamma}^\lambda_{\alpha\beta}\mathbf{a}_\lambda, \end{aligned} \tag{3.16}$$

where $\Gamma^\lambda_{\alpha\beta}$ is the surface Christoffel symbol given by

$$\Gamma^\lambda_{\alpha\beta} = \mathbf{a}^\lambda\cdot\mathbf{a}_{\alpha,\beta}. \tag{3.17}$$

It is clear that $\mathbf{v}_{|\alpha\beta}$ is symmetric in the indices (α, β) so that $\mathbf{v}_{|\alpha\beta} = \mathbf{v}_{|(\alpha\beta)}$, where the notation $T_{(\alpha\beta)}$ stands for

$$T_{(\alpha\beta)} = \tfrac{1}{2}[T_{\alpha\beta} + T_{\beta\alpha}]. \tag{3.18}$$

The third-order velocity gradient $\mathbf{v}_{|\alpha\beta\gamma}$ is symmetric only in (α, β). In fact, by a formula of the form (2.14),

$$\mathbf{v}_{|\alpha\beta\gamma} - \mathbf{v}_{|\alpha\gamma\beta} = R^\lambda_{\cdot\alpha\beta\gamma}\mathbf{v}_{|\lambda}, \tag{3.19}$$

and with the help of (3.19) and $(2.15)_2$, we have

$$\mathbf{v}_{|\alpha\beta\gamma} = \mathbf{v}_{|(\alpha\beta\gamma)} + \tfrac{1}{3}[b_{\alpha\gamma}b^\lambda_\beta + b_{\beta\gamma}b^\lambda_\alpha - 2b_{\alpha\beta}b^\lambda_\gamma]\mathbf{v}_{|\lambda}, \tag{3.20}$$

where $\mathbf{v}_{|(\alpha\beta\gamma)}$ is completely symmetric in (α, β, γ).

We can easily obtain various expressions for the components of $\mathbf{v}_{|\alpha\beta}$ and $\mathbf{v}_{|\alpha\beta\gamma}$, namely, $v_{\lambda;\alpha\beta}$, $v_{3;\alpha\beta}$ and $v_{\lambda;\alpha\beta\gamma}$, $v_{3;\alpha\beta\gamma}$ similar to those in (3.7), but we do not record these here.

4. SUPERPOSED RIGID BODY MOTIONS

We consider a second motion of $\mathfrak{s}$ which differs from the previous motion (3.1) only by a superposed rigid body motion when the surface has the same orientation as in the first motion. Let the velocity vector at time t, corresponding to the second motion, be denoted by $\mathbf{v}^*$. Then

$$\begin{aligned} \mathbf{v}^* &= \mathbf{v} + [\mathbf{v}_0 + \boldsymbol{\omega}\times(\mathbf{r} - \mathbf{r}_0)] \\ &= \mathbf{v} + [\mathbf{b} + \boldsymbol{\omega}\times\mathbf{r}], \end{aligned} \tag{4.1}$$

where the quantity in the bracket represents an arbitrary velocity due to rigid motion, and $\mathbf{b} = \mathbf{v}_0 - \boldsymbol{\omega}\times\mathbf{r}_0$, $\mathbf{v}_0$ and $\boldsymbol{\omega}$ are vector functions of time only.

From (4.1), we have

$$\mathbf{v}^*_{|\alpha} = \mathbf{v}^*_{,\alpha} = \mathbf{v}_{,\alpha} + \boldsymbol{\omega} \times \mathbf{a}_\alpha. \tag{4.2}$$

Also, with the help of (2.13), we obtain from (4.2)

$$\mathbf{v}^*_{|\alpha\beta} = \mathbf{v}_{|\alpha\beta} + b_{\alpha\beta}\boldsymbol{\omega} \times \mathbf{a}_3, \tag{4.3}$$

$$\mathbf{v}^*_{|\alpha\beta\gamma} = \mathbf{v}_{|\alpha\beta\gamma} - b_{\alpha\beta}b^\nu_\gamma \boldsymbol{\omega} \times \mathbf{a}_\nu + b_{\alpha\beta|\gamma}\boldsymbol{\omega} \times \mathbf{a}_3. \tag{4.4}$$

We now define the kinematical quantities

$$\begin{aligned}\boldsymbol{\eta}_{\alpha\beta} = \boldsymbol{\eta}_{(\alpha\beta)} &= \mathbf{v}_{|\alpha\beta} + b_{\alpha\beta}(\mathbf{v}_{|\lambda} \cdot \mathbf{a}_3)\mathbf{a}^\lambda \\ &= \mathbf{v}_{|\alpha\beta} + b_{\alpha\beta}v_{3;\lambda}\mathbf{a}^\lambda,\end{aligned} \tag{4.5}$$

$$\begin{aligned}\boldsymbol{\eta}_{\alpha\beta\gamma} &= \boldsymbol{\eta}_{(\alpha\beta\gamma)} \\ &= \mathbf{v}_{|\alpha\beta\gamma} + b_{\alpha\beta}b^\sigma_\gamma \mathbf{v}_{|\sigma} + b_{\alpha\beta|\gamma}(\mathbf{v}_{|\lambda} \cdot \mathbf{a}_3)\mathbf{a}^\lambda,\end{aligned} \tag{4.6}$$

and note that (4.6) is completely symmetric in the indices (α, β, γ) in view of (3.20). Using (4.2)–(4.6) and (3.13), and noting the identity

$$[(\boldsymbol{\omega} \times \mathbf{a}_\lambda) \cdot \mathbf{a}_3]\mathbf{a}^\lambda = -[(\boldsymbol{\omega} \times \mathbf{a}_3) \cdot \mathbf{a}_\lambda]\mathbf{a}^\lambda = -\boldsymbol{\omega} \times \mathbf{a}_3, \tag{4.7}$$

it can be easily verified that

$$\boldsymbol{\eta}^*_\alpha = \boldsymbol{\eta}_\alpha, \quad \boldsymbol{\eta}^*_{\alpha\beta} = \boldsymbol{\eta}_{\alpha\beta}, \quad \boldsymbol{\eta}^*_{\alpha\beta\gamma} = \boldsymbol{\eta}_{\alpha\beta\gamma}. \tag{4.8}$$

Hence the kinematic variables $\boldsymbol{\eta}_\alpha$, $\boldsymbol{\eta}_{\alpha\beta}$, $\boldsymbol{\eta}_{\alpha\beta\gamma}$ are invariant under superposed rigid body motions when the surface has the same orientation as in the first motion.

From (4.5) and (3.16), we have

$$\begin{aligned}\boldsymbol{\eta}_{\alpha\beta} = \boldsymbol{\eta}_{(\alpha\beta)} &= \mathbf{v}_{|\alpha\beta} + b_{\alpha\beta}(\mathbf{v}_{|\lambda} \cdot \mathbf{a}_3)\mathbf{a}^\lambda \\ &= b_{\alpha\beta}\mathbf{a}_3 + \Gamma^\lambda_{\alpha\beta}\mathbf{a}_\lambda = \eta_{i(\alpha\beta)}\mathbf{a}^i\end{aligned} \tag{4.9}$$

and

$$\eta_{\lambda(\alpha\beta)} = a_{\lambda\gamma}\dot{\Gamma}^\gamma_{\alpha\beta}, \quad \eta_{3(\alpha\beta)} = \dot{b}_{\alpha\beta}. \tag{4.10}$$

Recalling that

$$b_{\alpha\beta|\gamma} = b_{\alpha\beta,\gamma} - \Gamma^\sigma_{\alpha\gamma}b_{\sigma\beta} - \Gamma^\sigma_{\beta\gamma}b_{\sigma\alpha}, \tag{4.11}$$

then

$$\begin{aligned}\dot{\overline{(b_{\alpha\beta|\gamma})}} &= (\dot{b}_{\alpha\beta})_{|\gamma} - b_{\sigma\beta}a^{\sigma\lambda}\eta_{\lambda(\alpha\gamma)} - b_{\sigma\alpha}a^{\sigma\lambda}\eta_{\lambda(\gamma\beta)} \\ &= \eta_{3(\alpha\beta)|\gamma} - b^\lambda_\beta\eta_{\lambda(\alpha\gamma)} - b^\lambda_\alpha\eta_{\lambda(\gamma\beta)},\end{aligned} \tag{4.12}$$

$$\begin{aligned}\dot{\overline{(b_{\alpha\beta|\gamma\delta})}} &= \eta_{3(\alpha\beta)|\gamma\delta} - b^\lambda_\beta\eta_{\lambda(\alpha\gamma)|\delta} - b^\lambda_\alpha\eta_{\lambda(\gamma\beta)|\delta} \\ &\quad - [(b^\lambda_{\beta|\gamma}\eta_{\lambda(\alpha\delta)} + b^\lambda_{\alpha|\delta}\eta_{\lambda(\beta\gamma)}) \\ &\quad + (b^\lambda_{\beta|\delta}\eta_{\lambda(\alpha\gamma)} + b^\lambda_{\alpha|\gamma}\eta_{\lambda(\beta\delta)}) + a^{\lambda\nu}b_{\alpha\beta|\lambda}\eta_{\nu(\gamma\delta)}].\end{aligned} \tag{4.13}$$

Now, differentiate (4.9) to obtain

$$\begin{aligned}&\mathbf{v}_{|\alpha\beta\gamma} + b_{\alpha\beta|\gamma}(\mathbf{v}_{|\lambda} \cdot \mathbf{a}_3)\mathbf{a}^\lambda + b_{\alpha\beta}[(\mathbf{v}_{|\lambda\gamma} \cdot \mathbf{a}_3)\mathbf{a}^\lambda \\ &\qquad - \lambda b^\sigma_\gamma(\mathbf{v}_{|\lambda} \cdot \mathbf{a}_\sigma)\mathbf{a}^\lambda + b^\lambda_\gamma(\mathbf{v}_{|\lambda} \cdot \mathbf{a}_3)\mathbf{a}_3] \\ &= (\dot{b}_{\alpha\beta})_{|\gamma}\mathbf{a}_3 - \dot{b}_{\alpha\beta}b^\sigma_\gamma\mathbf{a}_\sigma + (\dot{\Gamma}^\lambda_{\alpha\beta})_{|\gamma}\mathbf{a}_\lambda + \dot{\Gamma}^\lambda_{\alpha\beta}b_{\lambda\gamma}\mathbf{a}_3.\end{aligned} \tag{4.14}$$

From (4.14), together with (4.5), (4.6), and (4.12), we deduce

$$\begin{aligned}&\boldsymbol{\eta}_{(\alpha\beta\gamma)} + b_{\alpha\beta}[\boldsymbol{\eta}_{(\lambda\gamma)} \cdot \mathbf{a}_3]\mathbf{a}^\lambda \\ &\qquad - b_{\alpha\beta}[b^\lambda_\gamma(\mathbf{v}_{|\lambda} \cdot \mathbf{a}_\sigma)\mathbf{a}^\sigma + b^\sigma_\gamma(\mathbf{v}_{|\lambda} \cdot \mathbf{a}_\sigma)\mathbf{a}^\lambda] \\ &\qquad = (b_{\alpha\beta})_{|\gamma}\mathbf{a}_3 - \dot{b}_{\alpha\beta}b^\sigma_\gamma\mathbf{a}_\sigma + (\dot{\Gamma}^\lambda_{\alpha\beta})_{|\gamma}\mathbf{a}_\lambda + \dot{\Gamma}^\lambda_{\alpha\beta}b_{\lambda\gamma}\mathbf{a}_3\end{aligned} \tag{4.15}$$

and

$$\begin{aligned}&\boldsymbol{\eta}_{(\alpha\beta\gamma)} = \eta_{i(\alpha\beta\gamma)}\mathbf{a}^i, \\ &\eta_{\lambda(\alpha\beta\gamma)} = \eta_{\lambda(\alpha\beta)|\gamma} - [b_{\lambda\gamma}\eta_{3(\alpha\beta)} + b_{\alpha\beta}\eta_{3(\lambda\gamma)}] + 2b_{\alpha\beta}b^\sigma_\gamma\eta_{\sigma\lambda}, \\ &\eta_{3(\alpha\beta\gamma)} = \dot{\overline{(b_{\alpha\beta|\gamma})}} + b^\lambda_\beta\eta_{\lambda(\alpha\gamma)} + b^\lambda_\alpha\eta_{\lambda(\gamma\beta)} + b^\lambda_\gamma\eta_{\lambda(\alpha\beta)},\end{aligned} \tag{4.16}$$

where we have also used

$$\dot{\Gamma}^\lambda_{\alpha\beta} = a^{\lambda\nu}\eta_{\nu(\alpha\beta)}, \tag{4.17}$$

which follows from $(4.10)_1$.

We close this section by calculating suitable expressions for the material derivative of gradients (with respect to $a_{\alpha\beta}$) of $A_{\alpha\beta}$ which will be needed subsequently. Recalling that

$$A_{\alpha\beta|\gamma} = A_{\alpha\beta,\gamma} - \Gamma^\lambda_{\alpha\gamma}A_{\lambda\beta} - \Gamma^\lambda_{\beta\gamma}A_{\lambda\alpha}, \tag{4.18}$$

with the help of (4.17) and the fact that $\dot{A}_{\alpha\beta} = 0$, we obtain

$$\dot{\overline{(A_{\alpha\beta|\gamma})}} = -a^{\lambda\nu}[A_{\lambda\beta}\eta_{\nu(\alpha\gamma)} + A_{\lambda\alpha}\eta_{\nu(\beta\gamma)}]. \tag{4.19}$$

In a similar manner, from $(A_{\alpha\beta|\gamma\delta})$, we deduce the result

$$\begin{aligned}\dot{\overline{(A_{\alpha\beta|\gamma\delta})}} &= -a^{\lambda\nu}[A_{\lambda\beta}\eta_{\nu(\alpha\gamma)|\delta} + A_{\lambda\alpha}\eta_{\nu(\beta\gamma)|\delta}] \\ &\quad - a^{\lambda\nu}\{[A_{\lambda\beta|\delta}\eta_{\nu(\alpha\gamma)} + A_{\lambda\beta|\gamma}\eta_{\nu(\alpha\delta)}] \\ &\quad + [A_{\lambda\alpha|\delta}\eta_{\nu(\beta\gamma)} + A_{\lambda\alpha|\gamma}\eta_{\nu(\beta\delta)}] + A_{\alpha\beta|\lambda}\eta_{\nu(\gamma\delta)}\}.\end{aligned} \tag{4.20}$$

5. A THEORY OF A DEFORMABLE SURFACE

Let σ, the area of $\mathfrak{s}$ at time t, be bounded by a closed curve c and let $\boldsymbol{\nu}$ be the outward unit normal to c lying in the surface. If $\mathbf{n}$ is a three-dimensional vector field defined on $\mathfrak{s}$ and if, for all arbitrary velocity fields $\mathbf{v}$, the scalar $\mathbf{n} \cdot \mathbf{v}$ is a rate of work per unit length of c, then $\mathbf{n}$ is called a *contact force* per unit length. If $\mathbf{n}^\alpha$ are contravariant vectors (under transformation of surface coordinates x^α) and if, for all arbitrary velocity gradients $\mathbf{v}_{|\alpha}$, the scalar $\mathbf{n}^\alpha \cdot \mathbf{v}_{|\alpha}$ is a rate of work per unit length of c, then $\mathbf{n}^\alpha$ will be called a *simple dipolar contact force* per unit length. Similarly, if $\mathbf{n}^{\alpha\beta}$ are contravariant tensors (under transformation of surface coordinates x^α) and

if, for all arbitrary second-order velocity gradients $\mathbf{v}_{|\alpha\beta}$, the scalar $\mathbf{n}^{\alpha\beta} \cdot \mathbf{v}_{|\alpha\beta}$ is a rate of work per unit length of c, then $\mathbf{n}^{\alpha\beta}$ will be called a *simple tripolar contact force* per unit length.

Similar definitions can be provided for the three-dimensional *assigned force* $\mathbf{f}$ per unit mass of $\mathfrak{s}$ and the *simple dipolar force* $\mathbf{f}^{\alpha}$ and *simple tripolar force* $\mathbf{f}^{\alpha\beta}$, each per unit mass, through their (scalar) rate of work per unit area of $\mathfrak{s}$. The above definitions can be easily generalized for *simple multipolar contact forces* per unit length ($\mathbf{n}^{\alpha_1 \cdots \alpha_m}$, $m = 1, 2, \cdots n$) and *simple multipolar forces* per unit mass of $\mathfrak{s}$($\mathbf{f}^{\alpha_1 \cdots \alpha_m}$).

We assume that the forces $\mathbf{f}$, $\mathbf{f}^{\alpha}$, and $\mathbf{f}^{\alpha\beta}$, per unit mass, act throughout σ (an arbitrary area of $\mathfrak{s}$ at time t) and that the contact forces $\mathbf{n}$, $\mathbf{n}^{\alpha}$, $\mathbf{n}^{\alpha\beta}$, all per unit length, act across c (a boundary curve of σ). In addition, we assume the existence of an internal energy function U per unit mass, an entropy function S per unit mass, a heat supply function r per unit mass and per unit time, a local temperature T (> 0), and a heat flux h per unit length of c. The equation for balance of energy and the Clausius–Duhem inequality may then be written as

$$\frac{D}{Dt}\int_\sigma \rho[U + \tfrac{1}{2}v^2 + K]\, d\sigma$$

$$= \int_\sigma \rho[r + \mathbf{f}\cdot\mathbf{v} + \mathbf{f}^{\alpha}\cdot\mathbf{v}_{|\alpha} + \mathbf{f}^{\alpha\beta}\cdot\mathbf{v}_{|\alpha\beta}]\, d\sigma$$

$$+ \int_c [\mathbf{n}\cdot\mathbf{v} + \mathbf{n}^{\alpha}\cdot\mathbf{v}_{|\alpha} + \mathbf{n}^{\alpha\beta}\cdot\mathbf{v}_{|\alpha\beta}]\, dc - \int_c h\, dc \quad (5.1)$$

and

$$\frac{D}{Dt}\int_\sigma \rho S\, d\sigma - \int_\sigma \rho\frac{r}{T}\, d\sigma + \int_c \frac{h}{T}\, dc \geq 0, \quad (5.2)$$

where ρ is the mass density, v is the magnitude of $\mathbf{v}$, D/Dt denotes the material derivative, K in (5.1), representing the contributions to the kinetic energy due to velocity gradients, is taken in the form

$$K = \tfrac{1}{2}[y^{\alpha\beta}\mathbf{v}_{|\alpha}\cdot\mathbf{v}_{|\beta} + y^{\alpha\beta\gamma\delta}\mathbf{v}_{|\alpha\beta}\cdot\mathbf{v}_{|\gamma\delta}]. \quad (5.3)$$

with $y^{\alpha\beta}$ and $y^{\alpha\beta\gamma\delta}$ assumed to be functions of the surface coordinates x^{α} only, so that $\dot{y}^{\alpha\beta} = 0$ and $\dot{y}^{\alpha\beta\gamma\delta} = 0$. Without loss of generality the coefficients $y^{\alpha\beta}$, $y^{\alpha\beta\gamma\delta}$ have obvious symmetry properties. Product terms in $\mathbf{v}$, $\mathbf{v}_{|\alpha}$, $\mathbf{v}_{|\alpha\beta}$ could be included but we restrict attention to (5.3).

Recalling that the rate of change of an element of surface area $d\sigma$ is given by

$$\dot{d\sigma} = \tfrac{1}{2}a^{-1}\dot{a}\, d\sigma, \quad d\sigma = a^{\frac{1}{2}}\, dx^1\, dx^2, \quad (5.4)$$

then after carrying out the indicated differentiation, (5.1) becomes

$$\int_\sigma \rho(\dot{U} + \dot{\mathbf{v}}\cdot\mathbf{v})\, d\sigma + \int_\sigma [U + \tfrac{1}{2}v^2 + K]a^{-\frac{1}{2}}\dot{(\rho a^{\frac{1}{2}})}\, d\sigma$$

$$- \int_\sigma \rho\mathbf{f}\cdot\mathbf{v}\, d\sigma - \int_c \mathbf{n}\cdot\mathbf{v}\, dc$$

$$= \int_\sigma \rho[r + \bar{\mathbf{f}}^{\alpha}\cdot\mathbf{v}_{|\alpha} + \bar{\mathbf{f}}^{\alpha\beta}\cdot\mathbf{v}_{|\alpha\beta}]\, d\sigma$$

$$+ \int_c [\mathbf{n}^{\alpha}\cdot\mathbf{v}_{|\alpha} + \mathbf{n}^{\alpha\beta}\cdot\mathbf{v}_{|\alpha\beta} - h]\, dc, \quad (5.5)$$

where a superposed dot stands for material derivative and where we have put

$$\bar{\mathbf{f}}^{\alpha} = \mathbf{f}^{\alpha} - y^{\alpha\beta}\dot{(\mathbf{v}_{|\beta})}, \quad \bar{\mathbf{f}}^{\alpha\beta} = \mathbf{f}^{\alpha\beta} - y^{\alpha\beta\gamma\delta}\dot{(\mathbf{v}_{|\gamma\delta})}. \quad (5.6)$$

Since $\mathbf{v}_{|\alpha\beta}$ is symmetric, it is clear from (5.5) that the tensor $\mathbf{n}^{\alpha\beta}$ in $\mathbf{n}^{\alpha\beta}\cdot\mathbf{v}_{|\alpha\beta}$ may be taken to be symmetric in (α, β) without loss in generality. A similar remark is applicable to $\bar{\mathbf{f}}^{\alpha\beta}$ in (5.5).

We now assume that ρ, U, h, r, $\mathbf{n}$, $\mathbf{n}^{\alpha}$, $\mathbf{n}^{\alpha\beta}$, $(\mathbf{f} - \dot{\mathbf{v}})$, $\bar{\mathbf{f}}^{\alpha}$, $\bar{\mathbf{f}}^{\alpha\beta}$, are unaltered by arbitrary superposed uniform rigid body translational velocities and proceed to consider the invariance of (5.5) under such motions. Thus, in view of (4.1), if we replace $\mathbf{v}$ by $(\mathbf{v} + \mathbf{b})$ in (5.5), then after subtraction follows the equation

$$\mathbf{b}\cdot\left\{\int_\sigma [\rho(\dot{\mathbf{v}} - \mathbf{f}) + \mathbf{v}a^{-\frac{1}{2}}\dot{(\rho a^{\frac{1}{2}})}]\, d\sigma - \int_c \mathbf{n}\, dc\right\}$$

$$+ \tfrac{1}{2}(\mathbf{b}\cdot\mathbf{b})\left\{\int_\sigma a^{-\frac{1}{2}}\frac{D}{Dt}(\rho a^{\frac{1}{2}})\, d\sigma\right\} = 0, \quad (5.7)$$

which must hold for all arbitrary constant vectors $\mathbf{b}$. By replacing $\mathbf{b}$ by $\beta\mathbf{b}$, β being a scalar, it then follows from (5.7) that

$$\int_\sigma a^{-\frac{1}{2}}\frac{D}{Dt}(\rho a^{\frac{1}{2}})\, d\sigma = 0, \quad (5.8)$$

$$\int_\sigma \left[\rho(\dot{\mathbf{v}} - \mathbf{f}) + \mathbf{v}a^{-\frac{1}{2}}\frac{D}{Dt}(\rho a^{\frac{1}{2}})\right] d\sigma - \int_c \mathbf{n}\, dc = 0. \quad (5.9)$$

Since (5.8) holds for all arbitrary areas σ, it follows that

$$a^{-\frac{1}{2}}(D/Dt)(\rho a^{\frac{1}{2}}) = 0, \quad (5.10)$$

or alternatively we have

$$\dot{\rho} + \tfrac{1}{2}\rho\frac{\dot{a}}{a} = \dot{\rho} + \rho\eta^{\lambda}_{\lambda} = 0, \quad (5.11)$$

where (3.15) has been used. In view of (5.10), (5.9) reduces to

$$\int_\sigma [\rho(\dot{\mathbf{v}} - \mathbf{f})]\, d\sigma - \int_c \mathbf{n}\, dc = 0, \quad (5.12)$$

and the energy equation (5.5) assumes the simpler form

$$\int_\sigma \rho[\dot{U} - r + (\dot{\mathbf{v}} - \mathbf{f})\cdot\mathbf{v} - \bar{\mathbf{f}}^\alpha \cdot \mathbf{v}_{|\alpha} - \bar{\mathbf{f}}^{\alpha\beta}\cdot\mathbf{v}_{|\alpha\beta}]\, d\sigma$$
$$= \int_c [\mathbf{n}\cdot\mathbf{v} + \mathbf{n}^\alpha\cdot\mathbf{v}_{|\alpha} + \mathbf{n}^{\alpha\beta}\cdot\mathbf{v}_{|\alpha\beta} - h]\, dc, \quad (5.13)$$

where $\bar{\mathbf{f}}^\alpha$ and $\bar{\mathbf{f}}^{\alpha\beta}$ given by (5.6) represent the difference between the assigned forces per unit mass and the inertia terms due to higher-order velocity gradients.

The simple contact forces $\mathbf{n}$, $\mathbf{n}^\alpha$, $\mathbf{n}^{\alpha\beta}$, when referred to the base vectors $\mathbf{a}_i$, can be expressed as

$$\begin{aligned}\mathbf{n} &= n^i\mathbf{a}_i = n^\lambda\mathbf{a}_\lambda + n^3\mathbf{a}_3,\\ \mathbf{n}^\alpha &= n^{i\alpha}\mathbf{a}_i = n^{\lambda\alpha}\mathbf{a}_\lambda + n^{3\alpha}\mathbf{a}_3,\\ \mathbf{n}^{\alpha\beta} &= n^{i\alpha\beta}\mathbf{a}_i = n^{\lambda\alpha\beta}\mathbf{a}_\lambda + n^{3\alpha\beta}\mathbf{a}_3.\end{aligned} \quad (5.14)$$

Similarly, the simple force multiples $\mathbf{f}$, $\mathbf{f}^\alpha$, $\mathbf{f}^{\alpha\beta}$, as well as $\bar{\mathbf{f}}^\alpha$, $\bar{\mathbf{f}}^{\alpha\beta}$ may be written in the forms

$$\begin{aligned}\mathbf{f} = f^i\mathbf{a}_i, \quad \mathbf{f}^\alpha = f^{i\alpha}\mathbf{a}_i, \quad \mathbf{f}^{\alpha\beta} = f^{i\alpha\beta}\mathbf{a}_i,\\ \bar{\mathbf{f}}^\alpha = \bar{f}^{i\alpha}\mathbf{a}_i, \quad \bar{\mathbf{f}}^{\alpha\beta} = \bar{f}^{i\alpha\beta}\mathbf{a}_i.\end{aligned} \quad (5.15)$$

Over a curve with unit normal $\boldsymbol{\nu} = \nu_\alpha\mathbf{a}^\alpha$, the (physical) contact force vector is $\mathbf{n}$. If $\mathbf{n}'^\alpha$ are the (physical) force vectors over each coordinate line, then application of (5.12) to a curvilinear triangle yields

$$\mathbf{n} = \sum_\alpha \nu_\alpha \mathbf{n}'^\alpha (a^{\alpha\alpha})^{\frac{1}{2}} = \mathbf{N}^\alpha\nu_\alpha, \quad \mathbf{N}^\alpha = \mathbf{n}'^\alpha(a^{\alpha\alpha})^{\frac{1}{2}}. \quad (5.16)$$

Hence, $\mathbf{N}^\alpha$ transforms as a contravariant surface vector and we can write

$$\mathbf{N}^\alpha = N^{i\alpha}\mathbf{a}_i = N^{\lambda\alpha}\mathbf{a}_\lambda + N^{3\alpha}\mathbf{a}_3 \quad (5.17)$$

and by $(5.14)_1$

$$n^i = N^{i\alpha}\nu_\alpha, \quad (5.18)$$

where $N^{\lambda\alpha}$ and $N^{3\alpha}$ are surface tensors under transformation of surface coordinates.

Substituting (5.16) into (5.12), then under the usual smoothness assumptions and making use of Stokes' theorem, we transform the line integral into a surface integral and deduce

$$\int_\sigma [\rho(\dot{\mathbf{v}} - \mathbf{f}) - \mathbf{N}^\alpha_{|\alpha}]\, d\sigma = 0, \quad (5.19)$$

which holds for arbitrary σ. Hence, we have the equations of motion

$$\mathbf{N}^\alpha_{|\alpha} + \rho\mathbf{f} = \rho\dot{\mathbf{v}}. \quad (5.20)$$

Before further considerations of the energy equation (5.13), we return to the entropy production inequality (5.2) and with the help of (5.10) deduce the inequality

$$\int_\sigma \rho\left[\dot{S} - \frac{r}{T}\right] d\sigma + \int_c \frac{h}{T}\, dc \geq 0. \quad (5.21)$$

If h^α is the value of the flux of heat across the x^α curves, then applying (5.21) to an arbitrary curvilinear triangle on $\mathfrak{s}$, bounded by coordinate curves through the point x^α and by c with unit normal ν_α, we obtain

$$h \geq \nu_\alpha q^\alpha, \quad q^\alpha = h^\alpha(a^{\alpha\alpha})^{\frac{1}{2}} \quad (5.22)$$

which holds for all ν_α. Now, let h^- be the flux of heat across c corresponding to $-\nu_\alpha$, then by $(5.22)_1$

$$h^- \geq -\nu_\alpha q^\alpha, \quad (5.23)$$

since q^α is the same for a given thermodynamic process. Combining the inequalities $(5.22)_1$ and (5.23), we get

$$-h^- \leq \nu_\alpha q^\alpha \leq h. \quad (5.24)$$

But, $h = -h^-$ since the heat flux must be continuous across all curves on $\mathfrak{s}$. Hence, we must have[14]

$$h = \nu_\alpha q^\alpha. \quad (5.25)$$

Denoting by $\mathbf{n}'^{\eta\alpha}$, $\mathbf{n}'^{\eta\alpha\beta}$ the (physical) simple force multipoles over the x^η curves, then application of the energy equation (5.13) to a curvilinear triangle on $\mathfrak{s}$, in view of (5.20) and (5.25), results in

$$\bar{\mathbf{n}}^\alpha\cdot\mathbf{v}_{|\alpha} + \bar{\mathbf{n}}^{\alpha\beta}\cdot\mathbf{v}_{|\alpha\beta} = 0, \quad (5.26)$$

where we have set

$$\bar{\mathbf{n}}^\alpha = \mathbf{n}^\alpha - \mathbf{N}^{\eta\alpha}\nu_\eta, \qquad \mathbf{N}^{\eta\alpha} = \mathbf{n}'^{\eta\alpha}(a^{\eta\eta})^{\frac{1}{2}}, \quad (5.27)$$

$$\bar{\mathbf{n}}^{\alpha\beta} = \mathbf{n}^{\alpha\beta} - \mathbf{N}^{\eta\alpha\beta}\nu_\eta, \quad \mathbf{N}^{\eta\alpha\beta} = \mathbf{n}'^{\eta\alpha\beta}(a^{\eta\eta})^{\frac{1}{2}}. \quad (5.28)$$

Introducing the notations

$$\begin{aligned}\mathbf{G}^\alpha &= \mathbf{N}^\alpha + \mathbf{N}^{\eta\alpha}_{|\eta} + \rho\bar{\mathbf{f}}^\alpha,\\ \mathbf{G}^{\alpha\beta} &= \mathbf{N}^{\alpha\beta} + \mathbf{N}^{\eta\alpha\beta}_{|\eta} + \rho\bar{\mathbf{f}}^{\alpha\beta},\\ \mathbf{G}^{\alpha\beta\gamma} &= \mathbf{N}^{\alpha\beta\gamma},\end{aligned} \quad (5.29)$$

then substituting (5.26) into (5.13), transforming the line integrals into surface integrals and after rearranging the results and using (5.20), we obtain

$$\rho r - q^\alpha_{|\alpha} - \rho\dot{U} + \mathbf{G}^\alpha\cdot\mathbf{v}_{|\alpha} + \mathbf{G}^{\alpha\beta}\cdot\mathbf{v}_{|\alpha\beta} + \mathbf{G}^{\gamma\alpha\beta}\cdot\mathbf{v}_{|\alpha\beta\gamma} = 0. \quad (5.30)$$

Next, we consider motions which differ from the given motion by a superposed uniform rigid body angular velocity, the surface $\mathfrak{s}$ occupying the same position at time t. We assume that ρ, U, r, q^α, $\bar{\mathbf{n}}^\alpha$, $\bar{\mathbf{n}}^{\alpha\beta}$, $\mathbf{G}^\alpha$, $\mathbf{G}^{\alpha\beta}$, $\mathbf{G}^{\alpha\beta\gamma}$, are all unaltered when the surface $\mathfrak{s}$ is subjected to a superposed uniform rigid body angular velocity. Thus, using the results (4.2) to (4.6) and (3.13), we deduce from (5.26) and (5.30) the equations[15]

$$\bar{\mathbf{n}}^\alpha\times\mathbf{a}_\alpha + b_{\alpha\beta}\bar{\mathbf{n}}^{\alpha\beta}\times\mathbf{a}_3 = 0, \quad (5.31)$$

[14] The result (5.25) deduced from (5.22) can also be shown in the case of a recent paper by Green, Naghdi, and Wainwright[7] which deals with a general theory of a Cosserat surface.

[15] It can be shown that the components of $\boldsymbol{\eta}_\alpha$, $\boldsymbol{\eta}_{\alpha\beta}$, $\boldsymbol{\eta}_{\alpha\beta\gamma}$ with respect to $\mathbf{a}^i$ are unaltered under all superposed rigid body motions.

and

$$(\mathbf{G}^{\nu} - b_{\alpha\beta}b_{\eta}^{\nu}\mathbf{G}^{\eta\alpha\beta}) \times \mathbf{a}_{\nu} + (b_{\alpha\beta}\mathbf{G}^{\alpha\beta} + b_{\alpha\beta|\gamma}\mathbf{G}^{\gamma\alpha\beta}) \times \mathbf{a}_3 = 0, \quad (5.32)$$

together with

$$\bar{\mathbf{n}}^{\alpha} \cdot \boldsymbol{\eta}_{\alpha} + \bar{\mathbf{n}}^{\alpha\beta} \cdot \dot{\boldsymbol{\eta}}_{(\alpha\beta)} = 0 \quad (5.33)$$

and

$$\rho r - q^{\alpha}_{|\alpha} - \rho\dot{U} + \mathbf{G}^{\alpha} \cdot \boldsymbol{\eta}_{\alpha} + \mathbf{G}^{\alpha\beta} \cdot \boldsymbol{\eta}_{\alpha\beta} + \mathbf{G}^{\gamma\alpha\beta} \cdot \boldsymbol{\eta}_{\alpha\beta\gamma} = 0. \quad (5.34)$$

We note here that since $\mathbf{n}^{\alpha\beta} = \mathbf{n}^{(\alpha\beta)}$ and since $\mathbf{N}^{\eta\alpha\beta}$ (being the values of $\mathbf{n}^{(\alpha\beta)}$ over the coordinate curves x^{η}) is symmetric in (α, β), then $\bar{\mathbf{n}}^{\alpha\beta}$ given by (5.28) is symmetric in (α, β). Also, by (5.29), we have

$$\mathbf{G}^{\alpha[\beta\gamma]} = \tfrac{1}{2}[\mathbf{G}^{\alpha\beta\gamma} - \mathbf{G}^{\alpha\gamma\beta}] = 0. \quad (5.35)$$

We complete our general theory by returning to the entropy production inequality (5.21). With the help of (5.25) and under suitable smoothness assumptions, after transforming the line integral into a surface integral, we deduce from (5.21) the (local) inequality

$$\rho T\dot{S} - \rho r + q^{\alpha}_{|\alpha} - q^{\alpha}(T_{,\alpha}/T) \geq 0. \quad (5.36)$$

The basic field equations of the theory then consist of: the (local) equation (5.11) for conservation of mass, the equations of motion (5.20), the energy equation (5.34), as well as the remaining basic equations and symmetry conditions (5.33), (5.31), (5.32), together with the inequality (5.36). Inspection of (5.33) and (5.34) suggests that for a complete theory constitutive equations must be found for q^{α}, U, $\bar{\mathbf{n}}^{\alpha}$, $\bar{\mathbf{n}}^{\alpha\beta}$, $\mathbf{G}^{\alpha}$, $\mathbf{G}^{\alpha\beta}$, and $\mathbf{G}^{\alpha\beta\gamma}$ and these can be reduced to a canonical form with the use of invariance conditions for each equation which keeps the left-hand sides of (5.33) and (5.34) unaltered by all superposed rigid body motions. If we assume that $\bar{\mathbf{n}}^{\alpha}$, $\bar{\mathbf{n}}^{\alpha\beta}$ are (invariant) vector functions under transformation of surface coordinates, it then follows from (5.27), (5.28) and (5.29) that $\mathbf{N}^{\eta\alpha}$, $\mathbf{N}^{\eta\alpha\beta}$, $\mathbf{G}^{\alpha}$, $\mathbf{G}^{\alpha\beta}$, etc., transform as contravariant tensors under transformation of surface coordinates. Moreover, if we set

$$\mathbf{G}^{\eta\alpha} = G^{i\eta\alpha}\mathbf{a}_i = G^{\lambda\eta\alpha}\mathbf{a}_{\lambda} + G^{3\eta\alpha}\mathbf{a}_3, \quad (5.37)$$

$$\mathbf{G}^{\eta\alpha\beta} = G^{i\eta\alpha\beta}\mathbf{a}_i = G^{\lambda\eta\alpha\beta}\mathbf{a}_{\lambda} + G^{3\eta\alpha\beta}\mathbf{a}_3, \quad (5.38)$$

then $G^{\lambda\eta\alpha}$, $G^{3\eta\alpha}$, $G^{\lambda\eta\alpha\beta}$, $G^{3\eta\alpha\beta}$ are surface tensors.

6. ALTERNATIVE FORMS OF THE BASIC FIELD EQUATIONS

Although the derivation of the basic theory of Sec. 5 in vector form is simple and attractive, it conceals the relative complexity of the resulting equations. For this reason and for future reference, we collect here the basic equations of Sec. 5 in terms of tensor components.

By taking the scalar products of (5.20) with $\mathbf{a}^{\beta}$ and again with $\mathbf{a}_3$ and using (5.17) and (2.13), we obtain the equations of motion in the form

$$\begin{aligned} N^{\beta\alpha}_{|\alpha} - b^{\beta}_{\alpha}N^{3\alpha} + \rho f^{\beta} &= \rho c^{\beta}, \\ N^{3\alpha}_{|\alpha} + b_{\alpha\beta}N^{\beta\alpha} + \rho f^{3} &= \rho c^{3}, \end{aligned} \quad (6.1)$$

where $c^i = \{c^{\beta}, c^3\}$ are the components of acceleration. Also, from (5.29), we have the differential equations

$$\begin{aligned} G^{\lambda\alpha} - \rho f^{\lambda\alpha} &= N^{\lambda\alpha} + N^{\lambda\eta\alpha}_{|\eta} - b^{\lambda}_{\eta}N^{3\eta\alpha}, \\ G^{3\alpha} - \rho f^{3\alpha} &= N^{3\alpha} + N^{3\eta\alpha}_{|\eta} + b_{\nu\eta}N^{\nu\eta\alpha}, \end{aligned} \quad (6.2)$$

$$\begin{aligned} G^{\lambda\alpha\beta} - \rho f^{\lambda(\alpha\beta)} &= N^{\lambda\alpha\beta} + N^{\lambda\eta(\alpha\beta)}_{|\eta} - b^{\lambda}_{\eta}N^{3\eta(\alpha\beta)}, \\ G^{3\alpha\beta} - \rho f^{3(\alpha\beta)} &= N^{3\alpha\beta} + N^{3\eta(\alpha\beta)}_{|\eta} + b_{\lambda\eta}N^{\lambda\eta(\alpha\beta)}, \end{aligned} \quad (6.3)$$

$$G^{\lambda\alpha(\beta\gamma)} = N^{\lambda\alpha(\beta\gamma)}, \quad G^{3\alpha(\beta\gamma)} = N^{3\alpha(\beta\gamma)}, \quad (6.4)$$

where $G^{i\alpha}$, $G^{i(\alpha\beta)}$, $G^{i\alpha(\beta\gamma)}$ are defined by (5.37) and (5.38).

With $\bar{n}^{i\alpha} = \bar{\mathbf{n}}^{\alpha} \cdot \mathbf{a}^i$, $\bar{n}^{i\alpha\beta} = \bar{\mathbf{n}}^{\alpha\beta} \cdot \mathbf{a}^i$, the component form of (5.31) may be written as

$$\epsilon_{i\alpha k}\bar{n}^{i\alpha} + \epsilon_{i3k}b_{\alpha\beta}\bar{n}^{i\alpha\beta} = 0, \quad (6.5)$$

or equivalently

$$\begin{aligned} \epsilon_{\lambda\alpha}\bar{n}^{\lambda\alpha} &= 0, \\ \bar{n}^{3\nu} - b_{\alpha\beta}\bar{n}^{\nu\alpha\beta} &= 0. \end{aligned} \quad (6.6)$$

Similarly, from (5.32), we obtain

$$\begin{aligned} \epsilon_{\lambda\alpha}[G^{\lambda\alpha} + G^{\alpha\epsilon(\beta\gamma)}b_{\beta\gamma}b^{\lambda}_{\epsilon}] &= 0, \\ G^{3\nu} - b_{\alpha\beta}G^{\nu\alpha\beta} - b^{\nu}_{\gamma}b_{\alpha\beta}G^{3\gamma(\alpha\beta)} - b_{\alpha\beta|\gamma}G^{\nu\gamma(\alpha\beta)} &= 0. \end{aligned} \quad (6.7)$$

The component form of (5.33) is given by

$$\bar{n}^{\lambda\alpha}\eta_{(\lambda\alpha)} + \bar{n}^{i\alpha\beta}\eta_{i(\alpha\beta)} = 0 \quad (6.8)$$

and the rate of energy equation (5.34) may be expressed in the form

$$\rho r - q^{\alpha}_{|\alpha} - \rho(T\dot{S} + \dot{T}S - \dot{A}) + G^{\lambda\alpha}\eta_{(\lambda\alpha)} + G^{i\alpha\beta}\eta_{i(\alpha\beta)} + G^{i\gamma\alpha\beta}\eta_{i(\alpha\beta\gamma)} = 0, \quad (6.9)$$

with A, defined by

$$A = U - TS, \quad (6.10)$$

being the Helmholtz free energy function per unit mass.

From (6.6) and (6.7) follow the results

$$\begin{aligned} n^{[\lambda\alpha]} &= N^{[\lambda(\epsilon)\alpha]}\nu_{\epsilon}, \\ (n^{3\lambda} - b_{\alpha\beta}n^{\lambda(\alpha\beta)}) &= (N^{3\epsilon\lambda} - b_{\alpha\beta}N^{\lambda\epsilon(\alpha\beta)})\nu_{\epsilon} \end{aligned} \quad (6.11)$$

and

$$\begin{aligned} G^{[\lambda\alpha]} + b_{\beta\gamma}b^{[\lambda}_{\epsilon}G^{\alpha]\epsilon(\beta\gamma)} &= 0, \\ G^{3\nu} - b_{\alpha\beta}G^{\nu\alpha\beta} - b^{\nu}_{\gamma}b_{\alpha\beta}G^{3\gamma(\alpha\beta)} - b_{\alpha\beta|\gamma}G^{\nu(\alpha\beta\gamma)} &= 0, \end{aligned} \quad (6.12)$$

where $n^{[\lambda\alpha]}$ is the usual notation for the antisymmetric part of $n^{\lambda\alpha}$. Also, by (5.35), we have

$$G^{i\alpha[\beta\gamma]} = 0. \tag{6.13}$$

Supplementary to the above field equations, we also have the equation (5.11) for conservation of mass and the inequality (5.36).

7. AN ELASTIC SURFACE

Guided by the form of the rate of energy equation (6.9) and other basic field equations in Sec. 6, and since the constitutive equations for an elastic continuum are rate independent, the constitutive equations for an elastic surface may be defined to depend on the first and second fundamental forms (in both deformed and the initial configurations), as well as their gradients. However, since $a_{\alpha\beta|\gamma} = 0$ and $\Gamma^{\lambda}_{\alpha\beta}$ is not a tensor, for the tripolar theory of this paper, we adopt as the kinematic variables $e_{\alpha\beta}$, $\kappa_{\alpha\beta}$, $A_{\alpha\beta}$, $A_{\alpha\beta|\gamma}$, $A_{\alpha\beta|\gamma\delta}$, $b_{\alpha\beta|\gamma}$, $B_{\alpha\beta}$, the first two of which are given by (3.12). Thus, we define an elastic surface by the constitutive assumptions[16]

$$A = A(T, e_{\alpha\beta}, \kappa_{\alpha\beta}, A_{\alpha\beta|\gamma}, A_{\alpha\beta|\gamma\delta}, b_{\alpha\beta|\gamma}), \tag{7.1}$$

together with assumptions of similar form for S, $N^{i\eta}$, $N^{i\eta\alpha}$, $N^{i\eta\alpha\beta}$, the constitutive assumptions

$$q^{\nu} = q^{\nu}(T, T_{,\alpha}, e_{\alpha\beta}, \kappa_{\alpha\beta}, A_{\alpha\beta|\gamma}, A_{\alpha\beta|\gamma\delta}, b_{\alpha\beta|\gamma}), \tag{7.2}$$

$$n^{i\alpha} = n^{i\alpha}(T, e_{\alpha\beta}, \kappa_{\alpha\beta}, A_{\alpha\beta|\gamma}, A_{\alpha\beta|\gamma\delta}, b_{\alpha\beta|\gamma}, \nu_{\gamma}), \tag{7.3}$$

and an assumption similar to (7.3) for $n^{i(\alpha\beta)}$. Although not shown explicitly, the dependence of the above constitutive functions on the initial values of the first and the second fundamental forms $A_{\alpha\beta}$ and $B_{\alpha\beta}$ are understood.

Recalling that (6.8) holds for all arbitrary values of the velocity gradients and since the constitutive assumptions are independent of the rate of change of the kinematic variables, it follows from (6.8) that

$$n^{i\alpha} = N^{i\eta\alpha}\nu_{\eta}, \tag{7.4}$$

$$n^{i(\alpha\beta)} = N^{i\eta(\alpha\beta)}\nu_{\eta}. \tag{7.5}$$

Equations (7.4) and (7.5) also show that $N^{i\eta\alpha}$ and $N^{i\eta(\alpha\beta)}$ are surface tensors with respect to the index η. Moreover, in view of (7.4) and (7.5), Eqs. (6.11) are now identically satisfied.

Combining (6.9) with (5.36), we obtain

$$G^{(\alpha\beta)}\eta_{(\alpha\beta)} + G^{i(\alpha\beta)}\eta_{i(\alpha\beta)} + G^{i(\alpha\beta\gamma)}\eta_{i(\alpha\beta\gamma)} - \rho\dot{A} - \rho S\dot{T} - \frac{1}{T}q^{\alpha}T_{,\alpha} \geq 0. \tag{7.6}$$

From (7.1) and using (3.12), (3.14), (4.10), (4.16), (4.19), and (4.20), the rate of change of the free energy A may be expressed as[17]

$$\begin{aligned}\dot{A} = {} & \frac{\partial A}{\partial T}\dot{T} + \frac{\partial A}{\partial e_{\alpha\beta}}\eta_{(\alpha\beta)} - \frac{\partial A}{\partial \kappa_{\alpha\beta}}\eta_{3(\alpha\beta)} + \frac{\partial A}{\partial(b_{\alpha\beta|\gamma})} \\ & \times [\eta_{3(\alpha\beta\gamma)} - (b^{\lambda}_{\alpha}\eta_{\lambda(\beta\gamma)} + b^{\lambda}_{\beta}\eta_{\lambda(\gamma\alpha)} + b^{\lambda}_{\gamma}\eta_{\lambda(\alpha\beta)})] \\ & - \frac{\partial A}{\partial(A_{\alpha\beta|\gamma})}a^{\lambda\nu}[A_{\lambda\alpha}\eta_{\nu(\beta\gamma)} + A_{\lambda\beta}\eta_{\nu(\alpha\gamma)}] - \frac{\partial A}{\partial(A_{\alpha\beta|\gamma\delta})} \\ & \times a^{\lambda\nu}\{A_{\lambda\beta}[\eta_{\nu(\alpha\gamma\delta)} + (b_{\nu\delta}\eta_{3(\alpha\gamma)} + b_{\alpha\gamma}\eta_{3(\nu\delta)}) \\ & - 2b_{\alpha\gamma}b^{\sigma}_{\delta}\eta_{(\sigma\nu)}] + A_{\lambda\alpha}[\eta_{\nu(\beta\gamma\delta)} + b_{\nu\delta}\eta_{3(\beta\gamma)} \\ & + b_{\beta\gamma}\eta_{3(\nu\delta)} - 2b_{\beta\gamma}b^{\sigma}_{\delta}\eta_{(\sigma\nu)}] + A_{\alpha\beta|\lambda}\eta_{\nu(\gamma\delta)} \\ & + [A_{\lambda\alpha|\delta}\eta_{\nu(\beta\gamma)} + A_{\lambda\beta|\delta}\eta_{\nu(\alpha\gamma)}] \\ & + [A_{\lambda\alpha|\gamma}\eta_{\nu(\beta\delta)} + A_{\lambda\beta|\gamma}\eta_{\nu(\alpha\delta)}]\}.\end{aligned} \tag{7.7}$$

Substituting (7.7) into (7.6) and taking account of (7.2) and the constitutive assumption for S, we obtain an inequality which holds for all arbitrary values of $\eta_{(\alpha\beta)}$, $\eta_{i(\alpha\beta)}$, $\eta_{i(\alpha\beta\gamma)}$ and $\dot{T}$ at time t. Then, in a manner similar to that of Coleman and Noll,[18] we deduce the constitutive equations

$$S = -\partial A/\partial T, \tag{7.8}$$

$$G^{(\alpha\beta)} = \rho\frac{\partial A}{\partial e_{\alpha\beta}} + 2\rho(a^{\sigma\alpha}b^{\beta}_{\delta} + a^{\sigma\beta}b^{\alpha}_{\delta})b^{\lambda}_{\gamma}A_{\sigma\nu}\frac{\partial A}{\partial(A_{\lambda\nu|\gamma\delta})}, \tag{7.9}$$

$$\begin{aligned}G^{3(\alpha\beta)} = {} & -\rho\frac{\partial A}{\partial \kappa_{\alpha\beta}} - \rho a^{\lambda\nu}A_{\lambda\gamma}b_{\nu\delta} \\ & \times \left[\frac{\partial A}{\partial(A_{\alpha\gamma|\beta\delta})} + \frac{\partial A}{\partial(A_{\beta\gamma|\alpha\delta})}\right] - \rho A_{\lambda\delta}b_{\nu\gamma} \\ & \times \left[a^{\lambda\alpha}\frac{\partial A}{\partial(A_{\nu\delta|\gamma\beta})} + a^{\lambda\beta}\frac{\partial A}{\partial(A_{\nu\delta|\gamma\alpha})}\right],\end{aligned} \tag{7.10}$$

$$\begin{aligned}G^{\lambda(\alpha\beta)} = {} & -\rho a^{\lambda\nu}A_{\nu\gamma}\left[\frac{\partial A}{\partial(A_{\beta\gamma|\alpha})} + \frac{\partial A}{\partial(A_{\alpha\gamma|\beta})}\right] \\ & - 3\rho b^{\lambda}_{\gamma}\frac{\partial A}{\partial(b_{\alpha\beta|\gamma})} - \rho a^{\lambda\nu} \\ & \times \left\{\tfrac{1}{2}A_{\gamma\delta|\nu}\left(\frac{\partial A}{\partial(A_{\gamma\delta|\alpha\beta})} + \frac{\partial A}{\partial(A_{\gamma\delta|\beta\alpha})}\right)\right. \\ & + A_{\nu\delta|\gamma}\left(\frac{\partial A}{\partial(A_{\beta\delta|\gamma\alpha})} + \frac{\partial A}{\partial(A_{\alpha\delta|\gamma\beta})}\right) \\ & \left. + A_{\nu\gamma|\delta}\left(\frac{\partial A}{\partial(A_{\beta\gamma|\alpha\delta})} + \frac{\partial A}{\partial(A_{\alpha\gamma|\beta\delta})}\right)\right\},\end{aligned} \tag{7.11}$$

[16] The difference $(\Gamma^{\lambda}_{\alpha\beta} - {}_0\Gamma^{\lambda}_{\alpha\beta})$, where ${}_0\Gamma^{\lambda}_{\alpha\beta}$ is the Christoffel symbol of the initial surface, is a tensor and could be used instead of $A_{\alpha\beta|\gamma}$ in (7.1). Also, variables of the type $B_{\alpha\beta|\gamma}$ are not included since they can be expressed in terms of $B_{\lambda\nu}$, $A_{\lambda\nu}$, $A_{\lambda\nu|\gamma}$ and the covariant derivatives of $B_{\lambda\nu}$ with respect to the initial surface.

[17] In evaluating $\partial A/\partial e_{\alpha\beta}$, the tensor $e_{\alpha\beta}$ in A is understood to stand for $\frac{1}{2}(e_{\alpha\beta} + e_{\beta\alpha})$. A similar remark holds for $\partial A/\partial\kappa_{\alpha\beta}$, etc. Also, in evaluating $\partial A/\partial(b_{\alpha\beta|\gamma})$, $b_{\alpha\beta|\gamma}$ in A is written in a form which is completely symmetric in (α, β, γ).

[18] B. D. Coleman and W. Noll, Arch. Ratl. Mech. Anal. 13, 167 (1963).

$$G^{3(\alpha\beta\gamma)} = \rho \frac{\partial A}{\partial(b_{\alpha\beta|\gamma})}, \tag{7.12}$$

$$G^{\lambda(\alpha\beta\gamma)} = -\tfrac{1}{3}\rho a^{\lambda\nu} A_{\nu\delta} \times \left\{ \left[\frac{\partial A}{\partial(A_{\delta\alpha|\beta\gamma})} + \frac{\partial A}{\partial(A_{\delta\beta|\gamma\alpha})} + \frac{\partial A}{\partial(A_{\delta\gamma|\alpha\beta})} \right] + \left[\frac{\partial A}{\partial(A_{\delta\alpha|\gamma\beta})} + \frac{\partial A}{\partial(A_{\delta\beta|\alpha\gamma})} + \frac{\partial A}{\partial(A_{\delta\gamma|\beta\alpha})} \right] \right\}, \tag{7.13}$$

as well as

$$-q^{\alpha} T_{,\alpha} \geq 0. \tag{7.14}$$

Also, with the use of (7.9) to (7.14), the rate of energy equation (6.9) reduces to

$$\rho r - \rho T \dot{S} - q^{\alpha}_{|\alpha} = 0, \tag{7.15}$$

which may be used to determine the temperature.

The constitutive relations (7.9) to (7.14) are valid for an elastic surface which is anisotropic in some preferred state, usually taken to be the initial undeformed state. We can now study the effect of material symmetries in restricting the form of the constitutive equations along the lines discussed by Green and Adkins.[19] In the special case when the surface is initially isotropic with a center of symmetry and if we assume that A is a polynomial in the arguments indicated in (7.1), we can then express A as a function of joint invariants of the arguments in (7.1), as well as B^{α}_{β}. This may be accomplished in a manner similar to that discussed by Green, Naghdi, and Wainwright,[7] where other references to more general and recent developments on symmetry restrictions may also be found.

8. SPECIAL CASES

In this section, we discuss special cases of the theory of an elastic surface with *simple force tripoles*. We recall that for an elastic surface, in addition to the constitutive equations (7.9) to (7.13), we have the symmetry restrictions (6.12) and (6.13), the transformation relations (5.18), (7.4), and (7.5), the equations of motion (6.1), and the equations (6.2), (6.3), and (6.4).

The number of components of surface forces $N^{i\alpha}$, $N^{i\alpha\beta}$, $N^{i\gamma(\alpha\beta)}$ which are present in the above field equations is $6 + 12 + 18 = 36$. However, not all of these are independent of each other; the restriction (6.13) is identically satisfied and (6.12) provides $1 + 2 = 3$ equations which various combinations of $G^{i\alpha}$, $G^{i\alpha\beta}$, $G^{i\alpha\beta\gamma}$ must satisfy. We thus have $36 - 3 = 33$ independent components of forces and the constitutive relations (7.9) to (7.13) supply $3 + 3 + 6 + 4 + 8 = 24$ equations relating these to the kinematic variables, thereby leaving $33 - 24 = 9$ components of forces indeterminate. Alternatively, we may regard (7.9) and (6.12) as equations for $G^{(\lambda\alpha)}$, $G^{[\lambda\alpha]}$ and $G^{3\alpha}$, (7.10) to (7.13) and (6.13) as constitutive relations for $G^{i(\alpha\beta)}$, $G^{i(\alpha\beta\gamma)}$, and $G^{i\alpha[\beta\gamma]}$. If we write $G^{i\gamma(\alpha\beta)}$ in the form

$$G^{i\gamma(\alpha\beta)} = \tfrac{1}{3}\{G^{i\gamma(\alpha\beta)} + G^{i\alpha(\beta\gamma)} + G^{i\beta(\gamma\alpha)}\} + \tfrac{1}{3}[G^{i\gamma(\alpha\beta)} - G^{i\alpha(\beta\gamma)}] + \tfrac{1}{3}[G^{i\gamma(\alpha\beta)} - G^{i\beta(\gamma\alpha)}] = G^{i(\gamma\alpha\beta)} + G^{i\langle\gamma\alpha\beta\rangle}, \tag{8.1}$$

where we have set

$$G^{i\langle\gamma\alpha\beta\rangle} = \tfrac{1}{3}[2G^{i\gamma(\alpha\beta)} - G^{i\alpha(\beta\gamma)} - G^{i\beta(\gamma\alpha)}], \tag{8.2}$$

then we see that $G^{i[\alpha\beta]}$ with 3 independent components and $G^{i\langle\alpha\beta\gamma\rangle}$ with $24 - 6 - 12 = 6$ independent components remain unaccounted for, and the total number of indeterminate components is again $3 + 6 = 9$.

In order to provide a determinate theory of simple force tripoles, we set

$$G^{i[\alpha\beta]} = 0, \quad G^{i\langle\alpha\beta\gamma\rangle} = 0. \tag{8.3}$$

It then follows from (8.3), (6.13), (7.12), and (7.13) that

$$G^{i\alpha\beta} = G^{i(\alpha\beta)}, \quad G^{i\gamma\alpha\beta} = G^{i(\gamma\alpha\beta)}, \tag{8.4}$$

and instead of the restrictions $(6.12)_1$ and (6.13), and the equations (6.2) to (6.4), we now have

$$N^{(\lambda\alpha)} + N^{(\lambda[\mu]\alpha)}_{|\mu} - b^{\lambda}_{\mu} N^{3\mu\alpha} + \rho \bar{f}^{(\lambda\alpha)} = G^{(\lambda\alpha)}, \tag{8.5}$$

$$N^{[\lambda\alpha]} + N^{[\lambda(\mu)\alpha]}_{|\mu} - b^{[\lambda}_{\mu} N^{3\mu\alpha]} + \rho \bar{f}^{[\lambda\alpha]} = G^{[\lambda\alpha]}, \tag{8.6}$$

$$N^{3\alpha} + N^{3\mu\alpha}_{|\mu} + b_{\lambda\mu} N^{\lambda\mu\alpha} + \rho \bar{f}^{3\alpha} = G^{3\alpha}, \tag{8.7}$$

$$N^{\lambda(\alpha\beta)} + N^{\lambda(\mu\alpha\beta)}_{|\mu} - b^{\lambda}_{\mu} N^{3(\mu\alpha\beta)} + \rho \bar{f}^{\lambda(\alpha\beta)} = G^{\lambda(\alpha\beta)}, \tag{8.8}$$

$$N^{\lambda[\alpha\beta]} + \rho \bar{f}^{\lambda[\alpha\beta]} = 0, \tag{8.9}$$

$$N^{3(\alpha\beta)} + N^{3(\mu\alpha\beta)}_{|\mu} + b_{\lambda\mu} N^{\lambda(\mu\alpha\beta)} + \rho \bar{f}^{3(\alpha\beta)} = G^{3(\alpha\beta)}, \tag{8.10}$$

$$N^{3[\alpha\beta]} + \rho \bar{f}^{3[\alpha\beta]} = 0, \tag{8.11}$$

$$N^{3(\alpha\beta\gamma)} = G^{3(\alpha\beta\gamma)}, \quad N^{3\alpha\beta\gamma} - N^{3(\alpha\beta\gamma)} \equiv N^{3\langle\alpha\beta\gamma\rangle} = 0, \tag{8.12}$$

$$N^{\lambda(\alpha\beta\gamma)} = G^{\lambda(\alpha\beta\gamma)}, \quad N^{\lambda\alpha\beta\gamma} - N^{\lambda(\alpha\beta\gamma)} \equiv N^{\lambda\langle\alpha\beta\gamma\rangle} = 0, \tag{8.13}$$

where $G^{(\lambda\alpha)}$, $G^{i(\alpha\beta)}$, $G^{i(\alpha\beta\gamma)}$ are specified by the constitutive equations (7.9) to (7.13), $G^{3\alpha}$ is given by the relation $(6.12)_2$ and $G^{[\lambda\alpha]}$ is determined from $(6.12)_1$.

We return now to the general theory without using (8.3) and no longer admit the simple surface and body force tripoles, i.e., we set

$$n^{i\alpha\beta} = f^{i\alpha\beta} \equiv 0. \tag{8.14}$$

[19] A. E. Green and J. E. Adkins, *Large Elastic Deformations* (Clarendon Press, Oxford, England, 1960).

Then, by (7.5) and (6.4),

$$G^{i\alpha\beta\gamma} \equiv N^{i\alpha\beta\gamma} = 0. \tag{8.15}$$

It follows from (7.12) and (7.13) that in this case

$$\partial A/\partial(b_{\alpha\beta|\gamma}) = \partial A/\partial(A_{\delta\alpha|\beta\gamma}) = 0,$$

so that (7.1) assumes the form[20]

$$A = A(T, e_{\alpha\beta}, \kappa_{\alpha\beta}, A_{\alpha\beta|\gamma}), \tag{8.16}$$

all other constitutive assumptions become independent of $b_{\alpha\beta|\gamma}$ and $A_{\alpha\beta|\gamma\delta}$, and the constitutive equations for $G^{(\lambda\alpha)}$, $G^{i(\alpha\beta)}$, the relations (6.12), and the system of equations (6.2) to (6.4) reduce to

$$G^{(\lambda\alpha)} = \rho\frac{\partial A}{\partial e_{\lambda\alpha}}, \quad G^{[\lambda\alpha]} = 0, \quad G^{3\lambda} = b_{\alpha\beta}G^{\lambda\alpha\beta}, \tag{8.17}$$

$$\begin{aligned} G^{\lambda(\alpha\beta)} &= -\rho a^{\lambda\nu}A_{\nu\gamma}\left[\frac{\partial A}{\partial(A_{\beta\gamma|\alpha})} + \frac{\partial A}{\partial(A_{\alpha\gamma|\beta})}\right], \\ G^{3(\alpha\beta)} &= -\rho\frac{\partial A}{\partial\kappa_{\alpha\beta}}, \end{aligned} \tag{8.18}$$

$$N^{(\lambda\alpha)} + N^{(\lambda[\mu]\alpha)}_{|\mu} - b^{(\lambda}_{\mu}N^{3\mu\alpha)} + \rho\bar{f}^{(\lambda\alpha)} = G^{(\lambda\alpha)}, \tag{8.19}$$

$$N^{[\lambda\alpha]} + N^{[\lambda(\mu)\alpha]}_{|\mu} - b^{[\lambda}_{\mu}N^{3\mu\alpha]} + \rho\bar{f}^{[\lambda\alpha]} = 0, \tag{8.20}$$

$$N^{3\alpha} + N^{3\mu\alpha}_{|\mu} + b_{\lambda\mu}N^{\lambda\mu\alpha} + \rho\bar{f}^{3\alpha} = G^{3\alpha}, \tag{8.21}$$

$$N^{i\alpha\beta} = G^{i\alpha\beta}. \tag{8.22}$$

It follows from an examination of (8.17) to (8.22) that in a theory in which (8.14) is assumed, the three components of $G^{i[\alpha\beta]}$ remain indeterminate. A determinate theory of simple force dipoles may be obtained, however, by setting

$$G^{i[\alpha\beta]} = 0 \tag{8.23}$$

from which, together with (8.22), we have

$$N^{i[\alpha\beta]} = 0.$$

Thus, for a determinate simple dipolar theory of an elastic surface, the relevant field equations are the equations of motion (6.1) and the system of equations (8.17) to (8.23).

Within the scope of a simple dipolar theory [corresponding to (8.14)] if, instead of (8.16), we admit a restricted assumption for the free energy A in the form

$$A = A(T, e_{\alpha\beta}, \kappa_{\alpha\beta}) \tag{8.24}$$

and also assume that all other constitutive assumptions are independent of $A_{\alpha\beta|\gamma}$, we can then conclude from $(8.18)_1$ that $G^{\lambda(\alpha\beta)} = 0$. If we also admit (8.23) by (8.22) and (7.4), we then have

$$\begin{aligned} N^{\lambda[\alpha\beta]} &= 0, \quad n^{\lambda\alpha} = 0, \\ N^{3[\alpha\beta]} &= 0, \quad n^{3\beta} = N^{3(\alpha\beta)}\nu_{\alpha}. \end{aligned} \tag{8.25}$$

Recalling the rate of energy equation (5.13), in view of (8.25) and for consistency we should also put[21]

$$\rho\bar{f}^{\lambda\alpha} = 0. \tag{8.26}$$

We call the theory in which A has the form (8.24), the restricted simple dipolar theory. The field equations of the determinate restricted (dipolar) theory are given by (6.1), (8.17), $(8.18)_2$, (8.19) to (8.21), (8.25), (8.26), and with

$$N^{\lambda\alpha\beta} = G^{\lambda\alpha\beta} = 0. \tag{8.27}$$

These field equations of the restricted dipolar theory (aside from differences in notations) are the same as those obtained previously[7] in the special case that the directors are identified with the unit normals to the surface at all times.

We return again to the original theory without specifying the indeterminate functions. If in addition to (8.14) we also set

$$n^{i\alpha} = f^{i\alpha} = 0, \tag{8.28}$$

then the special case of monopolar theory will result. From (8.28), together with (7.4) and $(5.29)_2$, we have

$$N^{i\alpha\beta} = G^{i\alpha\beta} = 0. \tag{8.29}$$

It then follows from (8.29) and (8.18) that for the monopolar theory,

$$A = A(T, e_{\alpha\beta}), \tag{8.30}$$

and we can easily recover the equations of the membrane theory as given by Green and Adkins.[19]

ACKNOWLEDGMENTS

The results reported here were obtained in the course of research supported by the U.S. Office of Naval Research under Contract Nonr 222(69) with the University of California, Berkeley. Part of this work was carried out when one of us (AEG) held a visiting appointment at the University of California, Berkeley.

[20] The dependence of A in (8.16) on initial values of the first and second fundamental forms is understood.

[21] Our motivation for requiring (8.26) stems from the fact that the term $\mathbf{n}^{\alpha}\cdot\mathbf{v}_{|\alpha}$ in (5.13), on account of $(8.25)_2$, is now reduced to $\mathbf{n}^{\alpha}\cdot\mathbf{v}_{|\alpha} = n^{3\alpha}v_{3;\alpha}$ [using the notation of (3.7)]. Hence for consistency, $\rho\bar{\mathbf{f}}^{\alpha}\cdot\mathbf{v}_{|\alpha}$ in (5.13) should be replaced with $\rho\bar{f}^{3\alpha}v_{3;\alpha}$ and this, in turn, implies (8.26).

Bargmann–Wigner Equations

CHARLES L. CRITCHFIELD
Los Alamos Scientific Laboratory, University of California, Los Alamos, New Mexico

(Received 22 March 1966)

The redundancy of the Bargmann–Wigner equation for free particles of spin S greater than one-half is analyzed. Only $2(2S+1)$ equations involve an essential time derivative, as is well known. An additional $\frac{1}{8}(S+3)(4S^2-1)$ is required to define all components of the representation. The remaining $(S+1)(4S^2-1)$ are time derivatives and divergence conditions; the latter occurring doubly.

IF one has a state comprising N independent, free particles that obey the Dirac equation the wavefunction would be a product

$$\Psi = \prod_{\nu=1}^{N} \psi_\nu(\xi_\nu^i, p_{\nu k}), \tag{1}$$

where ξ_ν^i give the four spin functions ($i = 1, 2, 3, 4$) and $p_{\nu k}$ the four-momentum of the νth particle. Each would have its own γ operators and mass m_ν and the wave equation is

$$\gamma_\nu^k p_{\nu k}\Psi = m_\nu \Psi \quad (\nu = 1, 2, \cdots, N). \tag{2}$$

Bargmann and Wigner[1] have proposed that the relativistic wave equation for a single particle of total spin S and mass m be derived from Eq. (2) by setting all m_ν equal to m, all $p_{\nu k}$ equal to a single four-vector p_k and requiring the wavefunction, Eq. (1), to be symmetric in spin labels ν. The γ_ν^k for different ν (which is now to be considered as a position number in a polynomial) still commute. The wave equation is then

$$\begin{aligned}&(\gamma_\nu^k p_k - m)\Psi_{1,2\cdots\nu\cdots N} = 0,\\ &(\nu = 1, 2, \cdots, N), \quad N = 2S.\end{aligned} \tag{3}$$

Since the wavefunction is symmetric the normalized basis states are uniquely designated by the numbers n_i which give the power to which the spin state ξ^i occurs in the polynomial. Each such polynomial contains

$$C(n_1, n_2, n_3, n_4) = N!/n_1!\, n_2!\, n_3!\, n_4! \tag{4}$$

orthogonal terms so the normalizing factor is $C(n_1, n_2, n_3, n_4)^{-\frac{1}{2}}$. The component belonging to a normalized basis function is denoted by $|n_1, n_2, n_3, n_4\rangle$. The total number of components belonging to a given spin S is just the number of choices of the n_i so as to keep

$$\sum_i n_i = N = 2S.$$

First consider the subset of states in which exactly q of the n_i are not zero, so that there are four such subsets, one for each q value of 1 to 4. The number of choices of i to be represented in each case is then

$$\binom{4}{q} = \frac{24}{q!\,(4-q)!}.$$

The number of partitions of N into q parts is

$$(N-1)!/(q-1)!\,(N-q)!.$$

The total number of states is therefore

$$\begin{aligned}D(N) &= \sum_{q=1}^{4} \frac{24(N-1)!}{q!\,(q-1)!\,(4-q)!\,(N-q)!}\\ &= \tfrac{1}{6}(N+1)(N+2)(N+3).\end{aligned} \tag{5}$$

The apparent number of equations is obtained by multiplying each term by q before summing;

$$\begin{aligned}E_A &= \sum_{q=1}^{4} \frac{24(N-1)!}{[(q-1)!]^2(4-q)!\,(N-q)!}\\ &= \tfrac{2}{3}N(N+1)(N+2).\end{aligned} \tag{6}$$

The ratio of number of equations to number of states approaches 4 for large S. The equations are therefore redundant and it is the purpose of this note to identify the redundancy, part of which is mere duplication of equations.

Let us define the operator $|k\rangle\langle m|$ that acts upon a state $|\cdots n_k \cdots n_m \cdots\rangle = |n_1, n_2, n_3, n_4\rangle$ so as to produce

$$\begin{aligned}&|k\rangle\langle m|\,|\cdots n_k, \cdots n_m, \cdots\rangle\\ &\qquad = |\cdots n_k + 1, \cdots n_m - 1, \cdots\rangle\\ &\text{or} \qquad = 0, \quad \text{if} \quad n_m = 0.\end{aligned}$$

Taking the normalization factors into account we may write Eq. (3) more explicitly

$$\begin{aligned}&\{n_1^{\frac{1}{2}}(p_0 - m) - (n_4+1)^{\frac{1}{2}}|4\rangle\langle 1|\,(p_x - ip_y)\\ &\qquad - (n_3+1)^{\frac{1}{2}}|3\rangle\langle 1|\,p_z\}\,|n_1, n_2, n_3, n_4\rangle = 0,\end{aligned} \tag{7a}$$

[1] V. Bargmann and E. P. Wigner, Proc. Natl. Acad. Sci. U.S. 34, 211 (1948).

$$\{n_2^{\frac{1}{2}}(p_0 - m) - (n_3 + 1)^{\frac{1}{2}} |3\rangle\langle 2| (p_x + ip_y) + (n_4 + 1)^{\frac{1}{2}} |4\rangle\langle 2| p_z\} |n_1, n_2, n_3, n_4\rangle = 0, \quad (7b)$$

$$\{-n_3^{\frac{1}{2}}(p_0 + m) + (n_2 + 1)^{\frac{1}{2}} |2\rangle\langle 3| (p_x - ip_y) + (n_1 + 1)^{\frac{1}{2}} |1\rangle\langle 3| p_z\} |n_1, n_2, n_3, n_4\rangle = 0, \quad (7c)$$

$$\{-n_4^{\frac{1}{2}}(p_0 + m) + (n_1 + 1)^{\frac{1}{2}} |1\rangle\langle 4| (p_x + ip_y) - (n_2 + 1)^{\frac{1}{2}} |2\rangle\langle 4| p_z\} |n_1, n_2, n_3, n_4\rangle = 0. \quad (7d)$$

Eliminating $p_0 - m$ from Eqs. (7a) and (7b) we get

$$n_2^{\frac{1}{2}}(n_4 + 1)^{\frac{1}{2}}(p_x - ip_y) |n_1 - 1, n_2, n_3, n_4 + 1\rangle + n_2^{\frac{1}{2}}(n_3 + 1)^{\frac{1}{2}} p_z |n_1 - 1, n_2, n_3 + 1, n_4\rangle - n_1^{\frac{1}{2}}(n_3 + 1)^{\frac{1}{2}}(p_x + ip_y) |n_1, n_2 - 1, n_3 + 1, n_4\rangle + n_1^{\frac{1}{2}}(n_4 + 1)^{\frac{1}{2}} p_z |n_1, n_2 - 1, n_3, n_4 + 1\rangle = 0. \quad (8)$$

Similarly, eliminating $p_0 + m$ from Eqs. (7c) and (7d) (and using r_i in place of n_i) we find

$$r_4^{\frac{1}{2}}(r_2 + 1)^{\frac{1}{2}}(p_x - ip_y) |r_1, r_2 + 1, r_3 - 1, r_4\rangle + r_4^{\frac{1}{2}}(r_1 + 1)^{\frac{1}{2}} p_z |r_1 + 1, r_2, r_3 - 1, r_4\rangle - r_3^{\frac{1}{2}}(r_1 + 1)^{\frac{1}{2}}(p_x + ip_y) |r_1 + 1, r_2, r_3, r_4 - 1\rangle + r_3^{\frac{1}{2}}(r_2 + 1)^{\frac{1}{2}} p_z |r_1, r_2 + 1, r_3, r_4 - 1\rangle = 0. \quad (9)$$

If we now substitute in Eq. (9),

$$r_1 = n_1 - 1, \quad r_2 = n_2 - 1, \quad r_3 = n_3 + 1, \quad r_4 = n_4 + 1,$$

it becomes identical with Eq. (8). This means that for every function in which both n_1 and n_2 do not vanish there is another giving the same equation. The same holds if neither n_3 nor n_4 vanish (even though n_1, n_2 do so). The total number of such equations in which n_1 and n_2 are not zero is $D(N-2)$

$$D(N-2) = \tfrac{1}{6}N(N^2 - 1)$$

plus an equal number in which n_3 and n_4 do not vanish. These equations may be written

$$p \cdot B = 0 \quad (10)$$

with

$$iB_x(n_1, n_2, n_3, n_4) = 2^{\frac{1}{2}} n_2^{\frac{1}{2}}(n_4 + 1)^{\frac{1}{2}} |n_1 - 1, n_2, n_3, n_4 + 1\rangle - n_1^{\frac{1}{2}}(n_3 + 1)^{\frac{1}{2}} |n_1, n_2 - 1, n_3 + 1, n_4\rangle,$$

$$iB_y(n_1, n_2, n_3, n_4) = -2^{\frac{1}{2}} i\{n_2^{\frac{1}{2}}(n_4 + 1)^{\frac{1}{2}} |n_1 - 1, n_2, n_3, n_4 + 1\rangle + n_1^{\frac{1}{2}}(n_3 + 1)^{\frac{1}{2}} |n_1, n_2 - 1, n_3 + 1, n_4\rangle\}, \quad (11)$$

$$iB_z(n_1, n_2, n_3, n_4) = 2^{\frac{1}{2}} n_2^{\frac{1}{2}}(n_3 + 1)^{\frac{1}{2}} |n_1 - 1, n_2, n_3 + 1, n_4\rangle + n_1^{\frac{1}{2}}(n_4 + 1)^{\frac{1}{2}} |n_1, n_2 - 1, n_3, n_4 + 1\rangle.$$

One can see from the general form of Eqs. (7) that p_0, i.e., the time derivative, can be eliminated from all but a few sets of equations and the function solved for directly. Those states for which p_0 must be retained are those with $q = 1$ (only one n_i not zero) of which there are four and those with $q = 2$ such that either n_1 and n_2 or n_3 and n_4 are the nonvanishing numbers. Then the properly weighted sum would seem appropriate to obtain equations in $(p_0 - m)$ or $p_0 + m$, respectively. Since there are $N - 1$ partitions into two parts of N there are $2(N-1)$ such functions, or with the four with only one n_i nonzero a total of $2(2S + 1)$ equations involving a time derivative. After substitution from the equations that do not contain p_0 these are equivalent to Hamiltonians[2] or propagators of the $2(2S + 1)$-dimensional representations.[3]

In all remaining combinations of n_i, Eqs. (7) can be solved for $m |n_1, n_2, n_3, n_4\rangle$ directly to give a complete set of $\frac{1}{6}(N+1)(N+2)(N+3)$ equations for as many functions. There exist in addition the same number as for $m |\cdots\rangle$, viz., $\frac{1}{6}(N+6)(N^2-1)$, for $p_0 |n_1, n_2, n_3, n_4\rangle$ plus those of Eq. (10), i.e., $\frac{1}{3}N(N^2-1)$, counting each twice. The sum

$$\tfrac{1}{6}(N+1)(N+2)(N+3) + \tfrac{1}{6}(N+6)(N^2-1) + \tfrac{1}{3}N(N^2-1) = \tfrac{8}{3}N(N+1)(N+2) = E_A.$$

As a simple illustration we determine the Proca equations. The functions of spin one that have $q = 1$ and $q = 2$ with i equal to 1 and 2, and 3 and 4, are six in number, viz.,

$$a_1 = |2000\rangle, \quad a_2 = |0200\rangle, \quad a_3 = |0020\rangle,$$
$$a_4 = |0002\rangle, \quad b_{12} = |1100\rangle, \quad b_{34} = |0011\rangle,$$

and may be arranged as follows:

$$mA_x = -i\{a_1 - a_2 - a_3 + a_4\},$$
$$mA_y = +\{a_1 + a_2 - a_3 - a_4\},$$
$$mA_z = i\sqrt{2}\,\{b_{12} - b_{34}\},$$
$$E_x = a_1 - a_2 + a_3 - a_4,$$
$$E_y = i\{a_1 + a_2 + a_3 + a_4\},$$
$$E_z = -\sqrt{2}\,\{b_{12} + b_{34}\}.$$

Replacing $p_0 = i\partial/\partial t$, $\mathbf{p} \to -i\nabla$ in Eqs. (7) the six essential equations with p_0 are, in vector form:

$$(\partial \mathbf{A}/\partial t) + \mathbf{E} + \nabla\varphi = 0, \quad (\partial \mathbf{E}/\partial t) - \nabla \times \mathbf{B} - m^2 A = 0, \quad (12)$$

with $\mathbf{B}$ from Eq. (11) and

$$m\varphi = +i\sqrt{2}\,\{|1001\rangle - |0110\rangle\}. \quad (13)$$

[2] D. L. Weaver, C. L. Hammer, and R. N. Good, Jr., Phys. Rev. 135, B241 (1964).
[3] S. Weinberg, Phys. Rev. 133, B1318 (1964).

The four equations in $m\,|n_1, n_2, n_3, n_4\rangle$ are

$$\mathbf{B} - \nabla \times \mathbf{A} = 0, \quad m^2\varphi + \nabla E = 0, \tag{14}$$

completing the necessary 10. As suggested by Foldy,[4] one can substitute from Eq. (14) into Eq. (12) and obtain a form of Hamiltonian in the six-dimensional representation. This Hamiltonian is quadratic in space derivations. Equations (7) shows that there are 16 apparent equations and of these the divergence equation occurs twice, viz.,

$$\nabla \cdot B = 0, \tag{15}$$

which corresponds to Eq. (10) and which follows from Eq. (14). The remaining four are readily seen to be

$$\begin{aligned} (\partial \mathbf{B}/\partial t) + \nabla \times \mathbf{E} &= 0, \\ (\partial \varphi/\partial t) + \nabla \cdot A &= 0, \end{aligned} \tag{16}$$

i.e., time derivatives that are derivable from Eqs. (12) and (14).

The Bargmann–Wigner equations have the property that the Klein–Gordon wave equation follows at once from Eq. (3) and

$$\gamma^k\gamma^l + \gamma^l\gamma^k = 2g^{kl}. \tag{17}$$

They remain linear in p but require a representation larger than the minimum $2(2S + 1)$. The large representation is generated not only by the Lorentz transformation on spin of which the infinitesimal operator is the six vector S^{kl} but by the four γ^k of Eq. (3). The algebra of this group is given by ($[a, b] = ab - ba$)

$$\begin{aligned} [\gamma^k, \gamma^l] &= -4iS^{kl}, \\ [S^{mn}, \gamma^k] &= i(g^{nk}\gamma^m - g^{mk}\gamma^n), \\ [S^{mn}, S^{kl}] &= i\{g^{kn}S^{ml} - g^{km}S^{nl} - g^{ln}S^{mk} + g^{lm}S^{mn}\}. \end{aligned} \tag{18}$$

The group is of rank two and order ten and the algebra was designated originally as B_2 which is isomorphic with what is known also as C_2. Bhabha[5] has investigated the relationship between relativistic waves of higher spin and representations of B_2. Here we prefer to use the notation more commonly applied to the C algebras. An irreducible representation of C_2 is given by two integers, $\{N_1, N_2\}$, which constitute its highest weight. The maximum spin in such a representation is

$$S = \tfrac{1}{2}(N_1 - N_2). \tag{19}$$

The other diagonal quantum number is that of

$$\gamma^4 = N_1 + N_2.$$

Substates of $\{N_1, N_2\}$, $(N_1 \geq N_2)$, are derived from it by permitted subtractions of the "simple roots,"[6] $(1, -1)$ and $(0, 2)$ and with proper attention to multiplicity of weights. The point here is that symmetric functions such as considered above have highest weights of the form $\{N, 0\}$, i.e., $N_2 = 0$. For a given S there are, according to Eq. (19), infinitely many representations and it might well be more appropriate to use $\{2, 1\}$ in place of $\{1, 0\}$ when $S = \frac{1}{2}$ if the system is made up of three fermions. The dimensionality is given by Weyl's formula and is

$$\begin{aligned} &D(N_1, N_2) \\ &= \tfrac{1}{6}(N_1 - N_2 + 1)(N_1 + 2)(N_1 + N_2 + 3)(N_2 + 1). \end{aligned} \tag{20}$$

A particular application of nonsymmetric representations would be to use the five states of $\{1, 1\}$ to get wave equations for a particle of spin zero.

ACKNOWLEDGMENT

This work was done under the auspices of the U.S. Atomic Energy Commission.

[4] L. L. Foldy, Phys. Rev. **102**, 568 (1956).

[5] H. J. Bhabha, Rev. Mod. Phys. **17**, 200 (1945).

[6] E. B. Dynkin, Am. Math. Soc. Transl., Ser. II **6**, 319 (1957).

Neutron Transport from a Point Source in Two Adjacent, Dissimilar, Semi-Infinite Media

R. C. ERDMANN
University of California, Department of Engineering, Los Angeles, California

(Received 29 September 1966)

A solution to the steady-state, two-dimensional Boltzmann equation is obtained for the flux due to a point source of neutrons located in a system of two adjacent half-spaces having the same mean-free path. Isotropic scattering and monoenergetic neutrons are assumed. The solution for an arbitrary source location is related to that for an interface source which is obtained by the Wiener–Hopf technique. Asymptotic expansions for the flux far from the source, both along and away from the interface, are derived and compared with approximate theory.

SOLVING neutron transport problems which have one spatial variable can be accomplished, for instance, by using the Case[1] technique. When more than one such variable is present, the differential Boltzmann equation has not yet been successfully treated with this method. An alternate method of solution—the Wiener–Hopf technique—can, however, be applied to some two spatial variable problems when the integral form of the Boltzmann equation is used. Elliot[2] has applied this technique to the integral Boltzmann equation to obtain the flux from a point source of neutrons located in a single half-space. In the present paper, the solution for the flux from a point source of neutrons located in either of two adjacent half-spaces is obtained under certain special conditions. Here again the Wiener–Hopf technique is applied to the integral equation describing the neutron field.

I. MATHEMATICAL DESCRIPTION OF THE PROBLEM

Consider a point source of neutrons located in one of two adjacent half-spaces as sketched in Fig. 1. Assuming a monoenergetic system with scattering isotropic in the laboratory system, the neutron flux can be written in the form

$$\rho_{z_0}(r, z) = \iiint_{-\infty}^{+\infty} dx'\, dy'\, dz' \times \left\{\left[\frac{c(z')\sigma(z')}{4\pi}\,\rho_{z_0}(r', z') + q\delta(x')\delta(y')\delta(z' - z_0)\right] \times \left[\frac{e^{-\tau(\mathbf{R},\mathbf{R}')}}{|\mathbf{R} - \mathbf{R}'|^2}\right]\right\}. \quad (1)$$

Here $c(z')$ is the probability of scatter,

$$c(z) = \begin{cases} c_1, & z > 0, \\ c_2, & z < 0; \end{cases}$$

$\sigma(z)$ is the total cross section in the same notation; q is a measure of source strength; and

$$r = (x^2 + y^2)^{\frac{1}{2}}, \quad R = (r^2 + z^2)^{\frac{1}{2}}.$$

$\tau(\mathbf{R}, \mathbf{R}')$, the optical depth,[3] is the total cross section integrated from $\mathbf{R}'$ to $\mathbf{R}$:

$$\tau(\mathbf{R}, \mathbf{R}') = \int_0^{|\mathbf{R}-\mathbf{R}'|} \sigma(R'')\, dR''.$$

To solve the problem exactly, it was found necessary to add the restriction $\sigma_1 = \sigma_2 = \sigma$, which gives $\tau(\mathbf{R}, \mathbf{R}') = \sigma|\mathbf{R} - \mathbf{R}'|$. Then Eq. (1) can be written, in units of mean free path, as

$$\rho_{z_0}(r, z) = \iiint_{-\infty}^{+\infty} dx'\, dy'\, dz' \times \left[\frac{c(z')\rho_{z_0}(r', z')}{4\pi} + q\delta(x')\delta(y')\delta(z' - z_0)\right] \times \frac{e^{-|\mathbf{R}-\mathbf{R}'|}}{|\mathbf{R} - \mathbf{R}'|^2}. \quad (2)$$

It is shown in Sec. IV that the solution for $\rho_{z_0}(r, z)$ can be written in terms of that for $\rho_0(r, z)$. Thus, we need solve only the simpler problem for the flux from an interface source.

Removing r by a Fourier transformation yields

$$\rho_0^{\dagger}(h, z) = 2\pi \int_{-\infty}^{+\infty} dz' \left[\frac{c(z')\rho_0^{\dagger}(h, z')}{4\pi} + q\delta(z')\right] \times I(|z - z'|), \quad (3)$$

[1] K. M. Case, Ann. Phys. (N.Y.) 9, 1 (1960).
[2] J. P. Elliot, Proc. Roy. Soc. (London) A228, 424 (1955).
[3] B. Davison, *Neutron Transport Theory* (Oxford University Press, New York, 1958).

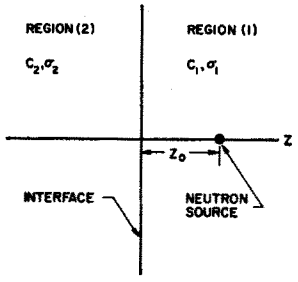


FIG. 1. Geometry and source location.

where

$$\rho_0^{\dagger}(h, z) = \int_{-\infty}^{+\infty}\int e^{ih\cdot r}\rho_0(r, z)\,dx\,dy,$$

$$I(|z - z'|) = \int_1^{\infty} \frac{dv}{v} e^{-v|z-z'|} J_0[h\,|z - z'|\,(v^2 - 1)^{\frac{1}{2}}]. \tag{4}$$

II. SOLUTION OF EQ. (3) BY THE WIENER-HOPF TECHNIQUE

Let us define

$$\rho_+^{\dagger}(h, z) = \begin{cases} \rho_0^{\dagger}(h, z), & z > 0, \\ 0, & z < 0, \end{cases}$$

$$\rho_-^{\dagger}(h, z) = \begin{cases} 0, & z > 0, \\ \rho_0^{\dagger}(h, z), & z < 0. \end{cases}$$

Then Eq. (3) can be written as

$$\rho_+^{\dagger}(h, z) + \rho_-^{\dagger}(h, z) = 2\pi\int_{-\infty}^{+\infty} dz'\left[\frac{c_1\rho_+^{\dagger}(h, z')}{4\pi} + \frac{c_2\rho_-^{\dagger}(h, z')}{4\pi} + q\delta(z')\right] \times I(|z - z'|). \tag{5}$$

Removing the z dependence by another Fourier transformation yields

$$\rho_+^{\dagger\dagger}(h, h_z) + \rho_-^{\dagger\dagger}(h, h_z) = [c_1\rho_+^{\dagger\dagger}(h, h_z) + c_2\rho_-^{\dagger\dagger}(h, h_z) + 4\pi q] \times \left[\frac{1}{2iH}\ln\frac{1 + iH}{1 - iH}\right], \tag{6}$$

where

$$\rho_{\pm}^{\dagger\dagger}(h, h_z) = \int_{-\infty}^{+\infty} e^{ih_z z}\rho_{\pm}^{\dagger}(h, z)\,dz,$$

$$H = (h^2 + h_z^2)^{\frac{1}{2}}.$$

Defining $S = q/(c_1 - c_2)$, and

$$\tau_n(H) = 1 - \frac{c_n}{2iH}\ln\frac{1 + iH}{1 - iH}, \quad n = 1, 2,$$

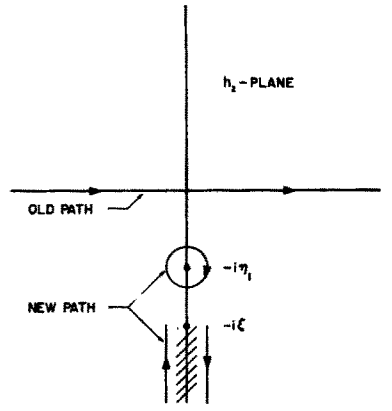


FIG. 2. Inversion path for $z > 0$.

permits us to write for Eq. (6),

$$[\rho_+^{\dagger\dagger}(h, h_z) + 4\pi S]\tau_1(H) = -[\rho_-^{\dagger\dagger}(h, h_z) - 4\pi S]\tau_2(H). \tag{7}$$

Equation (7) is easily solved by the Wiener–Hopf technique. The result is

$$\rho_{\pm}^{\dagger\dagger}(h, h_z) = \pm 4\pi S\{[\tau_{2\pm}(h_z)/\tau_{1\pm}(h_z)] - 1\},$$

where

$$\tau_{n+}(h_z) = \left(\frac{h_z + i\eta_n}{h_z + i\xi}\right)e^{q_{n+}(h_z)},$$

$$\tau_{n-}(h_z) = \left(\frac{h_z - i\xi}{h_z - i\eta_n}\right)e^{q_{n-}(h_z)},$$

$$q_{n\pm}(h_z) = \frac{1}{2\pi i}\int_{-\infty\mp ib}^{+\infty\mp ib} \frac{d\zeta}{\zeta - h_z}\ln\left\{\frac{\tau_n(\zeta)(\zeta^2 + \xi^2)}{(\zeta^2 + \eta_n^2)}\right\}, \quad 0 < b < \xi,$$

$$\xi = (1 + h^2)^{\frac{1}{2}}, \quad \eta_n = (\tau_{0n}^2 + h^2)^{\frac{1}{2}}, \quad \frac{\tau_{0n}}{c_n} = \tanh^{-1}\tau_{0n}.$$

To recover the z dependence, we have

$$\rho_{\pm}^{\dagger}(h, z) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-ih_z z}\rho_{\pm}^{\dagger\dagger}(h, h_z)\,dh_z = \pm 2S\int_{-\infty}^{+\infty} e^{-ih_z z}\left[\frac{\tau_{2\pm}(h_z)}{\tau_{1\pm}(h_z)} - 1\right]dh_z. \tag{8}$$

Equation (8) can be simplified by appropriately deforming the path of integration. For example, when $z > 0$, we change paths as indicated in Fig. 2. Since the only structure in the lower half h_z plane is a simple pole at $h_z = -i\eta_1$, and a branch cut extending from $-i\xi$ to ∞ along the imaginary h_z axis, we can write Eq. (8) for $z > 0$ as

$$\rho_+^{\dagger}(h, z) = 4\pi q\left\{\left(\frac{\eta_2 - \eta_1}{c_1 - c_2}\right)\exp\left[-z\eta_1 + q_{2+}(-i\eta_1) - q_{1+}(-i\eta_1)\right] + 2\int_{\xi}^{\infty} \frac{dt e^{-zt}e^{\{P_1(t) - P_2(t)\}/\pi}(t - \eta_2)(t^2 - h^2)^{\frac{1}{2}}}{(t - \eta_1)\left[\prod_{n=1}^{2}((\pi c_n)^2 + \{2[(t^2 - h^2)^{\frac{1}{2}} - c_n\coth^{-1}(t^2 - h^2)^{\frac{1}{2}}]\}^2)\right]^{\frac{1}{2}}}\right\}, \tag{9}$$

where now

$$q_{n+}(h_z) = -\frac{1}{\pi}\int_{\xi}^{\infty}\frac{ds}{s - ih_z}\,\theta_n(s),$$

$$\tan\theta_n(s) = \frac{\frac{1}{2}\pi c_n}{(s^2 - h^2)^{\frac{1}{2}} - c_n\coth^{-1}(s^2 - h^2)^{\frac{1}{2}}}$$

and

$$P_n(t) = P\int_{\xi}^{\infty}\frac{\theta_n(s)\,ds}{s - t}$$

($P \Rightarrow$ principal value integration).

To obtain the result for $z < 0$ one need only let $z \to -z$ and interchange the (1) and (2) subscripts in Eq. (9) to yield the proper $\rho_{-}^{\dagger}(h, z)$.

We can modify Eq. (9), thereby casting it into a form more closely identified with the Case[1] solutions in one dimension. To do this we change the integration variable in Eq. (9) by setting

$$\nu = (t^2 - h^2)^{-\frac{1}{2}}.$$

This yields

$$\rho_{+}^{\dagger}(h, z) = 4\pi q\left[a(h)e^{-z\eta_1} + \int_0^1 A(h, \nu)e^{-z(1+h^2\nu^2)^{\frac{1}{2}}/\nu}\,d\nu\right], \tag{10}$$

with

$$a(h) = \left(\frac{\eta_2 - \eta_1}{c_1 - c_2}\right)\exp\left[q_{2+}(-i\eta_1) - q_{1+}(-i\eta_1)\right],$$

$$A(h, \nu) = \frac{2}{\nu(1 + h^2\nu^2)^{\frac{1}{2}}}\left[\frac{\eta_2\nu - (1 + h^2\nu^2)^{\frac{1}{2}}}{\eta_1\nu - (1 + h^2\nu^2)^{\frac{1}{2}}}\right] \times \frac{\exp\left(\{P_1[(1 + h^2\nu^2)^{\frac{1}{2}}/\nu] - P_2[(1 + h^2\nu^2)^{\frac{1}{2}}/\nu]\}/\pi\right)}{\left[\prod_{n=1}^{2}((\pi c_n)^2 + \{2[1 - c_n\nu\coth^{-1}(1/\nu)]\}^2)\right]^{\frac{1}{2}}}.$$

A similar form exists for $z < 0$.

III. THE TOTAL FLUX FOR THE INTERFACE SOURCE

With the transformed solutions known, one can write the complete solution for the interface point source as

$$\rho_{\pm}(r, z) = \frac{1}{2\pi}\int_0^{\infty} h\,dh J_0(hr)\rho_{\pm}^{\dagger}(h, z), \tag{11}$$

where $\rho_{\pm}^{\dagger}(h, z)$ is given by either Eq. (9) or the appropriate form of Eq. (10).

By making both half-spaces identical, it is possible to simplify Eq. (11) and arrive at the single infinite medium solution (cf., Ref. 3). Setting $c_2 = 0$ effectively removes the half-space $z < 0$ and in this case the solution, as given by Eq. (11), can be reduced to that found in Ref. 2 for the half-space problem.

IV. SHIFTING THE SOURCE OFF-AXIS[4]

When the source is located in Region (1) ($z_0 > 0$), we can write the counterpart of Eq. (3) as

$$\rho_{z_0}^{\dagger}(h, z) = \int_{-\infty}^{+\infty}\frac{c(z')}{2}\rho_{z_0}^{\dagger}(h, z')I(|z - z'|) + 2\pi qI(|z - z_0|). \tag{12}$$

Making the change of variables $z \to z - \eta$ and differentiating yields

$$\frac{\partial\rho_{z_0-\eta}^{\dagger}(h, z-\eta)}{\partial\eta} = \frac{c_2}{2}\int_{-\infty}^{\eta}\frac{\partial\rho_{z_0-\eta}^{\dagger}(h, z' - \eta)}{\partial\eta}I(|z - z'|)\,dz' + \frac{c_1}{2}\int_{\eta}^{\infty}\frac{\partial\rho_{z_0-\eta}(h, z' - \eta)}{\partial\eta}I(|z - z'|)\,dz' + \frac{c_2 - c_1}{2}\rho_{z_0-\eta}^{\dagger}(h, 0)I(|z - \eta|). \tag{13}$$

But, when $z_0 = 0$, Eq. (12) can be written as

$$\rho_0^{\dagger}(h, z - \eta) = \frac{c_2}{2}\int_{-\infty}^{\eta}\rho_0^{\dagger}(h, z - \eta)I(|z - z'|)\,dz' + \frac{c_1}{2}\int_{\eta}^{\infty}\rho_0^{\dagger}(h, z' - \eta)I(|z - z'|)\,dz' + 2\pi qI(|z - \eta|). \tag{14}$$

Hence,

$$\frac{\partial\rho_{z_0-\eta}^{\dagger}(h, z - \eta)}{\partial\eta} = \left(\frac{c_2 - c_1}{2}\right)\frac{\rho_{z_0-\eta}^{\dagger}(h, 0)\rho_0^{\dagger}(h, z - \eta)}{2\pi q} = \left(\frac{c_2 - c_1}{2}\right)\frac{\rho_0^{\dagger}(h, z_0 - \eta)\rho_0^{\dagger}(h, z - \eta)}{2\pi q}. \tag{15}$$

The last step is a consequence of optical reciprocity.[3] Thus we have the desired relation,

$$\rho_{z_0}^{\dagger}(h, z) = \rho_0(h, |z - z_0|) + \frac{c_1 - c_2}{4\pi q}\int_0^{\min(z, z_0)}\rho_0^{\dagger}(h, z - \eta)\rho_0^{\dagger}(h, z_0 - \eta)\,d\eta. \tag{16}$$

A particularly simple consequence of Eq. (16)—one which also follows from optical reciprocity—is that the interface flux due to an off-interface source is related to the flux from an interface source by

$$\rho_{z_0}^{\dagger}(h, 0) = \rho_0^{\dagger}(h, z_0). \tag{17}$$

Also, one can see that the integration required in Eq. (16) can be easily carried out to yield a $\rho_{z_0}^{\dagger}(h, z)$ which contains discrete, continuous and mixtures of the two types of terms.

[4] The method used in this section follows closely the ideas expounded in Davison, Ref. 3.

V. ASYMPTOTIC EXPANSIONS WITH THE INTERFACE SOURCE

The complicated nature of Eq. (11) can be reduced when either $|z|$ or r is large. For z large and positive, the continuous term in Eq. (9) is small compared to the discrete term and its neglect in Eq. (11) yields the asymptotic estimate

$$\rho_+(r,z) \sim \left(\frac{2q}{c_1-c_2}\right)\left[\frac{(3a_1)^{\frac{1}{2}}}{c_1}\right]\left[\left(\frac{3a_2}{c_2}\right)^{\frac{1}{2}} - \left(\frac{3a_1}{c_1}\right)^{\frac{1}{2}}\right] \times (e^{q_{2+}-q_{1-}})\left(\frac{e^{-(3a_1/c_1)^{\frac{1}{2}}[z+(r^2/2z)]}}{z}\right), \quad z \gg r > 0, \quad a_n < c_n, \tag{18}$$

where $a_n = 1 - c_n$ and the $q_{n\pm}$ are evaluated for $h = 0$.

In the case of weak absorption ($a_n \ll c_n < 1$), Eq. (18) can be further reduced to

$$\rho_+(r,z) \sim \left[\frac{6q}{1+(a_2/a_1)^{\frac{1}{2}}}\right]\left[\frac{e^{-(3a_1)^{\frac{1}{2}}[z+(r^2/2z)]}}{z}\right], \quad z \gg r > 0. \tag{19}$$

Similarly, for $z \ll 0$ and weak absorption, we find

$$\rho_-(r,z) \sim \left[\frac{6q}{1+(a_1/a_2)^{\frac{1}{2}}}\right]\left[\frac{e^{-(3a_2)^{\frac{1}{2}}[|z|+(r^2/2|z|)]}}{|z|}\right], \quad -z \gg r > 0. \tag{20}$$

When there is no absorption in the half-plane $z > 0$, Eq. (11) has a somewhat different expansion for large z. Here we find

$$\rho_+(r,z) \sim \left[\frac{2\sqrt{3}\,q}{(a_2)^{\frac{1}{2}}}\right]\left(\frac{z-\{f(1)-f(c_2)\}}{\{[z-\{f(1)-f(c_2)\}]^2+r^2\}^{\frac{3}{2}}}\right), \quad z \gg r > 0, \quad c_1 = 1, \tag{19a}$$

where $f(k)$ is defined as

$$f(k) = \frac{k}{2}\int_0^1 t\,dt\left[1+\frac{kt^2}{1-t^2}\right] \times \left[\frac{1}{(1-k\tanh^{-1}t)^2+(\frac{1}{2}\pi kt)^2}\right]$$

This $f(k)$ function is similar in structure to the extrapolation distance integral for z_0 found in Ref. 5. The difference of the two $f(k)$ functions in Eq. (19a) can thus be thought of as a difference in two extrapolation distances. For $z \ll 0$, Eq. (20) is still asymptotically correct in this case ($c_1 = 1.0$) provided we set $a_1 = 0$.

For large r, a simple nonzero asymptotic expansion is found only in the case of one nonabsorbing and one absorbing half-space. Assuming $c_1 = 1$, one finds here that

$$\rho_+(r,z) \sim \frac{2q}{r^3}\left[\frac{\sqrt{3}}{(a_2)^{\frac{1}{2}}}\left\{z-[f(1)-f(c_2)]+\frac{1}{\tau_{02}}\right\} - \frac{1}{2}\int_1^\infty \frac{e^{-pz}(p-\tau_{02})e^{(1/\pi)[P_1(p)-P_2(p)]}}{\left[\prod_{n=1}^{2}\{(p-c_n\coth^{-1}p)^2+(\frac{1}{2}\pi c_n)\}\right]^{\frac{1}{2}}}\Bigg|_{h=0}\right] \tag{21}$$

When $a_n \neq 0$ in any of the half-spaces the asymptotic expansions for large r in these regions contain integrals involving Bessel functions which vanish. Hence, no useful asymptotic solution is available in these cases.

VI. COMPARISON WITH APPROXIMATE THEORY

It is possible to solve the two region interface source problem in diffusion theory. Here one has for the interface source,

$$-\tfrac{1}{3}\nabla^2\rho_D(r,z) + a_n\rho_D(r,z) = 4\pi q\delta(x)\delta(y)\delta(z). \tag{22}$$

The solution to Eq. (22) is readily found to be

$$\rho_D(r,z) = 6q\int_0^\infty \frac{h\,dhJ_0(hr)e^{-\alpha_n|z|}}{\alpha_1+\alpha_2} \tag{23}$$

with

$$\alpha_n = \begin{cases}\alpha_1 = (h^2+3a_1)^{\frac{1}{2}}, & z > 0,\\ \alpha_2 = (h^2+3a_2)^{\frac{1}{2}}, & z < 0.\end{cases}$$

Expanding Eq. (23) for large $|z|$, when absorption is present in both half-spaces, yields Eqs. (19) and (20) for $z \gg 0$ and $z \ll 0$, respectively. This is consistent since neutron diffusion theory has, as one of its assumptions, weak absorption. When absorption is absent for $z > 0$, diffusion theory yields the asymptotic form

$$\rho_D(r,z) \sim \frac{2\sqrt{3}\,q}{(a_2)^{\frac{1}{2}}}\frac{z}{(z^2+r^2)^{\frac{3}{2}}}, \quad z \gg 0, \tag{24}$$

which is similar to the exact expansion, Eq. (19a), save for the factor $[f(1)-f(c_2)]$. However, when absorption is made weak in the half-space $z < 0$, this factor tends to zero and the two expressions agree to the first term in their respective series.

Asymptotic expansions for large r in the diffusion theory approximation also have a nonzero value only in the case of no absorption in one of the half-spaces. For the case of $c_1 = 1$, we find here that

$$\rho_D(r,z) \sim \frac{2\sqrt{3}\,q}{(a_2)^{\frac{1}{2}}}\left[\frac{z+(3a_2)^{-\frac{1}{2}}}{r^3}\right], \quad r \gg z > 0. \tag{25}$$

[5] K. M. Case, F. De Hoffmann, and G. Placzek, "Introduction to the Theory of Neutron Diffusion," Los Alamos Scientific Laboratories (1953), Vol. I.

Comparing this with Eq. (21), we see that diffusion theory gives only a discrete term and this term compares favorably with exact theory only when absorption is weak for $z < 0$. The exact result contains an additional continuous term which is expected to contribute strongly to the flux near the interface.

VII. REMARKS

Based on this and previous work, it is now certain that any half-space problems in mono-energetic transport theory with isotropic scattering can be solved exactly if the mean free path is spatially invariant. Since the kernel that results when this last restriction is relaxed is quite complicated, it is not expected that the analytical method used above will be successful in the more general, unrestricted cross-section, problem. An analysis based on the differential form of the Boltzmann equation may be required before this problem can be done exactly. However, one may be able to approximate the solution to the more general problem by following the suggestions given in Ref. 6 on degenerate kernels. Work is now proceeding along this line. Some numerical work is also being done using the results of this paper.

[6] L. V. Kantorovich and V. I. Krylov, *Approximate Methods of Higher Analysis* (Interscience Publishers, Inc., New York, 1958).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 5 MAY 1967

Bounds for Certain Thermodynamic Averages*

A. Brooks Harris
Department of Physics, University of Pennsylvania, Philadelphia, Pennsylvania

(Received 12 November 1966)

Upper and lower bounds for thermodynamic averages of the form $\langle\{A, A^\dagger\}\rangle$ are presented.

THE purpose of this brief note is to present bounds for thermodynamic averages of the form $\langle\{A, A^\dagger\}\rangle$. From the lower bound we can derive a special case of the Bogoliubov[1] inequality. Our lower bound when applied to the one- and two-dimensional isotropic Heisenberg magnet[2] yields the same result as found by Mermin and Wagner.[3] Explicitly our lower bound is

$$\langle\{A, A^\dagger\}\rangle \geq \langle[A, A^\dagger]\rangle \coth(\tfrac{1}{2}\beta\langle\omega\rangle), \tag{1}$$

where $\langle\omega\rangle$ is an average frequency computed from sum rules as described below. An upper bound for $\langle\{A, A^\dagger\}\rangle$ is given by Eq. (19) below.

The derivation of Eq. (1) is elementary. Let n label the eigenstates of $\hat{\mathcal{K}} \equiv \mathcal{K} - \mu N$, where $\mathcal{K}$ is the Hamiltonian, μ the chemical potential, and N the number operator (for simplicity we assume a homogeneous system), so that

$$\hat{\mathcal{K}}\,|n\rangle = \hat{E}_n\,|n\rangle. \tag{2}$$

Using a grand canonical ensemble one has

$$\langle A^\dagger A\rangle = \sum_{mn} W_m \langle m|\, A^\dagger\, |n\rangle\langle n|\, A\, |m\rangle, \tag{3}$$

where

$$W_m = e^{-\beta\hat{E}_m} \Big/ \sum_n e^{-\beta\hat{E}_n}. \tag{4}$$

Equation (3) may be manipulated to give

$$\langle A^\dagger A\rangle = \beta^{-1} \sum_{nm} \frac{W_n - W_m}{\hat{E}_m - \hat{E}_n} |\langle n|\, A\, |m\rangle|^2 \frac{\beta(\hat{E}_m - \hat{E}_n)}{e^{\beta(\hat{E}_m - \hat{E}_n)} - 1}. \tag{5}$$

Define $\varphi(x) = x/(e^x - 1)$, in which case Eq. (5) is of the form

$$\langle A^\dagger A\rangle = \beta^{-1} \sum_i p(x_i)\varphi(\beta x_i), \tag{6}$$

where

$$p(x_i) = \sum_{nm} \frac{W_n - W_m}{\hat{E}_m - \hat{E}_n} |\langle n|\, A\, |m\rangle|^2 \delta_{x_i, \hat{E}_m - \hat{E}_n}, \tag{7}$$

where δ_{x_i,x_j} is a Kronecker delta. From (7) one sees that $p(x)$ is nonnegative and one can easily verify that

* Supported in part by the Advanced Research Projects Agency. This paper is a contribution of the Laboratory for Research on the Structure of Matter, University of Pennsylvania.

[1] N. N. Bogoliubov, Phys. Abhandl. *SU6*, 1, 113, 229 (1962); see also Ref. 3.

[2] The proof that the spontaneous magnetization for these systems vanishes for nonzero temperature is similar to that found in Ref. 3. We obtained the same upper bounds for the low-field magnetization as in Ref. 3.

[3] N. D. Mermin and H. Wagner, Phys. Rev. Letters 17, 1133 (1966).

Comparing this with Eq. (21), we see that diffusion theory gives only a discrete term and this term compares favorably with exact theory only when absorption is weak for $z < 0$. The exact result contains an additional continuous term which is expected to contribute strongly to the flux near the interface.

VII. REMARKS

Based on this and previous work, it is now certain that any half-space problems in mono-energetic transport theory with isotropic scattering can be solved exactly if the mean free path is spatially invariant. Since the kernel that results when this last restriction is relaxed is quite complicated, it is not expected that the analytical method used above will be successful in the more general, unrestricted cross-section, problem. An analysis based on the differential form of the Boltzmann equation may be required before this problem can be done exactly. However, one may be able to approximate the solution to the more general problem by following the suggestions given in Ref. 6 on degenerate kernels. Work is now proceeding along this line. Some numerical work is also being done using the results of this paper.

[6] L. V. Kantorovich and V. I. Krylov, *Approximate Methods of Higher Analysis* (Interscience Publishers, Inc., New York, 1958).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 5 MAY 1967

Bounds for Certain Thermodynamic Averages*

A. Brooks Harris
Department of Physics, University of Pennsylvania, Philadelphia, Pennsylvania

(Received 12 November 1966)

Upper and lower bounds for thermodynamic averages of the form $\langle\{A, A^\dagger\}\rangle$ are presented.

THE purpose of this brief note is to present bounds for thermodynamic averages of the form $\langle\{A, A^\dagger\}\rangle$. From the lower bound we can derive a special case of the Bogoliubov[1] inequality. Our lower bound when applied to the one- and two-dimensional isotropic Heisenberg magnet[2] yields the same result as found by Mermin and Wagner.[3] Explicitly our lower bound is

$$\langle\{A, A^\dagger\}\rangle \geq \langle[A, A^\dagger]\rangle \coth(\tfrac{1}{2}\beta\langle\omega\rangle), \tag{1}$$

where $\langle\omega\rangle$ is an average frequency computed from sum rules as described below. An upper bound for $\langle\{A, A^\dagger\}\rangle$ is given by Eq. (19) below.

The derivation of Eq. (1) is elementary. Let n label the eigenstates of $\hat{\mathcal{K}} \equiv \mathcal{K} - \mu N$, where $\mathcal{K}$ is the Hamiltonian, μ the chemical potential, and N the number operator (for simplicity we assume a homogeneous system), so that

$$\hat{\mathcal{K}}\,|n\rangle = \hat{E}_n\,|n\rangle. \tag{2}$$

Using a grand canonical ensemble one has

$$\langle A^\dagger A\rangle = \sum_{mn} W_m \langle m|\, A^\dagger\, |n\rangle\langle n|\, A\, |m\rangle, \tag{3}$$

where

$$W_m = e^{-\beta\hat{E}_m} \Big/ \sum_n e^{-\beta\hat{E}_n}. \tag{4}$$

Equation (3) may be manipulated to give

$$\langle A^\dagger A\rangle = \beta^{-1} \sum_{nm} \frac{W_n - W_m}{\hat{E}_m - \hat{E}_n} |\langle n|\, A\, |m\rangle|^2 \frac{\beta(\hat{E}_m - \hat{E}_n)}{e^{\beta(\hat{E}_m - \hat{E}_n)} - 1}. \tag{5}$$

Define $\varphi(x) = x/(e^x - 1)$, in which case Eq. (5) is of the form

$$\langle A^\dagger A\rangle = \beta^{-1} \sum_i p(x_i)\varphi(\beta x_i), \tag{6}$$

where

$$p(x_i) = \sum_{nm} \frac{W_n - W_m}{\hat{E}_m - \hat{E}_n} |\langle n|\, A\, |m\rangle|^2 \delta_{x_i, \hat{E}_m - \hat{E}_n}, \tag{7}$$

where δ_{x_i,x_j} is a Kronecker delta. From (7) one sees that $p(x)$ is nonnegative and one can easily verify that

* Supported in part by the Advanced Research Projects Agency. This paper is a contribution of the Laboratory for Research on the Structure of Matter, University of Pennsylvania.

[1] N. N. Bogoliubov, Phys. Abhandl. *SU6*, 1, 113, 229 (1962); see also Ref. 3.

[2] The proof that the spontaneous magnetization for these systems vanishes for nonzero temperature is similar to that found in Ref. 3. We obtained the same upper bounds for the low-field magnetization as in Ref. 3.

[3] N. D. Mermin and H. Wagner, Phys. Rev. Letters 17, 1133 (1966).

$\varphi(x)$ is convex, i.e., $d^2\varphi/dx^2 \geq 0$. Under these conditions one can easily show that

$$\sum_i p(x_i)\varphi(\beta x_i) \geq \varphi(\beta\langle x\rangle)\sum_i p(x_i), \tag{8}$$

where

$$\langle x\rangle = \sum_i p(x_i)x_i \Big/ \sum_i p(x_i). \tag{9}$$

Applying Eq. (8) to the case of Eq. (5), we obtain

$$\langle A^\dagger A\rangle \geq \sum_{nm} \frac{W_n - W_m}{\hat{E}_m - \hat{E}_n} |\langle n|\, A\, |m\rangle|^2 \frac{\langle\omega\rangle}{e^{\beta\langle\omega\rangle} - 1}, \tag{10}$$

where

$$\langle\omega\rangle = \frac{\sum_{nm}(W_n - W_m)\,|\langle n|\, A\, |m\rangle|^2}{\sum_{nm} \frac{W_n - W_m}{\hat{E}_m - \hat{E}_n}\,|\langle n|\, A\, |m\rangle|^2}. \tag{11}$$

Combining Eqs. (10) and (11) we find

$$\langle A^\dagger A\rangle \geq \langle [A, A^\dagger]\rangle/(e^{\beta\langle\omega\rangle} - 1), \tag{12}$$

from which Eq. (1) follows immediately.

It is interesting to note that Eq. (1) can be weakened to give (a) more tractable results and (b) a special case of the Bogoliubov[1] inequality. We note that $\langle\omega\rangle$ as defined by Eq. (11) may not be a convenient quantity since the denominator is not so easy to handle.[4] However, by the Cauchy–Schwarz inequality $\langle\omega\rangle^2 \leq \langle\omega^2\rangle$ so that the inequality (12) remains valid when $\langle\omega\rangle$ is replaced by $\langle\omega^*\rangle \equiv \langle\omega^2\rangle/\langle\omega\rangle$ since $\langle [A, A^\dagger]\rangle$, $\langle\omega\rangle$, and $\langle\omega^*\rangle$ all have the same algebraic sign. This is easily seen by writing

$$\langle\omega\rangle = \sum_i x_i p(x_i) \Big/ \sum_i p(x_i), \tag{13a}$$

$$\langle\omega^*\rangle \equiv \frac{\langle\omega^2\rangle}{\langle\omega\rangle} = \sum_i x_i^2 p(x_i) \Big/ \sum_i x_i p(x_i), \tag{13b}$$

$$\langle [A, A^\dagger]\rangle = \sum_i x_i p(x_i), \tag{13c}$$

and using the nonnegativity of $p(x)$. Equation (13b) can be written as

$$\langle\omega^*\rangle = \frac{\sum_{nm}(W_n - W_m)(\hat{E}_m - \hat{E}_n)\,|\langle n|\, A\, |m\rangle|^2}{\sum_{nm}(W_n - W_m)\,|\langle n|\, A\, |m\rangle|^2}, \tag{14a}$$

or simply as

$$\langle\omega^*\rangle = \frac{\langle [[A, \hat{\mathcal{H}}], A^\dagger]\rangle}{\langle [A, A^\dagger]\rangle} \tag{14b}$$

Thus a weaker but possibly more convenient inequality than (1) is

$$\langle\{A, A^\dagger\}\rangle \geq \langle [A, A^\dagger]\rangle \coth(\tfrac{1}{2}\beta\langle\omega^*\rangle). \tag{15}$$

[4] The denominator in Eq. (12) is rather convenient, but a sharper use of the Cauchy–Schwarz inequality shows that this denominator can be replaced by $\sum_{nm} |W_n - W_m|\,|\langle n|\, A\, |m\rangle|$.

But $|\coth x| \geq |1/x|$ so that a still weaker inequality is

$$\tfrac{1}{2}\langle\{A, A^\dagger\}\rangle \geq kT\langle [A, A^\dagger]\rangle^2/\langle [[A, \hat{\mathcal{H}}], A^\dagger]\rangle, \tag{16}$$

where we have used Eq. (14). The Bogoliubov inequality[1] may be written as

$$\tfrac{1}{2}\langle\{A, A^\dagger\}\rangle \geq kT\,|\langle [C, A]\rangle|^2/\langle [[C^\dagger, \hat{\mathcal{H}}], C]\rangle, \tag{17}$$

which is identical to Eq. (16) for the special case of $C = A^\dagger$. The reason the inequality (15) gives the same upper bound for the low-field magnetization of one- and two-dimension Heisenberg magnets as Mermin and Wagner found[3] using the Bogoliubov[1] inequality is that the dominant contribution to the magnetization comes from low-energy excitations in which

$$\text{limit } \coth(\tfrac{1}{2}\beta\omega) \approx (\tfrac{1}{2}\beta\omega)^{-1},$$

so that nothing is lost in going from Eq. (15) to Eq. (16).

Finally an upper bound for $\langle\{A, A^\dagger\}\rangle$ is obtained as follows. We may write

$$\langle\{A, A^\dagger\}\rangle = \sum_{nm}(W_m + W_n)\,|\langle n|\, A\, |m\rangle|^2 \tag{18a}$$

$$= \frac{2}{\beta}\sum_{nm}\frac{W_n - W_m}{\hat{E}_m - \hat{E}_n}\,|\langle n|\, A\, |m\rangle|^2 \times \left\{\frac{W_n + W_m}{W_n - W_m}\right\}\frac{\beta(\hat{E}_m - \hat{E}_n)}{2}. \tag{18b}$$

But

$$\left\{\frac{W_n + W_m}{W_n - W_m}\right\}\frac{\beta(\hat{E}_m - \hat{E}_n)}{2} = \frac{\beta(\hat{E}_m - \hat{E}_n)}{2}\coth\frac{\beta(\hat{E}_m - \hat{E}_n)}{2} \tag{19a}$$

$$\leq 1 + \tfrac{1}{3}(\tfrac{1}{2}\beta)^2(\hat{E}_m - \hat{E}_n)^2 \tag{19b}$$

so that

$$\langle\{A, A^\dagger\}\rangle \leq \frac{2}{\beta}\sum_{nm}\frac{W_n - W_m}{\hat{E}_m - \hat{E}_n}\,|\langle n|\, A\, |m\rangle|^2 \times \left(1 + \frac{\beta^2}{12}(\hat{E}_m - \hat{E}_n)^2\right) \tag{20a}$$

$$\leq \frac{2}{\beta}\sum_{nm}\frac{W_n - W_m}{\hat{E}_m - \hat{E}_n}\,|\langle n|\, A\, |m\rangle|^2 + \frac{\beta}{6}\langle [[A, \hat{\mathcal{H}}], A^\dagger]\rangle. \tag{20b}$$

Taking $A = [B, \hat{\mathcal{H}}]$ would enable one to eliminate the energy denominators, since $\langle n|\, A\, |m\rangle = \langle n|\, B\, |m\rangle \times (\hat{E}_m - \hat{E}_n)$; however, the commutators are slightly more tedious to evaluate.

ACKNOWLEDGMENT

The author would like to thank Professor J. R. Schrieffer for calling attention to the pre-published work of Mermin and Wagner.

Spinor Fields as Distortions of Space-Time

DAVID HESTENES*
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey

(Received 22 March 1966)

Spinor fields are introduced into Riemannian space-time in a new way. This approach admits a simple geometrical interpretation of spinor fields. A linear connection for space-time is derived which describes both gravitational and nongravitational forces. It is consistent with a straightforward generalization of the Dirac equation. This theory also entails a physical interpretation of inertial coordinates. The spin and current vectors of a spinor field are not in general orthogonal. This lack of orthogonality provides an absolute measure of one component of the gravitational field as seen in an inertial system.

1. INTRODUCTION

In previous papers[1] (hereafter called I and II) spinor fields were expressed entirely in terms of the real Dirac algebra, and the attempt was made to provide a *physical* interpretation for all the invariants of a spinor field and its derivatives. Here we try to develop a coherent *geometrical* interpretation of the same theory. By representing physical quantities in the geometric algebra of space-time and giving these quantities a geometrical interpretation, we hope to establish an isomorphism between physical and geometrical interpretations of the equations of physics. To the extent that the analogy between physics and geometry is completed, a geometrical theory of physics is achieved.

In I we showed that a spinor field ψ determines a map of an inertial frame $\{\gamma^\mu; \mu = 0, 1, 2, 3\}$ into a frame of physical vectors $\{J^\mu\}$:

$$\gamma^\mu \to J^\mu(x) = \psi(x)\gamma^\mu\tilde{\psi}(x) = \rho(x)e^\mu(x). \quad (1.1)$$

The J^μ are bilinear vector invariants of the spinor field ψ. The Dirac theory provides a physical interpretation of J^0 and J^3. To provide a physical interpretation for all the J^μ, in II we interpreted the J^μ as isospin currents. This theory was developed in Minkowski space-time. What changes are necessary for it to hold in space-time curved by gravitational forces? The answer depends on what properties of the γ^μ in (1.1) are physically significant.

In Minkowski space-time the γ^μ are gradients of *inertial* coordinate functions x^μ, i.e.,

$$\gamma^\mu = \Box x^\mu. \quad (1.2)$$

The γ^μ are also orthonormal. In curved space-time the γ^μ cannot have both these properties. If we suppose that the orthonormality property of the γ^μ is fundamental, then in curved space-time the γ^μ cannot be associated with a system of coordinates unless an additional assumption is introduced. The theory of spinor fields and gravitational interactions that follows has been discussed many times in the literature. A treatment using the real Dirac algebra has been given elsewhere (hereafter called STA).[2,3] If, on the other hand, we assume that (1.2) is fundamental, we relate spinor fields to curved space-time in a way which has never before been considered. This is the approach examined in this paper. It has the advantage that it leads to a geometrical interpretation of (1.1) without further assumption. It also leads to a unique determination of inertial coordinates in curved space-time, but we defer discussion of this point until later.

In STA the gravitational field in Einstein's theory is represented by the γ^μ, rather than by the metric tensor $g^{\mu\nu} = \gamma^\mu \cdot \gamma^\nu$. Thus the $\gamma^\mu = \gamma^\mu(x)$ characterize space-time and its distortion by gravitational forces. Now let us reconsider the transformation (1.1). It is composed of a "rigid" Lorentz rotation of the tangent space at each space-time point:

$$\gamma^\mu(x) \to e^\mu(x) = R(x)\gamma^\mu\tilde{R}(x) \quad (1.3)$$

and a change of scale by a factor $\rho(x)$. Clearly we can interpret (1.3) as a mapping from a space-time characterized by the γ^μ to a space-time manifold characterized by the e^μ. In physico-geometrical terms, (1.3) is a distortion of space-time due to the presence of a matter (spinor) field.

We would also like to interpret the scale transformation as a distortion of space-time. But here we come into conflict with the physical interpretation of $\rho(x)$, for, loosely speaking, $\rho(x)$ describes the matter density,

* Present address: Arizona State University, Tempe, Arizona.

[1] D. Hestenes, I, J. Math. Phys. **8**, 798 (1967); II, *ibid.* **8**, 809 (1967). We employ notations and conventions set down in these papers.

[2] D. Hestenes, *Space-Time Algebra* (Gordon and Breach Science Publishers, New York, 1966).

[3] The discussion of spinors in Sec. 24 of STA can be related to Eq. (1.1) by using the methods of I.

so in regions where matter is not present a scale transformation by $\rho(x)$ annihilates vector fields because $\rho(x)$ vanishes there. This is not permissible from a geometric point of view. We cannot interpret $\rho(x)$ as a scale factor unless we "renormalize" scale transformations so that they do not vanish in any region of space–time. Evidently we can do this by requiring that $\rho(x)$ to go to a finite constant where no matter is present. From now on we avoid complications in the interpretation of scale transformations by not talking about them. The Lorentz rotation (1.2) suffices to show how we interpret spinor fields as distortions of space–time.

The geometrical interpretation we have given to ψ leads to a theoretical understanding of why physical observables are bilinear in ψ. We have supposed that nongravitational fields "make themselves felt" geometrically by producing, among other effects, a rotation of the tangent space at each point of space–time. Since Hamilton's time it has been natural to represent rotations by bilinear algebraic quantities. So the bilinearity of the corresponding physical observables follows from assuming an isomorphism between the physical and geometrical interpretations of the spinor fields. Only bilinear functions of ψ have a geometrical interpretation, so only bilinear functions of ψ have a physical significance.

2. TORSION

We have suggested a geometrical interpretation of spinors. From this we find the geometrical significance of the physical equations of motion for a spinor field.

In Riemannian space–time, the directional (covariant) derivatives of the γ^μ can be written[4] as

$$\Box_\alpha \gamma^\mu = L^\mu_{\alpha\beta}\gamma^\beta, \tag{2.1}$$

where[5]

$$L^\mu_{\alpha\beta} = \tfrac{1}{2}g^{\mu\nu}(\partial_\nu g_{\alpha\beta} - \partial_\alpha g_{\nu\beta} - \partial_\beta g_{\nu\alpha}). \tag{2.2}$$

The $g_{\alpha\beta}$ are determined by the condition $g^{\mu\nu}g_{\nu\beta} = \delta^\mu_\beta$. Equation (2.1) determines what in differential geometry is called a *linear connection* for the space–time manifold.

To find the *connection* for the space–time manifold obtained by the distortion (1.3), we merely compute the directional derivatives of the e^μ. By differentiating $R\tilde{R} = 1$ we can show that the directional derivatives of R can always be written in the form

$$\Box_\alpha R = \tfrac{1}{2}\omega_\alpha R, \tag{2.3}$$

where the ω_α are bivectors. It follows that

$$\Box_\alpha e^\mu = L^\mu_{\alpha\beta}e^\beta + \tfrac{1}{2}[\omega_\alpha, e^\mu]. \tag{2.4}$$

It is natural to call the ω_α the *torsion bivectors* of the frame $\{e^\mu\}$. For (1.3) is a kind of space–time twisting, and the "rate" of this twisting, which we call *torsion*,[6] is determined by the ω_μ.

Physically, in the spinor equation (2.3) nongravitational interactions are described by the $L^\mu_{\alpha\beta}$. The distortion of space–time by both gravitational and nongravitational forces is described by Eq. (2.4).

The curvature of space–time can be described by the *curvature bivectors* $C_{\alpha\beta}$

$$[\Box_\alpha, \Box_\beta]e^\mu = \tfrac{1}{2}[C_{\alpha\beta}, e^\mu]. \tag{2.5}$$

By using (2.4), $C_{\alpha\beta}$ can be separated into a gravitational curvature $L_{\alpha\beta}$ and a torsion curvature $\Omega_{\alpha\beta}$.

$$C_{\alpha\beta} = L_{\alpha\beta} + \Omega_{\alpha\beta}, \tag{2.6}$$

$$L_{\alpha\beta} = \tfrac{1}{2}L^\mu_{\alpha\beta\nu}e^\nu e_\mu, \tag{2.7a}$$

$$L^\mu_{\alpha\beta\nu} = \partial_\alpha L^\mu_{\beta\nu} - \partial_\beta L^\mu_{\alpha\nu} + L^\mu_{\beta\rho}L^\rho_{\alpha\nu} - L^\mu_{\alpha\rho}L^\rho_{\beta\nu}, \tag{2.7b}$$

$$\Omega_{\alpha\beta} = \Box_\alpha\omega_\beta - \Box_\beta\omega_\alpha - \tfrac{1}{2}[\omega_\alpha, \omega_\beta]. \tag{2.8}$$

From the $C_{\alpha\beta}$ one can construct equations for the gravitational field in some analogy to Einstein's equations. But the appropriate prescription for this certainly involves physical and geometrical ideas other than those we have already introduced, so we do not attempt to discuss it here. Solution of this problem may well require a perfect understanding of stress–energy–momentum in purely geometrical terms—something that has not yet been achieved.

3. INERTIAL SYSTEMS

The γ^μ are related to a system of coordinates x^μ by (1.2). But these cannot be just any coordinates because the γ^μ are related to the J^μ by (1.1), and the J^μ represent observable currents. We know that in Minkowski space–time the x^μ are called inertial coordinates, so it is reasonable to use the same name in curved space–time. To preserve the physical interpretation of the J^μ, we must limit the allowable x^μ by a mathematical condition. This amounts to a physical definition of inertial coordinates. In this section we find such a definition.

An important physical feature of the J^μ is their relation to conservation laws. To get a convenient expression of this we use the fact that the distortion (1.3) induces a mapping of the gradient operator

$$\Box = \gamma^\mu\Box_\mu \tag{3.1}$$

[4] Equation (20.9) of STA.
[5] Equation (21.4) of STA.
[6] Our use of "torsion" is related but not equivalent to uses of the same word common in differential geometry.

into the gradient operator

$$\square' = e^\mu \square_\mu \tag{3.2}$$

for the distorted space–time manifold. The divergence of the J^μ can be calculated if we know the gradient of the spinor field. According to our discussion in I and II, in all inertial systems the divergence of J^μ must have the form

$$\square' \cdot J^\mu = \rho\pi^\mu. \tag{3.3}$$

A detailed expression for the π^μ in terms of physical fields can be read off from Eq. (2.9) of II.[7] We have summarized this expression to emphasize that the model of strong interactions discussed in II associates the π^μ with components of the pion field. It should be realized, however, that this particular interpretation is not essential to the ideas discussed here.

If we calculate the divergence of the J^μ directly, we get

$$\square' \cdot J^\mu = \rho\pi^\mu + \rho\square \cdot \gamma^\mu. \tag{3.4}$$

To preserve the physical interpretation of the J^μ, (3.4) must agree with (3.3). Therefore, inertial coordinate functions satisfy the equation

$$\square \cdot \gamma^\mu = \square^2 x^\mu = 0. \tag{3.5}$$

Coordinates satisfying (3.5) are called harmonic by V. Fock. Fock suggested long ago that harmonic coordinates should be identified with inertial coordinates in curved space–time.[8] But heretofore there has been no physical consequence of this identification to make it compelling.

As an alternative to (3.5), we can use (3.3) to write the condition for harmonic coordinates in terms of $\square'$:

$$\square' \cdot e^\mu = \square' \cdot \square' x^\mu = \pi^\mu - \square^\mu \ln \rho. \tag{3.6}$$

Thus, inertial coordinates are distinguished by the property that the divergence of the associated metric vector potentials depends only on nongravitational fields.

We have distinguished inertial coordinates by insisting on the conservation (or "partial conservation") laws (3.3) for spinor fields. The physical significance this gives to inertial coordinates can be grandly summarized in the *principle of global relativity*[9]: Conservation laws have the same form and interpretation in all inertial systems. This is a generalization of the principle of relativity introduced by Einstein in his special theory of relativity.

Physical identification of inertial coordinates makes possible an absolute measure of gravitational force, namely, the gravitational force as it appears in an inertial system. To see this, consider a "test particle" with current vectors J^μ. From (1.1)

$$J^\mu \cdot J^\nu = \rho\psi\gamma^\mu \cdot \gamma^\nu\tilde{\psi} = \rho^2 g^{\mu\nu}. \tag{3.7}$$

Thus, if we can measure the J^μ, we can obtain the gravitational potentials as seen in an inertial system. Unfortunately, detection of this effect is well beyond our present-day capacity.

4. COMPARISON WITH OTHER THEORIES

Many attempts have been made to generalize the linear connection of a Riemannian manifold to account for nongravitational forces. None have had any clear-cut success. One of the first and most ingenious theories of this kind was invented by Weyl.[10] Our theory is mathematically related to his, but differs profoundly in physical interpretation. In Weyl's theory the nongravitational interaction is determined entirely by the "divergence vector"

$$a^\mu \equiv \square' \cdot e^\mu = \pi^\mu - \square^\mu \ln \rho. \tag{4.1}$$

Weyl interpreted the π^μ as components of the electromagnetic field, whereas we relate them to components of the pion field. He interpreted the term $\square^\mu \ln \rho$ as arising from a scale transformation, just as we did in Sec. 1. However, Weyl made his theory "independent of gauge" by assuming invariance under arbitrary scale transformations, thus making the $\square^\mu \ln \rho$ term physically unobservable in principle. By contrast, in our theory ρ is a physical quantity describing the presence of matter.

Shortly after the invention of wave mechanics Weyl abandoned his theory because it could not come to terms with the new wave equation for the electron. But what was disastrous for his theory is advantageous to ours. For it is one of the strong points of our theory that the Dirac equation participates in the very definition of the connection of space–time. Other attempts have been made to use the Dirac equation to determine a connection for space–time. But they founder on the fact that $(-1)^{\frac{1}{2}}$ appears explicitly in the Dirac equation. For this fact seems to imply that space–time has a "complex" connection. What can that mean in what appears to be a real world? This problem is resolved in I where $(-1)^{\frac{1}{2}}$ is given a geometrical interpretation. The geometrical $(-1)^{\frac{1}{2}}$ is

[7] Equation (2.9) of II uses $\square$ instead of the $\square'$ used in (3.3). It seems to make little difference which is used, but this point is worthy of more careful study. We have used $\square'$ in (3.3) because it leads to the simplest coordinate condition—(3.5).

[8] For further discussion of harmonic coordinates see Sec. 23 of STA.

[9] Cf. Sec. 23 of STA.

[10] H. Weyl, in *The Principle of Relativity* (Dover Publications, Inc., New York, 1923); *Space, Time and Matter* (Dover Publications, Inc., New York, 1922). An excellent account and criticism of Weyl's theory is given by L. Pauli, *Theory of Relativity* (Pergamon Press, New York, 1958).

a critical ingredient in our theory. Without it we could not have written down even the first equation in this paper.

5. PHYSICS AS GEOMETRY

In this paper we have proposed a geometrical theory of "matter" which is in harmony with Dirac's equation and Einstein's theory of gravitation. Although the theory is incomplete in many respects, the underlying presumptions are simple and definite, and we have not detected any signs that this graft of Dirac's theory to Einstein's will not take. Our method has been to set up a strict analogy between physics and geometry, by expressing physical quantities in the geometric algebra of space–time and by interpreting the resulting algebraic quantities geometrically. To provide direction for further application of this method, it is worthwhile to speculate on how the analogy between physics and geometry might be pursued to completion.

The overriding problem is to assimilate or explain the fundamental notions of quantum theory. We have gone a long way in that direction by interpreting the Dirac equation and the bilinearity of spinor observables in geometric terms. In doing so we have ignored the probability interpretation of wavefunctions. Yet it may be questioned whether the probability interpretation is fundamental. Convenient as it is when applied to the Schrödinger theory, it has many difficulties when applied to relativistic theories, and *it has not led to any predictions*—the supreme scientific test of a useful concept. It has done no more than provide a psychological resolution of our conflicting images of waves and particles. The probability concept certainly does not account for the wide variety of conservation laws found in physics. On the other hand, whenever we have a conservation law, we can introduce a probability density merely by normalization. Thus, among electrons charge is conserved, so by normalizing the charge current density we have a probability current density. We could just as well have derived probability conservation from energy conservation. In fact, when it comes to photons that is what we have to do, because energy is the only conserved physical quantity available. Even then the probability notion applies in momentum space but not in space–time. In every case the probability notion comes in after the fact. It is not claimed that by these remarks we have dispensed with the probability interpretation of wavefunctions. We have merely brought it into question in recognition of the need to reconcile it with the geometrical interpretation presented in this paper.

A more important problem is to give a geometrical account of the discrete nature of physical quantities. Any such account must be in terms of *global geometry*. For in spite of the frequent association of quantum phenomena with the very small, discreteness is a global (or integral) physical property. It arises in many cases because of boundary conditions on the wavefunction. The description of a many-particle system including creation and annihilation of particles seems to be an exception to this rule, and the formalism of second quantization was developed to represent it. But Feynman's boundary condition,[11] which requires that waves of positive (negative) energy propagate only forward (backward) in time, accounts for features involving creation and annihilation of particles without formal quantization. This strongly suggests that a deeper understanding of the global properties of fields may provide a description of particles. To pursue this idea, we must adopt notions of field and particle which correspond to notions of local and global geometry. A field has a *local* aspect because it assigns a definite amplitude at each space–time point. But a particle must be thought of as a *global* property of a field, so that it does not make sense to speak of a particle at a point, but it does make sense to speak of a particle in a box. The box may indeed be very small, but the property of containing a particle is a property of the whole box, not of a point in the box. Thus the notions of field and particle refer respectively to local and global aspects of the same phenomenon.

The geometrical theory of "matter" developed in this paper is entirely local; it deals only with properties of space–time in the neighborhood of a point. Like any field theory it can be quantized. Although this procedure permits a physical description of particles, it conflicts with the general program of constructing an analogy between physics and geometry so long as we cannot provide a geometrical interpretation of quantization. Further, if particles are manifestations of global geometry, then quantization is suspect, because it makes little sense to represent global properties with local operators. If we are to construct a geometrical theory of physics, it appears we must look for the reflection of the wave-particle duality of physics in the local-global duality of geometry. Nevertheless, we do not know enough to say that quantization is not the most appropriate algebraic expression of this duality.[12]

A nontrivial application of global geometry to

---

[11] R. Feynman, Phys. Rev. 76, 749, 769 (1949).

[12] M. Schönberg has suggested that the algebra of quantum field theory does indeed have its origin in topological features of space–time. See especially Sec. 9 in Nuovo Cimento Suppl. 6, 356 (1957).

physics has never been truly successful and has seldom been attempted.[13] But the mathematical possibilities are most promising. In recent years mathematicians have made profound and surprising discoveries in global geometry.[14] Unfortunately, the fact that physicists are generally unfamiliar with the language and methods used in these researches is a big barrier to possible physical application. If the results of modern global geometry are to be used to further develop the theory in this paper, it is necessary to reformulate them in the language of Clifford algebra. This should be a straightforward and enlightening task. If it has any relevance to physics, we can be assured that the relationship is profound.

[13] Noteworthy is the use of de Rham's theorem by C. W. Misner and J. A. Wheeler to account for electric charge, Ann. Phys. (N.Y.) **2**, 525 (1957); reprinted in *Geometrodynamics* (Academic Press Inc., New York, 1962).

[14] I am sorry to say that this statement is based more on hearsay than it is on familiarity with the discoveries themselves. Having only one foot over the barrier separating physical and mathematical geometers, I can give only an inadequate list of references, especially in regard to the most recent work. I mention only de Rham's theorem (already cited in Ref. 11), and generalizations of the Gauss–Bonnet formula. See S. Chern, Acad. Brazil. Ciencias. Anais **35**, 17 (1963), and A. Avez, Compt. Rend. **255**, 2049 (1962).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 5 MAY 1967

Regularization and Peratization of Singular Potentials

WILLIAM A. GALE*
Rice University, Houston, Texas

(Received 28 September 1966)

We prove that the commonly used regularizations for singular potentials are successful. This means that one can investigate the peratization properties with confidence. A general argument for the success of peratization, as an approximation procedure, is presented. The class of failures of peratization is extended to a series of arbitrarily weakly singular perturbations on the inverse fourth potential. The two results present an unresolved contradiction; some resolutions are considered.

I. INTRODUCTION

THE technique of peratization is designed to give meaning to a series in which each term is a divergent function of some parameter. This approximation has been studied in the calculation of the scattering lengths of singular potentials. We restrict the problem in several ways. The potential must have its leading singularity repulsive. At infinite distance we require that the potential fall off more rapidly than inverse cubic. We consider only the zero-energy problem. Due to the bound states which exist for any negative coupling constant, the radius of convergence of expansions in the coupling constant must be zero. To overcome this a *regulated* potential is introduced. A regulated potential $V(g, r, \alpha)$ satisfies

$$\begin{aligned} &(1)\ V(g, r, \alpha) \text{ is nonsingular for } \alpha > 0, \\ &(2)\ V(g, r, 0) = V(g, r). \end{aligned} \tag{1}$$

The most common regularizations are

$$V(g, r, \alpha) = \vartheta(r - \alpha)V(g, r), \tag{2}$$

$$V(g, r, \alpha) = V(g, r, +\alpha). \tag{3}$$

We prove in the following section that the two common regularizations are successful. We then present a general argument for peratization, with a specific example, and a class of counterexamples.

II. REGULARIZATION

Since $V(g, r, \alpha)$ is nonsingular for each $\alpha > 0$, there is for each α, an analytic expression for the scattering length $A(\alpha, g) = \sum a_n(\alpha)g^n$. It may be possible to sum the series to obtain $A(\alpha, g)$. If now $\lim A(\alpha, g) = A(g)$, then the process of regularization has succeeded. The common regularizations have not been known to fail, but there have been no adequate theorems of sufficient conditions or necessary conditions.[1–3] Some regularizations have been invented which fail.[2] We consider the regularization of (3) first.

We write the solutions to the unregulated Schrödinger equation as $\psi_r(r) = r\Phi_r(r)$, the regular solution

* National Science Foundation Trainee in physics.

[1] N. N. Khuri and A. Pais, Rev. Mod. Phys. **36**, 590 (1964).
[2] F. Calogero, Phys. Rev. **139**, B602 (1965).
[3] M. Cornille, Nuovo Cimento **38**, 1243 (1965).

physics has never been truly successful and has seldom been attempted.[13] But the mathematical possibilities are most promising. In recent years mathematicians have made profound and surprising discoveries in global geometry.[14] Unfortunately, the fact that physicists are generally unfamiliar with the language and methods used in these researches is a big barrier to possible physical application. If the results of modern global geometry are to be used to further develop the theory in this paper, it is necessary to reformulate them in the language of Clifford algebra. This should be a straightforward and enlightening task. If it has any relevance to physics, we can be assured that the relationship is profound.

[13] Noteworthy is the use of de Rham's theorem by C. W. Misner and J. A. Wheeler to account for electric charge, Ann. Phys. (N.Y.) **2**, 525 (1957); reprinted in *Geometrodynamics* (Academic Press Inc., New York, 1962).

[14] I am sorry to say that this statement is based more on hearsay than it is on familiarity with the discoveries themselves. Having only one foot over the barrier separating physical and mathematical geometers, I can give only an inadequate list of references, especially in regard to the most recent work. I mention only de Rham's theorem (already cited in Ref. 11), and generalizations of the Gauss–Bonnet formula. See S. Chern, Acad. Brazil. Ciencias. Anais **35**, 17 (1963), and A. Avez, Compt. Rend. **255**, 2049 (1962).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 5 MAY 1967

Regularization and Peratization of Singular Potentials

WILLIAM A. GALE*
Rice University, Houston, Texas

(Received 28 September 1966)

We prove that the commonly used regularizations for singular potentials are successful. This means that one can investigate the peratization properties with confidence. A general argument for the success of peratization, as an approximation procedure, is presented. The class of failures of peratization is extended to a series of arbitrarily weakly singular perturbations on the inverse fourth potential. The two results present an unresolved contradiction; some resolutions are considered.

I. INTRODUCTION

THE technique of peratization is designed to give meaning to a series in which each term is a divergent function of some parameter. This approximation has been studied in the calculation of the scattering lengths of singular potentials. We restrict the problem in several ways. The potential must have its leading singularity repulsive. At infinite distance we require that the potential fall off more rapidly than inverse cubic. We consider only the zero-energy problem. Due to the bound states which exist for any negative coupling constant, the radius of convergence of expansions in the coupling constant must be zero. To overcome this a *regulated* potential is introduced. A regulated potential $V(g, r, \alpha)$ satisfies

$$\begin{aligned} &(1)\ V(g, r, \alpha) \text{ is nonsingular for } \alpha > 0, \\ &(2)\ V(g, r, 0) = V(g, r). \end{aligned} \tag{1}$$

The most common regularizations are

$$V(g, r, \alpha) = \vartheta(r - \alpha)V(g, r), \tag{2}$$

$$V(g, r, \alpha) = V(g, r, +\alpha). \tag{3}$$

We prove in the following section that the two common regularizations are successful. We then present a general argument for peratization, with a specific example, and a class of counterexamples.

II. REGULARIZATION

Since $V(g, r, \alpha)$ is nonsingular for each $\alpha > 0$, there is for each α, an analytic expression for the scattering length $A(\alpha, g) = \sum a_n(\alpha)g^n$. It may be possible to sum the series to obtain $A(\alpha, g)$. If now $\lim A(\alpha, g) = A(g)$, then the process of regularization has succeeded. The common regularizations have not been known to fail, but there have been no adequate theorems of sufficient conditions or necessary conditions.[1–3] Some regularizations have been invented which fail.[2] We consider the regularization of (3) first.

We write the solutions to the unregulated Schrödinger equation as $\psi_r(r) = r\Phi_r(r)$, the regular solution

* National Science Foundation Trainee in physics.

[1] N. N. Khuri and A. Pais, Rev. Mod. Phys. **36**, 590 (1964).
[2] F. Calogero, Phys. Rev. **139**, B602 (1965).
[3] M. Cornille, Nuovo Cimento **38**, 1243 (1965).

and $\psi_s(r) = r\Phi_s(r)$, the singular solution with

$$\lim_{r\to 0} \frac{\psi_r(r)}{\psi_s(r)} = 0. \tag{4}$$

We know that the asymptotic behavior of the solution is $\psi_r(r) \backsim_\infty N\cdot(r + A)$, hence there is an expansion about infinity and

$$A = \Phi_r'(\infty)/\Phi_r(\infty). \tag{5}$$

We then turn to the regulated problem which is solved

$$\psi(r, \alpha) = \psi_s(\alpha)\psi_r(r + \alpha) - \psi_r(\alpha)\psi_s(r + \alpha) \tag{6}$$

$$\backsim_\infty (r + \alpha)[\psi_s(\alpha)\Phi_r(\infty) - \psi_r(\alpha)\Phi_s(\infty)] + [\psi_s(\alpha)\Phi_r'(\infty) - \psi_r(\alpha)\Phi_s'(\infty)]. \tag{7}$$

Since the coefficient of r must be finite, and $\Phi_r(\infty)$ is defined, $\Phi_s(\infty)$ is also defined; and again since the scattering length exists for the regulated potential, $\Phi_s'(\infty)$ exists. Hence

$$A(\alpha) = \frac{[\psi_s(\alpha)/\psi_r(\alpha)]\Phi_r'(\infty) - \Phi_s'(\infty)}{[\psi_s(\alpha)/\psi_r(\alpha)]\Phi_r(\infty) - \Phi_s(\infty)} + \alpha. \tag{8}$$

Then as $\alpha \to 0$ the correct scattering length is regained; hence the regularization procedure $V(g, r, \alpha) = V(g, r + \alpha)$ has succeeded for any singular potential. In a similar fashion, we obtain for the regularization (2)

$$A(\alpha) = \frac{\Phi_s'(\alpha)\Phi_r'(\infty) - \Phi_r'(\alpha)\Phi_s'(\infty)}{\Phi_s'(\alpha)\Phi_r(\infty) - \Phi_r'(\alpha)\Phi(\infty)}. \tag{9}$$

Now if one function is more singular than another, its derivative will be more singular than that of the other. Thus the ϑ regularization also works.

III. PERATIZATION

For a series, each of whose terms diverges as a function of a parameter, the first peratization approximation is made by summing the leading singularity as a function of α in each order of g, then taking the limit of the resulting expression. The second peratization approximation is made by summing the two greatest singularities, etc. We term a series, each of whose terms diverges as a function of a parameter, *peratizable* if in each order of approximation, the sum is finite, and the approximation is improved.

The following theorem can be proved: The sum or product of an analytic function with the peratizable representation of another function is again peratizable. The same result no longer applies in general if both functions are in a peratizable representation. Conditions on the orders of the singularities can be set up which allow this, in particular if the singularity structure is the same, then the sums and products of two peratizable functions are peratizable. We now extend the argument given previously for regularization.

We recall Eq. (8) and suppose that the sum of the leading singularities in each order for $[\psi_s(\alpha)/\psi_r(\alpha)]$ diverges as $\alpha \to 0$ (which, considering its behavior as $\alpha \to 0$, means that it can be peratized). If we perform all the multiplications and additions before taking leading singularities, then some of the nonsingular terms of Φ_r' or Φ_r may be neglected in their order of g. As long as only a finite number of singular terms occurs in each order of g, we recover this information, in successive approximations. We expect to be able to divide the two peratizable functions in the numerator and denominator because of their similarity of singularities, though after multiplication they may no longer be associated with the same order of g.

A number of examples of solvable cases have been given previously[4-6]; in each case the argument can be followed explicitly. As a definite example we consider here the potential $g(r^{-5} + \phi r^{-4})$, which is similar to the case studied by Pais and Wu,[6] but solvable in readily visualized functions:

$$\psi(r) = r\,\mathrm{Ai}\,(g^{\frac{1}{3}}(r^{-1} + \phi)), \tag{10}$$

where Ai and Bi are independent solutions of the Airy equation. The scattering length is

$$A(g, \phi) = g^{\frac{1}{3}}\,\mathrm{Ai}'\,(g^{\frac{1}{3}}\phi)/\mathrm{Ai}\,(g^{\frac{1}{3}}\phi). \tag{11}$$

The scattering length for the ϑ regulated potential is

$$A(g, \phi, \alpha) = \gamma \frac{\mathrm{Bi}'\,(\gamma(\alpha^{-1} + \phi))\,\mathrm{Ai}'\,(\gamma\phi) - \mathrm{Ai}'\,(\gamma(\alpha^{-1} + \phi))\,\mathrm{Bi}'\,(\gamma\phi)}{\mathrm{Bi}'\,(\gamma(\alpha^{-1} + \phi))\,\mathrm{Ai}\,(\gamma\phi) - \mathrm{Ai}'\,(\gamma(\alpha^{-1} + \phi))\,\mathrm{Bi}\,(\gamma\phi)}, \tag{12}$$

where $\gamma = |g^{\frac{1}{3}}|$. As $\alpha \to 0$, $A(g, \phi, \alpha) \to A(g, \phi)$. The summation of the leading singularities of the directly calculated Born series gives

$$\gamma \frac{\mathrm{Bi}'\,(\gamma\alpha^{-1})\,\mathrm{Ai}'\,(0) - \mathrm{Ai}'\,(\gamma\alpha^{-1})\,\mathrm{Bi}'\,(0)}{\mathrm{Bi}'\,(\gamma\alpha^{-1})\,\mathrm{Ai}\,(0) - \mathrm{Ai}'\,(\gamma\alpha^{-1})\,\mathrm{Bi}\,(0)}, \tag{13}$$

as can be most readily checked by Eqs. (10.4.2) of Ref. 7. This gives a limit of $\gamma\,\mathrm{Ai}'\,(0)/\mathrm{Ai}\,(0)$, the scattering length of gr^{-5}.

The answer is no longer exact; however, for small ϕ, positive or negative, it is excellent approximation.

[4] F. Calogero and M. Cassandro, Nuovo Cimento 37, 760 (1965).
[5] W. A. Gale, J. Math. Phys. 7, 2171 (1966).
[6] A. Pais and T. T. Wu, J. Math. Phys. 5, 799 (1964).
[7] *Handbook of Mathematical Functions*, M. Abramowitz and L. A. Stegun, Eds. (U.S. Government Printing Office, Washington D.C., 1964).

The summation of the second terms gives a term linear in ϕ, or the slope of A near $\phi = 0$ which improves the approximation. The region in which the power series in ϕ for A converge is the circle about the origin with the radius equal to the distance of the nearest singularity: $|\phi| < g^{\frac{1}{3}}\,2.338$. We see that this is exactly the nearest zero energy bound state.

Calogero and Cassandro[4] considered a potential with an exponential singularity. Their results can be similarly interpreted.

The evidence is not so complete, nor so convincing for the weak singularities multiplied by an inverse power.[8,9] We examine the Born series of $gr^{-4}\varphi(r)$, where $\varphi(+\epsilon) > 0$, $r^{\pm 1}\varphi(r) \underset{0}{\backsim}_{\infty} r^{\pm 1}$, and $r\varphi'(r)$ is less singular than $\varphi(r)$. Examples of such functions are $[\ln (1/r)]^{\beta}$, $\beta > 1$; $\ln [\ln (r+1)]$; $\ln \{\ln [\ln (r+e)]\}$; giving a series of extremely weak singularities. The leading singularities of the ϑ regulated Born series are

$$-A(\alpha) = g\frac{\varphi(\alpha)}{\alpha} - \tfrac{1}{3}g^2\frac{\varphi^2(\alpha)}{\alpha^3} + \tfrac{2}{15}g^3\frac{\varphi^3(\alpha)}{\alpha^5} - \cdots . \tag{14}$$

This is established by one integration by parts, the residual integral being less singular than the product removed. Thus

$$A(\alpha) = -g^{\frac{1}{2}}\varphi^{\frac{1}{2}}(\alpha)\tanh\left[(g\varphi(\alpha)/\alpha)^{\frac{1}{2}}\right]\backsim_0 - g^{\frac{1}{2}}\varphi^{\frac{1}{2}}(\alpha), \tag{15}$$

which lacks a limit in the cases cited.

Unfortunately the cases cited are not solvable so that the breakdown of the general argument presented cannot be localized. We note with Cornille[3] that the expansion above for the hyperbolic tangent does not have an infinite radius of convergence, and the continuation assumed may well be incorrect for small α. Secondly, the essential quality of $\varphi(r)$ above is the weakness of its singularity; Wu[9] has suggested a reinterpretation of the cutoff parameter in one such case, and the suggestion is easily generalized for such weak perturbations.

IV. CONCLUSION

The success of the peratization program assumes the previous success of the regularization process. We have proved that the two common regularizations are successful.

Building on this result we have shown why we expect the peratization to be successful. We have, on the other hand, extended the class of potentials for which peratization appears to fail. These are a series of arbitrarily weak singular perturbations on an inverse fourth power potential.

We expect the peratization program to be valid as an approximation procedure at least for potentials without singularities weaker than a pole. In the exceptional cases cited, an additional prescription due to Wu gives an approximate answer.

ACKNOWLEDGMENTS

It is a pleasure to thank Professor I. M. Duck for discussion and initial direction toward this problem. This paper is based on part of the author's Master's thesis.

This work was supported in part by the U.S. Atomic Energy Commission.

[8] H. H. Aly, Riazuddin, and A. H. Zimerman, Phys. Rev. 136, B1174 (1964).
[9] T. T. Wu, Phys. Rev. 136, B1176 (1964).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 5 MAY 1967

Radiation of a Point Charge Moving Uniformly over an Infinite Array of Conducting Half-Planes

JOHN LAM
California Institute of Technology, Pasadena, California

(Received 31 October 1966)

The excitation of an infinite array of parallel semi-infinite metallic plates by a uniformly moving point charge is studied by the Wiener–Hopf method. The problem is treated as a boundary-value problem for the potential of the induced electromagnetic field, and is formulated in terms of a dual integral equation for the current density induced on the plates. The solution of the dual integral equation gives exact expressions for the induced current density and the induced field in the form of Fourier integrals. The Poynting vector is calculated, and the radiation shows that the array of plates behaves both like a diffraction grating and a series of parallel-plate waveguides.

I. INTRODUCTION

THE problem of the excitation of a conductor by a moving point charge has considerable practical importance. In many cases the radiations emitted during the process supply valuable information on the motion of the point charge. On the other hand, this process can be used as a means of generating electromagnetic radiations. Of special interest are conductors possessing periodic geometries, since in such cases the emitted radiations show regular characteristics reflecting this periodicity.

In this work we take the conductor to be made up of an infinite array of conducting half-planes spaced evenly with a constant separation d, so that if we set up a rectangular coordinate system the half-planes, or plates, can all be located in the lower half-space $y \leq 0$ at $x = 0, \pm d, \pm 2d, \cdots$. The edges of the plates all lie in the z–x plane, parallel to the z axis. To simplify the problem we consider the plates to be infinitely thin and perfectly conducting. Such idealizations are, of course, not realizable in the laboratory, and are thus valid only under restricted conditions. In this case if we resolve the electromagnetic field into its frequency components, the assumptions of zero thickness and perfect conductivity hold if $\lambda \gg T \gg \delta$, where T is the thickness of the plates and δ is the skin depth corresponding to the wavelength λ.

The point charge carrying a charge e moves in the upper half-space $y \geq 0$ with constant velocity v in the positive x direction. Its trajectory lies at a distance a above the x axis. We do not consider a nonzero y component of the velocity, since this brings in the further complication of transition radiations emitted when the charge pierces the plates. Also the z component of the velocity is taken to be zero, since its effect is merely a uniform shifting of the whole situation in the z direction. The assumption of uniform motion of the point charge implies either that the constant velocity is maintained by an external agent, or that the interaction between the point charge and the plates does not alter the motion of the former appreciably. The second statement is clearly not true; for since the plates are infinite in number, any effect of the force exerted by one plate on the point charge, however small, will multiply indefinitely as the charge traverses the plates in succession. Nevertheless, in the laboratory we can only construct a finite array of plates to which the infinite array in our calculations is a convenient approximation. So in practice when the transition time of the point charge across the finite array of plates is short, the velocity of the point charge may well be considered uniform.

It is well known that a point charge moving uniformly in vacuum does not radiate. The radiation is therefore attributed entirely to the current density induced on the plates by the passage of the charge. Clearly, our problem is solved if we can calculate this induced current density. In the following the induced current density is found to satisfy a dual integral equation which can be solved by the Wiener–Hopf method. In the 1940's the Wiener–Hopf method was introduced by Copson[1] and Schwinger into the study of electromagnetic diffraction by semi-infinite metallic structures. This method was immediately used by several authors[2] in a series of papers to obtain rigorous solutions to the problems of the excitation of an infinite set of metallic plates by monochromatic plane waves. The polarizations of the plane waves were so chosen that the essentially vectorial problems became two-dimensional and thereby scalarizable.

The problem we are going to solve is really an

[1] E. J. Copson, Quart. J. Math. 17, 19 (1946).
[2] J. F. Carlson, A. E. Heins, Quart. Appl. Math. 4, 313 (1947); 5, 82 (1947); A. E. Heins, *ibid.* 8, 281 (1950).

electromagnetic boundary-value problem. The boundary conditions to be satisfied are three in number. Firstly, there is the outgoing wave condition which requires the solution to assume the form of an outgoing wave at large distances from the sources. Secondly, there is the well-known condition on the electromagnetic field at the surface of a perfect conductor. Finally, there is the edge condition which is peculiar to problems involving surfaces with sharp edges.[3] It prescribes definite behavior on the induced quantities in the vicinity of the edges, and is intimately connected with the uniqueness of the solution. In the following the first two conditions are directly applied in the formulation of the problem. The edge condition is found to be automatically satisfied by the solution obtained.

II. FORMULATION OF THE PROBLEM

Without loss of generality we can take the point charge to be at the point $x = 0$, $y = a$, $z = 0$ at $t = 0$. Then its motion produces a current density

$$j_\mu^0(x, y, z, t) = ec\delta(x - vt)\delta(y - a)\delta(z) \quad (\beta, 0, 0, i), \tag{1}$$

where $\beta = v/c$ and where the superscript in $j_\mu^0(x, y, z, t)$ designates quantities associated with the point charge. It is now easy to calculate the Fourier transform

$$j_\mu^0(x, y, k_z, \omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} j_\mu^0(x, y, z, t)e^{-ik_z z + i\omega t}\,dz\,dt$$
$$= \frac{e}{2\pi\beta}e^{i(\omega/v)x}\delta(y - a) \quad (\beta, 0, 0, i). \tag{2}$$

This current density gives rise to a 4-potential which is the particular solution of the equation

$$(\partial^2/\partial x^2 + \partial^2/\partial y^2 + p^2)A_\mu^0(x, y, k_z, \omega) = -\mu_0 j_\mu^0(x, y, k_z, \omega), \tag{3}$$

where $p^2 = k^2 - k_z^2$ and $k = \omega/c$. The solution is given by

$$A_\mu^0(x, y, k_z, \omega) = \frac{\mu_0}{4\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} i\pi \times H_0^{(1)}\{p[(x - x')^2 + (y - y')^2]^{\frac{1}{2}}\} \times j_\mu^0(x', y', k_z, \omega)\,dx'\,dy', \tag{4}$$

where $i\pi H_0^{(1)}$ is the outgoing Green's function of the two-dimensional Helmholtz equation. Here we must take

$$p = (k^2 - k_z^2)^{\frac{1}{2}}, \quad k^2 \geq k_z^2,$$
$$= i(k_z^2 - k^2)^{\frac{1}{2}}, \quad k_z^2 \geq k^2.$$

[3] J. Meixner, Ann. Physik 6, 2 (1949).

Substituting (2) into (4) and using the relation

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} i\pi H_0^{(1)}[p(x^2 + y^2)^{\frac{1}{2}}]e^{itx}\,dx = \frac{e^{-|y|(-p^2+t^2)^{\frac{1}{2}}}}{(-p^2 + t^2)^{\frac{1}{2}}} \tag{5}$$

we get

$$A_\mu^0(x, y, k_z, \omega) = (\mu_0/4\pi)(e/q\beta)e^{i(\omega/v)x - q|y-a|} \quad (\beta, 0, 0, i), \tag{6}$$

where

$$q = [-p^2 + (\omega/v)^2]^{\frac{1}{2}}.$$

Let us denote the induced current density by $j_\mu(x, y, z, t)$ and the induced potential by $A_\mu(x, y, z, t)$. Then their Fourier transforms are connected by an equation analogous to (4). $j_\mu(x, y, k_z, \omega)$ can be analyzed into a sum of surface current densities

$$j_\mu(x, y, k_z, \omega) = \sum_{n=-\infty}^{\infty}\delta(x - nd)K_{n\mu}(y, k_z, \omega), \tag{7}$$

where $K_{n\mu}(y, k_z, \omega)$ is the surface current density induced on the plate at $x = nd$. The periodicity of our system yields the relation

$$K_{n\mu}(y, z, nd/v) = K_{0\mu}(y, z, 0).$$

In terms of the Fourier transforms, this takes the form

$$K_{n\mu}(y, k_z, \omega) = e^{i(\omega/v)nd}K_{0\mu}(y, k_z, \omega). \tag{8}$$

Thus all the surface current densities differ from that on the plate at $x = 0$ by a phase factor only. Substituting (7) and (8) into (4) and writing

$$K_{0\mu}(y, k_z, \omega) = \frac{1}{(2\pi)^{\frac{1}{2}}}\int_{-\infty}^{\infty} K_{0\mu}(k_y, k_z, \omega)e^{ik_y y}\,dk_y$$

and using (5) we get

$$A_\mu(x, y, k_z, \omega) = \frac{\mu_0}{4\pi}(2\pi)^{\frac{1}{2}}i\int_{-\infty}^{\infty}\sum_{n=-\infty}^{\infty} e^{i(\omega/v)nd + iw|x-nd|} \times \frac{K_{0\mu}(k_y, k_z, \omega)}{w}e^{ik_y y}\,dk_y,$$

where

$$w = (p^2 - k_y^2)^{\frac{1}{2}}, \quad p^2 \geq k_y^2,$$
$$= i(k_y^2 - p^2)^{\frac{1}{2}}, \quad k_y^2 \geq p^2.$$

Carrying out the summation we finally obtain

$$A_\mu(x, y, k_z, \omega) = -\frac{\mu_0}{4\pi}(2\pi)^{\frac{1}{2}}e^{i(\omega/v)md} \times \int_{-\infty}^{\infty}\frac{\sin w[x - (m + 1)d] - e^{i(\omega/v)d}\sin w[x - md]}{\cos wd - \cos(\omega/v)d} \times \frac{K_{0\mu}(k_y, k_z, \omega)}{w}e^{ik_y y}\,dk_y, \tag{9}$$

where m is an integer such that $md < x < (m + 1)d$.

If we set $x = md$ in (9), we get an integral representation for the boundary value of the induced potential on the plate at $x = md$:

$$A_\mu(md, y, k_z, \omega) = \frac{\mu_0}{4\pi}(2\pi)^{\frac{1}{2}} e^{i(\omega/v)md} \times \int_{-\infty}^{\infty} \frac{\sin wd}{w} \frac{1}{\cos wd - \cos(\omega/v)d} \times K_{0\mu}(k_y, k_z, \omega) e^{ik_y y} dk_y. \quad (10)$$

We can set up an integral equation for the unknown induced surface current density $K_{0\mu}(k_y, k_z, \omega)$ if we can find a relation between the potentials in (6) and (10). This is furnished by the boundary conditions on the total electromagnetic field on the plates at $x = md$, $y \leq 0$:

$$\begin{aligned} E_y + E_y^0 &= 0, \\ H_z + E_z^0 &= 0, \\ H_x + H_x^0 &= 0. \end{aligned} \quad (11)$$

These conditions can be translated into conditions on the potentials[4] by virtue of the relations

$$\begin{aligned} \mathbf{E} &= -\nabla\phi - \partial\mathbf{A}/\partial t, \\ \mathbf{B} &= \nabla \times \mathbf{A}, \end{aligned} \quad (12)$$

$$\nabla \cdot \mathbf{A} + (1/c^2)(\partial\phi/\partial t) = 0. \quad (13)$$

Substituting (12) into (11) and using (13) we find that the field boundary conditions in (11) transform into a set of second-order partial differential equations for the induced potential, one for each component. In particular, for the scalar potential, we find

$$(\partial^2/\partial y^2 + p^2)\phi(md, y, k_z, \omega) = (-\partial^2/\partial y^2 + k_z^2)\phi^0(md, y, k_z, \omega). \quad (14)$$

The other equations are similar in form. Using (6) we get the general solution for (14)

$$\phi(md, y, k_z, \omega) = e^{i(\omega/v)md}\left[be^{-ipy} - \frac{\mu_0}{4\pi}\frac{ec(1-\beta^2)}{q\beta}e^{q(y-a)}\right]. \quad (15)$$

Here the constant of integration in the complementary solution has been factorized so that $\phi(md, y, k_z, \omega)$ satisfies a periodicity condition analogous to (8), and b is an unknown function of k_z and ω to be determined later on. We have discarded a term proportional to e^{ipy} since it violates the outgoing wave condition. The boundary values of the other components of the induced potentials are

$$\begin{aligned} A_y(md, y, k_z, \omega) &= e^{i(\omega/v)md}\left[-b\frac{p}{\omega}e^{-ipy} - \frac{\mu_0}{4\pi}\frac{iec\beta}{\omega}e^{q(y-a)}\right], \\ A_z(md, y, k_z, \omega) &= e^{i(\omega/v)md}\left[b\frac{k_z}{\omega}e^{-ipy} + \frac{\mu_0}{4\pi}\frac{ec\beta k_z}{q\omega}e^{q(y-a)}\right], \end{aligned} \quad (16)$$

$A_x(md, y, k_z, \omega)$ is clearly identically zero.

Equating (10) to (15) and (16) we get a set of integral equations of the form

$$\int_{-\infty}^{\infty} K(k_y) K_{0\mu}(k_y, k_z, \omega) e^{ik_y y} dk_y = B_\mu e^{-ipy} + C_\mu e^{qy}, \quad y \leq 0, \quad (17)$$

where the kernel is defined by

$$K(k_y) = (\sin wd/w)[\cos wd - \cos(\omega/v)d]^{-1}. \quad (18)$$

For $y \geq 0$ the unknowns $K_{0\mu}(k_y, k_z, \omega)$ satisfy another set of equations

$$K_{0\mu}(y, k_z, \omega) = 0$$

or, equivalently,

$$\int_{-\infty}^{\infty} K_{0\mu}(k_y, k_z, \omega) e^{ik_y y} dk_y = 0, \quad y \geq 0. \quad (19)$$

Equations (17) and (19) constitute a dual integral equation soluble by the Wiener–Hopf method.

III. SOLUTION OF THE EQUATIONS

Equation (17) can easily be converted into a functional equation of the Wiener–Hopf type by Fourier transformation

$$K(k_y)f(k_y) = h(k_y) + \frac{1}{2\pi}\frac{iB}{k_y + p + i\epsilon} + \frac{1}{2\pi}\frac{iC}{k_y + iq}, \quad (20)$$

where $f(k_y)$ represents any component of $K_{0\mu}(k_y, k_z, \omega)$ and

$$h(k_y) = \frac{1}{2\pi}\int_0^{\infty} H(y) e^{-ik_y y} dy \quad (21)$$

with

$$H(y) = \int_{-\infty}^{\infty} K(k_y) f(k_y) e^{ik_y y} dk_y, \quad y \geq 0. \quad (22)$$

To simplify the mathematical analysis we have given p a small positive imaginary part

$$p \to p + i\epsilon, \quad \epsilon > 0. \quad (23)$$

Equation (20) contains two unknown functions $f(k_y)$ and $h(k_y)$. If it can be shown that these functions are analytic in the upper and lower complex k_y planes, respectively, the Wiener–Hopf method can be applied to find the solution. We in fact have such

[4] G. A. Grinberg and Yu. V. Pimenov, Zh. Tekh. Fiz. **27**, 2326 (1957) [English transl.: Soviet Phys.—Tech. Phys. **2**, 2160 (1957)].

analytic properties since we are dealing with half-planes. Let us define

$$F(y) = \int_{-\infty}^{\infty} f(k_y) e^{ik_y y}\, dk_y, \quad y \leq 0 \tag{24}$$

and assume that

$$|F(y)| \sim A_1 e^{\tau_1 y}, \quad y \to -\infty \tag{25}$$

with $A_1, \tau_1 > 0$. Then it follows from (19) and (25) that $f(k_y)$ is analytic inside the upper half-plane.

$$-\tau_1 < \operatorname{Im} k_y < \infty.$$

Similarly, if we assume that

$$|H(y)| \sim A_2 e^{-\tau_2 y}, \quad y \to \infty \tag{26}$$

with $A_2, \tau_2 > 0$, we find that $h(k_y)$ is analytic inside the lower half-plane

$$-\infty < \operatorname{Im} k_y < \tau_2.$$

The kernel can be shown to be analytic inside the strip

$$-\epsilon < \operatorname{Im} k_y < \epsilon$$

by virtue of (23). Thus the conclusion is that the terms in (20) have a common strip of analyticity

$$-\sigma < \operatorname{Im} k_y < \sigma, \quad \sigma < \epsilon, \tau_1, \tau_2.$$

In the following when we speak of the upper and lower k_y planes, we refer to the regions $-\sigma < \operatorname{Im} k_y < \infty$ and $-\infty < \operatorname{Im} k_y < \sigma$, respectively.

The solution of (20) now depends on the possibility of factorizing an analytic function into two members, one analytic and nonzero in the upper k_y plane and the other analytic and nonzero in the lower k_y plane, as well as the possibility of separating another function into two terms, one analytic in the upper k_y plane and the other analytic in the lower k_y plane. First we rewrite (18) as follows:

$$\begin{aligned} K(k_y) &= \frac{2}{d}\frac{1}{k_y^2 + q^2}\frac{\sin wd}{wd}\frac{\frac{1}{2}d(w - \omega/v)}{\sin \frac{1}{2}d(w - \omega/v)} \\ &\qquad \times \frac{\frac{1}{2}d(w + \omega/v)}{\sin \frac{1}{2}d(w + \omega/v)} \\ &= \frac{2}{d}\frac{K_+(k_y)K_-(k_y)}{k_y^2 + q^2}, \end{aligned} \tag{27}$$

where $K_+(k_y)$ and $K_-(k_y)$ are analytic and nonzero in the upper and lower k_y planes, respectively. Using the infinite product representation

$$\frac{\sin z}{z} = \prod_{\substack{n=-\infty \\ n\neq 0}}^{\infty} \left(1 - \frac{z}{n\pi}\right) e^{z/n\pi}$$

we get

$$\sin wd/wd = L(k_y)L(-k_y),$$

where

$$L(k_y) = \prod_{n=1}^{\infty} \left\{ \left[1 - \left(\frac{pd}{n\pi}\right)^2 \right]^{\frac{1}{2}} - i\frac{k_y d}{n\pi} \right\} e^{i(k_y d/n\pi)}. \tag{28}$$

Then $L(k_y)$ and $L(-k_y)$ are analytic and nonzero in the upper and lower k_y planes, respectively, with the provision that we take

$$\operatorname{Im}\,[1 - (pd/n\pi)^2]^{\frac{1}{2}} < 0.$$

Similarly

$$\frac{\sin \frac{1}{2}d(w - \omega/v)}{\frac{1}{2}d(w - \omega/v)} \frac{\sin \frac{1}{2}d(w + \omega/v)}{\frac{1}{2}d(w + \omega/v)} = G(k_y)G(-k_y),$$

where

$$\begin{aligned} G(k_y) = \prod_{n=1}^{\infty} &\left\{ \left[\left(1 + \frac{\omega d}{2n\pi v}\right)^2 - \left(\frac{pd}{2n\pi}\right)^2 \right]^{\frac{1}{2}} - i\frac{k_y d}{2n\pi} \right\} \\ &\times e^{-(d/2n\pi)(\omega/v - ik_y)} \\ \times \prod_{n=1}^{\infty} &\left\{ \left[\left(1 - \frac{\omega d}{2n\pi v}\right)^2 - \left(\frac{pd}{2n\pi}\right)^2 \right]^{\frac{1}{2}} - i\frac{k_y d}{2n\pi} \right\} \\ &\times e^{(d/2n\pi)(\omega/v + ik_y)}. \end{aligned} \tag{29}$$

Then $G(k_y)$ and $G(-k_y)$ are analytic and nonzero in the upper and lower k_y planes, respectively, if we take

$$\operatorname{Im}\left[\left(1 + \frac{\omega d}{2n\pi v}\right)^2 - \left(\frac{pd}{2n\pi}\right)^2 \right]^{\frac{1}{2}} < 0,$$

$$\operatorname{Im}\left[\left(1 - \frac{\omega d}{2n\pi v}\right)^2 - \left(\frac{pd}{2n\pi}\right)^2 \right]^{\frac{1}{2}} < 0.$$

We can write

$$\begin{aligned} K_+(k_y) &= J(k_y)L(k_y)/G(k_y), \\ K_-(k_y) &= L(-k_y)/J(k_y)G(-k_y), \end{aligned} \tag{30}$$

where $J(k_y)$ is a nonzero entire function as yet undetermined. The infinite products in (28) and (29) in general have exponential growth in the appropriate half-planes. We want to choose $J(k_y)$ so that $K_+(k_y)$ and $K_-(k_y)$ have algebraic growth.

Consider the limit $|k_y| \to \infty$,

$$\begin{aligned} L(k_y) &\sim \prod_{n=1}^{\infty} \left[1 - i\frac{k_y d}{n\pi} \right] e^{i(k_y d/n\pi)} \\ &= \frac{e^{i\gamma(k_y d/\pi)}}{-i(k_y d/\pi)\Gamma[-i(k_y d/\pi)]} \\ &\sim [1/(2\pi)^{\frac{1}{2}}] e^{-1(i-\gamma)(k_y d/\pi)} [-i(k_y d/\pi)]^{i(k_y d/\pi) - \frac{1}{2}}, \\ &\qquad \operatorname{Im} k_y > -\sigma, \end{aligned}$$

where $\gamma = 0.5772\cdots$ is Euler's constant. Similarly,

$$G(k_y) \sim (1/2\pi) e^{-i(1-\gamma)(k_y d/\pi)} [-i(k_y d/2\pi)]^{i(k_y d/\pi) - 1}, \quad \operatorname{Im} k_y > -\sigma,$$

Hence

$$L(k_y)/G(k_y) \sim Ak_y^{\frac{1}{2}} e^{i(k_y d/\pi)\ln 2}, \quad |k_y| \to \infty, \quad \text{Im } k_y > -\sigma.$$

Thus if we define

$$J(k_y) = e^{-i(k_y d/\pi)\ln 2}, \tag{31}$$

$K_+(k_y)$ and $K_-(k_y)$ will have algebraic growth

$$K_\pm(k_y) \sim A_\pm k_y^{\pm\frac{1}{2}}, \quad |k_y| \to \infty. \tag{32}$$

These prescribed asymptotic behaviors are essential in the solution of the Wiener–Hopf equation. The functions $K_+(k_y)$ and $K_-(k_y)$ are now determined up to an arbitrary multiplicative constant which is taken to be unity, since it does not affect the final solution.

Next we want to separate the function

$$\psi(k_y) = \frac{i}{2\pi}\frac{k_y - iq}{K_-(k_y)}\left[\frac{B}{k_y + p + i\epsilon} + \frac{C}{k_y + iq}\right] \tag{33}$$

into two terms

$$\psi(k_y) = \psi_+(k_y) + \psi_-(k_y)$$

such that $\psi_+(k_y)$ and $\psi_-(k_y)$ are analytic in the upper and lower k_y planes, respectively. This is done by writing $\psi(k_y)$ as an integral along a closed rectangular contour bounding its region of analyticity $-\sigma < \text{Im } k_y < \sigma$. We have

$$\psi(k_y) = \frac{1}{2\pi i}\int_{-\infty-i\sigma}^{\infty-i\sigma} \frac{\psi(k_y')}{k_y' - k_y}\,dk_y' + \frac{-1}{2\pi i}\int_{-\infty+i\sigma}^{\infty+i\sigma} \frac{\psi(k_y')}{k_y' - k_y}\,dk_y'.$$

The first integral can be identified with $\psi_+(k_y)$ and the second one with $\psi_-(k_y)$. Evaluating the first integral by the method of residues, we get

$$\psi_+(k_y) = \frac{-i}{2\pi}\left[\frac{B}{K_-(-p)}\frac{p + iq}{k_y + p + i\epsilon} + \frac{C}{K_-(-iq)}\frac{2iq}{k_y + iq}\right]. \tag{34}$$

This separation is unique up to an arbitrary additive entire function which is taken to be zero, since it does not affect the final solution.

We rearrange (20) as follows:

$$\frac{2}{d}\frac{K_+(k_y)f(k_y)}{k_y + iq} - \psi_+(k_y) = \frac{(k_y - iq)h(k_y)}{K_-(k_y)} + \psi_-(k_y). \tag{35}$$

Each side of (35) defines a function $I(k_y)$, but since the left and right sides are analytic in the upper and lower k_y planes, respectively, $I(k_y)$ is analytic in the whole k_y plane and is therefore an entire function.

Clearly our equation is solved if we can determine $I(k_y)$. From physical considerations, the induced surface current density $K_{0\mu}(y, k_z, \omega)$ is integrable with respect to y. This implies that $f(k_y)$ is asymptotic to zero in the upper k_y plane. Thus the left side of (35) is asymptotic to zero in the upper k_y plane. Similarly, it can be shown that the right side is also asymptotic to zero in the lower k_y plane. So $I(k_y)$ is asymptotic to zero in all directions. Since, by the maximum modulus theorem, the modulus of an entire function assumes its maximum at the point at infinity, $I(k_y)$ must be identically zero. We therefore obtain the solution of our equation:

$$f(k_y) = \frac{-id}{4\pi K_+(k_y)}\left[\frac{B(p + iq)}{K_-(-p)}\frac{k_y + iq}{k_y + p + i} + \frac{2iCq}{K_-(-iq)}\right]. \tag{36}$$

Here B still contains an unknown constant b in (15) and (16). We adjust b so that the solution satisfies the boundary condition

$$K_{0y}(y, k_z, \omega) = 0, \quad y = 0.$$

We list the final solution as follows:

$$\begin{aligned}
K_{0y}(k_y, k_z, \omega) &= \frac{-iec\beta d}{(2\pi)^{\frac{3}{2}}} \times \frac{e^{-qa}}{K_-(-iq)K_+(k_y)}\frac{q}{\omega}\left[\frac{p - iq}{k_y + p + i\epsilon}\right], \\
K_{0z}(k_y, k_z, \omega) &= \frac{-iec\beta d}{(2\pi)^{\frac{3}{2}}}\frac{e^{-qa}}{K_-(-iq)K_+(k_y)}\frac{k_z}{\omega} \times \left[\frac{q}{p}\frac{k_y + iq}{k_y + p + i\epsilon} + i\right], \\
c\sigma_0(k_y, k_z, \omega) &= \frac{-iec\beta d}{(2\pi)^{\frac{3}{2}}}\frac{e^{-qa}}{K_-(-iq)K_+(k_y)}\frac{1}{c} \times \left[\frac{q}{p}\frac{k_y + iq}{k_y + p + i\epsilon} - i\frac{1 - \beta^2}{\beta^2}\right].
\end{aligned} \tag{37}$$

This solution can be shown to satisfy the assumptions made in (25) and (26). Moreover, the asymptotic behaviors

$$\begin{aligned}
K_{0y}(k_y, k_z, \omega) &\sim B_1 k_y^{-\frac{3}{2}}, \\
K_{0z}(k_y, k_z, \omega) &\sim B_2 k_y^{-\frac{1}{2}}, \\
c\sigma_0(k_y, k_z, \omega) &\sim B_3 k_y^{-\frac{1}{2}}, \quad k_y \to \infty,
\end{aligned}$$

imply that

$$\begin{aligned}
K_{0y}(y, k_z, \omega) &\sim C_1 y^{\frac{1}{2}}, \\
K_{0z}(y, k_z, \omega) &\sim C_2 y^{-\frac{1}{2}}, \\
c\sigma_0(y, k_z, \omega) &\sim C_3 y^{-\frac{1}{2}}, \quad y \to -0.
\end{aligned}$$

These are precisely the edge conditions in electromagnetic diffraction.[3] We notice that K_{0y} vanishes at the edge, while K_{0z} and $c\sigma_0$ diverge there.

IV. INDUCED FIELDS

We can calculate the induced potential $A_\mu(x, y, k_z, \omega)$ by substituting (37) into (9). The integration can easily be carried out, since the only singularities of the integrand on the k_y plane are simple poles. For $y \leq 0$, we complete the contour in the lower k_y plane. The integral is reduced to a sum of residues at the poles at (i) $k_y = -iq$, (ii) $k_y = -p - i\epsilon$, and (iii) $k_y = -[p^2 - (n\pi/d)^2]^{\frac{1}{2}} - i\epsilon$, $n = 1, 2, 3, \cdots$. The contribution from each pole may be interpreted as a characteristic solution of the wave equation. For $y \geq 0$ we complete the contour in the upper k_y plane and collect contributions from poles at (i) $k_y = iq$ and (ii) $k_y = \{p^2 - [(2n\pi/d) - (\omega/v)]^2\}^{\frac{1}{2}} + i\epsilon$, $n = \pm 1, \pm 2, \pm 3, \cdots$. From $A_\mu(x, y, k_z, \omega)$ the induced fields $\mathbf{E}(x, y, k_z, \omega)$ and $\mathbf{B}(x, y, k_z, \omega)$ are obtained by standard operations.

Let us first consider $y \leq 0$. The results are summarized as follows:

(i) Contribution from pole at $k_y = -iq$:

$$\mathbf{E}(x, y, k_z, \omega) = \frac{\mu_0}{4\pi}\frac{ec}{\beta} e^{i(\omega/v)x + q(y-a)} \times \left[i\frac{1-\beta^2}{\beta}\frac{k}{q}, 1, i\frac{k_z}{q}\right], \quad (38)$$

$$\mathbf{B}(x, y, k_z, \omega) = \frac{\mu_0}{4\pi} e\, e^{i(\omega/v)x + q(y-a)}\left(0, -i\frac{k_z}{q}, 1\right).$$

These fields exactly cancel the fields of the point charge as calculated from (6).

(ii) Contribution from pole at $k_y = -p - i\epsilon$:

$$\mathbf{E}(x, y, k_z, \omega) = P_0(k_z, \omega) e^{i(\omega/v)md} e^{-ipy}(q/p, 0, 0),$$
$$\mathbf{B}(x, y, k_z, \omega) = (1/c) P_0(k_z, \omega) e^{i(\omega/v)md} e^{-ipy}(0, qk_z/pk, q/k), \quad (39)$$

where m is an integer such that $md < x < (m+1)d$, and

$$P_0(k_z, \omega) = \frac{\mu_0}{4\pi}\frac{2ec\beta}{d}[1 - e^{i(\omega/v)d}]\frac{K_-(-p)}{K_-(-iq)}\frac{e^{-qa}}{p + iq}.$$

These fields satisfy the relations

$$\mathbf{E}\cdot\mathbf{B} = 0, \quad E = cB. \quad (40)$$

They represent a *TEM* wave propagating downward between the plates at $x = md$ and $(m+1)d$ with wave vector

$$\mathbf{k} = (0, -p, k_z)$$

(iii) Contribution from pole at

$$k_y = -[p^2 - (n\pi/d)^2]^{\frac{1}{2}} - i\epsilon, \quad n = 1, 2, 3, \cdots.$$

$$\mathbf{E}(x, y, k_z, \omega) = P_n(k_z, \omega) e^{i(\omega/v)md} e^{ik_{yn}y} \times \left\{\left[\frac{q}{p}\frac{k_{yn} + iq}{k_{yn} + p} - i\frac{1-\beta^2}{\beta^2}\right]\cos\frac{n\pi}{d}x, \left[i\frac{k_{yn}d}{n\pi}\left(\frac{q}{p}\frac{k_{yn} + iq}{k_{yn} + p} - \frac{1-\beta^2}{\beta^2}\right) - i\frac{qd}{n\pi}\frac{p - iq}{k_{yn} + p}\right] \times \sin\frac{n\pi}{d}x, \frac{1}{\beta^2}\frac{k_z d}{n\pi}\sin\frac{n\pi}{d}x\right\},$$

$$\mathbf{B}(x, y, k_z, \omega) = \frac{1}{c} P_n(k_z, \omega) e^{i(\omega/v)md} e^{ik_{yn}y} \quad (41)$$
$$\times \left\{\left[-i\frac{k_{yn}d}{n\pi}\frac{k_z}{k}\left(\frac{q}{p}\frac{k_{yn} + iq}{k_{yn} + p} + i\right) + i\frac{qd}{n\pi}\frac{k_z}{k}\frac{p - iq}{k_{yn} + p}\right] \times \sin\frac{n\pi}{d}x, \frac{k_z}{k}\left(\frac{q}{p}\frac{k_{yn} + iq}{k_{yn} + p} + i\right)\cos\frac{\pi n}{d}x, -\frac{q}{k}\frac{p - iq}{k_{yn} + p}\cos\frac{n\pi}{d}x\right\},$$

where $md < x < (m+1)d$,

$$k_{yn} = -[p^2 - (n\pi/d)^2]^{\frac{1}{2}} - i\epsilon,$$

and

$$P_n(k_z, \omega) = \frac{\mu_0}{4\pi}\frac{2ec\beta}{d}(-1)^{mn}[1 - (-1)^n e^{i(\omega/v)d}] \times \frac{K_-(k_{yn})}{K_-(-iq)}\frac{(n\pi/d)^2}{(\omega/v)^2 - (n\pi/d)^2}\frac{e^{-qa}}{k_{yn}}.$$

These fields also satisfy the relations (40). They represent higher excited modes propagating in the waveguide formed by the plates at $x = md$ and $(m+1)d$. For sufficiently large n, k_{yn} is imaginary and the fields are exponentially damped in the negative y direction.

For $y \geq 0$ the fields take on the following forms:

(i) Contribution from pole at $k_y = iq$:

$$\mathbf{E}(x, y, k_z, \omega) = Q_0(k_z, \omega) e^{i(\omega/v)x - qy} \times \left(-i\frac{k}{q\beta}\left[\frac{q}{p}\frac{2q}{p + iq} - \frac{1-\beta^2}{\beta^2}\right], \frac{q}{p}\frac{2q}{p + iq} + \frac{p - iq}{p + iq} - \frac{1-\beta^2}{\beta^2}, i\frac{k_z}{q\beta^2}\right),$$

$$\mathbf{B}(x, y, k_z, \omega) = \frac{1}{c} Q_0(k_z, \omega) e^{i(\omega/v)x - qy} \times \left(-\frac{k_z}{k}\left[\frac{q}{p}\frac{2q}{p + iq} + \frac{p - iq}{p + iq} + 1\right] - i\frac{k_z}{q\beta}\left[\frac{q}{p}\frac{2q}{p + iq} + 1\right], \frac{1}{\beta}\frac{p - iq}{p + iq}\right),$$

where

$$Q_0(k_z, \omega) = (\mu_0/4\pi)ec\beta[K_-(iq)/K_-(-iq)]e^{-qa}.$$

Since $k_x = \omega/v$, we deduce that the group velocity is of magnitude v in the positive x direction. Thus these fields are simply dragged along by the moving point charge. They do not represent any outflow of energy.

(ii) Contribution from pole at

$$k_y = \{p^2 - [(2n\pi/d) - (\omega/v)]^2\}^{\frac{1}{2}} + i\epsilon,$$

$n = \pm 1, \pm 2, \pm 3, \cdots$:

$$\mathbf{E}(x, y, k_z, \omega) = Q_n(k_z, \omega)e^{i[(\omega/v)-(2n\pi/d)]x+ik'_{yn}y}$$
$$\times\left[-i\left(\frac{\omega}{v} - \frac{2n\pi}{d}\right)\left(\frac{q}{p}\frac{k'_{yn} + iq}{k'_{yn} + p} - i\frac{1-\beta^2}{\beta^2}\right),\right.$$
$$-ik'_{yn}\left(\frac{q}{p}\frac{k'_{yn} + iq}{k'_{yn} + p} - i\frac{1-\beta^2}{\beta^2}\right)$$
$$\left. + iq\frac{p - iq}{k'_{yn} + p}, -\frac{k_z}{\beta^2}\right],$$

$$\mathbf{B}(x, y, k_z, \omega) = \frac{1}{c}Q_n(k_z, \omega)e^{i[(\omega/v)-(2n\pi/d)]x+ik_{yn}'y} \quad (42)$$
$$\times\left[ik'_{yn}\frac{k_z}{k}\left(\frac{q}{p}\frac{k'_{yn} + iq}{k'_{yn} + p} + i\right) - iq\frac{k_z}{k}\frac{p - iq}{k'_{yn} + p},\right.$$
$$-i\left(\frac{\omega}{v} - \frac{2n\pi}{d}\right)\frac{k_z}{k}\left(\frac{q}{p}\frac{k'_{yn} + iq}{k'_{yn}} + i\right),$$
$$\left. i\left(\frac{\omega}{v} - \frac{2n\pi}{d}\right)\frac{q}{k}\frac{p - iq}{k'_{yn} + p}\right],$$

where $k'_{yn} = \{p^2 - [(2n\pi/d) - (\omega/v)]^2\}^{\frac{1}{2}} + i\epsilon$,

$$Q_n(k_z, \omega) = \frac{\mu_0}{4\pi}\frac{2ec\beta}{(\omega/v)^2 - [(2n\pi/d) - (\omega/v)]^2}$$
$$\times\frac{\text{Res } K_-(k'_{yn})}{K_-(-iq)}e^{-qa}$$

and Res $K_-(k'_{yn})$ is the residue of $K_-(k_y)$ at the pole $k_y = k'_{yn}$. These fields satisfy the relations (40). They represent a wave with propagation vector

$$\mathbf{k} = [\omega/v - (2n\pi/d), k'_{yn}, k_z).$$

For sufficiently large n, they are exponentially damped in the positive y direction.

V. POYNTING'S VECTOR

The property of the radiation is studied by calculating the Poynting vector

$$\mathbf{S}(x, y, z, \omega) = \mathbf{E}(x, y, z, \omega) \times \mathbf{H}^*(x, y, z, \omega). \quad (43)$$

To this end we must first carry out the k_z inversion on the fields calculated in the previous section. This amounts to evaluating integrals of the form

$$I = \frac{1}{(2\pi)^{\frac{1}{2}}}\int_{-\infty}^{\infty} f(k_z)\exp[\pm i(\lambda^2 - k_z^2)^{\frac{1}{2}}y + ik_z z]\,dk_z. \quad (44)$$

This integral cannot be evaluated exactly, since $f(k_z)$ has a very complex structure. But if we write

$$y = \pm\rho\cos\phi, \quad z = \rho\sin\phi, \quad -\tfrac{1}{2}\pi \le \phi \le \tfrac{1}{2}\pi,$$

where the upper and lower signs correspond to the half-spaces $y \ge 0$ and $y \le 0$, respectively, then for $\rho \to \infty$ we can use the method of steepest descent to get an expression for the far fields:

$$I \sim f(\lambda\sin\phi)\lambda\cos\phi(\lambda\rho)^{-\frac{1}{2}}e^{i\lambda\rho - i(\frac{1}{4}\pi)}. \quad (45)$$

This is in the form of a cylindrical wave.

We calculate the Poynting vector from the far fields and summarize the results below. It is clear that only undamped fields need be considered. For $y \le 0$ we get

(i) Contribution from pole at $k_y = -p - i\epsilon$:

$$\mathbf{S}(x, y, z, \omega) = \frac{\mu_0}{4\pi}\frac{2e^2v^2}{\pi cd^2}\left(1 - \cos\frac{\omega}{v}d\right)$$
$$\times(1 - \beta^2\cos^2\phi)\left|\frac{K_-(-p)}{K_-(-iq)}\right|^2\frac{e^{-2qa}}{k\rho}\mathbf{e}_\rho. \quad (46)$$

In this expression we must put $k_z = k\sin\phi$.

(ii) Contribution from pole at

$$k_y = -[p^2 - (n\pi/d)^2]^{\frac{1}{2}} - i\epsilon, \quad n = 1, 2, 3, \cdots:$$

In this case it is more appropriate to give the Poynting vector averaged over one period d:

$$\bar{\mathbf{S}}(x, y, z, \omega) = \frac{1}{d}\int_{md}^{(m+1)d}\mathbf{S}(x, y, z, \omega)\,dx.$$

We get

$$\bar{\mathbf{S}}(x, y, z, \omega) = \frac{\mu_0}{4\pi}\frac{e^2\omega^2}{\pi cd^2}\left[1 - (-1)^n\cos\frac{\omega}{v}d\right]$$
$$\times\frac{(n\pi/d)^2}{[(\omega/v)^2 - (n\pi/d)^2]^2}\left|\frac{K_-(k_{yn})}{K_-(-iq)}\right|^2$$
$$\times\frac{p(n\pi/d)^2[(1-\beta^2)/\beta^2] + 2(k_z/\beta)^2(k_{yn} + p)}{p(k_{yn} + p)^2}\frac{e^{-2qa}}{k\rho}\mathbf{e}_\rho. \quad (47)$$

In this expression we must put

$$k_z = [k^2 - (n\pi/d)^2]^{\frac{1}{2}}\sin\phi.$$

For $y \geq 0$ we get contributions only from poles at

$$k_y = \{p^2 - [(2n\pi/d) - (\omega/v)]^2\}^{\frac{1}{2}} + i\epsilon, \qquad n = \pm 1, \pm 2, \pm 3, \cdots,$$

$$\mathbf{S}(x, y, z, \omega) = \frac{\mu_0}{4\pi}\frac{e^2\omega^2}{\pi c}\frac{k^2 - [(2n\pi/d) - (\omega/v)]^2}{\{(\omega/v)^2 - [(2n\pi/d) - (\omega/v)]^2\}^2} \times \left|\frac{\text{Res } K_-(k'_{yn})}{K_-(-iq)}\right|^2 \times \frac{p[(2n\pi/d) - (\omega/v)]^2[(1-\beta^2)/\beta^2] + 2(k_z/\beta)^2(k'_{yn} + p)}{p(k'_{yn} + p)^2} \times \frac{e^{-2qa}}{k\rho}\frac{\cos^2\phi}{\{k^2 - [(2n\pi/d) - (\omega/v)]^2\}^{\frac{1}{2}}}\mathbf{k}, \tag{48}$$

where

$$\mathbf{k} = \left(\frac{\omega}{v} - \frac{2n\pi}{d}\right)\mathbf{e}_x + \left[k^2 - \left(\frac{2n\pi}{d} - \frac{\omega}{v}\right)^2\right]^{\frac{1}{2}}\mathbf{e}_\rho. \tag{49}$$

In (48) we must put

$$k_z = \{k^2 - [(2n\pi/d) - (\omega/v)]^2\}^{\frac{1}{2}} \sin\phi.$$

From Eq. (49) we see that at fixed frequency ω, the x component of the wave vector in the upper half-space $y \geq 0$ has only a finite number of discrete values:

$$k_x = (\omega/v) - (2n\pi/d), \quad n = \pm 1, \pm 2, \pm 3, \cdots. \tag{50}$$

We recognize this to be the von Laue condition for wave diffraction by a periodic structure if we recall that ω/v is the x component of the wave vector of the incident wave (6). If we let ω vary and observe the radiation at a fixed angle θ with respect to the x axis, then Eq. (50) gives

$$\omega = (2n\pi v/d)/(1 - \beta\cos\theta), \qquad n = \pm 1, \pm 2, \pm 3, \cdots. \tag{51}$$

Equation (51) states that the frequency spectrum at fixed θ consists of discrete lines corresponding to integral multiples of $2\pi v/d$ which measures the number of plates traversed by the point charge per unit time, shifted by the Doppler factor $(1 - \beta\cos\theta)^{-1}$.

The presence of infinite products in (46), (47), and (48) obscures the properties of the Poynting vector, in particular, its angular dependence. However, recently Kazantsev and Surdutovich[5] solved the problem of the excitation of a conducting half-plane by a uniformly moving point charge. These authors found a Poynting vector of the form

$$\mathbf{S}(x, y, z, \omega) = \frac{\mu_0}{4\pi}\frac{e^2 v}{8\pi^2 r^2} \times \frac{\cos^2\theta(1 + \sin\phi) + (1 - \beta^2\sin^2\theta)(1 - \sin\phi)}{\sin\theta(1 - \beta^2\sin^2\theta)(1 - \beta^2\sin^2\theta\cos^2\phi)} \times \exp\left\{-\frac{2a|\omega|}{v}[1 - \beta^2\sin^2\theta]^{\frac{1}{2}}\right\}, \tag{52}$$

where $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$, $z = r\cos\theta$ and the half-plane is taken to be at $x = 0$, $y \leq 0$. In general the radiation is concentrated along the y axis at high frequencies, along the z axis at low frequencies, and along the x axis at high velocities ($\beta \simeq 1$). In the approximation of zero interaction among the plates, the radiation pattern of N plates spaced evenly with a separation d near the plane $x = 0$ can be synthesized from that of a single plate. We get a Poynting vector equal to (52) multiplied by an "array factor"

$$\frac{\sin^2(\frac{1}{2}Nd)(k_x - \omega/v)}{\sin^2(\frac{1}{2}d)(k_x - \omega/v)},$$

which is sharply peaked in directions satisfying (50).

ACKNOWLEDGMENTS

The author is indebted to Professor C. H. Papas for his guidance, stimulating discussions and encouragement; and to Professor J. Mathews and Dr. K. S. H. Lee for their interest and criticism.

This work was supported by the U.S. Air Force Office of Scientific Research under Contract AF49 (638)–1266.

[5] A. P. Kazantsev and G. I. Surdutovich, Dokl. Akad. Nauk. SSSR 147, 74 (1962) [English transl.: Soviet Phys.—Dokl. 7, 990 (1963)].

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 5 MAY 1967

Absence of Ordering in Certain Classical Systems

N. D. MERMIN*
Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York

(Received 6 December 1966)

A classical inequality giving lower bounds for fluctuations about ordered states is derived. The inequality, analogous to a quantum result due to Bogoliubov, is established by a purely classical argument which makes explicit the nature of the surface boundary conditions required, a point which is rather obscure in the quantum derivations. As in the quantum case the inequality is useful in excluding certain kinds of phase transitions in one- and two-dimensional systems. This is illustrated for several kinds of classical spin systems.

AN exact inequality due to Bogoliubov[1] has recently been used to prove that several kinds of phase transitions cannot occur in one- and two-dimensional quantum systems.[2,3] The derivation of the inequality relies heavily on the formal apparatus of quantum mechanics, but if the quantum phase transitions have classical analogs, similar conclusions about the classical systems follow from taking the $\hbar \to 0$ limit of the final quantum result.

One would nevertheless like a purely classical proof of the Bogoliubov inequality for several reasons. It is, if nothing else, distasteful to be forced to appeal to the classical limit of a quantum mechanical result, in an otherwise purely classical argument. Furthermore, the detailed evaluation of terms in the inequality in particular cases is frequently less laborious classically, since terms that ultimately vanish in the $\hbar \to 0$ limit of the quantum result are dropped from the beginning, and equipartition of energy can frequently be exploited. Finally, the effect of surface conditions on the Bogoliubov inequality is somewhat hard to ascertain in the quantum theory, while in the classical theory the surface can lead directly to explicit correction terms in the basic inequality.

For these reasons a direct classical proof of the Bogoliubov inequality is given below. However, one should realize that the inequality is proved only for Hamiltonian systems, and these are always classical analogs of quantum systems. Therefore, aside from esthetic and calculational matters, the primary point of this note is to establish the validity of the naive $\hbar \to 0$ limit and to emphasize (as has not been done in the quantum derivations) the importance spatial boundary conditions can have on the form of the inequality.

Consider, then, a classical system described by canonical variables $q_1 \cdots q_n$, $p_1 \cdots p_n$, and Hamiltonian $H(q_1 \cdots p_n)$. We define the canonical ensemble average $\langle A \rangle$ of any function $A(q_1 \cdots p_n)$ in the usual way:

$$\langle A \rangle = \int d\Gamma e^{-\beta H} A \Big/ \int d\Gamma e^{-\beta H} = \int d\Gamma e^{-\beta(H-F)} A,$$

where $d\Gamma$ is the phase space volume element, $d\Gamma = dq_1 \cdots dp_n$, and $\beta = 1/k_B T$.

For any functions A and B for which the ensemble averages $\langle |A|^2 \rangle$, $\langle |B|^2 \rangle$, and $\langle A^* B \rangle$ converge, we have

$$\langle |A - \langle AB^* \rangle B / \langle |B|^2 \rangle |^2 \rangle \geq 0,$$

and hence the Schwartz inequality,

$$\langle |A|^2 \rangle \geq |\langle A^* B \rangle|^2 / \langle |B|^2 \rangle \tag{1}$$

(with equality if and only if A is a constant times B).

Classical versions of the Bogoliubov inequality emerge when the function B is of the form $[C, H]$, where the bracket is a Poisson bracket:

$$[A, B] = \sum_{i=1}^{n} \left(\frac{\partial A}{\partial q_i} \frac{\partial B}{\partial p_i} - \frac{\partial B}{\partial q_i} \frac{\partial A}{\partial p_i} \right). \tag{2}$$

It is then easily verified that

$$\langle A^* [C, H] \rangle = k_B T \langle [C, A^*] \rangle + \Delta_1 - \Delta_2, \tag{3}$$

where Δ_1 and Δ_2 are given by

$$\Delta_1 = k_B T \sum_{i=1}^{n} \int d\Gamma \frac{\partial}{\partial q_i} \left(A^* \frac{\partial C}{\partial p_i} e^{-\beta(H-F)} \right), \tag{4}$$

$$\Delta_2 = k_B T \sum_{i=1}^{n} \int d\Gamma \frac{\partial}{\partial p_i} \left(A^* \frac{\partial C}{\partial q_i} e^{-\beta(H-F)} \right). \tag{5}$$

For each i the integration over q_i in (4) and p_i in (5) can be done explicitly, showing that Δ_1 and Δ_2 are surface corrections to the basic result:

$$\langle A^* [C, H] \rangle = k_B T \langle [C, A^*] \rangle. \tag{6}$$

* Alfred P. Sloan Foundation Fellow.

[1] N. N. Bogoliubov, Phys. Abhandl. S.U. 6, 113 (1962); see also H. Wagner, Z. Physik 195, 273 (1966).

[2] P. C. Hohenberg, Phys. Rev. (to be published).

[3] N. D. Mermin and H. Wagner, Phys. Rev. Letters 17, 1133 (1966).

In a variety of cases the surface corrections to (6) vanish identically, either because the relevant functions are periodic, or because the statistical weight $e^{-\beta H}$ vanishes as any p_i or q_i goes to infinity. We assume this to be the ordinary state of affairs, in which case (6) and (1) imply, with $B = [C, H]$ that:

$$\langle |A|^2 \rangle \geq \frac{k_B T \, |\langle [C, A^*] \rangle|^2}{\langle [C, [C^*, H]] \rangle}. \tag{7}$$

This is precisely the naive classical limit of the quantum Bogoliubov inequality in the form given in Ref. 3, with commutators replaced by Poisson brackets times $i\hbar$.[4]

Equation (7) is the main result of this note. The question of when it has to be modified by nonvanishing surface terms must be taken up in each particular application of (7). Here we only illustrate its use in a few simple cases in which the surface terms vanish. [It is at first surprising that in the classical theory the validity of (7) depends on the vanishing of various surface terms, since no such problem appears to arise in the quantum theory; it is, however, swept under the rug[5] in the implicit assumption that the Hamiltonian is Hermitian, which allows one to integrate by parts in matrix elements like $\langle E | HA | E' \rangle$ to get $E \langle E | A | E' \rangle$ and no surface terms.]

Example 1:

$$H = \sum_{i=1}^{n} p_i^2/2m_i + U(q_1 \cdots q_n).$$

This form of H is relevant in discussing the possibility of crystalline ordering in one and two dimensions. The details of this problem will be presented elsewhere,[6] and here only the following elementary points are made:

Terms of the Δ_2 type [Eq. (5)] vanish for reasonable A and C because of the quadratic dependence of H on momentum. Terms of the Δ_1 type [Eq. (4)] are somewhat trickier. In using the Bogoliubov inequality one always works in a finite volume, reserving the thermodynamic limit for the end of the argument. If the finite volume is maintained by an explicit potential well term in U, then Δ_1 terms will also vanish at infinity in configuration space for well-behaved A and C. However, in this case the contribution of the well term in U must be kept in the denominator of (7). Alternatively, one can avoid such terms in the denominator of (7) by working in a configuration space of volume V, but then the Δ_1 surface terms do not vanish, and care must be taken to demonstrate that the arguments excluding ordering in one or two dimensions remain valid in their presence. Finally, one can avoid the problem by imposing the analog of the quantum periodic boundary conditions—i.e., artificially redefining U to have the macroscopic periodicity of the box of volume V, and only considering functions A and C which have this periodicity in configuration space.

Example 2: Models involving classical three-dimensional spins, on lattices of arbitrary dimensionality.

Consider the classical statistical mechanical system defined by the following free energy:

$$e^{-\beta F} = \int P(\{S(\mathbf{R})\}) \prod_{\mathbf{R}} d\mathbf{S}(\mathbf{R}) e^{-\beta H},$$
$$H = -\sum_{\mathbf{R}\mathbf{R}'} J(\mathbf{R} - \mathbf{R}') \mathbf{S}(\mathbf{R}) \cdot \mathbf{S}(\mathbf{R}') - h \sum_{\mathbf{R}} S_z(\mathbf{R}), \tag{8}$$

where the index $\mathbf{R}$ runs over the sites of a Bravais lattice of N sites with the usual periodic boundary conditions, and the *a priori* weight function P depends only on the magnitude $S(\mathbf{R})$ of each of the spins. Two special cases are of particular interest[7]:

(1) If

$$P(\{S(\mathbf{R})\}) \to \prod_{\mathbf{R}} \delta(S - S(\mathbf{R})),$$

then each spin has the same fixed magnitude S, and the integrations are only over the directions of each spin. This is just the classical Heisenberg model.

(2) If

$$P(\{S(\mathbf{R})\}) \to \delta\left(NS^2 - \sum_{\mathbf{R}} S^2(\mathbf{R})\right),$$

then (8) gives the free energy for a spherical model of three-dimensional spins.

Provided that the *a priori* probability of any single spin being infinite in magnitude vanishes and provided that J is of finite range, the system (8) can be shown to have no spontaneous magnetization in one and two dimensions. The trick in constructing the classical proof is to define the Poisson bracket. Suppose we first fix the magnitude of each spin. The system can then be described by the $2N$ canonical variables $\theta(\mathbf{R})$ and $S_z(\mathbf{R})$, where $\theta(\mathbf{R})$ is the angle between the projection of $\mathbf{S}(\mathbf{R})$ in the x–y plane and the x axis, and $S_z(\mathbf{R})$ is

[4] Note that one has equality in (7) (provided surface terms can be ignored) if and only if A is a constant times $[C, H]$. This is not the case in the quantum version of (7) since its derivation requires both the Schwartz inequality and the replacement of the quantum fluctuation dissipation theorem by its classical form (which maintains the inequality).

[5] The author is indebted to J. Langer for directing his attention to the correct rug.

[6] N. D. Mermin (to be published).

[7] The introduction of P slightly simplifies the proof at the expense of introducing a singular *a priori* distribution function in the two cases of chief interest. However, direct proofs that do not use P can easily be constructed in these cases.

the projection of $\mathbf{S}(\mathbf{R})$ on the z axis. The Poisson bracket is therefore

$$[A, B] = \sum_{\mathbf{R}} \left(\frac{\partial A}{\partial \theta(\mathbf{R})} \frac{\partial B}{\partial S_z(\mathbf{R})} - \frac{\partial A}{\partial S_z(\mathbf{R})} \frac{\partial B}{\partial \theta(\mathbf{R})} \right), \quad (9)$$

where the derivatives are to be taken at fixed spin magnitudes.

Rather than rewriting H in terms of the canonical variables and the dynamically inert spin magnitudes, it is simpler to rewrite the Poisson brackets in terms of the variables S_x, S_y, and S_z. Thus if

$$S_x = (S^2 - S_z^2)^{\frac{1}{2}} \cos\theta,$$
$$S_y = (S^2 - S_z^2)^{\frac{1}{2}} \sin\theta,$$

then

$$\frac{\partial A}{\partial \theta} \frac{\partial B}{\partial S_z} - \frac{\partial B}{\partial \theta} \frac{\partial A}{\partial S_z} = \left(\frac{\partial A}{\partial S_x} \frac{\partial S_x}{\partial \theta} + \frac{\partial A}{\partial S_y} \frac{\partial S_y}{\partial \theta} \right) \times \left(\frac{\partial B}{\partial S_x} \frac{\partial S_x}{\partial S_z} + \frac{\partial B}{\partial S_y} \frac{\partial S_y}{\partial S_z} + \frac{\partial B}{\partial S_z} \right) - (A \leftrightarrow B)$$
$$= \mathbf{S} \cdot \frac{\partial A}{\partial \mathbf{S}} \times \frac{\partial B}{\partial \mathbf{S}}$$

and therefore[8]

$$[A, B] = \sum_{\mathbf{R}} \mathbf{S}(\mathbf{R}) \cdot \frac{\partial A}{\partial \mathbf{S}(\mathbf{R})} \times \frac{\partial B}{\partial \mathbf{S}(\mathbf{R})}. \quad (10)$$

Equation (6) without any surface corrections can now be proved as follows:

$$\langle A^*[C, H] \rangle = \sum_{\mathbf{R}'} \int \prod_{\mathbf{R}} d\mathbf{S}(\mathbf{R}) P(\{S(\mathbf{R})\}) e^{-\beta(H-F)} A^* \mathbf{S}(\mathbf{R}') \cdot \frac{\partial C}{\partial \mathbf{S}(\mathbf{R}')} \times \frac{\partial H}{\partial \mathbf{S}(\mathbf{R}')}$$
$$= -\frac{1}{\beta} \sum_{\mathbf{R}'} \int \prod_{\mathbf{R}} d\mathbf{S}(\mathbf{R}) A^* \mathbf{S}(\mathbf{R}') \cdot \frac{\partial C}{\partial \mathbf{S}(\mathbf{R}')} \times \frac{\partial}{\partial \mathbf{S}(\mathbf{R}')} [P(\{S(\mathbf{R})\}) e^{-\beta(H-F)}]. \quad (11)$$

[The last line of (11) is justified by observing that

$$[C, Pe^{-\beta H}] = [C, e^{-\beta H}]P + [C, P]e^{-\beta H}$$

and

$$[C, P] = 0,$$

since P depends only on the spin magnitudes.] We may rewrite (11) as

$$\langle A^*[C, H] \rangle = -\frac{1}{\beta} \sum_{\mathbf{R}'} \int \prod_{\mathbf{R}} d\mathbf{S}(\mathbf{R}) A^* \left[\frac{1}{2} \frac{\partial}{\partial \mathbf{S}(\mathbf{R}')} S(\mathbf{R}')^2 \right] \cdot \frac{\partial C}{\partial \mathbf{S}(\mathbf{R}')} \times \frac{\partial}{\partial \mathbf{S}(\mathbf{R}')} (Pe^{-\beta(H-F)})$$
$$= \frac{1}{2\beta} \sum_{\mathbf{R}'} \int \prod_{\mathbf{R}} d\mathbf{S}(\mathbf{R}) S^2(\mathbf{R}') \frac{\partial}{\partial \mathbf{S}(\mathbf{R}')} A^* \cdot \frac{\partial C}{\partial \mathbf{S}(\mathbf{R}')} \times \frac{\partial}{\partial \mathbf{S}(\mathbf{R}')} (Pe^{-\beta(H-F)}), \quad (12)$$

where the integration by parts is justified by the assumed rapid vanishing of P for large spin magnitudes.

Now the integrand in (12) is invariant under cyclic permutation of A^*, C, and $Pe^{-\beta(H-F)}$, as must therefore be the last line of (11). But the permutation $A^* \to C \to Pe^{-\beta(H-F)} \to A^*$ reduces the last line of (11) to $(-1/\beta)[A^*, C]$, which proves (6).

With the Bogoliubov inequality (7) established for this classical spin system, the proof that there is no ferromagnetism in one and two dimensions is essentially the same as that given in Ref. 3:

We define

$$\mathbf{S}(\mathbf{k}) = \sum_{\mathbf{R}} e^{-i\mathbf{k}\cdot\mathbf{R}} S(\mathbf{R}), \quad J(\mathbf{k}) = \sum_{\mathbf{R}} e^{-i\mathbf{k}\cdot\mathbf{R}} J(\mathbf{R})$$

and let $C = S_+(\mathbf{k})$, $A = S_-(-\mathbf{k})$. Equation (7) now leads to the inequality[9]

$$(1/N^2)\langle |S_+(\mathbf{k})|^2 \rangle \geq \frac{1}{N} \frac{k_B T s_z^2}{\left[\frac{1}{N^2} \sum_{\mathbf{k}'} (J(\mathbf{k}') - J(\mathbf{k}' - \mathbf{k})) \times \langle |S_z(\mathbf{k}')|^2 + \frac{1}{2}|S_+(\mathbf{k}')|^2 \rangle + \frac{1}{2} h s_z \right]},$$

$$s_z = \frac{1}{N} \sum_{\mathbf{R}} \langle S_z(\mathbf{R}) \rangle.$$

But since

$$\sum_{\mathbf{k}} \frac{1}{N^2} \langle |S_+(\mathbf{k})|^2 \rangle = \frac{1}{N} \sum_{\mathbf{R}} \langle |S_+(\mathbf{R})|^2 \rangle < S^2,$$

$$\sum_{\mathbf{k}} \frac{1}{N^2} (\langle |S_z(\mathbf{k})|^2 + \tfrac{1}{2}|S_+(\mathbf{k})|^2 \rangle) < S^2,$$

where S^2 is the mean square magnitude of an individual spin, we have

$$S^2 \geq 2k_B T s_z^2 \frac{1}{N} \sum_{\mathbf{k}} \frac{1}{k^2 S^2 \sum_{\mathbf{R}} R^2 |J(\mathbf{R})| + |h|\,|s_z|}.$$

[8] Compare N. D. Mermin, Phys. Rev. 134, A112 (1964), Appendix A.

[9] Sums over k in both spin examples are restricted to the first Brillouin zone.

In the limit of an infinite system the sum becomes an integral, which diverges in one and two dimensions as h goes to zero, thereby proving that s_z must vanish with vanishing h.

Example 3: Lattice of plane rotors.[10]

If we take not three-dimensional spins, as in Example 2, but two-dimensional ones, the argument simplifies considerably. Let

$$e^{-\beta F} = \int_0^{2\pi} \prod_{\mathbf{R}} d\theta(\mathbf{R}) e^{-\beta H},$$

$$H = -\sum_{\mathbf{R}'} J(\mathbf{R}-\mathbf{R}') \cos\big(\theta(\mathbf{R}) - \theta(\mathbf{R}')\big) - h\sum_{\mathbf{R}} \cos\theta(\mathbf{R}). \tag{13}$$

The canonical variables are now $\theta(\mathbf{R})$ and $P(\mathbf{R})$, the angular momentum perpendicular to the plane of rotation. We take

$$A = \sum_{\mathbf{R}} \sin\theta(\mathbf{R}) e^{-i\mathbf{k}\cdot\mathbf{R}},$$

$$C = \sum_{\mathbf{R}} P(\mathbf{R}) e^{-i\mathbf{k}\cdot\mathbf{R}},$$

and hence

$$[C, H] = -\sum_{\mathbf{R}} e^{-i\mathbf{k}\cdot\mathbf{R}} \frac{\partial H}{\partial\theta(\mathbf{R})}.$$

In this model ensemble averages involve integrations only over each $\theta(\mathbf{R})$ between 0 and 2π. However, from the form of $[C, H]$ it is directly evident that (6) is still valid, since only an integration by parts with respect to $\theta(R)$ is required to prove it, and all functions in the integrand are periodic.

Therefore (7) again holds without any surface corrections, and the proof that this model is not ferromagnetic in one or two dimensions proceeds as follows:

$$\begin{aligned}\langle [C, [C^*, H]]\rangle &= \sum_{\mathbf{R}\mathbf{R}'} e^{i\mathbf{k}\cdot(\mathbf{R}-\mathbf{R}')} \left\langle \frac{\partial^2 H}{\partial\theta(\mathbf{R})\partial\theta(\mathbf{R}')} \right\rangle \\ &= 2\sum_{\mathbf{R}\mathbf{R}'} J(\mathbf{R}-\mathbf{R}')(1 - e^{i\mathbf{k}\cdot(\mathbf{R}-\mathbf{R}')})\langle\cos\big(\theta(\mathbf{R}) - \theta(\mathbf{R}')\big)\rangle \\ &\qquad + h\sum_{\mathbf{R}} \langle\cos\theta(\mathbf{R})\rangle \\ &\leq N\Big(\sum_{\mathbf{R}} R^2 |J(\mathbf{R})|\, k^2 + |h|\,|m|\Big),\end{aligned} \tag{14}$$

where m is the magnetization per particle,

$$m = \frac{1}{N}\sum_{\mathbf{R}} \langle\cos\theta(\mathbf{R})\rangle.$$

Furthermore $\langle [A^*, C]\rangle$ is just Nm. Therefore from (7) we have

$$\sum_{\mathbf{R}\mathbf{R}'} \langle\sin\theta(\mathbf{R})\sin\theta(\mathbf{R}')\rangle e^{i\mathbf{k}\cdot(\mathbf{R}-\mathbf{R}')} \geq \frac{Nk_BTm^2}{k^2\big(\sum R^2 |J(\mathbf{R})|\big) + |h|\,|m|}.$$

But

$$\frac{1}{N^2}\sum_{\mathbf{k}} \left(\sum_{\mathbf{R}\mathbf{R}'} \langle\sin\theta(\mathbf{R})\sin\theta(\mathbf{R}')\rangle e^{i\mathbf{k}\cdot(\mathbf{R}-\mathbf{R}')}\right) = \frac{1}{N}\sum_{\mathbf{R}} \langle\sin^2\theta(\mathbf{R})\rangle < 1,$$

and hence

$$1 > k_BTm^2 \frac{1}{N}\sum_{\mathbf{k}} [k^2\big(\sum R^2 |J(\mathbf{R})|\big) + |h|\,|m|]^{-1}. \tag{15}$$

In the thermodynamic limit this becomes ($n = N/V$)

$$1 > \frac{k_BTm^2}{n}\int \frac{d\mathbf{k}}{(2\pi)^d} [k^2\big(\sum R^2 |J(\mathbf{R})|\big) + |h|\,|m|]^{-1},$$

(where d is the number of dimensions) which again requires that m vanish as h goes to zero in one and two dimensions.

[10] This model has been studied by M. Kac, as described in a recent lecture at Cornell. The analysis of Kac's model given here was constructed at the suggestion and with the vigorous assistance of M. E. Fisher.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation under Contract GP-5517.

Three-Dimensional Formulation of Gravitational Null Fields. II

R. M. MISRA AND R. A. SINGH
Department of Physics, University of Gorakhpur, Gorakhpur, India

(Received 4 May 1966)

In the previous paper of this series, it was shown that three types of gravitational null fields may be characterized on the analogy of the electromagnetic field. The first two types were discussed in the previous paper. This paper discusses the remaining third type (C) of gravitational null field. The necessary and sufficient condition that the gravitational field be of type C is obtained. The formalism is also extended to include nonempty gravitational fields. It is shown that the nonempty space–time may also admit three types of gravitational null fields under certain circumstances. A typical case is discussed as an example.

1. INTRODUCTION

IN the first paper of this series,[1] we developed a criteria for the characterization of vacuum gravitational null fields and showed how this criteria leads to three distinct types of null fields designated as fields of type A, B, and C. In I, we considered types A and B in detail under the assumption that $R_{ab} = 0$. In this paper, we consider vacuum null fields of type C. We also extend our theory to include the nonempty gravitational fields as well.

Since in I we considered the theory in detail, we give only a brief outline here. In a vacuum normal hyperbolic Riemann space–time (i.e., $R_{ab} = 0$), we can construct two types of dual tensors from the curvature tensor in the following manner[2]:

$$^{*}R_{ijkl} = \tfrac{1}{2}\epsilon_{ijmn}R^{mn}_{kl}, \tag{1.1}$$

$$^{*}{}^{*}R_{ijkl} = \tfrac{1}{4}\epsilon_{ijmn}\epsilon_{klpq}R^{mnpq} \tag{1.2}$$

with the following property[3]:

$$^{*}R_{ijkl} = R^{*}_{ij}{}_{lk} \tag{1.3}$$

and

$$^{*}{}^{*}R_{ijkl} + R_{ijkl} = 0, \tag{1.4}$$

where $^{*}R$ and R^{*} are the left and the right duals (indices suppressed), respectively. Now, the following tensors are formed from the curvature tensor and its duals:

$$G_{ij} = R_{ikjl}u^k u^l, \tag{I.2.7}$$

$$H_{ij} = {}^{*}R_{ikjl}u^k u^l, \tag{I.2.8}$$

$$K_{ij} = {}^{*}{}^{*}R_{ikjl}u^k u^l, \tag{I.2.9}$$

where u^k is a timelike unit velocity vector representing the world line of the observer and satisfies

$$u^k u_k = -1. \tag{1.5}$$

The tensors G, H, and K satisfy the following conditions[4]:

$$G^k_k = H^k_k = K^k_k = 0, \tag{I.2.10}$$

$$G_{[ij]} = H_{[ij]} = K_{[ij]} = 0. \tag{I.2.11}$$

It should be remarked here that Eq. (I.2.10) is valid only in case of vacuum gravitational field, whereas Eq. (I.2.11) states that these tensors are symmetric. This can be easily seen from the symmetry properties of the curvature tensor. Now assuming the equalities of the eigenvalues of G_{ij} and H_{ij}, and the definite relative orientation of their eigenvectors [cf., assumptions (A) and (B) in I] leads to three types of gravitational null fields. Gravitational null fields of types A and B have already been considered in I, we study here the null fields of type C.

2. GRAVITATIONAL NULL FIELDS OF TYPE C

In case of gravitational null fields of type C all the eigenvalues of G_{ij} are zero and so also that of H_{ij}. The following equations are then satisfied:

$$G_{ij}G^{ij} = H_{ij}H^{ij}, \tag{2.1}$$

$$G_{ij}G^j_k G^{ki} = H_{ij}H^j_k H^{ki} = 0. \tag{2.2}$$

These equations are direct consequences of assumptions (A) and (B) of I.

For further discussion it is convenient to use the nonholonomic coordinate frame defined with the help of the unit eigenvectors e^i_μ of G_{ij} and u^i. The vectors

[1] R. M. Misra and R. A. Singh, J. Math. Phys. 7, 1836 (1966), referred to here as I.

[2] The signature of the space–time has been taken $+++-$. The range of small Latin indices is from 1 to 4 and those of Greek indices is from 1 to 3. Latin indices are used as tensor indices and Greek indices are used as labels. Summation convention is used throughout.

[3] In general, the left dual of R_{abcd} equals the right dual when and only when $S_{ab} \equiv R_{ab} - \frac{1}{4}g_{ab}R = 0$. This condition is obviously satisfied for vacuum gravitational field, i.e., $R_{ab} = 0$.

[4] The symbol ( ) has been used for symmetrization and the symbol [ ] for skew-symmetrization.

satisfy the following conditions

$$\underset{\alpha}{e_i}e^{\beta_i} = \delta^\beta_\alpha, \tag{2.3a}$$

$$\underset{\alpha}{e_i}u^i = 0, \tag{2.3b}$$

$$g_{ij} = \underset{\alpha}{e_i}\underset{\alpha}{e_j} - u_iu_j. \tag{2.4}$$

The nonholonomic components of an arbitrary tensor T_{ij} are then defined as

$$T_{\alpha\beta} = \underset{\alpha}{T_{ij}}\underset{\beta}{e^i}e^j. \tag{2.5}$$

We now state and prove the following theorem.

Theorem (2.1): If we define a tensor Q_{abcd} in the following manner

$$Q_{abcd} = R_{abpq}R^{pqrs}R_{rscd}, \tag{2.6}$$

then the necessary and sufficient condition that a vacuum space–time admits a null gravitational field of type C is that

$$Q_{abcd} = 0. \tag{2.7}$$

Proof: In order to establish this theorem, we first express the right-hand side of (3.6) in terms of G_{ij} and H_{ij} through Eq. (I.2.13). Then from equation (2.7), after contracting by u^au^l and simplifying, we obtain

$$G_{bq}\{G^{qs}G_{sd} - H^q_sH^s_d\} - H^j_b\{G_{jy}H^y_d + H^y_jG_{yd}\} = 0. \tag{2.8}$$

On further contraction of this equation by $\underset{1}{e^b}\underset{1}{e^d}$ and making use of the fact that

$$G_{11} = G_{22} = G_{33} = 0, \tag{2.9}$$

we get

$$G_{12}G_{13}G^{23} - G_{13}H^3_2H^2_1 - G_{12}H^2_3H^3_1 - H^3_1G_{23}H^2_1 = 0. \tag{2.10}$$

Similarly, contracting Eq. (2.8) by different pairs of $\underset{\alpha}{e^b}\underset{\beta}{e^d}$, we obtain a set of three more equations. Also Eqs. (2.1) and (2.2) are equivalent to

$$G^2_{12} + G^2_{23} + G^2_{31} = H^2_{12} + H^2_{23} + H^2_{31} \tag{2.11}$$

and

$$G_{12}G^{23}G_{31} = H_{12}H^{23}H_{31} = 0. \tag{2.12}$$

But these equations are not linearly independent. However, when these are solved simultaneously, we obtain a set of solutions, a typical one of which is

$$G^2_{12} = H^2_{13} = M^2 \quad \text{(say)}, \tag{2.13}$$

the remaining components of G and H being zero. This is equivalent to choosing a particular coordinate system in which

$$G_{ij} = \begin{pmatrix} 0 & M & 0 \\ M & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{2.14}$$

$$H_{ij} = \begin{pmatrix} 0 & 0 & -M \\ 0 & 0 & 0 \\ -M & 0 & 0 \end{pmatrix}. \tag{2.15}$$

The two three-tensors satisfy assumptions (A) and (B) and hence the statement of the theorem.

We now show that the gravitational field satisfying (2.7) admits a null vector in agreement with the properties of the gravitational null fields.[5]

Theorem (2.2): If the gravitational field is a null field of type C, then it defines a null vector given by

$$k_a = u_a - \underset{1}{e_a}. \tag{2.16}$$

The vector k_a satisfies the equation

$$R_{abc[d}k_{e]}k^c = 0. \tag{2.17}$$

Proof: According to the Debever's[6,7] theorem, the multiplicity of the null directions which a nonvanishing Riemann tensor may admit should be four, so we define the null directions as

$$k_a = A\underset{1}{e_a} + B\underset{2}{e_a} + C\underset{3}{e_a} + Du_a,$$

where A, B, C, and D are scalars whose values are to be determined. On substitution of this value of k_a in (2.17) and making use of (I.2.13), (2.1), (2.2), (2.14), and (2.15), we obtain a set of equations in the scalars A, B, etc. On solving these equations we easily find that

$$B = C = 0, \quad A = -D.$$

Thus we obtain

$$k_a = D(u_a - \underset{1}{e_a}).$$

However, the scalar D may be chosen to be unity without any loss of generality; hence the theorem.

This completes our study of null gravitational fields for an empty space–time. In the next section we consider the existence of null fields in a general space–time for which the energy–momentum tensor does not vanish.

[5] R. K. Sachs, Proc. Roy. Soc. (London) **A264**, 309 (1961).
[6] R. Debever, Bull Soc. Math. Belg. **10**, 112 (1959).
[7] R. Penrose, Ann. Phys. (N.Y.) **10**, 171 (1960).

3. NULL FIELDS IN NONEMPTY SPACE-TIME

We characterize the nonempty space-time with an energy-momentum tensor T_{ab} so that the field equation is

$$R_{ab} - \tfrac{1}{2}g_{ab}R = -T_{ab}, \tag{3.1}$$

where gravitational constant has been taken to be unity. Now, since in this case the field equation (3.1) is constructed with the help of the Ricci tensor, the metric tensor, and the energy-momentum tensor, it is no longer possible to base the investigations only on the properties of the curvature tensor. Therefore, we define an object used first by Petrov[8,9] called space-time-matter tensor in the following manner.

$$P_{abcd} = R_{abcd} + g_{a[d}T_{c]b} + g_{b[c}T_{d]a} + \sigma g_{a[d}g_{c]b}, \tag{3.2}$$

where σ is a coordinate dependent function. The tensor P_{abcd} has the symmetry properties of the Riemann tensor, i.e.,

$$P_{(ab)cd} = P_{ab(cd)} = P_{a[bcd]} = 0, \tag{3.3a}$$

$$P_{ab} = \omega g_{ab}. \tag{3.3b}$$

The necessity of introducing a coordinate dependent function σ is connected with the question of the uniqueness of the mass-energy tensor. It is because Eq. (3.1), satisfied by T^{ab} which has the following properties:

(a) T^{ab} is symmetric;

(b) $\nabla_b T^{ab} = 0$, where ∇ derivative is used for covariant differentiation;

(c) T^{ab} depends only on field potentials and g_{ab}; is also satisfied by T'^{ab}, where

$$T'^{ab} = T^{ab} + ag^{ab}$$

"a" being a scalar.[10]

Now, we define tensors $\bar{G}_{ab}$, $\bar{H}_{ab}$, and $\bar{K}_{ab}$ with the help of the equations

$$\bar{G}_{ac} = P_{abcd}u^b u^d, \tag{3.4}$$

$$\bar{H}_{ac} = {}^{*}P_{abcd}u^b u^d, \tag{3.5}$$

$$\bar{K}_{ac} = {}^{*}{}^{*}P_{abcd}u^b u^d, \tag{3.6}$$

where *operation has the usual meaning. The tensors $\bar{G}$, $\bar{H}$, and $\bar{K}$ are the analogs of G, H, and K respectively for a nonempty space-time. They satisfy

$$\bar{G}_{[ab]} = \bar{H}_{[ab]} = \bar{K}_{[ab]} = 0, \tag{3.7}$$

$$\bar{G}^a_a = -\bar{K}^a_a = \omega, \quad \bar{H}^a_a = 0, \tag{3.8}$$

where

$$\omega = \tfrac{3}{2}\sigma + T, \quad T = T^a_a. \tag{3.9}$$

It should be noted that Eq. (3.7) states that $\bar{G}$, $\bar{H}$, and $\bar{K}$ are symmetric, a consequence of the symmetry properties of P_{abcd}, and that (3.8) holds in view of (3.1). In general, the rank of $\bar{G}$, $\bar{H}$, and $\bar{K}$ is three and they lie in a space orthogonal to u^p. It may easily be seen that

$$P_{abcd} + {}^{*}{}^{*}P_{abcd} = 0. \tag{3.10}$$

In view of this equation we have the important relation

$$\bar{G}_{ab} + \bar{K}_{ab} = 0. \tag{3.11}$$

Now the tensor P_{abcd} in view of (3.3b) has ten algebraically different components. These are all contained in $\bar{G}$, $\bar{H}$, and $\bar{K}$. Thus the components of Riemann tensor have been partitioned into three symmetric tensors (3.4), (3.5), and (3.6). The Riemann tensor can be recovered from $\bar{G}$, $\bar{H}$, and $\bar{K}$ through the following equations

$$\begin{aligned}(P + i{}^{*}P)_{abcd} &= (g + i\epsilon)_{abpq}(g + i\epsilon)_{cdrs} \\ &\quad \times u^p u^r(\bar{G} + i\bar{H})^{qs} \\ &= -(g + i\epsilon)_{abpq}(g + i\epsilon)_{cdrs} \\ &\quad \times u^p u^r(\bar{K} - i\bar{H})^{qs}. \end{aligned} \tag{3.12}$$

We thus note that our formalism for the nonempty space-time is entirely analogous to that of vacuum fields. All the relations for nonempty Riemannian manifold are obtained from those of empty-space-time merely substituting P in place of R and $\bar{G}$, $\bar{H}$, and $\bar{K}$ for G, H, and K, respectively. But we further note one important difference, which is that, whereas G, H, and K for vacuum fields are all traceless, this property no longer holds true for $\bar{G}_{ab}$ and $\bar{K}_{ab}$. We observe from Eq. (3.8) that the trace of $\bar{G}_{ab}$ is equal to a scalar ω given by Eq. (3.9). In general this scalar is not zero. However, the choice of ω, σ, and T is connected with physical consideration. If the motion and distribution of the matter are given so that T_{ab} is known and if P_{abcd} is known for some specified choice of σ, then the curvature tensor is uniquely determined. Again, because of the freedom in choice of σ we can always make ω to vanish. However, for a general discussion we do not specify the choice of ω.

Now, for the characterization of gravitational null fields for a nonempty Riemannian manifold, we have to see whether assumptions (A) and (B) may also be satisfied in this case. If such a space-time

[8] A. Z. Petrov, Scientific Notices Kazan State University **114**, 55 (1954).

[9] A. Z. Petrov, in *Recent Developments in General Relativity* (Pergamon Press, Inc., New York, 1962), p. 371.

[10] V. Fock, in *Conference internationale sur les theories relativistes de la gravitation* (Pergamon Press, Inc., New York, 1954).

admits a null gravitational field then assumption (A) states that the eigenvalues of $\bar{G}_{ab}$ and $\bar{H}_{ab}$ should be equal, and assumption (B) states that their eigenvectors should have a required relative orientation. But we conclude from Eq. (3.8) that equality of eigenvalues of $\bar{G}_{ab}$ and $\bar{H}_{ab}$ cannot hold, and therefore assumption (A) cannot be satisfied. We hence have the following theorem.

Theorem (3.1): A nonempty Riemannian manifold for which Eqs. (3.1), (3.2), and (3.7)–(3.9) are satisfied cannot admit a gravitational null field in general.

Proof: In view of the above discussion the theorem is evident.

Thus we conclude that in this formalism a nonempty space–time with an arbitrary energy–momentum tensor does not admit null gravitational fields in general. However, we remarked earlier that, in particular, a proper choice of the coordinate system can be made and a given distribution and motion of matter may be considered for which the scalar ω vanishes. Equivalently, if trace of $\bar{G}_{ab}$ and $\bar{K}_{ab}$ is zero, then assumption (A) may be satisfied. In order to be more specific about the above statement, we consider an energy–momentum tensor whose trace vanishes and choose a coordinate system such that σ is zero. In particular, we investigate the following field equation

$$R_{ab} = -\mathcal{F}_{ab}, \tag{3.13}$$

where $\mathcal{F}_{ab}$ is the stress–energy tensor of the electromagnetic field defined in the following manner

$$\mathcal{F}_a^b = f_{ca}f^{cb} - \tfrac{1}{4}\delta_a^b f_{pq}f^{pq}, \tag{3.14}$$

where f_{ab} is the electromagnetic field tensor. In this case we obtain

$$\omega = F + \tfrac{3}{2}\sigma = 0, \tag{3.15}$$

$$P_{ab} = 0, \tag{3.16}$$

$$\bar{G}_a^a = \bar{H}_a^a = \bar{K}_a^a = 0, \tag{3.17}$$

where $F = \mathcal{F}_a^a = 0$. Thus in view of (3.8), (3.9), (3.15), and (3.16) assumptions (A) and (B) may be satisfied and we may characterize gravitational null fields in nonempty space–time exactly in the same manner as we did for vacuum fields. Accordingly, we characterize the gravitational field to be a null field of type A, B, and C, respectively, if:

(a) all the eigenvalues of $\bar{G}_{ij}$ and $\bar{H}_{ij}$ are nonvanishing;
(b) one of the eigenvalues of $\bar{G}_{ij}$ and $\bar{H}_{ij}$ is vanishing;
(c) all the eigenvalues of $\bar{G}_{ij}$ and $\bar{H}_{ij}$ are zero.

Thus the gravitational null fields in nonempty space–time may be discussed like those of vacuum null field if Eqs. (3.13), (3.15), and (3.16) remain satisfied. All the existence theorems for this case may be recovered from those of corresponding theorems for vacuum case merely by replacing R_{abcd} by P_{abcd}. We illustrate this with the help of the following theorem.

Theorem (3.2): If we define a tensor $\bar{Q}_{abcd}$ as

$$\bar{Q}_{abcd} = P_{abpq}P^{pqrs}P_{rscd}, \tag{3.18}$$

the necessary and sufficient condition that a nonempty space–time with vanishing curvature scalar admits a gravitational null field of type C, is that

$$\bar{Q}_{abcd} = 0 \tag{3.19}$$

or equivalently

$$(R - {}^{*}{}^{*}R)_{abpq}(R - {}^{*}{}^{*}R)^{pqrs}(R_{rscd} - {}^{*}{}^{*}R_{rscd}) = 0. \tag{3.20}$$

Proof: Proceeding on the lines of the proof of Theorem (2.1) we may easily establish this theorem. The equivalence of Eqs. (3.19) and (3.20) may be shown by recalling the well-known identity,

$$R_{abcd} + {}^{*}{}^{*}R_{abcd} = S_{ad}g_{bc} + S_{bc}g_{ad} - S_{ac}g_{bd} - S_{bd}g_{ac}. \tag{3.21}$$

In view of Eqs. (3.21) and (3.15) we get from (3.2)

$$P_{abcd} = \tfrac{1}{2}(R_{abcd} - {}^{*}{}^{*}R_{abcd}). \tag{3.22}$$

Making use of this equation in (3.18) and (3.19), it is a straightforward matter to obtain Eq. (3.20).

In conclusion it may be stressed that our formalism for characterization of gravitational null fields may be extended to any physical situation provided assumptions (A) and (B) remain satisfied.

ACKNOWLEDGMENT

The authors are grateful to Professor D. Sharma for constant encouragement throughout this investigation.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 5 MAY 1967

Realizations of Lie Algebras in Classical Mechanics

N. MUKUNDA*
Physics Department, Syracuse University, Syracuse, New York

(Received 4 October 1966)

Classical Poisson bracket realizations of semisimple Lie algebras are considered. An attempt is made to determine the minimum number of canonical degrees of freedom needed to find a realization of a given Lie algebra. Under the restriction to the symmetric traceless tensor representations of the orthogonal groups, and the symmetric tensor representations of the unimodular unitary groups, it is shown that with n pairs of canonical variables one can find realizations of the Lie algebras of $O(n+2)$ and $SU(n+1)$, but no higher groups.

INTRODUCTION AND SUMMARY

RECENT work on the group theoretical properties of classical dynamical systems has revealed some new and very interesting properties shared by large classes of systems.[1] It has long been known that the two most familiar systems in classical dynamics, the three-dimensional Kepler system and the three-dimensional isotropic harmonic oscillator, possess invariances under groups larger than the kinematical three-dimensional rotation group, these groups being, respectively, O_4, the real orthogonal group in four dimensions, and SU_3, the group of unitary, unimodular 3×3 matrices. The papers of Ref. 1 have demonstrated, however, that invariance under O_4 and SU_3 obtains for *all* classical Hamiltonians involving a centrally symmetric potential. Thus to a limited extent, and in the context of classical mechanics, the special status of the Keplerian and oscillator systems as the only ones possessing higher symmetry has been destroyed.

In a generalization of the above results, and using somewhat different methods, the present author has shown that, in fact, all classical Hamiltonians involving three degrees of freedom automatically possess invariance under both an O_4 and an SU_3 algebra, independent of the functional form of the Hamiltonian.[2] This result makes it clear that the possible symmetries of a classical Hamiltonian system are determined largely by the number of degrees of freedom of the system. Of course, it should be understood that the statement that a system with three degrees of freedom possesses SU_3 and O_4 invariances is, in general, only a statement about local properties. It implies only the existence of a set of constants of motion, which may not be real over all of phase space, but whose Poisson bracket algebra coincides with the Lie algebra of SU_3 or O_4. In particular, it may not be always possible to generate finite canonical transformations leaving the Hamiltonian invariant and furnishing a realization of the group SU_3 or O_4 as a whole.

In the present paper we examine in a preliminary way the connection between the number of degrees of freedom of a classical system, and the Lie algebras for which it may be possible to obtain realizations in terms of functions on the phase space of the system. We restrict attention to the classical semisimple compact Lie algebras.[3] In Sec. I, to show roughly how the arguments might go in general, we explain why a system with three degrees of freedom cannot possess invariance under the algebras of either the group O_5 or the group G_2 (these are, together with O_4 and SU_3, the only four compact semisimple Lie algebras of rank two). Section II deals with realizations of the algebras of the orthogonal rotation groups. Under restriction to the symmetric traceless tensor type representations, we prove that with n degrees of freedom we can obtain a realization of the Lie algebra of the group $O(n+2)$, but no higher orthogonal group. In Sec. III, under restriction to the symmetric tensor type representations, we show that with n degrees of freedom we can obtain a realization of the Lie algebra of the unitary unimodular group $SU(n+1)$ but no higher unitary group.

The analysis of more general realizations of the orthogonal and unitary groups, and realizations of the unitary symplectic groups, will be left to a later publication.

* On leave of absence from Tata Institute for Fundamental Research, Bombay, India.

[1] D. M. Fradkin, Progr. Theoret. Phys. (Kyoto) (to be published); H. Bacry, H. Ruegg, and J. M. Souriau, Commun. Math. Phys. 3, 323 (1966).

[2] N. Mukunda, Phys. Rev. 155, 1383 (1967).

[3] For a very useful review of the properties of these Lie algebras, see, for instance, R. E. Behrends, J. Dreitlein, C. Fronsdal, and B. W. Lee, Rev. Mod. Phys. 34, 1 (1962).

I. NON-INVARIANCE OF A SYSTEM WITH THREE DEGREES OF FREEDOM UNDER GROUP O_5 OR G_2

We recall here the main steps of the proof that all systems with three degrees of freedom possess SU_3 and O_4 symmetry.[2] Given a Hamiltonian $H(q_i, p_i)$ as a function of three q's and three p's, it is always possible to choose additional functions Ω, Q_1, Q_2, P_1, P_2 such that the following Poisson bracket relations hold[4]:

$$\begin{aligned} &\{H, \Omega\} \equiv \sum_{i=1}^{3} \left(\frac{\partial H}{\partial q_i} \frac{\partial \Omega}{\partial p_i} - \frac{\partial H}{\partial p_i} \frac{\partial \Omega}{\partial q_i} \right) = +1; \\ &\{H, Q_i\} = \{H, P_i\} = \{\Omega, Q_i\} = \{\Omega, P_i\} = 0; \\ &\{Q_i, Q_j\} = \{P_i, P_j\} = 0; \quad \{Q_i, P_j\} = \delta_{ij}. \end{aligned} \tag{1}$$

In other words, any H can be chosen as the first coordinate of a canonical set of variables, and then supplemented by additional variables to make up the whole set. (Of course, Ω and Q_i, P_i will not be unique.) Any function of q_i, p_i can (at least locally) be expressed as a function of H, Ω, Q_i, P_i; and further any time-independent constant of motion is a function of H, Q_i, P_i alone. It has been shown in Ref. 2 that one can find functions of H, Q_i, P_i whose Poisson bracket algebra coincides with the Lie algebra of O_4 or SU_3. Since both of these algebras are semisimple Lie algebras of rank two, one might wonder whether one can find functions of H, Q_i, P_i yielding realizations of the other two semisimple rank two Lie algebras, namely, O_5 and G_2. If this were possible, then every system with three degrees of freedom would exhibit invariance under O_5 and/or G_2, in addition to O_4 and SU_3. However, we show that this cannot be done.

Case of O_5: The Lie algebra of O_5 (ten elements) is made up of an O_4 algebra (six elements) and four more elements transforming as a four-vector with respect to the O_4 algebra. The Lie bracket relations of the O_4 algebra may be written

$$\begin{aligned} &\{M_i, M_j\} = \epsilon_{ijk} M_k; \quad \{N_i, N_j\} = \epsilon_{ijk} N_k. \\ &\{M_i, N_j\} = 0; \quad ijk = 1, 2, 3. \end{aligned} \tag{2}$$

Let us denote the remaining generators of O_5 by B_μ, $\mu = 1, 2, 3, 4$. B_μ has nonzero brackets with both M_i and N_i, transforming as a spin $\frac{1}{2}$ tensor under each $O(3)$ algebra. Now suppose we have a realization of the O_5 algebra via functions of H, Q_i, P_i, alone. (Note that for all practical purposes, H can be treated as a pure number.) Out of the elements of the O_4 algebra, we construct the functions

$$\begin{aligned} &\psi_1 = M_3, \quad \phi_1 = \tan^{-1}(M_1/M_2); \\ &\psi_2 = N_3, \quad \phi_2 = \tan^{-1}(N_1/N_2). \end{aligned} \tag{3}$$

One can check that they obey the Poisson bracket relations

$$\{\psi_i, \psi_j\} = \{\phi_i, \phi_j\} = 0; \quad \{\psi_i, \phi_j\} = \delta_{ij}. \tag{4}$$

Therefore the ψ_i and ϕ_i are four independent functions of Q_i, P_i, and in fact form a canonical set. Locally, Q_i, P_i can be expressed as functions of ψ_i, ϕ_i. Any function of Q_i, P_i having zero Poisson bracket with ψ_i, ϕ_i must also have zero bracket with Q_i, P_i, and is therefore a pure number (or a function of H alone). In particular this is true for the two variables

$$M^2 = M_1^2 + M_2^2 + M_3^2, \quad N^2 = N_1^2 + N_2^2 + N_3^2 \cdots. \tag{5}$$

But this in turn implies that M^2 and N^2 have zero bracket with the generators B_μ, which is impossible since the latter transform in a nontrivial way under the $O(3)$ transformations generated by M_i and N_i. [Alternatively, one may note that all the irreducible representations of the O_5 algebra by finite-dimensional Hermitian matrices (with commutator brackets for Lie brackets) are reducible into a sum of at least two irreducible matrix representations of the O_4 algebra, so that in these representations the operators M^2 and N^2 can never be constant multiples of the identity.][5] Therefore, it is not possible to construct a Poisson bracket realization of the O_5 algebra using functions of just two pairs of canonical variables.

Case of G_2: The 14-parameter Lie algebra of G_2 also contains the $O_4 \equiv O(3) \times O(3)$ algebra, and eight other elements that transform as a combined tensor of rank $\frac{3}{2}$ with respect to the first $O(3)$, and rank $\frac{1}{2}$ with respect to the second $O(3)$ in O_4. The arguments in this case are exactly parallel to the case of O_5, and show the impossibility of a Poisson bracket realization of G_2 with just two degrees of freedom.

At this point the following conjecture may be made. It is known that there cannot be more than n independent functions of n q's and n p's such that the Poisson bracket between any pair of these functions

[4] See, for instance, L. P. Eisenhart, *Continuous Groups of Transformations* (Dover Publications, Inc., New York, 1963), Chap. VI, pp. 281–291.

[5] For every value of n, a finite dimensional unitary irreducible matrix representation of the group $O(2n+1)$ contains at least two irreducible representations of the subgroup $O(2n)$, $O(2n)$ being embedded in $O(2n+1)$ in the usual canonical way. For details, see I. M. Gel'fand and M. L. Tseitlin, Dokl. Akad. Nauk. SSR 71, 825, 1017 (1950).

vanishes.[6] At most one can find sets of n such functions. This suggests that if the rank of a semisimple Lie algebra exceeds n, one cannot find a realization of it using functions of n q's and n p's only. A weakened form of this conjecture is verified for the unitary and real orthogonal groups in the next two sections.

II. ALGEBRAS OF ORTHOGONAL ROTATION GROUP FOR SYSTEMS WITH n DEGREES OF FREEDOM

The real orthogonal group in $(n+2)$ dimensions, $O(n+2)$, has $\frac{1}{2}(n+1)(n+2)$ generators and is of rank $\frac{1}{2}(n+2)$ or $\frac{1}{2}(n+1)$ according as n is even or odd. The generators form the components of a real antisymmetric second-rank tensor $J_{\mu\nu}$, $\mu, \nu = 1, 2, \cdots, n+2$. The Lie bracket relations are

$$\{J_{\mu\nu}, J_{\lambda\sigma}\} = \delta_{\mu\lambda}J_{\nu\sigma} - \delta_{\nu\lambda}J_{\mu\sigma} + \delta_{\mu\sigma}J_{\lambda\nu} - \delta_{\nu\sigma}J_{\lambda\mu}. \quad (6)$$

It is convenient to separate the $J_{\mu\nu}$ into the generators, J_{rs}, $r, s = 1 \cdots n$, of the subgroup $O(n)$, and the rest:

$$J_{r,n+1} = A_r; \quad J_{r,n+2} = B_r; \quad J_{n+1,n+2} = S. \quad (7)$$

The bracket relations (6) take the form

$$\begin{aligned}
&\{J_{rs}, J_{tu}\} = \delta_{rt}J_{su} - \delta_{st}J_{ru} + \delta_{ru}J_{ts} - \delta_{su}J_{tr};\\
&\{J_{rs}, A_t\} = \delta_{rt}A_s - \delta_{st}A_r; \quad \{A_r, A_s\} = J_{rs};\\
&\{J_{rs}, B_t\} = \delta_{rt}B_s - \delta_{st}B_r; \quad \{A_r, B_s\} = \delta_{rs}S;\\
&\{J_{rs}, S\} = 0; \quad \{A_r, S\} = -B_r;\\
&\{B_r, B_s\} = J_{rs};\\
&\{B_r, S\} = A_r.
\end{aligned} \quad (8)$$

A solution for these generators in terms of n q's and n p's may be constructed as follows:

$$\begin{aligned}
J_{rs} &= q_r p_s - q_s p_r\\
A_r &= (m^2 - p^2)^{\frac{1}{2}} q_r\\
B_r &= (\alpha^2 - J^2)^{\frac{1}{2}} p_r/m\\
S &= (\alpha^2 - J^2)^{\frac{1}{2}} (m^2 - p^2)^{\frac{1}{2}}/m\\
p^2 &= p_r p_r; \quad J^2 = \tfrac{1}{2}J_{rs}J_{rs} + A_r A_r = m^2 q^2 - (q_r p_r)^2.
\end{aligned} \quad (9)$$

Here, m and α are any two real numbers. These expressions may be verified to obey the bracket relations, Eq. (8). This solution is obtained as follows: We start with the expressions for the generators of the inhomogeneous Lorentz group in three space and one time dimension, given in terms of three q's and three p's.[7] These expressions can be modified and extended to $(n+1)$ dimensions to yield generators for the Euclidean group in $(n+1)$ dimensions, $E(n+1)$, in terms of n q's and n p's. The $E(n+1)$ generators consist of the $O(n+1)$ generators, (J_{rs} and A_r), together with $(n+1)$ generators (analogous to B_r and S) that have zero bracket with one another and transform as a vector under $O(n+1)$. Denoting the Casimir invariant of $O(n+1)$ by J^2, we multiply all the components of this $(n+1)$ vector by a common function of J^2, to get a new vector under $O(n+1)$, namely B_r and S. We determine this function by requiring that the Poisson brackets of B_r and S with themselves give back J_{rs}, A_r according to Eq. (8).

The question now is the following: Given the generators $J_{\mu\nu}$ of $O(n+2)$, obeying the Poisson bracket relations (6), can one define n variables q_i and n variables p_i as functions of $J_{\mu\nu}$, such that the Poisson bracket relations of the $J_{\mu\nu}$ imply that the q_i and p_i are canonical variables? We can show that this is possible, provided that the given realization of $J_{\mu\nu}$ belongs to the symmetric traceless tensor type representations of $O(n+2)$. This means that the following identity among the $J_{\mu\nu}$ holds[8]:

$$J_{\mu\nu}J_{\lambda\sigma} + J_{\mu\lambda}J_{\sigma\nu} + J_{\mu\sigma}J_{\nu\lambda} = 0. \quad (10)$$

[The solution given in Eq. (9) obeys Eq. (10).] In this case, define

$$\begin{aligned}
q_r &= (B^2 + S^2)^{\frac{1}{2}}(A_r/S); \quad p_r = B_r/(B^2 + S^2)^{\frac{1}{2}},\\
B^2 &= B_r B_r.
\end{aligned} \quad (11)$$

Using Eq. (8) above, one finds

$$\begin{aligned}
\{q_r, q_s\} &= \frac{(B^2 + S^2)}{S^3}\\
&\quad \times [J_{rs}J_{n+1,n+2} + J_{r,n+1}J_{n+2,s} + J_{r,n+2}J_{s,n+1}],\\
\{p_r, p_s\} &= \sum_{a=1}^{n+1} \frac{J_{a,n+2}}{(B^2 + S^2)^2}\\
&\quad \times [J_{rs}J_{a,n+2} + J_{ra}J_{n+2,s} + J_{r,n+2}J_{sa}],\\
\{q_r, p_s\} &= \delta_{rs} + \frac{A_r B_t}{S^2(B^2 + S^2)}\\
&\quad \times [J_{ts}J_{n+1,n+2} + J_{t,n+1}J_{n+2,s} + J_{t,n+2}J_{s,n+1}].
\end{aligned} \quad (12)$$

If we now use (10), we see that q_r, p_s form a canonical set.

This demonstration also proves that, subject to the restriction to symmetric traceless tensor representations, we cannot have realizations of orthogonal groups larger than $O(n+2)$ by means of functions of n q's and p's only. This can be seen in the following way. Notice that the functional forms of q_r and p_r in terms of $J_{\mu\nu}$, as given in Eq. (11) ensure that those functions of $J_{\mu\nu}$ always make sense and cannot be identically zero or infinite (hence meaningless) everywhere, as long as the $J_{\mu\nu}$ are real. Therefore, if we had a realization of the generators of $O(N)$, say, in terms of n pairs of canonical variables, we could recover from these generators exactly $N-2$ pairs of canonical variables by means of Eq. (11); the

[6] L. P. Eisenhart, Ref. 4, p. 283.

[7] See, for instance, L. L. Foldy, Phys. Rev. **102**, 568 (1956).

[8] The simplest solution for $J_{\mu\nu}$ using $(n+2)$ coordinates and momenta, $J_{\mu\nu} = q_\mu p_\nu - q_\nu p_\mu$, obeys this identity. On the other hand, this solution of $J_{\nu\mu}$ gives rise to the spherical harmonic type representations of $O(n+2)$, which is just the set of all traceless symmetric tensor representations of $O(n+2)$.

point is that, in spite of special identities that may happen to hold among the $J_{\mu\nu}$ in the particular realization these $N-2$ pairs of canonical variables cannot fail to exist. It follows that $N-2 \leq n$, or $N \leq n+2$. By a similar argument, one can see that any two such realizations of $O(n+2)$ in terms of n pairs of canonical variables can be transformed into one another by a suitable canonical transformation.

III. ALGEBRAS FOR n DEGREES OF FREEDOM RESTRICTED TO SYMMETRIC TENSOR REPRESENTATIONS

The unitary unimodular group in $(n+1)$ dimensions, $SU(n+1)$, has $n(n+2)$ generators and is of rank n. Its generators may be written B_α^β, and obey

$$\begin{gathered}(B_\alpha^\beta)^* = B_\beta^\alpha; \quad B_\alpha^\alpha = 0; \\ -i\{B_\alpha^\beta, B_\lambda^\mu\} = \delta_\lambda^\beta B_\alpha^\mu - \delta_\alpha^\mu B_\lambda^\beta; \\ \alpha\beta = 1, \cdots, n+1.\end{gathered} \tag{13}$$

It is convenient to split B_α^β into the generators, A_i^j, of the $SU(n)$ subgroup, and the rest

$$\begin{gathered}B_{n+1}^{n+1} = B; \quad B_i^{n+1} = B_i; \quad B_{n+1}^i = B^i; \\ B_i^j = A_i^j - \delta_i^j(B/n); \quad i, j, \cdots = 1, \cdots, n.\end{gathered} \tag{14}$$

In terms of these, we have

$$\begin{aligned}-i\{A_i^j, A_k^l\} &= \delta_k^j A_i^l - \delta_i^l A_k^j; \\ -i\{A_i^j, B_k\} &= \delta_k^j B_i - \delta_i^j(B_k/n); \\ -i\{A_i^j, B^k\} &= -\delta_i^k B^j + \delta_i^j(B^k/n); \quad \{A_i^j, B\} = 0; \\ -i\{B_j, B\} &= B_j; \quad -i\{B^j, B\} = -B^j; \\ -i\{B_j, B^k\} &= A_j^k - [(n+1)/n]\delta_j^k B; \\ \{B_i, B_j\} &= \{B^i, B^j\} = 0.\end{aligned} \tag{15}$$

A simple (harmonic oscillator type) solution in terms of n q's and n p's is obtained as follows: Instead of q_i, p_i we use the variables

$$a_j = (p_j - iq_j)/\sqrt{2}, \quad a_j^* = (p_j + iq_j)/\sqrt{2}. \tag{16}$$

The canonical Poisson bracket rules are

$$\begin{gathered}\{a_j, a_k\} = \{a_j^*, a_k^*\} = 0; \\ \{a_j, a_k^*\} = -i\delta_j^k.\end{gathered} \tag{17}$$

The solution is then

$$\begin{aligned}A_i^j &= a_j^* a_i - (N/n)\delta_i^j; \\ B_i &= (\alpha - N)^{\frac{1}{2}} a_i; \quad B^i = (\alpha - N)^{\frac{1}{2}} a_i^*; \\ B &= [n/(n+1)]\alpha - N; \\ N &= \sum_1^n a_j^* a_j.\end{aligned} \tag{18}$$

Here α is any real positive number.

Once again it is possible to reverse the procedure and, given generators B_α^β of $SU(n+1)$, to define variables p_i, q_i which are canonical, provided the B_α^β obey identities characteristic of the symmetric tensor representations. In this case, the identities may be written in the following way: Let $C_2 = B_\alpha^\beta B_\beta^\alpha$ be the quadratic Casimir invariant of $SU(n+1)$; define a variable Ω_1 by

$$C_2 = n(n+1)\Omega_1^2. \tag{19}$$

Then the relevant identities are[9]

$$\begin{aligned}B_\alpha^\beta B_\mu^\lambda - B_\alpha^\lambda B_\mu^\beta &= \Omega_1^2(\delta_\alpha^\lambda \delta_\mu^\beta - \delta_\alpha^\beta \delta_\mu^\lambda) \\ &\quad + \Omega_1(B_\alpha^\lambda \delta_\mu^\beta + B_\mu^\beta \delta_\alpha^\lambda - B_\alpha^\beta \delta_\mu^\lambda - B_\mu^\lambda \delta_\alpha^\beta).\end{aligned} \tag{20}$$

The a_i and a_i^* can be defined by

$$a_i = B_i/(B + \Omega_1)^{\frac{1}{2}}, \quad a_i^* = B^i/(B + \Omega_1)^{\frac{1}{2}} \tag{21}$$

as functions of the $SU(n+1)$ generators. Using (15) alone, we find

$$\begin{aligned}\{a_i, a_j\} &= \{a_i^*, a_j^*\} = 0; \\ \{a_i, a_j^*\} &= [-i/(B + \Omega_1)^2]\{B_i B^j - (B + \Omega_1)A_i^j \\ &\quad + [(n+1)/n]B(B + \Omega_1)\delta_i^j\}.\end{aligned} \tag{22}$$

The identity (20), for the case $\alpha = i \leq n$, $\lambda = j \leq n$, $\beta = \mu = n+1$, gives

$$B_i B^j = (B + \Omega_1)\{A_i^j + \delta_i^j[\Omega_1 - (B/n)]\}. \tag{23}$$

Using this in (22) reduces the right-hand side of the last Poisson bracket to $-i\delta_i^j$.

Thus from the algebra of the symmetric tensor representations of $SU(n+1)$ we can recover $2n$ canonical variables. At the same time, this proves that as far as such representations are concerned, one cannot go beyond the $SU(n+1)$ algebra using $2n$ canonical variables. The arguments are similar to those given at the end of the previous section; they hinge on the fact that the functions of the generators B_α^β appearing on the right-hand sides of Eq. (21) will always make sense, since neither B_i nor $B + \Omega_1$ can be identically zero in a nontrivial realization of $SU(n+1)$. Finally we see in the same way as before that any two such realizations of $SU(n+1)$ using n pairs of canonical variables are canonically equivalent.

ACKNOWLEDGMENTS

The author would like to thank Professor E. C. G. Sudarshan for many fruitful discussions.

This work was supported in part by the U.S. Atomic Energy Commission.

[9] The simplest solution for B_α^β using $(n+1)$ coordinates and momenta, $B_\alpha^\beta = a_\beta^* a_\alpha - [\delta_\alpha^\beta/(n+1)]a_\gamma^* a_\gamma$, obeys this identity. On the other hand, this solution for B_α^β gives rise to all the states of the isotropic $(n+1)$-dimensional harmonic oscillator, which is just the set of all symmetric tensor representations of $SU(n+1)$.

Properties of a Harmonic Crystal in a Stationary Nonequilibrium State*

Z. RIEDER, J. L. LEBOWITZ, AND E. LIEB†
Belfer Graduate School of Science, Yeshiva University, New York, New York

(Received 14 July 1966)

The stationary nonequilibrium Gibbsian ensemble representing a harmonic crystal in contact with several idealized heat reservoirs at different temperatures is shown to have a Gaussian Γ space distribution for the case where the stochastic interaction between the system and heat reservoirs may be represented by Fokker–Planck-type operators. The covariance matrix of this Gaussian is found explicitly for a linear chain with nearest-neighbor forces in contact at its ends with heat reservoirs at temperatures T_1 and T_N, N being the number of oscillators. We also find explicitly the covariance matrix, but not the distribution, for the case where the interaction between the system and the reservoirs is represented by very "hard" collisions. This matrix differs from that for the previous case only by a trivial factor. The heat flux in the stationary state is found, as expected, to be proportional to the temperature difference $(T_1 - T_N)$ rather than to the temperature gradient $(T_1 - T_N)/N$. The kinetic temperature of the jth oscillator $T(j)$ behaves, however, in an unexpected fashion. $T(j)$ is essentially constant in the interior of the chain *decreasing* exponentially in the direction of the *hotter* reservoir rising only at the end oscillator in contact with that reservoir (with corresponding behavior at the other end of the chain). No explanation is offered for this paradoxical result.

1. INTRODUCTION

IN a series of papers[1–3] Lebowitz and Bergmann developed a general formalism for describing the time evolution of a Gibbs ensemble representing a system in contact with one or more idealized heat reservoirs (temperature baths). They imagine the reservoirs made up of an infinite number of identical noninteracting components each of which interacts with the system at most once. This interaction is impulsive and it is assumed that prior to this interaction the components of each reservoir have an equilibrium distribution with some specified temperature T_α, where $\alpha = 1, \cdots, n$, specifies the different reservoirs. Under these conditions the Γ space ensemble density of the system $\mu(x, t)$ satisfies the generalized Liouville equation

$$\frac{\partial \mu(x,t)}{\partial t} + (\mu, H) = \sum_{\alpha=1}^{n} \int [K_\alpha(x, x')\mu(x', t) - K_\alpha(x', x)\mu(x, t)]\, dx'. \quad (1.1)$$

Here $x = (\mathbf{q}_1, \cdots, \mathbf{q}_N, \mathbf{p}_1, \cdots, \mathbf{p}_N)$ is a point in the phase space of the system, $H(x)$ is the Hamiltonian of the system, (μ, H) is the Poisson bracket between μ and H, and the right side of (1.1) represents the effect of collisions with reservoir components on the evolution of μ. $K(x, x')\, dx\, dt$ is the conditional probability that when the system is at the point x' in its Γ space it will suffer a collision in the time interval dt as a result of which it will jump to the region $(x, x + dx)$.

Under very general conditions $\mu(x, t)$ approaches, as $t \to \infty$, a stationary distribution $\mu_s(x)$. This stationary distribution will correspond to the system being in equilibrium if the temperature of all the reservoirs is the same; otherwise $\mu_s(x)$ will represent a stationary nonequilibrium state in which there are heat currents flowing through the system. (More general nonequilibrium situations may also be represented in this manner.[1,4]) It is to be expected for a *physical* system of macroscopic size, whose interaction with the heat reservoirs is confined to specified "surface regions," that its bulk properties in the stationary state will depend only on the temperature of the reservoirs and not on the details of the interaction (this, of course, is expected to be true when the reservoirs all have the same temperature); e.g., the properties of a "long" metal bar should not depend on whether its ends are in contact with water or with wine "heat reservoirs" at temperature T_1 and T_2. (We are assuming here "good" heat contact between reservoirs and system so that regions of the system in direct contact with a given reservoir are essentially at the "temperature" of that reservoir.) This belief justifies the idealization of the reservoirs already made in deriving (1.1) and the further drastic simplification made below, and thus, we expect, for realistic systems,

* Based in part on a Ph.D thesis submitted by Z. Rieder to Yeshiva University.

† Present address: Physics Department, Northeastern University, Boston, Massachusetts.

[1] P. G. Bergmann and J. L. Lebowitz, Phys. Rev. 99, 578 (1955); J. L. Lebowitz and P. G. Bergmann, Ann. Phys. (N.Y.) 1, 1 (1957).

[2] J. L. Lebowitz, Phys. Rev. 114, 1192 (1959).

[3] J. L. Lebowitz, Rend. Sculola Intern. Fis. XIV Corso Bologna, Italy (1961).

[4] J. L. Lebowitz and A. Shimony, Phys. Rev. 128, 1945 (1962); E. P. Gross and J. L. Lebowitz, *ibid.* 104, 1528 (1956).

that the stationary state found from our model will correctly represent, in the Gibbs ensemble sense, such a physical system in a steady nonequilibrium state.

To obtain an explicit simple form for the right side of (1.1) we imagine the system to contain at its surface n pistons of mass M_α. The αth reservoir will consist of point particles of mass m_α at uniform densities ρ_α *always* having a Maxwellian velocity distribution at temperature T_α prior to a collision with the αth piston. During such an elastic collision there will be an exchange of momentum in some specified direction. Under these conditions the kernel $K_\alpha(x, x')$ may be specified explicitly [cf. Eq. (2.3), Ref. 2]. Still further simplification is achieved when $m_\alpha \ll M_\alpha$ so that the piston velocity is changed very little during a collision. The effect of the collisions with the reservoirs on the time evolution of $\mu(x, t)$ may then be represented by a Fokker–Planck-type term,[2,5] and (1.1) assumes the form

$$\frac{\partial \mu(x,t)}{\partial t} + (\mu, H) = \sum_{\alpha=1}^{n} \lambda_\alpha \frac{\partial}{\partial P_\alpha}\left[P_\alpha \mu + kT_\alpha M_\alpha \frac{\partial}{\partial P_\alpha}\mu\right]. \tag{1.2}$$

Here (Q_α, P_α) are the coordinates and momentum of the αth piston (the pistons being part of the system) and λ_α is the "friction constant" of the αth piston given by [Eq. (3.3), Ref. 2]

$$\lambda_\alpha = \rho_\alpha A_\alpha (8 m_\alpha k T_\alpha / \pi M_\alpha^2)^{\frac{1}{2}}, \tag{1.3}$$

where A_α is the collision cross section (or area) of the αth piston. It is easy to show[2] that $\mu(x, t)$ satisfying (1.2) will in general approach, as $t \to \infty$, a stationary value $\mu_s(x)$.

Up to now we have not specified the nature of our system which determines $H(x)$. We now consider the case where our system is a harmonic crystal and the pistons are just some of the particles of the system. (Their location need not be specified at the moment.) It is then shown in Sec. 2 that the stationary solution of (1.2), which $\mu(x, t)$ will approach as $t \to \infty$, is a Gaussian in the coordinates and momenta of the system (corresponding to the canonical distribution when the temperatures of all the reservoirs are equal; $T_\alpha = T$). The explicit form of the stationary distribution, i.e., the covariance matrix of the Gaussian, is found in Sec. 3 for the special case of a one-dimensional crystal of N particles with nearest-neighbor interactions in which the first particle is in contact with a reservoir at temperature T_1 and the last with a reservoir at temperature T_N and $\lambda_1 = \lambda_N$. The general form of the distribution for this system, including its time dependence, has also been discussed independently by Bils.[6] Since we are interested in this system solely as a model, we do not worry about the drastic simplifications made in the right side of (1.1) to arrive at (1.2) or the further one, $\lambda_1 = \lambda_N$, required to obtain an explicit stationary nonequilibrium ensemble in Γ space. What is unfortunate, however, is that the harmonic crystal is *not* a realistic physical system. As is well known,[7] the harmonic crystal has an "infinite" heat conductivity; i.e., the heat flux is not proportional to the temperature gradient when one considers the relaxation of this system from some initial nonequilibrium state. This is reflected in the true stationary state considered here by the fact that the heat flux is proportional (when $N \gg 1$, or strictly speaking in the limit $N \to \infty$), to the temperature difference between the ends of the system, $(T_1 - T_N)$, rather than to the temperature gradient $(T_1 - T_N)/N$, which would be the case if there was any anharmonic coupling. This is also reflected in the form of $T(j)$, the kinetic temperature of the jth harmonic oscillator which is uniform throughout the linear chain, being equal to $\frac{1}{2}(T_1 + T_N)$, except near the edges where it varies exponentially in a backward way; i.e., with $T_1 > T_N$, $T(j)$ will *decrease* from its mean value as $j \to 1$ jumping to a higher value, close to T_1, for $j = 1$. Also the heat flux, $J(\lambda)$, will vary with λ, the strength of the coupling to the reservoirs, in an unphysical way, reaching a maximum at $\lambda = \frac{1}{2}\sqrt{3}\,\omega$, (where $m\omega^2$ is the force constant between the oscillators), and vanishing as ω^2/λ for $\lambda \to \infty$. This may perhaps be understood as a mismatching between the frequencies of the reservoirs and the oscillators. We have no explanation for the abnormal behavior of $T(j)$.

An alternate idealization of the stochastic interaction between the reservoirs and the system is to imagine that after *each* collision with a component of the αth reservoir the momentum P_α will have a Maxwellian distribution at the temperature T_α,

$$h_\alpha(P) = (M_\alpha k T_\alpha / 2\pi)^{\frac{1}{2}} \exp\left[-P^2/2M_\alpha k T_\alpha\right]. \tag{1.4}$$

This is an opposite extreme of the small momentum transfer considered before and corresponds to the pistons and reservoir components having the same mass.[1,2] We simplify this further by assuming that

[5] J. L. Lebowitz and P. Resibois, Phys. Rev. 139, A1101 (1965).

[6] O. Bils, "On the Non-Stationary Equilibrium of a Finite Chain of Coupled Oscillators" (to be published).

[7] G. Klein and I. Prigogine, Physica 19, 1053 (1953); P. C. Hemmer, Kgl. Norske Videnskab. Selskab Fork. 33, 101 (1960); E. I. Takizawa and K. Kobayasi, Chinese J. Phys. 1, 59 (1963); E. Teramoto, Progr. Theoret. Phys. (Kyoto) 28, 1059 (1962); R. Rubin, Phys. Rev. 131, 964 (1963).

the probability of a collision with a reservoir component in a time interval dt is given by $\lambda_\alpha' dt$, independent of the state of the system. These assumptions lead to a modified Krook type of collision kernel[4] and Eq. (1.2) assumes the form

$$\frac{\partial \mu(x,t)}{\partial t} + (\mu, H) = \sum_\alpha \lambda_\alpha' \Big\{ h_\alpha(P_\alpha) \int \mu(x,t)\, dP_\alpha - \mu(x,t) \Big\}. \quad (1.5)$$

While the stationary solution of (1.5) for a harmonic crystal is no longer a Gaussian the stationary covariance matrix of the linear chain and hence the kinetic temperature and heat flux, is of the same form as before. The only change is that the system acts as if the temperature difference $(T_1 - T_N)$ was reduced by the factor $[1 + (\omega^2/\lambda^2)\varphi_1]$, where φ_1 depends on ω and λ. These assertions about the covariance matrix are proved in Sec. 3.

2. STATIONARY STATE OF A HARMONIC CRYSTAL

The Hamiltonian of a harmonic crystal containing $\mathcal{N}$ particles, each being s dimensional, may be written in the general form[8]

$$H = \tfrac{1}{2}\sum_{i=N}^{2N} x_i^2 + \tfrac{1}{2}\sum_{i,j=1}^{N} \Phi_{ij} x_i x_j; \quad N = s\mathcal{N}. \quad (2.1)$$

Here the x_i, $i = 1, \cdots, N$, are the Cartesian coordinates of the particles, (relative to their equilibrium positions), while x_j, $j = i + N$, is the momentum conjugate to x_i (we have set the mass of the particles equal to unity). The generalized Liouville equation (1.2) now has the form

$$\frac{\partial \mu(x,t)}{\partial t} = \sum_{i=1}^{2N} \frac{\partial}{\partial x_i}(\xi_i \mu) + \tfrac{1}{2}\sum_{i,j=1}^{2N} \frac{\partial^2}{\partial x_i \partial x_j}(d_{ij}\mu), \quad (2.2)$$

where

$$\xi_i = \sum_{j=1}^{2N} a_{ij} x_j \quad (2.3)$$

and a_{ij} and d_{ij} are elements of $2N$ by $2N$ matrices $\mathbf{a}$ and $\mathbf{d}$ which we write in the partitioned form

$$\mathbf{a} = \begin{pmatrix} \mathbf{0} & -\mathbf{I} \\ \mathbf{\Phi} & \mathbf{\mathcal{R}} \end{pmatrix}, \quad \mathbf{d} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\epsilon} \end{pmatrix}. \quad (2.4)$$

Here $\mathbf{0}$ and $\mathbf{I}$ are the null and unit N by N matrices, Φ_{ij} is defined in (2.1), $\mathcal{R}_{ij} = \lambda_\alpha \delta_{\alpha i}\delta_{ij}$ [λ_α given in (1.3) with $M_\alpha = 1$] and $\epsilon_{ij} = 2kT_i\mathcal{R}_{ij}$. The general time-dependent solution of (2.2) may be found[6] by diagonalizing the right side of (2.2) as was done by Wang and Uhlenbeck[9] for fluctuations in electrical circuits. (Wang and Uhlenbeck consider only the case corresponding to all the T_α being the same.) It is clear, however, from an inspection of (2.2) that its stationary solution μ_s [corresponding to setting $\partial\mu/\partial t = 0$ in (2.2)], which is all that is of interest to us in this problem, has the general form

$$\mu_s(x) = (2\pi)^{-N} \operatorname{Det}[\mathbf{b}^{-\frac{1}{2}}] \exp\Big[-\tfrac{1}{2}\sum_{i,j=1}^{2N} b_{ij}^{-1} x_i x_j \Big]. \quad (2.5)$$

The matrix $\mathbf{b}$ is the positive definite covariance matrix, and is related to expectation values in the stationary state by

$$b_{ij} = \langle x_i x_j \rangle = \int \mu_s(x) x_i x_j\, dx \quad (2.6)$$

and we have

$$A_i = \langle x_i \rangle = \int \mu_s(x) x_i\, dx = 0. \quad (2.7)$$

Substituting (2.5) into (2.2) and equating terms yields the basic, necessary, and sufficient equation

$$\mathbf{a}\cdot\mathbf{b} + \mathbf{b}\cdot\mathbf{a}^\dagger = \mathbf{d}, \quad (2.8)$$

where $\mathbf{a}^\dagger$ is the transpose of $\mathbf{a}$. Once $\mathbf{b}$ is known all the properties of the stationary state, e.g., heat flux, local kinetic temperature, etc., are readily available. [It is clear that when all the $T_\alpha = T$ then $\boldsymbol{\epsilon} = 2kT\mathcal{R}$ and $\mu_s(x) \sim e^{-\beta H(x)}$, $\beta = (kT)^{-1}$; i.e, the stationary state is the equilibrium state at temperature T.]

The uniqueness of the stationary solution $\mu_s(x)$ for the case where the coupling with the reservoirs does not vanish and the phase space of the crystal is not divided into different isolated parts (i.e., the representative phase point of the system can move between any two regions via a combination of its natural motion and collision with the reservoirs) follows from the general results of Ref. 1, explicitly verifiable here, that an *arbitrary* initial distribution will approach a unique $\mu_s(x)$ as $t \to \infty$. For the harmonic crystal in which there are no "torn bonds" isolating some parts this condition of ergodicity is clearly satisfied. The uniqueness of μ_s for the linear chain is shown explicitly in the next section.

Equations (2.7) and (2.8) are consequences of the general equations satisfied by the time-dependent expectation values $A_i(t)$ and $b_{ij}(t)$ defined with $\mu_s(x) \to \mu(x,t)$ in (2.6)–(2.7). We then have from (1.2)

$$(d/dt)\mathbf{A}(t) = -\mathbf{a}\cdot\mathbf{A}(t) \quad (2.9)$$

[8] Cf., for example, A. A. Maradudin, E. W. Montroll, and G. H. Weiss, *Theory of Lattice Dynamics in the Harmonic Approximation* (Academic Press Inc., New York, 1963); E. W. Montroll, Third Berkeley Symp. Math. Stat. and Prob. 3, 209 (1957).

[9] M. C. Wang and G. E. Uhlenbeck, Rev. Mod. Phys. 17, 323 (1945).

and

$$(d/dt)\mathbf{b}(t) = \mathbf{d} - \mathbf{a}\cdot\mathbf{b}(t) - \mathbf{b}(t)\cdot\mathbf{a}^\dagger. \quad (2.10)$$

For the case where $\mu(x, t)$ satisfies Eq. (1.5) the expectation values of the coordinates and momenta $\mathbf{A}'(t)$ again satisfy (2.9) (with λ_α replaced by λ'_α) while the covariance matrix $\mathbf{b}'(t)$ now satisfies the equation

$$(d/dt)\mathbf{b}'(t) = \tfrac{1}{2}\mathbf{d} - \mathbf{a}\cdot\mathbf{b}'(t) - \mathbf{b}'(t)\cdot\mathbf{a}^\dagger + \mathbf{r}\cdot\mathbf{b}'(t)\cdot\mathbf{r} \quad (2.11)$$

with

$$\mathbf{r} = \begin{pmatrix} 0 & 0 \\ 0 & \mathcal{R}^{\frac{1}{2}} \end{pmatrix}. \quad (2.12)$$

In the stationary state $\mathbf{A}'$ and $\mathbf{A}$ again vanish while $\mathbf{b}'$ or $\mathbf{b}$ satisfy (2.11) or (2.10) with the left sides set equal to zero.

3. EXPLICIT SOLUTION FOR A LINEAR CHAIN

We consider now a one-dimensional harmonic crystal (chain of pistons) with nearest-neighbor interactions, whose ends are rigidly fixed.[8] The interaction with the reservoirs takes place at the first and last piston, $\alpha = 1, N$ and we set

$$\lambda_1 = \lambda_N = \lambda, \quad T_1 = T(1+\eta), \quad T_N = T(1-\eta); \quad |\eta| \le 1. \quad (3.1)$$

The N by N matrices $\mathbf{\Phi}$, $\mathcal{R}$, and $\boldsymbol{\epsilon}$ now have the form

$$\mathbf{\Phi} = \omega^2\mathbf{G}; \quad G_{ij} = 2\delta_{ij} - \delta_{i+1,j} - \delta_{i,j+1} \quad (\text{for, } j = 1 \text{ through } N-1), \quad (3.2)$$

$$G_{Nj} = G_{jN} = \begin{cases} 0, & j < N-1, \\ -1, & j = N-1, \\ 2, & j = N, \end{cases}$$

$$\mathcal{R} = \lambda\mathbf{R}; \quad \mathbf{R}_{ij} = \delta_{ij}(\delta_{i1} + \delta_{iN}), \quad (3.3)$$

$$\boldsymbol{\epsilon} = 2kT\lambda(\mathbf{R} + \eta\mathbf{E}); \quad E_{ij} = \delta_{ij}(\delta_{i1} - \delta_{iN}). \quad (3.4)$$

We now write the $2N$ by $2N$ covariance matrix $\mathbf{b}$ in the partitioned form

$$\mathbf{b} = \begin{pmatrix} \mathbf{x} & \mathbf{z} \\ \mathbf{z}^\dagger & \mathbf{y} \end{pmatrix}. \quad (3.5)$$

The N by N matrices $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$ give, respectively, the correlations in the stationary state, among the coordinates, momenta and between the coordinates and momenta

$$x_{ij} = \langle q_i q_j\rangle, \quad y_{ij} = \langle p_i p_j\rangle, \quad z_{ij} = \langle q_i p_j\rangle. \quad (3.6)$$

To obtain the deviation of these correlations from their equilibrium values at uniform temperature T, corresponding to $\eta = 0$, we write

$$\mathbf{x} = (kT/\omega^2)[\mathbf{G}^{-1} + \eta\mathbf{X}], \quad (3.7)$$

$$\mathbf{y} = kT[\mathbf{I} + \eta\mathbf{Y}], \quad (3.8)$$

$$\mathbf{z} = \lambda^{-1}kT\eta\mathbf{Z}. \quad (3.9)$$

Using now (2.8) we find the following equations for $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$:

$$\mathbf{Z} = -\mathbf{Z}^\dagger, \quad (3.10)$$

$$\mathbf{Y} = \mathbf{XG} + \mathbf{ZR}, \quad (3.11)$$

$$2\mathbf{E} - \mathbf{YR} - \mathbf{RY} = \nu[\mathbf{GZ} - \mathbf{ZG}]. \quad (3.12)$$

In addition, $\mathbf{X}$ and $\mathbf{Y}$ are required to be symmetric

$$\mathbf{X} = \mathbf{X}^\dagger, \quad \mathbf{Y} = \mathbf{Y}^\dagger, \quad (3.13)$$

while $\mathbf{b}$ is required to be positive definite. The quantity ν in (3.12) is $\nu = \omega^2/\lambda^2$, and is the only dimensionless parameter to remain in the problem.

To obtain an explicit solution of (3.10)–(3.13) we first note that the left side of (3.12), $2\mathbf{E} - \mathbf{YR} - \mathbf{RY}$, is a bordered matrix (it has nonvanishing elements only in the first and last rows and columns). Hence $\mathbf{GZ} - \mathbf{ZG}$ must also be bordered. Using the explicit form of G, (3.2), together with the antisymmetry requirement (3.10), it is easy to show that $\mathbf{Z}$ is necessarily a skew-symmetric Toeplitz matrix when $\mathbf{GZ} - \mathbf{ZG}$ is a bordered matrix, and $\mathbf{Z}$ may therefore be written in the form

$$\mathbf{Z} = \begin{pmatrix} 0 & \varphi_1 & \varphi_2 & \cdots & \varphi_{N-2} & \varphi_{N-1} \\ -\varphi_1 & & & & & \varphi_{N-2} \\ -\varphi_2 & & & & & \vdots \\ \vdots & & & & & \varphi_2 \\ & & & & & \varphi_1 \\ -\varphi_{N-1} & \cdots & -\varphi_2 & -\varphi_1 & 0 & \end{pmatrix}. \quad (3.14)$$

The quantities $\varphi_1, \cdots, \varphi_N$ are simply related to the entries in the bordered matrix in the left side of (3.12) and turn out to be

$$\nu\varphi_j = \delta_{j1} - Y_{1j} = \delta_{j1} + Y_{N,N-j+1}, \quad (3.15)$$

where $\varphi_N \equiv 0$ by definition. Equation (3.15) implies certain obvious restrictions on $\mathbf{Y}$ in order that (3.12) has a solution.

Next, Eq. (3.11), together with the fact that $\mathbf{Y}$ is symmetric implies that

$$\mathbf{XG} - \mathbf{GX} = -(\mathbf{RZ} + \mathbf{ZR}). \quad (3.16)$$

Once again, the right side of (3.16) is a bordered matrix, which is known in terms of the φ's. Unlike $\mathbf{Z}$,

however, $\mathbf{X}$ is required to be symmetric and we find that one solution is a Hankel matrix:

$$\mathbf{X} = \begin{pmatrix} \varphi_1 & \varphi_2 & \varphi_3 & \cdots & \cdots & \varphi_{N-2} & \varphi_{N-1} & 0 \\ & & & & & & & -\varphi_{N-1} \\ & & & & & & & \vdots \\ & & & & & & & -\varphi_2 \\ & & & & & & & -\varphi_1 \end{pmatrix}. \tag{3.17}$$

The solution to (3.16) is, however, not unique. We could add to (3.17) any symmetric matrix that commutes with $\mathbf{G}$. Because all the eigenvalues of $\mathbf{G}$ are nondegenerate, such a matrix must be of the form $\tilde{\mathbf{X}} = P(\mathbf{G})$, where P is any polynomial. Nevertheless, $\tilde{\mathbf{X}}$ is required to vanish as a consequence of our last condition, (3.11).

Substitution of (3.14) and (3.17) into the right side of (3.11) gives an expression for $\mathbf{Y}$ in terms of the φ's. This, when combined with (3.15) yields an equation for the φ's, namely,

$$\sum_{j=1}^{N-1} K_{ij}\varphi_j = \delta_{i1}, \tag{3.18}$$

where $\mathbf{K}$ is the $(N-1)$-square matrix

$$\mathbf{K} = \begin{pmatrix} \nu+2 & -1 & & & & \\ -1 & \nu+2 & -1 & & & \\ & -1 & \ddots & \ddots & & \\ & & \ddots & \ddots & -1 & \\ & & & -1 & \nu+2 & -1 \\ & & & & -1 & \nu+2 \end{pmatrix}. \tag{3.19}$$

The matrix $\mathbf{Y}$ is then the Hankel matrix

$$\mathbf{Y} = \mathbf{E} - \nu\mathbf{X}. \tag{3.20}$$

It is to be noted that both $\mathbf{X}$ and $\mathbf{Y}$ are antisymmetric about the "cross" diagonal [i.e., the $(1, N)$–$(N, 1)$ diagonal], a state of affairs that reflects the fact that changing the sign of η corresponds to interchanging the reservoirs at the ends of the chain. Were we to add a matrix $\tilde{\mathbf{X}} = P(\mathbf{G})$ to $\mathbf{X}$, as discussed above, then $\mathbf{X}$ would acquire a symmetric component about the "cross" diagonal. By (3.11) $\mathbf{Y}$, too, would acquire such a component. But (3.15) precludes this possibility.

We conclude then, that the solution to (3.10)–(3.13) is given *uniquely* by (3.14), (3.17), and (3.20), assuming, of course that (3.18) has a unique solution. That this is so follows from the fact that $\mathbf{K}$ is positive definite for all $\nu \geq 0$.

To find φ we use Cramer's rule. Consider the general M-square version of K and let D_M denote its determinant. Then

$$\varphi_j = (D_{N-1})^{-1} \times (\text{cofactor of } K_{1j} \text{ in } D_{N-1}). \tag{3.21}$$

But, it is easily seen that the cofactor of K_{1j} is simply D_{N-1-j}. The computation of D_M is simple. We observe that if we consider D_M to be a function of the parameter $x = 1 + \frac{1}{2}\nu$, then $D_M(x) = 2xD_{M-1}(x) - D_{M-2}(x)$, while $D_0(x) = 1$ and $D_1(x) = 2x$. This is just the recursion relation for $U_M(x)$, the Chebyshev polynomial of the second kind. Using the fact that $U_M(\cosh\theta) = \sinh(M+1)\theta/\sinh\theta$, and defining α by

$$\cosh\alpha = 1 + \tfrac{1}{2}\nu \tag{3.22}$$

we obtain

$$\varphi_j = \sinh(N-j)\alpha/\sinh N\alpha. \tag{3.23}$$

For large N and fixed j we have the asymptotic formula:

$$\varphi_j = e^{-j\alpha} = (\varphi_1)^j, \tag{3.24}$$

and

$$\varphi_1 = e^{-\alpha} = 1 + \tfrac{1}{2}\nu - \tfrac{1}{2}(4\nu + \nu^2)^{\frac{1}{2}}. \tag{3.25}$$

For $\nu \to 0$ or ∞ we have

$$\varphi_1(\nu) \to \begin{cases} 1 - \nu^{\frac{1}{2}}, & \text{as } \nu \to 0, \\ \nu^{-1}, & \text{as } \nu \to \infty. \end{cases} \tag{3.26}$$

Finally, we note that $\mathbf{b}$ is necessarily positive definite, a property that can be proved directly from (2.8) using the fact that $\mathbf{\Phi}$ is positive definite.

For the case of hard collisions, $\mu(x, t)$ satisfying (1.5) and $\mathbf{b}'(t)$ satisfying (2.11), the corresponding matrices $\mathbf{X}'$, $\mathbf{Y}'$, and $\mathbf{Z}'$ will also satisfy Eq. (3.10) and (3.11) while (3.12) is now replaced by

$$\mathbf{E} - \mathbf{Y}'\mathbf{R} - \mathbf{R}\mathbf{Y}' + \mathbf{R}\mathbf{Y}'\mathbf{R} = \nu[\mathbf{G}\mathbf{Z}' - \mathbf{Z}'\mathbf{G}] \tag{3.27}$$

with ν now given by λ'^2/ω^2. These equations may be brought into the same form as (3.10)–(3.12) by the replacements

$$\mathbf{X}' = (1 + \nu\varphi_1)^{-1}\mathbf{X}, \quad \mathbf{Y}' = (1 + \nu\varphi_1)^{-1}\mathbf{Y}, \quad \mathbf{Z}' = (1 + \nu\varphi_1)^{-1}\mathbf{Z}, \tag{3.28}$$

which corresponds simply to the replacement of the temperature difference η by η^*, where

$$\eta^* = \frac{\eta}{(1 + \nu\varphi_1)} \to \begin{cases} \eta & \text{as } \nu \to 0, \\ \frac{1}{2}\eta & \text{as } \nu \to \infty. \end{cases} \tag{3.29}$$

4. PROPERTIES OF THE STATIONARY STATE

Kinetic Temperature

The kinetic temperature of the jth particle is given by

$$kT(j, \nu; N) = \langle p_j^2 \rangle = kT[1 + \eta Y_{jj}] = \langle x_j\, \partial H/\partial x_j \rangle \tag{4.1}$$

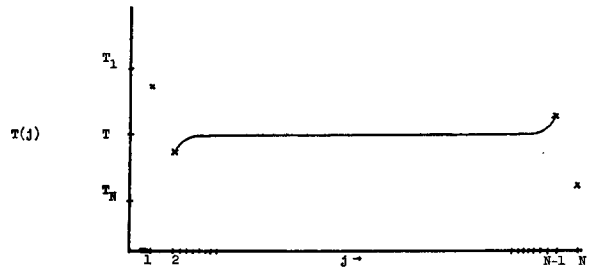


FIG. 1. Typical behavior of $T(j)$.

the last equality holding for all $i = 1, \cdots, 2N$ indicating some kind of equipartition for the stationary nonequilibrium state. In the limit of $N \to \infty$ we have

$$T(j, \nu) = \begin{cases} T[1 - \eta\nu(\varphi_1)^{2j-1}], & 1 < j < \frac{1}{2}N, \\ T[1 + \eta\nu(\varphi_1)^{2j'-1}], & 1 < j' = N - j < \frac{1}{2}N, \end{cases} \tag{4.2}$$

and

$$\begin{aligned} T(1, \nu) &= T_1 - \nu\varphi_1 T\eta, \\ T(N, \nu) &= T_N + \nu\varphi_1 T\eta, \end{aligned} \tag{4.3}$$

with φ_1 given by (3.25) and $T_1 = T(1 + \eta)$, $T_N = T(1 - \eta)$. The temperature of the linear chain thus deviates from its average value $T = \frac{1}{2}(T_1 + T_N)$ only at its edges where it changes exponentially over a length $\ell(\nu)$;

$$\ell(\nu) \to \begin{cases} \frac{1}{2}\nu^{-\frac{1}{2}}, & \nu \to 0, \\ \frac{1}{2}(\ln \nu)^{-1}, & \nu \to \infty. \end{cases}$$

It is a totally unexpected result of this model that the kinetic temperature *drops* below the average value at the second particle and then *increases* exponentially as we move *away* from the hot reservoir (cf. Fig. 1).

For the case of hard collisions (4.2) is unchanged except for the replacement of η by η^*, while (4.3) now has the form

$$\begin{aligned} T'(1, \nu) &= T_1 - 2\nu\varphi_1 T\eta^*, \\ T'(N, \nu) &= T_N - 2\nu\varphi_1 T\eta^*. \end{aligned} \tag{4.5}$$

Heat Flux

It is easy to show that the energy flux across a plane separating the $(i - 1)$th particle from the ith particle is given by[7]

$$j_{i-1,i} = \omega^2\langle q_{i-1}p_i\rangle = \omega^2 Z_{12} \equiv J, \quad i = 2, \cdots, N - 1.$$

The equalities hold in the stationary state where the flux is constant throughout the system and coincides with the energy flux $j_1 = -j_N$, coming from the reservoir at the left and going into the reservoir at the right, which is given by[2] $\lambda k[T_1 - T(1, \nu)]$. We then have using (3.25)

$$\begin{aligned} J(\lambda, \omega) = (\omega^2/\lambda)kT\varphi_1\eta &\xrightarrow[N\to\infty]{} \tfrac{1}{2}(\omega^2/\lambda)kT[1 + (\nu/2) \\ &\quad - \tfrac{1}{2}\nu(1 + 4/\nu)^{\frac{1}{2}}](T_1 - T_N) \\ &= \tfrac{1}{2}(\omega^2/\lambda)k(T_1 - T_N) \quad \text{for } \lambda \gg \omega \\ &= \tfrac{1}{2}\lambda k(T_1 - T_N) \qquad \text{for } \lambda \ll \omega. \end{aligned} \tag{4.6}$$

As expected the heat flux for the harmonic crystal is proportional to $(T_1 - T_N)$ rather than to the temperature gradient $(T_1 - T_N)/N$, i.e., the "heat" conductivity[7] is proportional to the size of the system.

The behavior of $J(\lambda, \omega)$ for fixed ω is very peculiar. For small λ, $\lambda \ll \omega$, J is proportional to λ as it should be, and is independent of ω, the whole chain behaving as if it were just one tight piston. As λ increases J reaches a maximum at $\lambda = \frac{1}{2}\sqrt{3}\,\omega$, $J_{\max} = \frac{1}{8}k(T_1 - T_N)\omega$, and then decreases, vanishing as λ^{-1} when $\lambda \to \infty$; the system now behaves, in the stationary state, as if the oscillators at the two ends are at the temperatures of the corresponding reservoirs, while the remainder are in equilibrium at temperature T. This latter behavior is quite unexpected. For a physical system (anharmonic coupling) we would expect $J(\lambda)$ to reach a limiting value proportional to the heat conductivity of the system times the temperature gradient.

For the case of hard collisions

$$J'(\lambda, \omega) = (1 + \nu\varphi_1)^{-1}J(\lambda, \omega) \to \begin{cases} J(\lambda, \omega), & \nu \to 0, \\ \frac{1}{2}J(\lambda, \omega), & \nu \to \infty. \end{cases}$$

ACKNOWLEDGMENTS

One of us (J. L. L.) would like to thank Dr. O. Bils for stimulating correspondence.

This work was supported by the U.S. Air Force Office of Scientific Research under Grant No. 508-66.

Eigenfunction Expansions Associated with the Second-Order Invariant Operator on Hyperboloids and Cones. III

N. LIMIĆ,* J. NIEDERLE,† AND R. RĄCZKA‡
International Center for Theoretical Physics, Trieste, Italy

(Received 14 July 1966)

The eigenfunction expansions associated with the second-order invariant operator on hyperboloids and cones are derived. The global unitary irreducible representations of the $SO_0(p, q)$ groups related to hyperboloids and cones are obtained. The decomposition of the quasi-regular representations into the irreducible ones is given and the connection with the Mautner theorem and nuclear spectral theory is discussed.

1. INTRODUCTION

IN our two previous works[1,2] the most degenerate irreducible infinitesimal representations of an arbitrary noncompact rotation group $SO_0(p, q)$ have been derived.[3,4] These representations have been related to the homogeneous spaces $SO_0(p, q)/SO_0(p-1, q)$, $SO_0(p, q)/SO_0(p, q-1)$, and $SO_0(p, q)/T^{p+q-2}\,\boxed{S}\,SO(p-1, q-1)$ which can be represented by the hyperboloids H_q^p, H_p^q and by the cone C_q^p, respectively. These homogeneous spaces are of rank one under the action of the group. The infinitesimal representations have been constructed by means of the sets of harmonic functions associated with the second-order invariant operators related to the above-mentioned manifolds.

The completeness of these sets of harmonic functions has been proved in the present work by using the classical Titchmarsh[5]–Kodaira[6] eigenfunction expansion theory associated with an ordinary second-order differential equation.

In Sec. 2 the review of the main results in the form convenient for applications is given. The other sections are devoted to the proofs. Thus Sec. 3 contains the proof of the essential self-adjointness of the Laplace–Beltrami operator on the linear manifold $\mathfrak{D}(X)$, which was introduced earlier.[2] The proof of the completeness of the harmonic functions which were constructed in our previous papers[1,2] is given in Sec. 4. In Sec. 5 we prove the unitarity and irreducibility of the representations of the group $SO_0(p, q)$ and consider the decomposition of the quasi-regular representations into irreducible ones. In Sec. 6 we show an interesting connection of our approach to the eigenfunction expansion and the general theory of the eigenfunction expansion developed by Gel'fand and Kostiuchenko,[7] Maurin,[8] and Gårding.[9] Appendix I contains some auxiliary computations. In Appendix II we review the main results on the representations of the compact group $SO(p)$.

2. REVIEW OF MOST DEGENERATE REPRESENTATIONS OF $SO_0(p, q)$ GROUPS

We have considered in our previous papers[1,2] three homogeneous spaces X of rank one under the action of the noncompact rotation group $SO_0(p, q)$ [see (2.2)[1] and (4.1)[2]]. They can be represented by the hyperboloids H_q^p and H_p^q and the cone C_q^p:

$$(x^1)^2 + \cdots + (x^p)^2 - (x^{p+1})^2 - \cdots - (x^{p+q})^2 = \begin{cases} 1 & \text{for the hyperboloid } H_q^p, \quad p \geq q, \\ 0 & \text{for the cone } C_q^p, \\ -1 & \text{for the hyperboloid } H_p^q, \quad p \geq q. \end{cases} \tag{2.1}$$

* On leave of absence from Institute Rudjer Bošković, Zagreb.

† On leave of absence from Institute of Physics of the Czechoslovak Academy of Sciences, Prague.

‡ On leave of absence from Institute of Nuclear Research, Warsaw.

[1] R. Rączka, N. Limić, and J. Niederle, J. Math. Phys. 7, 1861 (1966).

[2] N. Limić, J. Niederle, and R. Rączka, J. Math. Phys. 7, 2026 (1966).

[3] The representations of the compact $SO(p)$ group related to the sphere and the Lorentz-type group $SO_0(p, 1)$ related to the real Lobachevski space (H_p^1 in our notation) are given in the book by N. J. Vilenkin: "Special Functions and Theory of Group Representation," "Nauka," Moscow (1965) (in Russian).

[4] For example, the representations of some lower-dimensional $SO(p, q)$ groups are given in: J. A. C. Alcaras and P. L. Ferreira, J. Math. Phys. 6, 578 (1965); V. Bargmann, Ann. Math. 48, 568 (1947); C. G. Bollini and J. J. Giambiagi, "Unified Representations of the Rotation and Lorentz (2 + 1) Groups," preprint, University of Buenos Aires (1965); J. Dixmier, Bull. Soc. Math. France 89, 9 (1961); A. Z. Dolginov, Soviet Phys.—JETP 3, 589 (1956); A. Z. Dolginov and J. N. Toptygin, *ibid.* 10, 1022 (1960); A. Z. Dolginov and A. N. Moskalev, *ibid.* 10, 1202 (1960); J. B. Ehrman, Proc. Cambridge Phil. Soc. 53, 290 (1956); N. T. Evans, J. Math. Phys. 8, 170 (1967); A. Kihlberg, Arkiv Fysik 27, 373 (1964), 28, 121 (1964), and 30, 121 (1965); A. Kihlherg and S. Ström, Arkiv Fysik 31, 491 (1966); A. Kihlberg, *ibid.* 30, 121 (1965); L. H. Thomas, Ann. Math. 42, 113 (1941).

[5] E. C. Titchmarsh, *Eigenfunction Expansions* (Clarendon Press, Oxford, England, 1962), Part I.

[6] K. Kodaira, Am. J. Math. 71, 921 (1949).

[7] I. M. Gel'fand and A. G. Kostiuchenko, Dokl. Akad. Nauk SSSR 103, 349 (1955).

[8] K. Maurin, Bull. Acad. Polon. Sci. 7, 471 (1959).

[9] L. Gårding, Seminar on Applied Mathematics, Boulder, Colorado (1957), pp. 1–30.

These homogeneous spaces are imbedded in the $(p+q)$-dimensional Minkowski space $M^{p,q}$. The considered homogeneous spaces are parametrized by use of biharmonic coordinates Ω which are defined in Sec. 3[1] for the hyperboloids and in Sec. 4[2] for the cone.

Let $d\mu(\Omega)$ be the Riemannian left-invariant measure on X and let $\mathfrak{H}(X)$ be the Hilbert space of $L^2(\mu)$ type. The quasi-regular representation of the rotation group is determined by

$$SO_0(p,q) \ni g \to (U_g f)(\Omega) = f(g^{-1}\Omega), \quad f \in \mathfrak{H}(X). \tag{2.2}$$

The quasi-regular representation defined by (2.2) is unitary but reducible. Its irreducible parts are given below for any particular choice of the homogeneous space X. Since the homogeneous spaces H_q^p, H_p^q, and C_q^p are of rank one under the action of the $SO_0(p,q)$ group, the ring of the invariant operators of the representation of the corresponding Lie algebra $\mathfrak{R}(p,q)$ is generated only by one operator.[1]

The generators x_{ij} of the Lie algebra $\mathfrak{R}(p,q)$ are represented by the unbounded operators X_{ij} in the Hilbert space $\mathfrak{H}(X)$ [see (6.2),[1] (6.3),[1] (4.4),[2] and (4.5)[2]]. Hence, if we consider the representations of the generators and their polynomials (for instance, the Casimir operator) we must restrict their domain of definition to some dense linear manifold in $\mathfrak{H}(X)$. The most convenient choice is the common invariant domain $\mathfrak{D}(X)$ for all the operators X_{ij}, i.e., $\mathfrak{D}(X)$ has the property $X_{ij}\mathfrak{D}(X) \subset \mathfrak{D}(X)$ for any X_{ij}.

Let $\mathfrak{D}(X)$ be the linear manifold determined by vectors $f \in \mathfrak{H}(X)$ of the form

$$f(\Omega) = P(x^1, \cdots, x^{p+q}) \exp\left\{-\sum_{i=1}^{p+q}(x^i)^2\right\},$$

where $P(x^1, \cdots, x^{p+q})$ is a polynomial in $x^1, \cdots, x^{p+q}$ and the x^i, $i = 1, 2, \cdots, p+q$ are expressed in biharmonic coordinates Ω as in Sec. 3[1] and Sec. 4.[2] Then the linear manifold $\mathfrak{D}(X)$ is a dense invariant domain of the operators X_{ij}.

The proof can be found in Appendix I. Let us review now the particular cases in detail.

A. Hyperboloid H_q^p, $p \geq q > 2$

(i) The spectrum[10] $S(\Delta(H_q^p))$ of the Laplace–Beltrami operator $\Delta(H_q^p)$ [see (3.10)[1]] consists of the discrete spectrum[11] $PS(\Delta(H_q^p))$: $-L(L+p+q-2)$, $L = -\{\frac{1}{2}(p+q-4)\} - \{\frac{1}{2}(p+q-4)\} + 1, \cdots$ and the continuous spectrum $CS(\Delta(H_q^p))$: $\Lambda^2 + [\frac{1}{2}(p+q-2)]^2$, $\Lambda \in [0, \infty)$.

(ii) The corresponding eigenfunctions are (see Sec. 3[1] and Sec. 2[2])

$$Y^{L,l_2,\cdots,l_{\{p/2\}},\bar{l}_2,\cdots,\bar{l}_{\{q/2\}}}_{m_1,\cdots,m_{[p/2]},\tilde{m}_1,\cdots,\tilde{m}_{[q/2]}}(\Omega) = V^L_{l_{\{p/2\}},\bar{l}_{\{q/2\}}}(\theta) Y^{l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}}(\omega) Y^{\bar{l}_2,\cdots,\bar{l}_{\{q/2\}}}_{\tilde{m}_1,\cdots,\tilde{m}_{[q/2]}}(\tilde{\omega}),$$

$$Y^{\Lambda,l_2,\cdots,l_{\{p/2\}},\bar{l}_2,\cdots,\bar{l}_{\{q/2\}}}_{m_1,\cdots,m_{[p/2]},\tilde{m}_1,\cdots,\tilde{m}_{[q/2]}}(\Omega) = V^\Lambda_{l_{\{p/2\}}\,\bar{l}_{\{q/2\}}}(\theta) Y^{l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}}(\omega) Y^{\bar{l}_2,\cdots,\bar{l}_{\{q/2\}}}_{\tilde{m}_1,\cdots,\tilde{m}_{[q/2]}}(\tilde{\omega}),$$

where

$$V^L_{l_{\{p/2\}},\bar{l}_{\{q/2\}}}(\theta) = (1/N^{\frac{1}{2}})(\tanh\theta)^{\bar{l}_{\{q/2\}}}(\cosh\theta)^{-(L+p+q-2)}\,{}_2F_1$$
$$\times [\tfrac{1}{2}(p+q-2+l_{\{p/2\}}+\bar{l}_{\{q/2\}}+L), \tfrac{1}{2}(L+q+\bar{l}_{\{q/2\}}-l_{\{p/2\}}); \bar{l}_{\{q/2\}}+\tfrac{1}{2}q; \tanh^2\theta],$$

$$N = \frac{\Gamma[\frac{1}{2}(l_{\{p/2\}}-\bar{l}_{\{q/2\}}-L-q+2)]\Gamma^2(\bar{l}_{\{q/2\}}+\frac{1}{2}q)\Gamma[\frac{1}{2}(L-\bar{l}_{\{q/2\}}+l_{\{p/2\}}+p)]}{2[L+\frac{1}{2}(p+q-2)]\Gamma[\frac{1}{2}(l_{\{p/2\}}+\bar{l}_{\{q/2\}}+L+p+q-2)]\Gamma[\frac{1}{2}(l_{\{p/2\}}+\bar{l}_{\{q/2\}}-L)]},$$

$$V^\Lambda_{l_{\{p/2\}},\bar{l}_{\{q/2\}}}(\theta) = (1/M^{\frac{1}{2}})(\tanh\theta)^{\bar{l}_{\{q/2\}}}(\cosh\theta)^{i\Lambda-\frac{1}{2}(p+q-2)}\,{}_2F_1$$
$$\times \{\tfrac{1}{2}[\bar{l}_{\{q/2\}}+l_{\{p/2\}}-i\Lambda+\tfrac{1}{2}(p+q-2)], \tfrac{1}{2}[\bar{l}_{\{q/2\}}-l_{\{p/2\}}-i\Lambda+\tfrac{1}{2}(q-p+2)];\ {}_{\{q/2\}}+\tfrac{1}{2}q, \tanh^2\theta\},$$

$$M = 2\pi\left|\frac{\Gamma(\bar{l}_{\{q/2\}}+q/2)\Gamma(i\Lambda)}{\Gamma\{\frac{1}{2}[i\Lambda+l_{\{p/2\}}+\bar{l}_{\{q/2\}}+\frac{1}{2}(p+q-2)]\}\Gamma\{\frac{1}{2}[i\Lambda+\bar{l}_{\{q/2\}}-l_{\{p/2\}}+\frac{1}{2}(q-p+2)]\}}\right|^2$$

The functions $Y^{l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}}(\omega)$ are defined in Appendix II in the expressions (A1)–(A6). The same expressions hold for the functions $Y^{\bar{l}_2,\cdots,\bar{l}_{\{q/2\}}}_{\tilde{m}_1,\cdots,\tilde{m}_{[q/2]}}(\tilde{\omega})$ except in tilde variables and tilde indices. The variable Λ is independent of the indices l_i, m_i, $\bar{l}_i$, and $\tilde{m}_i$, whereas L has the following dependence:

$$L - l_{\{p/2\}} + \bar{l}_{\{q/2\}} = -q - 2n, \quad n = 0, 1, 2, \cdots. \tag{2.3}$$

[10] In fact, we mean the spectrum of the self-adjoint extension of $\Delta(H_q^p)$. (For details see Secs. 3 and 4.)

[11] We use notation: $[a/2]$ is $\frac{1}{2}a$ if $a = 2r$ and $\frac{1}{2}(a-1)$ if $a = 2r+1$, $r = 1, 2, \cdots$; $\{a/2\}$ is $\frac{1}{2}a$ if $a = 2r$ and $\frac{1}{2}(a+1)$ if $a = 2r+1$, $r = 1, 2, \cdots$.

(iii) The completeness relations

$$\int_{H_q^p} \overline{Y^{L,l_2,\cdots,l_{\{p/2\}},\tilde{l}_2,\cdots,\tilde{l}_{\{q/2\}}}_{m_1,\cdots,m_{[p/2]},\tilde{m}_1,\cdots,\tilde{m}_{[q/2]}}(\Omega)} Y^{L',l_2',\cdots,l_{\{p/2\}}',\tilde{l}_2',\cdots,\tilde{l}_{\{q/2\}}'}_{m_1',\cdots,m_{[p/2]}',\tilde{m}_1',\cdots,\tilde{m}_{[q/2]}'}(\Omega)\, d\mu(\Omega) = \delta_{LL'} \prod_{k=2}^{\{p/2\}} \delta_{l_k l_k'} \prod_{k=1}^{[p/2]} \delta_{m_k m_k'} \prod_{k=2}^{\{q/2\}} \delta_{\tilde{l}_k \tilde{l}_k'} \prod_{k=1}^{[q/2]} \delta_{\tilde{m}_k \tilde{m}_k'},$$

$$\int_{H_q^p} \overline{Y^{\Lambda,l_2,\cdots,l_{\{p/2\}},\tilde{l}_2,\cdots,\tilde{l}_{\{q/2\}}}_{m_1,\cdots,m_{[p/2]},\tilde{m}_1,\cdots,\tilde{m}_{[q/2]}}(\Omega)} Y^{\Lambda',l_2',\cdots,l_{\{p/2\}}',\tilde{l}_2',\cdots,\tilde{l}_{\{q/2\}}'}_{m_1',\cdots,m_{[p/2]}',\tilde{m}_1',\cdots,\tilde{m}_{[q/2]}'}(\Omega)\, d\mu(\Omega)$$
$$= \delta(\Lambda - \Lambda') \prod_{k=2}^{\{p/2\}} \delta_{l_k l_k'} \prod_{k=1}^{[p/2]} \delta_{m_k m_k'} \prod_{k=2}^{\{q/2\}} \delta_{\tilde{l}_k \tilde{l}_k'} \prod_{k=1}^{[q/2]} \delta_{\tilde{m}_k \tilde{m}_k'},$$

where the left-invariant measure $d\mu(\Omega)$ on H_q^p is defined by

$$d\mu(\Omega) = \cosh^{p-1}\theta \sinh^{q-1}\theta\, d\theta\, d\mu(\omega_{p-1})\, d\mu(\tilde{\omega}_{q-1}),\ \theta \in [0, \infty)$$

and $d\mu(\omega)$ is defined in Appendix II by (A9).

$$\sum_{L=-\{\frac{1}{2}(p+q-4)\}}^{\infty} \sum_{\mathcal{N}_L} \overline{Y^{L,l_2,\cdots,l_{\{p/2\}},\tilde{l}_2,\cdots,\tilde{l}_{\{q/2\}}}_{m_1,\cdots,m_{[p/2]},\tilde{m}_1,\cdots,\tilde{m}_{[q/2]}}(\Omega)} Y^{L,l_2,\cdots,l_{\{p/2\}},\tilde{l}_2,\cdots,\tilde{l}_{\{q/2\}}}_{m_1,\cdots,m_{[p/2]},\tilde{m}_1,\cdots,\tilde{m}_{[q/2]}}(\Omega')$$
$$+ \int_0^{\infty} d\Lambda \sum_{\mathcal{N}_\Lambda} \overline{Y^{\Lambda,l_2,\cdots,l_{\{p/2\}},\tilde{l}_2,\cdots,\tilde{l}_{\{q/2\}}}_{m_1,\cdots,m_{[p/2]},\tilde{m}_1,\cdots,\tilde{m}_{[q/2]}}(\Omega)} Y^{\Lambda,l_2,\cdots,l_{\{p/2\}},\tilde{l}_2,\cdots,\tilde{l}_{\{q/2\}}}_{m_1,\cdots,m_{[p/2]},\tilde{m}_1,\cdots,\tilde{m}_{[q/2]}}(\Omega') = \delta(\Omega - \Omega'; \mu),$$

where[12] $\mathcal{N}_L$ is the set of values of indices $l_2, \cdots, \tilde{m}_{[q/2]}$, which are restricted by the conditions (A4), (A5), and (2.3) and $\mathcal{N}_\Lambda$ the set restricted by the conditions (A4) and (A5).

(iv) Fourier transform:

$$\mathfrak{H}(H_q^p) \ni f \to \chi = Ff$$
$$:= \left\{ \int_{H^p} \overline{Y^{\sigma,l_2,\cdots,l_{\{p/2\}},\tilde{l}_2,\cdots,\tilde{l}_{\{q/2\}}}_{m_1,\cdots,m_{[p/2]},\tilde{m}_1,\cdots,\tilde{m}_{[q/2]}}(\Omega)} f(\Omega)\, d\mu(\Omega); \right.$$
$$\left. \sigma \in \Sigma,\ l_2, \cdots, \tilde{m}_{[q/2]} \in \mathcal{N}_\sigma \right\} \in \mathfrak{H}(S),$$

where σ denotes either the variable L or Λ and Σ is the range of σ corresponding to the whole spectrum of $\Delta(H_q^p)$.[10]

(v) The complete classification of the irreducible representations can be made by means of the spectrum of the operator $\Delta(H_q^p)$ and the spectrum of the operator P, where P is defined in (5.4).[2] [On the subspace $\mathfrak{H}^L \subset \mathfrak{H}(S)$ corresponding to a point of $PS(\Delta(H_q^p))$ the operator P has a unique eigenvalue $(-)^{L+q}$, whereas on the linear subset $\mathfrak{H}^\Lambda$ corresponding to a point of $CS(\Delta(H_q^p))$, P has the "eigenvectors" with both its eigenvalues.] We denote by γ a pair σ, p where p is the eigenvalue of $P: p = \pm 1$. Let $\mathcal{N}_\gamma$ be that subset of $\mathcal{N}_\sigma$ for which either $l_{\{p/2\}} + \tilde{l}_{\{q/2\}}$ is even and $p = 1$ or $l_{\{p/2\}} + \tilde{l}_{\{q/2\}}$ is odd and $p = -1$. (The set $\mathcal{N}_L$ is already of this type, while $\mathcal{N}_\Lambda$ has two proper subsets of this type.) We consider the unitary space $\mathfrak{D}^\gamma$ of l^2 type determined by vectors

$$\chi^\gamma := \{\langle Y^{\sigma,l_2,\cdots,l_{\{p/2\}},\tilde{l}_2,\cdots,\tilde{l}_{\{q/2\}}}_{m_1,\cdots,m_{[p/2]},\tilde{m}_1,\cdots,\tilde{m}_{[q/2]}}, f\rangle;\ l_2, \cdots, \tilde{m}_{[q/2]} \in \mathcal{N}_\gamma, f \in \mathfrak{D}(H_q^p)\},$$

where

$$\langle Y^{\sigma,l_2,\cdots,l_{\{p/2\}},\tilde{l}_2,\cdots,\tilde{l}_{\{q/2\}}}_{m_1,\cdots,m_{[p/2]},\tilde{m}_1,\cdots,\tilde{m}_{[q/2]}}, f\rangle := \int \overline{Y^{\sigma,l_2,\cdots,l_{\{p/2\}},\tilde{l}_2,\cdots,\tilde{l}_{\{q/2\}}}_{m_1,\cdots,m_{[p/2]},\tilde{m}_1,\cdots,\tilde{m}_{[q/2]}}(\Omega)} f(\Omega)\, d\mu(\Omega).$$

Let us keep in mind that the scalar product in $\mathfrak{D}^\gamma$ is determined by

$$(\chi^\gamma, \psi^\gamma)_\gamma = \sum_{\mathcal{N}_\gamma} \overline{\langle Y^{\sigma,l_2,\cdots,l_{\{p/2\}},\tilde{l}_2,\cdots,\tilde{l}_{\{q/2\}}}_{m_1,\cdots,m_{[p/2]},\tilde{m}_1,\cdots,\tilde{m}_{[q/2]}}, f\rangle} \langle Y^{\sigma,l_2,\cdots,l_{\{p/2\}},\tilde{l}_2,\cdots,\tilde{l}_{\{q/2\}}}_{m_1,\cdots,m_{[p/2]},\tilde{m}_1,\cdots,\tilde{m}_{[q/2]}}, g\rangle.$$

The completion of the unitary space $\mathfrak{D}^\gamma$ with respect to the norm $\|\cdot\|_\gamma$ is the Hilbert space $\mathfrak{H}^\gamma$. The unitary irreducible representation of the group $SO_0(p, q)$ and the irreducible representation of the algebra $\mathfrak{R}(p, q)$ are defined by

$$SO_0(p, q) \ni g \to U_g^\gamma \chi^\gamma := \{\langle Y^{\sigma,l_2,\cdots,l_{\{p/2\}},\tilde{l}_2,\cdots,\tilde{l}_{\{q/2\}}}_{m_1,\cdots,m_{[p/2]},\tilde{m}_1,\cdots,\tilde{m}_{[q/2]}}, U_g f\rangle;\ l_2, \cdots, \tilde{m}_{[q/2]} \in \mathcal{N}_\gamma, f \in \mathfrak{D}(H_q^p)\} \in \mathfrak{H}^\gamma$$

$$\mathfrak{R}(p, q) \ni x_{ij} \to X_{ij}^\gamma \chi^\gamma := \{\langle Y^{\sigma,l_2,\cdots,l_{\{p/2\}},\tilde{l}_2,\cdots,\tilde{l}_{\{q/2\}}}_{m_1,\cdots,m_{[p/2]},\tilde{m}_1,\cdots,\tilde{m}_{[q/2]}}, X_{ij} f\rangle;\ l_2, \cdots, \tilde{m}_{[q/2]} \in \mathcal{N}_\gamma, f \in \mathfrak{D}(H_q^p)\} \in \mathfrak{D}^\gamma$$

[12] The δ function $\delta(\Omega - \Omega'; \mu)$ is defined by:

$$\int \delta(\Omega - \Omega'; \mu) f(\Omega')\, d\mu(\Omega') = f(\Omega).$$

where $U_g f$ is defined by (2.2) and $X_{ij} f$ by (6.2),[1] (6.3),[1] (4.4),[2] and (4.5).[2]

(vi) The quasi-regular representation in (2.2) decomposes into the irreducible representations in the following way:

$$U_g = \sum \oplus U_g^{L,(-)^{L+q}} \oplus \int \oplus d\Lambda U_g^{\Lambda,+} \oplus \int \oplus d\Lambda U_g^{\Lambda,-}$$

in the Hilbert space (see Secs. 4 and 5)

$$\mathfrak{H}(S) = \sum \oplus \mathfrak{H}^{L,(-)^{L+q}} \oplus \int \oplus d\Lambda \mathfrak{H}^{\Lambda,+} \oplus \int \oplus d\Lambda \mathfrak{H}^{\Lambda,-}.$$

B. Hyperboloid H_p^q, $p \geq q > 2$

Changing $q \rightleftarrows p$ in all the expressions of A and removing the tilde from any variable or index which previously had it and at the same time placing a tilde over any variable or index which previously did not have it, we obtain the corresponding expressions related to the hyperboloid H_p^q.

C. Hyperboloid H_2^p, $p > 2$

Replacing the function $Y^{l_2,\cdots,l_{[q/2]}}_{\tilde{m}_1,\cdots,\tilde{m}_{[q/2]}}(\tilde{\omega})$ by the function $(2\pi)^{-\frac{1}{2}} \exp(i\tilde{m}_1\tilde{\varphi}_1)$ and the index $\tilde{l}_{[q/2]}$ by $\tilde{m}_1$ (as $\tilde{l}_1 \equiv \tilde{m}_1$) and putting $q = 2$ in A we obtain the desired expressions related to the hyperboloid H_2^p.

D. Hyperboloid H_p^2, $p \geq 2$

First we change $q \rightleftarrows p$ in all the expressions of A and remove the tilde from any variable or index which previously had it and at the same time place a tilde over any variable or index which previously did not have it and put $q = 2$. Then we replace $Y^{l_2,\cdots,l_{[q/2]}}_{\tilde{m}_1,\cdots,\tilde{m}_{[q/2]}}(\tilde{\omega})$ by $(2\pi)^{-\frac{1}{2}} \exp(i\tilde{m}_1\tilde{\varphi}_1)$ and $\tilde{l}_{[q/2]}$ by $|\tilde{m}_1|$. The expressions obtained of the points (i)–(iv) are the desired expressions of the points (i)–(iv) of the present case. The points (v) and (vi) for H_p^2, $p \geq 2$ differ essentially from those in A and we therefore write them explicitly.

(v) The complete classification of the irreducible representations can be obtained using the spectra of the operators $\Delta(H_p^2)$, P, and T. The operator T is defined by $TY^{L,l_2,\cdots,l_{[p/2]},l_2,\cdots,l_{[q/2]}}_{m_1,\cdots,m_{[p/2]},\tilde{m}_1,\cdots,\tilde{m}_{[q/2]}}(\Omega) = \operatorname{sign} \tilde{m}_1 \cdot Y^{L,l_2,\cdots,l_{[p/2]},l_2,\cdots,l_{[q/2]}}_{m_1,\cdots,m_{[p/2]},\tilde{m}_1,\cdots,\tilde{m}_{[q/2]}}(\Omega)$ and $Tf = 0$ on the subspace $\mathfrak{H}^C(H_p^2) \subset \mathfrak{H}(H_p^2)$ corresponding to the continuous spectrum of the operator $\Delta(H_p^2)$. Let us denote by δ the triplet of indices σ, p, t, where t is the eigenvalue of T and by $\mathcal{N}_\delta$ that subset of $\mathcal{N}_\gamma$ for which the function

$$Y^{\sigma,l_2,\cdots,l_{[p/2]},l_2,\cdots,l_{[q/2]}}_{m_1,\cdots,m_{[p/2]},\tilde{m}_1,\cdots,\tilde{m}_{[q/2]}}(\Omega),\ l_2,\cdots,\tilde{m}_{[q/2]} \in \mathcal{N}_\gamma$$

belongs to the fixed "eigenspace" corresponding to the eigenvalue t. We define the unitary space $\mathfrak{D}^\delta$ and the representation of the group and algebra analogously as in A.

(vi) The decomposition of the quasi-regular representation looks like

$$U_g = \sum \oplus U_g^{L,(-)^{L+p},+} \oplus \sum \oplus U_g^{L,(-)^{L+p},-} \oplus \int \oplus d\Lambda U_g^{\Lambda,+,0} \oplus \int \oplus d\Lambda U_g^{\Lambda,-,0}$$

on the space

$$\mathfrak{H}(S) = \sum \oplus \mathfrak{H}^{L,(-)^{L+p},+} \oplus \sum \oplus \mathfrak{H}^{L,(-)^{L+p},-} \oplus \int \oplus d\Lambda \mathfrak{K}_g^{\Lambda,+,0} \oplus \int \oplus d\Lambda \mathfrak{K}_g^{\Lambda,-,0}.$$

E. Hyperboloid H_1^p, $p > 1$

(i) The spectrum of the Laplace–Beltrami operator[10] $\Delta(H_1^p)$ consists of the discrete spectrum $PS(\Delta(H_1^p))$: $-L(L+p-1)$, $L = -\{\frac{1}{2}(p-3)\}, -\{\frac{1}{2}(p-3)\}+1, \cdots$ and the continuous spectrum $CS(\Delta(H_1^p))$: $\Lambda^2 + (\frac{1}{2}(p-1))^2$, $\Lambda \in [0, \infty)$.

(ii) In the biharmonic coordinates (Sec. 5[1], Sec. 4[2]) the eigenfunctions are [keep in mind that $\theta \in (-\infty, \infty)$ in the present case]

$${}_1Y^{L,l_2,\cdots,l_{[p/2]}}_{m_1,\cdots,m_{[p/2]}}(\Omega) = \frac{-2\tanh\theta}{({}_1N)^{\frac{1}{2}}} \cosh\theta^{-(L+p-1)}\, {}_2F_1\left(\frac{L + l_{[p/2]} + p}{2}, \frac{L - l_{[p/2]} + 2}{2}; \frac{3}{2}, \tanh^2\theta\right) Y^{l_2,\cdots,l_{[p/2]}}_{m_1,\cdots,m_{[p/2]}}(\omega)$$

with

$$L - l_{[p/2]} = -(2n+2), \quad n = 0, 1, 2, \cdots. \tag{2.4}$$

$${}_2Y^{L,l_2,\cdots,l_{[p/2]}}_{m_1,\cdots,m_{[p/2]}}(\Omega) = \frac{1}{({}_2N)^{\frac{1}{2}}} \cosh\theta^{-(L+p-1)}\, {}_2F_1\left(\frac{L + l_{[p/2]} + p - 1}{2}, \frac{L - l_{[p/2]} + 1}{2}; \frac{1}{2}, \tanh^2\theta\right) Y^{l_2,\cdots,l_{[p/2]}}_{m_1,\cdots,m_{[p/2]}}(\omega)$$

with

$$L - l_{[p/2]} = -(2n+1), \quad n = 0, 1, 2, \cdots \tag{2.5}$$

$${}_1Y^{\Lambda,l_2,\cdots,l_{[p/2]}}_{m_1,\cdots,m_{[p/2]}}(\Omega) = \frac{-2\tanh\theta}{({}_1K)^{\frac{1}{2}}} \cosh\theta^{-[\frac{1}{2}(p-1)+i\Lambda]} \times {}_2F_1\left(\frac{i\Lambda + l_{[p/2]} + \frac{1}{2}(p+1)}{2}, \frac{i\Lambda - l_{[p/2]} - \frac{1}{2}(p-5)}{2}; \frac{3}{2}, \tanh^2\theta\right) Y^{l_2,\cdots,l_{[p/2]}}_{m_1,\cdots,m_{[p/2]}}(\omega),$$

$$ {}_2Y^{\Lambda,l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}}(\Omega) = \frac{1}{({}_2K)^{\frac{1}{2}}} \cosh\theta^{-[\frac{1}{2}(p-1)+i\Lambda]} \times {}_2F_1\left(\frac{i\Lambda + l_{\{p/2\}} + \frac{1}{2}(p-1)}{2}, \frac{i\Lambda - l_{\{p/2\}} - \frac{1}{2}(p-3)}{2}; \frac{1}{2}, \tanh^2\theta\right) Y^{l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}} $$

where

$$ {}_2N = \frac{\pi\Gamma[\frac{1}{2}(l_{\{p/2\}} - L + 1)]\Gamma[\frac{1}{2}(l_{\{p/2\}} + L + p)]}{[L + \frac{1}{2}(p-1)]\Gamma[\frac{1}{2}(l_{\{p/2\}} + L + p - 1)]\Gamma[\frac{1}{2}(l_{\{p/2\}} - L)]}, $$

$$ {}_1N = \frac{4\pi\Gamma[\frac{1}{2}(l_{\{p/2\}} - L)]\Gamma[\frac{1}{2}(l_{\{p/2\}} + L + p - 1)]}{(l_{\{p/2\}} - L - 1)(2L + p - 1)\Gamma[\frac{1}{2}(l_{\{p/2\}} + L + p)]\Gamma[\frac{1}{2}(l_{\{p/2\}} - L - 1)]}, $$

$$ {}_1K = \frac{\pi[\cosh(\pi\Lambda) - (-)^{l_{\{p/2\}}}\cos\frac{1}{2}(p-1)\pi]\,|\Gamma\{\frac{1}{2}[i\Lambda + l_{\{p/2\}} + \frac{1}{2}(p-1)]\}|^2}{\sinh(\pi\Lambda)\,|\Gamma\{\frac{1}{2}[i\Lambda + l_{\{p/2\}} + \frac{1}{2}(p+1)]\}|^2}, $$

$$ {}_2K = \frac{\pi[\cosh(\pi\Lambda) + (-)^{l_{\{p/2\}}}\cos\frac{1}{2}(p-1)\pi]\,|\Gamma\{\frac{1}{2}[i\Lambda + l_{\{p/2\}} + \frac{1}{2}(p+1)]\}|^2}{\sinh(\pi\Lambda)\,|\Gamma\{\frac{1}{2}[i\Lambda + l_{\{p/2\}} + \frac{1}{2}(p-1)]\}|^2}. $$

(iii) Completeness relations have the form

$$ \int_{H_1^p} {}_\alpha\overline{Y^{\sigma,l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}}(\Omega)}\,{}_\alpha Y^{\sigma',l_2',\cdots,l_{\{p/2\}}'}_{m_1',\cdots,m_{[p/2]}'}(\Omega)\,d\mu(\Omega) = \delta_{\alpha\alpha'}\,\delta_{\sigma\sigma'}\prod_{k=2}^{\{p/2\}}\delta_{l_k l_k'}\prod_{k=1}^{[p/2]}\delta_{m_k m_k'}, $$

where $\delta_{\sigma\sigma'} = \delta_{LL'}$ if $\sigma = L$, $\sigma' = L'$; $\delta_{\sigma\sigma'} = \delta(\Lambda - \Lambda')$ if $\sigma = \Lambda$, $\sigma' = \Lambda'$ and $\delta_{\sigma\sigma'} = 0$ otherwise.

$$ \sum_{L=-\{(p-3)/2\}}^{\infty}\sum_{\mathcal{N}_L} {}_\alpha\overline{Y^{L,l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}}(\Omega)}\,{}_\alpha Y^{L,l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}}(\Omega') + \int_0^\infty d\Lambda \sum_{\mathcal{N}_\Lambda} {}_\alpha\overline{Y^{\Lambda,l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}}(\Omega)}\,{}_\alpha Y^{\Lambda,l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}}(\Omega') = \delta(\Omega - \Omega';\mu), $$

where $\mathcal{N}_L$ is the set of values of indices $\alpha, l_2, \cdots, m_{[p/2]}$ restricted by (A4), (A5), (2.4), and (2.5) and $\mathcal{N}_\Lambda$ the set of values of indices $\alpha, l_2, \cdots, m_{[p/2]}$ restricted by (A4) and (A5).[12] The left invariant measure $d\mu(\Omega)$ is defined by

$$ d\mu(\Omega) = \cosh^{p-1}\theta\,d\theta\,d\mu(\omega_{p-1}), \quad \theta \in (-\infty, \infty). $$

(iv) The Fourier transform:

$$ \mathfrak{H}(H_1^p) \ni f \to \chi = Ff := \{\langle {}_\alpha Y^{\sigma,l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}}, f\rangle;\ \sigma \in \Sigma,\ l_2, \cdots, m_{[p/2]} \in \mathcal{N}_\sigma\} \in \mathfrak{H}(S). $$

(v) The unitary irreducible representations are classified by using both of the spectra of the operators $\Delta(H_1^p)$ and P. (See Sec. 3[2].) We denote by ϵ the pair α, p and by $\mathcal{N}_\epsilon$ that subset of the set $\mathcal{N}_\sigma$ for which either $\alpha + l_{\{p/2\}}$ is even and $p = 1$ or $\alpha + l_{\{p/2\}}$ is odd and $p = -1$. We consider the unitary space $\mathfrak{D}^\epsilon$ determined by the sequences

$$ \chi^\epsilon := \{\langle {}_\alpha Y^{\sigma,l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}}, f\rangle,\ \alpha, l_2, \cdots, m_{[p/2]} \in \mathcal{N}_\epsilon, f \in \mathfrak{D}(H_1^p)\}. $$

The unitary irreducible representation of the group $SO_0(p, 1)$ and the irreducible representation of the corresponding Lie algebra are defined on $\mathfrak{D}^\epsilon$ in an analogous way as in (a).

(vi) The decomposition of the quasi-regular representation is given by

$$ U_g = \sum\oplus U_g^{L,(-)^{L+1}} \oplus \int\oplus d\Lambda U_g^{\Lambda,+} \oplus \int\oplus d\Lambda U_g^{\Lambda,-} $$

on the Hilbert space

$$ \mathfrak{H}(S) = \sum\oplus \mathfrak{H}^{L,(-)^{L+1}} \oplus \int\oplus d\Lambda\mathfrak{H}^{\Lambda,+} \oplus \int\oplus d\Lambda\mathfrak{H}^{\Lambda,-} $$

F. Hyperboloid H_p^1, $p > 1$ (Upper sheet of H_p^1)

The Laplace–Beltrami operator[10] $\Delta(H_p^1)$ has only a continuous spectrum $S(\Delta(H_p^1))$: $\Lambda^2 + [\frac{1}{2}(p-1)]^2$, $\Lambda \in [0, \infty)$. The corresponding eigenfunctions are

$$ Y^{\Lambda,l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}}(\Omega) = \frac{1}{N^{\frac{1}{2}}} \tanh\theta^{|l_{\{p/2\}}|}\cosh\theta^{i\Lambda - \frac{1}{2}(p-1)} \times {}_2F_1\left(\frac{|l_{\{p/2\}}| - i\Lambda + \frac{1}{2}(p-1)}{2}, \frac{|l_{\{p/2\}}| - i\Lambda + \frac{1}{2}(p+1)}{2}; |l_{\{p/2\}}| + \frac{1}{2}p; \tanh^2\theta\right) Y^{l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}}(\omega), $$

$$ N = \left|\frac{(2\pi)^{\frac{1}{2}}\Gamma(i\Lambda)\cdot\Gamma(|l_{\{p/2\}}| + \frac{1}{2}p)}{\Gamma\{\frac{1}{2}[i\Lambda + \frac{1}{2}(p-1) + |l_{\{p/2\}}|]\}\Gamma\{\frac{1}{2}[i\Lambda + \frac{1}{2}(p+1) + |l_{\{p/2\}}|]\}}\right|^2. $$

The completeness relations and the Fourier transform have the same structure as in A. The unitary irreducible representations of $SO_0(p, 1)$ related to the upper sheet of the hyperboloid H_p^1 are classified by Λ, $\Lambda \in [0, \infty)$. The unitary space $\mathfrak{D}^\Lambda$ is determined by the sequences

$$\chi^\Lambda := \{\langle Y^{\Lambda, l_2, \cdots, l_{[p/2]}}_{m_1, \cdots, m_{[p/2]}}, f\rangle, \quad l_2, \cdots, m_{[p/2]} \in \mathcal{N}_\Lambda, f \in \mathfrak{D}(H_p^1)\}.$$

The unitary irreducible representations of the group and the irreducible representation of the algebra are defined as in A. The decomposition of the quasi-regular representation is

$$U_g = \int \oplus\, d\Lambda U_g^\Lambda \quad \text{on} \quad \mathfrak{H}(S) = \int \oplus\, d\Lambda \mathfrak{H}^\Lambda.$$

G. Cone C_q^p (Upper sheet of C_1^p, $p > 1$)

(i) The second-order invariant Q_2 (4.2)^2 has only the continuous spectrum $S(Q_2)$: $\Lambda^2 + [\frac{1}{2}(p + q - 2)]^2$, $\Lambda \in (-\infty, \infty)$.

(ii) The corresponding eigenfunctions are

$$Y^{\Lambda, l_2, \cdots, l_{[p/2]}, l_2, \cdots, l_{[q/2]}}_{m_1, \cdots, m_{[p/2]}, \tilde{m}_1, \cdots, \tilde{m}_{[q/2]}}(\Omega) = (2\pi)^{-\frac{1}{2}} r^{i\Lambda - \frac{1}{2}(p+q-2)} Y^{l_2, \cdots, l_{[p/2]}}_{m_1, \cdots, m_{[p/2]}}(\omega) Y^{l_2, \cdots, l_{[q/2]}}_{\tilde{m}_1, \cdots, \tilde{m}_{[q/2]}}(\tilde{\omega}),$$

where $r \in (0, \infty)$.

(iii) The completeness relations have the form[12]

$$\int_{C_q^p} \overline{Y^{\Lambda, l_2, \cdots, l_{[p/2]}, l_2, \cdots, l_{[q/2]}}_{m_1, \cdots, m_{[p/2]}, \tilde{m}_1, \cdots, \tilde{m}_{[q/2]}}(\Omega)} Y^{\Lambda', l_2', \cdots, l_{[p/2]}', l_2', \cdots, l_{[q/2]}'}_{m_1', \cdots, m_{[p/2]}', \tilde{m}_1', \cdots, \tilde{m}_{[q/2]}'}(\Omega)\, d\mu(\Omega) = \delta(\Lambda - \Lambda') \prod_{k=2}^{[p/2]} \delta_{l_k l_k'} \prod_{k=1}^{[p/2]} \delta_{m_k m_k'} \prod_{k=2}^{[q/2]} \delta_{l_k l_k'} \prod_{k=1}^{[q/2]} \delta_{\tilde{m}_k \tilde{m}_k'},$$

$$\int_0^\infty d\Lambda \sum_{\mathcal{N}_\Lambda} \overline{Y^{\Lambda, l_2, \cdots, l_{[p/2]}, l_2, \cdots, l_{[q/2]}}_{m_1, \cdots, m_{[p/2]}, \tilde{m}_1, \cdots, \tilde{m}_{[q/2]}}(\Omega)} Y^{\Lambda, l_2, \cdots, l_{[p/2]}, l_2, \cdots, l_{[q/2]}}_{m_1, \cdots, m_{[p/2]}, \tilde{m}_1, \cdots, \tilde{m}_{[q/2]}}(\Omega') = \delta(\Omega - \Omega'; \mu),$$

where the left invariant measure on C_q^p is

$$d\mu(\Omega) = r^{p+q-3}\, dr\, d\mu(\omega_{p-1})\, d\mu(\tilde{\omega}_{q-1}), \quad r \in (0, \infty).$$

(iv) The Fourier transform is defined by

$$\mathcal{H}(C_q^p) \ni f \to \chi = Ff := \{\langle Y^{\Lambda, l_2, \cdots, l_{[p/2]}, l_2, \cdots, l_{[q/2]}}_{m_1, \cdots, m_{[p/2]}, \tilde{m}_1, \cdots, \tilde{m}_{[q/2]}}, f\rangle; \Lambda \in (-\infty, \infty), l_2, \cdots, \tilde{m}_{[q/2]} \in \mathcal{N}_\Lambda\} \in \mathfrak{H}(S).$$

(v) The unitary irreducible representations are classified by Λ and if $p > q = 1$ also by an eigenvalue of P. Let us denote by ξ the pair Λ, p, $\Lambda \in (-\infty, \infty)$, $p \in \{\pm 1\}$, and let us consider the unitary space $\mathfrak{D}^\xi$ determined by the sequences

$$\chi^\xi := \{\langle Y^{\Lambda, l_2, \cdots, l_{[p/2]}, l_2, \cdots, l_{[q/2]}}_{m_1, \cdots, m_{[p/2]}, \tilde{m}_1, \cdots, \tilde{m}_{[q/2]}}, f\rangle, \quad l_2, \cdots, \tilde{m}_{[q/2]} \in \mathcal{N}_\xi, f \in \mathfrak{D}(C_q^p)\},$$

where $\mathcal{N}_\xi$ is the subset of the set $\mathcal{N}_\Lambda$ restricted by the condition that either $l_{[p/2]} + \tilde{l}_{[q/2]}$ is even and $p = 1$ or $l_{[p/2]} + \tilde{l}_{[q/2]}$ is odd and $p = -1$. The unitary irreducible representation of the group and the irreducible representation of the algebra are defined as in A.

(vi) The decomposition of the quasi-regular representation into the irreducible ones has the form

$$U_g = \int_{-\infty}^{+\infty} \oplus\, d\Lambda U_g^{\Lambda,+} \oplus \int_{-\infty}^{+\infty} \oplus\, d\Lambda U_g^{\Lambda,-}$$

on

$$\mathfrak{H}(S) = \int_{-\infty}^{+\infty} \oplus\, d\Lambda \mathfrak{H}^{\Lambda,+} \oplus \int_{-\infty}^{+\infty} \oplus\, d\Lambda \mathfrak{H}^{\Lambda,-}$$

3. ESSENTIAL SELF-ADJOINTNESS OF $\Delta(X)$ ON $\mathfrak{D}(X)$

To simplify the notations we denote by l, m, $\tilde{l}$, and $\tilde{m}$ the sets $l_2, \cdots, l_{[p/2]}; m_1, \cdots, m_{[p/2]}, \tilde{l}_2, \cdots, \tilde{l}_{[q/2]}$ and $\tilde{m}_1, \cdots, \tilde{m}_{[q/2]}$, respectively. The set of values of the indices $l_2, \cdots, l_{[p/2]}, m_1, \cdots, m_{[p/2]}, \tilde{l}_2, \cdots, \tilde{l}_{[q/2]}$, $\tilde{m}_1, \cdots, \tilde{m}_{[q/2]}$, i.e., the set of integers restricted by the conditions (A4), (A5), we denote by $\mathcal{N}$.

The restriction of the differential operator $\Delta(X)$ which was derived in Refs. 1 and 2 to the dense invariant linear manifold $\mathfrak{D}(X)$ is denoted by Δ^s. The operator Δ^s is symmetric on $\mathfrak{D}(X)$. We prove in this section that Δ^s is the essentially self-adjoint operator on $\mathfrak{D}(X)$ for any particular case of the homogeneous space X.

A. Hyperboloid H_q^p, $p \geq q > 1$

The operator

$$P_{lml\tilde{m}}: \mathfrak{D}(X) \ni f \to P_{lml\tilde{m}} f := \{Y^l_m(\omega) Y^{\tilde{l}}_{\tilde{m}}(\tilde{\omega}) \times \int_X \overline{Y^l_m(\omega') Y^{\tilde{l}}_{\tilde{m}}(\tilde{\omega}')} f(\Omega')\, d\mu(\omega')\, d\mu(\tilde{\omega}'), \Omega \in X\} \in \mathfrak{D}(X)$$

is a bounded symmetric operator on $\mathfrak{D}(X)$ satisfying $P^2_{lml\tilde{m}} = P_{lml\tilde{m}}$ on $\mathfrak{D}(X)$. As $\mathfrak{D}(X)$ is a dense linear manifold in $\mathfrak{H}(X)$, $P_{lml\tilde{m}}$ has the unique extension to the projector $P_{lml\tilde{m}}$ on $\mathfrak{H}(X)$, the range of which is the closed subspace $\mathfrak{H}_{lml\tilde{m}}(X)$. The subspaces $\mathfrak{H}_{lml\tilde{m}}(X)$ with a different set of indices are orthogonal to each other and their sum over all l, m, $\tilde{l}$, $\tilde{m} \in \mathcal{N}$ is equal to the space $\mathfrak{H}(X)$. This follows from the density of $\mathfrak{D}(X)$ in $\mathfrak{H}(X)$ and the completeness of the orthogonal set of harmonic functions $Y^l_m(\omega)$ in the Hilbert space $\mathfrak{H}(S^{p-1})$. [For the completeness of the set of $Y^l_m(\omega)$ see Appendix II.]

As $P_{lml\tilde{m}}$ maps $\mathfrak{D}(X)$ into $\mathfrak{D}(X)$ and $\Delta^s P_{lml\tilde{m}} - P_{lml\tilde{m}} \Delta^s = 0$ on $\mathfrak{D}(X)$ every subspace $\mathfrak{H}_{lml\tilde{m}}(X)$

reduces Δ^s to the symmetric operator

$$D^s_{lml\tilde{m}}:\mathfrak{D}(D^s_{lml\tilde{m}})\ni f\to D^s_{lml\tilde{m}}f$$
$$:=\Big\{-\frac{1}{\cosh^{p-1}\theta\sinh^{q-1}\theta}\frac{\partial}{\partial\theta}\cosh^{p-1}\theta\sinh^{q-1}\theta\frac{\partial}{\partial\theta}f(\Omega)$$
$$+\frac{l_{\{q/2\}}(l_{\{q/2\}}+q-2)}{\sinh^2\theta}f(\Omega)$$
$$-\frac{l_{\{p/2\}}(l_{\{p/2\}}+p-2)}{\cosh^2\theta}f(\Omega),\ \Omega\in X\Big\}\in\mathfrak{H}_{lml\tilde{m}}, \tag{3.1}$$

where $\mathfrak{D}(D^s_{lml\tilde{m}})=\mathfrak{D}(X)\cap\mathfrak{H}_{lml\tilde{m}}(X)$.

Proposition 3.1: The operator $\mathfrak{D}^s_{lml\tilde{m}}$ is essentially self-adjoint on $\mathfrak{D}(D^s_{lml\tilde{m}})$.

Proof: Let us define the mapping

$$V:\mathfrak{H}_{lml\tilde{m}}(X)\cap\mathfrak{D}(X)\ni f\to Vf$$
$$:=\Big\{[d\mu(\theta)/d\theta]^{\frac{1}{2}}\int Y^l_m(\omega)Y^l_{\tilde{m}}(\tilde{\omega})f(\Omega)\,d\mu(\omega)\,d\mu(\tilde{\omega}),$$
$$\theta\in(0,\infty)\Big\}\in\mathfrak{L}^2(0,\infty),$$

where by $\mathfrak{L}^2(a,b)$ we denote the Hilbert space of L^2 type over the interval (a,b). The mapping V can be extended to the isometric mapping of the Hilbert space $\mathfrak{H}_{lml\tilde{m}}(X)$ on to the Hilbert space $\mathfrak{L}^2(0,\infty)$ such that $V^{-1}V=I$ on $\mathfrak{H}_{lml\tilde{m}}(X)$ and $VV^{-1}=I$ on $\mathfrak{L}^2(0,\infty)$. The symmetric operator $D^s_{lml\tilde{m}}$ on $\mathfrak{D}(D^s_{lml\tilde{m}})$ is mapped by V to the symmetric operator

$$A^s:\mathfrak{D}(A^s)\ni f\to A^sf$$
$$:=\Big\{\Big(-\frac{d^2}{d\theta^2}+K^{p,q}_{l_{\{p/2\}}l_{\{q/2\}}}(\theta)\Big)f(\theta),\ \theta\in(0,\infty)\Big\}$$
$$\in\mathfrak{L}^2(0,\infty), \tag{3.2}$$

where $\mathfrak{D}(A^s):=V\mathfrak{D}(D^s_{lml\tilde{m}})$ and

$$K^{p,q}_{l_{\{p/2\}}l_{\{q/2\}}}(\theta)=\frac{l_{\{q/2\}}(l_{\{q/2\}}+q-2)+[\frac{1}{2}(q-2)]^2-\frac{1}{4}}{\sinh^2\theta}$$
$$-\frac{l_{\{p/2\}}(l_{\{p/2\}}+p-2)+[\frac{1}{2}(p-2)]^2-\frac{1}{4}}{\cosh^2\theta}$$
$$+\Big(\frac{p+q-2}{2}\Big)^2. \tag{3.3}$$

As the properties of the corresponding objects in the Hilbert spaces $\mathfrak{H}_{lml\tilde{m}}(X)$ and $\mathfrak{L}^2(0,\infty)$ are equivalent, we study from now on the symmetric operator A^s on $\mathfrak{D}(A^s)$ for which the well-established theory of the second-order differential operators can be applied.

Let A be the restriction of A^s to

$$\mathfrak{D}(A):=\{f\in\mathfrak{D}(A^s)\mid f(0)=0,\ f'(0)=0\}.$$

The closure $\tilde{A}$ is self-adjoint on $\mathfrak{D}(\tilde{A})$ if and only if the differential equation

$$d^2\psi(\theta)/d\theta=(K^{p,q}_{l_{\{p/2\}},l_{\{q/2\}}}-\lambda)\psi(\theta) \tag{3.4}$$

for any nonreal λ and any $c>0$ possesses only one solution $\psi_0(\cdot)\in\mathfrak{L}^2(0,c)$ and only one solution $\psi_\infty(\cdot)\in\mathfrak{L}^2(c,\infty)$.[13] If $l_{\{q/2\}}\neq 0$, Eq. (3.4) has the property mentioned so that A^s is essentially self-adjoint on $\mathfrak{D}(A^s)$ in this case. If $l_{\{q/2\}}=0$ the assertion still holds if $q>3$. Finally if $l_{\{q/2\}}=0$ and $q=2$ or $q=3$, both linearly independent solutions of Eq. (3.4) are from $\mathfrak{L}^2(0,c)$. This means that $\tilde{A}$ is not self-adjoint on $\mathfrak{D}(\tilde{A})$ and it does not follow automatically that A^s is self-adjoint on $\mathfrak{D}(A^s)$.

We now prove that A^s is essentially self-adjoint on $\mathfrak{D}(A^s)$ even in two exceptional cases. Let us consider the case $q=2$. The domain $\mathfrak{D}(A^*)$ of the adjoint operator A^* is determined by vectors $g:=\{g(\theta),\ \theta\in(0,\infty)\}\in\mathfrak{L}^2(0,\infty)$ for which $g(\theta)$ is differentiable, $g'(\theta)$ is absolutely continuous and

$$A^*g:=\{-g''(\theta)+K^{p,q}_{l_{\{p/2\}}l_{\{q/2\}}}(\theta)g(\theta),\ \theta\in(0,\infty)\}$$
$$\in\mathfrak{L}^2(0,\infty).$$

Let us derive the behavior of the function $g(\theta)$, $g\in\mathfrak{D}(A^*)$ at the origin. Every continuous solution $g(\theta)$ of the differential equation

$$g''(\theta)-K^{p,q}_{l_{\{p/2\}}l_{\{q/2\}}}(\theta)g(\theta)=h(\theta),\quad h\in\mathfrak{L}^2(0,\infty) \tag{3.5}$$

in the interval $(0,1]$ has the form

$$g(\theta)=au_1(\theta)+bu_2(\theta)$$
$$+\int_0^\theta\frac{u_1(\theta)u_2(\theta')-u_2(\theta)u_1(\theta')}{W(u_1,u_2)}h(\theta')\,d\theta', \tag{3.6}$$

where $u_1(\theta)$ and $u_2(\theta)$ are two linearly independent solutions of Eq. (3.5) for $h(\theta)\equiv 0$, $W(u_1,u_2)=u_1'(\theta)u_2(\theta)-u_2'(\theta)u_1(\theta)$, $\theta\in(0,1]$ and a, b are complex constants. The existence of the solution (3.6) in any closed interval $[\theta_0,1]$, $\theta_0>0$, can be proved using analyticity of functions $u_{1,2}(\theta)$, integrability of $h(\theta)$ in $(0,1)$ and the Titchmarsh lemma.[14]

Let us choose as $u_1(\theta)$ and $u_2(\theta)$ those solutions of $\psi''-K\psi=0$ for which

$$\lim_{\theta\to0}\frac{u_1(\theta)}{\theta^{\frac{1}{2}}}=1,\quad\lim_{\theta\to0}\frac{u_2(\theta)}{\theta^{\frac{1}{2}}\ln\theta}=1,$$
$$\lim_{\theta\to0}2\theta^{\frac{1}{2}}u_1'(\theta)=1,\quad\lim_{\theta\to0}2\frac{\theta^{\frac{1}{2}}}{\ln\theta}u_2'(\theta)=1.$$

One can easily verify by direct calculation the possibility of this choice. Then by rough estimate of the third term in (3.6) we find that this term vanishes faster than $\theta^{\frac{1}{2}}$ at the origin and its derivative faster than $1/\theta^{\frac{1}{2}}$, so that the first two terms describe the behavior of $g(\theta)$ at the origin. The domain $\mathfrak{D}(\tilde{A})$ of the closure $\tilde{A}$ is the linear manifold determined by those $f\in\mathfrak{D}(A^*)$ for which $B(f,g)=0$, $g\in\mathfrak{D}(A^*)$, where[13]

$$B(f,g)=\lim_{\theta\to0}[\overline{f(\theta)}\,g'(\theta)-g(\theta)\overline{f'(\theta)}].$$

[13] M. H. Stone, *Linear Transformations in Hilbert Space* (American Mathematical Society, New York, 1964), Colloq. Publ., Vol. XV, Chap. X; M. A. Neumark, *Lineare Differentialoperatoren* (Akademir-Verlag, Berlin, 1963), Chap. V.

[14] Lemma 5.6 of Ref. 5.

As the domain $\mathfrak{D}(\tilde{A}^s)$ of the closure $\tilde{A}^s$ of the symmetric operator A^s on $\mathfrak{D}(A^s)$ contains $\mathfrak{D}(A^s)$, it has to be $B(f_0, g) = 0$ for

$$f_0 := \{(\sinh\theta)^{\frac{1}{2}}\cosh^{\frac{1}{2}(p-1)}\theta P(\cosh\theta) \times \exp(-2\cosh^2\theta),\ \theta\in(0,\infty),\ P(1)=1\}\in\mathfrak{D}(A^s).$$

This condition implies $b = 0$ in (3.6). Then any two vectors $g_1, g_2 \in \mathfrak{D}(A^{s*})$ satisfy $B(g_1, g_2) = 0$, i.e., the adjoint A^{s**} of A^{s*} is defined on $\mathfrak{D}(A^{s*})$, i.e., A^{s**} is self-adjoint on $\mathfrak{D}(A^{s**})$. In other words, A^s is essentially self-adjoint on $\mathfrak{D}(A^s)$. The analogous proof holds for $q = 3$, $l_2 = 0$.

From Proposition 3.1 and the decomposition

$$\mathfrak{H}(X) = \sum_{\mathcal{N}} \mathfrak{H}_{lml\tilde{m}}(X)$$

follows Proposition 3.2.

Proposition 3.2: The operator Δ^s is essentially self-adjoint on $\mathfrak{D}(X)$. The domain $\mathfrak{D}(\Delta^{sa})$ of the self-adjoint operator $\Delta^{sa} = \tilde{\Delta}^s$ is determined by vectors $f \in \mathfrak{H}(X)$ for which $P_{lml\tilde{m}} f \in \mathfrak{D}(D^{sa}_{lml\tilde{m}})$ and $\Sigma \| D^{sa}_{lml\tilde{m}} P_{lml\tilde{m}} f\|^2 < \infty$. The operator Δ^{sa} on $\mathfrak{D}(\Delta^{sa})$ is defined by $\Delta^{sa} f = \sum D^{sa}_{lml\tilde{m}} P_{lml\tilde{m}} f$.

B. Other Hyperboloids H^p_q and H^q_p

As the form (3.2) of the operator A^s holds for any type of hyperboloid, the proof of the essential self-adjointness of Δ^s on $\mathfrak{D}(X)$ can be carried out in an analogous way to A.

C. Cone C^p_q, $p \geq q \geq 1$

The operator A^s in the present case is defined by

$$\mathfrak{D}(A^s) = V\big(\mathfrak{D}(X) \cap \mathfrak{H}_{lml\tilde{m}}(X)\big) \ni f \to A^s f := \left\{ -\frac{d}{dr} r^2 \frac{d}{dr} f(r) + \left[\left(\frac{p+q-2}{2}\right)^2 - \frac{1}{4}\right] f(r),\ r\in(0,\infty)\right\} \in \mathfrak{D}(A^s).$$

The associated differential equation

$$\frac{d}{dr} r^2 \frac{d}{dr}\psi(r) - \left[\left(\frac{p+q-2}{2}\right)^2 - \frac{1}{4} - \lambda\right]\psi(r) = 0 \tag{3.7}$$

has the solutions

$$\psi^\lambda_{1,2}(r) = \exp\{(-\tfrac{1}{2} \pm \{[\tfrac{1}{2}(p+q-2)]^2 - \lambda\}^{\frac{1}{2}})\ln r\}.$$

Again $\tilde{A}^s$ is the self-adjoint on $\mathfrak{D}(\tilde{A}^s)$ as for nonreal λ one of $\psi^\lambda_i(r)$, $i = 1, 2$, has the property $\psi^\lambda_i(\cdot) \in \mathfrak{L}^2(0, c)$ for every $c > 0$, whereas the other has not this property and the same relation holds at infinity.

4. EIGENFUNCTION EXPANSION

In two previous works[1,2] we used the set of those eigenfunctions $Y^{\lambda l l}_{m\tilde{m}}(\Omega)$ with the eigenvalue λ of the differential operator $\Delta(X)$ or Q_2 which were restricted by the boundary condition $Y^{\lambda l l}_{m\tilde{m}}(\Omega_0) = 0$, where $\Omega_0 = \{\theta = 0, \omega, \tilde{\omega}\}$. Let us prove now that this set constitutes the complete set of eigenfunctions of the self-adjoint operator Δ^{sa} which was defined in the previous section.

A. Hyperboloid H^p_q, $p \geq q > 1$

We have shown that the operator Δ^{sa} is reduced to the operator $D^{sa}_{lml\tilde{m}}$ on every subspace $\mathfrak{H}_{lml\tilde{m}}(X)$. Then by the mapping V the operator $D^{sa}_{lml\tilde{m}}$ has been mapped to the operator A^{sa} (3.2) and the space $\mathfrak{H}_{lml\tilde{m}}(X)$ on to the space $\mathfrak{L}^2(0, \infty)$. In this way the problem of the eigenfunction expansions associated with the operator Δ^{sa} is reduced to the investigation of the eigenfunction expansions associated with the operator A^{sa} or, in other words, associated with the second-order ordinary differential equation (3.4). Except for $l_{\{q/2\}} = 0$, $q = 2$ and $l_{\{q/2\}} = 0$, $q = 3$ there is the unique solution of the eigenfunction expansion associated with the differential equation (3.4)[6] and this expansion is just that one which is associated with the self-adjoint operator A^{sa}. In two exceptional cases we have to impose boundary conditions in such a way that the corresponding expansions are associated with the self-adjoint operator A^{sa}. For $l_{\{q/2\}} = 0$, $q = 2$, the behavior of the function $g(\theta)$, $g \in \mathfrak{D}(A^{sa})$, is described by (3.6) where $b = 0$ as proved. Since $B(u_1, g) = 0$ according to the proof of Proposition 3.1, we conclude that[6] $f_0(\lambda) = \infty$ and consequently only the function $u_1(\theta)$ enters the expression for the eigenfunction expansion.[6] A similar conclusion holds for the case $l_{\{q/2\}} = 0$, $q = 3$: only that eigenfunction $\psi^\lambda(\theta)$ of (3.4) enters the expression for the eigenfunction expansions for which $\psi^\lambda(0) = 0$.

The differential equation (3.4) has the form of the Schrödinger differential equation and we prefer to use the method of Jost and Kohn[15] adopted for such equations rather than the general method.[5,6] The solution of the differential equation (3.4) which is regular at the origin has the form

$$\psi(\theta,\lambda) = (\tanh\theta)^{l_{\{q/2\}}+\frac{1}{2}(q-1)}(\cosh\theta)^{-\{[\frac{1}{2}(p+q-2)]^2-\lambda\}^{\frac{1}{2}}} \times F(a_+, b_+; c; \tanh^2\theta), \tag{4.1}$$

where $c = l_{\{q/2\}} + q/2$ and

$$\begin{aligned} 2a_\pm &= l_{\{q/2\}} + l_{\{p/2\}} + \tfrac{1}{2}(p+q-2) \\ &\quad \pm \{[\tfrac{1}{2}(p+q-2)]^2 - \lambda\}^{\frac{1}{2}}, \\ 2b_\pm &= l_{\{q/2\}} - l_{\{p/2\}} + \tfrac{1}{2}(p+q-2) \\ &\quad \pm \{[\tfrac{1}{2}(p+q-2)]^2 - \lambda\}^{\frac{1}{2}}. \end{aligned} \tag{4.2}$$

The solutions of (3.4) which behave like

$$\exp\{\mp\theta([\tfrac{1}{2}(p+q-2)]^2 - \lambda)^{\frac{1}{2}}\}$$

are

$$V_\pm(\theta,\lambda) = (\tanh\theta)^{l_{\{q/2\}}+\frac{1}{2}(q-1)}(\cosh\theta)^{\mp\{[\frac{1}{2}(p+q-2)]^2-\lambda\}^{\frac{1}{2}}} \times F\left(a_\pm, b_\pm; 1 \pm \{[\tfrac{1}{2}(p+q-2)]^2 - \lambda\}^{\frac{1}{2}}; \frac{1}{\cosh^2\theta}\right). \tag{4.3}$$

[15] R. Jost and W. Kohn, Phys. Rev. **87**, 977 (1952).

These two solutions are generally singular at the origin. The connection among the solutions (4.1) and (4.3) can be easily described by introducing the variable $\mu = -\frac{1}{2}(p+q-2) + \{[\frac{1}{2}(p+q-2)]^2 - \lambda\}^{\frac{1}{2}}$ and denoting $\varphi(\theta,\mu) = \psi(\theta,\lambda)$, $u_\pm(\theta,\mu) = V_\pm(\theta,\lambda)$.

$$\varphi(\theta,\mu) = W_-(\mu)u_-(\theta,\mu) + W_+(\mu)u_+(\theta,\mu), \quad (4.4)$$

where

$$W_\pm(\mu) = 2^{\pm[\frac{1}{2}(p+q-2)+\mu]} \times \frac{\Gamma(c)\Gamma\{\pm[\frac{1}{2}(p+q-2)+\mu]\}}{\Gamma[a_\pm(\mu)]\Gamma[b_\pm(\mu)]}. \quad (4.5)$$

Let us introduce now the function

$$G(\mu,\theta,\theta') = \begin{cases} \dfrac{\varphi(\theta,\mu)u_-(\theta',\mu)}{W_+(\mu)} & \text{for } \theta' \geq \theta, \\ \dfrac{\varphi(\theta',\mu)u_-(\theta,\mu)}{W_+(\mu)} & \text{for } \theta' \leq \theta. \end{cases} \quad (4.6)$$

The function $G(\mu,\theta,\theta')$ is analytic in the half-plane $\operatorname{Re}\mu > -\frac{1}{2}(p+q-2)$ having poles at zeros of the function $W_+(\mu)$, i.e., at the points $b_+(\mu) = -n$. The number of poles is finite and we choose R in the contour C of Fig. 1 large enough to enclose all the poles. Let us multiply the equality

$$\oint_C G(\mu,\theta,\theta')\,d\mu = 2\pi i \sum_n \operatorname{Re}\, G(\mu_n,\theta,\theta')$$

by $f(\theta')$ and integrate over the interval $(0,\infty)$, where $f(\theta')$ is an infinite differentiable function vanishing faster than any power in θ' for large θ'. The limiting value of the expression obtained as $R\to\infty$ has the form

$$f(\theta) = \sum_n \frac{\varphi(\theta,\mu_n)}{N}\int \frac{\varphi(\theta',\mu_n)}{N} f(\theta')\,d\theta' + \int_0^\infty \frac{d\sigma}{2\pi\,|W_+[-\frac{1}{2}(p+q-2)+i\sigma]|^2} \times \varphi[\theta,-\tfrac{1}{2}(p+q-2)+i\sigma] \times \int_0^\infty \overline{\varphi[\theta',-\tfrac{1}{2}(p+q-2)+i\sigma]} f(\theta')\,d\theta'. \quad (4.7)$$

Here the function $W_\pm(\mu)$, $\mu = -\frac{1}{2}(p+q-2) + \{[\frac{1}{2}(p+q-2)]^2 - \lambda\}^{\frac{1}{2}}$ is essentially the Titchmarsh spectral function $m(\mu)$ which determines the spectrum of the associated differential operator.

Now we can write the eigenfunction expansion of an arbitrary function $f(\Omega)$, $f\in\mathfrak{D}(X)$. The decomposition

$$f = \sum_{\mathcal{N}} f_{lml\tilde{m}}, \quad f\in\mathfrak{D}(X), \quad f_{lml\tilde{m}}\in\mathfrak{D}(X)\cap\mathfrak{H}_{lml\tilde{m}}(X)$$

has a finite number of terms different from zero. Hence for the eigenfunction expansion of the function $f(\Omega)$, $f\in\mathfrak{D}(X)$ it is sufficient to consider the eigenfunction expansion of $f_{lml\tilde{m}}(\Omega)$, $f_{lml\tilde{m}}\in\mathfrak{D}(X)\cap\mathfrak{H}_{lml\tilde{m}}(X)$. This latter problem is already solved, as the eigenfunction expansion of $f_{lml\tilde{m}}(\Omega)$ is an image under V^{-1} of the equality (4.7):

$$f_{lml\tilde{m}}(\Omega) = \sum Y^{Lil}_{m\tilde{m}}(\Omega)\int \overline{Y^{Lil}_{m\tilde{m}}(\Omega')} f_{lml\tilde{m}}(\Omega')\,d\mu(\Omega') + \int d\Lambda\, Y^{\Lambda il}_{m\tilde{m}}(\Omega)\int \overline{Y^{\Lambda il}_{m\tilde{m}}(\Omega')} f_{lml\tilde{m}}(\Omega')\,d\mu(\Omega'), \quad (4.8)$$

where the functions $Y^{Lil}_{m\tilde{m}}(\Omega)$ and $Y^{\Lambda il}_{m\tilde{m}}(\Omega)$ are defined in Sec. 2. We remark here that the equality (4.8) is valid for every point $\Omega\in X$.

The eigenfunction expansion of an arbitrary $f(\Omega)$, $f\in\mathfrak{H}(X)$ can be obtained using Cauchy sequence $f_n\to f$, $f_n\in\mathfrak{D}(X)$.

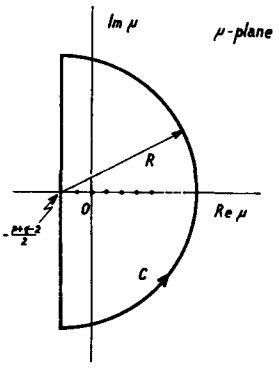


FIG. 1. The contour of the integration for the function $G(\mu,\theta,\theta')$.

B. Hyperboloids H^p_1 and H^1_p

The eigenfunction expansion associated with the operator $\Delta^{sa}(H^1_p)$ is the special case of the eigenfunction expansions derived in A. The eigenfunction expansions associated with the operator $\Delta^{sa}(H^p_1)$ can be derived in an analogous way to that of A. In the course of the proof the required eigenfunction expansion associated with the operator A^{sa} is already calculated in Example 4.19 of Ref. 3.

C. Cone C^p_q, $p\geq q\geq 1$

The eigenfunction expansion associated with the differential equation (3.7) can again be constructed using the general theory. But we can avoid the tedious computations of both the Titchmarsh spectral function and the corresponding linearly independent solutions if we start from the Mellin transform

$$f(r) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} ds r^{-s}\int_0^\infty r'^{s-1} f(r')\,dr'. \quad (4.9)$$

Choosing $c=\frac{1}{2}$ and $s=\frac{1}{2}+i\Lambda$ and denoting $\varphi(r,\Lambda) = r^{-\frac{1}{2}+i\Lambda}$ we obtain the following eigenfunction expansion associated with the differential equation (3.7):

$$f(r) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} d\Lambda\varphi(r,\Lambda)\int_0^\infty \varphi(r',\Lambda) f(r')\,dr', \quad (4.10)$$

where $f(r')$ is a function which has a derivative of any

order and vanishes at infinity faster than any power in r'.

Now the eigenfunction expansion of an arbitrary function $f(\Omega), f \in \mathfrak{D}(X)$, can be obtained using (4.10), the mapping V^{-1} and the decomposition $\mathfrak{H}(X) = \sum \mathfrak{H}_{lmlm}(X)$. The expansion of a function $f(\Omega)$, $f \in \mathfrak{H}(X)$, has to be understood in the sense of the Cauchy convergence $f_n \to f, f_n \in \mathfrak{D}(X)$.

In the following considerations we deal always with the hyperboloid H_q^p, $p \geqslant q > 2$. However, our statements are also true for other cases.

The nondecreasing function on $S(\Delta(H_q^p)), p \geqslant q > 2$

$$\mu(\lambda) = \begin{cases} \{\lambda - [\frac{1}{2}(p+q-2)]^2\}^{\frac{1}{2}} & \text{for} \quad \lambda \geq [\frac{1}{2}(p+q-2)]^2, \\ 0 & \text{for} \quad \lambda_{L_{\min}} \leqslant \lambda < [\frac{1}{2}(p+q-2)]^2, \\ -(L - L_{\min} + 1) & \text{for} \quad \lambda_{L+1} \leqslant \lambda < \lambda_L, \end{cases} \tag{4.11}$$

where $\lambda_L = -L(L+p+q-2)$, defines a measure on the semi-ring consisting of the empty set ϕ and the cells $(\lambda_1, \lambda_2]$: $\mu((\lambda_1, \lambda_2]) = \mu(\lambda_2) - \mu(\lambda_1)$.[16] The extension procedure furnishes the σ ring of all σ-measurable sets $\mathbf{E}$ with the Stieltjes–Lebesque measure $\mu(\mathbf{E})$. The set of all objects

$$\chi := \{\chi_{m\tilde{m}}^{\lambda l \bar{l}};\ \lambda \in S(\Delta^{sa}),\ l, m, \bar{l}, \tilde{m} \in \mathcal{N}_\lambda\},$$

where $\chi_{m\tilde{m}}^{\lambda l \bar{l}}$ are equivalent classes of complex μ-measurable functions such that

$$\int_{S(\Delta^{sa})} \sum_{\mathcal{N}_\lambda} |\chi_{m\tilde{m}}^{\lambda l \bar{l}}|^2 \, d\mu(\lambda) < \infty$$

is a Hilbert space $\mathfrak{H}(S)$ of the $L^2(\mu)$ type. Here $\mathcal{N}_\lambda \subset \mathcal{N}$ is defined in Sec. 2 for particular cases.

Then the results of this section may be put in the form of Proposition 4.1.

Proposition 4.1: The mappings

$$F: \mathfrak{H}(X) \ni f \to \chi = Ff := \left\{ s.\lim \int_X \overline{Y_{m\tilde{m}}^{\lambda l \bar{l}}(\Omega)} f(\Omega)\, d\mu(\Omega);\ \lambda \in S(\Delta^{sa}),\ l, m, \bar{l}, \tilde{m} \in \mathcal{N}_\lambda \right\} \in \mathfrak{H}(S), \tag{4.12}$$

$$F^{-1}: \mathfrak{H}(S) \ni \chi \to f = F^{-1}\chi := \left\{ s.\lim \int_{S(\Delta^{sa}),\mathcal{N}_\lambda} \sum Y_{m\tilde{m}}^{\lambda l \bar{l}}(\Omega) \chi_{m\tilde{m}}^{\lambda l \bar{l}}\, d\mu(\lambda);\ \Omega \in X \right\} \in \mathfrak{H}(X) \tag{4.13}$$

are, one to the other, inverse isometric mappings: $F^{-1}F = I$ on $\mathfrak{H}(X)$ and $FF^{-1} = I$ on $\mathfrak{H}(S)$. The symbol $s.\lim$ in (4.12) means the strong convergence $f_n \to f$, $f_n \in \mathfrak{D}(X)$ and $s.\lim$ in (4.13) means the strong convergence $\chi_n \to \chi$, $\chi_n := \{\langle Y_{m\tilde{m}}^{\lambda l \bar{l}}, f_n \rangle;\ f_n \in \mathfrak{D}(X),\ \lambda \in S(\Delta^{sa}), l, m, \bar{l}, \tilde{m} \in \mathcal{N}_\lambda\}$. The operator Δ^{sa} is diagonally (or spectrally) represented in $\mathfrak{H}(S)$.

5. UNITARITY, IRREDUCIBILITY, AND MAUTNER'S DECOMPOSITION

The unitary representation $SO_0(p,q) \ni g \to U_g\chi := \{\chi_{m\tilde{m}}^{\lambda l \bar{l}}(U_g f);\ \lambda \in S(\Delta^{sa}),\ l,\ m,\ \bar{l},\ \tilde{m} \in \mathcal{N}_\lambda\} \in \mathfrak{H}(S)$ is reducible. [This representation and the quasi-regular representation (2.2) are unitarily equivalent with respect to the operator F of Proposition 4.1.] The carrier space of an irreducible representation is necessarily an eigenspace $\mathfrak{L}^{\lambda, ir}$ with an eigenvalue $\lambda \in S(\Delta^{sa})$. If the eigenspace $\mathfrak{L}^{\lambda, ir}$ corresponds to the proper eigenvalue, i.e., to $\lambda \in PS(\Delta^{sa})$, then $\mathfrak{L}^{\lambda, ir}$ is a proper subspace of the Hilbert space $\mathfrak{H}(S)$. (The image of these spaces under F^{-1} are the irreducible spaces considered in Ref. 1.) In the case of the representation of the noncompact group the invariant operator may have a continuous spectrum besides a discrete one as in our case. The "eigenspace" $\mathfrak{L}^{\lambda}$ corresponding to $\lambda \in CS(\Delta^{sa})$ is not a proper subspace of the Hilbert space $\mathfrak{H}(S)$ and one needs to extend such eigenspace to a certain Hilbert space $\mathfrak{K}^{\lambda, ir}$ in order to obtain a representation on the Hilbert space $\mathfrak{K}^{\lambda, ir}$ Then one expects that the Hilbert space $\mathfrak{H}(S)$ will be decomposed into the sum over $\mathfrak{H}^{\lambda}$, $\lambda \in PS(\Delta^{sa})$ and an integral of $\mathfrak{H}^{\lambda}$, $\lambda \in CS(\Delta^{sa})$. Such decomposition of the Hilbert space $\mathfrak{H}(S)$ into a so-called direct integral of Hilbert spaces was investigated by von Neumann.[17]

Let us try to construct the desired decomposition of the Hilbert space $\mathfrak{H}(S)$. The structure of the Hilbert space $\mathfrak{H}(S)$ leads us to consider a unitary space of l^2 type determined by vectors $\chi^{\lambda} := \{\chi_{m\tilde{m}}^{\lambda l \bar{l}}(f); f \in \mathfrak{H}(X), l, m, \bar{l}, \tilde{m} \in \mathcal{N}_\lambda\}$, where $\chi_{m\tilde{m}}^{\lambda l \bar{l}}(f) = \langle Y_{m\tilde{m}}^{\lambda l \bar{l}}, f \rangle$. As

$$\left(\bigcup_{\lambda \in S(\Delta^{sa})} (\lambda \times \mathcal{N}_\lambda) \right) \times \mathfrak{H}(X) \to \chi_{m\tilde{m}}^{\lambda l \bar{l}}(f)$$

is not a bounded mapping we restrict ourselves to the mapping

$$\left(\bigcup_{\lambda \in S(\Delta^{sa})} (\lambda \times \mathcal{N}_\lambda) \right) \times \mathfrak{D}(X) \to \chi_{m\tilde{m}}^{\lambda l \bar{l}}(f)$$

and justify this by the following two propositions.

Proposition 5.1: For every $l, m, \bar{l}, \tilde{m} \in \mathcal{N}_\lambda$, $f \in \mathfrak{D}(X)$ the function $\chi_{m\tilde{m}}^{\lambda l \bar{l}}(f)$ is analytic in the λ plane cut along the segment $(-\infty, [\frac{1}{2}(p+q-2)]^2]$ and tends to zero faster than any polynomial in λ along the positive real axis. Moreover, for every bounded subset $\mathfrak{B} \subset \bigcup_{\lambda \in S(\Delta^{sa})} (\lambda \times \mathcal{N}_\lambda)$ there exists a $f \in \mathfrak{D}(X)$ such that $\chi_{m\tilde{m}}^{\lambda l \bar{l}}(f) \not\equiv 0$ on $\mathfrak{B}$.

[16] A. C. Zaanen, *An Introduction to the Theory of Integration* (North-Holland Publishing Company, Amsterdam, 1961).

[17] J. von Neumann, Ann. Math. **50**, 401 (1949).

Proof: Let Σ be any bounded convex domain in the cut λ plane. Using expressions (4.1) and (4.3) and the definition of the mapping of Sec. 3 we easily find the uniform bound of the functions $Y^{\lambda l l'}_{m\tilde{m}}(\Omega)$ in the domain $\Sigma \times X$:

$$|Y^{\lambda l, l'}_{m\tilde{m}}(\Omega)| < A(\Sigma) \exp\{B(\Sigma)\theta\} |Y^{l}_{m}(\omega) Y^{l'}_{\tilde{m}}(\tilde{\omega})|, \quad (5.1)$$

where the constants $A(\Sigma)$ and $B(\Sigma)$ depend on the domain Σ. The analyticity of $\chi^{\lambda l l'}_{m\tilde{m}}(f)$ in the cut λ plane follows from the analyticity of the function $Y^{\lambda l l'}_{m\tilde{m}}(\Omega)$ in the cut λ plane for any $\Omega \in X$, the bound (5.1) and the bound

$$|f(\Omega)| < C \cosh^{\rho} \theta \exp(-2 \cosh^2 \theta),$$

$f \in \mathfrak{D}(X)$.[18] The asymptotic behavior of $\chi^{\lambda l l'}_{m\tilde{m}}(f)$ along the positive real axis follows from the equality

$$\lambda^n \chi^{\lambda l l'}_{m\tilde{m}}(f) = \int \overline{Y^{\lambda l l'}_{m\tilde{m}}(\Omega)} [(\Delta^{sa})^n f](\Omega)\, d\mu(\Omega),$$

the invariance of $\mathfrak{D}(X)$ under Δ^{sa} and the uniform bound $|Y^{\lambda l l'}_{m\tilde{m}}(\Omega)| < K \exp(M\theta)$ on $R \times X$, where $R := \{\lambda; [\tfrac{1}{2}(p+q-2)]^2 \le \lambda < \infty\}$. The second part of the proposition follows from its first part and the density of $\mathfrak{D}(X)$ in $\mathfrak{H}(X)$.

Using Proposition 5.1 the following is easily proved.

Proposition 5.2: The linear manifold $\mathfrak{D}^{\lambda}$ determined by vectors $\chi^{\lambda} := \{\chi^{\lambda l l'}_{m\tilde{m}}(f);\ l, m, l', \tilde{m} \in \mathcal{N}_{\lambda}, f \in \mathfrak{D}(X)\}$ is dense in the Hilbert space $\mathfrak{H}^{\lambda}$ of l^2 type, vectors of which are sequences $\chi^{\lambda} := \{\chi^{\lambda l l'}_{m\tilde{m}};\ l, m, l', \tilde{m} \in \mathcal{N}_{\lambda}\}$ for which $\|\chi^{\lambda}\|^2_{\lambda} = \sum_{\mathcal{N}_{\lambda}} |\chi^{\lambda l l'}_{m\tilde{m}}|^2 < \infty$.

Let $O := \{f_1, f_2, \cdots\}$ be an orthonormal complete sequence in the Hilbert space $\mathfrak{H}(X)$ which is obtained by the Smith orthogonalization process from a sequence $g_1, g_2, \cdots$ determining $\mathfrak{D}(X)$. In our case the denumerable set of vector fields $S(\Delta^{sa}) \times O \ni (\lambda, f_n) \to \chi^{\lambda}_n := \{\chi^{\lambda l l'}_{m\tilde{m}}(f_n);\ l,\ m,\ l',\ \tilde{m} \in \mathcal{N}_{\lambda},\ f_n \in O\}$ forms a so-called fundamental family of vector fields.[19,20] The vector field $S(\Delta^{sa}) \ni \lambda \to \psi^{\lambda} := \{\psi^{\lambda l l'}_{m\tilde{m}};\ l, m, l', \tilde{m} \in \mathcal{N}_{\lambda}\}$ is called μ-measurable if

$$(\chi^{\lambda}_n, \psi^{\lambda})_{\lambda} = \sum_{\mathcal{N}_{\lambda}} \overline{\chi^{\lambda l l'}_{m\tilde{m}}(f_n)} \psi^{\lambda l l'}_{m\tilde{m}}$$

is a μ-measurable function for all n. The direct integral

$$\hat{\mathfrak{H}} = \int_{S(\Delta^{sa})} \oplus \hat{\mathfrak{H}}^{\lambda}\, d\mu(\lambda) \quad (5.2)$$

of the Hilbert spaces $\hat{\mathfrak{H}}^{\lambda}$ is the Hilbert space $\hat{\mathfrak{H}}$ of equivalent classes of the μ-measurable vector fields χ^{λ} for which $\int_{S(\Delta^{sa})} \|\chi^{\lambda}\|^2_{\lambda}\, d\mu(\lambda) < \infty$ and where the scalar product is defined by

$$(\chi; \psi) = \int_{S(\Delta^{sa})} (\chi^{\lambda}, \psi^{\lambda})_{\lambda}\, d\mu(\lambda). \quad (5.3)$$

The Hilbert spaces $\mathfrak{H}(X)$ and $\mathfrak{H}(S)$ are isometrically isomorphic with respect to the Fourier transform F, and $\mathfrak{H}(S)$ is represented as the direct integral of Hilbert spaces $\hat{\mathfrak{H}}^{\lambda}$.

Unitarity: We obtained the Hilbert spaces $\hat{\mathfrak{H}}^{\lambda}$, closed subspaces $\hat{\mathfrak{H}}^{\lambda, ir}$ of which may be the carrier spaces of the unitary irreducible representations of the group $SO_0(p, q)$. We first prove the unitarity of the representations on $\hat{\mathfrak{H}}^{\lambda}$ for every $\lambda \in S(\Delta^{sa})$. (This induces the unitarity of the representation in any of $\hat{\mathfrak{H}}^{\lambda, ir}$.) For the proof of the unitarity we need the auxiliary:

Proposition 5.3: The properties stated in Proposition 5.1 for the function $\chi^{\lambda l l'}_{m\tilde{m}}(f), f \in \mathfrak{D}(X)$ are also valid for the function $\chi^{\lambda l l'}_{m\tilde{m}}(U_g f), f \in \mathfrak{D}(X)$.

Proof: From the definition of the function $\chi^{\lambda l l'}_{m\tilde{m}}(U_g f)$ and left invariance of the measure $d\mu(\Omega)$ it follows

$$\chi^{\lambda l l'}_{m\tilde{m}}(U_g f) = \int_X \overline{Y^{\lambda l l'}_{m\tilde{m}}(g\Omega)} f(\Omega)\, d\mu(\Omega). \quad (5.4)$$

Now, the proof of Proposition 5.1 may be applied as $Y^{\lambda l l'}_{m\tilde{m}}(g\Omega)$ has essentially the same bound as in the proof of Proposition 5.1 because the finite point Ω under the action of $g \in SO_0(p, q)$ is transformed again to a finite point $\Omega' = g\Omega \in X$.

Proposition 5.4: The representation

$$SO_0(p, q) \ni g \to U^{\lambda}_g \chi^{\lambda} := \{\chi^{\lambda l l'}_{m\tilde{m}}(U_g f),\ l, m, l', \tilde{m} \in \mathcal{N}_{\lambda}, f \in \mathfrak{D}(X)\} \in \mathfrak{D}^{\lambda} \quad (5.5)$$

is norm preserving in $\mathfrak{D}^{\lambda}$ and has unique unitary extension V^{λ}_g *in* $\mathfrak{H}^{\lambda} = \tilde{\mathfrak{D}}^{\lambda}$ for every $\lambda \in S(\Delta^{sa})$.

Proof: The unitarity of the representations of the group $SO_0(p, q)$ on $\mathfrak{H}^{\lambda}$, $\lambda \in PS(\Delta^{sa})$ was considered in Ref. 1. (In fact, there we worked with the spaces $F^{-1}\mathfrak{H}^{\lambda}$). Now we give the general proof valid also for $\lambda \in CS(\Delta^{sa})$.

Because of the strong commutativity of U_g and Δ^{sa}, for any μ-measurable function $F(\lambda)$, uniformly bounded on $S(\Delta^{sa})$, we have

$$(U_g f, F(\Delta^{sa}) U_g h) - (f, F(\Delta^{sa}) h) = 0, \quad f, h \in \mathfrak{D}(X). \quad (5.6)$$

According to Proposition 4.1 we can rewrite (5.6) in the form

$$\int_{S(\Delta^{sa})} |F(\lambda)|^2 \{(U^{\lambda}_g \chi^{\lambda}, U^{\lambda}_g \psi^{\lambda})_{\lambda} - (\chi^{\lambda}, \psi^{\lambda})_{\lambda}\}\, d\mu(\lambda) = 0, \quad \chi^{\lambda}, \psi^{\lambda} \in \mathfrak{D}^{\lambda}. \quad (5.7)$$

Because of the continuity of $(U^{\lambda}_g \chi^{\lambda}, U^{\lambda}_g \psi^{\lambda})_{\lambda} - (\chi^{\lambda}, \psi^{\lambda})$

[18] M. A. Lavrentiev and B. V. Shabat, "Methods of Theory of Functions of Complex Variables," "Nauka," Moscow (1965) (in Russian).

[19] K. Maurin, *The Methods of Hilbert Space* (Mir, Moscow, 1965) (in Russian).

[20] K. Maurin, "General Eigenfunction Expansions and Group Representations" (Lecture Notes), ICTP, Trieste, preprint IC/66/12.

on $CS(\Delta^{sa})$ (Propositions 5.1 and 5.3) and boundedness on $PS(\Delta^{sa})$ and because of the arbitrariness of the uniformly bounded function $F(\lambda)$ on $S(\Delta^{sa})$ we conclude $(U_g^\lambda \chi^\lambda, U_g^\lambda \psi^\lambda)_\lambda = (\chi^\lambda, \psi^\lambda)_\lambda$, $\lambda \in S(\Delta^{sa})$, $\chi^\lambda, \psi^\lambda \in \mathfrak{D}^\lambda$. As $\mathfrak{D}^\lambda$ is the dense linear manifold in $\mathfrak{H}^\lambda$, U_g^λ can be uniquely extended to the unitary operator V_g^λ in $\mathfrak{H}^\lambda = \tilde{\mathfrak{D}}^\lambda$. [The Hilbert spaces $\mathfrak{H}^\lambda$ are representatives of $\hat{\mathfrak{H}}^\lambda$ and V_g^λ is the representative of $\hat{U}_g^\lambda$ of the decomposable operator $U_g = \int_{S(\Delta^{sa})} \hat{U}_g^\lambda \, d\mu(\lambda)$ in $\hat{\mathfrak{H}} = \int_{S(\Delta^{sa})} \hat{\mathfrak{H}}^\lambda \, d\mu(\lambda)$.]

The following proposition is not directly connected with our considerations but it shows a deeper property of the linear manifold $\mathfrak{D}^\lambda$.

Proposition 5.5: The linear manifold $\mathfrak{D}^\lambda$ consists of analytic vectors. The operators X_{ij}^λ are essentially skew-adjoint on $\mathfrak{D}^\lambda$.

Proof: The operators X_{ij}^λ are skew-symmetric operators on $\mathfrak{D}^\lambda$ as $V_{g_{ij}(t)}$ is unitary in $\mathfrak{H}^\lambda$ and $\|[t^{-1}(V_{g_{ij}(t)} - I) - X_{ij}^\lambda]\chi^\lambda\|_\lambda \to 0$ for $\chi^\lambda \in \mathfrak{D}^\lambda$ as $t \to 0$. Let us first show that any vector $\chi^\lambda \in \mathfrak{D}^\lambda$ is an analytic vector for the operators L_{ij}^λ (L_{ij}^λ is the representation of the generator l_{ij} of the one-parameter compact subgroup[1]). First, one proves that $-\sum_{i<j} L_{ij}^2$ is an invariant operator using the commutation rules. Hence it is proportional to $-\Delta(S^{p-1}) - \Delta(S^{q-1})$ as the representations are the most degenerate ones. We do not lose on generality, supposing that the constant of the proportionality is equal to one. For any $\chi^\lambda \in \mathfrak{D}^\lambda$ the decomposition

$$\chi^\lambda = \sum \chi^\lambda_{l_{\{p/2\}}, l_{\{q/2\}}}, \quad \chi^\lambda_{l_{\{p/2\}} l_{\{q/2\}}} \in \mathfrak{H}^\lambda_{l_{\{p/2\}}, l_{\{q/2\}}},$$

where $\mathfrak{H}^\lambda_{l_{\{p/2\}}, l_{\{q/2\}}}$ are finite-dimensional vector spaces determined by vectors $\chi^{\lambda l \bar{l}}_{m\bar{m}}(f)$, $f \in \mathfrak{D}(X)$ with fixed $l_{\{p/2\}}, \bar{l}_{\{q/2\}} \in \mathcal{N}_\lambda$, has a finite number of nonvanishing terms. We easily find the bound

$$\|L_{ij}^\lambda \chi^\lambda\|_\lambda^2 \le [l_{\max}(l_{\max} + p - 2) + \bar{l}_{\max}(\bar{l}_{\max} + q - 2)] \, \|\chi^\lambda\|_\lambda^2$$

from the expression

$$-\sum (L_{ij}^\lambda)^2 \chi^\lambda = \sum_{l_{\{p/2\}}, \bar{l}_{\{q/2\}}} [\bar{l}_{\{q/2\}}(\bar{l}_{\{q/2\}} + q - 2) + l_{\{p/2\}}(l_{\{q/2\}} + p - 2)] \chi^\lambda_{l_{\{p/2\}}, \bar{l}_{\{q/2\}}}.$$

Now, using the irreducibility of the algebra determined by L_{ij} in the subspaces $\mathfrak{H}^\lambda_{l_{\{p/2\}}, \bar{l}_{\{q/2\}}}$ we find the estimate

$$\left\| \prod_{\alpha=1}^n L^\lambda_{i_\alpha j_\alpha} \chi^\lambda \right\|_\lambda \le [l_{\max}(l_{\max} + p - 2) + \bar{l}_{\max}(\bar{l}_{\max} + q - 2)]^{\frac{1}{2}n} \|\chi^\lambda\|_\lambda, \quad \chi^\lambda \in \mathfrak{D}^\lambda. \tag{5.8}$$

From the inequality (5.8) we easily see that $\chi^\lambda \in \mathfrak{D}^\lambda$ is an analytic vector of L_{ij}^λ and consequently L_{ij}^λ is essentially skew-adjoint on $\mathfrak{D}^\lambda$.[21]

The estimate for $\|\prod_{\alpha=1}^n B^\lambda_{i_\alpha j_\alpha} \chi^\lambda\|_\lambda$, $\chi^\lambda \in \mathfrak{D}^\lambda$ can be obtained in an analogous way starting with the invariant operator

$$\sum_{i<j} (B_{ij}^\lambda)^2 - \sum_{i<j} (L_{ij}^\lambda)^2$$

on $\mathfrak{D}^\lambda$:

$$\left\| \prod_{\alpha=1}^n B^\lambda_{i_\alpha j_\alpha} \chi^\lambda \right\|_\lambda \le [\lambda + (l_{\max} + n + 1)(l_{\max} + p + n + 1) + (\bar{l}_{\max} + n + 1)(\bar{l}_{\max} + p + n + 1)]^{\frac{1}{2}n} \|\chi^\lambda\|_\lambda, \quad \chi^\lambda \in \mathfrak{D}^\lambda. \tag{5.9}$$

It follows again that $\chi^\lambda \in \mathfrak{D}^\lambda$ is an analytic vector for B_{ij}^λ and that B_{ij}^λ is essentially skew-adjoint on $\mathfrak{D}^\lambda$.

Irreducibility: Let us try to find the invariant subspaces $\mathfrak{H}^{\lambda, ir}$ of the space $\mathfrak{H}^\lambda$ with respect to the representation of the group (5.5).

In our previous work[1,2] we have considered the representations of the Lie algebra $\mathfrak{R}(p, q)$ of the group $SO_0(p, q)$ on the Lie algebra of the unbounded operators X_{ij}^λ in the Hilbert space $\mathfrak{H}^\lambda$, $\lambda \in S(\Delta^{sa})$. The irreducibility with respect to the algebra of the operators X_{ij}^λ has been defined in the following way: The representation

$$\mathfrak{R}(p, q) \ni X_{ij} \to X_{ij}^\lambda \chi^\lambda := \{\chi^{\lambda l \bar{l}}_{m\bar{m}}(X_{ij} f); \, l, m, \bar{l}, \bar{m} \in \mathcal{N}_\lambda^{ir}, \, f \in \mathfrak{D}(X)\}$$

is called irreducible on a common invariant domain $\mathfrak{D}^{\lambda, ir}$ if for any two $\chi^\lambda, \psi^\lambda \in \mathfrak{D}^\lambda$ an operator A^λ exists in the enveloping algebra such that $(\chi^\lambda, A^\lambda \psi^\lambda)_\lambda \neq 0$. The irreducible representations of the Lie algebra were found in Refs. 1 and 2 and they are reviewed in Sec. 2 of the present work. It was established there that the common invariant domain $\mathfrak{D}^{\lambda, ir}$ has the following structure:

$$\mathfrak{D}^{\lambda, ir} = \sum_{l_{\{p/2\}}, \bar{l}_{\{q/2\}} \in \mathcal{N}_\lambda^{ir}} \oplus \, \mathfrak{H}^\lambda_{l_{\{p/2\}}, \bar{l}_{\{q/2\}}},$$

where $\mathfrak{H}^\lambda_{l_{\{p/2\}}, \bar{l}_{\{q/2\}}}$ are finite-dimensional vector spaces determined by vectors $\chi^{\lambda l \bar{l}}_{m\bar{m}}(f)$, $f \in \mathfrak{D}(X)$ with fixed $l_{\{p/2\}}, \bar{l}_{\{q/2\}} \in \mathcal{N}_\lambda^{ir}$ and $\mathcal{N}_\lambda^{ir}$ is a certain subset of the set $\mathcal{N}_\lambda$.

Proposition 5.6: The representation (5.5) is irreducible on the closed subspace $\mathfrak{H}^{\lambda, ir} = \tilde{\mathfrak{D}}^{\lambda, ir}$ of the space $\mathfrak{H}^\lambda$.

Proof: Every finite-dimensional subspace $\mathfrak{H}^\lambda_{l_{\{p/2\}}, \bar{l}_{\{q/2\}}}$ is irreducible with respect to the representation L_{ij} of the Lie algebra l_{ij} of the maximal compact subgroup $SO(p) \times SO(q)$.[1] This fact and the continuity of the operators L_{ij} in the Hilbert spaces $\mathfrak{H}^\lambda_{l_{\{p/2\}}, \bar{l}_{\{q/2\}}}$ induces the irreducibility of the representation $SO(p) \times SO(q) \ni g \to U_g^\lambda \chi^\lambda \in \mathfrak{H}^\lambda_{l_{\{p/2\}}, \bar{l}_{\{q/2\}}}$, as, for instance, $A^\lambda U_g^\lambda - U_g^\lambda A^\lambda = 0$ in $\mathfrak{H}^\lambda_{l_{\{p/2\}}, \bar{l}_{\{q/2\}}}$ induces $AL_{ij} -$

[21] E. Nelson, Ann. Math. 70, 572 (1959).

$L_{ij}A = 0$ in $\mathfrak{H}^{\lambda}_{l_{\{p/2\}},l_{\{q/2\}}}$. Hence any invariant subspace $\mathfrak{H}^{\lambda}_{k}$, $k = 1, 2, \cdots$ of the space $\mathfrak{H}^{\lambda,ir}$ with respect to the representation of the group $SO_0(p, q)$ must have the form $\mathfrak{H}^{\lambda}_{k} = \sum_{l_{\{p/2\}},l_{\{q/2\}}\in\mathcal{N}_k} \oplus \mathfrak{H}^{\lambda}_{l_{\{p/2\}},l_{\{q/2\}}}$, where $\mathcal{N}_k$ is an infinite subset of the set $\mathcal{N}^{ir}_{\lambda}$. Then we can introduce the operator K^{λ}, eigenspaces of which are $\mathfrak{H}^{\lambda}_{k}$, with the eigenvalue $1/K$. From $K^{\lambda}U^{\lambda}_{g} - U^{\lambda}_{g}K^{\lambda} = 0$ in $\mathfrak{H}^{\lambda,ir} = \Sigma \oplus \mathfrak{H}^{\lambda}_{k}$ and the fact that $\mathfrak{D}^{\lambda,ir}$ is the common invariant domain of K^{λ}, L^{λ}_{ij}, B^{λ}_{ij} we conclude $K^{\lambda}B^{\lambda}_{ij} - B^{\lambda}_{ij}K^{\lambda} = 0$ in $\mathfrak{D}^{\lambda,ir}$ in contradiction to the results of Refs. 1 and 2.

Mautner's Decomposition: We have determined the unitary irreducible representations of the group $SO_0(p, q)$ related to three homogeneous spaces H^{p}_{q}, H^{q}_{p}, and C^{p}_{q}. Let us now discuss the consequences which follow from the comparison of our results with the general theory. Let $\mathfrak{R}_X$ be the smallest weakly closed * algebra generated by the set of elements of the group ring associated to the representation U_g and let $\mathfrak{R}'_X$ be the commutant of $\mathfrak{R}_X$. We know that the center $\mathfrak{Z}_X = \mathfrak{R}_X \cap \mathfrak{R}'_X$ is generated by the operators $F(\Delta^{sa})$, where $F(\lambda) = (\lambda + \frac{1}{8})^{-1}$. [The operator Δ^{sa} is unbounded so that we have to consider its function $F(\Delta^{sa})$ which is a bounded operator in $\mathfrak{H}(X)$.] The decomposition of the Hilbert space $\mathfrak{H}(S)$ into the direct integral (5.2) is so-called central decomposition. We found that the Hilbert spaces $\hat{\mathfrak{H}}^{\lambda}$ in (5.2) are generally reducible so that the central decomposition does not lead generally to decomposition into the irreducible representations. We have to add a certain number of discrete operators in order to classify the irreducible representations. (Let us recall that the ring of the invariant operators in the enveloping algebra is generated by Δ^{sa}.) According to Mautner's theorem[22] this means that the maximal commuting subalgebra $\mathfrak{M}_X$ of the commutant $\mathfrak{R}_X$ is generally generated by more operators than $F(\Delta^{sa})$. In particular cases we have the following situations:

(a) For $SO(p, q)$, $p \geq q > 2$: $\mathfrak{M}_{H_q{}^p} = \{F(\Delta^{sa}), P\}$, $\mathfrak{M}_{H_p{}^q} = \{F(\Delta^{sa}), P\}$, and $\mathfrak{M}_{C_q{}^p} = \{F(\Delta^{sa}), P\}$, where the parity operator P is defined by (5.4).[2]

(b) For $SO(p, q)$, $p \geq q = 2$: $\mathfrak{M}_{H_2{}^p} = \{F(\Delta^{sa}), P, T\}$, where the operator T is defined in point (v) of Sec. 2D, $\mathfrak{M}_{H_2{}^p} = \{F(\Delta^{sa}), P\}$, $\mathfrak{M}_{C_2{}^p} = \{F(\Delta^{sa}), P\}$.

(c) For $SO_0(p, 1)$, $p > 1$: $\mathfrak{M}_{H_1{}^p} = \{F(\Delta^{sa}), P\}$, $\mathfrak{M}_{H_p{}^1} = \{F(\Delta^{sa})\}$, $\mathfrak{M}_{C_1{}^p} = \{F(\Delta^{sa})\}$.

As the spectra of the operators P and T have only a finite number of points, it is easy to find the Mautner decomposition of the Hilbert space $\mathfrak{H}(S)$ with respect to the maximal commutative algebra $\mathfrak{M}_X$. The decomposition for any particular case is written in Sec. 2.

[22] F. I. Mautner, Ann. Math. 51, 1 (1950).

In Ref. 1 we have given the maximal set of commuting operators A (7.8) for every irreducible representation. These operators are unbounded and we have to replace them by the bounded operators $F(A)$, $F(\lambda) = (\lambda + \frac{1}{8})^{-1}$. The operators A of the set (7.8)[1] are in the enveloping algebra and according to the preceding discussion we have in general to complete the operators $F(A)$ with a certain number of discrete operators in order to generate the maximal commutative algebra on $\mathfrak{H}(X)$.

6. NUCLEAR SPECTRAL THEORY

The existence of the complete set of generalized eigenfunctions for a self-adjoint operator on a separable Hilbert space $\mathfrak{H}$ was first established by Gel'fand and Kostiuchenko.[7] They introduced a nuclear space ϕ, such that the Hilbert space $\mathfrak{H}$ is the completion of ϕ with respect to one of the norms in ϕ and the dual ϕ' of continuous functionals and they considered the triplet $\phi \subset \mathfrak{H} \subset \phi'$ which we call the Gel'fand–Kostiuchenko[20] triplet. Their theory can be put in a very elegant form using Maurin's construction of the nuclear space ϕ for an arbitrary, strongly commutative denumerable set of normal operators A_{β}.[8,19] Let us state now the nuclear spectral theorem[19,20,23] and then let us show how the results derived in the present work may be a nice illustration of this general theory.

Nuclear spectral theorem:

(1) *Such a nuclear space ϕ exists that $\phi \subset \mathfrak{H} \subset \phi'$ is a Gel'fand–Kostiuchenko triplet.*

(2) *Each A_{β} maps ϕ continuously into ϕ.*

(3) *There exist a direct integral $\hat{\mathfrak{H}} = \int_D \hat{\mathfrak{H}}^{\lambda}\, d\sigma(\lambda)$ and such a Hilbert isomorphism $F: \mathfrak{H} \to \hat{\mathfrak{H}}$ that for σ—almost all λ, $\hat{\mathfrak{H}}^{\lambda} \subset \phi'$, and $\hat{\mathfrak{H}}^{\lambda}$ are common generalized eigenspaces of all A_{β}: $\langle e^{\lambda}, A_{\beta}\varphi\rangle = \hat{A}^{\lambda}_{\beta}\langle e^{\lambda}, \varphi\rangle$ identically in $\varphi \in \phi$, where $\{\hat{A}^{\lambda}_{\beta}, \lambda \in D\}$ is the spectrum of A_{β}. Taking in each $\hat{\mathfrak{H}}^{\lambda}$ the orthonormal basis e^{λ}_{n}, $n = 1, 2, \cdots, \dim \hat{\mathfrak{H}}^{\lambda}$ we obtain the generalized Fourier–Plancherel equation*

$$(\varphi, \psi) = \int_D \sum_{n=1}^{\dim \hat{\mathfrak{H}}^{\lambda}} \overline{\langle e^{\lambda}_{n}, \varphi\rangle}\langle e^{\lambda}_{n}, \psi\rangle\, d\sigma(\lambda), \quad \varphi, \psi \in \phi.$$

(4) *The isomorphism F is defined by $\mathfrak{H} \supset \phi \ni \varphi \to F\varphi := \{(F\varphi)(\lambda), \lambda \in D\} = \{\hat{\varphi}^{\lambda}, \lambda \in D\} \in \hat{\mathfrak{H}}^{\lambda}$, where*

$$\hat{\varphi}^{\lambda} = \sum_{n=1}^{\dim \hat{\mathfrak{H}}^{\lambda}} \overline{\langle e^{\lambda}_{n}, \varphi\rangle} e^{\lambda}_{n}.$$

(5) *The spectral synthesis of $\varphi \in \phi$ is given by $\varphi = \int_D \hat{\varphi}^{\lambda}\, d\sigma(\lambda)$.*

The nuclear space ϕ of the spectral theorem can be obtained using Maurin's construction.[8,19,23] According

[23] K. Maurin, manuscript of the forthcoming book: *Unitary Representations of Non-Compact Groups and Associated Eigenfunction Expansions.*

to this construction as a nuclear space ϕ, for the operator Δ^{sa} (or the maximal commutative algebra $\mathfrak{M}_X$ of $\mathfrak{R}_X$) may be taken to be the linear manifold $\mathfrak{D}(X)$, which has been exploited so far very extensively, furnished, of course, with certain nuclear topology (for details see Ref. 19). The dual space $\mathfrak{D}'$ is the space of the continuous functionals

$$\chi^\lambda: \mathfrak{D}(X) \ni f \to \chi^\lambda := \{\chi^{\lambda m l}_{\tilde{l}\tilde{m}}(f);\ l, m, \tilde{l}, \tilde{m} \in \mathcal{N}_\lambda\} \in \hat{\mathfrak{H}}^\lambda. \quad (6.1)$$

It was proved in Sec. 4 that the set of functionals (6.1) constitute a complete set of eigenfunctionals. The direct integral, the Fourier transform F and the spectral synthesis were explicitly written in Sec. 4 if only the operator Δ^{sa} is considered. It is easy to generalize the expressions of Sec. 4 for the maximal commutative algebra.

ACKNOWLEDGMENTS

The authors would like to thank Professor A. Salam and Professor P. Budini, and the IAEA for hospitality kindly extended to them at the International Centre for Theoretical Physics, Trieste.

The authors are grateful to Professor K. Maurin not only for much valuable insight and many private communications but also for placing at their disposal the manuscript of his forthcoming book. They are grateful to Dr. J. Fischer for some remarks and Dr. S. Woronowicz for interesting discussions.

APPENDIX I

We prove in this appendix the density of $\mathfrak{D}(X)$ in $\mathfrak{H}(X)$. Let us suppose that some $g \in \mathfrak{H}(X)$, $g \neq 0$, exists, such that $(f, g)_X = 0$ for every $f \in \mathfrak{D}(X)$. We introduce the local coordinates, y_g, $k = 1, 2, \cdots, p+q$, in the Minkowski space $M^{p,q}$ round the hyperboloids H^p_q and H^q_p by $y^k = \rho x^k$, $\rho > 0$, where x^k are biharmonic coordinates[1] and round the cone C^p_q by

$$y^k = (r+\rho)x^k,\ k = 1, 2, \cdots, p, \quad y^k = (r-\rho)x^k,\ k = p+1, p+2, \cdots, p+q, \quad \rho > 0,$$

r real. For the hyperboloid we define the function $h(\rho, \theta) = \exp(2\rho^2 \cosh^2\theta - 2\cosh^2\theta)$ for $\rho \in (\frac{1}{3}, \frac{1}{2})$ and zero otherwise. The function $v(y^1, y^2, \cdots, y^{p+q}) = (hg)(\rho, \Omega)$ has the property

$$0 < \|v\|^2_{M^{p,q}} = \int_{M^{p,q}} |(hg)(\rho, \Omega)|^2\, d\nu(\rho)\, d\mu(\Omega) < \infty,$$

where $d\nu(\rho) = dy^1 \cdots dy^{p+q}/d\mu(\Omega)$. As the set of functions

$$u(y^1, y^2, \cdots, y^{p+q}) = P(y^1, y^2, \cdots, y^{p+q}) \times \exp\left\{-\sum_{k=1}^{p+q}(y^k)^2\right\},$$

where $P(y^1, y^2, \cdots, y^{p+q})$ is a polynomial, determine a dense linear manifold in $\mathfrak{L}^2(M^{p,q})$, such a function $u(y^1, y^2, \cdots, y^{p+q})$ exists that $(u, v)_{M^{p,q}} \neq 0$. But this contradicts our assumption as

$$(u, v)_{M^{p,q}} = \int d\nu(\rho) \int d\mu(\Omega) h(\rho, \theta) \overline{P(\rho, \Omega)} \times \exp(-2\rho^2 \cosh^2\theta) g(\Omega) = 0.$$

An analogous proof holds for the cone.

The proof of the invariance of $\mathfrak{D}(X)$ with respect to X_{ij}. The invariance follows from the equality

$$(L_{ij}f)(\Omega) = \left(y^i \frac{\partial}{\partial y^j} - y^j \frac{\partial}{\partial y^i}\right) f(y^1, y^2, \cdots, y^{p+q})\Big|_{\rho=1}, \quad f \in \mathfrak{D}(X).$$

An analogous equality holds for B_{ij}. Similarly, the skew-symmetry of the operator X_{ij} can be proved and consequently the symmetry of $\Delta(H^p_q)$ on $\mathfrak{D}(X)$.

APPENDIX II. REVIEW OF MOST DEGENERATE REPRESENTATIONS OF $SO(p)$ GROUPS

(i) The Laplace–Beltrami operator $\Delta(S^{p-1})$ defined in (A6),[1] (A7)[1] has only a discrete spectrum $S(\Delta(S^{p-1})): -l_{\{p/2\}}(l_{\{p/2\}} + p - 2)$, $l_{\{p/2\}} = 0, 1, 2, \cdots$, $l_1 \equiv m_1$.

(ii) The corresponding eigenfunctions in our biharmonic system are[1,24]

$$Y^{l_2, \cdots, l_{\{p/2\}}}_{m_2, \cdots, m_{[p/2]}}(\omega) = \begin{cases} \dfrac{1}{(N_r)^{\frac{1}{2}}} \displaystyle\prod_{k=2}^{r} \sin^{2-k}\vartheta^k\, d^{J_k}_{M_k, M_{k'}}(2\vartheta^k) \times \prod_{k=1}^{r} \exp(im_k\varphi^k), & p = 2r, \\ \dfrac{1}{(N_{r+1})^{\frac{1}{2}}} \sin^{1-r}\vartheta^{r+1}\, d^{J_{r+1}}_{M_{r+1}, 0}(\vartheta^{r+1}) \times \displaystyle\prod_{k=2}^{r} \sin^{2-k}\vartheta^k\, d^{J_k}_{M_k, M_{k'}}(2\vartheta^k) \prod_{k=1}^{r} \exp(im_k\varphi^k), & p = 2r+1, \end{cases} \quad (A1)$$

where ω is the set of variables $\vartheta^2, \cdots, \vartheta^{\{p/2\}}$, $\varphi^1, \cdots, \varphi^{[p/2]}$: $\vartheta^k \in [0, \frac{1}{2}\pi)$ $k = 2, \cdots, r$, $\varphi^l \in [0, 2\pi)$, $l = 1, \cdots, r$, $\vartheta^{r+1} \in [0, \pi]$, and the indices in (A1) are:

$$J_k = \tfrac{1}{2}(l_k + k - 2), \quad M_k = \tfrac{1}{2}(m_k + l_{k-1} + k - 2), \quad M'_k = \tfrac{1}{2}(m_k - l_{k-1} - k + 2) \quad (A2)$$

for $k = 2, 3, \cdots$;

$$J_{r+1} = l_{r+1} + r - 1,\ M_{r+1} = l_r + r - 1, \quad (A3)$$

l_k are nonnegative integers, m_k are integers. Here $l_2, \cdots, l_{\{p/2\}}, |m_1|, \cdots, |m_{[p/2]}|$ and $l_1 \equiv m_1$ are restricted by the conditions

$$|m_2| + |m_1| = l_2 - 2n_2, |m_3| + l_2 = l_3 - 2n_3, \cdots, |m_r| + l_{r-1} = l_r - 2n_r, \quad (A4)$$

$$l_r = l_{r+1} - n_{r+1} \quad \begin{matrix} n_k = 0, 1, \cdots, \{\frac{1}{2}l_k\}, \\ k = 2, 3, \cdots, r, \\ n_{r+1} = 0, \cdots, l_{r+1}. \end{matrix} \quad (A5)$$

[24] We use d function as defined in: M. E. Rose, *Elementary Theory of Angular Momentum* (John Wiley & Sons, Inc., New York, 1961), p. 53.

The normalization factors in (A1) are

$$N_r = 2\pi^r \prod_{k=2}^{r} \frac{1}{l_k + k - 1},$$

$$N_{r+1} = 4\pi^r \frac{1}{2(l_{r+1} + r) - 1} \prod_{k=2}^{r} \frac{1}{l_k + k - 1}. \quad \text{(A6)}$$

(iii) The completeness relations[25]:

$$\int_{S^{p-1}} \overline{Y^{l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}}(\omega)} Y^{l_2',\cdots,l_{\{p/2\}}'}_{m_1',\cdots,m_{[p/2]}'}(\omega)\, d\mu(\omega) = \prod_{k=2}^{\{p/2\}} \delta_{l_k l_k'} \prod_{k=1}^{[p/2]} \delta_{m_k m_k'}, \quad \text{(A7)}$$

$$\sum_{\mathcal{N}} \overline{Y^{l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}}(\omega)} Y^{l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}}(\omega') = \delta(\omega - \omega'; \mu),^{12} \quad \text{(A8)}$$

where $\mathcal{N}$ is the set of values of indices $l_2, \cdots, m_{[p/2]}$ restricted by (A4) and (A5), and $d\mu(\omega)$ is the left invariant measure on S^{p-1} defined by

$$d\mu(\omega_{p-1}) = \begin{cases} \prod_{k=2}^{r} \cos\vartheta^k \sin^{2k-3}\vartheta^k\, d\vartheta^k \prod_{k=1}^{r} d\varphi^k, & p = 2r, \\ \sin^{2r-1}\vartheta^{r+1}\, d\vartheta^{r+1} \prod_{k=2}^{r} \cos\vartheta^k \times \sin^{2k-3}\vartheta^k\, d\vartheta^k \prod_{k=1}^{r} d\varphi^k, & p = 2r+1, \end{cases} \quad \text{(A9)}$$

(iv) There is no need to go to the space of Fourier transforms to define the unitary irreducible representations. Therefore, we consider the space $\mathfrak{H}(S^{p-1})$. The unitary irreducible representations are classified by the number $l_{\{p/2\}}$, $l_{\{p/2\}} = 0, 1, 2, \cdots$ and they are defined by

$$SO(p) \ni g \to (U_g Y^{l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}})(\omega) = Y^{l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}}(g^{-1}\omega) \in \mathfrak{H}^{l_{\{p/2\}}}(S^{p-1}),$$

where $\mathfrak{H}^{l_{\{p/2\}}}(S^{p-1})$ is the subspace of the space $\mathfrak{H}(S^{p-1})$ determined by the vectors $Y^{l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}}(\omega)$ with fixed $l_{\{p/2\}}$.

(v) The decomposition of the quasi-regular representation (2.2) into the irreducible ones is trivial as

$$\mathfrak{H}(S^{p-1}) = \sum_{l_{\{p/2\}}=0}^{\infty} \oplus \mathfrak{H}^{l_{\{p/2\}}}(S^{p-1}).$$

[25] The proof of the completeness of the set of harmonic functions $Y^{l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}}(\omega)$ in $\mathfrak{H}(S^{p-1})$ in another coordinate system is given in Ref. 3.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 5 MAY 1967

Recursive Evaluation of Some Atomic Integrals

R. A. SACK,* C. C. J. ROOTHAAN, AND W. KOŁOS†
Laboratory of Molecular Structure and Spectra, Department of Physics, University of Chicago, Chicago, Illinois

(Received 26 September 1966)

A simple recursion scheme is set up for the higher derivatives of functions of the form $f(\alpha, \beta)(\alpha + \beta)^{-1}$ and $f(\alpha, \beta, \gamma)(\alpha + \beta + \gamma)^{-1}$ and applied to the rapid calculation of the Calais–Löwdin two-particle atomic integrals.

IN various branches of theoretical physics integrals occur which can be expressed in the general form

$$\Gamma_{lm}(\alpha, \beta) = \left(-\frac{\partial}{\partial\alpha}\right)^l \left(-\frac{\partial}{\partial\beta}\right)^m \frac{f(\alpha, \beta)}{\alpha + \beta}. \quad (1)$$

Application of Leibnitz' theorem to the differentiation with respect to both variables yields

$$\Gamma_{lm} = \sum \frac{l!\, m!\, (\lambda' + \mu')!}{\lambda!\, \lambda'!\, \mu!\, \mu'!\, (\alpha + \beta)^{\lambda'+\mu'+1}} \times \left(-\frac{\partial}{\partial\alpha}\right)^{\lambda} \left(-\frac{\partial}{\partial\beta}\right)^{\mu} f(\alpha, \beta), \quad \lambda + \lambda' = l, \quad \mu + \mu' = m. \quad (2)$$

The recurrence relation for the binomial coefficients

$$\frac{(\lambda' + \mu')!}{\lambda'!\, \mu'!} = \binom{\lambda' + \mu'}{\lambda'} = \binom{\lambda' + \mu' - 1}{\lambda'} + \binom{\lambda' + \mu' - 1}{\lambda' - 1} \quad (3)$$

can be applied to each term in the sum (2) except the term with $\lambda = \mu = 0$, with the result

$$\Gamma_{lm} = \frac{1}{\alpha + \beta}\left[l\Gamma_{l-1,m,n} + m\Gamma_{l,m-1,n} + \left(-\frac{\partial}{\partial\alpha}\right)^l \left(-\frac{\partial}{\partial\beta}\right)^m f(\alpha, \beta)\right]. \quad (4)$$

This recurrence formula can be used to particular advantage in the evaluation of atomic two-electron integrals for the case of exponential correlation

$$\Gamma_{lmn}(\alpha, \beta, \gamma) = \iint \exp(-\alpha r_1 - \beta r_2 - \gamma r_{12}) r_1^{l-1} r_2^{m-1} r_{12}^{n-1}\, d^3\mathbf{r}_1\, d^3\mathbf{r}_2, \quad (5)$$

where the integration is to be taken over all possible positions $\mathbf{r}_1$ and $\mathbf{r}_2$ of the two electrons and r_{12} denotes their separation. These integrals were discussed by

* Permanent address: Department of Mathematics, University of Salford, England.

† Permanent address: Department of Theoretical Chemistry, University of Warsaw, ul. Pasteura 1, Warsaw 22, Poland.

The normalization factors in (A1) are

$$N_r = 2\pi^r \prod_{k=2}^{r} \frac{1}{l_k + k - 1},$$

$$N_{r+1} = 4\pi^r \frac{1}{2(l_{r+1} + r) - 1} \prod_{k=2}^{r} \frac{1}{l_k + k - 1}. \quad \text{(A6)}$$

(iii) The completeness relations[25]:

$$\int_{S^{p-1}} \overline{Y^{l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}}(\omega)} Y^{l_2',\cdots,l_{\{p/2\}}'}_{m_1',\cdots,m_{[p/2]}'}(\omega)\, d\mu(\omega) = \prod_{k=2}^{\{p/2\}} \delta_{l_k l_k'} \prod_{k=1}^{[p/2]} \delta_{m_k m_k'}, \quad \text{(A7)}$$

$$\sum_{\mathcal{N}} \overline{Y^{l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}}(\omega)} Y^{l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}}(\omega') = \delta(\omega - \omega'; \mu),^{12} \quad \text{(A8)}$$

where $\mathcal{N}$ is the set of values of indices $l_2, \cdots, m_{[p/2]}$ restricted by (A4) and (A5), and $d\mu(\omega)$ is the left invariant measure on S^{p-1} defined by

$$d\mu(\omega_{p-1}) = \begin{cases} \prod_{k=2}^{r} \cos\vartheta^k \sin^{2k-3}\vartheta^k\, d\vartheta^k \prod_{k=1}^{r} d\varphi^k, & p = 2r, \\ \sin^{2r-1}\vartheta^{r+1}\, d\vartheta^{r+1} \prod_{k=2}^{r} \cos\vartheta^k \times \sin^{2k-3}\vartheta^k\, d\vartheta^k \prod_{k=1}^{r} d\varphi^k, & p = 2r+1, \end{cases} \quad \text{(A9)}$$

(iv) There is no need to go to the space of Fourier transforms to define the unitary irreducible representations. Therefore, we consider the space $\mathfrak{H}(S^{p-1})$. The unitary irreducible representations are classified by the number $l_{\{p/2\}}$, $l_{\{p/2\}} = 0, 1, 2, \cdots$ and they are defined by

$$SO(p) \ni g \to (U_g Y^{l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}})(\omega) = Y^{l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}}(g^{-1}\omega) \in \mathfrak{H}^{l_{\{p/2\}}}(S^{p-1}),$$

where $\mathfrak{H}^{l_{\{p/2\}}}(S^{p-1})$ is the subspace of the space $\mathfrak{H}(S^{p-1})$ determined by the vectors $Y^{l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}}(\omega)$ with fixed $l_{\{p/2\}}$.

(v) The decomposition of the quasi-regular representation (2.2) into the irreducible ones is trivial as

$$\mathfrak{H}(S^{p-1}) = \sum_{l_{\{p/2\}}=0}^{\infty} \oplus \mathfrak{H}^{l_{\{p/2\}}}(S^{p-1}).$$

[25] The proof of the completeness of the set of harmonic functions $Y^{l_2,\cdots,l_{\{p/2\}}}_{m_1,\cdots,m_{[p/2]}}(\omega)$ in $\mathfrak{H}(S^{p-1})$ in another coordinate system is given in Ref. 3.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 5 MAY 1967

Recursive Evaluation of Some Atomic Integrals

R. A. SACK,* C. C. J. ROOTHAAN, AND W. KOŁOS†
Laboratory of Molecular Structure and Spectra, Department of Physics, University of Chicago, Chicago, Illinois

(Received 26 September 1966)

A simple recursion scheme is set up for the higher derivatives of functions of the form $f(\alpha, \beta)(\alpha + \beta)^{-1}$ and $f(\alpha, \beta, \gamma)(\alpha + \beta + \gamma)^{-1}$ and applied to the rapid calculation of the Calais–Löwdin two-particle atomic integrals.

IN various branches of theoretical physics integrals occur which can be expressed in the general form

$$\Gamma_{lm}(\alpha, \beta) = \left(-\frac{\partial}{\partial\alpha}\right)^l \left(-\frac{\partial}{\partial\beta}\right)^m \frac{f(\alpha, \beta)}{\alpha + \beta}. \quad (1)$$

Application of Leibnitz' theorem to the differentiation with respect to both variables yields

$$\Gamma_{lm} = \sum \frac{l!\, m!\, (\lambda' + \mu')!}{\lambda!\, \lambda'!\, \mu!\, \mu'!\, (\alpha + \beta)^{\lambda'+\mu'+1}} \times \left(-\frac{\partial}{\partial\alpha}\right)^{\lambda} \left(-\frac{\partial}{\partial\beta}\right)^{\mu} f(\alpha, \beta), \quad \lambda + \lambda' = l, \quad \mu + \mu' = m. \quad (2)$$

The recurrence relation for the binomial coefficients

$$\frac{(\lambda' + \mu')!}{\lambda'!\, \mu'!} = \binom{\lambda' + \mu'}{\lambda'} = \binom{\lambda' + \mu' - 1}{\lambda'} + \binom{\lambda' + \mu' - 1}{\lambda' - 1} \quad (3)$$

can be applied to each term in the sum (2) except the term with $\lambda = \mu = 0$, with the result

$$\Gamma_{lm} = \frac{1}{\alpha + \beta}\left[l\Gamma_{l-1,m,n} + m\Gamma_{l,m-1,n} + \left(-\frac{\partial}{\partial\alpha}\right)^l \left(-\frac{\partial}{\partial\beta}\right)^m f(\alpha, \beta)\right]. \quad (4)$$

This recurrence formula can be used to particular advantage in the evaluation of atomic two-electron integrals for the case of exponential correlation

$$\Gamma_{lmn}(\alpha, \beta, \gamma) = \iint \exp(-\alpha r_1 - \beta r_2 - \gamma r_{12}) r_1^{l-1} r_2^{m-1} r_{12}^{n-1}\, d^3\mathbf{r}_1\, d^3\mathbf{r}_2, \quad (5)$$

where the integration is to be taken over all possible positions $\mathbf{r}_1$ and $\mathbf{r}_2$ of the two electrons and r_{12} denotes their separation. These integrals were discussed by

* Permanent address: Department of Mathematics, University of Salford, England.

† Permanent address: Department of Theoretical Chemistry, University of Warsaw, ul. Pasteura 1, Warsaw 22, Poland.

Calais and Löwdin[1] and expressed as series expansions on the basis of the relations

$$\Gamma_{000} = (4\pi)^2[(\alpha+\beta)(\alpha+\gamma)(\beta+\gamma)]^{-1}, \quad (6)$$

$$\Gamma_{lmn} = (-\partial/\partial\alpha)^l(-\partial/\partial\beta)^m(-\partial/\partial\gamma)^n\Gamma_{000}. \quad (7)$$

The same integrals had already been employed by the writers in their treatment of the S states of helium-like atoms and ions,[2] but they were evaluated recursively by the method described in this note. Application of (4) to (6) and (7) yields

$$\Gamma_{lmn} = (\alpha+\beta)^{-1}[l\Gamma_{l-1,m,n} + m\Gamma_{l,m-1,n} + B_{l,m;n}], \quad (8)$$

where

$$B_{l,m;n} = \left(-\frac{\partial}{\partial\alpha}\right)^l\left(-\frac{\partial}{\partial\beta}\right)^m\left(-\frac{\partial}{\partial\gamma}\right)^n\frac{(4\pi)^2}{(\alpha+\gamma)(\beta+\gamma)}. \quad (9)$$

The same relation (4) applied to the variables α and γ in (9) yields

$$B_{l,m;n} = (\alpha+\gamma)^{-1}[lB_{l-1,m;n} + nB_{l,m;n-1} + A_{l;m,n}], \quad (10)$$

where

$$A_{l;m,n} = \left(-\frac{\partial}{\partial\alpha}\right)^l\left(-\frac{\partial}{\partial\beta}\right)^m\left(-\frac{\partial}{\partial\gamma}\right)^n\frac{(4\pi)^2}{\beta+\gamma} = \frac{(4\pi)^2\delta_{l0}(m+n)!}{(\beta+\gamma)^{m+n+1}}. \quad (11)$$

Alternatively one can use the ratios

$$\Gamma'_{lmn} = \Gamma_{lmn}/(l!\,m!\,n!), \quad B'_{lm;n} = B_{lm;n}/(l!\,m!\,n!) \quad (12)$$

for which (8) becomes

$$\Gamma'_{lmn} = (\alpha+\beta)^{-1}[(1-\delta_{l0})\Gamma'_{l-1,m,n} + (1-\delta_{m0})\Gamma'_{l,m-1,n} + B'_{lm;n}] \quad (13)$$

and correspondingly for (10). Evaluation of a set of such integrals with $0 \leq l, m, n \leq 12$ on the Remington Rand Univac Scientific 1103 A computer took approximately 10 sec, the time required for their evaluation by the series expansion[1] had been estimated to as many minutes. In addition, the terms on the right-hand sides of (8), (10), and (13) are all positive so that the maximum accuracy of rounded numbers remains preserved.

Another set of integrals to which the relations (4) can be usefully applied are the overlap integrals between two spherically symmetric Slater-type orbitals centered on two atoms A and B a distance R apart.[3,4] The integrals

$$\Gamma_{l,m}(\alpha,\beta) = \int \exp(-\alpha r_A - \beta r_B) r_A^{l-1} r_B^{m-1}\, d^3\mathbf{r} \quad (14)$$

satisfy the relation (1) with

$$\Gamma_{00}(\alpha,\beta) = \frac{4\pi}{\alpha+\beta}\,\frac{e^{-\beta R} - e^{-\alpha R}}{\alpha R - \beta R}. \quad (15)$$

The recurrence relations resulting from the application of (4) to (14) and (15) have already been derived by one of the authors[3] by means of integration by parts, but to obtain the result he made use of spheroidal coordinates which are not required in the present approach. The derivatives of the second factor in (15) have recently been discussed in detail by Ruedenberg, O-Ohata, and Wilson[4]; provided these are given accurately, the relations (4) provide a stable process for obtaining the Γ_{lm} for large l and m beginning with the formula (15).

A number of generalizations of (4) follow straightforwardly. Thus for

$$\Gamma_{lmn}(\alpha,\beta,\gamma) = \left(-\frac{\partial}{\partial\alpha}\right)^l\left(-\frac{\partial}{\partial\beta}\right)^m\left(-\frac{\partial}{\partial\gamma}\right)^n\frac{f(\alpha,\beta,\gamma)}{\alpha+\beta+\gamma}, \quad (16)$$

the formula becomes

$$\Gamma_{lmn} = (\alpha+\beta+\gamma)^{-1}\Big[l\Gamma_{l-1,m,n} + m\Gamma_{l,m-1,n} + n\Gamma_{l,m,n-1} + \left(-\frac{\partial}{\partial\alpha}\right)^l\left(-\frac{\partial}{\partial\beta}\right)^m \times \left(-\frac{\partial}{\partial\gamma}\right)^n f(\alpha,\beta,\gamma)\Big]; \quad (17)$$

and correspondingly for more variables. One or other of the indices l or m may be negative, corresponding to a repeated integration with regard to the relevant variables; the recurrence relation (4) is still valid in this case, but the intervention of the factor 0 prevents any connection between expressions with a negative index l and those with $l \geq 0$. However, if a few such integrals are evaluated by other means, e.g., by numerical quadrature, the remainder can be evaluated from (4); but as the terms may now be of either sign, care must be taken to avoid accumulation of roundoff errors. Similar considerations apply to expressions with fractional indices l and m.

ACKNOWLEDGMENTS

This work was supported by the U.S. Office of Naval Research and by a grant from the National Science Foundation.

[1] J. L. Calais and P. O. Löwdin, J. Mol. Spectry. **8**, 203 (1962), cf. also P. J. Roberts, J. Chem. Phys. **43**, 3547 (1965).
[2] W. Kołos, C. C. J. Roothaan, and R. A. Sack, Rev. Mod. Phys. **32**, 178, (1960).
[3] C. C. J. Roothaan, J. Chem. Phys. **24**, 947 (1956).
[4] K. Ruedenberg, K. O-Ohata, and D. G. Wilson, J. Math. Phys. **7**, 539 (1966).

Exact Distributions of the Reduced-Width Amplitude

NAZAKAT ULLAH
Tata Institute of Fundamental Research, Bombay, India

(Received 21 February 1966)

The invariance hypothesis is used to derive the various multivariate distributions of the reduced-width amplitude. Simple expressions are given for the multi-level and multi-channel distributions valid for all dimensions of the random orthogonal matrix.

I. INTRODUCTION

In an earlier paper[1] we derived the asymptotic reduced-width amplitude distributions using the method of moments. This derivation has the advantage that it does not use the weak assumption of level independence.[2] It was shown that, in the limit of large dimension of the random orthogonal matrix, the distribution of the reduced-width amplitude is the same as the one obtained using the level-independence hypothesis.[3] In this paper we show that exact multi-level and multi-channel distributions of the reduced-width amplitude, valid for all dimensions of the random orthogonal matrix, can be obtained.

From the earlier paper[1] we write the reduced-width amplitude $\gamma_{\lambda c}$ for level λ and channel c as

$$\gamma_{\lambda c} = \sum_{\mu=1}^{N} a_{\mu\lambda} J_{\mu c}, \tag{1}$$

where $a_{\mu\lambda}$ are the elements of $N \times N$ random orthogonal matrix and $J_{\mu c}$ are the overlap integrals of the channel function and the wavefunctions of the basis set, which is used for the expansion of the compound-nucleus wavefunction.

The joint probability distribution $P(\{a_{\mu\lambda}\})$ of the elements of random orthogonal matrix is given by[4,5]

$$P(\{a_{\mu\lambda}\}) \prod_{\mu,\lambda=1}^{N} da_{\mu\lambda} = K^{-1}\left[\prod_{\lambda=1}^{N} \delta\left(\sum_{\mu=1}^{N} a_{\mu\lambda}^2 - 1\right)\right] \times \left[\prod_{\lambda<\lambda'}^{N} \delta\left(\sum_{\mu=1}^{N} a_{\mu\lambda} a_{\mu\lambda'}\right)\right] \prod_{\mu,\lambda=1}^{N} da_{\mu\lambda}, \tag{2}$$

where K is the normalization integral.

II. MULTI-LEVEL DISTRIBUTION

Let us consider a single channel c and levels $\lambda = 1, 2, \cdots, n$, where $n < N$. The joint probability distribution $P(\{\gamma_{\lambda c}\})$, using (1) and (2), can be written as

$$P(\{\gamma_{\lambda c}\}) = K^{-1} \int \left[\prod_{\lambda=1}^{n} \delta\left(\gamma_{\lambda c} - \sum_{\mu=1}^{N} a_{\mu\lambda} J_{\mu c}\right)\right] \times \left[\prod_{\lambda=1}^{N} \delta\left(\sum_{\mu=1}^{N} a_{\mu\lambda}^2 - 1\right)\right] \times \left[\prod_{\lambda<\lambda'}^{N} \delta\left(\sum_{\mu=1}^{N} a_{\mu\lambda} a_{\mu\lambda'}\right)\right] \prod_{\mu,\lambda=1}^{N} da_{\mu\lambda}. \tag{3}$$

Integrating over the column vectors $\lambda = n+1, \cdots, N$, Eq. (3) gives

$$P(\{\gamma_{\lambda c}\}) = K^{-1} \int \left[\prod_{\lambda=1}^{n} \delta\left(\gamma_{\lambda c} - \sum_{\mu=1}^{N} a_{\mu\lambda} J_{\mu c}\right)\right] \times \left[\prod_{\lambda=1}^{n} \delta\left(\sum_{\mu=1}^{N} a_{\mu\lambda}^2 - 1\right)\right] \times \left[\prod_{\lambda<\lambda'}^{n} \delta\left(\sum_{\mu=1}^{N} a_{\mu\lambda} a_{\mu\lambda'}\right)\right] \prod_{\mu=1,\lambda=1}^{N\,n} da_{\mu\lambda}, \tag{4}$$

where the normalization integral K is now the integral over n column vectors.

Let us make an orthogonal transformation on the variables $a_{\mu\lambda}$

$$a'_{\nu\lambda} = \sum_{\mu=1}^{N} C_{\nu\mu} a_{\mu\lambda}, \tag{5}$$

and choose

$$C_{1\mu} = J_{\mu c}\left(\sum_{\mu=1}^{N} J_{\mu c}^2\right)^{-\frac{1}{2}}, \quad \mu = 1, \cdots, N. \tag{6}$$

Since C is an orthogonal matrix, the scalar products and the volume element remain invariant. Using this transformation in expression (4) and carrying out the integrations we get

$$P(\{\gamma_{\lambda c}\}) = \{\Gamma(\tfrac{1}{2}N)/\Gamma[\tfrac{1}{2}(N-n)]\}(\pi N\langle\gamma_c^2\rangle)^{-\frac{1}{2}n} \times \left[1 - \left(\sum_{\lambda=1}^{n} \gamma_{\lambda c}^2 \Big/ N\langle\gamma_c^2\rangle\right)\right]^{\frac{1}{2}(N-n-2)}, \tag{7}$$

where

$$\langle\gamma_c^2\rangle \equiv \langle\gamma_{\lambda c}^2\rangle = \frac{1}{N}\sum_{\mu=1}^{N} J_{\mu c}^2.$$

For $n = N$, we get

$$P(\{\gamma_{\lambda c}\}) = \Gamma(\tfrac{1}{2}N)(\pi N\langle\gamma_c^2\rangle)^{-\frac{1}{2}N}\delta\left[1 - \left(\sum_{\lambda=1}^{N} \gamma_{\lambda c}^2 \Big/ N\langle\gamma_c^2\rangle\right)\right]. \tag{8}$$

[1] N. Ullah, J. Math. Phys. **6**, 1102 (1965).
[2] F. J. Dyson, J. Math. Phys. **3**, 140 (1962).
[3] T. J. Krieger and C. E. Porter, J. Math. Phys. **4**, 1272 (1963).
[4] N. Rosenweig, Phys. Letters **6**, 123 (1963); International Conference on Nuclear Physics with Reactor Neutrons, ANL 6797 (1963), p. 302.
[5] N. Ullah, Nucl. Phys. **58**, 65 (1964).

Expressions (7), (8) give the exact multi-level single-channel distribution of the reduced-width amplitude. Earlier results[1,3,4] can be obtained by expanding (7) for large values of N.

III. MULTI-CHANNEL DISTRIBUTION

We next consider a single level and m channels ($m < N$), $c1, c2, \cdots, cm$. Using (1) and (2) the multi-channel distribution $P(\{\gamma_c\})$ can be written as

$$P(\{\gamma_c\}) = K^{-1}\int\left[\prod_{j=1}^{m}\delta\left(\gamma_{cj} - \sum_{\mu=1}^{N} a_\mu J_{\mu cj}\right)\right] \times \delta\left(\sum_{\mu=1}^{N} a_\mu^2 - 1\right)\prod_{\mu=1}^{N} da_\mu, \quad (9)$$

where we dropped the level index λ, and K is again the appropriate normalization integral.

By taking proper linear combinations of the quantities $J_{\mu cj}$

$$J'_{\mu ci} = \sum_{j=1}^{m} T_{ij} J_{\mu cj}, \quad (10)$$

we can construct $J'_{\mu ci}$ which form m orthonormal vectors in N-dimensional space. Expression (9) then becomes

$$P(\{\gamma_c\}) = K^{-1}\int\left[\prod_{j=1}^{m}\delta\left(\gamma_{cj} - \sum_{\mu=1,k=1}^{N\ m} a_\mu (T^{-1})_{jk} J'_{\mu ck}\right)\right] \times \delta\left(\sum_{\mu=1}^{N} a_\mu^2 - 1\right)\prod_{\mu=1}^{N} da_\mu. \quad (11)$$

As we had done in Sec. II, we make an orthogonal transformation on the variables a_μ and now choose

$$C_{\nu\mu} = J'_{\mu c\nu}, \quad \nu = 1, \cdots, m. \quad (12)$$

Using this transformation and carrying out the integrations in (11) we can write

$$P(\{\gamma_c\}) = \frac{\Gamma(\frac{1}{2}N)\,|T|}{\Gamma[\frac{1}{2}(N-m)]\pi^{\frac{1}{2}m}}\,[1 - (\gamma, \tilde{T}T\gamma)]^{\frac{1}{2}(N-m-2)}, \quad (13)$$

where $|T|$ is the determinant of the matrix T.

Introducing the covariance matrix[6] Σ

$$\Sigma = \langle\gamma\tilde{\gamma}\rangle = N^{-1}(\tilde{T}T)^{-1}, \quad (14)$$

expression (13) becomes

$$P(\{\gamma_c\}) = \frac{\Gamma(\frac{1}{2}N)}{\Gamma[\frac{1}{2}(N-m)](\pi N)^{\frac{1}{2}m}|\Sigma|^{\frac{1}{2}}} \times \left[1 - \frac{1}{N}(\gamma, \Sigma^{-1}\gamma)\right]^{\frac{1}{2}(N-m-2)}. \quad (15)$$

For $m = N$ we get

$$P(\{\gamma_c\}) = \frac{\Gamma(\frac{1}{2})N}{(\pi N)^{\frac{1}{2}N}|\Sigma|^{\frac{1}{2}}}\,\delta\left[1 - \frac{1}{N}(\gamma, \Sigma^{-1}\gamma)\right]. \quad (16)$$

For $m > N$, $P(\{\gamma_c\})$ involves additional δ functions. This is because in an N-dimensional space we cannot choose more than N linearly independent J'_c vectors.

IV. MULTI-LEVEL, MULTI-CHANNEL DISTRIBUTION

The joint probability distribution $P(\{\gamma_{\lambda c}\})$ for the multi-level, multi-channel case is more difficult to write down explicitly. We give an expression for the case of two levels $\lambda = 1, 2$ and two channels $c1, c2$. It is given by

$$P(\gamma_{1c1}, \gamma_{1c2}, \gamma_{2c1}, \gamma_{2c2}) = \frac{(N-2)(N-3)}{4\pi^2 N^2 |\Sigma|}\left[1 - \frac{1}{N}\{(\gamma_1, \Sigma^{-1}\gamma_1) + (\gamma_2, \Sigma^{-1}\gamma_2)\} + \frac{1}{N^2|\Sigma|}(\gamma_{1c1}\gamma_{2c2} - \gamma_{1c2}\gamma_{2c1})^2\right]^{\frac{1}{2}(N-5)}. \quad (17)$$

Similar expressions can also be derived for the case of unitary and symplectic ensembles defined by Dyson.[2]

[6] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis* (John Wiley & Sons, Inc., New York, 1958), Chap. II.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 5 MAY 1967

Wigner Method in Quantum Statistical Mechanics*

KAYA İMRE† AND ERCÜMENT ÖZIZMIR‡
Çekmece Nuclear Research and Training Center, Istanbul, Turkey

MARCOS ROSENBAUM§
Department of Nuclear Engineering, The University of Michigan, Ann Arbor, Michigan

P. F. ZWEIFEL‖
Department of Theoretical Physics, Middle East Technical University, Ankara, Turkey, and Department of Nuclear Engineering, The University of Michigan, Ann Arbor, Michigan

(Received 28 March 1966)

The Wigner method of transforming quantum-mechanical operators into their phase-space analogs is reviewed with applications to scattering theory, as well as to descriptions of the equilibrium and dynamical states of many-particle systems. Inclusion of exchange effects is discussed.

I. INTRODUCTION

THE ensemble expectation value of a quantum-mechanical operator A is expressed in the usual formulation of quantum statistics, as the trace of ρA, namely,

$$\langle A \rangle = \mathrm{Tr}\, \rho A, \tag{1}$$

where ρ is the von Neumann density matrix[1] defined by

$$\rho = \sum_i w_i |\phi_i(t)\rangle\langle\phi_i(t)|, \tag{2}$$

where w_i is the probability that the system will be in the state $|\phi_i(t)\rangle$. The density matrix satisfies the equation

$$i\hbar(\partial\rho/\partial t) = H\rho - \rho H, \tag{3}$$

where H is the Hamiltonian of the system under consideration. In equilibrium, for a canonical ensemble one has

$$\rho = e^{-\beta H}/\mathrm{Tr}\, e^{-\beta H}. \tag{4}$$

In calculating expectation values of physical interest one may choose any convenient representation in which to work. Wigner[2] in 1932 introduced a method for evaluating expectation values which is particularly suitable for "almost classical" systems in that it expresses the expectation values as a power series expansion with respect to Planck's constant. For such systems the expansion may be expected to converge rapidly. Another important advantage of this method is that there are direct classical analogs of the quantities and operations used. In particular, the analog of classical phase space is introduced into quantum statistics. In this way the expectation values of physical variables may be expressed in terms of a phase-space integration. The purpose of the present study is to review various fields of application of this method.[3]

In Sec. II, we demonstrate that the Wigner method can be defined as a means for associating a c-number function in phase space with every operator which is a function of position and momentum operators. This rule, in fact, is the inverse of Weyl's rule, which is used to calculate quantum-mechanical operators from classical quantities. However, it is interesting to observe that there are various equivalent ways of stating the association which are, in many cases, simpler than Weyl's rule. Various properties and applications of this correspondence can be found in Sec. III. Section IV is devoted to the application to scattering theory. A method of inclusion of exchange effects in the previous results is given in Sec. V, where the second quantized formalism is also discussed.

Application to the equilibrium case (for Boltzmann statistics) is given in Appendixes A and B, where the equation of state is derived up to the order $\hbar^4$.

* The portion of this work carried out at the University of Michigan was supported by the U.S. Atomic Energy Commission. Some of this material is based on portions of a thesis of one of the authors (M. R.) presented to the University of Michigan in partial fulfillment of the requirements for the Ph.D. degree.

† Present address: Conductron Corporation, Ann Arbor, Michigan.

‡ Present address: Plasma Physics Laboratory, Princeton, New Jersey.

§ Present address: General Electric Company, TEMPO, Santa Barbara, California.

‖ On sabbatical leave, academic year 1964–1965 at Middle East Technical University. Permanent address: Department of Nuclear Engineering, The University of Michigan, Ann Arbor, Michigan.

[1] J. von Neumann, *Mathematical Foundation of Quantum Mechanics* (Princeton University Press, Princeton, New Jersey, 1955).

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

[3] There are several papers published which deal with the Wigner distribution function. Some of the basic references are: H. J. Groenewold, Physica **12**, 405 (1946); J. E. Moyal, Proc. Cambridge Phil. Soc. **45**, 99 (1949); J. H. Irving and R. W. Zwanzig, J. Chem. Phys. **19**, 1173 (1951); H. Mori, I. Oppenheim, and J. Ross, in *Studies in Statistical Mechanics*, J. de Boer and G. E. Uhlenbeck, Eds. (North-Holland Publishing Company, Amsterdam, 1962), Vol. 1.

II. WIGNER DISTRIBUTION FUNCTION

We restrict ourselves to Boltzmann statistics in this section, so that exchange effects are ignored. Also, we assume that the Hamiltonian of our system depends only on the position and momentum operators R and P.[4]

Wigner[2] defines a distribution function $f_w(r, p)$ as the Fourier transform of the off-diagonal elements of the density matrix

$$f_w(r, p) = (2\pi\hbar)^{-3N}\int dz \exp(ip \cdot z/\hbar) \times \left\langle r - \frac{z}{2}\middle|\rho\middle| r + \frac{z}{2}\right\rangle \quad (5a)$$

$$= (2\pi\hbar)^{-3N}\int dk \exp(-ir \cdot k/\hbar) \times \left\langle p - \frac{k}{2}\middle|\rho\middle| p + \frac{k}{2}\right\rangle. \quad (5b)$$

It is easily verified that f_w has the following properties:

$$\int dp f_w(r, p) = \langle r|\rho|r\rangle, \quad (6a)$$

$$\int dr f_w(r, p) = \langle p|\rho|p\rangle, \quad (6b)$$

and, evidently

$$\int dr\, dp f_w(r, p) = \mathrm{Tr}\,\rho = 1. \quad (6c)$$

Corresponding to a quantum operator $A(R, P)$, we define a function $A_w(r, p)$ by an equation analogous to (5):

$$A_w(r, p) = \int dz \exp(ip \cdot z/\hbar)\left\langle r - \frac{z}{2}\middle| A \middle| r + \frac{z}{2}\right\rangle \quad (7a)$$

$$= \int dk \exp(-ir \cdot k/\hbar)\left\langle p - \frac{k}{2}\middle| A \middle| p + \frac{k}{2}\right\rangle, \quad (7b)$$

which we call the Wigner equivalent of A. Thus we see that f_w is simply $(2\pi\hbar)^{-3N}$ times the Wigner equivalent of the density matrix ρ:

$$f_w = (2\pi\hbar)^{-3N}\rho_w. \quad (8)$$

Furthermore, from (5) and (7)

$$\mathrm{Tr}\,\rho A = \int dr\, dp A_w(r, p) f_w(r, p). \quad (9)$$

Equation (9) is the key result of the Wigner method, since it expresses the ensemble average of an operator A as a phase space integral.

The rules (5) or (7) for Wigner equivalent operators are actually equivalent to Weyl's rule[5] for defining the classical analog of a quantum operator. This rule may be stated most conveniently starting with the Fourier transform of the classical function $A_w(r, p)$:

$$\alpha(\sigma, \tau) = \int dr\, dp \exp[-i(\sigma \cdot r + \tau \cdot p)/\hbar] A_w(r, p). \quad (10a)$$

Then $A(R, P)$ is defined from

$$A(R, P) = \left(\frac{1}{2\pi\hbar}\right)^{6N} \times \int d\sigma\, d\tau \exp[i(\sigma \cdot R + \tau \cdot P)/\hbar]\alpha(\sigma, \tau). \quad (10b)$$

That the Weyl rules (10) relating A and A_w are identical with the Wigner rules, (7) is shown below. [Also, we show that Eqs. (10) "work both ways". Given A_w one may determine A, and vice versa.] There has apparently been some confusion in the literature, in which one frequently finds the statement that Eq. (9) holds with f_w defined through Eq. (5) and A_w through (10).

The equivalence of (7) and (10) may be proved as follows. We begin by proving the completeness (and orthogonality) of the operators $\exp[i(\sigma \cdot R + \tau \cdot P)]/\hbar$ in the space of operators of the form $A = A(R, P)$. We first write

$$\exp[i(\sigma \cdot R + \tau \cdot P)/\hbar] = \exp[i\sigma \cdot R/\hbar] \exp[i\tau \cdot P/\hbar] \exp[i\sigma \cdot \tau/2\hbar], \quad (11)$$

by making use of the identity $e^{A+B} = e^A e^B e^{\frac{1}{2}[B,A]}$ (true if the commutator $[B, A]$ commutes with both A and B). Since $\exp(i\tau \cdot p/\hbar)\,|r\rangle = |r - \tau\rangle$, we have

$$\langle r'|\exp[\pm i(\sigma \cdot R + \tau \cdot P)/\hbar]\,|r\rangle = \exp[\pm i\sigma \cdot (r \mp \tfrac{1}{2})/\hbar]\delta(r' - r \pm \tau), \quad (12)$$

which implies that

$$\mathrm{Tr}\exp[-i(\sigma \cdot R + \tau \cdot P)/\hbar] = (2\pi\hbar)^{3N}\delta(\sigma)\delta(\tau). \quad (13)$$

Therefore

$$\mathrm{Tr}\exp[-i(\sigma' \cdot R + \tau' \cdot P)/\hbar]\exp[i(\sigma \cdot R + \tau \cdot P)/\hbar] = \mathrm{Tr}\exp\{-i[(\sigma' - \sigma)R + (\tau' - \tau)P]/\hbar\} \times \exp\left[-\frac{1}{2}\frac{i}{\hbar}(\tau' \cdot \sigma - \sigma' \cdot \tau)\right], \quad (14a)$$

$$= (2\pi\hbar)^{3N}\delta(\sigma' - \sigma)\delta(\tau' - \tau). \quad (14b)$$

[4] In our notation, r, p represent $3N$-dimensional vector c numbers for position and momentum variables and R, P represent the corresponding vector operators. A $3N$-dimensional scalar product is written as $R.P$ or $r.p$. Also, R_i, P_i: r_i, p_i, etc. denote ordinary three-dimensional vectors associated with the ith particle.

[5] H. Weyl, *The Theory of Groups and Quantum Mechanics* (Dover Publications, New York, 1950).

[Equation (14a) was obtained trivially using the identity below Eq. (11).]

We thus have proved that the operators

$$(2\pi\hbar)^{-3N/2}\exp[i(\sigma\cdot R+\tau\cdot p)/\hbar]$$

are orthonormal. To prove completeness we attempt to expand an arbitrary operator $A(R,P)$ in terms of these functions:

$$A(R,P)=\int d\sigma\, d\tau\alpha(\sigma,\tau)\exp[i(\sigma\cdot R+\tau\cdot P)/\hbar]. \quad (15)$$

If the expansion exists then the coefficient $\alpha(\sigma,\tau)$ can be calculated by multiplying Eq. (15) by $\exp[-i(\sigma\cdot R'+\tau\cdot p')]$ and using Eq. (14):

$$\alpha(\sigma,\tau)=(2\pi\hbar)^{-3N}\times \mathrm{Tr}\{A(R',P')\exp[-i(\sigma\cdot R'+\tau\cdot P')/\hbar]\}. \quad (16)$$

To prove that the expansion (15) exists, i.e., to prove completeness, we substitute from (10) into (15) and prove that the result is an identity, say by taking matrix elements in the position representation.

$$\langle r|A|r'\rangle=(2\pi\hbar)^{-3N}\int d\sigma\, d\tau\, dr''\, dr'''\,\langle r''|A|r'''\rangle\times\langle r'''|\exp[-i(\sigma\cdot R'+\tau\cdot P')/\hbar]|r''\rangle\times\langle r|\exp[i(\sigma\cdot R+\tau\cdot P)/\hbar]|r'\rangle. \quad (17)$$

Making use of Eq. (13) and carrying out the trivial integration, Eq. (17) is seen to reduce to the identity

$$\langle r|A|r'\rangle=\langle r|A|r'\rangle, \quad (18)$$

which proves completeness (in the weak topological sense at least).

This proof permits us to consider Eqs. (10) to work in either direction, i.e., given $A(R,P)$, then $A_w(r,p)$ can be found and vice versa. Thus there exists a one-to-one correspondence between $A(R,P)$ and the c-number functions $A_w(r,p)$.

Finally, it is trivial to prove now that definitions (7) and (10) are equivalent. It is only necessary to substitute the expansion (10b) into (7a), use Eq. (12), and carry out the trivial integration, whereupon the Fourier inverse of (10a) is obtained.

In the next section we consider various properties of the correspondence between $A(R,P)$ and $A_w(r,p)$. In particular, we show a third rule for defining this equivalence (Groenewold's rule) which frequently is simpler to apply than either Wigner's or Weyl's rules [Eqs. (7) and (10), respectively]. In summarizing the results of the present section we have shown that given any two operators $A(R,P)$ and $B(R,P)$ that

$$\mathrm{Tr}\,AB=(2\pi\hbar)^{-3N}\int dr\, dp A_w(r,p)B_w(r,p), \quad (19)$$

where A_w and B_w are related to A and B through Wigner's rule (7) or equivalently Weyl's rule (10). In particular, the Wigner distribution function, $f_w(r,p)$ is simply $(2\pi\hbar)^{-3N}\rho_w$, where ρ_w is the Wigner equivalent of the density matrix.

III. WIGNER EQUIVALENT OF OPERATORS

From the results of the previous section, we immediately deduce the following properties:

(a) if $A=A(P)$ (i.e., independent of R), then $A_w=A(p)$;
(b) if $A=A(R)$, then $A_w=A(r)$;
(c) if $A=\text{const}$, then $A_w=A$;
(d) $\mathrm{Tr}\,A=(2\pi\hbar)^{-3N}\int dr\, dp A_w(r,p)$;
(e) $\int dp A_w(r,p)=(2\pi\hbar)^{3N}\langle r|A|r\rangle$;
(f) $\int dr A_w(r,p)=(2\pi\hbar)^{3N}\langle p|A|p\rangle$;
(g) $\langle r|A|r'\rangle=(2\pi\hbar)^{-3N}\int dp\exp[ip\cdot(r-r')/\hbar]\times A_w(\frac{1}{2}(r+r'),p)$
$=(2\pi\hbar)^{-6N}\int d\sigma\exp[i\sigma\cdot(r+r')/2\hbar]\times\alpha(\sigma,r'-r)$; (20)

where $\alpha(\sigma,\tau)$ is the Fourier transform of $A_w(r,p)$ as in Eq. (10b).

Wigner Equivalent of Products

Next we consider the Wigner equivalent of a product of operators AB, and derive a formula which expresses $(AB)_w$ in terms of A_w and B_w. We have

$$(AB)_w=\int dz\exp(ip\cdot z/\hbar)\langle r-\tfrac{1}{2}z|AB|r+\tfrac{1}{2}z\rangle \quad (21a)$$

$$=\int dz\, dr'\exp(ip\cdot z/\hbar)\times\langle r-\tfrac{1}{2}z|A|r'\rangle\langle r'|B|r+\tfrac{1}{2}z\rangle, \quad (21b)$$

or

$$(AB)_w=(2\pi\hbar)^{-12N}\int dz\, dr'\exp(ip\cdot z/\hbar)\times\int d\sigma\, d\sigma'\exp[i\sigma\cdot(r+r'-\tfrac{1}{2}z)/2\hbar]\times\alpha(\sigma,r'-r+\tfrac{1}{2}z)\times\exp[i\sigma'\cdot(r'+r+\tfrac{1}{2}z)/2\hbar]\times\beta(\sigma',r-r'+\tfrac{1}{2}z). \quad (22)$$

Here we have used Eq. (20g) for both A and B. Now, making the change of variables

$$\tau=r'-r+\tfrac{1}{2}z,\quad \tau'=r-r'+\tfrac{1}{2}z,$$

we readily obtain

$$(AB)_w = (2\pi\hbar)^{-12N}\int d\sigma\, d\sigma'\, d\tau\, d\tau' \times \exp[i(\sigma\cdot r + \tau\cdot p)/\hbar]\alpha(\sigma,\tau) \times \exp[i(\sigma'\cdot\tau - \sigma\cdot\tau')/2\hbar]\beta(\sigma',\tau') \times \exp[i(\sigma'\cdot r + \tau'\cdot p)/\hbar]. \quad (23)$$

The factor $\exp[i(\sigma'\cdot\tau - \sigma\cdot\tau')/2\hbar]$ in the latter integrand can be replaced by $\exp(\hbar\Lambda/2i)$, where Λ is the Poisson bracket operator, i.e.,

$$\Lambda = \overleftarrow{\nabla}_p\cdot\overrightarrow{\nabla}_r - \overleftarrow{\nabla}_r\cdot\overrightarrow{\nabla}_p, \quad (24)$$

so that $A_w\Lambda B_w = (B_w; A_w)$ is the standard classical Poisson bracket. (The arrows on the $3N$-dimensional gradient operators indicate the direction in which they operate.) We, therefore, obtain the formulas

$$(AB)_w = A_w(r,p)\exp(\hbar\Lambda/2i)B_w(r,p) \quad (25a)$$

$$= B_w(r,p)\exp(-\hbar\Lambda/2i)A_w(r,p) \quad (25b)$$

$$= A_w\big(r - (\hbar/2i)\nabla_p, p + (\hbar/2i)\nabla_r\big)B_w(r,p). \quad (25c)$$

The above result is due to Groenewold,[6] which, by successive application, along with (20a)–(20c) permits one to calculate the Wigner equivalent of any operator. This rule, then is equivalent to the two rules (Wigner's and Weyl's) described in the previous section. In particular, from Eq. (20d)

$$\mathrm{Tr}\,AB = (2\pi\hbar)^{-3N}\int dr\,dp A_w(r,p)\exp(\hbar\Lambda/2i)B_w(r,p) = (2\pi\hbar)^{-3N}\int dr\,dp A_w(r,p)B_w(r,p). \quad (26)$$

To obtain the latter, we have performed $6N$ partial integrations on the former, which flip the arrows pointing to the left to the right, thus making $\Lambda\to 0$. This result thus agrees with Eq. (19).

We see that A_w is expressed as a power series in $\hbar$. Similarly f_w is so expressed and, in particular, so are thermal expectation values.

Wigner Equivalent of Heisenberg Operators

For $A(t) = \exp(itH/\hbar)A(0)\exp(-itH/\hbar)$, one has

$$\partial A(t)/\partial t = (i/\hbar)(HA - AH). \quad (27)$$

Thus, upon forming the Wigner equivalent, we obtain

$$\partial A_w(t)/\partial t = (i/\hbar)[H_w\exp(\hbar\Lambda/2i)A_w - A_w\exp(\hbar\Lambda/2i)H_w] = (i/\hbar)[H_w\exp(\hbar\Lambda/2i)A_w - H_w\exp(-\hbar\Lambda/2i)A_w] \quad (28a)$$

or

$$\partial A_w(t)/\partial t = (2/\hbar)H_w\sin(\hbar\Lambda/2)A_w(t). \quad (28b)$$

Integration gives the formula

$$A_w(t) = \exp[(2t/\hbar)H_w\sin(\hbar\Lambda/2)]A_w(0). \quad (29)$$

To lowest order in $\hbar$ this is simply the classical equation of motion.

Quantum Liouville and Bloch Equations

The so-called quantum Liouville equation, which determines the time evolution of the Wigner distribution function, can be deduced readily by forming the Wigner equivalent of Eq. (3) with the aid of Eq. (8). By similar manipulations used to obtain Eq. (28b), we get

$$\partial f_w(t)/\partial t = -(2/\hbar)H_w\sin(h\Lambda/2)f_w(t), \quad (30)$$

which can be solved formally as

$$f_w(t) = \exp[-(2t/\hbar)H_w\sin(\hbar\Lambda/2)]f_w(0). \quad (31)$$

To the lowest order in $\hbar$, the above equations reduce to the classical Liouville equation:

$$\partial f_w^C/\partial t = -H_w\Lambda f_w^C \Rightarrow f_w^C = \exp(-tH_w\Lambda)f_w^C(0). \quad (32)$$

Equation (30) may be solved in powers of $\hbar^2$ starting from the classical distribution function. Such a procedure has, in fact, been followed by Wigner[2] and by Irving and Zwanzig.[3]

For a canonical ensemble in equilibrium, one has [cf. Eq. (4)],

$$\rho Z(\beta) = \Omega = e^{-\beta H}, \quad (33)$$

where Ω is the so-called unnormalized density matrix and $Z(\beta)$ is the partition function. Forming the Wigner equivalent, we get

$$\Omega_w = (e^{-\beta H})_w, \quad (34)$$

which may be evaluated as a power series expansion with respect to an appropriate parameter. However, there is another way to handle this evaluation which is suitable for almost-classical systems, and which makes use of the equation[7]

$$\partial\Omega/\partial\beta = -H\Omega = -\Omega H. \quad (35)$$

The Wigner equivalent of the latter is

$$\partial\Omega_w/\partial\beta = -H_w\exp(\hbar\Lambda/2i)\Omega_w = -\Omega_w\exp(\hbar\Lambda/2i)H_w = -H_w\exp(-\hbar\Lambda/2i)\Omega_w \quad (36)$$

or

$$\partial\Omega_w/\partial\beta = -H_w\cos(\hbar\Lambda/2)\Omega_w, \quad (37a)$$

[6] H. J. Groenewold, Ref. 3; several properties of the Wigner method have been first given in this work.

[7] This approach has been used by I. Oppenheim and J. Ross, Phys. Rev. 107, 28 (1957).

with

$$\Omega_w(\beta = 0) = 1. \tag{37b}$$

Equation (37a) is known as the Bloch equation, which provides a means for calculating Ω_w (and thus f_w) in powers of $\hbar^2$, which calculation may be found in Appendix A. In this appendix, this method is employed to deduce the $\hbar^4$ correction in the equation of state.

IV. APPLICATION TO SCATTERING THEORY

The differential scattering cross section (in first Born approximation) for a system of interacting particles can be written in the form[8]

$$\partial^2\sigma/\partial\Omega\partial\epsilon = CS(\mathbf{q}, \epsilon), \tag{38}$$

where C is a factor depending upon the momenta of the incoming and outgoing particles and upon the scattering potential for neutron scattering (which for neutron scattering may be taken as the Fermi pseudopotential[9]). Also, $\epsilon = \hbar w$ and $\mathbf{q} = \hbar\mathbf{x}$ are the energy and momentum changes in the scattering event. It is customary to express $S(\mathbf{q}, \epsilon)$ (called the "scattering law") in terms of Fourier transformed functions $\chi(\mathbf{q}, t)$ and/or $G(\mathbf{r}, t)$ as

$$S(\mathbf{q}, \epsilon) = (2\pi\hbar)^{-1}N\int dt \exp(-i\epsilon t/\hbar)\chi(\mathbf{q}, t) \tag{39a}$$

$$= (2\pi\hbar)^{-1}N\int d^3r\, dt \times \exp[i(\mathbf{q}\cdot\mathbf{r} - \epsilon t)\hbar]G(\mathbf{r}, t). \tag{39b}$$

Thus χ and G are related by

$$\chi(\mathbf{q}, t) = \int d^3r \exp(i\mathbf{q}\cdot\mathbf{r})G(\mathbf{r}, t) \tag{40a}$$

or

$$G(\mathbf{r}, t) = (2\pi\hbar)^{-3}\int d^3q \exp(-i\mathbf{q}\cdot\mathbf{r})\chi(\mathbf{q}, t). \tag{40b}$$

Explicitly χ and G are related to the density fluctuations of the scattering system:

$$G(\mathbf{r}, t) = \left\langle \frac{1}{N}\sum_{i,j=1}^{N}\int d^3r'\delta(\mathbf{r} + \mathbf{R}_i(0) - \mathbf{r}') \times \delta(\mathbf{r}' - \mathbf{R}_j(t))\right\rangle, \tag{41a}$$

$$\chi(\mathbf{q}, t) = \left\langle \frac{1}{N}\sum_{ij}\exp[-i\mathbf{q}\cdot\mathbf{R}_i(0)/\hbar] \times \exp[i\mathbf{q}\cdot\mathbf{R}_j(t)/\hbar]\right\rangle, \tag{41b}$$

where $\mathbf{R}_j(t)$ is the Heisenberg position operator corresponding to the jth scattering center (out of the total N).

For almost classical systems, it is desirable to relate the cross section to the classical time-dependent correlation function

$$G^C(\mathbf{r}, t) = \frac{1}{N}\sum_{i,j}\langle\delta(\mathbf{r} + \mathbf{r}_i(0) - \mathbf{r}_j(t))\rangle_C, \tag{42}$$

where $\langle\delta(\mathbf{r} + \mathbf{r}_i(0) - \mathbf{r}_j(t))\rangle_C$ denotes the classical thermal average. (In this way, scattering data can be used to give a physical picture of the scattering system[10] or alternatively, cross sections can be calculated from a knowledge of the classical mechanics of the scattering system.[11])

Following the approach of Aamodt *et al.*,[12] we consider from Eq. (41b)

$$\chi_{ij}(\mathbf{q}, t) = \mathrm{Tr}\,\rho\exp[-i\mathbf{q}\cdot\mathbf{R}_i(0)/\hbar] \times \exp[i\mathbf{q}\cdot\mathbf{R}_j(t)/\hbar] \tag{43a}$$

$$= \int dp'\, dr' A_i(r', p')B_j(r', p', t), \tag{43b}$$

where, from the application of the Wigner rules derived in the previous section, we have

$$A_i(r, p) = f_w(r, p)\exp(\hbar\Lambda/2i)\exp(-i\mathbf{q}\cdot\mathbf{r}_i/\hbar) \tag{44a}$$

$$= \exp(-i\mathbf{q}\cdot\mathbf{r}_i/\hbar)\exp(-\tfrac{1}{2}\mathbf{q}\cdot\nabla\mathbf{p}_i')f_w(r, p), \tag{44b}$$

and

$$B_j(r, p, t) = \exp[(2t/\hbar)H_w\sin(\hbar\Lambda/2)]\exp(i\mathbf{q}\cdot\mathbf{r}_j/\hbar), \tag{45a}$$

$$= \exp\{(t/m)\mathbf{p}\cdot\overrightarrow{\nabla}_r - (2t/\hbar)\Phi(r) \times \sin(\hbar/2)\overleftarrow{\nabla}_r\cdot\overrightarrow{\nabla}_p\}\exp(i\mathbf{q}\cdot\mathbf{r}_j/\hbar). \tag{45b}$$

[In the last line we have used a special form for the Hamiltonian, namely, $H_w = p^2/2m + \Phi(r)$.] Further, we observe

$$\chi(\mathbf{q}, t) = \frac{1}{N}\sum_{i,j}\chi_{ij}(\mathbf{q}, t). \tag{46}$$

Here we consider only the lowest-order contributions in $\hbar$,[13] so that f_w can be taken proportional to $e^{-\beta H_w}$ (cf. Appendix A). We then can write

$$A_i(r, p) = f_w\exp(-\beta\mathbf{q}^2/8m) \times \exp(\beta\mathbf{q}\cdot\mathbf{p}_i/2m)\exp(-i\mathbf{q}\cdot\mathbf{r}_i/\hbar), \tag{47}$$

[8] L. Van Hove, Phys. Rev. **95**, 249 (1954).

[9] E. Fermi, Ric. Sci. 7, 13 (1938); G. C. Summerfield, Ann. Phys. (N.Y.) **26**, 72 (1964).

[10] B. N. Brockhouse in *Proceedings of the Symposium on Inelastic Scattering of Neutrons in Solids and Liquids* (International Atomic Energy Commission, Vienna, 1960).

[11] R. Nossal, Phys. Rev. **135**, A1579 (1964).

[12] R. Aamodt, K. M. Case, M. Rosenbaum, and P. F. Zweifel, Phys. Rev. **126**, 1165 (1962).

[13] Higher-order corrections are studied in a paper by M. Rosenbaum and P. F. Zweifel, Phys. Rev. **137**, B271 (1965); Also see M. Rosenbaum, Doctoral Thesis, University of Michigan (1964).

where we have used the Taylor's series property

$$\exp(-a\cdot\nabla p)f(p)=f(p-a). \tag{48}$$

To the same order

$$B_j(r,p,t)=\exp(tH_w\Lambda)\exp(i\mathbf{q}\cdot\mathbf{r}_j/\hbar) \tag{49a}$$

$$=\exp[i\mathbf{q}\cdot\mathbf{r}_j(t)/\hbar]. \tag{49b}$$

Thus to lowest order

$$\chi(\mathbf{q},t)=\frac{1}{N}\sum_{ij}\int dp'\,dr' f_w\exp(-\beta\mathbf{q}^2/8m)\times\exp(\beta\mathbf{q}\cdot\mathbf{p}_i/2m)\exp(-i\mathbf{q}\cdot\mathbf{r}_i/\hbar)\times\exp(i\mathbf{q}\cdot\mathbf{r}_j(t)/\hbar], \tag{50}$$

and

$$G(\mathbf{r},t)=\frac{1}{N}(2\pi\hbar)^{-3}\int d^3q\exp(-i\mathbf{q}\cdot\mathbf{r})\times\exp(-\beta\mathbf{q}^2/8m)\langle\exp(\beta\mathbf{q}\cdot\mathbf{p}_i/2m)\times\exp(-i\mathbf{q}\cdot\mathbf{r}_i/\hbar)\exp[i\mathbf{q}\cdot\mathbf{r}_j(t)/\hbar]\rangle. \tag{51}$$

Suppose, in Eq. (51), we replace $\mathbf{r}_i(0)+i\mathbf{p}_i(0)\beta\hbar/2m$ by $\mathbf{r}_i(\frac{1}{2}i\beta\hbar)$, which is correct to order $\hbar^2$. Then

$$G(\mathbf{r},t)=(2\pi\hbar)^{-3}N^{-1}\int d^3q\exp(-i\mathbf{q}\cdot\mathbf{r})\times\exp(-\beta\mathbf{q}^2/8m)\langle\exp[i\mathbf{q}\cdot\mathbf{r}_j(t)/\hbar]\times\exp[-i\mathbf{q}\cdot\mathbf{r}_i(\tfrac{1}{2}i\beta\hbar)]\rangle \tag{52a}$$

$$=(2\pi\hbar)^{-3}N^{-1}\int d^3q\exp(-i\mathbf{q}\cdot\mathbf{r})\times\exp(-\beta\mathbf{q}^2/8m)\langle\exp[-i\mathbf{q}\cdot\mathbf{r}_i(0)]\times\exp[i\mathbf{q}\cdot\mathbf{r}_j(t-\tfrac{1}{2}i\beta\hbar)]\rangle. \tag{52b}$$

To obtain Eq. (52b) time-translational invariance has been employed.

Referring now to (39b), we see, after some trivial manipulations, that

$$S(\mathbf{q},\epsilon)=\exp(\beta\epsilon/2)\exp(-\beta q^2/8m)S_C(\mathbf{q},\epsilon), \tag{53}$$

where $S_C(\mathbf{q},\epsilon)$ is related to $G_C(\mathbf{r},t)$ through Eq. (39b). This is the desired result (to lowest order in $\hbar$) since it expresses the cross section in terms of the Fourier transform of the classical time-dependent correlation function. It is referred to as the "quasi-classical" approximation.

The integration of Eqs. (52) can be carried out explicitly, to give

$$G(\mathbf{r},t)=\frac{1}{N}\left(\frac{2m}{\pi\hbar}\right)^{\frac{3}{2}}\times\sum_{ij}\langle\exp\{-(2m/\beta\hbar^2)[\mathbf{r}+\mathbf{r}_i(0)-\mathbf{r}_j(t-\tfrac{1}{2}i\beta\hbar)]\}\rangle. \tag{54}$$

Now let us consider the scattering law as $\hbar\to 0$;

$$S(\mathbf{q},\epsilon)=(2\pi\hbar)^{-1}\int dt\exp(-i\epsilon t/\hbar)\times\sum_{i,j}\int dp\,dr A_iB_j. \tag{55}$$

Changing the dummy variable by $t=\hbar\tau$, and utilizing (44) and (45), we obtain

$$S(q,\epsilon)=\exp(-\beta q^2/8m)\sum_{i,j}\int dp\,dr f_w\times\exp(\beta\mathbf{q}\cdot\mathbf{p}_i/2m)\exp(-i\mathbf{q}\cdot\mathbf{r}_i/\hbar)\times(2\pi)^{-1}\int d\tau\exp(-i\epsilon\tau)\times\exp[i\mathbf{q}\cdot\mathbf{r}_j(\tau\hbar)/\hbar]. \tag{56}$$

Assuming the limit $\hbar\to 0$ can be taken before the integrations, we can replace $\mathbf{r}_j(\tau\hbar)\to\mathbf{r}_j+\tau\hbar\mathbf{p}_j/m$ to get

$$S(\mathbf{q},\epsilon)=\exp(-\beta q^2/8m)\exp(\beta\epsilon/2)\sum_{i,j}\int dp\,dr f_w\times\exp[i\mathbf{q}\cdot(\mathbf{r}_j-\mathbf{r}_i)/\hbar]\delta(\epsilon-\mathbf{q}\cdot\mathbf{p}_i/m) \tag{57a}$$

$$=\frac{1}{N}S^0(\mathbf{q},\epsilon)\sum_{i,j}\int dr n_N\times\exp[i\mathbf{q}\cdot(\mathbf{r}_j-\mathbf{r}_i)/\hbar], \tag{57b}$$

where

$$S^0(q,\epsilon)=N(m\beta/2\pi q^2)^{\frac{1}{2}}\exp(-\beta q^2/8m)\times\exp(\beta\epsilon/2)\exp(-m\beta\epsilon^2/2q^2) \tag{58}$$

is the scattering law corresponding to the ideal gas, and $n_N=\int dpf_w$. It is noted that the "self-terms", i.e., $i=j$, give the ideal gas result. For the case of binary central potential we obtain (with $\rho=N/V$)

$$S(\mathbf{q},\epsilon)=S^0(\mathbf{q},\epsilon)\left[1+\rho^{-1}\int d^3r n_2(\mathbf{r},0)\times\exp(i\mathbf{q}\cdot\mathbf{r}/\hbar)\right]$$

$$=S^0(\mathbf{q},\epsilon)\left[1+\rho\int d^3r g(r)\exp(i\mathbf{q}\cdot\mathbf{r}/\hbar)\right], \tag{59}$$

where $n_s(\mathbf{r}_1,\cdots,\mathbf{r}_s)$ is the s-particle reduced distribution function in configuration space:

$$n_s(\mathbf{r}_1,\cdots,\mathbf{r}_s)=[N!/(N-s)!]\int n_N\,d^3r_{s+1}\cdots d^3r_N, \tag{60}$$

and $g(r)=\rho^{-2}n_2(\mathbf{r},0)$ is the usual radial distribution function. In deriving (59), we have also used the translational invariance property, viz., for any $\mathbf{a}$,

$$n_s(\mathbf{r},\cdots,\mathbf{r}_s)=n_s(\mathbf{r}+\mathbf{a},\cdots,\mathbf{r}_s+\mathbf{a}), \tag{61}$$

which follows from the homogeneity of the scattering system.

For an ideal gas $n_N = V^{-N}$; $n_2 = N(N-1)V^{-2}$; thus $S(\mathbf{q}, \epsilon) = S^0(\mathbf{q}, \epsilon)$ if one drops the contribution of the "distinct terms" ($i \neq j$) which corresponds to the incident beam ($\mathbf{q} = 0, \epsilon = 0$) in the first Born approximation.[12,13]

The point here is that if the limit $\hbar \to 0$ is taken before the t (or τ) or ($\mathbf{r}$, t) integrations are performed, one gets a manifestly incorrect result, i.e., basically the ideal gas result for any system. It is important to go to the limit correctly, i.e., to use the "quasi-classical" approximation.

V. SOME CONSIDERATIONS FOR THE INCLUSION OF EXCHANGE EFFECTS

In previous sections, we have assumed the Boltzmann statistics to hold in the system at hand, so that the exchange effects due to the symmetry of the system have been ignored. We now outline a method which enables one to modify the previous results to include such effects.

For simplicity, we assume that the particles in the system are all identical, so that the state vectors (for example, in the coordinate representation) can be written as

$$|r\rangle^\theta = (N!)^{-\frac{1}{2}} \sum_P \theta^{|P|} |Pr\rangle, \tag{62}$$

where the summation is over all permutations P of $\mathbf{r}, \cdots, \mathbf{r}_N$, $|P|$ is the parity of P, and $\theta = 1$ for bosons and $\theta = -1$ for fermions.

The symmetrized (or antisymmetrized) Wigner equivalent of an operator A is then written as

$$\begin{aligned} A_w^\theta(r, p) &= \int dz \exp(ip \cdot z/\hbar) \\ &\quad \times \left\langle r - \frac{z}{2} \middle| A \middle| r + \frac{z}{2} \right\rangle^\theta \Big/ N! \\ &= \int dk \exp(-ir \cdot k/\hbar) \\ &\quad \times \left\langle p - \frac{k}{2} \middle| A \middle| p + \frac{k}{2} \right\rangle^\theta \Big/ N!. \end{aligned} \tag{63}$$

Our aim in this section is to express A_w^θ in terms of the previously defined quantity A_w.

Let us associate with every P an operator U_P defined as

$$U_P |r\rangle = |Pr\rangle. \tag{64}$$

It is readily seen that U_P is unitary, namely $U_P^\dagger = U_{P^{-1}} = U_P^{-1}$. Also, let

$$I^\theta = (1/N!) \sum_P \theta^{|P|} U_P. \tag{65}$$

It can be verified that I^θ is Hermitian, and also that

$$U_P I^\theta = \theta^{|P|} I^\theta = I^\theta U_P, \tag{66a}$$

$$I^\theta I^\theta = I^\theta, \tag{66b}$$

$$I^\theta A I^\theta = A^S I^\theta = I^\theta A^S, \tag{66c}$$

where $A^S = (1/N!) \sum_P U_P A$. If A is symmetrical (that is, $U_P A = A$, for all P) then $A^S = A$ (as is the case for all observable operators). We thus drop the superscript S.

In this notation, the matrix elements

$$\begin{aligned} {}^\theta\langle r| A |r'\rangle^\theta &= N! \langle r| I^\theta A I^\theta |r'\rangle \qquad (67a) \\ &= N! \langle r| A I^\theta |r'\rangle. \qquad (67b) \end{aligned}$$

Therefore, from Eq. (63), we obtain, using (25a),

$$A_w^\theta(r, p) = A_w(r, p) \exp(\hbar\Lambda/2i) I_w^\theta(r, p), \tag{68}$$

where $I_w^\theta(r, p)$ is the symmetrized Wigner equivalent of the identity operator [i.e., Eq. (63) for $A = I$], or it is the ordinary Wigner equivalent of I^θ.

In order to calculate the expectation value of an operator A, we consider the definition

$$\langle A \rangle = \sum_i w_i \langle \phi_i| A |\phi_i\rangle. \tag{69}$$

Because of the symmetry of the system, the state vectors in the position representation possess the invariance property

$$\begin{aligned} \langle r \mid \phi_i\rangle &= \theta^{|P|} \langle Pr \mid \phi_i\rangle \qquad (70a) \\ &= \theta^{|P|} \langle r| U_p |\phi_i\rangle \qquad (70b) \end{aligned}$$

for any P. Thus

$$\begin{aligned} \langle A \rangle = (1/N!)^2 \sum_{P'} \sum_P \theta^{|P|+|P'|} \sum_i w_i \int dr\, dr' \langle Pr \mid \phi_i\rangle \\ \times \langle \phi_i \mid P'r'\rangle \langle r'| A |r\rangle, \end{aligned} \tag{71a}$$

$$\langle A \rangle = (1/N!) \int dr\, dr'\, {}^\theta\langle r'| \rho |r\rangle^\theta \langle r| A |r'\rangle, \tag{71b}$$

where we have used the definition (2). With Eq. (67), we thus have

$$\langle A \rangle = \operatorname{Tr} \rho I^\theta A, \tag{72}$$

since ρ is symmetrical. Using our previous results, we get

$$\langle A \rangle = \int dr\, dp f_w^\theta(r, p) A_w(r, p), \tag{73}$$

where

$$\begin{aligned} f_w^\theta &= (2\pi\hbar)^{-3N} \rho_w(r, p) \exp(\hbar\Lambda/2i) I_w^\theta(r, p) \qquad (74a) \\ &= (2\pi\hbar)^{-3N} I_w^\theta(r, p) \cos\frac{\hbar\Lambda}{2} \rho_w(r, p). \qquad (74b) \end{aligned}$$

In the last step we have used the relation $I^\theta \rho = \rho I^\theta$. Alternatively we can write

$$\langle A \rangle = \int dr\, dp f_w(r, p) A_w^\theta(r, p). \tag{75}$$

We thus have reduced the calculation of the exchange effects to the evaluation of the quantity

$$N!\, I_w^\theta = \sum_P \theta^{|P|} \int dz \exp(ip\cdot z/\hbar) \times \delta(r - Pr - \tfrac{1}{2}z - \tfrac{1}{2}Pz) \quad (76a)$$

$$= \sum_P \theta^{|P|} \int dk \exp(-ik\cdot r/\hbar) \times \delta(p - Pp - \tfrac{1}{2}k - \tfrac{1}{2}Pk). \quad (76b)$$

Let us note that I_w^θ is an even function with respect to both r and p. This can be seen by using the property $\delta(r - Pr') = \delta(r' - P^{-1}r)$ and changing the dummy variables $z \to -z, P \to P^{-1}$ in (76a). This manipulation shows that the factor $\exp(ip\cdot z/\hbar)$ in the integrand can be replaced by $\cos p\cdot z/\hbar$ (therefore I_w^θ is real). Similar manipulations on (76b) yield the symmetry property with respect to r.

Secondly, we observe that the term in Eqs. (76) corresponding to the identity permutation is 1, which corresponds in turn to Boltzmann statistics. [The factor $N!$ is to be replaced by $(1!)^N = 1$ in the latter statistics, since all particles are distinct.] The evaluation of I_w^θ as a power series in $\hbar$ is not appropriate because of the essential singularity at $\hbar = 0$. One can reduce this evaluation to the calculation of the contributions of cyclic permutations.

Second Quantized Approach[14]

An alternative way to handle the problem of inclusion of the exchange effects is to utilize the second quantized formalism. The (anti)symmetrized position ket vectors can be written as

$$I^\theta |r\rangle = (N!)^{-\frac{1}{2}} \psi^\dagger(\mathbf{r}_1)\cdots\psi^\dagger(\mathbf{r}_N)|0\rangle, \quad (77)$$

where $\psi(r)$ and its Hermitian conjugate $\psi^\dagger(\mathbf{r})$ are the annihilation and creation operators which satisfy the usual (anti)commutation relations corresponding to (fermions) bosons. The matrix elements

$${}^\theta\langle r'|\, A\, |r\rangle^\theta$$

$$= \langle 0|\, \psi(\mathbf{r}'_{N\to 1})\cdots\psi(\mathbf{r}'_{N\to 1}) A \psi^\dagger(\mathbf{r}_1)\cdots\psi^\dagger(\mathbf{r}_N)\,|0\rangle \quad (78a)$$

$$= \mathrm{Tr}\, A\psi^\dagger(\mathbf{r}_1)\cdots\psi^\dagger(\mathbf{r}_N)\psi(\mathbf{r}'_N)\cdots\psi(\mathbf{r}'_1), \quad (78b)$$

where we have used the fact that the total number of particles, N, in the system is constant. Substituting in Eq. (73), and utilizing (20g) we obtain

$$\langle A\rangle = \int dr\, dp f_w^\theta(r, p, t) A_w(r, p), \quad (79)$$

where

$$f_w^\theta(r, p, t) = (2\pi\hbar)^{-3N}(N!)^{-1}\int dz \exp(ip\cdot z/\hbar) \times \mathrm{Tr}\,\rho(t)\psi^\dagger(\mathbf{r}_1 + \tfrac{1}{2}\mathbf{z}_1)\cdots\psi^\dagger(\mathbf{r}_N + \tfrac{1}{2}\mathbf{z}_N) \times \psi(\mathbf{r}_N - \tfrac{1}{2}\mathbf{z}_N)\cdots\psi(\mathbf{r}_1 - \tfrac{1}{2}\mathbf{z}_1), \quad (80)$$

which represents the Wigner distribution function in the second quantized formalism.

Equation (80) can also be written in the Heisenberg picture

$$f_w^\theta(r, p, t) = (N!)^{-1}(2\pi\hbar)^{-3N}\int dz \exp(ip\cdot z/\hbar) \times \mathrm{Tr}\,\rho(0)\Gamma^\dagger(r + \tfrac{1}{2}z, t)\Gamma(r - \tfrac{1}{2}z, t), \quad (81)$$

where

$$\Gamma(r, t) = \psi(\mathbf{r}_N, t)\cdots\psi(\mathbf{r}_1, t). \quad (82)$$

Let us assume, for simplicity, that

$$H_w = p^2/2m + \Phi(r), \quad (83a)$$

$$\Phi(r) = \sum_{i<j}\phi(r_i - r_j). \quad (83b)$$

(We also assume that the interaction potential is real.) One then has

$$i\hbar\partial\psi(\mathbf{r}, t)/\partial t = -(\hbar^2/2m)\nabla_r^2\psi(\mathbf{r}, t) + \hat{\phi}(\mathbf{r})\psi(\mathbf{r}, t), \quad (84)$$

where

$$\hat{\phi}(\mathbf{r}) = \int d^3r'\phi(\mathbf{r} - \mathbf{r}')\psi^\dagger(\mathbf{r}')\psi(\mathbf{r}'). \quad (85)$$

Thus, $\Gamma(r, t)$ obeys the equation (with $\nabla_r^2 = \sum_i \nabla_{\mathbf{r}_i}^2$),

$$i\hbar\partial\Gamma(r, t)/\partial t = -(\hbar^2/2m)\nabla_r^2\Gamma(r, t) + \Phi(r)\Gamma(r, t) + \sum_{j=1}^N \hat{\phi}(\mathbf{r}_j)\Gamma(r, t)\theta^{N-j}, \quad (86)$$

where we have successively used the (anti)commutator relation

$$[\psi(\mathbf{r}), \hat{\phi}(\mathbf{r}')]_\theta = \psi(\mathbf{r})\phi(\mathbf{r}' - \mathbf{r}). \quad (87)$$

The last term in the right-hand side of Eq. (86) vanishes identically since there is no $(N+1)$-particle state. One readily finds

$$i\hbar\frac{\partial}{\partial t}\Gamma^\dagger(r + \tfrac{1}{2}z, t)\Gamma(r - \tfrac{1}{2}z, t) = [\hbar^2/2m\nabla_z\cdot\nabla_r - 2\Phi(r)\sinh\tfrac{1}{2}\overleftarrow{\nabla}_r\cdot z] \times \Gamma^\dagger\left(r + \frac{z}{2}, t\right)\Gamma\left(r - \frac{z}{2}, t\right). \quad (88)$$

Equation (88) enables one to calculate the time rate of change of f_w^θ to obtain

$$\left(\frac{\partial}{\partial t} + \frac{1}{m}p\cdot\nabla_r - \frac{2}{\hbar}\phi(r)\sin\frac{\hbar}{2}\overleftarrow{\nabla}_r\cdot\overrightarrow{\nabla}_p\right) \times f_w^\theta(r, p, t) = 0, \quad (89)$$

which is nothing but the quantum Liouville equation [cf. Eq. (30)].

In kinetic theory, one introduces the reduced distribution functions by integrating f_w^θ with respect to all state variables $\mathbf{r}$, $\mathbf{p}$ but those belonging to one,

[14] For a general review, see W. E. Brittin and W. R. Chappell, Rev. Mod. Phys. 34, 620 (1962); also R. Balescu, *Statistical Mechanics of Charged Particles* (Interscience Publishers, Inc., New York, 1963), Part II.

two, etc., particles. Let us define with $i \triangleq (\mathbf{r}_i, \mathbf{p}_i)$,

$$f_s^0(1, \cdots, s; t) = \frac{N!}{(N-s)!} \times \int d^3r_{s+1} \cdots d^3r_N \, d^3p_{s+1} \cdots d^3p_N f_w^0,$$

where we have used the "generic" normalization. By integrating Eq. (89) one obtains the quantum BBGKY hierarchy:

$$\left(\frac{\partial}{\partial t} + \frac{1}{m} p \cdot \nabla_r - \frac{2}{\hbar} \sum_{i<j\leq s} \phi(\mathbf{r}_i - \mathbf{r}_j) \sin \tfrac{1}{2}\hbar \overleftarrow{\nabla}_r \cdot \overrightarrow{\nabla}_p \right) \times f_s^0(1, \cdots, s; t) = \frac{2}{\hbar} \sum_{i=1}^{s} \int d^3r_{s+1} \, d^3p_{s+1} \phi(\mathbf{r}_i - \mathbf{r}_{s+1}) \times \sin (\tfrac{1}{2}\hbar \overleftarrow{\nabla}_r \cdot \overrightarrow{\nabla}_p) f_{s+1}^0(1, \cdots, s+1; t). \quad (90)$$

The right-hand side of Eq. (90) stems from the contribution of the last term in Eq. (86) which does not vanish when $s < N$.

Equation (90) [rather the first one (or two) of the chain] has been studied for various systems in which the radiation field, photon field, spin and relativistic effects are included.[15] For further details, the reader is referred to Ref. 15.

ACKNOWLEDGMENTS

Two of the authors (M. R. and P. F. Z.) gratefully acknowledge many stimulating discussions with Professor R. K. Osborn and Professor F. C. Shure during the early stages of this work. In addition, P. F. Z. wishes to express his gratitude to the Ford Foundation for financial support and to the Çekmece Nuclear Center for its hospitality during the time this paper was being written.

APPENDIX A

Solution of Bloch Equation

Although Eqs. (37) can be formally solved to obtain

$$\Omega_w = \exp(-\beta H_w \cos \hbar\Lambda/2), \quad (A1)$$

the evaluation of this result is somewhat lengthy. Instead, here, for simplicity, we restrict ourselves to the case for which

$$H_w = p^2/2m + \Phi(r), \quad (A2)$$

where $\Phi(r)$ is the interaction potential which is usually assumed to be pairwise additive; i.e.,

$$\Phi(r) = \sum_{i<j} \phi(|\mathbf{r}_i - \mathbf{r}_j|). \quad (A3)$$

[15] R. K. Osborn and E. H. Klevans, Ann. Phys. (N.Y.) 15, 105 (1961); E. Özizmir, Doctoral thesis, University of Michigan (1962); R. K. Osborn, Phys. Rev. 130, 2142 (1963).

The Bloch equation for this case can be written as

$$\partial\chi/\partial\beta = (\hbar^2/8m) e^{\beta\phi} \nabla_r^2 e^{-\beta\phi} \chi + \phi e^{\beta p^2/2m} \times (1 - \cos \tfrac{1}{2}\hbar \overleftarrow{\nabla}_r \cdot \overrightarrow{\nabla}_p) e^{-\beta p^2/2m}, \quad (A4)$$

where we have set

$$\chi = e^{\beta H^w} \Omega_w. \quad (A5)$$

The first term in the right-hand side of (4) can be written as

$$(\hbar^2/8m)[\nabla_r^2 - 2\beta \nabla_r\phi \cdot \nabla_r - \beta \nabla_r^2 \phi + \beta^2 (\nabla_r \phi)^2]\chi.$$

We also have

$$\begin{aligned} &\exp(\beta p^2/2m) \cos(\tfrac{1}{2}\hbar) \overleftarrow{\nabla}_r \cdot \overrightarrow{\nabla}_p \exp(-\beta p^2/2m) \\ &= \exp(\beta p^2/2m) \operatorname{Re} \exp(\tfrac{1}{2} i\hbar \overleftarrow{\nabla}_r \cdot \overrightarrow{\nabla}_p) \exp(-\beta p^2/2m) \\ &= \exp(\beta p^2/2m) \operatorname{Re} \exp[-\beta(p + \tfrac{1}{2} i\hbar \overleftarrow{\nabla}_r)/2m] \times \exp(\tfrac{1}{2} i\hbar \overleftarrow{\nabla}_r \cdot \overrightarrow{\nabla}_p) \\ &= \exp(\beta\hbar^2 \overleftarrow{\nabla}_r^2/8m) \operatorname{Re} \exp(-i\beta\hbar \overleftarrow{\nabla}_r \cdot p/2m) \times \exp(i\hbar \overleftarrow{\nabla}_r \cdot \overrightarrow{\nabla}_p/2). \end{aligned} \quad (A6)$$

Hence, one obtains, with $\chi(\beta = 0) = 1$,

$$\begin{aligned} \partial\chi/\partial\beta = (\hbar^2/8m)[&\nabla_r^2 \chi - 2\beta \nabla_r\phi \cdot \nabla_r \chi \\ &- \beta(\nabla_r^2 \phi)\chi + \beta^2 (\nabla_r \phi)^2 \chi] + \phi\chi \\ &- [\exp(\hbar^2 \beta \nabla_r^2/8m)\phi][\cos(\hbar\beta/2m) \overleftarrow{\nabla}_r \cdot p \cos(\tfrac{1}{2}\hbar) \overleftarrow{\nabla}_r \cdot \overrightarrow{\nabla}_p + \sin(\hbar\beta/2m) \overleftarrow{\nabla}_r \cdot p \sin(\tfrac{1}{2}\hbar) \overleftarrow{\nabla}_r \cdot \overrightarrow{\nabla}_p]\chi. \end{aligned} \quad (A7)$$

This equation can be solved by expanding χ in powers of $\hbar^2$. To the lowest order, $\chi_0 = 1$, which corresponds to the classical case. To first order, one gets

$$\chi_1 = (8m)^{-1}[-\beta^2 \nabla^2 \phi + \tfrac{1}{3}\beta^3 (p \cdot \nabla)^2 \phi], \quad (A8)$$

which was first derived by Wigner.[2] The second-order contribution can also be carried out although the calculation is somewhat lengthy. The result is

$$\begin{aligned} \chi_2 = (\beta^3/64m^2)\{&-\tfrac{1}{2}\nabla^4\phi + \beta[\tfrac{1}{6}\nabla^2(\nabla\phi)^2 \\ &+ (1/3m)(p \cdot \nabla)^2 \nabla^2 \phi + \tfrac{1}{3}\nabla\phi \cdot \nabla(\nabla^2\phi) + \tfrac{1}{2}(\nabla^2\phi)^2] \\ &- \beta^2[\tfrac{2}{15}\nabla\phi \cdot \nabla(\nabla\phi)^2 + (2/15m)\nabla\phi \cdot \nabla(p \cdot \nabla)^2 \phi \\ &+ \tfrac{1}{3}(\nabla^2\phi)(\nabla\phi)^2 + (1/3m)(\nabla^2\phi)(p \cdot \nabla)^2 \phi \\ &+ (4/15m)(\nabla(p \cdot \nabla\phi))^2 + \tfrac{1}{30}(p \cdot \nabla)^4 \phi] \\ &+ \beta^3[\tfrac{1}{15}(\nabla\phi)^4 + (1/9m)(\nabla\phi)^2 (p \cdot \nabla)^2 \phi \\ &+ (1/6m^2)((p \cdot \nabla)^2 \phi)^2]\}. \end{aligned} \quad (A9)$$

In order to calculate the partition function Z we consider the relation

$$\int dp \, dr \Omega_w = Z(2\pi\hbar)^{3N} \quad (A10)$$

or

$$Z = (2\pi\hbar)^{-3N} \langle\chi\rangle_C;$$

one obtains

$$Z = Z_C[1 + \hbar^2 C_1 + \hbar^4 C_2 + \cdots], \quad \text{(A11)}$$

where Z_C is the classical partition function. One has

$$C_1 = (1/8m)[-\beta^2\langle\nabla^2\phi\rangle_C + \tfrac{1}{3}\beta^3\langle(\nabla\phi)^2\rangle_C + \tfrac{1}{3}\beta^2\langle\nabla^2\phi\rangle_C] = -(\beta^2/24m)\langle\nabla^2\phi\rangle_C. \quad \text{(A12)}$$

and

$$C_2 = (\beta^3/4m^2)\Big[\frac{\beta}{288}\langle(\nabla\phi)^4\rangle_C - \frac{\beta^2}{72}\langle(\nabla^2\phi)(\nabla\phi)^2\rangle_C - \frac{\beta^2}{120}\langle\nabla\phi\cdot\nabla(\nabla\phi)^2\rangle_C + \frac{\beta}{72}\langle(\nabla^2\phi)^2\rangle_C + \frac{\beta}{45}\langle\nabla\phi\cdot\nabla(\nabla^2\phi)\rangle_C + \frac{\beta}{180}\langle\nabla^2(\nabla\phi)^2\rangle_C - \frac{1}{60}\langle\nabla^4\phi\rangle_C\Big],$$
$$= \frac{\beta^3}{144m^2}\Big[\frac{\beta}{8}\langle(\nabla^2\phi)^2\rangle_C + \frac{\beta}{40}\langle\nabla^2(\nabla\phi)^2\rangle_C - \frac{7}{40}\langle\nabla^4\phi\rangle_C\Big]. \quad \text{(A13)}$$

The evaluation of C_2 has been carried out by Goldberger and Adams[16] using field theoretic methods. Our result, however, differs by a factor of $\frac{7}{8}$ from theirs in the middle term of the last equation.

The unsatisfactory part of the above analysis is that these coefficients blow up the limit $N \to \infty$. This can be seen by writing

$$\langle\nabla^2\phi\rangle_C = N\rho^{-1}\int d^3r n_2^C(r, 0)\nabla^2\phi(r), \quad \text{(A14)}$$

where $n_2^C(\mathbf{r}_1, \mathbf{r}_2) = n_2^C(\mathbf{r}_1 - \mathbf{r}_2, 0)$ is the classical reduced distribution function in configuration space defined as in Eq. (60), and $\rho = N/V$ (= finite). This difficulty was first observed by Mayer and Band.[17]

In order to remedy this difficulty, one can use the well-known property that the free energy per particle F/N is a finite quantity in this limit, as was shown by Kahn and Uhlenbeck[18] in general terms. In other words, $Z^{1/N}$ remains finite as $N \to \infty$ [cf. (A15)]. This can be observed to be true for the case of the harmonic oscillator; cf. Eq. (B10). Green[19] demonstrated that if one expands F/N, instead of Z, to which it is related through the relation

$$Z = e^{-\beta F}, \quad \text{(A15)}$$

then the contribution to order $\hbar^2$ remains finite as $N \to \infty$. We now show that this is also the case for the term of order $\hbar^4$.

To do this, we consider

$$(Z/Z_C)^{1/N} = (1 + \hbar^2 C_1 + \hbar^4 C_2 + \cdots)^{1/N} = 1 + \hbar^2(1/N)C_1 + \hbar^4[(1/N)C_2 + (1/2N)(1/N - 1)C_1^2] + O(\hbar^6). \quad \text{(A16)}$$

We have already seen that

$$(1/N)C_1 = -(\beta^2/24m\rho)\int d^3r n_2^C(\mathbf{r}, 0)\nabla^2\phi(r), \quad \text{(A17)}$$

which is bounded as $N \to \infty$, where $r = |\mathbf{r}|$ (not to be confused with $3N$-dimensional vector r).

We can write from (A13) (by making use of a vector identity and an integration by parts)

$$(1/N)C_2 = (\beta^3/1152m^2N) \times \Big[\beta\langle(\nabla^2\phi)^2\rangle_C + \frac{2\beta}{5}\langle\phi(\overleftarrow{\nabla}\cdot\overrightarrow{\nabla})^2\phi\rangle_C - \langle\nabla^4\phi\rangle_C\Big]. \quad \text{(A18)}$$

It is readily seen that

$$\frac{1}{N}\langle\nabla^4\phi\rangle_C = \frac{2}{\rho}\int d^3r n_2^C(r, 0)\nabla^4\phi(r) \quad \text{(A19)}$$

and

$$\frac{1}{N}\langle\phi(\nabla\cdot\nabla)\rangle_C = \frac{2}{\rho}\int d^3r n_2^C(r, 0)[\phi(r)(\overleftarrow{\nabla}_r\cdot\overrightarrow{\nabla}_r)^2\phi(r)] + \frac{1}{\rho}\int d^3r\, d^3r' n_3^C(\mathbf{r}, \mathbf{r}', 0)[\phi(r)(\overleftarrow{\nabla}_r\cdot\overrightarrow{\nabla}_{r'})^2\phi(r')]. \quad \text{(A20)}$$

Therefore these quantities are both finite. Finally, we calculate

$$\frac{1}{N}\langle(\nabla^2\phi)^2\rangle_C = \frac{2}{\rho}\int d^3r n_2^C(r, 0)(\nabla^2\phi(r))^2 + \frac{4}{\rho}\int d^3r\, d^3r' n_3^C(\mathbf{r}, \mathbf{r}', 0)\nabla^2\phi(r)\nabla^2\phi(r') + \frac{1}{\rho}\int d^3r\, d^3r'\, d^3r'' n_4^C(\mathbf{r}, \mathbf{r}' + \mathbf{r}'', \mathbf{r}''', 0)\nabla^2\phi(r)\nabla^2\phi(r'). \quad \text{(A21)}$$

The last term in this equation contains a part which is proportional to N. This part, in fact, compensates the term $-(1/2N)C_1^2$ which also blows up as $N \to \infty$.

To see this, we introduce the cluster development[20]

[16] M. L. Goldberger and E. N. Adams, II, J. Chem. Phys. **20**, 240 (1952).
[17] J. E. Mayer and W. Band, J. Chem. Phys. **15**, 141 (1947).
[18] B. Kahn and G. E. Uhlenbeck, Physica **5**, 399 (1938).
[19] H. S. Green, J. Chem. Phys. **19**, 955 (1951).
[20] T. R. Hill, *Statistical Mechanics* (McGraw-Hill Book Company, Inc., New York, 1956).

as (by dropping the superscript C)

$$\begin{aligned} n_1 &= \rho = G_1, \\ n_2(\mathbf{r}_1, \mathbf{r}_2) &= G_2(\mathbf{r}_1, \mathbf{r}_2) + \rho^2, \\ n_3(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) &= G_3(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) + \rho[G_2(\mathbf{r}_1, \mathbf{r}_2) \\ &\quad + \text{cyclic terms}] + \rho^3, \quad \text{(A22)} \end{aligned}$$

etc., where G_s is the s-particle correlation function which vanishes as $|\mathbf{r}| \to \infty$. The last term in the right-hand side of (A21) becomes, after some manipulations,

$$\begin{aligned} &= \frac{1}{\rho}\int d^3r\, d^3r'\, d^3r'' G_4(\mathbf{r}, \mathbf{r}' + \mathbf{r}'', \mathbf{r}'', 0)\nabla^2\phi(r)\nabla^2\phi(r') \\ &\quad + \frac{2}{\rho}\int d^3r\, d^3r'\, d^3r'' G_2(\mathbf{r} - \mathbf{r}' - \mathbf{r}'', 0) \\ &\qquad \times G_2(\mathbf{r}'', 0)\nabla^2\phi(r)\nabla^2\phi(r') \\ &\quad + \frac{V}{\rho}\left[\int d^3r n_2^C(\mathbf{r}, 0)\nabla^2\phi(r)\right]^2. \quad \text{(A23)} \end{aligned}$$

The last term is the contribution of the part of the double-pair correlations which divides $n_4(\mathbf{r}, \mathbf{r}' + \mathbf{r}'', \mathbf{r}' + \mathbf{r}'', \mathbf{r}'', 0)$ as

$$G_2(\mathbf{r}, 0)G_2(\mathbf{r}' + \mathbf{r}'', \mathbf{r}'') = G_2(r, 0)G_2(r', 0).$$

This term blows up linearly with N. When (A23) is substituted in (A16) one finds that this term cancels the term $C_1^2/2N$; therefore the coefficient of $\hbar^4$ remains finite.

The free energy, F, can then be written as

$$\begin{aligned} F &= F_C - (N/\beta)\ln(Z/Z_C)^{1/N} \\ &= F_C - (\hbar^2/\beta)C_1 - (\hbar^4/\beta) \\ &\qquad \times [(C_2 - \tfrac{1}{2}C_1^2) - (1/N)C_1^2] + O(\hbar^6), \quad \text{(A24)} \end{aligned}$$

a result which may be used to calculate the quantum corrections to the equation of state.

To do this, we consider the definition of pressure[20]

$$\begin{aligned} p &= -\partial F/\partial V \\ &= p^C - \frac{\hbar^2}{\beta}\frac{\partial C_1^2}{\partial V} - \frac{\hbar^4}{\beta}\frac{\partial}{\partial V} \\ &\qquad \times \left[(C_2 - \tfrac{1}{2}C_1^2) - \frac{1}{N}C_1^2\right] + O(\hbar^6), \quad \text{(A25)} \end{aligned}$$

where p^C is the classical pressure which is to be calculated from the classical equation of state.

To evaluate the volume differentiations, one can consider the spatial integrations to be carried out in a box of length $V^{\frac{1}{3}}$. By changing the dummy variable $\mathbf{r} = \mathbf{x}V^{\frac{1}{3}}$, so that the volume of integration becomes unity, the differentiation can be performed on the integrand.[20] In general, one has, for an integrable function $h(\mathbf{r}_1, \cdots, \mathbf{r}_s, V)$,

$$\begin{aligned} &V\frac{\partial}{\partial V}\int d^3r_1 \cdots d^3r_s h(\mathbf{r}_1, \cdots, \mathbf{r}_s, V) \\ &= (s-1)\int d^3r_1 \cdots d^3r_s h(\mathbf{r}_1, \cdots, \mathbf{r}_s, V) \\ &\quad + \int d^3r_1 \cdots d^3r_s V\frac{\partial}{\partial V}h(\mathbf{r}_1, \cdots, \mathbf{r}_s, V) \\ &\quad + \tfrac{1}{3}\int d^3r_1 \cdots d^3r_s \sum_{i=1}^{s}\nabla_{\mathbf{r}_i}\cdot(\mathbf{r}_i \cdot h(\mathbf{r}_1, \cdots, \mathbf{r}_s, V)). \\ &\qquad \text{(A26)} \end{aligned}$$

The last term vanishes if h vanishes sufficiently rapidly for large $|\mathbf{r}_1|$.

Here for simplicity, we consider dilute systems; that is we ignore correlations involving more than two particles. We thus obtain

$$p = p^C + \hbar^2 A_1 + \hbar^4 A_2 + O(\hbar^6), \quad \text{(A25a)}$$

where

$$\begin{aligned} A_1 &= (\beta\rho^2/24m)\int d^3r g(r)\nabla^2\phi(r) \\ A_2 &= -(\beta^2\rho^2/1152m^2)\Big\{2\beta\int d^3r g(r)(\nabla^2\phi)^2 \\ &\quad + 4\beta\rho\int d^3r\, d^3r' g(|\mathbf{r} - \mathbf{r}'|)\nabla^2\phi(r)\nabla^2\phi(r') \\ &\quad + 6\beta\rho^2\int d^3r\, d^3r'\, d^3r'' g(|\mathbf{r} - \mathbf{r}' - \mathbf{r}''|) \\ &\quad \times g(r'')\nabla^2\phi(r)\nabla^2\phi(r') + \tfrac{4}{5}\beta\int d^3r g(r)[\phi(\overleftarrow{\nabla}\cdot\overrightarrow{\nabla})^2\phi] \\ &\quad + \tfrac{2}{5}\beta\rho\int d^3r\, d^3r' g(|\mathbf{r} - \mathbf{r}'|)[\phi(r)(\overleftarrow{\nabla}_{\mathbf{r}}\cdot\overrightarrow{\nabla}_{\mathbf{r}'})^2\phi(r')] \\ &\quad - \tfrac{4}{5}\beta\int d^3r g(r)\nabla^4\phi - 4\rho\left[\int d^3r g(r)\nabla^2\phi\right]^2\Big\}. \quad \text{(A27)} \end{aligned}$$

Here $g(r)$ is the radial distribution function defined as $\rho^2 g(r) = n_2(\mathbf{r}, 0)$. (We have also dropped the contribution of the volume differentiation of correlations, which may be small for dilute systems.)

Our result, to the order $\hbar^2$, coincides with those given by Uhlenbeck and Beth[21] and also by Green[19] if one further assumes that

$$g(r) = \exp[-\beta\phi(r)],$$

a relation which is valid in the low-density limit.[20] In the latter case, it is possible to simplify further the quantities in (A27).

[21] G. E. Uhlenbeck and E. Beth, Physica **3**, 729 (1936); **4**, 915 (1937). For an excellent review of this subject, see J. de Boer, Rept. Progr. Phys. **12**, 305 (1949).

APPENDIX B

Harmonic Oscillator

For the harmonic oscillator, which is characterized by the Hamiltonian

$$H_w = p^2/2m + \tfrac{1}{2}m\omega^2 r^2, \tag{B1}$$

the Bloch equation (37) reduces to

$$\partial\Omega_w/\partial\beta = -H_w\Omega_w + (\hbar^2/8m) \times [\nabla_r^2 + m^2\omega^2\nabla_p^2]\Omega_w. \tag{B2}$$

We seek a solution of (B2) of the form

$$\Omega_w = \exp[-A(\beta)H_w + B(\beta)], \tag{B3}$$

which yields the following equation:

$$-(dA/d\beta)H_w + (dB/d\beta) = -H_w - (\tfrac{1}{4}\hbar^2)\omega^2 A + (\tfrac{1}{4}\hbar^2)\omega^2 H_w. \tag{B4}$$

Since the only (r, p) dependence in (B4) occurs in H_w, its coefficient must vanish separately. Thus, we obtain

$$A = (2/\hbar\omega)\tanh \hbar\omega\beta/2, \tag{B5}$$

$$B = -\ln\cosh \hbar\omega\beta/2, \tag{B6}$$

where the initial condition $\Omega(\beta = 0) = 1$ has been used.

The partition function can now be calculated by using Eq. (A10), which yields

$$Z = (1/\hbar\omega A)^{3N}[1/\cosh(\hbar\omega\beta/2)], \tag{B7}$$

and, therefore

$$f_w = (1/2\pi\hbar)^{3N}(\tfrac{1}{2}\Omega_w) = (\omega A/2\pi)^{3N}e^{-A(\beta)H_w}. \tag{B8}$$

Let us note that the free energy per particle is bounded as $N \to \infty$, since

$$(F/N) = -(1/\beta)\ln Z^{(1/N)} = (1/\beta)\ln(\hbar\omega A)[\cosh(\tfrac{1}{2}\hbar\omega\beta)]^{1/N} \tag{B9}$$

as was pointed out in Appendix A. One obtains

$$\lim_{N\to\infty}(F/N) = (1/\beta)\ln[2\tanh(\tfrac{1}{2}\hbar\omega\beta)]. \tag{B10}$$

Symmetries of the Bethe–Salpeter Equation for Relativistic Bound-State Problem

S. N. BISWAS*
Department of Physics, University of California, Los Angeles, California

(Received 3 October 1966)

Possible dynamical symmetries of the Bethe–Salpeter equation in the ladder approximation for the bound state of two neutral scalar mesons interacting via a massless scalar boson have been investigated. It is shown that the Bethe–Salpeter equation exhibits an $O(5)$ or $O(4)$ symmetry according as E, the total energy of the system, is zero or nonvanishing and the corresponding no-invariance group being $O(5, 1)$ when $E = 0$ and the $O(5, 1)$ spectrum splits into $\sum \oplus O(4, 1)$ when $E \neq 0$ as pointed out by Salam *et al.* The method employed here differs from that of Cutkosky and of Salam *et al.*

1. INTRODUCTION

RECENTLY interest[1] in the study of the relativistic bound-state problem using the Bethe–Salpeter equation[2] has been revived in connection with the dynamical origin of higher symmetries. Such considerations have been thoroughly discussed[3] with respect to the nonrelativistic Schrödinger equation for the hydrogen atom problem and in the strong coupling limit of the static meson field theory.[4] The existence of an $O(4)$ symmetry for the hydrogen atom problem is well known.[5] In this note we consider the relativistic bound state problem in the Bethe–Salpeter (B–S) equation. In its simplest form we consider the equation for two neutral scalar mesons of equal mass m interacting via a massless scalar boson; the interaction Hamiltonian being of the form $g[\phi(x)]^2A(x)$, where g is the interaction strength, ϕ the meson of mass m, and $A(x)$ is the massless scalar field. Only the ladder diagrams are kept as usual. Applying a stereographic projection to the five-dimensional pseudosphere it was pointed out by Cutkosky[6] and later by Salam *et al.*[1] that the B–S equation for such an interacting system possesses an $O(5)$ symmetry when the total energy of the system is zero and an $O(4)$ symmetry for nonzero total energy. The purpose of this note is to obtain the same result by noting that the B–S equation can be separated in a suitable system of coordinates.[7] This immediately leads to the required degeneracy of the energy spectrum consistent with $O(5)$ and $O(4)$ symmetry of the relativistic B–S equation for zero and nonzero total energy, respectively.

2. ENERGY SPECTRUM AND SYMMETRIES OF THE B–S EQUATION

The B–S wavefunction, $X(x, y)$ for the two-body scalar problem is given by

$$X(x, y) = \langle 0| T(\phi(x)\phi(y)\,|k\rangle,$$

where k denotes half the total four-momentum of the system. We write the wavefunction as the product of two terms describing the center-of-mass motion and relative motion of the two particles, namely,

$$X(x, y) = e^{ik(x+y)}\phi(x - y),$$

where ϕ is the wavefunction in relative coordinates. The B–S equation in momentum space and in the ladder approximation is given by

$$\{(p_4 + E)^2 + \mathbf{p}^2 + m^2\} \times \{(p_4 - E)^2 + \mathbf{p}^2 + m^2\}\phi(p) = \psi(p),$$

$$\psi(p) = \pi^{-2}\lambda\int (p - q)^{-2}\phi(q)\, d^4q, \quad (1)$$

where $\phi(p)$ is the Fourier transform of $\phi(x - y)$. In obtaining Eq. (1) we have performed the Wick rotation in the relative time variable, i.e., we have replaced p_0 by $p_4 = ip_0$, and further specialized to the rest frame and used $k = (\mathbf{0}, iE)$. λ denotes the interaction strength ($\lambda = g^2/8\pi^2$).

Using the fact[8] that

$$\Box_p\left(\frac{1}{p - q}\right)^2 = \left(\frac{\partial^2}{\partial p_4^2} + \frac{\partial^2}{\partial \mathbf{p}^2}\right)\left(\frac{1}{p - q}\right)^2 = -4\pi^2\delta^4(p - q)$$

we obtain from (1) the following differential equation:

$$(\partial^2/\partial p_4^2 + \partial^2/\partial \mathbf{p}^2)\psi(p) = -4\lambda\phi(p). \quad (2)$$

* Permanent address: Centre for Advanced Study in Physics, Department of Physics, Delhi University, Delhi, India.

[1] R. Delbourge, Abdus Salam, and J. Strathedre, ICTP preprint No. IC/66/60; A. O. Barut, P. Budini, and C. Fronsdal, ICTP preprint No. IC/65/34.

[2] E. E. Salpeter and H. A. Bethe, Phys. Rev. **84**, 1232 (1951). For solutions, see G. C. Wick, *ibid.* **96**, 1124 (1954); R. E. Cutkosky, *ibid.* **96**, 1135 (1954); J. S. Goldstein, *ibid.* **91**, 1516 (1953); H. S. Green, *ibid.* **97**, 540 (1955); S. Okubo and A. Feldman, *ibid.* **117**, 279 (1960).

[3] N. Mukanda, L. O'Raifeartaigh, and E. C. G. Sudarshan, Phys. Rev. Letters **15**, 1041 (1965). Barut, Budini, Fronsdal (see Ref. 1).

[4] T. Cook, C. J. Goebel, and B. Sakita, Phys. Rev. Letters **15**, 35 (1965).

[5] V. Fock, Z. Physik, **98**, 145 (1935).

[6] See Ref. 2.

[7] H. S. Green, Nuovo Cimento **5**, 580 (1957); Author's Ph.D. thesis, Adelaide University (1957); S. N. Biswas, *ibid.* **8**, 540 (1958).

[8] H. S. Green, Phys. Rev. **97**, 540 (1955).

This differential equation is equivalent to the integral equation (1) provided $p^2\psi(p)$ is finite at large p and also $\psi(p)$ is finite at small p.

We now use the following set of transformation variables: $p_1 = p_s \sin\theta\cos\phi, p_2 = p_s \sin\theta \sin\phi, p_3 = p_s\cos\theta, p_s = \omega \sin\beta/(\cosh\xi - \cos\beta); p_4 = \omega \sinh\xi/(\cosh\xi - \cos\beta)$. The whole energy-momentum plane is contained in the ranges $0 \leq \theta \leq \pi, 0 \leq \phi \leq 2\pi$, and $-\infty \leq \xi \leq \beta$ and $\omega^2 = E^2 + m^2$. In this coordinate system we write

$$p_s\psi = f(\xi)g(\beta)Y_l^m(\theta, \phi),$$

where $Y_l^m(\theta, \phi)$ are the usual spherical harmonics. The differential equation (2) becomes

$$\left\{\frac{\partial^2}{\partial\xi^2} + \frac{\partial^2}{\partial\beta^2} - \frac{l(l+1)}{\sin^2\beta}\right\}(p_s\psi) = \lambda\frac{(p_s\psi)}{E^2\sinh^2\xi - \omega^2\cosh^2\xi}. \quad (3)$$

Separating the variables of the equation we have

$$(d^2g/d\beta^2) + \{n^2 - [l(l+1)/\sin^2\beta]g\} = 0 \quad (4)$$

and

$$(d^2f/d\xi^2) - \{n^2 - [\lambda/(\omega^2\cosh^2\xi - E^2\sinh^2\xi)]\}f = 0. \quad (5)$$

The boundedness conditions on $g(\beta)$ and $f(\xi)$ in the new coordinates are obtained by substituting them in the integral equation which is satisfied if $g(0) = g(\pi)$ and $f(\xi)$ to vanish at large ξ and f is an even function of ξ. Further f and g must be finite within the given ranges.

The solution of Eq. (4) is

$$g(\beta) = \text{const} \times \sin^{l+1}\beta(\partial/\partial\cos\beta)^l \times [\sin(n\beta + \delta)/\sin\beta].$$

As g must be finite at $\beta = 0$ then δ must vanish, the requirement $g(0) = g(\pi)$ demands that n must be an integer. Since $g \neq 0$ we must have $n \geq l + 1$. This solution can be written conveniently in terms of Gegenbauer polynomials,

$$g(\beta) = \text{const} \times \sin^{l+1}\beta C_{n-l-1}^{l+1}(\cos\beta). \quad (6)$$

The differential equation (5) for $f(\xi)$ determines the energy eigenvalues of the bound state. Here, as was mentioned at the beginning, we consider the case when total energy $E = 0$. The equation can then be solved exactly. Putting

$$\lambda/m^2 = N(N+1),$$

where N is an integer, Eq. (5) reduces to

$$d^2f/d\xi^2 = \{n^2 - [N(N+1)/\cosh^2\xi]\}f. \quad (7)$$

To solve Eq. (8) we put $\tanh\xi = \cos\chi$; χ then varies from 0 to π as $\xi \to \infty$ to $-\infty$. The solution satisfying the boundedness condition can be put as

$$f \simeq (-)^n\frac{(N+n)!}{n!\,(N-n)!}\sin^n\chi\, {}_2F_1 \times [n - N, N + n + 1, n + 1, \sin^2(\tfrac{1}{2}\chi)], \quad (8)$$

which are the usual Gegenbauer polynomials

$$f_{N(n)} = \text{const} \times (-)^n \sin^n\chi C_{N-n}^{n+\frac{1}{2}}(\cos\chi). \quad (9)$$

This is the same solution as given by Cutkosky (Appendix A of Ref. 2) and Okubo and Feldman[9] for the S-wave case when we put $n = 1$ in our solution (9). Here again for $f_{N(n)} \neq 0$ we get $N \geq n \geq l + 1$, $l > |m|$. The energy eigenfunctions $f_{N(n)}$ have then the degeneracy $\frac{1}{6}N(N+1)(2N+1)$. We now combine the solutions for $f(\chi)$, $g(\beta)$ and defining a new wavefunction Φ through[10]

$$\Phi = (1 - \cos\beta\sin\chi)^{-3}\phi,$$

where ϕ is the wavefunction occurring in (1) we obtain the eigenfunctions Φ_N's corresponding to the eigenvalue $\lambda_n = N(N+1)$ to be

$$\Phi_N \sim Y_{Nnlm} \quad (10)$$

apart from a constant.

The Y_{Nnlm} are the spherical harmonics in the 5-dim Euclidean space and are given by

$$Y_{Nnlm} = (\sin\chi)^{n-1}C_{N-n}^{n+\frac{1}{2}}(\cos\chi)\sin^l\beta C_{n-l-1}^{l+1}(\cos\beta) \times \sin^m\theta C_{l-m}^{m+\frac{1}{2}}(\cos\theta)e^{im\phi}, \quad (11)$$

with $N > n > l + 1$, $l > |m|$ and the degeneracy of Y_{Nnlm} is $\frac{1}{6}N(N+1)(2N+1)$. Hence the number of independent Φ_N's is $\frac{1}{6}N(N+1)(2N+1)$ which correspond to the component of the fully symmetric irreducible traceless tensors of the $O(5)$ group. We note down here in particular those generators of the group which will change l by one unit keeping n fixed and change n by one unit keeping N fixed. We have

$$N_l^+ g_{n(l)} = [\partial/\partial\beta - (l+1)\cot\beta]g_{n(l)} = g_{n(l+1)},$$
$$N_l^- g_{n(l+1)} = [\partial/\partial\beta + (l+1)\cot\beta]g_{n(l+1)} = (n - l - 1)(n + l + 1)g_{n(l)},$$

and

$$N_n^+ f_{N(n)} = (\partial/\partial\chi - n\cot\chi)f_{N(n)} = (2n+1)f_{N(n+1)},$$
$$N_n^- f_{N(n+1)} = [\partial/\partial\chi + (n+1)\cot\chi]f_{N(n+1)} = [(N-n)(N+n+1)/(2n+1)]f_{N(n)}.$$

[9] See Ref. 2.

[10] This definition of the new function corresponds to that introduced by Salam *et al.* [See Ref. 1; Eq. (4).]

If in addition one puts the infinite B–S energy levels characterized by $\lambda_1 \cdots \lambda_\infty$ in a single irreducible representation of a noncompact group, then the resulting group is $O(5, 1)$ having $O(5)$ as the maximum compact subgroup. $O(5, 1)$ is then the so-called noninvariance[11] group of the B–S equation.

3. OTHER REMARKS

In the case when $E \neq 0$, the differential equation (5) cannot be solved exactly at special limits. It is interesting to note that our Eq. (5) is equivalent to Eq. (15) of Ref. 1 (see Salam *et al.*) if we make the following substitution. Putting $f(\xi) = (\text{sech } \xi)^n$. $G(\xi)$ and finally writing $z = \tanh \xi$ we find the equation for G as follows:

$$(1 - z^2)\frac{d^2G}{dz^2} + 2(n-1)z\frac{dG}{dz} + \left\{\frac{(\lambda/m^2)}{(1 - \epsilon^2 + \epsilon^2 z^2)} - n(n-1)\right\}G = 0, \quad (12)$$

where ϵ is the rest mass of the bound system in units of m. This equation (12) has been studied by Salam *et al.*[1] and earlier by Cutkosky.[2] We have then obtained the final equation for symmetry studies in a much simplified manner without using the generalized stereographic projection method of Fock[5] as utilized by Salam *et al.*[1] and previously by Cutkosky.[2] However, following the same argument as that of Salam *et al.*[1] one then finds that the symmetry group of the B–S equation in this nonvanishing energy limit is $O(4)$ and the $E = 0$, $O(5, 1)$ spectrum splits into $\sum \oplus (O(4, 1))$ spectrum when $E \neq 0$, as in this case the energy levels are characterized by an extra quantum number denoting the number of zeros of the solution G.

[11] See N. Mukunda *et al.*, Ref. 3, and Salam *et al.*, Ref. 1.

ACKNOWLEDGMENTS

The author is extremely thankful to Professor D. Saxon for giving him the opportunity to work during the summer at University of California at Los Angeles. The author is grateful to Professor E. C. G. Sudarshan for clarifying many points and to Professor V. Singh and Dr. S. K. Bose, Dr. R. P. Saxena, and Dr. J. Kuriyan for interesting conversations. The author wishes to thank Professor R. C. Mazumdar for his interest. Finally, he would like to thank Dr. J. Grodsky for going through the manuscript and the National Science Foundation for support.

Bogoliubov–Parasiuk–Hepp Renormalization Theorem and Spacelike Regularization

RICHARD A. BRANDT*
Department of Physics and Astronomy, University of Maryland, College Park, Maryland

(Received 20 October 1966)

Hepp's recent proof that the Bogoliubov–Parasiuk renormalization prescription leads to finite integrals after the employed space–time regularization of the Green's function is removed is extended to a case in which only spacelike regularization is used.

1. INTRODUCTION

A RENORMALIZATION prescription for relativistic perturbation theory has been proposed by Bogoliubov and Shirkov[1] which avoids contour rotation difficulties by introducing the "α representation,"

$$\Delta_l^{\epsilon}(k) = \frac{iP_l(k)}{k^2 - m_l^2 + i\epsilon} = P_l(k)\int_0^{\infty} d\alpha e^{i\alpha(k^2 - m_l^2 + i\epsilon)}, \tag{1.1}$$

for the propagators. Here $P_l(k)$ is a spin matrix which is a polynomial in k and $m_l > 0$ is always assumed. The rigorous exposition of the Bogoliubov prescription, however, as put forth by Bogoliubov and Parasiuk,[2] is unfortunately rather vague. Their proofs, furthermore, are incomplete or incorrect in part. These defects have recently been removed by Hepp,[3] who has presented a more precise exposition of the Bogoliubov prescription and has completed and corrected the Bogoliubov–Parasiuk proofs. We refer to the resulting theory as the BPH theory.

Let us consider an arbitrary nth-order Feynman diagram $G = G(V_1, \cdots, V_n, \mathfrak{L})$ connecting the vertices $V_1, \cdots, V_n$ with the lines $\mathfrak{L} = \{l_1 \cdots l_L\}$. In x space, G corresponds to a regularized unrenormalized Green's function

$$I^{r,\epsilon} = \prod_{l\in\mathfrak{L}} \Delta_l^{r,\epsilon}(x_{i_l} - x_{f_l}), \tag{1.2}$$

where $\Delta_l^{r,\epsilon}(x)$ is the Fourier transform of a suitable regularization of the net (1.1). Hepp uses the simple regularization

$$\Delta_l^{r,\epsilon}(k) = P_l(k)\int_r^{\infty} d\alpha e^{i\alpha(k^2 - m_l^2 + i\epsilon)}, \tag{1.3}$$

whereas Bogoliubov uses

$$\Delta_l^{M,\epsilon}(k) = P_l(k)\int_0^{\infty} d\alpha I_l(\alpha) e^{i\alpha(k^2 - m_l^2 + i\epsilon)}, \tag{1.4}$$

where

$$I_l(\alpha) = 1 + \sum_j c_j e^{-i\alpha(M_j^2 - m_l^2)} \tag{1.5}$$

for suitably chosen c_j, M_j [see Ref. 2, Eq. (1.8)]. It follows[4] that

$$\lim_{\epsilon\downarrow 0}\lim_{r\downarrow 0} I^{r,\epsilon}$$

is a continuous linear functional on the subspace $\mathcal{S}_N$ of test functions $\varphi \in \mathcal{S} = \mathcal{S}(\mathbf{R}^{4n})$ which vanish sufficiently fast whenever any x_i, x_j coincide. Here $\mathcal{S}$ is the space of C^{∞} functions of fast decrease, appropriately[5] topologized, and $\mathcal{S}_N$ is a closed subspace with the induced topology. The above limit exists in the (weak) topology of the space $\mathcal{S}'$ of continuous linear functionals (tempered distributions) on $\mathcal{S}$. The BPH theory gives a prescription for "renormalizing" $I^{r,\epsilon}$ by subtracting counter terms which vanish on $\mathcal{S}_N$, so that the remainder $RI^{r,\epsilon}$ converges to a tempered distribution on all of $\mathcal{S}$.

Now, the regularizations in (1.3) and (1.4) "cut off" all the components of k and are Lorentz covariant. The purpose of this paper is to extend the above result to a case where the propagators (1.1) are regularized so that only the space components of k are cut off. Thus our regularized propagators will have the (noncovariant) form

$$\Delta_l^{\eta,\epsilon}(k) = iP_l(k) f_\eta^2(\mathbf{k})/(k^2 - m_l^2 + i\epsilon), \tag{1.6}$$

where $f_\eta(\mathbf{0}) = 1$, $f_0(\mathbf{k}) = 1$, and

$$f_\eta(\mathbf{k}) \xrightarrow[|\mathbf{k}|\to\infty]{} 0$$

rapidly for $\eta > 0$. We show that, for the case of an exponential cutoff

$$f_\eta(\mathbf{k}) = e^{-\eta\mathbf{k}^2}, \tag{1.7}$$

the regularized unrenormalized Green's functions $I^{\eta,\epsilon}$ constructed from the propagators (1.6) are finite

* National Science Foundation predoctoral fellow, 1963–1966.

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

[2] N. N. Bogoliubov and O. S. Parasiuk, Acta Math. **97**, 227 (1957).

[3] K. Hepp, Commun. Math. Phys. **2**, 301 (1966).

[4] Sec. 4 of Ref. 3. See also O. Parasiuk, Dokl. Akad. Nauk, SSSR **100**, 643 (1955).

[5] See, e.g., I. M. Gel'fand and G. E. Shelov, *Generalized Functions* (Academic Press Inc., New York, 1964), Vol. 1.

in all orders of ϕ^4 perturbation theory and that the regularized renormalized Green's functions $RI^{\eta,\epsilon}$ converge in the topology of $\mathcal{S}'$ to the usual expressions:

$$\lim_{\epsilon\downarrow 0}\lim_{\eta\downarrow 0} I^{\eta,\epsilon} = \lim_{\epsilon\downarrow 0}\lim_{r\downarrow 0} I^{r,\epsilon}. \tag{1.8}$$

The significance of this extension is twofold. Firstly, it provides an example of an non-Lorentz-invariant theory, depending on a cutoff parameter η, which converges to the usual Lorentz invariant theory as $\eta \downarrow 0$. Secondly, a cutoff such as $f_\eta(\mathbf{k})$ is the type which Nature herself might provide.[6] The propagator (1.6) would arise if the large spacelike momentum contributions to the interaction Hamiltonian were suitably cut off, corresponding in x space to a smearing of the region of interaction among the fields. This would mean that a spatial extension is ascribed to the particles associated with the fields so that they acquire a definite density distribution such as might be provided by interactions with other particles not explicitly considered or by the existence of a finite intrinsic "radius" associated with the particles. The regularized propagators (1.3) and (1.4), on the other hand, have no such interpretations and appear, in fact, to be without physical significance.

For simplicity, in this paper we explicitly consider only the case of a scalar field $\phi(x)$, self-coupled according to the renormalizable interaction[7] $\phi^4(x)$. The propagators

$$\Delta^{\eta,\epsilon}(k) = ie^{-2\eta\mathbf{k}^2}/(k^2 - m^2 + i\epsilon) \tag{1.9}$$

will result if the usual interaction

$$\int d\mathbf{p}\, d\mathbf{q}\, d\mathbf{p}'\, d\mathbf{q}'\phi(\mathbf{p})\phi(\mathbf{q})\phi(\mathbf{p}')\phi(\mathbf{q}')\delta(\mathbf{p}+\mathbf{q}-\mathbf{p}'-\mathbf{q}') \tag{1.10}$$

is replaced by

$$\int d\mathbf{p}\, d\mathbf{q}\, d\mathbf{p}'\, d\mathbf{q}'\phi(\mathbf{p})\phi(\mathbf{q})\phi(\mathbf{p}')\phi(\mathbf{q}') \times e^{-\eta(\mathbf{p}^2+\mathbf{q}^2+\mathbf{p}'^2+\mathbf{q}'^2)}\delta(\mathbf{p}+\mathbf{q}-\mathbf{p}'-\mathbf{q}'). \tag{1.11}$$

In x space, this corresponds to replacing the interaction

$$\lambda\int d\mathbf{x}\phi^4(\mathbf{x}, t) \tag{1.12}$$

by

$$\lambda\int d\mathbf{x}\left[\int d\mathbf{y}\phi(\mathbf{y}, t)f_\eta(\mathbf{y}-\mathbf{x})\right]^4. \tag{1.13}$$

The factor $f_\eta(\mathbf{y}-\mathbf{x})$, the Fourier transform of $f_\eta(\mathbf{k}) = e^{-\eta\mathbf{k}^2}$, has the significance of a form factor describing the spatial extension of the particle associated with $\phi(\mathbf{y}, t)$.

In Sec. 2 we outline aspects of the BPH theory relevant to our work and record some BPH theorems and equations which we need. In Sec. 3 we discuss the regularized unrenormalized Green's functions $I^{\eta,\epsilon}$ and show that they are finite for $\eta > 0$, $\epsilon > 0$. In Sec. 4 we discuss the renormalized functions $RI^{\eta,\epsilon}$, show that the renormalizations can be implemented by adding spacelike-regularized counter terms to the Hamiltonian, and establish the equality (1.8).

2. BOGOLIUBOV-PARASIUK-HEPP RENORMALIZATION

In this section we only write down the parts of the BPH formalism to which we must explicitly refer. The reader is referred to Hepp's paper[3] for a coherent presentation of the renormalization procedure and for proofs of the theorems. One begins by writing the regularized unrenormalized function (1.2) in momentum space and introducing the α representation (1.3) for the propagators. The k integrations are then performed by means of the relation[8]

$$\lim_{\delta\downarrow 0}\int dk e^{i(ak^2+b\cdot k)-\delta(k^{0^2}+\mathbf{k}^2)} = \frac{\pi^2}{ia^2}e^{-ib^2/4a}, \tag{2.1}$$

valid for $a > 0$. One thus obtains

$$I^{r,\epsilon}(p_1, \cdots, p_n) = \delta(\textstyle\sum p_a)\int_r^\infty d\alpha_1 \cdots d\alpha_L e(\alpha, \epsilon)F_0(\alpha) \times \exp\left[i\sum_{a,b=1}^{n} A_{ab}(\alpha)(p_a\cdot p_b)\right], \tag{2.2}$$

where

$$e(\alpha, \epsilon) = \exp\left(-im^2\sum_{l=1}^{L}\alpha_l - \epsilon\sum_{l=1}^{L}\alpha_l\right), \tag{2.3}$$

$F_0(\alpha)$ is a rational function of the α's with possible nonintegrable poles when some $\alpha_l = 0$ $(r = 0)$, (A_{ab}) is a positive semidefinite quadratic form satisfying $\sum_{ab} A_{ab}x_a x_b < \sum_l \alpha_l(\sum_a |x_a|)^2$, and $A_{ab}(\alpha)$ are rational and homogeneous of degree $+1$ in α.

Now, (2.2) is to be renormalized by subtracting from it terms corresponding to diagrams obtained from the original diagram G by combining subsets of the original set $\{V_1 \cdots V_n\}$ of vertices into "generalized vertices" in all possible ways. These generalized vertices $\{V_1' \cdots V_m'\}$ are inductively assigned

[6] For earlier discussions of cutoff theories see, e.g., E. Arnous, W. Heitler, and Y. Takahashi, Nuovo Cimento **16**, 671 (1960) and references cited therein.

[7] As K. Hepp has kindly pointed out to the author, spacelike regularization cannot be used for higher spin theories. Consider, for example, the second-order self-energy diagram in a theory with $\Delta(k) = k^2/(k^2 - m^2 + i\epsilon)$. Spacelike regularization can, however, be used in any renormalizable field theory.

[8] Reference 1, pp. 288–289.

vertex parts $\mathfrak{X}_{\mathfrak{L}}^{r,\epsilon}(V_1' \cdots V_m')$ as follows:

$$\mathfrak{X}_{\mathfrak{L}}^{r,\epsilon}(V_1' \cdots V_m') = \begin{cases} 1, & \text{if } m = 1, \\ 0, & \text{if } G(V_1' \cdots V_m', \mathfrak{L}) \text{ is weakly connected,} \\ -M\bar{\mathcal{R}}_{\mathfrak{L}}^{r,\epsilon}(V_1' \cdots V_m'), & \text{otherwise,} \end{cases}$$

$$\bar{\mathcal{R}}_{\mathfrak{L}}^{r,\epsilon}(V_1' \cdots V_m') = \sum_P{}' \prod_{j=1}^{k(P)} \mathfrak{X}_{\mathfrak{L}}^{r,\epsilon}(V_{j_1}^P \cdots V_{j_{r(j)}}^P) \prod_{\text{conn}} \Delta^{r,\epsilon}.$$

Here $\sum_P'$ is over all partitions $\{V_{j_1}^P \cdots V_{j_{r(j)}}^P\}$, $1 \leq j \leq k(P)$, of $\{V_1' \cdots V_m'\}$ into $1 < k(P) \leq m$ sets and $\prod_{\text{conn}}$ is over all $l \in \mathfrak{L}$ which connect different sets of the partition. The M operator is defined in p space by

$$\begin{aligned} M\bar{\mathcal{R}}_{\mathfrak{L}}^{r,\epsilon}(V_1' \cdots V_m') &= M\delta(p_1' + \cdots + p_m')F(p_1', \cdots, p_m') \\ &= \delta(p_1' + \cdots + p_m')T(p_1', \cdots, p_m'), \end{aligned} \tag{2.4}$$

where T is the Taylor series of F around $p_1' = \cdots = p_m' = 0$ up to order ν, the superficial divergence of $G(V_1' \cdots V_m', \mathfrak{L})$, if $\nu \geq 0$ and $T = 0$ if $\nu < 0$. The renormalized regularized Green's function is finally defined by

$$RI^{r,\epsilon} = \bar{\mathcal{R}}_{\mathfrak{L}}^{r,\epsilon}(V_1 \cdots V_n) + \mathfrak{X}_{\mathfrak{L}}^{r,\epsilon}(V_1 \cdots V_n). \tag{2.5}$$

The subtraction terms will thus be α integrals of the form (2.2), with the same $d\alpha e(\alpha, \epsilon)$ as I but with different F's and A's. We can therefore write the regularized renormalized function as

$$\begin{aligned} RI^{r,\epsilon}(p_1, \cdots, p_n) = \delta(\textstyle\sum p_a) \int_r^\infty d\alpha_1 \cdots d\alpha_L e(\alpha, \epsilon) \\ \times \Big\{ F_0(\alpha) \exp\,[i \textstyle\sum A_{ab}(\alpha) p_a p_b] \\ - \sum_{s=0}^{S} F_0^s(\alpha, p) \exp\,[i \textstyle\sum A_{ab}^s(\alpha) p_a p_b] \Big\}. \end{aligned} \tag{2.6}$$

The integral in (2.6) can be written as a sum of integrals over all possible sectors of α space of the form

$$\alpha_{l_1} \geq \cdots \geq \alpha_{l_L} \geq r. \tag{2.7}$$

Hepp proves that in each such sector the integrand in (2.6) can be written as a sum of terms of the form

$$\begin{aligned} \delta(\textstyle\sum p_a) e(\alpha, \epsilon) \int_0^1 d\tau_1 \cdots d\tau_N P(p) Q(\alpha, \tau) \\ \times \Big[\prod_{i=1}^{N} D^{(i)}(\alpha, \tau) R^{(i)}(\alpha, \tau) \Big] \exp\,[i \sum_{ab} \tau_1^2 A_{ab}(\alpha, \tau) p_a p_b], \end{aligned} \tag{2.8}$$

where

$$P(p) \text{ is a monomial in the } p_a\text{'s}, \tag{2.9}$$

$Q(\alpha, \tau)$ is a rational function in the α's and τ's, uniformly bounded in (2.7) for $r \geq 0$ and all $0 \leq \tau_i \leq 1$, (2.10)

$$D^{(i)}(\alpha, \tau) = \prod_{l \in \mathfrak{L}_i} D_l^{-2}(\alpha, \tau), \quad \mathfrak{L}_i \subset \mathfrak{L}, \tag{2.11}$$

$$\mathfrak{L}_i \cap \mathfrak{L}_j = \phi \quad \text{if} \quad i \neq j, \quad i = 1, \cdots, N,$$

$D_l(\alpha, \tau)$ is rational in α, τ and $D_l(\alpha, \tau) \geq \alpha_l$ for $r \geq 0$, (2.12)

$A(\alpha, \tau) = (A_{ab}(\alpha, \tau))$ is a positive semidefinite quadratic form, the $A_{ab}(\alpha, \tau)$ are rational in α, τ and homogeneous of degree $+1$ in α, (2.13)

and where the $R^{(i)}(\alpha, \tau)$ are rational functions of α, τ which are such that (2.8) is integrable over (2.7) for $r \downarrow 0$. $R^{(i)}(\alpha, \tau)$ is essentially a monomial in $A_{ba}^{(i)}(\alpha, \tau)$, where $A^{(i)}$ is a form with the properties of A and such that

$$\begin{aligned} |A_{ab}^{(i)}(\alpha, \tau)| \leq c_i \max\,\{\alpha_l \,|\, l \in \mathfrak{L}_i'|\}, \\ c_i < \infty, \quad \mathfrak{L}_i' \subset \mathfrak{L}. \end{aligned} \tag{2.14}$$

Equation (2.8) is convergent for $r \downarrow 0$ in (2.7) because $\alpha_l > \alpha_{l'}$ for $l \in \mathfrak{L}_i$, $l' \in \mathfrak{L}_i'$ in (2.7). We remark that the $D_l^{-2}(\alpha, \tau)$ factors arise from the use of (2.1) to perform the k_l integrations, where k_l is the momentum assigned to the line l and D_l corresponds to a. Furthermore, the conditions (2.12) and (2.14) are all that are needed to establish the integrability of (2.8) in (2.7). The τ's in (2.8) arise from the use of the expression

$$\frac{1}{j!} \int_0^1 d\tau (1 - \tau)^j \frac{\partial^{j+1}}{\partial \tau^{j+1}} f(\tau x_1, \cdots, \tau x_m) \tag{2.15}$$

for the remainder of the Taylor series of $f(x_1, \cdots, x_m)$ around $(0, \cdots, 0)$ up to order j.

It follows that

$$\lim_{r \downarrow 0} I^{r,\epsilon}(p_1, \cdots, p_n) = I^{0,\epsilon}(p_1, \cdots, p_n) \tag{2.16}$$

exists in the ordinary sense and is [apart from $\delta(\sum p_a)$] a polynomially bounded C^∞ function. Hepp finally shows that

$$\lim_{\epsilon \downarrow 0} I^{0,\epsilon}(p_1, \cdots, p_n)$$

exists in $\mathcal{S}'$ for the $p_1, \cdots, p_{n-1}$ restricted to any m-dimensional linear manifold, $0 \leq m \leq 4(n-1)$, so that the nth order Green's functions

$$\langle T\phi(x_1) \cdots \phi(x_m) \rangle_{(n)},$$

renormalized as above, are in p space [up to $\delta(\sum p_a)$] Lorentz covariant boundary values in $\mathcal{S}'(\mathbf{R}^{4(m-1)})$ of sums of Feynman integrals.

3. SPACELIKE REGULARIZATION

In this section we are concerned with the regularized unrenormalized Green's function $I^{n,\epsilon}$, corresponding

to the diagram $G = G(V_1, \cdots, V_n, \mathfrak{L})$, which replaces (2.2) when the spacelike regularized propagators (1.9) are used instead of (1.3). In order to perform the integrations over the internal momenta, we use the relation

$$\lim_{\delta \downarrow 0} \int dk e^{i(ak^2+b\cdot k)} e^{-\zeta \mathbf{k}^2 - \xi\cdot\mathbf{k}} e^{-\delta(k^{0^2}+\mathbf{k}^2)} = \frac{\pi^2}{i}[a(a-i\zeta)^3]^{-\frac{1}{2}} e^{-ib^2/4a} e^{[a(2b\cdot\xi - i\xi^2) - b^2\zeta]/4a(a-i\zeta)}, \tag{3.1}$$

valid provided

$$a > 0, \quad \mathrm{Re}\,(a - i\zeta) \equiv a + \zeta_2 > 0, \quad \mathrm{Re}\,\zeta \equiv \zeta_1 > 0, \tag{3.2}$$

where we wrote $\zeta = \zeta_1 + i\zeta_2$ with ζ_1 and ζ_2 real. Equation (3.1) is a generalization of, and can be derived in the same way as, Eq. (2.1). The simplicity of the relation (3.1) is the reason for our choice of the exponential cutoff (1.7). The factorization of the exponentials in (3.1) into a Lorentz invariant piece, the same as occurs in (2.1), and a non-Lorentz invariant piece will be especially helpful.

According to the Feynman rules, we have in momentum space

$$I^{\eta,\epsilon}(p_1 \cdots p_n) = \int K \prod_{a=1}^{n} \delta(\textstyle\sum k - p_a) \cdot \prod_{l\in\mathfrak{L}} \frac{ie^{-2\eta \mathbf{k}_l^2}}{k_l^2 - m^2 + i\epsilon} dk_l, \tag{3.3}$$

where K contains the vertex and 2π factors and the δ functions represent momentum conservation at each vertex ($\sum k$ is the sum of the momenta of the internal lines meeting at the vertex a to which the external momentum p_a corresponds). We introduce the α representation (1.1) to get

$$I^{\eta,\epsilon}(p_1 \cdots p_n) = K \int d\alpha_1 \cdots d\alpha_L \int \exp\left(i \sum_l \alpha_l k_l^2 - 2\eta \sum_l \mathbf{k}_l^2\right) \times e(\alpha, \epsilon) \prod_a \delta(\textstyle\sum k - p_a) \prod_l dk_l, \tag{3.4}$$

where $e(\alpha, \epsilon)$ is given by (2.3). We proceed formally and use (3.1) to perform the k_l integrations. We afterwards show that the conditions (3.2) are satisfied at each step and that the resulting α integrands are integrable so that our manipulations were justified. Thus we arrive at

$$I^{\eta,\epsilon}(p_1 \cdots p_n) = \delta(\textstyle\sum p_a) \int d\alpha_1 \cdots d\alpha_L e(\alpha, \epsilon) F_\eta(\alpha) \times \exp\left[i \sum A_{ab}(\alpha)(p_a \cdot p_b)\right] \times \exp\left[-\sum U_{ab}(\alpha, \eta)(\mathbf{p}_a \cdot \mathbf{p}_b)\right]. \tag{3.5}$$

Our notation indicates the crucial fact, which follows from comparison of (2.1) and (3.1), that functions $A_{ab}(\alpha)$ and

$$F_0(\alpha) = \lim_{\eta \downarrow 0} F_\eta(\alpha)$$

defined by (3.5) are the same as those defined by (2.2). It is furthermore clear that

$$\lim_{\eta \downarrow 0} U_{ab}(\alpha, \eta) = 0.$$

Although we have not yet shown that (3.5) is finite, the quadratic forms $A = (A_{ab})$ and $U = (U_{ab})$ are still well defined. All our results follow from the positive semidefiniteness of the following forms:

$$A, \quad \mathrm{Re}\, U, \quad A + \mathrm{Im}\, U. \tag{3.6}$$

That A is positive semidefinite (psd) has already been shown by Bogoliubov *et al.*[1,2] by the method of induction. They observe that A is psd for the simplest diagrams and that if the A corresponding to any diagram G has this property, then so do the A's corresponding to the diagrams G_a and G_b obtained by (a) adding a new internal line for a given number of vertices and (b) adding a new vertex together with one line joining this vertex to one of the old vertices. Their result then follows from the fact that any diagram can be constructed from the simplest diagrams by means of successive application of a suitable number of operations (a) and (b). We now use this same method to show that Re U and A + Im U are psd.

All diagrams for the ϕ^4 theory can be constructed as above from the simplest second-order diagram ($\lambda\Delta\lambda$)—for which $A = \alpha$ and $U = 2\eta$ have the desired properties. Thus we suppose that the A and U defined by (3.5), corresponding to the arbitrary diagram G, are such that Re U and A + Im U are psd and consider operation (a). We assume that an internal line l_a with momentum k_0 is added between, say, vertices 1 and 2 of G to create diagram G_a. This replaces the factors

$$\delta\Big(\sum_1 k + p_1\Big)\delta\Big(\sum_2 k + p_2\Big) \tag{3.7}$$

in (3.4) by

$$\frac{1}{(2\pi)^4} \int d\alpha_0 e^{i\alpha_0(k_0^2 - m^2 + i\epsilon)} e^{-2\eta \mathbf{k}_0^2} \times \delta\Big(\sum_1 k + p_1 - k_0\Big)\delta\Big(\sum_2 k + p_2 - k_0\Big) dk_0. \tag{3.8}$$

Thus, defining

$$\mathfrak{L}_a = \mathfrak{L} \cup \{l_a\}, \quad e_1 = 1, \quad e_2 = -1, \quad e_r = 0 \quad (r > 2), \tag{3.9}$$

we see that the transformation $G \to G_a$ induces the transformation

$$\delta(\sum p_a)\int d\alpha_1 \cdots d\alpha_L e_{\mathfrak{L}}(\alpha,\epsilon)F_\eta(\alpha) \times \exp[i\sum A_{ab}p_a\cdot p_b]\exp[-\sum U_{ab}\mathbf{p}_a\cdot\mathbf{p}_b]$$
$$\to \delta(\sum p_a)\int d\alpha_0\cdots d\alpha_L e_{\mathfrak{L}_a}(\alpha,\epsilon)F_\eta(\alpha)\frac{1}{(2\pi)^4} \times \int dk_0 \exp\{i[(\alpha_0+\sum A_{ab}e_ae_b)k_0^2 + 2\sum A_{ab}e_ak_0p_b + \sum A_{ab}p_ap_b]\}\exp(-1)[(2\eta+\sum U_{ab}e_ae_b)\mathbf{k}_0^2 + 2\sum U_{ab}e_a\mathbf{k}_0\cdot\mathbf{p}_b + \sum U_{ab}\mathbf{p}_a\cdot\mathbf{p}_b] \tag{3.10}$$

$$= \delta(\sum p_a)\int d\alpha_0\cdots d\alpha_L e_{\mathfrak{L}_a}(\alpha,\epsilon)F_\eta(\alpha)\frac{1}{(2\pi)^4} \times \int dk_0 \exp\{i[(\alpha_0+a)k_0^2 + 2k_0\cdot b + c]\} \times \exp(-1)[(2\eta+\zeta)\mathbf{k}_0^2 + 2\mathbf{k}_0\cdot\boldsymbol{\xi} + \nu] \tag{3.11}$$

$$= \delta(\sum p_a)\int d\alpha_0\cdots d\alpha_L e_{\mathfrak{L}_a}(\alpha,\epsilon)F_\eta(\alpha)\frac{\pi^2}{(2\pi)^4 i} \times \{(\alpha_0+a)[\alpha_0+a-i(2\eta+\zeta)]^3\}^{-\frac{1}{2}} \times \exp\left[i\left(c-\frac{b^2}{\alpha_0+a}\right)\right]\exp(-1)\mathfrak{U}. \tag{3.12}$$

In (3.11) we have introduced the definitions

$$a = \sum A_{ab}e_ae_b, \quad b = \sum A_{ab}e_ap_b, \quad c = \sum A_{ab}p_ap_b, \tag{3.13}$$

$$\zeta = \sum U_{ab}e_ae_b, \quad \boldsymbol{\xi} = \sum U_{ab}e_a\mathbf{p}_b, \quad \nu = \sum U_{ab}\mathbf{p}_a\cdot\mathbf{p}_b, \tag{3.14}$$

and in (3.12) we put

$$\mathfrak{U} = \frac{-(\alpha+a)(2\mathbf{b}\cdot\boldsymbol{\xi} - i\boldsymbol{\xi}^2) + \mathbf{b}^2(2\eta+\zeta)}{(\alpha+a)[\alpha+a-i(2\eta+\zeta)]} + \nu, \qquad \alpha = \alpha_0, \tag{3.15}$$

according to (3.1).

If we decompose ζ, $\boldsymbol{\xi}$, and ν into real and imaginary parts by writing

$$\zeta = \zeta_1 + i\zeta_2, \quad \boldsymbol{\xi} = \boldsymbol{\xi}_1 + i\boldsymbol{\xi}_2, \quad \nu = \nu_1 + i\nu_2, \tag{3.16}$$

it then follows from the induction assumption of the positive semidefiniteness of Re U that

$$\zeta_1 = \mathrm{Re}\sum U_{ab}e_ae_b \geq 0, \quad \nu_1 = \mathrm{Re}\sum U_{ab}\mathbf{k}_a\cdot\mathbf{k}_b \geq 0, \quad \nu_1\zeta_1 \geq \boldsymbol{\xi}_1^2. \tag{3.17}$$

The last inequality follows from the fact that the discriminate of the quadratic form $\mathrm{Re}\sum U_{ab}x_ax_b$ is nonnegative for $x_a = p_a^it_1 + e_at_2$ $(i = 1, 2, 3)$ so that

$$(\mathrm{Re}\sum U_{ab}p_a^ip_b^i)(\mathrm{Re}\sum U_{ab}e_ae_b) \geq (\mathrm{Re}\sum U_{ab}p_a^ie_a)^2. \tag{3.18}$$

Similarly, the positive semidefiniteness of $A + \mathrm{Im}\,U$ implies that

$$a+\zeta_2 \geq 0, \quad -c+\nu_2 \geq 0, \quad (a+\zeta_2)(-c+\nu_2) \geq (\mathbf{b}+\boldsymbol{\xi}_2)^2, \tag{3.19}$$

for $p_a = (0, \mathbf{p}_a)$ (note that $p_a\cdot p_b = p_a^0p_b^0 - \mathbf{p}_a\cdot\mathbf{p}_b$ so that $-c = +\sum A_{ab}\mathbf{p}_a\cdot\mathbf{p}_b \geq 0$). Finally, since A is psd, we have

$$a \geq 0, \quad -ca \geq -b^2. \tag{3.20}$$

Keeping $p_a = (0, \mathbf{p}_a)$, we will have shown that the operation (a) preserves the positive definiteness of Re U and $A + \mathrm{Im}\,U$ provided we show that

$$\mathrm{Re}\,\mathfrak{U} \geq 0 \tag{3.21}$$

and

$$-c + [b^2/(\alpha+a)] + \mathrm{Im}\,\mathfrak{U} \geq 0. \tag{3.22}$$

We have

$$[(\alpha+a+\zeta_2)^2 + (2\eta+\zeta_1)^2]\,\mathrm{Re}\,\mathfrak{U} = (2\eta+\zeta_1)[(\mathbf{b}+\boldsymbol{\xi}_2)^2 - \boldsymbol{\xi}_1^2] - 2(\alpha+a+\zeta_2)\boldsymbol{\xi}_1\cdot(\mathbf{b}+\boldsymbol{\xi}_2) + \nu_1[(\alpha+a+\zeta_2)^2 + (2\eta+\zeta_1)^2] \tag{3.23}$$

$$\geq (\nu_1\zeta_1 - \boldsymbol{\xi}_1^2)(2\eta+\zeta_1) + (2\eta+\zeta_1)(\mathbf{b}+\boldsymbol{\xi}_2)^2 - 2(\alpha+a+\zeta_2)\boldsymbol{\xi}_1\cdot(\mathbf{b}+\boldsymbol{\xi}_2) + \nu_1(\alpha+a+\zeta_2)^2 \tag{3.24}$$

$$\geq \zeta_1(\mathbf{b}+\boldsymbol{\xi}_2)^2 - 2(\alpha+a+\zeta_2)\boldsymbol{\xi}_1\cdot(\mathbf{b}+\boldsymbol{\xi}_2) + (\boldsymbol{\xi}_1^2/\zeta_1)(\alpha+a+\zeta_2)^2 \tag{3.25}$$

$$= (1/\zeta_1)[\zeta_1(\mathbf{b}+\boldsymbol{\xi}_2) - (\alpha+a+\zeta_2)\boldsymbol{\xi}_1]^2 \tag{3.26}$$

$$\geq 0. \tag{3.27}$$

Here (3.23) follows from (3.15), (3.24) from $2\eta\nu_1 \geq 0$, $\zeta_1 \geq 0$, and (3.25) from $\nu_1\zeta_1 - \boldsymbol{\xi}_1^2 \geq 0$. This establishes (3.21). We also have

$$[(\alpha+a+\zeta_2)^2 + (2\eta+\zeta_1)^2]\,\mathrm{Im}\,\mathfrak{U} = -(2\mathbf{b}\cdot\boldsymbol{\xi}_1 + 2\boldsymbol{\xi}_1\cdot\boldsymbol{\xi}_2)(2\eta+\zeta_1) - (2\mathbf{b}\cdot\boldsymbol{\xi}_2 - \boldsymbol{\xi}_1^2 + \boldsymbol{\xi}_2^2)(\alpha+a+\zeta_2) - [b^2/(\alpha+a)][\zeta_2^2 + (2\eta+\zeta_1)^2] - b^2\zeta_2 + \nu_2[(\alpha+a+\zeta_2)^2 + (2\eta+\zeta_1)^2] \tag{3.28}$$

so that

$$[(\alpha+a+\zeta_2)^2 + (2\eta+\zeta_1)^2]\left[-c + \frac{b^2}{\alpha+a} + \mathrm{Im}\,\mathfrak{U}\right] = -(2\mathbf{b}\cdot\boldsymbol{\xi}_1 + 2\boldsymbol{\xi}_1\cdot\boldsymbol{\xi}_2)(2\eta+\zeta_1) - (2\mathbf{b}\cdot\boldsymbol{\xi}_2 - \boldsymbol{\xi}_1^2 + \boldsymbol{\xi}_2^2)(\alpha+a+\zeta_2) + b^2(\alpha+a+\zeta_2) + (-c+\nu_2) \times [(\alpha+a+\zeta_2)^2 + (2\eta+\zeta_1)^2]$$

$$= [(a + \zeta_2)(-c + \nu_2) - (\mathbf{b}^2 + 2\mathbf{b}\cdot\boldsymbol{\xi}_2 + \boldsymbol{\xi}_2^2)]$$
$$\times (\alpha + a + \zeta_2) - 2(\mathbf{b} + \boldsymbol{\xi}_2)\cdot\boldsymbol{\xi}_1(2\eta + \zeta_1)$$
$$+ \boldsymbol{\xi}_1^2(\alpha + a + \zeta_2) + (-c + \nu_2)$$
$$\times [\alpha(\alpha + a + \zeta_2) + (2\eta + \zeta_1)^2]$$
$$\geq -2(\mathbf{b} + \boldsymbol{\xi}_2)\cdot\boldsymbol{\xi}_1(2\eta + \zeta_1) + \boldsymbol{\xi}_1^2(\alpha + a + \zeta_2)$$
$$+ \frac{(\mathbf{b} + \boldsymbol{\xi}_2)^2}{\alpha + a + \zeta_1}(2\eta + \zeta_1)^2 \tag{3.29}$$

$$= (\alpha + a + \zeta_2)\left[\frac{(2\eta + \zeta_1)(\mathbf{b} + \boldsymbol{\xi}_2)}{\alpha + a + \zeta_2} - \boldsymbol{\xi}_1\right]^2$$
$$\geq 0, \tag{3.30}$$

where (3.29) follows from (3.19) and (3.30) from $a + \zeta_2 \geq 0$. This establishes (3.22).

Next, we again suppose that A and U defined by (3.5) are such that Re U and $A + \text{Im } U$ are positive definite and consider operation (b) applied to G. We assume that the new vertex 0 is connected with vertex 1 by a line l_b of momentum p_0. In (3.4) this will give rise to the transformation

$$\delta\left(\sum_1 k + p_1\right) \to -\lambda \int d\alpha_0 e^{i\alpha_0(k_0^2 - m^2 + i\epsilon)} e^{-2\eta\mathbf{k}_0^2}$$
$$\times \delta\left(\sum_1 k + p_1 + k_0\right)\delta(p_0 - k_0)\, dk_0, \tag{3.31}$$

where λ is the ϕ^4 coupling constant. Thus, defining

$$\mathfrak{L}_b = \mathfrak{L} \cup \{l_b\}, \quad e_1 = 1, \quad e_a = 0 \quad (a \geq 2), \tag{3.32}$$

we see that $G \to G_b$ induces the transformation

$$\delta\left(\sum_{a=1}^n p_a\right)\int d\alpha_1 \cdots d\alpha_L e_{\mathfrak{L}}(\alpha, \epsilon) F_\eta(\alpha) \exp\left[i\sum A_{ab} p_a \cdot p_b\right]$$
$$\times \exp\left[-\sum U_{ab}\mathbf{p}_a \cdot \mathbf{p}_b\right]$$
$$\to \delta\left(\sum_{a=0}^n p_a\right)\int d\alpha_0 \cdots d\alpha_L e_{\mathfrak{L}_b}(\alpha, \epsilon) F_\eta(\alpha)(-\lambda)$$
$$\times \exp\{i[\alpha_0 p_0^2 + \sum A_{ab}(p_a + e_a p_0)(p_b + e_b p_0)\}$$
$$\times \exp(-1)[2\eta\mathbf{p}_0^2 + \sum U_{ab}(\mathbf{p}_a + e_a\mathbf{p}_0)\cdot(\mathbf{p}_b + e_b\mathbf{p}_0)]. \tag{3.33}$$

It follows immediately that the new Re U and the new $A + \text{Im } U$ are psd. This completes the induction proof and establishes the positive semidefiniteness of the forms (3.6) for an arbitrary diagram.

It follows from the above theorem and its proof that the conditions (3.2) are satisfied at each step in the evaluation of the k_l integrals in (3.4) so that use of (3.1) is justified. We simply observe that (3.2) hold for the simplest diagram and that if they hold at each step in the evaluation of (3.4) corresponding to the arbitrary diagram G, then they do likewise for G_a and G_b. Indeed, for the k_0 integration in (3.10) the conditions (3.2) become

$$\alpha_0 + a > 0, \quad a + \zeta_2 > 0, \quad 2\eta + \zeta_1 > 0, \tag{3.34}$$

and these follow from the positive definiteness of A, $A + \text{Im } U$, and Re U, whereas the k_0 integration in (3.31) is evaluated by means of $\delta(p_0 - k_0)$. This establishes the equality of (3.4) and (3.5).

Our final task in this section is to prove that the integrand in (3.5) is (Lebesgue) integrable so that (3.5) is well-defined and finite. This will also be shown to follow from the above theorem and its proof. We use the above method of induction to show that for an arbitrary diagram $G(V_1 \cdots V_n, \mathfrak{L})$ the corresponding function $F_\eta(\alpha)$ satisfies for fixed $\eta > 0$ the condition

$$|F_\eta(\alpha)| \leq \text{const} \prod_{l\in\tilde{\mathfrak{L}}} \alpha_l^{-\frac{1}{2}} \tag{3.35}$$

for some subset $\tilde{\mathfrak{L}}$ of $\mathfrak{L}$. This is immediate for the simplest diagram and so we suppose it holds for F_η in (3.5) corresponding to G and show it then holds for the F_η's corresponding to G_a and G_b.

We see by (3.12) that $G \to G_a$ induces the transformation

$$F_\eta(\alpha) \to F_{a\eta}(\alpha) = F_\eta(\alpha)[\pi^2/(2\pi)^4 i]$$
$$\times \{(\alpha_0 + a)[\alpha_0 + a - i(2\eta + \zeta)]^3\}^{-\frac{1}{2}}. \tag{3.36}$$

However, since $a \geq 0$ and $\zeta_1 \geq 0$, we have

$$|\{(\alpha_0 + a)[\alpha_0 + a - i(2\eta + \zeta)]^3\}^{-\frac{1}{2}}|^2$$
$$= (\alpha_0 + a)^{-1}[(\alpha_0 + a + \zeta_2)^2 + (2\eta + \zeta_1)^2]^{-\frac{3}{2}}$$
$$\leq (2\eta)^{-3}\alpha_0^{-1}, \tag{3.37}$$

so that

$$|F_{a\eta}(\alpha)| \leq \text{const} \prod_{l\in\tilde{\mathfrak{L}}_a} \alpha_l^{-\frac{1}{2}}, \tag{3.38}$$

where $\mathfrak{L}_a = \tilde{\mathfrak{L}} \cup \{l_a\}$. Next, we see by (3.33) that $G \to G_b$ induces the transformation

$$F_\eta(\alpha) \to F_{b\eta}(\alpha) = F_\eta(\alpha)(-\lambda) \tag{3.39}$$

so that

$$|F_{b\eta}(\alpha)| \leq \text{const} \prod_{l\in\tilde{\mathfrak{L}}} \alpha_l^{-\frac{1}{2}}. \tag{3.40}$$

This establishes (3.35) for an arbitrary diagram.

Now we can establish the integrability of the integrand in (3.5). This follows from the facts that it is measurable and, since (3.35) holds and

$$\text{Re}\left\{\sum_{ab} U_{ab}(\alpha, \eta)(\mathbf{p}_a \cdot \mathbf{p}_b)\right\} > 0,$$

it is bounded by the integrable (for $\eta > 0$ and $\epsilon > 0$) function

$$\text{const} \exp\left(-\epsilon \sum_{l=1}^L \alpha_l\right) \prod_{l\in\tilde{\mathfrak{L}}} \alpha_l^{-\frac{1}{2}}. \tag{3.41}$$

This shows that the spacelike regularized propagations (1.9) define a finite unrenormalized perturbation theory.

4. RENORMALIZATION AND REMOVAL OF REGULARIZATION

Now we are ready to discuss the renormalization of (3.5). We want the subtractions to arise from the addition of spacelike-regularized counterterms to the interaction Hamiltonian (1.13) and to be such that the equality (1.8) holds. There still remains, however, a large amount of arbitrariness in the choice of these subtractions. The simplest procedure would probably be to use the BPH prescription (2.5) with the M operator (2.4). This would correspond to spacelike-regularized counterterms but would require a rather involved proof of (1.8). Such a proof would require the use of more properties of (3.5) than we have already established and would, furthermore, not be applicable to more general cutoffs in which terms like η^2/α could appear in U.

We are thus led to consider a more general $\bar{M}$ operator which amounts to splitting U into two parts,

$$U = U' + U'', \tag{4.1}$$

where

$$U' \xrightarrow[\eta \downarrow 0]{} 0$$

and

$$U'' \xrightarrow[\eta \downarrow 0]{} 0,$$

and applying M only to U'. This $\bar{M}$ operation would still correspond to spacelike-regularized counterterms since U'', the only form to which M is not applied, involves only spacelike momentum components. In fact, using the methods developed in Ref. 9, we can write, for example, the charge-renormalization counter term [the analog of $(Z_1^{-1} - 1)j^\mu A^\mu$ in electrodynamics] in the form

$$\int \left[\prod_{a=1}^{4} d\mathbf{p}_a \phi(\mathbf{p}_a) f_\eta(\mathbf{p}_a)\right] \mathcal{F}_\eta(\mathbf{p}_1, \cdots, \mathbf{p}_4)\delta(\textstyle\sum \mathbf{p}_a), \tag{4.2}$$

where

$$\mathcal{F}_\eta(\mathbf{p}_1, \cdots, \mathbf{p}_4)\delta(\textstyle\sum \mathbf{p}_a) = \sum_G (\bar{M} S I_G^\eta)(\mathbf{p}_1, \cdots, \mathbf{p}_4). \tag{4.3}$$

Here $\sum_G$ is over all strongly connected Feynman diagrams G (I_G^η is the corresponding Green's function) with four external lines (we actually only sum up to some finite order, of course). S is the operator defined in Ref. 9 so that, if we write

$$I_G^\eta(p_1, p_2, p_3, p_4) = i\lambda \int \frac{dk}{(2\pi)^4} \frac{dk'}{(2\pi)^4} D_{G_1}^\eta(k) D_{G_2}^\eta(p_1 - k - k') D_{G_3}^\eta(k') \times J_G^\eta(k, p_1 - k - k', k', p_2, p_3, p_4), \tag{4.4}$$

where the D's are regularized unrenormalized propagators and J_G is the regularized unrenormalized function corresponding to a suitable diagram with six external lines, then

$$SI_G = i\lambda \int R D_{G_1} R D_{G_2} R D_{G_3} R J_G, \tag{4.5}$$

where R is defined by (2.5) using $\bar{M}$ instead of M. If we substitute the α representations of the factors in the integrand (see below), perform the k and k' integrations, and apply $\bar{M}$, we arrive at an expression for $\mathcal{F}_\eta(\mathbf{p}_1 \cdots \mathbf{p}_4)\delta(\sum \mathbf{p}_a)$ as a sum of terms of the form

$$\delta(\textstyle\sum \mathbf{p}_a) \sum_G \int d\tau_1 \cdots d\tau_{N_G} \int d\alpha_1 \cdots d\alpha_{L_G} e_G(\alpha, \epsilon) \times F_G(\alpha, \tau, \eta, \mathbf{p}) \exp\left[-\sum_{ab} U''_{Gab}(\alpha, \tau, \eta)\mathbf{p}_a \cdot \mathbf{p}_b\right]. \tag{4.6}$$

Now we want $A + \operatorname{Im} U'$, $\operatorname{Re} U''$, and $\operatorname{Im} U''$ to be psd forms and, in order to have a simple proof of (1.8), we also want the matrix elements U'_{ab} to be well behaved. We can achieve all this with the following choice of $\bar{M}$. Operating on any expression of the form (3.5), with A, $\operatorname{Re} U$, and $A + \operatorname{Im} U$ psd forms, our $\bar{M}$ operator is related to the BPH M operator by

$$\bar{M} \int d\alpha e(\alpha, \epsilon) F_\eta(\alpha, p) \exp\left[i \textstyle\sum A_{ab}(\alpha) p_a p_b\right] \times \exp\left[-\textstyle\sum U_{ab}(\alpha, \eta)\mathbf{p}_a \cdot \mathbf{p}_b\right] = \int d\alpha e(\alpha, \epsilon) \exp\left[-\textstyle\sum W_{ab}(\alpha, \eta)\mathbf{p}_a \cdot \mathbf{p}_b\right] M F_\eta(\alpha, p) \times \exp\left[i \textstyle\sum A_{ab}(\alpha) p_a p_b - i \sum V_{ab}(\alpha, \eta)\mathbf{p}_a \cdot \mathbf{p}_b\right], \tag{4.7}$$

where the real forms $\operatorname{Re} W$, $\operatorname{Im} W$, and V are defined by

$$\operatorname{Re} W = \operatorname{Re} U, \tag{4.8}$$

$$\operatorname{Im} W = [(\mathcal{A} - \mathcal{B})/(\mathcal{A} + \mathcal{B})]^2 B, \tag{4.9}$$

$$V = -A + [4\mathcal{A}\mathcal{B}/(\mathcal{A} + \mathcal{B})^2]B, \tag{4.10}$$

with

$$B = A + \operatorname{Im} U, \tag{4.11}$$

$$\mathcal{A} = \sum_{ab} |A_{ab}|, \quad \mathcal{B} = \sum_{ab} |B_{ab}|. \tag{4.12}$$

It is then immediate that

$$U = \operatorname{Re} W + i \operatorname{Im} W + iV = W + iV, \tag{4.13}$$

[9] R. A. Brandt, Ph.D. thesis, MIT (1966) (to be published).

that the forms

$$\text{Re } W, \quad \text{Im } W, \quad A + V, \tag{4.14}$$

are psd, that

$$W_{ab}(\alpha, 0) = V_{ab}(\alpha, 0) = 0, \tag{4.15}$$

and that

$$|V_{ab}(\alpha, \eta)| \leq 5\mathcal{A}(\alpha). \tag{4.16}$$

We see in particular that $\bar{M}$ could be applied to (3.5) even if U contained terms like $1/\alpha$, whereas then M could not be applied.

We extend $\bar{M}$ to operate on sums of α integrals by linearity. A further property of $\bar{M}$ is that it, like M, is idempotent: $\bar{M}^2 = \bar{M}$. This follows from the fact that if W corresponds to the pair A, $U = W + iV$, then it also corresponds to the pair 0, W. We should also mention, in connection with expressions of the form (2.8), which are actually sums of α integrands, that $\bar{M}$ does not commute with $\int_0^1 d\tau$ but that it does commute with

$$\int_0^1 d\tau (1-\tau)^j \left(\frac{\partial}{\partial \tau}\right)^{j+1};$$

i.e.,

$$\bar{M} \int_0^1 d\tau (1-\tau)^j \left(\frac{\partial}{\partial \tau}\right)^{j+1} \mathcal{K}(\tau, p) = \int_0^1 d\tau (1-\tau)^j \left(\frac{\partial}{\partial \tau}\right)^{j+1} \bar{M}\mathcal{K}(\tau, p). \tag{4.17}$$

Thus the τ differentiations implicit in (2.8) must be undone before $\bar{M}$ can be taken inside the τ integrals.

Now we define $RI^{\eta,\epsilon}$ analogously to the definition (2.5) of $RI^{r,\epsilon}$ except that the M operation is replaced by the $\bar{M}$ operation and the space–time regularization (1.3) is replaced by the spacelike regularization (1.9). Then, corresponding to each term (2.8) integrated over (2.7) contributing to $RI^{r,\epsilon}$, we will have a term

$$\begin{aligned}&\delta(\textstyle\sum p_a) e(\alpha, \epsilon) \int_0^1 d\tau_1 \cdots d\tau_N P(p) Q(\alpha, \tau) \\ &\times \left[\prod_{i=1}^N D^{(i)}(\alpha, \tau, \eta) R^{(i)}(\alpha, \tau)\right] \\ &\times \exp\left[i \textstyle\sum \tau_1^2 A_{ab}(\alpha, \tau) p_a p_b - i \sum \tau_1^2 V_{ab}(\alpha, \tau, \eta) \mathbf{p}_a \cdot \mathbf{p}_b\right] \\ &\times \exp\left[-\textstyle\sum W_{ab}(\alpha, \tau, \eta) \mathbf{p}_a \cdot \mathbf{p}_b\right]\end{aligned} \tag{4.18}$$

integrated over (2.7), with $r = 0$, contributing to $RI^{\eta,\epsilon}$, where

$$D^{(i)}(\alpha, \tau, \eta) = \prod_{l \in \mathfrak{L}_i} D_l^{-2}(\alpha, \tau, \eta), \tag{4.19}$$

$D_l(\alpha, \tau, \eta)$ is algebraic in α, τ, η and

$$|D_l(\alpha, \tau, \eta)| \geq \alpha_l \quad \text{for} \quad \eta \geq 0, \tag{4.20}$$

$$D_l(\alpha, \tau, 0) = D_l(\alpha, \tau), \tag{4.21}$$

$A(\alpha, \tau) + V(\alpha, \tau, \eta)$, Re $W(\alpha, \tau, \eta)$, Im $W(\alpha, \tau, \eta)$ are psd quadratic forms whose matrix elements are rational functions of α, τ, η, (4.22)

and

$$V(\alpha, \tau, 0) = W(\alpha, \tau, 0) = 0. \tag{4.23}$$

$RI^{\eta,\epsilon}$ will contain other terms corresponding to (2.8), each of which will have the form

$$\begin{aligned}&\delta(\textstyle\sum p_a) e(\alpha, \epsilon) \int_0^1 d\tau_1 \cdots d\tau_N P(p, \mathbf{p}) Q(\alpha, \tau) \\ &\times \left[\prod_{i=1}^N D^{(i)}(\alpha, \tau, \eta) R^{(i)}(\alpha, \tau, \eta)\right] \\ &\times \exp\left[i \textstyle\sum \tau_1^2 A(\alpha, \tau) pp - i \sum \tau_1^2 V(\alpha, \tau, \eta) \mathbf{pp} \right. \\ &\qquad \left. - \textstyle\sum W(\alpha, \tau, \eta) \mathbf{pp}\right],\end{aligned} \tag{4.24}$$

where

$$P(p, \mathbf{p}) \text{ is a monomial in the } p_a\text{'s and } \mathbf{p}_a\text{'s}, \tag{4.25}$$

and $R^{(i)}(\alpha, \tau, \eta)$ differs from $R^{(i)}(\alpha, \tau)$ by having some factors $A_{ab}^{(i)}(\alpha, \tau)$ replaced by factors $V_{ab}^{(i)}(\alpha, \tau, \eta)$, which are rational function satisfying

$$|V_{ab}^{(i)}(\alpha, \tau, \eta)| \leq 5 \sum_{ab} |A_{ab}^{(i)}(\alpha, \tau)| \equiv 5\mathcal{A}^{(i)}(\alpha, \tau)$$

and

$$V_{ab}^{(i)}(\alpha, \tau, 0) = 0. \tag{4.26}$$

The results stated in the above paragraph can be derived in exactly the same way as (2.8)–(2.14) were derived by Hepp. The required additional algebraic manipulations are precisely the same as those performed in Sec. III. We should emphasize that although the algebra is the same, Hepp's induction proof is much more subtle than the one used in Sec. III since it takes renormalization into account. We do not give a detailed derivation here, but mention the algebraic differences between such a derivation and that of Hepp and show how they lead to (4.18)–(4.25).

Hepp's proof amounts algebraically to establishing the invariance of the form (2.8) with the properties (2.9)–(2.14) under the operations (b) and (a) of Sec. III followed by an application of the M operator or, using (2.15), the $(1 - M)$ operator. We do the same thing for (4.18)–(4.26). We first observe that the operation (b) essentially only creates new quadratic forms A', V', W', whose matrix elements are linear combinations of the corresponding ones in (4.18) and α_0 and 2η and which trivially leave the properties (4.22) and (4.23) invariant. The invariance of (4.22) under operation (a) follows as in (3.27) and (3.30) if one first combines the W and $i\tau_1^2 V$ in (4.18) into a U and then, after application of (a), makes a new W'' and iV'' according to (4.7)–(4.11). Operation (a) also multiplies the integrand in (4.18) or (4.24) by

$$D_{l_a}^{-2}(\alpha, \tau, \eta) \equiv \{(\alpha_0 + a)[\alpha_0 + a - i(2\eta + \zeta)]^3\}^{-\frac{1}{2}}, \tag{4.27}$$

where

$$a = \sum A'_{ab} e_a e_b, \quad \zeta = \sum (W'_{ab} + iV'_{ab}) e_a e_b, \quad (4.28)$$

so that

$$a \geq 0, \quad a + \zeta_2 \geq 0. \quad (4.29)$$

Thus

$$\begin{aligned} |D_{l_a}(\alpha, \tau, \eta)|^8 &= |(\alpha_0 + a)[\alpha_0 + a - i(2\eta + \zeta)]^3|^2 \\ &= (\alpha_0 + a)^2[(\alpha_0 + a + \zeta_2)^2 + (2\eta + \zeta_1)^2]^3 \\ &\geq \alpha_0^2[\alpha_0^2]^3 = \alpha_0^8 \end{aligned} \quad (4.30)$$

or

$$|D_{l_a}(\alpha, \tau, \eta)| \geq \alpha_0. \quad (4.31)$$

Furthermore, comparison with Hepp shows that $D_{l_a}(\alpha, \tau, 0) = D_{l_a}(\alpha, \tau)$. Thus the properties (4.19)–(4.21) are invariant under (a).

Finally we apply $\bar{M}$ or $(1 - \bar{M})$, recalling (4.17). These operate on functions of the form

$$P(p) \exp [\sum (iA''pp - iV''\mathbf{pp} - W''\mathbf{pp}]. \quad (4.32)$$

$\bar{M}$ brings down factors $A''pp$ and $V''\mathbf{pp}$ whereas $(1 - \bar{M})$ changes (4.32) to

$$\frac{1}{j!} \int_0^1 d\tau (1 - \tau)^j \left(\frac{\partial}{\partial \tau}\right)^{j+1} P(\tau p) \times \exp [\sum (i\tau^2 A''pp - i\tau^2 V''\mathbf{pp} - W''\mathbf{pp}], \quad (4.33)$$

which brings down factors of $\tau A''pp$ and $\tau V''\mathbf{pp}$. In either case we get a term involving only $A''pp$ and this is absorbed into $P(p)Q(\alpha, \tau)R(\alpha, \tau)$. The remaining terms are absorbed into $P(p, \mathbf{p})Q(\alpha, \tau)R(\alpha, \tau, \eta)$. This completes our outline of the derivation of (4.18)–(4.26).

We can now show that the integral of (4.18) over (2.7) with $r = 0$ converges, as $\eta \downarrow 0$, to the integral of (2.8) over (2.7) with $r = 0$ and that the integral of (4.23) over (2.7) with $r = 0$ converges to zero. This will imply that

$$\lim_{\eta \downarrow 0} I^{\eta, \epsilon} = I^{0, \epsilon} = \lim_{r \downarrow 0} I^{r, \epsilon}, \quad (4.34)$$

from which (1.8) will follow. In view of (4.21), (4.23), and (4.26), we need only show that the η limits can be taken inside the α integrals of (4.18) and (4.24). This is easily shown to follow from Hepp's result that (2.8) is integrable over (2.7) for $r \downarrow 0$ whenever (2.9)–(2.13) hold and $R^i(\alpha, \tau)$ is a monomial of sufficiently high order in the $A^{(i)}_{ab}(\alpha, \tau)$, which satisfy (2.14).

First we consider (4.18). Taking (4.19), (4.20), and (4.22) into account, we see that it is bounded by the function

$$\delta(\sum p_a) e^{-\epsilon \Sigma \alpha_l} \int_0^1 d\tau_1 \cdots d\tau_N \, |P(p) Q(\alpha, \tau)| \times \left[\prod_{i=1}^N D^{(i)}(\alpha) \, |R^{(i)}(\alpha, \tau)|\right], \quad (4.35)$$

where

$$D^{(i)}(\alpha) = \prod_{l \in \mathfrak{L}_i} \alpha_l^{-2}. \quad (4.36)$$

It follows from the (absolute) integrability of (2.8), in particular whenever $D_l(\alpha, \tau) \geq \alpha_l$, that (4.35) is integrable over (2.7), $r = 0$. Hence, by the dominated convergence theorem, the η limit can be brought inside the integral of (4.18) over (2.7), $r = 0$, to give the integral of (2.8).

Next we consider (4.24). By virtue of (4.19), (4.20), (4.22), and (4.26), we see that it is bounded by the function

$$\delta(\sum p_a) e^{-\epsilon \Sigma \alpha_l} \int_0^1 d\tau_1 \cdots d\tau_N \, |P(p, \mathbf{p}) Q(\alpha, \tau)| \times \left[\prod_{i=1}^N D^{(i)}(\alpha) \, |\mathcal{R}^{(i)}(\alpha, \tau)|\right], \quad (4.37)$$

where $\mathcal{R}^{(i)}(\alpha, \tau)$ differs from $R^{(i)}(\alpha, \tau, \eta)$ by having its factors $V^{(i)}_{ab}(\alpha, \tau, \eta)$ replaced by $5\mathcal{A}^{(i)}(\alpha, \tau)$. It again follows from the (absolute) integrability of (2.8), in particular whenever $D_l(\alpha, \tau) \geq \alpha_l$ and (2.14) is satisfied, that (4.37) is integrable over (2.7), $r = 0$. Thus the dominated convergence theorem tells us that the η limit can be brought inside the integral of (4.24) over (2.7), $r = 0$, to give zero. We thus establish (1.13).

The simplicity of the step in the preceding paragraph is a consequence of our use of the $\bar{M}$ operator. Had we used instead the M operator, the V's in $R^{(i)}(\alpha, \tau, \eta)$ would have been replaced by U's, which do not satisfy an inequality like that in (4.26). Thus the existence of the $\eta \downarrow 0$ limit of the integral of (4.24) would depend on more detailed properties of (4.24) than we have used.

ACKNOWLEDGMENTS

The author is deeply indebted to Professor P. Federbush for the suggestion of this problem and for numerous helpful discussions. He thanks Professor C. Kacser for reading the manuscript.

This research was performed at the Laboratory for Nuclear Science of the Massachusetts Institute of Technology. This publication is supported by the U.S. Air Force on grant AFOSR 500-66.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 5 MAY 1967

Class of Nonanalytic Perturbations in Quantum Mechanics

WILLIAM M. FRANK
U.S. Naval Ordnance Laboratory, Silver Spring, Maryland
and
Bar-Ilan University, Israel

(Received 5 October 1966)

It is shown that the resolvent operator for the Hamiltonian which is the sum of the harmonic oscillator Hamiltonian H_0 and a polynomial perturbation $gP(x)$ of degree exceeding two (g is a coupling constant) is not expressible as a convergent power series in g. The source of this nonanalyticity is the failure of the anharmonic perturbation operator to be small in norm compared to H_0. The nature of the singularity at $g = 0$ is conjectured. The result makes clear that the divergence of the Ward–Hurst–Thirring model has nothing to do with the "difficulties" of field theory which are related to the infinite number of degrees of freedom of a field.

THE purpose of this paper is to indicate the phenomenon of nonanalytic perturbations (i.e., perturbed Hamiltonians whose eigenstates do not possess convergent expansions in the unperturbed states) in simple nonrelativistic quantum mechanics, as well as to indicate its relevance to this phenomenon as discussed in connection with field theories. It is shown that the expansion of an anharmonic oscillator system in harmonic oscillator states is never convergent. The result may be relevant to the physics of anharmonic systems *per se*, as in connection with crystal problems, but this aspect is not pursued here. Our purpose is rather to emphasize that these results clarify the significance of the well-known results of the divergence of the perturbation power series (p.s.) for a nonlinearly self-coupled boson field (the Ward–Hurst–Thirring model),[1] as well as a general three boson field coupling. Consequently the inference drawn by some[2] for general relativistic field theories from the property of the Ward–Hurst–Thirring model is unwarranted. A number of other simple nonanalytic models have been studied in recent literature,[3] also carrying with them insinuations regarding field theory. These models involve the basic type of structure studied in this paper, and can be understood in the light of these results. They do not pertain to field theory as suggested.

Another interesting conclusion that is suggested in Sec. IV is the phenomenon of "analyticity flip," i.e., that while there is no analyticity in g there is "essential analyticity" in $1/g$. [$f(z)$ is said to be "essentially analytic" at $z = 0$ if one can find α, β such that $z^\alpha f(z^\beta)$ is analytic at $z = 0$.] Such an analyticity flip has been noted in connection with a solvable field theory model.[4]

This paper presents a preliminary study of this situation and leads to interesting questions which should merit further investigation. The quantity studied is the resolvent operator, and the weak divergence (which is the *strongest* type of divergence) of its p.s. is demonstrated by combinatorial methods.

The result is first demonstrated for the ground-state expectation value of a class of anharmonic perturbations in Sec. II. The results are extended to all anharmonic perturbations and to all matrix elements in Sec. III. The nonanalytic behavior is classified in Sec. IV. In Sec. V the results are discussed in the context of Hilbert space theory. Other approaches and ramifications are discussed in Sec. VI.

I. THE PERTURBATION EXPANSION

We consider the one-dimensional system defined by the Hamiltonian

$$H = H_0 + gV. \tag{1}$$

The unperturbed Hamiltonian H_0 is that of a harmonic oscillator

$$H_0 = \tfrac{1}{2}(p^2 + \omega^2 q^2 - \omega) \tag{2}$$

($\hbar = 1$) with eigenstates $|m)$ such that

$$(H_0 - m\omega)\,|m) = 0 \quad (m = 0, 1, 2, \cdots). \tag{3}$$

The state $|m)$ is termed the "mth level" of the unperturbed oscillator [though it is actually the $(m + 1)$th level]. gV is an anharmonic perturbation of the form

$$gV = gP(x), \tag{4}$$

[1] C. A. Hurst, Proc. Cambridge Phil. Soc. **48**, 625 (1952); W. Thirring, Helv. Phys. Acad. **26**, 33 (1953); A. Petermann, Arch. Sci. Phys. Nat. **6**, 5 (1953); R. Utiyama and T. Imamura, Progr. Theoret. Phys. (Kyoto) **9**, 431 (1953); W. M. Frank, J. Math. Phys. **5**, 363 (1963).

[2] See, e.g., S. Frautschi, Progr. Theoret. Phys. (Kyoto) **22**, 882 (1952), p. 882 footnote*; A. Peres, J. Math. Phys. **4**, 332 (1963); G. A. Baker and R. Chisholm, J. Math. Phys. **7**, 1900 (1966).

[3] A. Peres, J. Math. Phys. **4**, 332 (1963); F. Calogero, Nuovo Cimento **30**, 916 (1963).

[4] W. M. Frank, Nuovo Cimento **38**, 1077 (1965).

where $P(x)$ is a polynomial in x of degree higher than two, and g is the coupling constant. The result is demonstrated for the more general form of V, a multinomial in a and $a^\dagger$ [the raising and lowering operators, defined in Eqs. (10a), (10b)], but we loosely speak of V as a polynomial in x. Only Hermitian V is considered, though this limitation is really not essential. The expansion is in powers of g. For $P(x)$ a linear or quadratic polynomial (or multinomial in a and $a^\dagger$) the power series respectively converge with infinite and finite radius of convergence. This fact follows from the convergence theorems discussed by Speisman.[5] It can also be derived by application of the Rellich–Kato theorem on regular perturbations,[6] or by the methods of this paper which have already been applied to a study of convergence in an earlier paper.[7]

The quantities whose perturbation expansion we consider are the matrix elements of the resolvent operator

$$R(E) = (E - H)^{-1} \tag{5}$$

taken between unperturbed eigenstates. We first construct the proof for the case of the oscillator ground-state expectation value of $R(E)$. Other matrix elements are discussed in Sec. III. The perturbation series is constructed by expansion of the resolvent operator $R(E)$ in powers of g which entails iteration of the unperturbed resolvent

$$r(E) = (E - H_0)^{-1} \tag{6}$$

(i.e., energy denominators). For this reason the values of E should be restricted to lie in the resolvent set (i.e., not in the spectrum) of H_0. Since $R(E)$ has a pole in E at a point in the spectrum of H, it is advisable to choose E in the resolvent set of H. For g real, complex E would be in this set, and if the spectrum of H is bounded from below sufficiently negative E should do. For purposes of simplicity negative E is chosen, though all the steps follow with only an obvious minor modification for complex E. These modifications in a more complicated form are made in Ref. 7 and discussed in Sec. III. While some of the anharmonic perturbations considered lead to spectra with no lower bound, the reasoning makes clear that the expansions in g are nonanalytic for complex E. In fact, the unboundedness of the spectrum from above and below in these cases is a manifestation of the nonanalyticity we are proving, since the perturbation can in no sense be small if it changes a spectrum (the unperturbed one) with a finite lower bound into one which has none. The nonanalyticity also follows for a class of perturbations leading to spectra which have a lower bound.

We write $-E = \lambda > 0$. The perturbation series for

$$R(-\lambda) = -\langle 0|\, 1/(\lambda + H)\, |0\rangle \tag{7}$$

is given by the formal series

$$R(-\lambda) \equiv \sum_{n=0}^{\infty} g^n R_n = -\langle 0|\, r(-\lambda) \sum_{n=0}^{\infty} (-g)^n T(-\lambda^n)\, |0\rangle, \tag{8}$$

where

$$T(-\lambda) = -Vr(-\lambda) = V[1/(\lambda + H_0)]. \tag{9}$$

II. POSITIVE ANHARMONIC COUPLING

In this section we study a class of anharmonic couplings, which we term "positive." These are defined in terms of "pure" anharmonic couplings. Anharmonic couplings where $P(x)$ of Eq. (4) is a normally ordered monomial, i.e., of the form $P(x) = :x^k:$ are termed "pure." If there exists a positive constant b such that the matrix elements of $P(x) - b :x^k:$ for some $k \geqslant 3$ are all nonnegative, then $P(x)$ is termed "positive." The demonstration is spelled out in detail for the case of a pure cubic term, though the method lends itself straightforwardly to all pure anharmonic couplings. The combinatorial technique of "words" used in an earlier paper[7] is applied. This involves a diagrammatic analysis based on the particle interpretation of harmonic oscillator states. The "lowering" or "annihilation" operator a is introduced.

$$a = (\tfrac{1}{2}\omega)^{\frac{1}{2}}(x + ip/\omega) \tag{10a}$$

as well as its adjoint "raising" or "creation" operator

$$a^\dagger = (\tfrac{1}{2}\omega)^{\frac{1}{2}}(x - ip/\omega) \tag{10b}$$

which obey $[a, a^\dagger] = 1$. It is assumed that the reader is familiar with the properties of these operators as well as their "field-theoretic" interpretation. In terms of them

$$H_0 = \omega a^\dagger a \tag{11}$$

and for pure cubic coupling

$$V = (2\omega)^{-\frac{3}{2}}(a^{\dagger 3} + a^3 + 3a^{\dagger 2}a + 3a^\dagger a^2) \equiv \sum_{i=1}^{4} V_i, \tag{12}$$

where the operator in parentheses is essentially a normally ordered x^3. The operators V_i $(i = 1, 2, 3, 4)$ are defined respectively according to their order of writing in the parentheses in Eq. (12). Clearly only

[5] G. Speisman, Phys. Rev. **107**, 1180 (1957).

[6] See, e.g., F. Rellich, "Perturbation Theory of Eigenvalue Problems," Lecture Notes, New York University (1953); T. Kato, J. Fac. Sci. Tokyo University **6**, 198 (1951).

[7] W. M. Frank, J. Math. Phys. **3**, 272 (1962).

even-order terms appear in the perturbation series for $R(-\lambda)$. The $2n$th-order term in the expansion has the form

$$g^{2n}R_{2n} = \frac{g^{2n}}{\lambda^2}(0|\, V \frac{1}{\lambda + H_0} V \cdots V \frac{1}{\lambda + H_0} V\,|0) \tag{13}$$

with V appearing $2n$ times. Each V is a sum of four operators V_i $(i = 1, 2, 3, 4)$ as given in Eq. (12), and by distributing this sum in the product in Eq. (13), R_{2n} is expressed as a sum of 16^n "terms." Each such term is positive since all the energy denominators are positive for positive λ and the matrix elements of the V_i are always nonnegative. R_{2n} clearly exceeds the contribution of that single term where the first n operators V (starting from the rightmost) are taken to be $V_1 = (2\omega)^{-\frac{3}{2}}a^{\dagger 3}$, and the remaining n are taken to be $V_2 = (2\omega)^{-\frac{3}{2}}a^3$. This particular term to be denoted by $\tilde{R}_{2n}$ is readily evaluated.

$$\begin{aligned}\tilde{R}_{2n} &= \frac{3\omega(\mu + n)(3n)!}{\lambda^2(72\omega^5)^n}\prod_{m=1}^{n}\frac{1}{(\mu + m)^2} \\ &= \frac{3\omega(\mu + n)(3n)!\,\Gamma^2(\mu)}{\lambda^2(72\omega^5)^n\Gamma^2(\mu + n + 1)},\end{aligned} \tag{14}$$

where $\mu = \lambda/3\omega$. It is quite clear that the sum

$$\begin{aligned}\tilde{R}(-\lambda) &\equiv \sum_{n=0}^{\infty} g^{2n}\tilde{R}_{2n} \\ &= \frac{3\omega\Gamma^2(\mu)}{\lambda^2}\sum_{n=0}^{\infty}\left(\frac{g^2}{72\omega^5}\right)^n \frac{(\mu + n)(3n)!}{\Gamma^2(\mu + n + 1)}\end{aligned} \tag{15}$$

is divergent. In fact it is clear that R_{2n} exceeds all of the other 16^n terms in R_{2n} for λ positive, so that

$$\tilde{R}_{2n} < R_{2n} < 16^n\tilde{R}_{2n} \tag{16}$$

and

$$\tilde{R}(g; -\lambda) < \tilde{R}(g; -\lambda) < R(4g; -\lambda). \tag{17}$$

It is seen that the $2n$ term in the expansion grows like $n!$ (to within geometrical factors). This method applies straightforwardly to pure anharmonic perturbations of the type

$$V = :x:^k \tag{18}$$

and gives the result that the $2n$ term in $R(-\lambda)$ grows (to within geometrical factors) like $n!^{k-2}$. The lack of convergence for g with any particular phase, of course, implies the same for g of larger modulus regardless of phase. The divergence follows for positive anharmonic couplings from the pure case since the nth-order perturbation coefficient for the positive case exceeds the corresponding coefficient for a pure case.

III. GENERAL ANHARMONIC COUPLINGS

In what follows, divergent power series are manipulated as if they were convergent, with the rationale that the relations found would be valid up to an arbitrarily high term if the expansion were terminated at some other appropriately larger finite term. More accurately, the expansions are to be viewed as asymptotic expansions, and the operations performed would then be justified.

In order to establish the divergences of the p.s. for the ground-state expectation value of the resolvent for an arbitrary polynomial, it is sufficient to show that for sufficiently large n, the coefficients of the expansion grow like a positive power of $n!$ despite alternations in sign of the contributions from different "diagrams." The anharmonic coupling V can be expanded (uniquely) in the form

$$V = \sum_{j=0}^{k}\alpha_j :x^j:. \tag{19}$$

The upper limit k of summation in Eq. (19) is termed the "degree" of V. Without loss of generality we take $\alpha_k > 0$. Because of the latter condition, there is an integer p such that

$$(m|\, V\,|n) \geq 0 \quad m; n \geq p \tag{20}$$

and we choose p to be the minimum such integer.

We abbreviate

$$\begin{aligned} r &= (\lambda + H_0)^{-1}; \\ V^n r^{n-1}; &= V(rV)^{n-1} \\ &= VrVrV\cdots V \quad (V\,n \text{ times}).\end{aligned}$$

Let

$$U_0(n) = (0|;\, V^n r^{n-1};\, |0). \tag{21}$$

Let Λ_0 be the rejection operator for the state $|0)$, i.e.,

$$\Lambda_0\,|m) = (1 - \delta_{m,0})\,|m),$$

$$r_0 = \Lambda_0 r = \Lambda_0/(\lambda + H_0)$$

and further define

$$T_0(n) = (0|;\, V^n r_0^{n-1};\, |0) \tag{22}$$

so that the intermediate state $|0)$ is forbidden. It is easy to see that the formal generating functions

$$\mathfrak{U}(z) = \frac{1}{\lambda} + \frac{1}{\lambda^2}\sum_{n=1} U_0(n)z^n, \tag{23}$$

$$\mathfrak{T}(z) = \sum_{n=1} T_0(n)z^n, \tag{24}$$

are related via the mass renormalization relation,

$$\mathfrak{U}(z) = [\lambda - \mathfrak{T}(z)]^{-1}. \tag{25}$$

It may be shown that the coefficient of z^n in $\mathfrak{T}(z)$ grows as $(n!)^{\frac{1}{2}k-1}$ in the sense that

$$0 < \overline{\lim_{n\to\infty}}\,|T_0(n)n!^{-\frac{1}{2}k+1}|^{1/n} < \infty. \tag{26}$$

We call $\frac{1}{2}k - 1$ the "exponent of factorial growth" (e.f.g.) of $\mathfrak{G}(z)$ written e.f.g. $[\mathfrak{G}(z)]$, and also speak of it as the e.f.g. of the sequence $\{T_0(n)\}$. It is easy to see that if ρ is the e.f.g. of $\mathfrak{G}(z)$ then it is also the e.f.g. of $\mathfrak{U}(z)$. From the formal expansion into a power series of the function $[\lambda - \mathfrak{G}(z)]^{-1}$ it is clear, allowing for the possibility of cancellations, that

$$\text{e.f.g. } [\mathfrak{U}(z)] \leq \text{e.f.g. } [\mathfrak{G}(z)]. \tag{27}$$

The inverted relation

$$\mathfrak{G}(z) = \lambda - \mathfrak{U}(z)^{-1} \tag{28}$$

implies by the same reasoning, the reverse inequality. Thus

$$\text{e.f.g. } [\mathfrak{G}(z)] = \text{e.f.g. } [\mathfrak{U}(z)]. \tag{29}$$

We now determine the e.f.g. of $\mathfrak{G}(z)$. Let the operators $V^{(+)}$ and $V^{(-)}$ be defined by

$$\begin{aligned} V^{(+)}|k) &= \begin{cases} 0, & \text{if } k \leq p, \\ V|k), & \text{if } k > p. \end{cases} \\ V^{(-)}|k) &= \begin{cases} 0, & \text{if } k \geq p, \\ V|k), & \text{if } k < p. \end{cases} \end{aligned} \tag{30}$$

Then

$$(n|\, V^{(+)}|m) \geq 0, \quad m \geq p. \tag{31}$$

Define

$$T_p(n) = (p|;\, V^n r_0^{n-1};\, |p), \tag{32}$$

$$I_p(n) = (p|;\, V^{(-)^n} r_0^{n-1};\, |0). \tag{33}$$

Further define

$$M_p^{(+)}(n) = (p|;\, V^{(+)^n} r_0^{n-1};\, |p), \tag{34}$$

$$M_p^{(-)}(n) = (p|;\, V^{(-)^n} r_0^{n-1};\, |p). \tag{35}$$

In $I_p(n)$, $M_p^{(+)}(n)$, and $M_p^{(-)}(n)$ the pth level cannot be an intermediate state. In $M_p^{(+)}(n)$ all the intermediate states are levels higher than p, and in $M_p^{(-)}(n)$ and $I_p(n)$ they are all lower. The arguments of Sec. II can be readily used to establish the e.f.g.'s of the sequences $M_p^{(+)}(n)$ and $M_p^{(-)}(n)$ for large n in terms of the degree k of V.

$$\text{e.f.g. } \{M_p^{(+)}(n)\} = \tfrac{1}{2}k - 1, \tag{36}$$

$$\text{e.f.g. } \{M_p^{(-)}(n)\} = 0. \tag{37}$$

We can decompose $T_p(n)$ according to the number of times that the level p appears as an intermediate state, and write

$$T_p(n) = \sum_{k=1}^{n} T_p(n;\, k), \tag{38}$$

where in $T_p(n;\, k)$ the intermediate state p appears exactly $k - 1$ times. $T_p(n;\, k)$ is thus a sum of products of factors each describing a transition from the state p to the state p through states unequal to p. Such transitions are described by the matrix elements $M_p^{(\pm)}(n)$ of Eqs. (34) and (35). To each term which contributes to $T_p(n;\, k)$ there corresponds a set of k ordered positive integers $\{N_1 \cdots N_k\}$ each denoting the "length," i.e., the number of interactions (i.e., V's) in the p to p transitions, as well as an ordered set of signs $\{s_1, s_2, \cdots, s_k\}$ each s having the value $+$ or $-$, according to whether the corresponding p to p transition is through levels, respectively, above or below p. The set of integers $\{N_1 \cdots N_k\}$ may contain multiplicities, i.e., identical entries, and we denote the unordered set of $\{N_1 \cdots N_k\}$ by $[n_1^{r_1}, n_2^{r_2}, \cdots, n_m^{r_m}]$ meaning that the integer n_1 appears r_1 times in $\{N_1 \cdots N_k\}$ n_2 appears r_2 times $\cdots$ n_m appears r_m times ($r_j > 0$), where m is the number of distinct integers among $N_1 \cdots N_k$. As each p to p transition may be through levels higher or lower than p, one would further decompose each r_j ($j = 1 \cdots m$) into a sum

$$r_j = r_j^{(+)} + r_j^{(-)}$$

such that $r_j^{(+)}$ of the p–p transitions of length r_j are through levels above p, and $r_j^{(-)}$ are through levels below p. Then we can write

$$\begin{aligned} T_p(n;\, k) = k! \sum_{m=1}^{k} \sum_{\substack{n_1 \cdots n_m \\ r_1(\pm) \cdots r_m(\pm)}} & \times \frac{(M_p^{(+)}(n_1))^{r_1^{(+)}}}{(r_1^{(+)})!} \cdots \frac{(M_p^{(+)}(n_m))^{r_m^{(+)}}}{(r_m^{(+)})!} \\ & \times \frac{[M_p^{(-)}(n_1)]^{r_1^{(-)}}}{(r_1^{(-)})!} \cdots \frac{[M_p^{(-)}(n_m)]^{r_m^{(-)}}}{(r_m^{(-)})!} \\ & \equiv k!\, M_p(n:\, k), \end{aligned}$$

where

$$\sum_{j,s=\pm} n_j r_j^{(s)} = n, \quad \sum_{j,s=\pm} r_j^{(s)} = k. \tag{39}$$

Let us now define a number of formal generating functions (to be understood in general as asymptotic series)

$$\mathfrak{I}(z) = 1 + \sum_{n=1}^{\infty} I_p(n) z^n, \tag{40}$$

$$\mathfrak{G}_p(z) = 1 + \sum_{n=1}^{\infty} T_p(n) z^n, \tag{41}$$

and

$$\mathfrak{M}_p^{(\pm)}(z) = 1 + \sum_{n=1}^{\infty} M_p^{(\pm)}(n) z^n, \tag{42}$$

with coefficients defined in Eqs. (32)–(35). Further let

$$T_0^{(-)}(n) = (0|;\, V^{(-)^n} r_0^{n-1};\, |0), \tag{43}$$

i.e., the contribution to $T_0(n)$ arising from intermediate

states all of which are below the level p. Let

$$\mathcal{C}_0^{(-)}(z) = \sum_{n=1}^{\infty} T_0^{(-)}(n) z^n. \tag{44}$$

Then one readily verifies that

$$\mathcal{C}(z) = \mathcal{C}_0^{(-)}(z) + [1/(\lambda + p\omega)]\mathfrak{I}^2(z) + [1/(\lambda + p\omega)^2]\mathfrak{I}^2(z)\mathcal{C}_p(z), \tag{45}$$

where the three expressions on the right, respectively, denote contributions from terms where the state p appears as an intermediate state zero, one or more than one time. It is clear that

$$\text{e.f.g.}\,[\mathfrak{I}^2(z)] = \text{e.f.g.}\,[\mathcal{C}_0^{(-)}(z)] = 0 \tag{46}$$

and we wish to determine e.f.g. $[\mathcal{C}_p(z)]$.

Let us define the double generating function

$$V_p(z:t) = 1 + \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} \frac{T_p(n:k)}{k!} t^k z^n. \tag{47}$$

Then formally

$$\int_0^{\infty} dt e^{-t} V_p(z:t) = 1 + \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} T_p(n:k) z^n = 1 + \sum_{n=1}^{\infty} T_p(n) z^n = \mathcal{C}_p(z), \tag{48}$$

in view of Eq. (38). The quantity $V_p(z:t)$ can, however, be formally summed in view of Eq. (39).

$$V_p(z:t) = 1 + \sum_{n=1}^{\infty}\sum_{k=1}^{n}\sum_{m=1}^{k} \sum_{\substack{n_1 \cdots n_m \\ r_1^{(\pm)} \cdots r_m^{(\pm)} \\ \sum_{j,s} n_j r_j^{(s)} = n \\ \sum_{j,s} r_j^{(s)} = k}} \times \frac{[tz^{n_1} M_p^{(+)}(n_1)]^{r_1^{(+)}}}{(r_1^{(+)})!} \cdots \frac{[tz^{n_m} M_p^{(-)}(n_m)]^{r_m^{(-)}}}{(r_m^{(-)})!}$$

$$= \exp\left[t \sum_{n=1}^{\infty} z^n [(M_p^{(+)}(n) + M_p^{(-)}(n)]\right] \tag{49}$$

and formally

$$\mathcal{C}_p(z) = \int_0^{\infty} dt e^{-t} V_p(z;t) = \int_0^{\infty} dt \exp\left\{-t\left[1 - \sum_{n=1} z^n (M_p^{(+)}(n) + M_p^{(-)}(n))\right]\right\} = \left\{1 - \sum_n z^n [M_p^{(+)}(n) + M_p^{(-)}(n)]\right\}^{-1} = [1 - \mathcal{M}_p(z)]^{-1}, \tag{50}$$

where $\mathcal{M}_p(z) = \mathcal{M}_p^{(+)}(z) + \mathcal{M}_p^{(-)}(z)$ as defined in Eq. (42). From the considerations leading to Eq. (29) we conclude that

$$\text{e.f.g.}\,[\mathcal{C}_p(z)] = \text{e.f.g.}\,[\mathcal{M}_p(z)] = \text{e.f.g.}\,[\mathcal{M}_p^{(+)}(z) + \mathcal{M}_p^{(-)}(z)] = \tfrac{1}{2}k - 1$$

the latter following from Eqs. (36), (37). This through Eqs. (45), (29) leads to the result

$$\text{e.f.g.}\,\{U_0(n)\} = \tfrac{1}{2}k - 1. \qquad \text{Q.E.D.}$$

The extension of these results to complex λ and to other than ground-state expectation values of the resolvent is quite straightforward. For $\lambda = \lambda_r + i\lambda_i$ ($\lambda \neq 0$) we are guaranteed the nonvanishing of energy denominators. The upper bounds of Secs. II and III remain valid if one takes into account that λ_r, if it is negative, will cancel a bounded part of the intermediate state energies, an effect which is inconsequential to the contribution from highly excited intermediate states, which determine the e.f.g. of the perturbation series. It is not difficult, though somewhat involved, to show that the complex energy denominators do not provide cancellations lowering the e.f.g. of the power series. As the energy of the intermediate excited states become large, the complex phase angle of each amplitude becomes small due to the bounded imaginary part of the energy denominator, and phase cancellations have no effect on the e.f.g. For other than ground-state expectation values, the same result for the e.f.g. is valid. Excited initial and final states only shift the energy denominators by a bounded quantity, which does not affect anything.

IV. BEHAVIOR NEAR $g = 0$

What can be said about the nature of the singularity near $g = 0$? We have seen that the formal series of Eq. (8), which we denoted by $R(g:\lambda)$ is now simply $R(g)$, has e.f.g. $\tau \equiv \frac{1}{2}k - 1$. The coefficients R_n of the power series will have a large n behavior of the form

$$R_n = p(n)n!^r \quad \text{or} \quad R_n \sim n!^r, \tag{51}$$

where

$$\overline{\lim_{n\to\infty}} \frac{|\ln p(n)|}{n \ln n} = 0. \tag{52}$$

We shall simple-mindedly conjecture the singularity suggested by the asymptotic series in Eq. (51). A proper study of the question will be carried out elsewhere.

Let us write formally

$$R_n = \Gamma(\sigma n + 1)c(n), \tag{53}$$

where $\tau < \sigma$ which is otherwise arbitrary. According to Eq. (51)

$$c(n) \sim 1/n!^{\sigma - r} \tag{54}$$

so that $C(g)$ is an entire function of g, where

$$c(g) \equiv \sum_{n=0}^{\infty} c(n) g^n.$$

From Eq. (54) we conclude according to standard theorems on entire functions that the "exponential order"[8] of $C(g)$ is $\rho \equiv (\sigma - \tau)^{-1}$. Substituting formally

$$\Gamma(\sigma n + 1) = \int_0^\infty du e^{-u} u^{\sigma n}$$

into

$$R(g) = \sum_{n=0} \Gamma(\sigma n + 1) c(n) g^n, \tag{55}$$

one obtains

$$R(g) = \int_0^\infty du e^{-u} C(u^\sigma g). \tag{56}$$

The knowledge of the exponential order of $C(g)$ tells us that there is an inequality like

$$|C(u^\sigma g)| < \exp\left[A\,|g|^{1/\sigma-\tau} \cdot u^{\sigma/\sigma-\tau}\right] \tag{57}$$

(A some constant). Under these conditions the integral in Eq. (57) would not be expected to converge, a reflection of the anticipated singularity of $R(g)$ near $g = 0$. However, the inequality Eq. (57) is too general. Frequently the function $C(g)$ either has sectors wherein it approaches zero exponentially, or is a sum of such functions (as, e.g., cosh z). According to results[9] on entire functions, in these sectors

$$C(z) \sim \exp\left(-B\,|z|^\rho\right) \tag{58}$$

if the exponential order is ρ. For g within such an "attenuating sector" the integral in Eq. (56) can be evaluated by expanding e^{-u} in a power series

$$\begin{aligned} R(g) &= \sum_{n=0}^\infty \frac{(-)^n}{n!} \int_0^\infty du u^n C(u^\sigma g) \\ &\sim \sum_{n=0}^\infty \frac{(-)^n}{n!} \int_0^\infty du u^n \exp\left(-[B\,|g|^{1/\sigma-\tau} u^{\sigma/\sigma-\tau}]\right) \end{aligned} \tag{59}$$

for large g. The latter quantity is evaluated straightforwardly, and results in

$$R(g) \sim \sum_{n=0}^\infty \frac{(-)^n}{n!} [B^{\sigma-\tau}\,|g|]^{-(n+1)/\sigma} \Gamma\left[(n+1)\left(1 - \frac{\tau}{\sigma}\right)\right], \tag{60}$$

which is a convergent expansion in negative non-integral powers of $|g|$. It thus can be analytically extended via this power series to the whole finite plane of $g^{-\sigma}$, and is thus "essentially analytic" in the sense indicated earlier. Its exponential order is σ/τ, which implies an essential singularity near $g = 0$ of the form

$$R(g) \sim \exp\left(Kg^{-1/\tau}\right) \tag{61}$$

(independent of σ as it should be since σ is arbitrary). We thus conjecture that the singularity of the resolvent $R(g; E)$ as a function of g is essential at $g = 0$ and of the nature given in Eq. (61). Independent calculations along different lines are being carried out to verify this point.

V. HILBERT SPACE CONSIDERATIONS

The results found are not at all difficult to understand. They simply say that perturbation theory will not work if the perturbing operator is not in some appropriate sense small compared to the unperturbed term. Such sufficiency criteria for the perturbation treatment of operators and spectra in Hilbert space are known and they in fact break down for the present example, suggesting the divergences presently found. The writer is not familiar enough with the necessary conditions which may possibly directly imply the results obtained. The important sufficiency condition applicable to problems of this kind is the Rellich–Kato theorem.[6] Definitions of the standard terms can be found in standard works on Hilbert space theory.[10]

Definition: λ is said to be in the resolvent set of a closed linear operator T, if $(T - \lambda)^{-1}$ is a bounded operator.

Definition: If $T(z)$ is a closed linear operator function of z then $T(z)$ is said to be regular in z if for some λ in the resolvent set of $T(0)$, $R(z)$ can be expanded for sufficiently small z in a power series in z which converges in the operator norm.[11]

Rellich–Kato Theorem: If $\|Vf\| < a\,\|H_0 f\| + b\,\|f\|$, ($a$, b constants) for all f in an everywhere dense subset of the Hilbert space, then

$$H(z) = H_0 + zV \tag{62}$$

is regular for $|z| < a^{-1}$.

Applied to our case it is clear that the conditions are not fulfilled. In fact $-\lambda = E$ negative, is in the resolvent set of H_0 and we have seen that

$$R_\lambda(g) = -1/(\lambda + H)$$

has a zero radius of convergence in the weak norm, i.e.,

$$(m|\, R_\lambda(g)\,|n) \tag{63}$$

[8] An entire function $f(z)$ is said to be of exponential order $\alpha > 0$, if α is the greatest lower bound of real numbers γ such that one can find a K for which

$$|f(z)| < \exp|Kz^\gamma|.$$

[9] See, e.g., E. C. Titchmarsh, *The Theory of Functions* (Oxford University Press, New York, 1939), 2nd ed., Sec. 8.7.

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

[11] The norm $\|A\|$ of the operator A is defined as

$$\|A\| = \max \|Af\|/\|f\|,$$

where f ranges over the domain of A. $A_n \Rightarrow A$ in the operator norm if $\|A_n - A\| \to 0$.

diverges for all m, n. It follows *a fortiori* that $R_\lambda(g)$ is divergent in the operator norm (since operator convergence implies weak convergence) for such λ, and this by the theorem implies divergence for all λ. The Rellich–Kato theorem is the basis of convergence criteria in quantum mechanics discussed by Speisman.[5]

VI. DISCUSSION

The result obtained is basically transparent. It is related to the result of Baym[12] concerning the lack of a lower bound in the spectrum of odd-order boson couplings. The Rellich–Kato theorem implies that a regular perturbation on a spectrum bounded from below, must also have a lower bound. The nonanalyticity we find also applies, however, to potentials of the form $+gx^4$ which do have a lower bound. The nonanalyticity for $+gx^4$ derives, however, from the lack of a lower bound for $-gx^4$, as suggested by Dyson.[13]

There are a number of further suggestions and implications of this result which are of interest. First of all as to method. This problem was studied via the creation–annihilation operator formalism. It is also related to other methods in the theory of differential equations, which have not been exploited. Thus by going over into the momentum representation $X = i\,d/dp$, the Schrödinger equation for a polynomial potential becomes in the neighborhood of $g = 0$ a singularly perturbed equation, i.e., one where the perturbation changes the order (i.e., the order of the highest derivative) of the equation. The machinery of such problems is not familiar to the author, but such machinery exists and may lead to further information on the structure of such theories. In another approach the variable change $X = \xi^{-1}$ can be made, leading to a strong singularity in the neighborhood of the origin. Some of the methods applied to singular potential problems may be applicable. Thirdly, the traditional WKB method should be applicable, and such an investigation will appear in a separate article.

The considerations of the Rellich–Kato theorem that $\|V\|$ is never smaller than $\|H_0\|$, lead one to look at the problem as if H_0 were the perturbation on V suggesting a convergent expansion in negative powers of g. This phenomenon of "analyticity flip" has been noted by the author in connection with a field theory model.[4] One in fact finds from the WKB approach that this is true. The results will be reported separately.

The original motivation for this problem was to show that the oft-derived divergences of the perturbation expansion for multiboson interactions (e.g., φ^3) has nothing whatsoever to do with field theory. In a field theory, infinitely many oscillators are anharmonically coupled, while we see that the anharmonic coupling of one single oscillator is also not analytic in the coupling constant. The divergence of various other field theory models[3] derives from the considerations presented here. The analyticity flip which follows from other arguments very likely applies to these field theory models and suggests a search for construction of solutions of such field theories via strong coupling approaches.

[12] G. Baym, Phys. Rev. **117**, 886 (1960).
[13] F. J. Dyson, Phys. Rev. **85**, 631 (1952).

Double Structure of a Combined Internal and Space–Time Symmetry Group and Mass Splitting

OWEN FLEISCHMAN* AND JOHN G. NAGEL†
Department of Physics, Syracuse University, Syracuse, New York

(Received 19 September 1966)

Assuming that the group E, obtained by combining the Poincaré group P with a *general* internal symmetry group S, has a direct-product-like structure as occurs in the McGlinn theorem, the conditions for the existence of an isomorphism of E with a nontrivial semidirect-product-like structure are examined. It is found that a nontrivial double structure can exist if and only if S contains an *internal* Lorentz group L, P or their covering groups $\bar{L}$, $\bar{P}$. Mass splitting then occurs if the Poincaré group of the semidirect-product-like structure is identified as the physical space–time group.

1. INTRODUCTION

THE problem of combining the (proper orthochronous) Poincaré group P with an internal symmetry Lie group S to form a larger Lie group E has been considered by many authors[1,2] who reached the conclusion that under rather general conditions $\mathfrak{E}$ is the direct sum

$$\mathfrak{E} = \mathfrak{S} + \mathfrak{P}, \tag{1.1}$$

where $\mathfrak{P}$, $\mathfrak{S}$, and $\mathfrak{E}$ are the Lie algebras of P, S, and E.

It was then concluded that the mass operator $m^2 = p_\mu p^\mu$, where p_μ is the four-momentum, must be degenerate within an irreducible representation of E, which then cannot contain particles of different mass.

However, it is quite possible that the $\mathfrak{E}$ given by Eq. (1.1) might also be isomorphic to a nontrivial semidirect sum with $\mathfrak{S}$ as the ideal (invariant Lie subalgebra), i.e.,

$$\mathfrak{E} = \mathfrak{S} + \mathfrak{P} \cong \mathfrak{S} \boxplus \mathfrak{P}',$$

where

$$\mathfrak{P}' \cong \mathfrak{P}. \tag{1.2}$$

This might give rise to mass splitting, if one identifies $\mathfrak{P}'$ with the physical Poincaré Lie algebra.

The possibility of double structure for $\mathfrak{E}$ was considered by Ottoson, Kihlberg, and Nilsson[3] (OKN) for the case of $\mathfrak{S}$ semisimple and compact. OKN and Kihlberg[4] discuss the possibility of mass splitting arising from double structure.

In this article we consider the problem of double structure of E for the case of a *general* internal symmetry group S, from a global point of view.

In Sec. 3 we find the condition under which a nontrivial double structure can exist. This condition, as given in Theorems A and B, is that $S \supset \bar{L}, L, \bar{P}$, or P, where L is the Lorentz group and $\bar{L}$, $\bar{P}$ are the covering groups of L, P, i.e., that S contains an *internal* Lorentz or Poincaré group or their covering groups, all of which are noncompact.

In Sec. 2 we review the theory of group extensions needed in Sec. 3.

In Sec. 4 we give in Lemmas A and B, the conditions for the double structures to be of physical interest.

In Sec. 5 we analyze the double structures and find in which cases they give rise to mass splitting.

In Sec. 6 we discuss in some detail the articles in Refs. 2–4 in order to clarify them. We point out that Theorems A and B rule out the possibility for a nontrivial double structure for the case of compact S, which is the problem considered in Refs. 3 and 4.

Finally, in Sec. 7 we discuss a simple example of nontrivial double structure.

2. GROUP EXTENSIONS

We first need some mathematical formalism from the theory of group extensions.

Definition: E is a group extension of Q by K, if K is an invariant subgroup of E, i.e., if $K \lhd E$ such that the factor group

$$E/K \cong Q.$$

Let

$$\varphi: G_1 \to G_2 \quad \text{or} \quad G_1 \xrightarrow{\varphi} G_2,$$

denote a single-valued mapping (but, in general, *not* a homomorphism) of a group G_1 *into* a group G_2 and let

$$\theta: G_1 \to G_2 \quad \text{or} \quad G_1 \xrightarrow{\theta} G_2$$

denote a homomorphism θ of G_1 *into* G_2, and let Im θ and Ker θ denote the image and the kernel of θ, respectively.

* Present address: Department of Physics, Boston University, Boston, Massachusetts.
† Present address: The Niels Bohr Institute, Copenhagen, Denmark.

[1] W. D. McGlinn, Phys. Rev. Letters **12**, 467 (1964); F. Coester, M. Hammermesh, and W. D. McGlinn, Phys. Rev. **135**, B451 (1964); M. E. Mayer, H. J. Schnitzer, E. C. G. Sudarshan, R. Acharya, and M. Y. Han, *ibid.* **136**, B888 (1964); C. W. Gardiner, Phys. Letters **11**, 258 (1964).
[2] E. C. G. Sudarshan, J. Math. Phys. **6**, 1329 (1965).
[3] U. Ottoson, A. Kihlberg, and J. Nilsson, Phys. Rev. **137**, B658 (1965).
[4] A. Kihlberg, Nuovo Cimento **37**, 217 (1965).

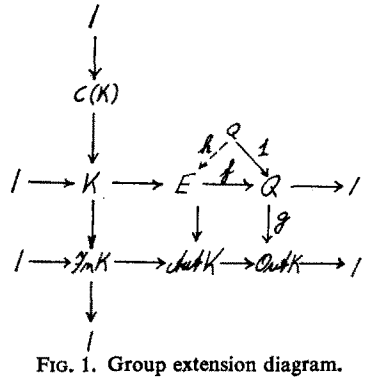


FIG. 1. Group extension diagram.

Definition: A sequence of homomorphisms

$$\cdots \longrightarrow G_{i-1} \xrightarrow{f_i} G_i \xrightarrow{f_{i+1}} G_{i+1} \longrightarrow \cdots,$$

is exact if

$$\text{Im } f_i = \text{Ker } f_{i+1} \quad \forall i.$$

An extension E of Q by K can then be symbolized by the exact sequence

$$1 \to K \to E \to Q \to 1,$$

where 1 denotes the trivial group of only one element.

This exact sequence can be implemented to the usual group extension diagram of commutative exact sequences (see, e.g., Ref. 5) shown in Fig. 1, where $C(K)$ is the center of K, Aut K is the group of automorphisms of K, In K and Out K the groups of inner and outer automorphisms of K, and h is a 1:1 mapping (not a monomorphism) such that

$$f \circ h = 1$$

the identity homomorphism.

The group extension problem is characterized by a triple $(Q, K, g: Q \to \text{Out } K)$ and can have none, one or more than one solution E for a given triple.

Definition: The extension is *central* if $g: Q \to$ Out K is the trivial homomorphism of Q onto the unit element of Out K, i.e., if $g = 0$.

For a central extension the group extension diagram of Fig. 1 reduces to the diagram shown in Fig. 2, where the mappings t and h' are defined later.

In Fig. 1 the mapping h defines, by the equation

$$h(q_1)h(q_2) = \omega(q_1, q_2)h(q_1q_2),$$

the factor system $\omega(q_1, q_2)$ of h which fulfills

$$\omega(q_1, q_2) \in K,$$

$$\omega(q_1, 1) = 1 = \omega(1, q_2),$$

and

$$\omega(q_1, q_2)\omega(q_1q_2, q_3) = q_1[\omega(q_2, q_3)]\omega(q_1, q_2q_3), \quad (2.1)$$

where

$$q[k] \equiv h(q)sh(q)^{-1}.$$

[5] L. Michel, in *Brandeis Summer Lectures in Theoretical Physics 1965* (Gordon and Breach Science Publishers, New York, 1966).

The elements e of E can be written as the ordered pair $e = (k, q)$, where $k \in K$ and $q \in Q$ with the group multiplication law

$$(k_1, q_1)(k_2, q_2) = (k_1q_1[k_2]\omega(q_1, q_2), q_1q_2). \quad (2.2)$$

In the case when E is a central extension, h can be chosen (see, e.g., Ref. 5) such that Eq. (2.2) takes the form

$$(k_1, q_1)(k_2, q_2) = (k_1k_2\omega(q_1, q_2), q_1q_2), \quad (2.3)$$

where

$$\omega(q_1, q_2) \in C(K)$$

and fulfills

$$\omega(q_1, q_2)\omega(q_1q_2, q_3) = \omega(q_2, q_3)\omega(q_1, q_2q_3). \quad (2.4)$$

In the case when E is a semidirect product, h can be chosen such that

$$\omega(q_1, q_2) = 1,$$

i.e., such that Eq. (2.2) takes the form

$$(k_1, q_1)(k_2, q_2) = (k_1q_1[k_2], q_1q_2). \quad (2.5)$$

For the case when E is a direct product, h can be chosen such that Eq. (2.2) takes the form

$$(k_1, q_1)(k_2, q_2) = (k_1k_2, q_1q_2). \quad (2.6)$$

3. DOUBLE STRUCTURE

1. Corresponding to δ given by Eq. (1.1) we have globally either the direct product

$$E = S \times \bar{P}, \quad (3.1a)$$

or the direct-product-like structure

$$E = (S \times \bar{P})/C_2. \quad (3.1b)$$

Here C_2 is the cyclic group of order two whose elements are $[1, (0, 1)]$ and $[\xi, (0, -1)]$, i.e.,

$$C_2 = \{[1, (0, 1)], [\xi, (0, -1)]\} \subset S \times \bar{P},$$

where

$$(0, -1) \in \bar{L}.$$

Also

$$\xi^2 = 1,$$

i.e.,

$$\xi \in {}_2C(S),$$

the subgroup of $C(S)$ of elements of order two. Equation (3.1b) includes, for $\xi = 1$, the case $E = S \times P$.

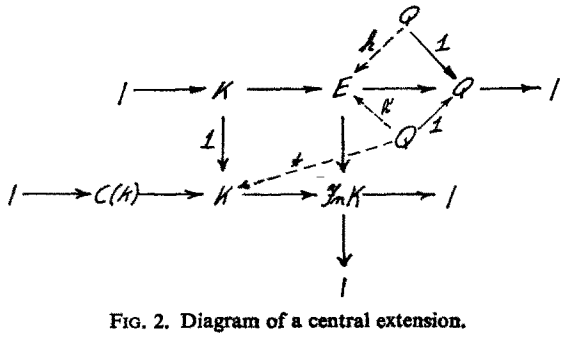


FIG. 2. Diagram of a central extension.

The E given by Eq. (3.1a) is a covering group of the E given by Eq. (3.1b).

The global statements of double structure corresponding to the Lie algebraic equation (1.2) are that the E defined by Eq. (3.1a) be isomorphic to a nontrivial semidirect product with $S \lhd E$

$$E = S \times \bar{P} \cong S \boxtimes \bar{P}', \tag{3.2a}$$

where

$$E/S \cong \bar{P}' \cong \bar{P}, \tag{3.3a}$$

or that the E defined by Eq. (3.1b) be isomorphic to a nontrivial semidirect-prcduct-like structure.

$$E = (S \times \bar{P})/C_2 \cong (S \boxtimes \bar{P}')/C_2', \tag{3.2b}$$

where

$$E/S \cong \bar{P}'/C_2' \cong P' \cong P \cong \bar{P}/C_2. \tag{3.3b}$$

Here

$$C_2' = \{[1, (0, 1)], [\xi', (0, -1)]\} \subset S \boxtimes \bar{P}',$$

where

$$\xi' \in {}_2C(S).$$

Equation (3.2b) includes, for $\xi = 1$, the case

$$E = S \times P \cong (S \boxtimes P')/C_2'$$

and, for $\xi' = 1$, the case

$$E = (S \times \bar{P})/C_2 \cong S \boxtimes P'.$$

2. From Eqs. (3.3) and from the definition of a central extension or from Eq. (2.3) it is clear that E given by Eq. (3.1a) or by Eq. (3.1b) is a central extension of $\bar{P}$ or P, respectively, by S. In order to find the conditions under which the double structures (3.2) exist, we therefore rewrite the group multiplication law (2.3) of a central extension.

Let us consider an arbitrary mapping

$$t: Q \dashrightarrow K,$$

as shown in Fig. 2. Then

$$\begin{aligned} t(q_1)t(q_2) &\equiv \bar{\omega}(q_1, q_2)t(q_1q_2) \\ &\equiv \omega'(q_1, q_2)\omega(q_1, q_2)^{-1}t(q_1q_2) \\ &= \omega'(q_1, q_2)t(q_1q_2)\omega(q_1, q_2)^{-1}, \end{aligned} \tag{3.4}$$

which defines $\bar{\omega}(q_1, q_2)$ and where

$$\omega'(q_1, q_2) \equiv \bar{\omega}(q_1, q_2)\omega(q_1, q_2). \tag{3.5}$$

We now map the elements (k, q) of E onto $(k, q)' \equiv (kt(q), q)$. Defining

$$\begin{aligned} q[k] &\equiv t(q)kt(q)^{-1} \\ &= t(q)h(q)kh(q)^{-1}t(q)^{-1} \\ &= h'(q)kh'(q)^{-1} \end{aligned}$$

we get for the multiplication law (2.3) of the central extension using Eq. (3.4)

$$\begin{aligned} (k_1, q_1)'(k_2, q_2)' &= (k_1t(q_1), q_1)(k_2t(q_2), q_2) \\ &= (k_1t(q_1)k_2t(q_2)\omega(q_1, q_2), q_1q_2) \\ &= (k_1t(q_1)k_2t(q_1)^{-1}t(q_1)t(q_2)\omega(q_1, q_2), q_1q_2) \\ &= (k_1q_1[k_2]\omega'(q_1, q_2)t(q_1q_2), q_1q_2) \\ &= (k_1q_1[k_2]\omega'(q_1, q_2), q_1q_2)', \end{aligned} \tag{3.6}$$

where we now only have to show that $\omega'(q_1, q_2)$ fulfills Eq. (2.1).

From Eq. (3.4) we have

$$\begin{aligned} &q_1[\bar{\omega}(q_2, q_3)]\bar{\omega}(q_1, q_2q_3) \\ &\quad = t(q_1)t(q_2)t(q_3)t(q_2q_3)^{-1}t(q_1)^{-1}t(q_1)t(q_2q_3)t(q_1q_2q_3)^{-1}) \\ &\quad = t(q_1)t(q_2)t(q_1q_2)^{-1}t(q_1q_2)t(q_3)t(q_1q_2q_3)^{-1} \\ &\quad = \bar{\omega}(q_1, q_2)\bar{\omega}(q_1q_2, q_3). \end{aligned} \tag{3.7}$$

From Eqs. (2.4) and (3.7) it then follows that

$$\begin{aligned} &q_1[\omega'(q_2, q_3)]\omega'(q_1, q_2q_3) \\ &\quad = q_1[\bar{\omega}(q_2, q_3)]q_1[\omega(q_2, q_3)]\bar{\omega}(q_1, q_2q_3)\omega(q_1, q_2q_3) \\ &\quad = q_1[\bar{\omega}(q_2, q_3)]\omega(q_2, q_3)\bar{\omega}(q_1, q_2q_3)\bar{\omega}(q_1, q_2q_3) \\ &\qquad\qquad \times \omega(q_1, q_2q_3) \\ &\quad = q_1[\bar{\omega}(q_2, q_3)]\bar{\omega}(q_1, q_2q_3)\omega(q_2, q_3)\omega(q_1, q_2q_3) \\ &\quad = \bar{\omega}(q_1, q_2)\bar{\omega}(q_1q_2, q_3)\omega(q_1, q_2)\omega(q_1q_2, q_3) \\ &\quad = \bar{\omega}(q_1, q_2)\omega(q_1, q_2)\bar{\omega}(q_1q_2, q_3)\omega(q_1q_2, q_3) \\ &\quad = \omega'(q_1, q_2)\omega'(q_1q_2, q_3), \end{aligned}$$

which proves our point.

3. For the E given by Eq. (3.2a) we see that

$$Q = \bar{P}, \quad K = S,$$

while from Eqs. (2.6) and (2.3)

$$\omega(\bar{p}_1, \bar{p}_2) = 1$$

and from Eqs. (2.5) and (3.6)

$$\omega'(\bar{p}_1, \bar{p}_2) = 1.$$

Hence from Eq. (3.5)

$$\bar{\omega}(p_1, p_2) = 1,$$

and it follows from Eq. (3.4) that t is a homomorphism. Since the only proper invariant subgroups of $\bar{P}$ are the translation group T_4 and $C_2 \equiv \{(0, 1), (0, -1)\}$, it follows that $t(\bar{P}) = 1$, $\bar{L}''$, L'', $\bar{P}''$, or P'', purely *internal* groups isomorphic to 1, $\bar{L}$, L, $\bar{P}$, or P, respectively. We therefore have the following condition for double structure for E.

Theorem A: Let

$$E = S \times \bar{P}.$$

Then

$$E \cong S \boxtimes \bar{P}', \quad \bar{P}' \cong \bar{P},$$

if and only if there exists

$$t: \bar{P} \to S, \quad t \neq 0,$$

i.e., if and only if $S \supset \bar{L}''$, L'', $\bar{P}''$ or $P'' \cong \bar{L}$, L, $\bar{P}$ or P, respectively.

The theorem can be illustrated by the diagram shown in Fig. 3.

We notice that there exists a homomorphism $j: \bar{P} \to InS$.

When we turn to the E given by Eq. (3.2b) we have

$$Q = P, \quad K = S,$$

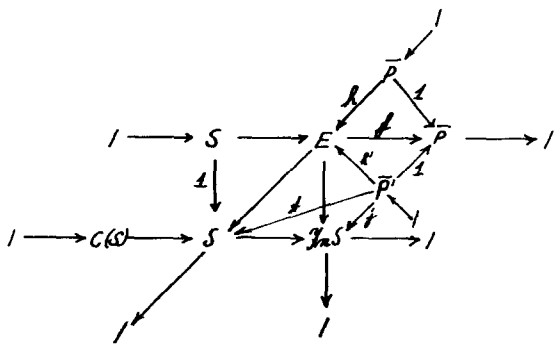


FIG. 3. $E = S \times \bar{P} \cong S \boxtimes \bar{P}'$.

while from Eqs. (2.3) and (3.6)

$$\omega(p_1, p_2) \in {}_2C(S), \quad \omega'(p_1, p_2) \in {}_2C(S).$$

Hence from Eq. (3.5)

$$\bar{\omega}(p_1, p_2) \in {}_2C(S), \tag{3.8}$$

and it follows from Eq. (3.4) that t is not a homomorphism.

However, from Eqs. (3.4) and (3.8) we have

$$\begin{aligned} t(p_1)t(p_2)st(p_2)^{-1}t(p_1)^{-1} &= \bar{\omega}(p_1, p_2)t(p_1p_2)st(p_1p_2)^{-1}\bar{\omega}(p_1, p_2)^{-1} \\ &= t(p_1p_2)st(p_1p_2)^{-1}, \end{aligned}$$

and hence there exists a homomorphism j shown below in the diagram of Fig. 4,

$$j: P \to \text{In } S,$$

such that

$$j = i \circ t$$

where i is defined by

$$1 \to C(S) \to S \xrightarrow{i} \text{In } S \to 1.$$

From Eqs. (3.4) and (3.8) it also follows that there exists a homomorphism,

$$\bar{t}: \bar{P} \to S,$$

such that

$$\bar{t}(\!(0, -1)\!) \in {}_2C(S),$$

and hence such that

$$\bar{t}(\bar{p})t(p)^{-1} = \bar{t}(\bar{p})[(t \circ \ell)(\bar{p})]^{-1} \in {}_2C(S), \tag{3.9}$$

where ℓ is defined by

$$1 \to C_2 \to \bar{P} \xrightarrow{\ell} P \to 1.$$

It now follows from Eq. (3.9) that

$$t(p)st(p)^{-1} = (t \circ \ell)(\bar{p})s[(t \circ \ell)(\bar{p})]^{-1} = \bar{t}(\bar{p})s\bar{t}(\bar{p})^{-1},$$

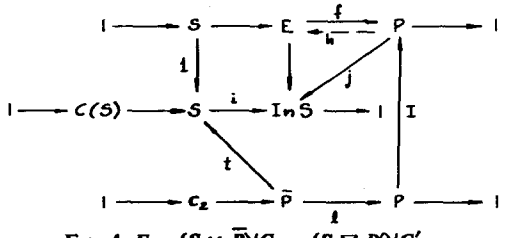


FIG. 4 $E = (S \times \bar{P})/C_2 \cong (S \boxtimes P')/C_2'$.

i.e.,

$$j \circ \ell = i \circ t \circ \ell = i \circ \bar{t},$$

so that we have a commutative diagram which is shown as a part of the diagram in Fig. 4.

We now have the following condition for double structure of E in this case.

Theorem B: Let

$$E = (S \times \bar{P})/C_2,$$

where

$$C_2 = \{[1, (0, 1)], [\xi, (0, -1)]\}, \quad \xi \in {}_2C(S).$$

Then

$$E \cong (S \boxtimes \bar{P}')/C_2', \quad \bar{P}' \cong \bar{P},$$

where

$$C_2' = \{[1, (0, 1)], [\xi', (0, -1)]\}, \quad \xi' \in {}_2C(S),$$

if there exists $\bar{t}: \bar{P} \to S$, $\bar{t} \neq 0$, i.e., if $S \supset \bar{L}''$, L'', $\bar{P}''$, or $P'' \cong \bar{L}$, L, $\bar{P}$, or P, respectively, such that

$$\bar{t}(\!(0, -1)\!) = \xi\xi' \in {}_2C(S).$$

We see that in both cases the condition for double structure is that $S \supset \bar{L}''$, L'', $\bar{P}''$, or P''.

4. PHYSICAL CONDITION ON DOUBLE STRUCTURE

If we want to obtain mass splitting from double structure, it is clear that $\bar{P}$ (or P) which occurs in the direct-product-like structures of Eqs. (3.2) and (3.3) cannot be identified with the physical Poincaré group. We therefore identify $\bar{P}'$ (or P') with the physical Poincaré group, while $\bar{P}$ (or P) describes a nonphysical Poincaré group.

In any theory of interest to particle physics we must ensure exact conservation of the generalized charges, i.e., the electric charge, the baryon, electron and muon numbers, etc., a condition used as a basis for the articles of Refs. 2–4. It follows that the homomorphisms t and $\bar{t}$ must satisfy the following mathematical conditions, as seen from Eq. (3.6).

Lemma A: The double structure

$$E = S \times \bar{P} \cong S \boxtimes \bar{P}'$$

is of physical interest if

$$\bar{p}[s_0] = s_0 \quad \forall \bar{p} \in \bar{P} \quad \forall s_0 \in S_0,$$

i.e., if

$$t(\bar{P}) \subset C_S(S_0),$$

where S_0 is the subgroup of S generated by the charges s_0 and where $C_S(S_0)$ is the centralizer of S_0 in S.

Lemma B: The double structure

$$E = (S \times \bar{P})/C_2 \cong (S \boxtimes \bar{P}')/C_2',$$

is of physical interest if

$$p[s_0] = s_0 \quad \forall p \in P \quad \forall s_0 \in S_0,$$

i.e., if

$$t(\bar{P}) \subset C_S(S_0).$$

5. MASS SPLITTING

We analyze the double structures given in Eqs. (3.2) in view of Theorems A and B to see whether or not they can give rise to mass splitting.

Since the mass operator is given in terms of the four-momentum we return to the Lie algebras and Eq. (1.2). The generators of the physical Poincaré group are then

$$p_i' \equiv p_i + p_i^0 \in \mathfrak{P}',$$

where $p_i \in \mathfrak{P}$ are the generators of the nonphysical Poincaré group, and the

$$p_i^0 \equiv t(p_i) \in \mathfrak{S},$$

are purely *internal*. In particular we have

$$[p_i, s_\rho] = 0. \tag{5.1}$$

For the physical mass operator we have

$$m'^2 = p'^2 = (p + p^0)^2 = p^2 + 2p \cdot p^0 + p^{0^2} = m^2 + 2p \cdot p^0 + m^{0^2}. \tag{5.2}$$

There are now two main cases to consider.

(1) $S \supset \bar{L}''$ *or* L'' but $S \not\supset \bar{P}''$ or P''. In this case

$$p_\mu^0 = 0,$$

and hence

$$m'^2 = m^2.$$

Since m^2 is the Casimir operator of $\bar{P}$ it follows that

$$[m'^2, p_i] = 0$$

and from Eq. (5.1) we have

$$[m'^2, s_\rho] = 0,$$

and hence there is no mass splitting.

(2) $S \supset \bar{P}''$ or P''. In this case $p_\mu^0 \neq 0$. In a given representation of E we now have the following possibilities:

(a) $p_\mu \neq 0$ and $p_\mu^0 \not\equiv 0$. We then have from Eq. (5.2)

$$[m'^2, l_{\mu\nu}] \neq 0,$$

where $l_{\mu\nu} \in \mathfrak{L} \subset \mathfrak{P}$ are the generators of the nonphysical Lorentz group, and

$$[m'^2, l_{\mu\nu}^0] \neq 0,$$

i.e., m'^2 transforms under both the nonphysical and the purely internal Lorentz transformations, so that there is mass splitting. However, a theorem by O'Raifeartaigh[6] shows that there can be no mass splitting in a given representation of E if the mass operator in this representation has a discrete point in its spectrum. Thus if there is any mass splitting, as in this case, it must be continuous and can only have physical application in a theory with a continuous mass spectrum.

(b) $p_\mu = 0$ and $p_\mu^0 \not\equiv 0$. Now

$$m'^2 = m^{0^2}.$$

From Eq. (5.1) we then have

$$[m'^2, p_i] = 0,$$

but in general

$$[m'^2, s_\rho] \neq 0,$$

since m^{0^2} is not in general a Casimir operator of S and we get in general a continuous spectrum of m'^2 as in Case (a).

(c) $p_\mu^0 \equiv 0$. We are here back to Case (1) and there is no mass splitting.

6. DISCUSSION

In Refs. 2–4 the problem of combining $\mathfrak{P}$ with $\mathfrak{S}$ is considered for the case when $\mathfrak{S}$ is semisimple and compact, and a semidirect sum, which might be trivially a direct sum, is obtained for $\mathfrak{E}$

$$\mathfrak{E} \cong \mathfrak{S} \dotplus \mathfrak{P}', \tag{6.1}$$

as a result of the conditions on a generalized charge which we discussed in Sec. IV.

The commutations relations of Eq. (6.1)

$$[s_\rho, s_\sigma] = c_{\rho\sigma}^\tau s_\tau, \quad s_\rho \in \mathfrak{S},$$
$$[p_i', p_j'] = c_{ij}^k p_k', \quad p_i' \in \mathfrak{P}',$$

and

$$[p_i', s_\rho] = c_{i\rho}^\sigma s_\sigma,$$

are explored, and it is shown that there exists a set of 10 linear combinations p_i^0 of the s_ρ

$$p_i^0 \equiv b_i^\rho s_\rho,$$

such that

$$[p_i^\rho, s_\rho] = c_{i\rho}^\sigma s_\sigma. \tag{6.2}$$

Furthermore, the p_i^0 satisfy the Poincaré Lie algebra

$$[p_i^0, p_j^0] = c_{ij}^k p_k^0. \tag{6.3}$$

The quantities

$$p_i \equiv p_i' - p_i^0$$

also satisfy the Poincaré Lie algebra

$$[p_i, p_j] = c_{ij}^k p_k,$$

and furthermore

$$[p_i, s_\rho] = 0,$$

so that the double structure, which might be trivial,

$$\mathfrak{E} = \mathfrak{S} \dotplus \mathfrak{P}' \cong \mathfrak{S} + \mathfrak{P},$$

has been obtained.

It does not appear to have been clearly noticed that the necessary and sufficient condition for the existence of a solution p_i^0 to Eq. (6.3) is that, corresponding to

[6] L. O'Raifeartaigh, Phys. Rev. Letters 14, 575 (1965).

our Theorems A and B, there exists a homomorphism

$$t:\mathfrak{T}' \to \mathcal{S}.$$

The 10 p_i^0 are linearly independent if and only if t is an isomorphism.

In Refs. 3 and 4 the problem of nontrivial double structure is considered. However, since the only homomorphic images of $\mathfrak{T}'$ are $t(\mathfrak{T}') = 0$, $\mathfrak{L}''$ or $\mathfrak{T}'' \cong 0$, $\mathfrak{L}$ or $\mathfrak{T}$, respectively, and a nontrivial double structure excludes $t(\mathfrak{T}') = 0$, we must in such a case have $t(\mathfrak{T}') = \mathfrak{L}''$ or $\mathfrak{T}'' \subset \mathcal{S}$. These purely internal Lorentz or Poincaré Lie algebras are noncompact, which clearly excludes the possibility of nontrivial double structure for $\mathcal{S}$ compact, the case considered in these articles.

In Ref. 2, Sudarshan proves for $\mathcal{S}$ compact that in all *unitary* representations of $\mathcal{E}$ no nontrivial double structure can exist. We have now generalized this theorem and shown it to be true in *all* representations of $\mathcal{E}$ for $\mathcal{S}$ compact.

7. EXAMPLE OF DOUBLE STRUCTURE

In this section we consider a simple example for S which gives rise to double structure which is at first sight in apparent contradiction to our theorems. Though this double structure does not give rise to mass splitting we do believe that the discussion and resolution of the paradox throws some light onto attempts in the literature to obtain mass splitting.

We for S take the *compact* group SO_3 or its covering group SU_2 with the Lie algebra

$$\mathcal{SO}_3 = \{K_\alpha\}, \quad \alpha = 1, 2, 3,$$

with the well-known commutation relations

$$[K_\alpha, K_\beta] = i\xi_{\alpha\beta\gamma}K_\gamma.$$

For the Poincaré group P or $\bar{P}$ we have

$$\mathfrak{T} = \{p_\mu, J_\alpha, N_\alpha\}, \quad \mu = 1, 2, 3, 4,$$

with commutation relations,

$$[J_\alpha, J_\beta] = i\xi_{\alpha\beta\gamma}J_\gamma,$$
$$[J_\alpha, N_\beta] = i\xi_{\alpha\beta\gamma}N_\gamma,$$
$$[N_\alpha, N_\beta] = -i\xi_{\alpha\beta\gamma}J_\gamma.$$

The combined algebra

$$\mathcal{E} = \mathcal{S} + \mathfrak{T},$$

is then described by

$$[p_\mu, K_\alpha] = 0, \quad [J_\alpha, K_\beta] = 0, \quad [N_\alpha, K_\beta] = 0.$$

The quantities,

$$\begin{aligned} p'_\mu &\equiv p_\mu, \\ J'_\alpha &\equiv J_\alpha + K_\alpha, \\ N'_\alpha &\equiv N_\alpha + iK_\alpha, \end{aligned} \tag{7.1}$$

close upon themselves with the following commutation relations:

$$\begin{aligned} [J'_\alpha, J'_\beta] &= i\xi_{\alpha\beta\gamma}J'_\gamma, \\ [J'_\alpha, N'_\beta] &= i\xi_{\alpha\beta\gamma}N'_\gamma, \\ [N'_\alpha, N'_\beta] &= -i\xi_{\alpha\beta\gamma}J'_\gamma, \end{aligned}$$

and hence constitute a Poincaré algebra $\mathfrak{T}'$, which we consider as the physical one. The commutation relations of P' with $\mathcal{SO}_3$ are

$$\begin{aligned} [p'_\mu, K_\alpha] &= 0, \\ [J'_\alpha, K_\beta] &= i\xi_{\alpha\beta\gamma}K_\gamma, \\ [N'_\alpha, K_\beta] &= i\xi_{\alpha\beta\gamma}iK_\gamma, \end{aligned} \tag{7.2}$$

which then indicate that $\mathcal{E}$ is isomorphic to the semidirect sum of $\mathfrak{T}'$ with $\mathcal{SO}_3$.

This is in apparent contradiction to our remarks in Sec. 6 of the impossibility of obtaining a double structure for $\mathcal{S}$ compact!

The clue to the resolution of the paradox lies in the i in the term iK_α of Eqs. (7.1) and (7.2) which indicates that we are not dealing with $\mathcal{SO}_3$ over the field R of real numbers, i.e., $\mathcal{SO}(3, R)$, as we started out by saying, but rather with $\mathcal{SO}_3$ over the field C of complex numbers, i.e., $\mathcal{SO}(3, C)$.

Furthermore, one has committed an error by combining the Lie algebra $\mathfrak{T}$ over R with the Lie algebra $\mathcal{SO}_3$ over C. In order to do things in a formally more correct manner we therefore have to put $\mathcal{SO}(3, C)$ in the form of a Lie algebra over R by doubling the number of generators, i.e., by defining

$$M_\alpha \equiv iK_\alpha.$$

The commutation relations of $\mathcal{SO}(3, C)$ over R then become

$$[K_\alpha, K_\beta] = i\xi_{\alpha\beta\gamma}K_\gamma, \tag{7.3a}$$
$$[K_\alpha, M_\beta] = i\xi_{\alpha\beta\gamma}M_\gamma, \tag{7.3b}$$
$$[M_\alpha, M_\beta] = -i\xi_{\alpha\beta\gamma}K_\gamma. \tag{7.3c}$$

Equations (7.1) and (7.2) become

$$\begin{aligned} N'_\alpha &= N_\alpha + M_\alpha, \\ [N'_\alpha, K_\beta] &= i\xi_{\alpha\beta\gamma}M_\gamma, \end{aligned}$$

which now are in a real form.

We now see from Eqs. (7.3) that

$$\mathcal{S} = \mathcal{SO}(3, C) \cong \mathcal{SO}(3, 1; R) = \mathfrak{L}, \tag{7.4}$$

so not only was the original $\mathcal{S}$ not compact, but we also have our homomorphism t

$$\mathfrak{T} \xrightarrow{t} \mathcal{SO}(3, 1; R) \to 0,$$

as required by the theorem!

Equation (7.4) corresponds globally to

$$SO(3, C) \cong SO(3, 1; R) = L$$

or

$$SU(2, C) \cong SL(2, C) = \bar{L}.$$

Note added in proof: Our attention has been drawn to the following two articles where the problem of double structure is also considered: J. Roskies, J. Math. Phys. **1**, 395 (1966); G. Cologna, Nuovo Cimento **41**, A603 (1966).

ACKNOWLEDGMENTS

We would like to express our thanks to Professor E. C. G. Sudarshan for the hospitality extended to us at Syracuse University, for his kind interest and in particular for suggesting the example considered in Sec. VII. We also thank our colleagues at the Irving Institute.

This work was supported in part by the U.S. Atomic Energy Commission.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 5 MAY 1967

Reduction of Relativistic Wavefunctions to the Irreducible Representations of the Inhomogeneous Lorentz Group. I. Nonzero Mass Components

H. E. MOSES

Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts*

(Received 5 October 1966)

When the space–time coordinates of a relativistic system undergo the transformations of the proper, orthochronous, inhomogeneous Lorentz group, the wavefunction of the system undergoes transformations which may be considered to constitute a representation of the group. We give a simple algorithm for reducing this representation to the irreducible unitary ray representations if we assume that only nonzero mass representations occur. The extension to cases in which zero mass representations occur will be given in a later paper. The form in which the reduction is given is an expansion of the wavefunction as given in configuration space in terms of a basis such that the coefficients transform in accordance with the Foldy–Shirokov realization of the irreducible representations. Any wave equation which the wavefunction satisfies and any auxiliary conditions, such as the Lorentz condition or reality conditions, eliminate or relate in a simple way some of the representations which can appear. As examples, we reduce the scalar wavefunction, the four-vector with and without the Lorentz condition, the Dirac wavefunction, the wavefunction which transforms like the electromagnetic field, and a wavefunction which transforms as a generalization of the Dirac wavefunction. In these examples it is also shown that if one replaces the amplitudes associated with the irreducible representations by annihilation and creation operators in a suitable manner, one obtains the usual canonical formalism for second quantization in configuration space. The reduction technique given herein is a simple application of the results of an earlier paper by the author and J. S. Lomont in which is shown how to reduce any unitary ray representation of the inhomogeneous Lorentz group.

1. INTRODUCTION

IN Ref. 1 a recipe was given and proved which enables one to reduce any unitary ray representation of the proper, orthochronous, inhomogeneous Lorentz group. The present paper is a simple application of Ref. 1 and is, in a sense, an extension of it.

In the present paper we consider a set of functions $\Psi(\mathbf{x}, t, \gamma)$ where the variable γ runs through a set of discrete or continuous values. At times it is convenient to suppress the variable γ and write $\Psi(\mathbf{x}, t) = \{\Psi(\mathbf{x}, t, \gamma)\}$. If γ is a discrete variable, it is also useful sometimes to regard $\Psi(\mathbf{x}, t)$ as being a column vector with components $\Psi(\mathbf{x}, t, \gamma)$.

We wish to regard $\Psi(\mathbf{x}, t)$ as being a wavefunction which transforms by means of a unitary transformation to another function $\Psi'(\mathbf{x}, t)$ when the space–time coordinates undergo any transformation of the proper, orthochronous inhomogeneous Lorentz group. These transformations form a ray representation of the group and we can use Ref. 1 to reduce these functions to the irreducible representations of the group. In the present paper we restrict ourselves to the case where only nonzero mass irreducible representations appear. In the next paper we consider zero mass representations.

An earlier paper on the reduction of wavefunctions of nonzero mass is Ref. 2. We believe that our technique for the reduction, in addition to being

* Operated with support from the U.S. Advanced Research Projects Agency.

[1] J. S. Lomont and H. E. Moses, J. Math. Phys. **8**, 838 (1967).

[2] D. L. Pursey, Ann. Phys. (N.Y.) **32**, 157 (1965).

Equation (7.4) corresponds globally to

$$SO(3, C) \cong SO(3, 1; R) = L$$

or

$$SU(2, C) \cong SL(2, C) = \bar{L}.$$

Note added in proof: Our attention has been drawn to the following two articles where the problem of double structure is also considered: J. Roskies, J. Math. Phys. **1**, 395 (1966); G. Cologna, Nuovo Cimento **41**, A603 (1966).

ACKNOWLEDGMENTS

We would like to express our thanks to Professor E. C. G. Sudarshan for the hospitality extended to us at Syracuse University, for his kind interest and in particular for suggesting the example considered in Sec. VII. We also thank our colleagues at the Irving Institute.

This work was supported in part by the U.S. Atomic Energy Commission.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 5 MAY 1967

Reduction of Relativistic Wavefunctions to the Irreducible Representations of the Inhomogeneous Lorentz Group. I. Nonzero Mass Components

H. E. MOSES

Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts*

(Received 5 October 1966)

When the space–time coordinates of a relativistic system undergo the transformations of the proper, orthochronous, inhomogeneous Lorentz group, the wavefunction of the system undergoes transformations which may be considered to constitute a representation of the group. We give a simple algorithm for reducing this representation to the irreducible unitary ray representations if we assume that only nonzero mass representations occur. The extension to cases in which zero mass representations occur will be given in a later paper. The form in which the reduction is given is an expansion of the wavefunction as given in configuration space in terms of a basis such that the coefficients transform in accordance with the Foldy–Shirokov realization of the irreducible representations. Any wave equation which the wavefunction satisfies and any auxiliary conditions, such as the Lorentz condition or reality conditions, eliminate or relate in a simple way some of the representations which can appear. As examples, we reduce the scalar wavefunction, the four-vector with and without the Lorentz condition, the Dirac wavefunction, the wavefunction which transforms like the electromagnetic field, and a wavefunction which transforms as a generalization of the Dirac wavefunction. In these examples it is also shown that if one replaces the amplitudes associated with the irreducible representations by annihilation and creation operators in a suitable manner, one obtains the usual canonical formalism for second quantization in configuration space. The reduction technique given herein is a simple application of the results of an earlier paper by the author and J. S. Lomont in which is shown how to reduce any unitary ray representation of the inhomogeneous Lorentz group.

1. INTRODUCTION

IN Ref. 1 a recipe was given and proved which enables one to reduce any unitary ray representation of the proper, orthochronous, inhomogeneous Lorentz group. The present paper is a simple application of Ref. 1 and is, in a sense, an extension of it.

In the present paper we consider a set of functions $\Psi(\mathbf{x}, t, \gamma)$ where the variable γ runs through a set of discrete or continuous values. At times it is convenient to suppress the variable γ and write $\Psi(\mathbf{x}, t) = \{\Psi(\mathbf{x}, t, \gamma)\}$. If γ is a discrete variable, it is also useful sometimes to regard $\Psi(\mathbf{x}, t)$ as being a column vector with components $\Psi(\mathbf{x}, t, \gamma)$.

We wish to regard $\Psi(\mathbf{x}, t)$ as being a wavefunction which transforms by means of a unitary transformation to another function $\Psi'(\mathbf{x}, t)$ when the space–time coordinates undergo any transformation of the proper, orthochronous inhomogeneous Lorentz group. These transformations form a ray representation of the group and we can use Ref. 1 to reduce these functions to the irreducible representations of the group. In the present paper we restrict ourselves to the case where only nonzero mass irreducible representations appear. In the next paper we consider zero mass representations.

An earlier paper on the reduction of wavefunctions of nonzero mass is Ref. 2. We believe that our technique for the reduction, in addition to being

* Operated with support from the U.S. Advanced Research Projects Agency.

[1] J. S. Lomont and H. E. Moses, J. Math. Phys. **8**, 838 (1967).

[2] D. L. Pursey, Ann. Phys. (N.Y.) **32**, 157 (1965).

more general, is simpler and more suitable for applications.

We present our result in the form of a simple algorithm which allows us to expand $\Psi(\mathbf{x}, t)$ in terms of amplitudes which transform according to the Foldy–Shirokov[3,4] realization of the irreducible representations of the group in the momentum representations.

It is to be noted that to reduce the wavefunction, only the transformation properties are necessary. While the requirement that Ψ satisfies a wave equation may determine the transformation properties of the function (that this is not always the case can be seen by noting that both the components of four-vector wavefunction and the scalar wavefunction can be required to satisfy the ordinary wave equation), the wave equation is not otherwise used to reduce the wavefunction. The effect of the wave equation is to restrict the number of independent irreducible representations which appear—for example, by permitting only one mass. Auxiliary conditions, such as the Lorentz condition, in the case of the four-vector, or reality conditions further eliminate or relate in a simple way some of the irreducible representations.

Finally, we also show how the negative energy representations may be replaced by positive energy representations (i.e., how to introduce "antiparticles") and, if the wavefunction satisfies a wave equation, how second quantization is to be introduced to agree with the canonical formalism in configuration space. Though we carry out the second quantization in some special cases only, we feel that the general procedure can be deduced from the special cases.

We reduce the following wavefunctions as illustrations of the procedure (a) the scalar wavefunction, (b) the wavefunction that transforms like a Dirac spinor, (c) the four-vector with and without the Lorentz condition, (d) a wavefunction that transforms like the electromagnetic field, and (e) a wavefunction which transforms as a generalization of the Dirac spinor.

2. ALGORITHM

It is convenient to regard any transformation of the proper, orthochronous, inhomogeneous Lorentz group as a product of three particular transformations. Let x^α ($\alpha = 0, 1, 2, 3$) denote the components of the space–time four-vector with $x^0 = -x_0 = t$, $x^1 = x_1$, $x^2 = x_2$, $x^3 = x_3$. We take $\hbar = c = 1$ in our units.

The first transformation is a translation $T(a^\alpha)$ characterized by the four-vector a^α in which the components of the x^α four-vector in the new frame of reference are given by

$$x'^\alpha = x^\alpha - a^\alpha. \tag{2.1}$$

The second transformation is a rotation $R(\boldsymbol{\theta})$ where the direction of the vector $\boldsymbol{\theta}$ gives the direction of the axis of rotation and $\theta = |\boldsymbol{\theta}|$ gives the angle of rotation. Under the rotation $R(\boldsymbol{\theta})$ the four-vector transforms in the following way:

$$\begin{aligned} x'^0 &= x^0, \\ \mathbf{x}' &= \mathbf{x}\cos\theta + \frac{1-\cos\theta}{\theta^2}(\boldsymbol{\theta}\cdot\mathbf{x})\boldsymbol{\theta} - \frac{\sin\theta}{\theta}(\boldsymbol{\theta}\times\mathbf{x}), \end{aligned} \tag{2.2}$$

where $\mathbf{x}$ is the space part of the four-vector x^α, i.e., $\mathbf{x} = \{x_1, x_2, x_3\}$.

The third transformation is the pure Lorentz transformation $L(\boldsymbol{\beta})$ where the direction of $\boldsymbol{\beta}$ is in the opposite direction of the moving frame of reference as observed in the original frame and the magnitude $\beta = |\boldsymbol{\beta}|$ is given by $\cosh\beta = [1 - v^2]^{-\frac{1}{2}}$. Under the transformation $L(\boldsymbol{\beta})$ we have

$$\begin{aligned} x'^0 &= x^0\cosh\beta + \boldsymbol{\beta}\cdot\mathbf{x}(\sinh\beta/\beta), \\ \mathbf{x}' &= \mathbf{x} + \boldsymbol{\beta}(\boldsymbol{\beta}\cdot\mathbf{x})[(\cosh\beta - 1)/\beta^2] + \boldsymbol{\beta}x^0(\sinh\beta/\beta). \end{aligned} \tag{2.3}$$

It is convenient to use the concept of column vector to describe the four-vector x^α. Accordingly, let the column vector x be defined by

$$x = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}. \tag{2.4}$$

Let us also define the three matrices $\hat{M}_i$ and the three matrices $\hat{N}_i$ ($i = 1, 2, 3$) by

$$\hat{M}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix},$$

$$\hat{M}_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{pmatrix}, \tag{2.5}$$

$$\hat{M}_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

[3] L. L. Foldy, Phys. Rev. 102, 568 (1956).

[4] Yu. M. Shirokov, Zh. Eksperim. i Teor. Fiz. 33, 1196 (1957) [English transl.: Soviet Phys.—JETP 13, 240 (1961)].

$$\hat{N}_1 = \begin{pmatrix} 0 & -i & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\hat{N}_2 = \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{2.6}$$

$$\hat{N}_3 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}.$$

Then the transformation $T(a^\alpha)$ can be written in terms of column vectors as

$$x' = x - a, \tag{2.1a}$$

while the transformations $R(\boldsymbol{\theta})$ and $L(\boldsymbol{\beta})$ can be written in terms of matrices as follows:

$$x' = \exp[i\boldsymbol{\theta} \cdot \hat{\mathbf{M}}]x, \tag{2.2a}$$

$$x' = \exp[i\boldsymbol{\beta} \cdot \hat{\mathbf{N}}]x. \tag{2.3a}$$

In (2.2a) and (2.3a) $\boldsymbol{\theta} \cdot \hat{\mathbf{M}} = \sum_i \theta_i \hat{M}_i$ and $\boldsymbol{\beta} \cdot \hat{\mathbf{N}} = \sum_i \beta_i \hat{N}_i$.

It is also convenient to write (2.1a), (2.2a), and (2.3a) as follows:

$$x' = T(a)x, \tag{2.1b}$$

$$x' = R(\boldsymbol{\theta})x, \tag{2.2b}$$

$$x' = L(\boldsymbol{\beta})x. \tag{2.3b}$$

As used in (2.2b) and (2.3b) $R(\boldsymbol{\theta})$ and $L(\boldsymbol{\beta})$ are matrices which should not be confused with the abstract transformations though the same notation is used.

We note here for later use some properties of $\hat{M}_i$ and $\hat{N}_i$. As is well known the matrices $\hat{M}_i$ and $\hat{N}_i$ satisfy the commutation rules of the infinitesimal generators of the proper, orthochronous, homogeneous Lorentz group, namely,

$$\begin{aligned} [\hat{M}_i, \hat{M}_j] &= i \sum_k \epsilon_{ijk} \hat{M}_k, \\ [\hat{M}_i, \hat{N}_j] &= i \sum_k \epsilon_{ijk} \hat{N}_k, \\ [\hat{N}_i, \hat{N}_j] &= -i \sum_k \epsilon_{ijk} \hat{M}_k, \end{aligned} \tag{2.7}$$

where ϵ_{ijk} is the usual antisymmetric three-index symbol.

We notice also that the matrices $\hat{M}_i$ by themselves satisfy the commutation relations for the infinitesimal generators of the rotation group. In fact, the matrices $\hat{M}_i$ appear in reduced form. Let us define the matrices $\hat{S}_i$ by

$$\hat{S}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \qquad \hat{S}_2 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix},$$

$$\hat{S}_3 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{2.8}$$

The matrices $\hat{S}_i$ constitute, as is well known, the irreducible representation of the generators of the rotation group corresponding to vector rotations. Let us denote the trivial one by one-dimensional representation of the generators by $\tilde{\mathbf{S}}$. Then, of course, $\tilde{S}_i = 0$.

The matrices $\hat{M}_i$ take the reduced form

$$\hat{M}_i = \begin{pmatrix} \tilde{S}_i & 0 \\ 0 & \hat{S}_i \end{pmatrix}. \tag{2.9}$$

We now consider the way that the wavefunctions change under the transformations of the inhomogeneous Lorentz group. Let us denote by $\Psi(x)$ the wavefunction which we have denoted in Sec. 1 by $\Psi(\mathbf{x}, t)$. Let us denote the wavefunction in the new frame of reference after any transformation of the inhomogeneous Lorentz group by $\Psi'(x)$. Under the transformation $T(a^\alpha)$ we require

$$\Psi'(x) = \Psi(x + a). \tag{2.10}$$

Under the transformations $R(\boldsymbol{\theta})$ and $L(\boldsymbol{\beta})$ we require

$$\Psi'(x) = \exp[i\boldsymbol{\theta} \cdot \mathbf{M}]\Psi(R(-\boldsymbol{\theta})x), \tag{2.11}$$

$$\Psi'(x) = \exp[i\boldsymbol{\beta} \cdot \mathbf{N}]\Psi(L(-\boldsymbol{\beta})x), \tag{2.12}$$

where M_i and N_i $(i = 1, 2, 3)$ are operators which operate on the γ variable of the wavefunction written as $\Psi(\mathbf{x}, t, \gamma)$ as in Sec. 1. We require that the matrices M_i and N_i satisfy the commutation rules for the infinitesimal generators of the homogeneous Lorentz group (2.7) and also that they be integrable in the sense that $\exp[i\boldsymbol{\theta} \cdot \mathbf{M}]$ and $\exp[i\boldsymbol{\beta} \cdot \mathbf{N}]$ exist and that they can be used to generate a ray representation of the homogeneous Lorentz group in an obvious way.

The transformation requirements which we have put on the wavefunctions are the usual ones which are put on wavefunctions to characterize them as being relativistically invariant.

In the usual cases of interest the variable γ runs through a finite set of values and, hence, the integrability conditions on M_i and N_i are satisfied. Furthermore, the operators M_i which correspond to the spin part of the angular momentum operators are usually taken to be Hermitian. As is well known from the theory of the homogeneous Lorentz group, if the variable γ runs through a finite set of values and the matrices M_i are Hermitian, then the matrices N_i cannot be.

If a suitable inner product is introduced—and we later show that one can always introduce such an inner product—the transformations on the wavefunction Ψ given by Eqs. (2.10)–(2.12) form, locally at least, a representation of the inhomogeneous Lorentz group in a sense to be given now.

Let the operators $P_i = P^i$ $(i = 1, 2, 3)$ and $P^0 = -P_0 = H$ be defined by the way that they act on the wavefunction:

$$P^\alpha\Psi(x) = -i(\partial/\partial x_\alpha)\Psi(x) \quad (\alpha = 0, 1, 2, 3). \tag{2.13}$$

The operators P^α are operators corresponding to the momentum four-vector. Let us define the operators J_i which correspond to the components of the angular momentum by

$$J_i\Psi(x) = \left[-i\sum_{jk}\epsilon_{ijk}x_j\frac{\partial}{\partial x_k} + M_i\right]\Psi(x) \qquad (i, j, k = 1, 2, 3). \tag{2.14}$$

[In (2.14) and later $i = (-1)^{\frac{1}{2}}$ should not be confused with the index i.]

Finally the operators which correspond to the space–time portion of the four-by-four relativistic angular momentum tensor

$$\mathcal{J}_i\Psi(x) = \left[i\left(x_i\frac{\partial}{\partial x^0} + x^0\frac{\partial}{\partial x_i}\right) + N_i\right]\Psi(x). \tag{2.15}$$

These operators satisfy the commutation rules for the infinitesimal generators of the inhomogeneous Lorentz group.

Let us define

$$\begin{aligned} J_{23} = -J_{32} = J_1, \quad J_{31} = -J_{13} = J_2, \\ J_{12} = -J_{21} = J_3. \\ J_{0i} = -J_{i0} = \mathcal{J}_i. \end{aligned} \tag{2.16}$$

Then

$$\begin{aligned} [P^\alpha, P^\beta] = 0, \quad [J_{\alpha\beta}, P_\lambda] = i[g_{\alpha\lambda}P_\beta - g_{\beta\lambda}P_\alpha], \\ [J_{\alpha\beta}, J_{\lambda\mu}] = i[g_{\alpha\lambda}J_{\beta\mu} - g_{\beta\lambda}J_{\alpha\mu} + g_{\alpha\mu}J_{\lambda\beta} - g_{\beta\mu}J_{\lambda\alpha}], \end{aligned} \tag{2.17}$$

where $g_{\alpha\beta}$ is the metric tensor defined by $g_{\alpha\beta} = 0$ if $\alpha \neq \beta$, $g_{00} = -g_{11} = -g_{22} = -g_{33} = -1$.

It is easily shown that the wavefunctions transforming under $T(a^\alpha)$, $R(\boldsymbol{\theta})$, $L(\boldsymbol{\beta})$ as in (2.10)–(2.12), respectively, can be written

$$\Psi'(x) = \exp\left[i\sum_\alpha a^\alpha P_\alpha\right]\Psi(x), \tag{2.10a}$$

$$\Psi'(x) = \exp\left[i\boldsymbol{\theta}\cdot\mathbf{J}\right]\Psi(x), \tag{2.11a}$$

$$\Psi'(x) = \exp\left[i\boldsymbol{\beta}\cdot\boldsymbol{\mathcal{J}}\right]\Psi(x). \tag{2.12a}$$

Our observation that the commutation rules (2.17) for the infinitesimal generators of the inhomogeneous Lorentz group are satisfied by the infinitesimal generators associated with the transformation of the wavefunction Ψ is the starting point of our investigation to reduce these operators to the irreducible representations of the inhomogeneous Lorentz group using the techniques of Ref. 1.

We now discuss the Foldy–Shirokov form for the irreducible representations of the inhomogeneous Lorentz group for particles of mass μ, spin s, and sign of energy ϵ $(\epsilon = \pm 1)$. For each representation there is an irreducible set of spin operators S_i such that $\mathbf{S}^2 = s(s+1)I$, where I is the identity operator. We introduce a set of complex functions of a vector $\mathbf{p}$ and a discrete variable λ which runs through $2s+1$ values and upon which the spin operators S_i act. Denoting a function of this set by $f(\mathbf{p}, \lambda)$ we introduce a norm

$$\sum_\lambda\int\frac{d\mathbf{p}}{\omega(\mu, p)}|f(\mathbf{p}, \lambda)|^2,$$

and a corresponding inner product, where

$$\omega(\mu, p) = [\mu^2 + p^2]^{\frac{1}{2}}, \quad p = |\mathbf{p}|.$$

We regard the set of functions $\{f(\mathbf{p}, \lambda)\}$ as constituting a representation of the states of a Hilbert space. We denote the infinitesimal generators of the inhomogeneous Lorentz group by $\hat{P}^\alpha, \hat{J}_i, \hat{\mathcal{J}}_i$ [which, of course satisfy (2.17)]. We have—on suppressing the λ variable in $f(\mathbf{p}, \lambda)$ for simplicity—

$$\begin{aligned} \hat{P}^0 f(\mathbf{p}) &= \hat{H}f(\mathbf{p}) = \epsilon\omega(\mu, p)f(\mathbf{p}), \\ \hat{P}^i f(\mathbf{p}) &= p_i f(\mathbf{p}), \\ \hat{J}_i f(\mathbf{p}) &= \left[-i\sum_{j,k}\epsilon_{ijk}p_j\frac{\partial}{\partial p_k} + S_i\right]f(\mathbf{p}), \\ \hat{\mathcal{J}}_i f(\mathbf{p}) &= \epsilon\left[i\omega(\mu, p)\frac{\partial}{\partial p_i} + \frac{1}{\omega(\mu, p)+\mu}\times\sum_{j,k}\epsilon_{ijk}p_jS_k\right]f(\mathbf{p}). \end{aligned} \tag{2.18}$$

[The carats used above the operators in (2.18) are simply a notation and are not meant to imply in general a relation of these operators to $\hat{M}_i$, $\hat{N}_i$, or $\hat{S}_i$.]

For the sake of completeness we show how the function $f(\mathbf{p})$, which may be regarded as a wavefunction for a particle of spin s, mass μ, and sign of energy ϵ changes when the frame of reference is changed under the finite transformations of the inhomogeneous Lorentz group. The transformations for the wavefunction are analogous to those for $\Psi(x)$ as given by Eqs. (2.10)–(2.12) and (2.10a)–(2.12a). They are given in Ref. 5.

For each three-vector $\mathbf{p}$ let us introduce a four-vector p whose space components are the components of $\mathbf{p}$ and whose time component p^0 is given by

$$p^0 = \epsilon\omega(\mu, p),$$
$$= -p_0.$$

We use the notation p_c for the column vector whose components are those of p.

Then denoting by $f'(\mathbf{p})$ the wavefunction in the new frame of reference, we have under the transformation $T(a^\alpha)$

$$f'(\mathbf{p}) = \exp\left[i\sum_\alpha a^\alpha \hat{P}_\alpha\right] f(\mathbf{p}) = \exp\left[i\sum_\alpha a^\alpha p_\alpha\right] f(\mathbf{p}). \tag{2.19}$$

Under the rotation $R(\boldsymbol{\theta})$ the wavefunction becomes

$$f'(\mathbf{p}) = \exp[i\boldsymbol{\theta}\cdot\hat{\mathbf{J}}] f(\mathbf{p})$$
$$= \exp[i\boldsymbol{\theta}\cdot\mathbf{S}] f(\mathbf{p}'), \tag{2.20}$$

where $\mathbf{p}'$ is the space part of the four-vector p' given by

$$p'_c = \exp[-i\boldsymbol{\theta}\cdot\hat{\mathbf{M}}] p_c \tag{2.20a}$$

[cf. (2.2a)]. Under the space–time transformation $L(\boldsymbol{\beta})$ the wavefunction becomes

$$f'(\mathbf{p}) = \exp[i\boldsymbol{\beta}\cdot\hat{\mathbf{J}}] f(\mathbf{p})$$
$$= \exp\left\{\frac{2i\epsilon}{k}\left[\left(\frac{\boldsymbol{\beta}}{\beta}\times\mathbf{p}\right)\cdot\mathbf{S}\right]\Phi(\boldsymbol{\beta},\mathbf{p})\right\} f(\mathbf{p}'), \tag{2.21}$$

where $\mathbf{p}'$ is the space part of the four-vector p' given by

$$p'_c = \exp[-i\boldsymbol{\beta}\cdot\hat{\mathbf{N}}] p_c \tag{2.21a}$$

[cf. (2.3a)],

$$k = \{p^2 - [(\boldsymbol{\beta}/\beta)\cdot\mathbf{p}]^2\}^{\frac{1}{2}} \tag{2.21b}$$

and the function Φ is given by

$$\tan\Phi(\boldsymbol{\beta},\mathbf{p}) = \frac{k\sinh(\frac{1}{2}\beta)}{[\omega(\mu,p)+\mu]\cosh(\frac{1}{2}\beta) - \epsilon(\boldsymbol{\beta}/\beta)\cdot\mathbf{p}\sinh(\frac{1}{2}\beta)}. \tag{2.21c}$$

Let us consider the functions $g(\mathbf{p}) = f^*(-\mathbf{p})$. From equations (2.19)–(2.21) one can show that the functions $g(\mathbf{p})$ transform like $f(\mathbf{p})$ but with ϵ replaced by $-\epsilon$ and the spin matrices S_i replaced by new spin matrices $\bar{S}_i$ with

$$\bar{S}_i = -S_i^*,$$

where the asterisk means complex conjugate. That is, the functions $g(\mathbf{p})$ transform like wavefunctions of a particle of mass μ, sign of energy $-\epsilon$, and spin s using the spin matrices $\bar{S}_i$ in the expressions for the infinitesimal or finite generators of the inhomogeneous Lorentz group. Later, by using this theorem, we are able to replace negative energy wavefunctions in the expansion of Ψ by the complex conjugate of positive energy wavefunctions, these positive energy wavefunctions playing the role of "antiparticle" wavefunctions. We are thus able to expand Ψ in terms of "physical" wavefunctions only, i.e., with wavefunctions corresponding to particles having positive energy.

In replacing negative energy wavefunctions with positive energy wavefunctions it is sometimes convenient to use the original spin operators S_i instead of $\bar{S}_i$. Since the set of matrices $\{\bar{S}_i\}$ correspond to the same irreducible representation of the infinitesimal generators of the rotation group as $\{S_i\}$ there exists a unitary matrix U such that

$$S_i = U\bar{S}_iU^{-1}.$$

Let $h(\mathbf{p}) = Ug(\mathbf{p})$ where U acts only on the suppressed λ variables in $g(\mathbf{p})$. Then it is not difficult to see that $h(\mathbf{p})$ transforms like the wavefunction of a particle of mass μ, sign of energy $-\epsilon$, and spin s using now the matrices S_i in the expressions for the infinitesimal or finite generators of the inhomogeneous Lorentz group.

It is our objective to expand $\Psi(x)$ in terms of wavefunctions $f(\mathbf{p})$ belonging to various values of μ, ϵ, and s. In this expression we require that if every $f(\mathbf{p})$ is replaced by $\hat{K}f(\mathbf{p})$, where $\hat{K}$ is an infinitesimal generator of the inhomogeneous Lorentz group, then this expansion equals $K\Psi(x)$, where K is the corresponding generator in configuration space. We have thus reduced the set of infinitesimal generators acting on the wavefunctions $\Psi(x)$. We can also express our requirement in "integrated" form. Let the frame of reference undergo one of the transformations $T(a^\alpha)$, $R(\boldsymbol{\theta})$, $L(\boldsymbol{\beta})$. Then in the expansion we replace $f(\mathbf{p})$ by $f'(\mathbf{p})$ given by one of the appropriate formulas (2.19)–(2.21). With this replacement, this expansion will equal $\Psi'(x)$ given by the appropriate formula (2.10)–(2.12). We now set up various definitions and conventions. First of all, we consider $\Psi(x)$ to be given completely, as far as its transformation

[5] H. E. Moses, Ann. Phys. (N.Y) 41, 158 (1967).

properties are concerned, when the matrices M_i and N_i are given. By the conditions which we have given on the spin operators M_i these operators are completely reducible. We assume, to make matters as simple as possible, that the matrices M_i are given in completely reduced form. If, then, Ψ were given in a form in which the matrices M_i were *not* reduced, we should first change our basis in the variable γ so that the matrices M_i are reduced. We then would also be given the matrices N_i in this basis. We note that γ is a discrete variable. We thus assume that the matrices M_i have the form

$$M_i = \begin{pmatrix} S_i^{(1)} & & & & \\ & S_i^{(2)} & & & \\ & & S_i^{(3)} & & \\ & & & \cdot & \\ & & & & \cdot \end{pmatrix}, \quad (2.22)$$

where $S_i^{(r)}$ are a set of irreducible matrices belonging to the rth block of this decomposition of M_i. We should mention that in many cases the matrices M_i are already given in the reduced form above. All the examples which we treat later are such cases.

Let us now consider the matrix $\exp[-i\nu \cdot \mathbf{N}]$ whose elements we write

$$\{\exp[-i\nu \cdot \mathbf{N}]\}_{\gamma,\gamma'}.$$

Now, for each set of matrices $\{S_i^{(r)}\}$ which appear in (2.22) and for every value of μ, ϵ, $\mathbf{p}$ we define a column vector $\chi^{(r)}(\mu, \epsilon, \mathbf{p}, \lambda)$ with components $\chi^{(r)}(\gamma \mid \mu, \epsilon, \mathbf{p}, \lambda)$ by

$$\chi^{(r)}(\gamma \mid \mu, \epsilon, \mathbf{p}, \lambda) = \{\exp[-i\nu \cdot \mathbf{N}]\}_{\gamma,\lambda}, \quad (2.23)$$

where λ is restricted to values corresponding to the columns in rth *block of* (2.22). Furthermore, the vector ν is related to μ, ϵ, $\mathbf{p}$ by

$$\mathbf{p} = -\epsilon\mu(\nu/\nu)\sinh\nu, \quad \nu = |\nu|, \quad (2.23a)$$

from which

$$p = \mu\sinh\nu, \quad \omega(\mu, p) = \mu\cosh\nu. \quad (2.23b)$$

It is usually convenient to relabel λ from 1 to $2s + 1$ where s is the spin corresponding to the matrices $S_i^{(r)}$.

We can now state the principal theorem of the present paper:

Let $f^{(r)}(\mu, \epsilon, \mathbf{p}, \lambda)$ be a function upon which the infinitesimal generators act as given by (2.18) using the spin matrices $S_i^{(r)}$ acting on the variable λ [which is suppressed in (2.18)]; or equivalently let $f^{(r)}$ transform as in (2.19)–(2.21) with mass μ, sign of energy ϵ, and spin matrices $S_i^{(r)}$, then the most general expansion of the function $\Psi(x)$ in terms of the irreducible unitary ray representations of the proper, orthochronous, inhomogeneous Lorentz group is

$$\Psi(x) = \sum_{\epsilon}\sum_{r}\sum_{\lambda}\int dM^{(r)}(\mu, \epsilon)\int\frac{d\mathbf{p}}{\omega(\mu, p)} \times \chi^{(r)}(\mu, \epsilon, \mathbf{p}, \lambda) f^{(r)}(\mu, \epsilon, \mathbf{p}, \lambda) \times \exp\{i[\mathbf{p}\cdot\mathbf{x} - \epsilon\omega(\mu, p)t]\}, \quad (2.24)$$

where $M^{(r)}(\mu, \epsilon)$ is an arbitrary measure in the Lesbegue–Stieltjes sense.

We want to emphasize that (2.24) is a necessary condition which reduces the infinitesimal generators to the irreducible *Hermitian* infinitesimal generators, which appear in the reduction. These can be considered as dynamical variables of particles of various masses and spins, as defined in the Wigner sense. Furthermore, since our theorem is a necessary condition, the expansion (2.24) should be verified for any particular case. In the cases which follow we have verified the expansions, but we do not reproduce the verification because of the length of proof and because the reader can carry out the proof himself without too much difficulty. [It should be mentioned, however, that some computations which the writer has made indicate that (2.24) may also be sufficient for a complete reduction.]

Equation (2.24) is the principal result of the paper. We prove it using the recipe for reducing reducible unitary ray representations given in Ref. 1.

One of the more remarkable aspects of the result is that the measures $M^{(r)}(\mu, \epsilon)$ is arbitrary, that is, the transformation properties—though they are frequently derived from the requirement that a wave equation be invariant—do not determine which masses and signs of energy appear. We see that the wave equations restrict the masses and, in general, eliminate some of the irreducible representations which appear in (2.24). Reality and subsidiary conditions also eliminate some of the irreducible representations.

Equation (2.24) enables us to define an inner product of two wavefunctions in configuration space which is invariant under the transformations of the inhomogeneous Lorentz group—such a definition being usually difficult to give. Let us define the inner product of $\Psi^{(1)}(x)$ and $\Psi(x)$ by

$$(\Psi^{(1)}, \Psi) = \sum_{\epsilon}\sum_{r}\sum_{\lambda}\int dM_1^{(r)}(\mu, \epsilon)\int\frac{d\mathbf{p}}{\omega(\mu, p)} \times f^{(1)(r)*}(\mu, \epsilon, \mathbf{p}, \lambda) f^{(r)}(\mu, \epsilon, \mathbf{p}, \lambda), \quad (2.25)$$

where $M_1^{(r)}(\mu, \epsilon)$ are appropriately chosen measure

functions compatible with the measure functions $M^{(r)}(\mu, \epsilon)$ of (2.24).

For the purposes of second quantization it is usually convenient to work only with positive energy representations. We make substitutions in accordance with our earlier discussion in which we showed that the complex conjugate of a positive energy wavefunction with reversed momentum transformed like a negative energy particle. To describe this situation we modify our notation somewhat.

Let us define

$$\begin{aligned} M^{(r)}(\mu) &= M^{(r)}(\mu, +1), \quad N^{(r)}(\mu) = M^{(r)}(\mu, -1), \\ \chi^{(r)}(\mu, \mathbf{p}, \lambda) &= \chi^{(r)}(\mu, +1, \mathbf{p}, \lambda), \\ \zeta^{(r)}(\mu, \mathbf{p}, \lambda) &= \sum_{\lambda'} \chi^{(r)*}(\mu, -1, -\mathbf{p}, \lambda') U^{(r)*}_{\lambda,\lambda'}, \\ f^{(r)}(\mu, \mathbf{p}, \lambda) &= f^{(r)}(\mu, +1, \mathbf{p}, \lambda), \\ h^{(r)}(\mu, \mathbf{p}, \lambda) &= \sum_{\lambda'} U^{(r)}_{\lambda,\lambda'} f^{(r)*}(\mu, -1, -\mathbf{p}, \lambda'), \end{aligned} \tag{2.26}$$

where the matrix $U^{(r)} = \{U^{(r)}_{\lambda,\lambda'}\}$ is the matrix which takes $\bar{S}^{(r)}_i$ into $S^{(r)}_i$ so that $h^{(r)}$ transforms as a particle of positive energy using the matrices $S^{(r)}_i$ as in our earlier discussion. Then (2.24) takes the form

$$\begin{aligned} \Psi(x) = &\sum_r \sum_\lambda \int dM^{(r)}(\mu) \int \frac{d\mathbf{p}}{\omega(\mu, p)} \chi^{(r)}(\mu, \mathbf{p}, \lambda) \\ &\times f^{(r)}(\mu, \mathbf{p}, \lambda) \exp\{i[\mathbf{p}\cdot\mathbf{x} - \omega(\mu, p)t]\} \\ &+ \sum_r \sum_\lambda \int dN^{(r)}(\mu) \int \frac{d\mathbf{p}}{\omega(\mu, p)} \zeta^{(r)*}(\mu, \mathbf{p}, \lambda) \\ &\times h^{(r)*}(\mu, \mathbf{p}, \lambda) \exp\{-i[\mathbf{p}\cdot\mathbf{x} - \omega(\mu, p)t]\}. \end{aligned} \tag{2.24a}$$

Generally, we prefer to work with (2.24a) instead of (2.24). We now proceed to discuss examples.

3. SCALAR FIELD

The case in which the wavefunction transforms as a scalar is the simplest. Nevertheless, the treatment of the scalar case serves as a prototype of the treatment of wavefunctions which transform in a more complicated way. Hence we go into the properties of the scalar wavefunction in considerable detail, despite the obviousness of some of the results.

(a) Reduction of the Wavefunction

For the scalar field the operators M_i and N_i of Eqs. (2.11) and (2.12) are zero and the variables γ and λ do not appear at all. Also, the representation label r is not needed and is dropped. Furthermore, the column vectors which are generally noted by $\chi^{(r)}$ and $\zeta^{(r)}$ are simply 1. Then Eqs. (2.24) and (2.24a) become

$$\begin{aligned} \Psi(x) = &\sum_\epsilon \int dM(\mu, \epsilon) \int \frac{d\mathbf{p}}{\omega(\mu, p)} f(\mu, \mathbf{p}, \epsilon) \\ &\times \exp\{i[\mathbf{p}\cdot\mathbf{x} - \epsilon\omega(\mu, p)t] \end{aligned} \tag{3.1}$$

and

$$\begin{aligned} \Psi(x) = &\int dM(\mu) \int \frac{d\mathbf{p}}{\omega(\mu, p)} f(\mu, \mathbf{p}) \\ &\times \exp\{i[\mathbf{p}\cdot\mathbf{x} - \omega(\mu, p)t]\} \\ &+ \int dN(\mu) \int \frac{d\mathbf{p}}{\omega(\mu, p)} h^*(\mu, \mathbf{p}) \\ &\times \exp\{-i[\mathbf{p}\cdot\mathbf{x} - \omega(\mu, p)t]\}. \end{aligned} \tag{3.2}$$

We work primarily with (3.2).

(b) Reality Condition

Let us see how reality conditions affect the representations which appear. Let us require that Ψ be real, i.e., $\Psi(x) = \Psi^*(x)$. From the linear independence of the exponents it follows that

$$dN(\mu)h(\mu, \mathbf{p}) = dM(\mu)f(\mu, \mathbf{p}). \tag{3.3}$$

Thus the reality condition essentially fixes the "antiparticle" representative in terms of the particle representative.

(c) Wave Equation

Let us now require that $\Psi(x)$ satisfy the wave equation

$$[(\partial^2/\partial t^2) - \nabla^2 + m^2]\psi(x) = 0, \tag{3.4}$$

where m is the mass of the scalar particle. (We do not assume the reality condition on ψ in the present discussion.) In terms of the infinitesimal generators this requirement is equivalent to

$$[H^2 - \mathbf{p}^2]\Psi(x) = m^2\Psi(x). \tag{3.5}$$

Let K be any of the infinitesimal generators. Then $K\Psi(x)$ is obtained from (3.2) by replacing $f(\mathbf{p})$ and $h(\mathbf{p})$ by $\hat{K}f(\mathbf{p})$ and $\hat{K}h(\mathbf{p})$, respectively, in (3.2). It follows that (3.5) is entirely equivalent to

$$\begin{aligned} &\int dM(\mu) \int \frac{d\mathbf{p}}{\omega(\mu, p)} (\mu^2 - m^2) f(\mu, \mathbf{p}) \\ &\quad\times \exp\{i[\mathbf{p}\cdot\mathbf{x} - \omega(\mu, p)t] \\ &\quad+ \int dN(\mu) \int \frac{d\mathbf{p}}{\omega(\mu, p)} (\mu^2 - m^2) h^*(\mu, \mathbf{p}) \\ &\quad\times \exp\{-i[\mathbf{p}\cdot\mathbf{x} - \omega(\mu, p)t]\} = 0. \end{aligned} \tag{3.6}$$

It follows that for $f(\mu, \mathbf{p})$ and $h(\mu, \mathbf{p})$ not to be identically zero, the measure functions $M(\mu)$ and

$N(\mu)$ must have a jump at $\mu = m$ and can be constant for all other values of μ. That is,

$$dM(\mu) = C\delta(\mu - m)\,d\mu, \qquad dN(\mu) = D\delta(\mu - m)\,d\mu. \tag{3.7}$$

In (3.7) C and D are real positive constants.

Then (3.2) becomes

$$\Psi(x) = C\int \frac{d\mathbf{p}}{\omega(p)} f(\mathbf{p}) \exp\{i[\mathbf{p}\cdot\mathbf{x} - \omega(p)t]\} + D\int \frac{d\mathbf{p}}{\omega(p)} h^*(\mathbf{p}) \exp\{-i[\mathbf{p}\cdot\mathbf{x} - \omega(p)t]\}, \tag{3.8}$$

where $f(\mathbf{p}) = f(m, \mathbf{p})$, $h(\mathbf{p}) = h(m, \mathbf{p})$, and $\omega(p) = \omega(m, p)$.

It is easy to see that (3.8) gives the general solution of the wave equation (3.5) so that our expansion in terms of the *unitary* irreducible representations completely solves the problem.

We can choose C and D so that the usual canonical formalism in terms of Lagrangian and Hamiltonian densities (see, e.g., Ref. 6) agrees with the particle interpretation. Let $H(\mathbf{x})$ be the Hamiltonian density of the field which leads to the wave equation (3.5). Let us define the energy of the field E to be

$$E = \int H(\mathbf{x})\,d\mathbf{x}. \tag{3.9}$$

It is easy to see that if $L(\mathbf{x})$ is a Lagrangian density which leads to a wave equation (without interactions), $-L(\mathbf{x})$ is also a Lagrangian which leads to the same wave equation. Consequently if $H(\mathbf{x})$ is a Hamiltonian density, so is $H'(\mathbf{x}) = -H(\mathbf{x})$. Of course, $H(\mathbf{x})$ depends on the wavefunction Ψ which is a solution of the wave equation. Our principle for choosing $H(\mathbf{x})$ or $H'(\mathbf{x})$ is that we require E to be positive when Ψ contains only one mode, i.e., when only one of the functions $f(\mathbf{p})$ or $h(\mathbf{p})$ is not identically zero. When we discuss nonscalar wavefunctions we use the same principle for choosing the sign of $H(\mathbf{x})$. That is, we set all wavefunctions except one identically equal to zero and choose $H(\mathbf{x})$ for this mode so that E is positive.

Let us first consider the case that ψ is complex. For both modes [i.e., either $h(\mathbf{p}) = 0$ or $f(\mathbf{p}) = 0$] we can choose $H(\mathbf{x})$ to be the usual functional of Ψ, namely,[6]

$$H(\mathbf{x}) = \dot{\Psi}^*\dot{\Psi} + \nabla\Psi^*\cdot\nabla\Psi + m^2\Psi^*\Psi. \tag{3.10}$$

Having picked $H(\mathbf{x})$ for each mode, we now require that E (which must now be positive) be the expectation value of the energy as viewed from the particle picture for each mode. Thus (using subscripts on E to designate the mode)

$$E_f = \int \frac{d\mathbf{p}}{\omega(p)} f^*(\mathbf{p})\omega(p)f(\mathbf{p}) \quad \text{for the mode} \quad h(\mathbf{p}) = 0,$$
$$E_h = \int \frac{d\mathbf{p}}{\omega(p)} h^*(\mathbf{p})\omega(p)h(\mathbf{p}) \quad \text{for the mode} \quad f(\mathbf{p}) = 0. \tag{3.11}$$

The requirements (3.11) lead to the following values for the constants C and D of the expansion (3.8):

$$C = D = (2)^{-\frac{1}{2}}(2\pi)^{-\frac{3}{2}}. \tag{3.12}$$

We have thus obtained a rationale for the usual formalism in the language of the reduction of the scalar wavefunction in terms of the representations of the inhomogeneous Lorentz group.

We define the expectation value of the total energy when the state is a superposition of modes as being $E = E_f + E_h$. In this case E corresponds to the total energy of the field when Ψ is *any* solution of the wave equation.

We can view the quadratic form which gives E as being a sort of weighted inner product. It is now natural to define the true inner product in configuration space as

$$(\Psi^1, \Psi) = \int \frac{d\mathbf{p}}{\omega(p)} f^{1*}(\mathbf{p})f(\mathbf{p}) + \int \frac{d\mathbf{p}}{\omega(p)} h^1(\mathbf{p})h^*(\mathbf{p}). \tag{3.13}$$

With this inner product the infinitesimal generators of Eqs. (2.13)–(2.15) are Hermitian in configuration space. Clearly this inner product is invariant.

We could equally well consider real scalar wavefunctions Ψ in which case we should have to restrict $h(\mathbf{p})$ in the expansion (3.8) to satisfy the requirement [cf. (3.3)]

$$Dh(\mathbf{p}) = Cf(\mathbf{p}) \tag{3.14}$$

and then use the Hamiltonian density corresponding to the real field,

$$H(\mathbf{x}) = \tfrac{1}{2}[\dot{\Psi}^2 + (\nabla\Psi)^2 + m^2\Psi^2]. \tag{3.15}$$

Then we require that the energy

$$E = \int H(\mathbf{x})\,d\mathbf{x}$$

be also given by

$$E = \int d\mathbf{p}\,|f(\mathbf{p})|^2. \tag{3.16}$$

We find that C and D have the same values as before.

[6] G. Wentzel, *Quantum Theory of Fields* (Interscience Publishers Inc., New York, 1949).

The inner product of two wavefunctions is then defined by

$$(\Psi^1, \Psi) = \int \frac{d\mathbf{p}}{\omega(p)} f^{1*}(\mathbf{p}) f(\mathbf{p}). \tag{3.17}$$

(d) Second Quantization

We now indicate how the theory is to be second quantized when Ψ satisfies the wave equation. The second quantization is carried out by replacing the amplitudes corresponding to the irreducible representations by annihilation operators and the complex conjugate of such amplitudes by creation operators. Let us first take Ψ to be complex. In the expansion for the operators Ψ and Ψ^* (the adjoint of Ψ) given by (3.8) we regard $f(\mathbf{p})$ and $h(\mathbf{p})$ as annihilation operators and their respective adjoints, denoted by $f^*(\mathbf{p})$ and $h^*(\mathbf{p})$, as creation operators. We take C and D to be given by (3.12). Thus assuming Bose statistics, we have the following commutation rules:

$$\begin{aligned}[f(\mathbf{p}), f(\mathbf{p}')] &= [f^*(\mathbf{p}), f^*(\mathbf{p}')] = [h(\mathbf{p}), h(\mathbf{p}')] \\ &= [h^*(\mathbf{p}), h^*(\mathbf{p}')] = [f(\mathbf{p}), h(\mathbf{p}')] \\ &= [f^*(\mathbf{p}), h(\mathbf{p}')] = 0,\end{aligned} \tag{3.18}$$

$$[f(\mathbf{p}), f^*(\mathbf{p}')] = [h(\mathbf{p}), h^*(\mathbf{p}')] = \omega(p)\delta(\mathbf{p} - \mathbf{p}'). \tag{3.19}$$

It is easily shown that the operator $\Psi(\mathbf{x}, t) = \Psi(x)$, $\Psi^*(\mathbf{x}, t) = \Psi^*(x)$ satisfy the following commutation relations:

$$\begin{aligned}&[\Psi(\mathbf{x}, t), \Psi(\mathbf{x}', t')] = [\Psi^*(\mathbf{x}, t), \Psi^*(\mathbf{x}', t')] = 0, \\ &[\Psi(\mathbf{x}, t), \Psi^*(\mathbf{x}', t')] = D(\mathbf{x} - \mathbf{x}', t - t'),\end{aligned} \tag{3.20}$$

where $D(\mathbf{x}, t)$ is the usual invariant function

$$D(\mathbf{x}, t) = (2\pi)^{-3} \int \frac{d\mathbf{p}}{\omega(p)} \exp[i\mathbf{p} \cdot \mathbf{x}] \sin \omega(p) t. \tag{3.21}$$

Of course, (3.20) are the usual canonical commutation rules for a complex scalar field (see Ref. 6).

We should note that $f^*(\mathbf{p})$ is the creation operator for a particle and $h^*(\mathbf{p})$ is the creation operator for the "antiparticle," both the particle and antiparticle having positive energy and the same mass. The particle and antiparticle are treated at the same level.

We can now give the unitary transformation under which the operators Ψ and Ψ^* transform when the frame of reference is changed. Let $\hat{A}$ be any one of the operators $\hat{H}$, $\hat{P}^i = \hat{P}_i$, $\hat{J}_i$, $\hat{\mathcal{J}}_i$ acting on the representatives as in (2.18) (with $S_i = 0$ in the present case). Let us define $\hat{A}f(\mathbf{p})$ when $f(\mathbf{p})$ is a destruction operator as the result of the operator $\hat{A}$ acting on the variable $\mathbf{p}$ in the operator $f(\mathbf{p})$ in the same way as though $f(\mathbf{p})$ were a representative as in (2.18). The second quantized operator $[A]$ corresponding to the infinitesimal generator $\hat{A}$ is defined by

$$[A] = \int \frac{d\mathbf{p}}{\omega(p)} f^*(p) \hat{A} f(p) + \int \frac{d\mathbf{p}}{\omega(p)} h^*(\mathbf{p}) \hat{A} h(\mathbf{p}). \tag{3.22}$$

It is easily seen that the second quantized operators $[H]$, $[P_i]$, $[J_i]$, $[\mathcal{J}_i]$ satisfy the same commutation rules as H, P_i, J_i, $\mathcal{J}_i$ [Eq. (2.17)]. The second-quantized operators constitute representation of the infinitesimal generators of the proper, orthochronous, inhomogeneous Lorentz group in the second-quantized theory.

Let us consider the Lorentz transformation $T(a^\alpha)$. Then using the notation $x = (t, \mathbf{x})$ the operator $\Psi(x)$ transforms to the operator $\Psi'(x)$, where

$$\begin{aligned}\Psi'(x) &= \Psi(x + a) \\ &= \exp\left\{-i \sum_\alpha a^\alpha [P_\alpha]\right\} \Psi(x) \exp\left\{i \sum_\alpha a^\alpha [P_\alpha]\right\}.\end{aligned} \tag{3.23}$$

From (3.23) it can be shown that Ψ satisfies the wave equation, as required.

Under the rotation $R(\boldsymbol{\theta})$ the operator Ψ transforms to Ψ' which is given by

$$\begin{aligned}\Psi'(x) &= \Psi(R(-\boldsymbol{\theta})x) \\ &= \exp\{-i\boldsymbol{\theta} \cdot [\mathbf{J}]\} \Psi(x) \exp\{i\boldsymbol{\theta} \cdot [\mathbf{J}]\},\end{aligned} \tag{3.24}$$

where $\boldsymbol{\theta} \cdot [\mathbf{J}] = \sum_i \theta_i [J_i]$.

Under the space–time transformation $L(\boldsymbol{\beta})$ the transformed operator Ψ' is given by

$$\begin{aligned}\Psi'(x) &= \Psi(L(-\boldsymbol{\beta})x) \\ &= \exp\{-i\boldsymbol{\beta} \cdot [\mathcal{J}]\} \Psi(x) \exp\{i\boldsymbol{\beta} \cdot [\mathcal{J}]\},\end{aligned} \tag{3.25}$$

where $\boldsymbol{\beta} \cdot [\mathcal{J}] = \sum \beta_i [\mathcal{J}_i]$. Equations (3.23)–(3.25) are proved using the well-known algorithm

$$e^{-A} B e^{A} = \sum_n \{B, A\}^{(n)} \frac{1}{n!}, \tag{3.26}$$

where A and B are any two operators and $\{B, A\}^{(n)}$ is defined inductively by

$$\{B, A\}^{(0)} = B, \quad \{B, A\}^{(n)} = [\{B, A\}^{(n-1)}, A]. \tag{3.26a}$$

Thus the second-quantized theory is also invariant under the transformations of the inhomogeneous Lorentz group, since unitary operators exist which transform $\Psi(x)$ in the required way under changes of frame of reference.

Let us now discuss the real field from the viewpoint of second quantization. In the expansion (3.8) we consider $f(\mathbf{p})$ to be a destruction operator as before. The function $h(\mathbf{p})$ is replaced by $f(\mathbf{p})$ also. The

constants C and D are given by (3.12). Thus $\Psi(x)$ is a Hermitian operator. The annihilation operator $f(\mathbf{p})$ and creation operator $f^*(\mathbf{p})$ are required to satisfy the same commutation rules as before, namely:

$$[f(\mathbf{p}), f(\mathbf{p}')] = 0, \quad [f(\mathbf{p}), f^*(\mathbf{p}')] = \omega(p)\delta(\mathbf{p} - \mathbf{p}'). \tag{3.27}$$

The infinitesimal generators for the second-quantized theory are obtained from those for the first-quantized theory in a manner analogous to (3.22):

$$[A] = \int \frac{d\mathbf{p}}{\omega(p)} f^*(\mathbf{p}) \hat{A} f(\mathbf{p}). \tag{3.28}$$

With these infinitesimal generators, (3.23)–(3.25) hold for transforming the Hermitian operator Ψ under changes of frames of reference.

4. WAVEFUNCTIONS WHICH TRANSFORM AS AN ANTISYMMETRIC TENSOR OR ELECTROMAGNETIC FIELD

In the present paper we reduce the antisymmetric tensors or, what is equivalent, electromagnetic fields. However, we do not require that wave equations be satisfied and we do not introduce second quantization in this section. We introduce second quantization later when we derive such skew-symmetric tensors from vector potentials.

(a) Characterization of the Wavefunction. Transformation Properties

Let us first consider a real antisymmetric tensor $F^{\alpha\beta} = -F^{\beta\alpha}$. We require that under the translation $T(a^\alpha)$ the components of the tensor $F^{\alpha\beta}(x)$ transform to the tensor whose components are $F^{\alpha\beta}(x + a)$. Furthermore, we require that $F^{\alpha\beta}$ transform as a tensor under the transformations of proper homogeneous Lorentz group. It is convenient, though by no means necessary, to introduce a wavefunction which is related more directly to the electromagnetic field description of an antisymmetric tensor. Accordingly we define

$$\begin{aligned} E_i &= F^{0i} \quad (i = 1, 2, 3), \\ H_1 &= F^{23}, \quad H_2 = F^{31}, \quad H_3 = F^{12}. \end{aligned} \tag{4.1}$$

We now introduce a column vector $\Psi(x)$ with components $\Psi(\mathbf{x}, t, \gamma)$ with $\gamma = 1, 2, 3$ given by

$$\Psi(\mathbf{x}, t, \gamma) = E_\gamma(\mathbf{x}, t) - iH_\gamma(\mathbf{x}, t). \tag{4.2}$$

It is clear that $F^{\alpha\beta}$ can be obtained from $\Psi(x)$.

Under the transformations $T(a^\alpha)$, $R(\boldsymbol{\theta})$, and $L(\boldsymbol{\beta})$, Ψ transforms according to (2.10)–(2.12) where the matrices M_i are given by

$$M_i = \hat{S}_i, \tag{4.3}$$

the matrices $\hat{S}_i$ being given by (2.8) and constitute the irreducible representation of the generators of the rotation group corresponding to vector rotations. The matrices N_i are given by

$$N_i = i\hat{S}_i. \tag{4.4}$$

The matrix $\exp[i\boldsymbol{\theta} \cdot \mathbf{M}] = \exp[i\boldsymbol{\theta} \cdot \hat{\mathbf{S}}]$ is simply the matrix which gives the new components of a vector in terms of the old components when the new frame of reference is obtained from the old one by a rotation described by $\boldsymbol{\theta}$.

Using the fact that $(\boldsymbol{\theta} \cdot \hat{\mathbf{S}})^3 = \theta^2(\boldsymbol{\theta} \cdot \hat{\mathbf{S}})$ we obtain the result

$$\exp[i\boldsymbol{\theta} \cdot \hat{\mathbf{S}}] = I + i(\boldsymbol{\theta} \cdot \hat{\mathbf{S}})(\sin\theta/\theta) + (\boldsymbol{\theta} \cdot \hat{\mathbf{S}})^2 \times [(\cos\theta - 1)/\theta^2]. \tag{4.5}$$

Defining the matrix $\hat{R}(\boldsymbol{\theta}) = \exp[i\boldsymbol{\theta} \cdot \hat{\mathbf{S}}]$ we have for the matrix elements of this matrix

$$[\hat{R}(\boldsymbol{\theta})]_{ij} = \delta_{ij}\cos\theta - \frac{\theta_i\theta_j}{\theta^2}(\cos\theta - 1) + \sum_k \epsilon_{ijk}\theta_k \frac{\sin\theta}{\theta}, \tag{4.6}$$

where ϵ_{ijk} is the usual three-index antisymmetric symbol.

Let us define the matrix $\hat{L}(\boldsymbol{\beta})$ by

$$\hat{L}(\boldsymbol{\beta}) = \exp[i\boldsymbol{\beta} \cdot \mathbf{N}]. \tag{4.7}$$

On using the fact that $(\boldsymbol{\beta} \cdot \mathbf{N})^3 = -\beta^2(\boldsymbol{\beta} \cdot \mathbf{N})$ we have

$$\hat{L}(\boldsymbol{\beta}) = I + i(\boldsymbol{\beta} \cdot \mathbf{N})(\sinh\beta/\beta) + (\boldsymbol{\beta} \cdot \mathbf{N})^2 \times [(1 - \cosh\beta)/\beta^2]. \tag{4.8}$$

The matrix elements are given by

$$[\hat{L}(\boldsymbol{\beta})]_{ij} = \delta_{ij}\cosh\beta - \frac{\beta_i\beta_j}{\beta^2}(\cosh\beta - 1) + i\sum_k \epsilon_{ijk}\beta_k \frac{\sinh\beta}{\beta}. \tag{4.9}$$

Equations (2.10)–(2.12), which show how the wavefunction transforms under the changes of frames of reference, take on a simpler form when we use a vector notation for the wavefunction Ψ, which up to now we have considered a column vector. Let us introduce the vector field $\boldsymbol{\Psi}(x)$ whose components are just the components of the column vector Ψ. Then under the transformation $T(a^\alpha)$ the new vector field is

$$\boldsymbol{\Psi}'(x) = \boldsymbol{\Psi}(x + a). \tag{4.10}$$

Under the rotation $R(\boldsymbol{\theta})$ the new vector field is

$$\boldsymbol{\Psi}'(x) = \boldsymbol{\Psi}\big(R(-\boldsymbol{\theta})x\big)\cos\theta - \boldsymbol{\theta}[\boldsymbol{\theta} \cdot \boldsymbol{\Psi}\big(R(-\boldsymbol{\theta})x\big)] \times \frac{\cos\theta - 1}{\theta^2} - [\boldsymbol{\theta} \times \boldsymbol{\Psi}\big(R(-\boldsymbol{\theta})x\big)]\frac{\sin\theta}{\theta} \tag{4.11}$$

[compare (2.2)].

Under the transformation $L(\boldsymbol{\beta})$ the new vector field is transformed to

$$\Psi'(x) = \Psi(L(-\boldsymbol{\beta})x)\cosh\beta - \boldsymbol{\beta}[\boldsymbol{\beta}\cdot\Psi(L(-\boldsymbol{\beta})x)] \times \frac{\cosh\beta - 1}{\beta^2} - i[\boldsymbol{\beta}\times\Psi(L(-\boldsymbol{\beta})x)]\frac{\sinh\beta}{\beta}. \tag{4.12}$$

Equation (4.12) is a remarkably simple expression for the space–time Lorentz transformation for the electromagnetic field. The use of Ψ to simplify the transformation properties of the electromagnetic field is discussed from a somewhat different point of view in Ref. 7.

(b) Expansion of the Wavefunction

Let us now return (for the moment) to the column vector notation for Ψ. Since the matrices $M_i = \hat{S}_i$ are irreducible, we do not need the label (r) in the expansion of Ψ in the manner of (2.24) and (2.24a). The components of the column vector $\chi(\mu, \epsilon, \mathbf{p}, \lambda)$ are obtained from (2.23), (4.7), and (4.9). Our result for χ is (on labeling λ from 1 to 3)

$$\chi(\gamma\,|\,\mu, \epsilon, \mathbf{p}, \lambda) = \frac{1}{\mu}\left[\omega(\mu, p)\delta_{\gamma\lambda} - \frac{p_\gamma p_\lambda}{\omega(\mu, p) + \mu} + i\epsilon\sum_k \epsilon_{\gamma\lambda k}p_k\right]. \tag{4.13}$$

The expansions corresponding to (2.24) and (2.24a) takes a neater form if we use a vector notation. As before we use $\Psi(x)$ for the vector constructed from the column vector $\Psi(x)$. Let us construct the vectors $\mathbf{f}(\mu, \epsilon, \mathbf{p})$, $\mathbf{f}(\mu, \mathbf{p}) = \mathbf{f}(\mu, +1, \mathbf{p})$, and $\mathbf{h}(\mu, \mathbf{p}) = \mathbf{f}^*(\mu, -1, -\mathbf{p})$, respectively, having the components $f(\mu, \epsilon, \mathbf{p}, \lambda)$, $f(\mu, \mathbf{p}, \lambda)$, and $h(\mu, \mathbf{p}, \lambda)$. Indeed, from (2.20) $\mathbf{f}(\mu, \epsilon, \mathbf{p})$, $\mathbf{f}(\mu, \mathbf{p})$, and $\mathbf{h}(\mu, \mathbf{p})$ will transform like vector fields under rotations of the coordinate system.

Then Eq. (2.24) becomes (on absorbing $1/\mu$ into the measure functions)

$$\Psi(x) = \sum_\epsilon \int dM(\mu, \epsilon)\int\frac{d\mathbf{p}}{\omega(\mu, p)} \times \exp\{i[\mathbf{p}\cdot\mathbf{x} - \epsilon\omega(\mu, p)t]\}\left\{\omega(\mu, p)\mathbf{f}(\mu, \epsilon, \mathbf{p}) - \frac{\mathbf{p}}{\omega(\mu, p) + \mu}[\mathbf{p}\cdot\mathbf{f}(\mu, \epsilon, \mathbf{p})] - i\epsilon[\mathbf{p}\times\mathbf{f}(\mu, \epsilon, \mathbf{p})]\right\}, \tag{4.14}$$

$$\Psi(x) = \int dM(\mu)\int\frac{d\mathbf{p}}{\omega(\mu, p)}\, e^{i[\mathbf{p}\cdot\mathbf{x} - \omega(\mu, p)t]}\left\{\omega(\mu, p)\mathbf{f}(\mu, \mathbf{p}) - \frac{\mathbf{p}}{\omega(\mu, p) + \mu}[\mathbf{p}\cdot\mathbf{f}(\mu, \mathbf{p})] - i[\mathbf{p}\times\mathbf{f}(\mu, \mathbf{p})]\right\} + \int dN(\mu)\int\frac{d\mathbf{p}}{\omega(\mu, p)}\, e^{-i[\mathbf{p}\cdot\mathbf{x} - \omega(\mu, p)t]}\left\{\omega(\mu, p)\mathbf{h}^*(\mu, \mathbf{p}) - \frac{\mathbf{p}}{\omega(\mu, p) + \mu}[\mathbf{p}\cdot\mathbf{h}^*(\mu, \mathbf{p})] - i[\mathbf{p}\times\mathbf{h}^*(\mu, \mathbf{p})]\right\}. \tag{4.15}$$

In deriving (4.15) we have used the fact that the matrix U whose elements appear in (2.26) can be chosen to be the identity, since $\hat{S}_i = \hat{S}_i$.

(c) Generalization to Complex Tensors

Up to now we have considered only real tensors $F^{\alpha\beta}$. However, the generalization to complex tensors is not difficult. Let us write

$$F^{\alpha\beta} = F^{\alpha}_R + iF^{\alpha\beta}_I, \tag{4.16}$$

where $F^{\alpha\beta}_R$ and $F^{\alpha\beta}_I$ are the real and imaginary parts of the tensor $F^{\alpha\beta}$. We can then construct the two 3-component column vectors $\Psi_R(x)$ and $\Psi_I(x)$ from $F^{\alpha\beta}_R$ and $F^{\alpha\beta}_I$, respectively, in the same way that Ψ was constructed from the real tensor $F^{\alpha\beta}$. Then the Ψ wavefunction is the six-component column vector formed by placing the three component column vector Ψ_R above Ψ_I. It is clear that the complex tensor $F^{\alpha\beta}$ can be reconstructed from the column vector Ψ.

It is further clear that the matrices M_i and N_i appear in completely reduced form in which the three-by-three reduced matrices are identical to those for the real tensor $F^{\alpha\beta}$ given by (4.3) and (4.4). It is further clear that the new function Ψ can be expanded precisely as the Ψ for the real tensor. Four irreducible representations of the inhomogeneous Lorentz group appear in the expansion for each value of μ because now the index (r) can take on two values with two signs for each value of (r). We refrain from details for the sake of brevity and because of the obviousness of the procedure.

We also postpone to later sections the discussion of how certain wave equations affect the irreducible representations which appear in Ψ.

5. WAVEFUNCTIONS WHICH TRANSFORM AS A FOUR-VECTOR

In the present section we show that the relativistic decomposition of the vector potential simplifies the discussion of such potentials considerably. We show that generally we have a perfectly well-defined theory which involves particles with two spins. When the

[7] H. E. Moses, Nuovo Cimento, Suppl. 7 (1), 1 (1958).

Lorentz condition is imposed, the representatives which correspond to one of the spins is identically zero. The usual difficulties which appear in the canonical formalism (see, e.g., Ref. 6) in which redundant variables appear are resolved in a very simple manner.

The results in this paper anticipate to some extent the results for the mass zero case where the vector potential with the Lorentz condition is interpreted as being an electromagnetic vector potential. (See Ref. 8.) There are, however, also some essential differences.

(a) Transformation Properties of the Wavefunction

We now discuss wavefunctions which transform like the four-vector $x = \{x^\alpha\}$ under the transformations of the homogeneous Lorentz group. Accordingly we label the index γ of the components $\Psi(\mathbf{x}, t, \gamma)$ of the column vector $\Psi(x)$ by taking γ to have the values $\gamma = 0, 1, 2, 3$. Then if $A^\gamma(\mathbf{x}, t)$ are the components of of a vector field which transforms as x^γ, we make the identification

$$\Psi(\mathbf{x}, t, \gamma) = A^\gamma(\mathbf{x}, t). \tag{5.1}$$

In the present section we use A^γ instead of Ψ in order to be closer to more conventional notation.

It is clear that for the present case

$$\begin{aligned} M_i &= \hat{M}_i, \\ N_i &= \hat{N}_i, \end{aligned} \tag{5.2}$$

where $\hat{M}_i$ and $\hat{N}_i$ are given by (2.5) and (2.6).

(b) Reduction of the Wavefunction

From (2.9) we see that $\hat{M}_i$ reduces to two irreducible representations of the rotation group, namely, the scalar representation and the vector representation. Thus the label (r) takes on two values which we call 0 and 1 for the scalar and vector representations, respectively.

Furthermore,

$$\exp[i\boldsymbol{\beta}\cdot\hat{\mathbf{N}}] = I + i(\boldsymbol{\beta}\cdot\hat{\mathbf{N}})(\sinh\beta/\beta) + (\boldsymbol{\beta}\cdot\hat{\mathbf{N}})^2 \times [(1-\cosh\beta)/\beta^2]. \tag{5.3}$$

It is perhaps interesting to note that (5.3) has the same form as (4.8).

It is now a straightforward matter to apply the algorithm to find the transformation functions $\chi^{(r)}$. For $r = 0$, the variable λ takes on only one value. Hence it need not be indicated.

Then $\chi^{(0)}(\gamma \mid \mu, \epsilon, \mathbf{p})$ is given by

$$\begin{aligned} \chi^{(0)}(0 \mid \mu, \epsilon, \mathbf{p}) &= \omega(\mu, p)/\mu \\ \chi^{(0)}(\gamma \mid \mu, \epsilon, \mathbf{p}) &= (\epsilon/\mu)p_\gamma \quad (\gamma = 1, 2, 3). \end{aligned} \tag{5.4}$$

For $r = 1$ the variable λ can take on three values which we take to be 1, 2, 3.

Then

$$\begin{aligned} \chi^{(1)}(0 \mid \mu, \epsilon, \mathbf{p}, \lambda) &= (\epsilon/\mu)p_\lambda, \\ \chi^{(1)}(\gamma \mid \mu, \epsilon, \mathbf{p}, \lambda) &= \frac{1}{\mu}\left[\mu\delta_{\gamma\lambda} + \frac{p_\gamma p_\lambda}{\omega(\mu, p) + \mu}\right] \quad (\gamma = 1, 2, 3). \end{aligned} \tag{5.5}$$

The expansions (2.24) and (2.24a) take a particularly simple form when a vector notation is used. Accordingly let us define the scalar $V(\mathbf{x}, t)$ by

$$V(\mathbf{x}, t) = A^0(\mathbf{x}, t) \tag{5.6}$$

and the vector $\mathbf{A}(\mathbf{x}, t)$ by

$$\mathbf{A}(\mathbf{x}, t) = \{A^1(\mathbf{x}, t), A^2(\mathbf{x}, t), A^3(\mathbf{x}, t)\}. \tag{5.7}$$

It is clear that $f^{(0)}(\mu, \epsilon, \mathbf{p})$ transforms as a scalar under rotations. Let us also introduce the vectors $\mathbf{f}^{(1)}(\mu, \epsilon, \mathbf{p})$, $\mathbf{f}^{(1)}(\mu, \mathbf{p})$, $\mathbf{h}^{(1)}(\mu, \mathbf{p})$ by

$$\begin{aligned} \mathbf{f}^{(1)}(\mu, \epsilon, \mathbf{p}) &= \{f^{(1)}(\mu, \epsilon, \mathbf{p}, 1), f^{(1)}(\mu, \epsilon, \mathbf{p}, 2), \\ &\qquad f^{(1)}(\mu, \epsilon, \mathbf{p}, 3)\}, \\ \mathbf{f}^{(1)}(\mu, \mathbf{p}) &= \{f^{(1)}(\mu, \mathbf{p}, 1), f^{(1)}(\mu, \mathbf{p}, 2), f^{(1)}(\mu, \mathbf{p}, 3)\}, \\ \mathbf{h}^{(1)}(\mu, \mathbf{p}) &= \{h^{(1)}(\mu, \mathbf{p}, 1), h^{(1)}(\mu, \mathbf{p}, 2), h^{(1)}(\mu, \mathbf{p}, 3)\}. \end{aligned} \tag{5.8}$$

Then, on absorbing the factor $1/\mu$ into the measure functions, Eq. (2.24) becomes the equations

$$\begin{aligned} V(\mathbf{x}, t) = \sum_\epsilon \int dM^{(0)}(\mu, \epsilon) \int \frac{d\mathbf{p}}{\omega(\mu, p)} e^{i[\mathbf{p}\cdot\mathbf{x} - \epsilon\omega(\mu,p)t]} \\ \times \omega(\mu, p) f^{(0)}(\mu, \epsilon, \mathbf{p}) + \sum_\epsilon \int dM^{(1)}(\mu, \epsilon) \int \frac{d\mathbf{p}}{\omega(\mu, p)} \\ \times \epsilon e^{i[\mathbf{p}\cdot\mathbf{x} - \epsilon\omega(\mu,p)t]} [\mathbf{p}\cdot\mathbf{f}^{(1)}(\mu, \epsilon, \mathbf{p})], \end{aligned} \tag{5.9}$$

$$\begin{aligned} \mathbf{A}(\mathbf{x}, t) = \sum_\epsilon \int dM^{(0)}(\mu, \epsilon) \int \frac{d\mathbf{p}}{\omega(\mu, p)} \\ \times \epsilon e^{i[\mathbf{p}\cdot\mathbf{x} - \epsilon\omega(\mu,p)t]} \mathbf{p} f^{(0)}(\mu, \epsilon, \mathbf{p}) \\ + \sum_\epsilon \int dM^{(1)}(\mu, \epsilon) \int \frac{d\mathbf{p}}{\omega(\mu, p)} e^{i[\mathbf{p}\cdot\mathbf{x} - \epsilon\omega(\mu,p)t]} \\ \times \left\{\mu\mathbf{f}^{(1)}(\mu, \epsilon, \mathbf{p}) + \frac{\mathbf{p}}{\omega(\mu, p) + \mu}[\mathbf{p}\cdot\mathbf{f}^{(1)}(\mu, \epsilon, \mathbf{p})]\right\}. \end{aligned} \tag{5.9a}$$

Equation (2.24a) becomes

$$\begin{aligned} V(\mathbf{x}, t) = \int dM^{(0)}(\mu) \frac{d\mathbf{p}}{\omega(\mu, p)} \\ \times e^{i[\mathbf{p}\cdot\mathbf{x} - \omega(\mu,p)t]} \omega(\mu, p) f^{(0)}(\mu, \mathbf{p}) \\ + \int dN^{(0)}(\mu) \int \frac{d\mathbf{p}}{\omega(\mu, p)} e^{-i[\mathbf{p}\cdot\mathbf{x} - \omega(\mu,p)t]} \omega(\mu, p) h^{(0)*}(\mu, \mathbf{p}) \\ + \int dM^{(1)}(\mu) \int \frac{d\mathbf{p}}{\omega(\mu, p)} e^{i[\mathbf{p}\cdot\mathbf{x} - \omega(\mu,p)t]} [\mathbf{p}\cdot\mathbf{f}^{(1)}(\mu, \mathbf{p})] \\ + \int dN^{(1)}(\mu) \int \frac{d\mathbf{p}}{\omega(\mu, p)} e^{-i[\mathbf{p}\cdot\mathbf{x} - \omega(\mu,p)t]} [\mathbf{p}\cdot\mathbf{h}^{(1)*}(\mu, \mathbf{p})], \end{aligned} \tag{5.10}$$

[8] H. E. Moses, Nuovo Cimento 42, 757 (1966).

$$\mathbf{A}(\mathbf{x}, t) = \int dM^{(0)}(\mu) \int \frac{d\mathbf{p}}{\omega(\mu, p)} e^{i[\mathbf{p}\cdot\mathbf{x} - \omega(\mu,p)t]} \mathbf{p} f^{(0)}(\mu, \mathbf{p})$$
$$+ \int dN^{(0)}(\mu) \int \frac{d\mathbf{p}}{\omega(\mu, p)} e^{-i[\mathbf{p}\cdot\mathbf{x} - \omega(\mu,p)t]} \mathbf{p} h^{(0)*}(\mu, \mathbf{p})$$
$$+ \int dM^{(1)}(\mu) \int \frac{d\mathbf{p}}{\omega(\mu, p)} e^{i[\mathbf{p}\cdot\mathbf{x} - \omega(\mu,p)t]}$$
$$\times \left\{ \mu \mathbf{f}^{(1)}(\mu, \mathbf{p}) + \frac{\mathbf{p}}{\omega(\mu, p) + \mu} [\mathbf{p} \cdot \mathbf{f}^{(1)}(\mu, \mathbf{p})] \right\}$$
$$+ \int dN^{(1)}(\mu) \int \frac{d\mathbf{p}}{\omega(\mu, p)} e^{-i[\mathbf{p}\cdot\mathbf{x} - \omega(\mu,p)t]}$$
$$\times \left\{ \mu \mathbf{h}^{(1)*}(\mu, \mathbf{p}) + \frac{\mathbf{p}}{\omega(\mu, p) + \mu} [\mathbf{p} \cdot \mathbf{h}^{(1)*}(\mu, \mathbf{p})] \right\}. \tag{5.10a}$$

We prefer to work with Eqs. (5.10) rather than (5.9) because only positive energy representatives appear in (5.10).

(c) Reality Condition

It is readily seen that a necessary and sufficient condition that $V(\mathbf{x}, t)$ and $A(\mathbf{x}, t)$ be real is that

$$dM^{(0)}(\mu) f^{(0)}(\mu, \mathbf{p}) = dN^{(0)}(\mu) h^{(0)}(\mu, \mathbf{p}),$$
$$dM^{(1)}(\mu) \mathbf{f}^{(1)}(\mu, \mathbf{p}) = dN^{(1)}(\mu) \mathbf{h}^{(1)}(\mu, \mathbf{p}). \tag{5.11}$$

(d) Relation to the Real Antisymmetric Tensor

In the previous section we have expanded the real antisymmetric tensor $F^{\alpha\beta}$ in terms of spin 1 representatives $\mathbf{f}(\mu, \mathbf{p})$ and $\mathbf{h}(\mu, \mathbf{p})$ [see Eq. (4.15)]. Let us now consider the real tensor $F^{\alpha\beta}$ when it is obtained from the real four-vector A^α through

$$F^{\alpha\beta} = (\partial A^\beta / \partial x_\alpha) - (\partial A^\alpha / \alpha x_\beta). \tag{5.12}$$

The following theorem is easily proved:

Theorem 1: A necessary and sufficient condition for (5.12) to hold is that

$$dM(\mu) \mathbf{f}(\mu, \mathbf{p}) = dN(\mu) \mathbf{h}(\mu, \mathbf{p}) = i\mu \, dM^{(1)}(\mu) \mathbf{f}^{(1)}(\mu, \mathbf{p}). \tag{5.13}$$

Of course, this theorem restricts considerably the expansion (4.15) for antisymmetric tensors. This theorem also enables one to construct a real four-vector A^α from $F^{\alpha\beta}$ when it is known that this tensor is obtained from a four-vector through (5.12).

It is very interesting to note that the scalar wavefunctions $f^{(0)}$ and $h^{(0)}$ play no role in the relationship between the tensor and the four-vector other than providing a "constant of integration" in the construction of the four-vector from the tensor. It is to be noted that we have not used the Lorentz condition to make this statement true.

The extension of the above results to complex tensors derived from complex four-vectors is straightforward but for the sake of brevity we do not go into the matter.

(e) Lorentz Condition

The following theorem can be proved by substituting (5.10) in an appropriate fashion:

Theorem 2: A necessary and sufficient condition for the Lorentz condition

$$\sum_{\alpha=0}^{3} \frac{\partial A^\alpha}{\partial x^\alpha} = 0 \tag{5.14}$$

is that

$$f^{(0)}(\mu, \mathbf{p}) = h^{(0)}(\mu, \mathbf{p}) = 0.$$

Thus the Lorentz condition is entirely equivalent to stating that there are no particles of spin zero in the expansion for A^α. Thus expansion (5.10) considerably simplifies the discussion of the effect of the Lorentz condition on four-vectors.

(f) Wave Equation

The following statements are easily proved to be true using the techniques which were used for analogous theorems for the scalar wavefunctions.

A necessary and sufficient condition that the components of a four-vector satisfy the wave equation is that

$$(\partial^2/\partial t^2 - \nabla^2 + m^2) A^\alpha(\mathbf{x}, t) = 0 \tag{5.15}$$

$$dM^{(i)}(\mu) = C^{(i)} \delta(\mu - m) \, d\mu,$$
$$dN^{(i)}(\mu) = D^{(i)} \delta(\mu - m) \, d\mu \quad (i = 0, 1), \tag{5.16}$$

where $C^{(i)}$ and $D^{(i)}$ are positive real constants.

Thus if the four-vector satisfies the wave equation (5.15), the most general expansion in terms of positive energy wavefunctions is given by

$$V(\mathbf{x}, t) = C^{(0)} \int d\mathbf{p} e^{i[\mathbf{p}\cdot\mathbf{x} - \omega(p)t]} f^{(0)}(\mathbf{p})$$
$$+ D^{(0)} \int d\mathbf{p} e^{-i[\mathbf{p}\cdot\mathbf{x} - \omega(p)t]} h^{(0)*}(\mathbf{p})$$
$$+ C^{(1)} \int \frac{d\mathbf{p}}{\omega(p)} e^{i[\mathbf{p}\cdot\mathbf{x} - \omega(p)t]} [\mathbf{p} \cdot \mathbf{f}^{(1)}(p)]$$
$$+ D^{(1)} \int \frac{d\mathbf{p}}{\omega(p)} e^{-i[\mathbf{p}\cdot\mathbf{x} - \omega(p)t]} [\mathbf{p} \cdot \mathbf{h}^{(1)*}(\mathbf{p})], \tag{5.17}$$

$$\mathbf{A}(\mathbf{x}, t) = C^{(0)} \int \frac{d\mathbf{p}}{\omega(p)} e^{i[\mathbf{p}\cdot\mathbf{x} - \omega(p)t]} \mathbf{p} f^{(0)}(\mathbf{p})$$
$$+ D^{(0)} \int \frac{d\mathbf{p}}{\omega(p)} e^{-i[\mathbf{p}\cdot\mathbf{x} - \omega(p)t]} \mathbf{p} h^{(0)*}(\mathbf{p})$$
$$+ C^{(1)} \int \frac{d\mathbf{p}}{\omega(p)} e^{i[\mathbf{p}\cdot\mathbf{x} - \omega(p)t]}$$
$$\times \left\{ m \mathbf{f}^{(1)}(\mathbf{p}) + \frac{\mathbf{p}}{\omega(p) + m} [\mathbf{p} \cdot \mathbf{f}^{(1)}(\mathbf{p})] \right\}$$
$$+ D^{(1)} \int \frac{d\mathbf{p}}{\omega(p)} e^{-i[\mathbf{p}\cdot\mathbf{x} - \omega(p)t]}$$
$$\times \left\{ m \mathbf{h}^{(1)*}(\mathbf{p}) + \frac{\mathbf{p}}{\omega(p) + m} [\mathbf{p} \cdot \mathbf{h}^{(1)*}(\mathbf{p})] \right\}, \tag{5.17a}$$

where $f^{(0)}(\mathbf{p}) = f^{(0)}(m, \mathbf{p})$, $h^{(0)}(\mathbf{p}) = h^{(0)}(m, \mathbf{p})$, $\mathbf{f}^{(1)}(\mathbf{p}) = \mathbf{f}^{(1)}(m, \mathbf{p})$, $\mathbf{h}^{(1)}(\mathbf{p}) = h^{(1)}(m, \mathbf{p})$, and $\omega(p) = \omega(m, p)$.

If the four-vector A^α is to be real, we have the requirement that

$$\begin{aligned} C^{(0)}f^{(0)}(\mathbf{p}) &= D^{(0)}h^{(0)}(\mathbf{p}), \\ C^{(1)}\mathbf{f}^{(1)}(\mathbf{p}) &= D^{(1)}\mathbf{h}^{(1)}(\mathbf{p}). \end{aligned} \tag{5.18}$$

The representations given by $h^{(0)}$ and $\mathbf{h}^{(1)}$ are thereby expressed in terms of those given by $f^{(0)}$ and $\mathbf{f}^{(1)}$.

If the Lorentz condition (5.13) is to be satisfied we must have

$$f^{(0)}(\mathbf{p}) = h^{(0)}(\mathbf{p}) = 0. \tag{5.19}$$

(g) Determination of the Constants $C^{(i)}$ and $D^{(i)}$ through the Use of the Canonical Formalism

We assume that the four-vector A^α satisfies the wave equation. By identifying the total energy as obtained from the Hamiltonian density with the expectation value of the energy when only one mode or wavefunction is not identically zero we can obtain the constants $C^{(i)}$ and $D^{(i)}$, as explained in some detail in Sec. 3. Let us first take A^α to be complex. Then for the cases (I) $f^{(0)} = h^{(0)} = \mathbf{h}^{(1)} = 0$ and (II) $f^{(0)} = h^{(0)} = \mathbf{f}^{(1)} = 0$ we use the obvious Hamiltonian density for a complex four-vector

$$H(\mathbf{x}) = \sum_\alpha \{\dot{A}^\alpha \dot{A}^*_\alpha + \nabla A^\alpha \cdot \nabla A^*_\alpha + m^2 A^\alpha A^*_\alpha\} \tag{5.20}$$

to obtain positive energies. [We have suppressed the appearance of the time variable in $H(\mathbf{x})$.] For the Cases (III) $h^{(0)} = \mathbf{f}^{(1)} = \mathbf{h}^{(1)} = 0$ and (IV) $f^{(0)} = \mathbf{f}^{(1)} = \mathbf{h}^{(1)} = 0$, we find that to get positive total energies we must use $-H(\mathbf{x})$ as the Hamiltonian density. With these choices for Hamiltonian density we obtain the following values for $C^{(i)}$ and $D^{(i)}$:

$$C^{(i)} = D^{(i)} = (2)^{-\frac{1}{2}}(2\pi)^{-\frac{3}{2}}. \tag{5.21}$$

On using these constants we have

$$\begin{aligned} \int H(\mathbf{x})\, d\mathbf{x} &= \int \mathbf{f}^{(1)*}(\mathbf{p}) \cdot \mathbf{f}^{(1)}(\mathbf{p})\, d\mathbf{p}, && \text{for Case I} \\ &= \int \mathbf{h}^{(1)*}(\mathbf{p}) \cdot \mathbf{h}^{(1)}(\mathbf{p})\, d\mathbf{p}, && \text{for Case II} \\ &= -\int |f^{(0)}(\mathbf{p})|^2\, d\mathbf{p}, && \text{for Case III} \\ &= -\int |h^{(0)}(\mathbf{p})|^2\, d\mathbf{p}, && \text{for Case IV.} \end{aligned} \tag{5.22}$$

For real fields we replace $D^{(1)}\mathbf{h}^{(1)}(\mathbf{p})$ by $C^{(1)}\mathbf{f}^{(1)}(\mathbf{p})$ and $D^{(0)}h^{(0)}(\mathbf{p})$ by $C^{(0)}f^{(0)}(\mathbf{p})$ in the expansion (5.17) in accordance with the discussion leading to (5.18). Then we need find only $C^{(0)}$ and $C^{(1)}$. The Hamiltonian density is taken to be $\frac{1}{2}H(\mathbf{x})$ for the case that $f^{(0)}(\mathbf{p}) = 0$ and $-\frac{1}{2}H(\mathbf{x})$ for the case that $\mathbf{f}^{(1)}(\mathbf{p}) = 0$. On using the requirement that the energy of the field equals the right-hand side of the first and third of Eqs. (5.22) for the two cases, respectively, we find that $C^{(0)}$ and $C^{(1)}$ are also given by (5.21).

(h) Definition of Inner Product of Vector Wavefunctions

We now define the inner product of two complex vector wavefunctions such that the Hermiticity of the infinitesimal generators in configuration space is assured and such that the inner product is invariant under all the transformations of the proper, inhomogeneous Lorentz group. Let A^α and A^α_1 be the components of two complex four-vectors which satisfy the wave equation. Then following the discussion of the inner product given for (2.25) we define the inner product (A_1, A) by

$$\begin{aligned} (A_1, A) = &\int \frac{d\mathbf{p}}{\omega(p)} f^{(0)*}_1(\mathbf{p}) f^{(0)}(\mathbf{p}) + \int \frac{d\mathbf{p}}{\omega(p)} h^{(0)*}(\mathbf{p}) h^{(0)}_1(\mathbf{p}) \\ &+ \int \frac{d\mathbf{p}}{\omega(\mathbf{p})} \mathbf{f}^{(1)*}_1(\mathbf{p}) \cdot \mathbf{f}^{(1)}(\mathbf{p}) \\ &+ \int \frac{d\mathbf{p}}{\omega(p)} \mathbf{h}^{(1)*}(\mathbf{p}) \cdot \mathbf{h}^{(1)}_1(\mathbf{p}), \end{aligned} \tag{5.23}$$

where the subscript 1 on the representative indicates that it appears in the expansion (5.17) for A^α_1.

For real fields the inner product is defined as above but with the terms involving $h^{(0)}$ and $\mathbf{h}^{(1)}$ omitted.

(i) Second Quantization without Lorentz Condition

We now second quantize the complex vector field without making use of the Lorentz condition. We thereby obtain a relativistic two-spin theory with a positive definite Hamiltonian and a positive definite norm for the Hilbert space.

Let us then replace the representatives $f^{(0)}(\mathbf{p})$ and $h^{(0)}(\mathbf{p})$ by destruction operators and $f^{(0)*}(\mathbf{p})$ and $h^{(0)*}(\mathbf{p})$ by creation operators. Likewise we replace the components $f^{(1)}(\mathbf{p}, \lambda)$ and $h^{(1)}(\mathbf{p}, \lambda)$ of the spin 1 representative vectors $\mathbf{f}^{(1)}(\mathbf{p})$ and $\mathbf{h}^{(1)}(\mathbf{p})$, respectively, by destruction operators and their complex conjugates by creation operators. Thus four particles are involved—a particle, antiparticle pair of spin zero and a particle, antiparticle pair of spin 1.

We assume Bose statistics and require that all commutators vanish except the following:

$$\begin{aligned} [f^{(0)}(\mathbf{p}), f^{(0)*}(\mathbf{p}')] &= \omega(p)\delta(\mathbf{p} - \mathbf{p}'), \\ [h^{(0)}(\mathbf{p}), h^{(0)*}(\mathbf{p}')] &= \omega(p)\delta(\mathbf{p} - \mathbf{p}'), \\ [f^{(1)}(\mathbf{p}, \lambda), f^{(1)*}(\mathbf{p}', \lambda')] &= \omega(p)\delta(\mathbf{p} - \mathbf{p}')\delta_{\lambda,\lambda'}, \\ [h^{(1)}(\mathbf{p}, \lambda), h^{(1)*}(\mathbf{p}', \lambda')] &= \omega(p)\delta(\mathbf{p} - \mathbf{p}')\delta_{\lambda,\lambda'}. \end{aligned} \tag{5.24}$$

We thus regard the components A^α as given by (5.17) with $C^{(i)}$ and $D^{(i)}$ given by (5.21) as an operator and denote the Hermitian adjoint of this operator by $A^{\alpha *}$. We readily derive the following commutation rules for the components of the vector potential:

$$[A^\alpha(\mathbf{x}, t), A^\beta(\mathbf{x}', t')] = 0,$$

$$[A^\alpha(\mathbf{x}, t), A^{\beta *}(\mathbf{x}', t')] = \frac{i}{m^2}\left[2\frac{\partial^2}{\partial x_\alpha \partial x_\beta} - m^2 g^{\alpha\beta}\right] D(\mathbf{x} - \mathbf{x}', t - t'), \tag{5.25}$$

where $D(\mathbf{x}, t)$ is defined by (3.21).

We now show that the theory is relativistically invariant. Let $\hat{A}$ be any one of the operators $\hat{H}$, $\hat{P}_i$, $\hat{J}_i$, $\hat{\mathfrak{J}}_i$ acting on the representatives corresponding to the irreducible representations of the inhomogeneous Lorentz group as in (2.16). Let us define the operators $\hat{A}f^{(0)}(\mathbf{p}), \hat{A}h^{(0)}(\mathbf{p}), \hat{A}f^{(1)}(\mathbf{p}, \lambda), \hat{A}h^{(1)}(\mathbf{p}, \lambda)$ as the operators formed when $\hat{A}$ acts on $\mathbf{p}$ and λ as though $f^{(0)}(\mathbf{p})$, $h^{(0)}(\mathbf{p}), f^{(1)}(\mathbf{p}, \lambda)$, $h^{(1)}(\mathbf{p}, \lambda)$ were representatives instead of destruction operators. For each operator $\hat{A}$ we define the second-quantized operator

$$[A] = \int \frac{d\mathbf{p}}{\omega(p)} f^{(0)*}(\mathbf{p})\hat{A}f^{(0)}(\mathbf{p}) + \int \frac{d\mathbf{p}}{\omega(p)} h^{(0)*}(\mathbf{p})\hat{A}h^{(0)}(\mathbf{p}) + \sum_\lambda \int \frac{d\mathbf{p}}{\omega(p)} f^{(1)*}(\mathbf{p}, \lambda)\hat{A}f^{(1)}(\mathbf{p}, \lambda) + \sum_\lambda \int \frac{d\mathbf{p}}{\omega(p)} h^{(1)*}(\mathbf{p}, \lambda)\hat{A}h^{(1)}(\mathbf{p}, \lambda). \tag{5.26}$$

The operators $[P_i]$, $[H]$, $[J_i]$, and $[\mathfrak{J}_i]$ constitute a representation of the infinitesimal generators of the inhomogeneous Lorentz group. The operators $[P_i]$ are the components of linear momentum, $[H]$ is the Hamiltonian, and $[J_i]$ and $[\mathfrak{J}_i]$ are the components of angular momentum.

We now show that the second-quantized theory is invariant under all the transformations of the inhomogeneous Lorentz group.

Let $A(x)$ denote the column vector whose components are $A^\alpha(\mathbf{x}, t)$. Then under the translation $T(a^\alpha)$ the set of operators called $A(x)$ transform to $A'(x)$ where

$$A'(x) = A(x + a) = \exp\{-i\sum_\alpha a^\alpha[P_\alpha]\}A(x)\exp\{i\sum_\alpha a^\alpha[P_\alpha]\}. \tag{5.27}$$

Equation (5.27) is to be regarded as true component by component as are the equations to follow.

Under the rotation $R(\boldsymbol{\theta})$ the new set of operators is given by

$$A'(x) = R(\boldsymbol{\theta})A(R(-\boldsymbol{\theta})x) = \exp\{-i\boldsymbol{\theta}\cdot[\mathbf{J}]\}A(x)\exp\{i\boldsymbol{\theta}\cdot[\mathbf{J}]\}. \tag{5.28}$$

In Eq. (5.28) $R(\boldsymbol{\theta})$ is the matrix defined by the four-vector rotation (2.2b).

Under the space–time transformation $L(\boldsymbol{\beta})$ the new set of operators is given by

$$A'(x) = L(\boldsymbol{\beta})A(L(-\boldsymbol{\beta})x) = \exp\{-i\boldsymbol{\beta}\cdot[\mathfrak{J}]\}A(x) \times \exp\{i\boldsymbol{\beta}\cdot[\mathfrak{J}]\}. \tag{5.29}$$

In (5.29) $L(\boldsymbol{\beta})$ is the matrix defined by the Lorentz transformation (2.3b).

To second quantize real fields we identify the operators $h^{(0)}(\mathbf{p})$ with $f^{(0)}(\mathbf{p})$ and $h^{(1)}(\mathbf{p}, \lambda)$ with $f^{(1)}(\mathbf{p}, \lambda)$ in the expansion (5.17). Thus $A(x)$ is a set of Hermitian operators. We need only consider commutation rules

$$[f^{(0)}(\mathbf{p}), f^{(0)}(\mathbf{p}')] = \omega(p)\delta(\mathbf{p} - \mathbf{p}'),$$
$$[f^{(1)}(\mathbf{p}, \lambda), f^{(1)*}(\mathbf{p}', \lambda')] = \omega(p)\delta(\mathbf{p} - \mathbf{p}')\delta_{\lambda,\lambda'}, \tag{5.30}$$
$$[f^{(0)}(\mathbf{p}), f^{(1)}(\mathbf{p}', \lambda)] = [f^{(0)}(\mathbf{p}), f^{(1)*}(\mathbf{p}', \lambda)] = 0.$$

The commutation rules for the operators A^α are

$$[A^\alpha(\mathbf{x}, t), A^\beta(\mathbf{x}', t')] = (i/m^2)\left[2\frac{\partial^2}{\partial x_\alpha \partial x_\beta} - m^2 g^{\alpha\beta}\right] \times D(\mathbf{x} - \mathbf{x}', t - t'). \tag{5.31}$$

Second-quantized operators are introduced as in (5.26) with terms containing $h^{(0)}(\mathbf{p})$ and $h^{(1)}(\mathbf{p}, \lambda)$ stricken out. Equations (5.27) through (5.29) hold.

(j) Second Quantization with Lorentz Condition

When the Lorentz condition is imposed we replace $h^{(0)}(\mathbf{p})$ and $f^{(0)}(\mathbf{p})$ in the expansion (5.17) by zero. For the complex field $f^{(1)}(\mathbf{p}, \lambda)$ and $h^{(1)}(\mathbf{p}, \lambda)$ satisfy the same commutation rules (5.24) as before. The commutation rules for the components of the vector potential are

$$[A^\alpha(\mathbf{x}, t), A^\beta(\mathbf{x}', t')] = 0,$$
$$[A^\alpha(\mathbf{x}, t), A^{\beta *}(\mathbf{x}', t')] = -i[g^{\alpha\beta} - (1/m^2)(\partial^2/\partial x_\alpha \partial x_\beta)]D(\mathbf{x} - \mathbf{x}', t - t'), \tag{5.32}$$

which are identical to the usual commutation rules (see Ref. 5). Hence our theory is equivalent to the usual one except that it is much easier to handle the Lorentz condition and is relativistic.

Second-quantized operators are introduced as in (5.26) except that terms containing $h^{(0)}$ and $f^{(0)}$ do not appear. Equations (5.27)–(5.29) continue to hold for transformations of the wavefunction under relativistic changes of frame.

To obtain the second-quantized theory of real fields we identify $h^{(1)}(\mathbf{p}, \lambda)$ with $f^{(1)}(\mathbf{p}, \lambda)$ in the expansion

(5.17). The commutation rules for the fields are

$$[A^\alpha(\mathbf{x}, t), A^\beta(\mathbf{x}', t')] = -i[g^{\alpha\beta} - (1/m^2)(\partial^2/\partial x_\alpha \partial x_\beta)]D(\mathbf{x} - \mathbf{x}', t - t'). \tag{5.33}$$

In second quantizing operators only the term involving $f^{(1)}(\mathbf{p}, \lambda)$ is used in the definition (5.26). The transformation equations (5.27)–(5.29) continue to hold.

6. WAVEFUNCTIONS WHICH TRANSFORM LIKE DIRAC SPINORS

We now reduce wavefunctions which transform like Dirac spinors. In a manner similar to our treatment of scalar wavefunctions and wavefunctions which transform as a vector, we first reduce spinor wavefunctions generally and then require that they satisfy the Dirac equation to eliminate some of the representations which appear. We then relate the solutions of the Dirac equation to the canonical formalism and second quantize the theory. We thus show that the usual second-quantization theory is recovered. However, the modes which we introduce transform in a simple manner in contrast to the modes which are usually introduced.

(a) Pauli Matrices, Dirac Matrices, Dirac Hamiltonian, and Dirac Equation

We wish first to introduce the Pauli and Dirac matrices in standard form. Accordingly we take the Pauli matrices $\sigma_i\,(i = 1, 2, 3)$ to be given in the conventional form

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{6.1}$$

We also define the Dirac matrices γ^α ($\alpha = 0, 1, 2, 3$) in a conventional way as a 4-by-4 matrix written as a two-by-two matrix with two-by-two matrix components as follows

$$\gamma^0 = -\gamma_0 = -i\begin{pmatrix} I_2 & 0_2 \\ 0_2 & -I_2 \end{pmatrix}, \qquad \gamma^i = \gamma_i = -i\begin{pmatrix} 0_2 & \sigma_i \\ -\sigma_i & 0_2 \end{pmatrix}. \tag{6.2}$$

In (6.2) 0_2 is the two-by-two zero matrix and I_2 is the two-by-two identity matrix. Also one should not confuse $i = (-1)^{\frac{1}{2}}$ with the subscript label i.

The Dirac-Hamiltonian H is given by

$$H = i\gamma^0(\boldsymbol{\gamma} \cdot \boldsymbol{\nabla} + m). \tag{6.3}$$

The Dirac wavefunction $\Psi(x)$ is a column vector with components $\Psi(\mathbf{x}, t, \gamma)$ in accordance with our usual notation, where the discrete variable γ runs from 1 to 4. (The discrete variable γ should not be confused with the Dirac matrices γ^α or γ_α which is always associated with a superscript or subscript.)

Of course, the Dirac equation is

$$H\Psi(x) = i(\partial/\partial t)\Psi(x). \tag{6.4}$$

(b) Transformation Properties of the Wavefunction

The requirement that the wave equation (6.4) be invariant leads to the transformation properties of the wavefunction $\Psi(x)$. For the time being we assume that the wavefunction does *not* necessarily satisfy (6.4) but that it has the same transformation properties required of the solution of (6.4). These transformation properties are summarized by giving the matrices M_i and N_i explicitly. These matrices are given in terms of the Pauli matrices as follows:

$$M_i = \begin{pmatrix} \frac{1}{2}\sigma_i & 0_2 \\ 0_2 & \frac{1}{2}\sigma_i \end{pmatrix}, \qquad N_i = \begin{pmatrix} 0_2 & -\frac{1}{2}i\sigma_i \\ -\frac{1}{2}i\sigma_i & 0_2 \end{pmatrix}. \tag{6.5}$$

We can see how $\Psi(x)$ transforms under the Lorentz transformations $T(a^\alpha)$, $R(\boldsymbol{\theta})$, and $L(\boldsymbol{\beta})$ by using (2.10)–(2.12) with

$$\exp[i\boldsymbol{\theta} \cdot \mathbf{M}] = I\cos\tfrac{1}{2}\theta + 2i(\boldsymbol{\theta} \cdot \mathbf{M})(\sin\tfrac{1}{2}\theta/\theta), \qquad \exp[i\boldsymbol{\beta} \cdot \mathbf{N}] = I\cosh\tfrac{1}{2}\beta + 2i(\boldsymbol{\beta} \cdot \mathbf{N})(\sinh\tfrac{1}{2}\beta/\beta). \tag{6.6}$$

In (6.6), I is the four-by-four identity matrix.

(c) Transformation to Irreducible Representations

From the first of Eqs. (6.5) we see that the label (r) takes on two values which we can take to be 1 and 2. For each value of (r) the label λ takes on two values which we also take to be 1 and 2. We then can expand $\Psi(x)$ as in (2.24) and (2.24a). On using (2.23) and the second of equations (6.6) we obtain as the components of the transformation vector $\chi^{(r)}(\mu, \epsilon, \mathbf{p}, \lambda)$

$$\chi^{(1)}(\mu, \epsilon, \mathbf{p}, 1) = (2\mu)^{-\frac{1}{2}}\begin{pmatrix} [\omega(\mu,p) + \mu]^{\frac{1}{2}} \\ 0 \\ [\omega(\mu,p) + \mu]^{-\frac{1}{2}}\epsilon p_3 \\ [\omega(\mu,p) + \mu]^{-\frac{1}{2}}\epsilon(p_1 + ip_2) \end{pmatrix},$$

$$\chi^{(1)}(\mu, \epsilon, \mathbf{p}, 2) = (2\mu)^{-\frac{1}{2}}\begin{pmatrix} 0 \\ [\omega(\mu,p) + \mu]^{\frac{1}{2}} \\ [\omega(\mu,p) + \mu]^{-\frac{1}{2}}\epsilon(p_1 - ip_2) \\ -[\omega(\mu,p) + \mu]^{-\frac{1}{2}}\epsilon p_3 \end{pmatrix}, \tag{6.7}$$

$$\chi^{(2)}(\mu, \epsilon, \mathbf{p}, 1) = (2\mu)^{-\frac{1}{2}} \begin{pmatrix} [\omega(\mu,p)+\mu]^{-\frac{1}{2}}\epsilon p_3 \\ [\omega(\mu,p)+\mu]^{-\frac{1}{2}}\epsilon(p_1+ip_2) \\ [\omega(\mu,p)+\mu]^{\frac{1}{2}} \\ 0 \end{pmatrix},$$

$$\chi^{(2)}(\mu, \epsilon, \mathbf{p}, 2) = (2\mu)^{-\frac{1}{2}} \begin{pmatrix} [\omega(\mu,p)+\mu]^{-\frac{1}{2}}\epsilon(p_1-ip_2) \\ -[\omega(\mu,p)+\mu]^{-\frac{1}{2}}\epsilon p_3 \\ 0 \\ [\omega(\mu,p)+\mu]^{\frac{1}{2}} \end{pmatrix}. \tag{6.8}$$

In the expansion (2.24a) we define, as usual, $\chi^{(r)}(\mu, \mathbf{p}, \lambda)$ by

$$\chi^{(r)}(\mu, \mathbf{p}, \lambda) = \chi^{(r)}(\mu, +1, \mathbf{p}, \lambda). \tag{6.9}$$

The matrix U whose elements are used in (2.26) is given by

$$U = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{6.10}$$

Then the transformation column vectors $\zeta^{(r)}(\mu, \mathbf{p}, \lambda)$, which are used in (2.24a), are given by

$$\zeta^{(1)}(\mu, \mathbf{p}, 1) = (2\mu)^{-\frac{1}{2}} \begin{pmatrix} 0 \\ [\omega(\mu,p)+\mu]^{\frac{1}{2}} \\ [\omega(\mu,p)+\mu]^{-\frac{1}{2}}(p_1+ip_2) \\ -[\omega(\mu,p)+\mu]^{-\frac{1}{2}}p_3 \end{pmatrix},$$

$$\zeta^{(1)}(\mu, \mathbf{p}, 2) = (2\mu)^{-\frac{1}{2}} \begin{pmatrix} -[\omega(\mu,p)+\mu]^{\frac{1}{2}} \\ 0 \\ -[\omega(\mu,p)+\mu]^{-\frac{1}{2}}p_3 \\ -[\omega(\mu,p)+\mu]^{-\frac{1}{2}}(p_1-ip_2) \end{pmatrix},$$

$$\zeta^{(2)}(\mu, \mathbf{p}, 1) = (2\mu)^{-\frac{1}{2}} \begin{pmatrix} [\omega(\mu,p)+\mu]^{-\frac{1}{2}}(p_1+ip_2) \\ -[\omega(\mu,p)+\mu]^{-\frac{1}{2}}p_3 \\ 0 \\ [\omega(\mu,p)+\mu]^{\frac{1}{2}} \end{pmatrix},$$

$$\zeta^{(2)}(\mu, \mathbf{p}, 2) = (2\mu)^{-\frac{1}{2}} \begin{pmatrix} -[\omega(\mu,p)+\mu]^{-\frac{1}{2}}p_3 \\ -[\omega(\mu,p)+\mu]^{-\frac{1}{2}}(p_1-ip_2) \\ -[\omega(\mu,p)+\mu]^{\frac{1}{2}} \\ 0 \end{pmatrix}. \tag{6.11}$$

As in our treatments of scalar and vector wavefunctions we work with (2.24a) instead of (2.24) in the next sections.

(d) Wave Equation

We now require the wavefunction $\Psi(x)$ to satisfy the Dirac equation (6.4). Since the wavefunctions which satisfy (6.4) also satisfy the wave equation

$$[(\partial^2/\partial t^2) - \nabla^2 + m^2]\Psi(x) = 0, \tag{6.12}$$

it follows in precisely the same fashion as for the scalar and vector wave equations that only the mass m appears in the expansion (2.24a), that is,

$$\begin{aligned} dM^{(r)}(\mu) &= C^{(r)}\delta(\mu - m)\, d\mu, \\ dN^{(r)}(\mu) &= D^{(r)}\delta(\mu - m), \end{aligned} \tag{6.13}$$

where $C^{(r)}$ and $D^{(r)}$ are positive constants. However, this restriction on the representations is not the only one for wavefunctions which satisfy (6.4). If we substitute the expansion (2.22a) into (6.4) and use (6.7), (6.8), (6.9), (6.11), and (6.13) we find that we must have

$$f^{(2)}(m, \mathbf{p}, \lambda) = h^{(1)}(m, \mathbf{p}, \lambda) = 0. \tag{6.14}$$

Thus wavefunctions which transform like Dirac spinors of one mass are more general than Dirac spinors which satisfy the Dirac equation.

In what follows we consider only solutions of the Dirac equations. For simplicity of notation we write

$$\begin{aligned} f^{(1)}(m, \mathbf{p}, \lambda) &= f(\mathbf{p}, \lambda), \\ h^{(2)}(m, \mathbf{p}, \lambda) &= h(\mathbf{p}, \lambda). \end{aligned} \tag{6.15}$$

$$\begin{aligned} C^{(1)} &= C, \\ D^{(2)} &= D. \end{aligned} \tag{6.16}$$

(e) Connection with the Canonical Formalism. Definition of Inner Product

We now find the constants C and D by identifying the total energy of the field as obtained from the canonical formalism with the expectation value of the energy in terms of the representatives corresponding to the irreducible representations.

Let us first set $h(\mathbf{p}, \lambda) = 0$ and take the usual expression for the Hamiltonian density, namely

$$H(\mathbf{x}) = \Psi^\dagger(x)H\Psi(x) = i\Psi^\dagger(x)\dot{\Psi}(x), \tag{6.17}$$

where $H\Psi(x)$ is the column vector obtained by operating with the Hamiltonian H on the vector $\Psi(x)$ and where $\Psi^\dagger$ is the complex conjugate of the row vector formed from Ψ. The first equation of (6.17) is thus meant to represent a bilinear form. We require that

$$\int H(\mathbf{x})\, d\mathbf{x} = \sum_\lambda \int |f(\mathbf{p}, \lambda)|^2\, d\mathbf{p}. \tag{6.18}$$

We then set $f(\mathbf{p}, \lambda) = 0$ and using $-H(\mathbf{x})$ as the Hamiltonian density require that

$$-\int H(\mathbf{x})\, d\mathbf{x} = \sum_{\lambda}\int |h(\mathbf{p}, \lambda)|^2\, d\mathbf{p}. \tag{6.19}$$

In this way we obtain the following results for C and D.

$$C = D = [m/(2\pi)^3]^{\frac{1}{2}}. \tag{6.20}$$

It is not difficult to define an inner product of two spinors which satisfy the Dirac equation. In complete analogy to the scalar and vector wavefunctions we define the inner product (Ψ_1, Ψ) by

$$(\Psi_1, \Psi) = \sum_{\lambda}\int \frac{d\mathbf{p}}{\omega(m, p)} \times [f_1^*(\mathbf{p}, \lambda) f(\mathbf{p}, \lambda) + h^*(\mathbf{p}, \lambda) h_1(\mathbf{p}, \lambda)]. \tag{6.21}$$

(f) Second Quantization

We now assume that the spinor wavefunction satisfies the Dirac equation and that C and D are given by (6.20). To second quantize the theory we now regard $f(\mathbf{p}, \lambda)$ and $h(\mathbf{p}, \lambda)$ as destruction operators and $f^*(\mathbf{p}, \lambda)$ and $h^*(\mathbf{p}, \lambda)$ as creation operators which satisfy the following fermion anticommutation rules:

$$\begin{aligned}
[f(\mathbf{p}, \lambda), f(\mathbf{p}', \lambda')]_+ &= [h(\mathbf{p}, \lambda), h(\mathbf{p}', \lambda')]_+ = 0,\\
[f(\mathbf{p}, \lambda), h(\mathbf{p}', \lambda')]_+ &= [f(\mathbf{p}, \lambda), h^*(\mathbf{p}', \lambda')]_+ = 0,\\
[f(\mathbf{p}, \lambda), f^*(\mathbf{p}', \lambda')]_+ &= [h(\mathbf{p}, \lambda), h^*(\mathbf{p}', \lambda')]_+\\
&= \omega(m, p)\delta(\mathbf{p} - \mathbf{p}')\delta_{\lambda,\lambda'}.
\end{aligned} \tag{6.22}$$

On using the expansion corresponding to (2.24a) for the operator $\Psi(\mathbf{x}, t, \gamma)$ and the Hermitian adjoint operator $\Psi^*(\mathbf{x}, t, \gamma)$ we obtain the usual commutation rules for these spinor components

$$\begin{aligned}
&[\Psi(\mathbf{x}, t, \gamma), \Psi(\mathbf{x}', t', \gamma')]_+ = 0,\\
&[\Psi(\mathbf{x}, t, \gamma), \Psi^*(\mathbf{x}', t', \gamma')]_+\\
&\quad = i[(1/i)(\partial/\partial t)I - H]_{\gamma,\gamma'} D(\mathbf{x} - \mathbf{x}', t - t').
\end{aligned} \tag{6.23}$$

The subscript on the square bracket labels matrix elements in the usual fashion. I is the identity operator and H is the Hamiltonian (6.3) which operates on the $\mathbf{x}$ variable of the invariant function.

For every infinitesimal generator $\hat{A}$ of the inhomogeneous Lorentz group we can introduce a second-quantized operator $[A]$ in a manner similar to that for the vector and scalar wavefunctions. These second-quantized operators also constitute a representation of the infinitesimal generators of the inhomogeneous Lorentz group. In particular $[H]$ is the second-quantized energy which is positive definite. The discussion of the invariance of the second-quantized theory is close to that for the scalar and vector wave equations and we therefore do not give it for the sake of brevity.

7. A MULTIPLE SPIN DIRAC EQUATION

There have been several generalizations of the Dirac equation. One of the best known which describes particles of higher spin is given in Ref. 9. Such equations are shown to give rise to wavefunctions which transform according to irreducible representations of the inhomogeneous Lorentz group. (It might be mentioned that, using the techniques of the present paper, the representations given in Ref. 9 can be converted into the Foldy–Shirokov representations.)

In the present section, however, we consider a generalization of the Dirac equation given in Ref. 10. In Ref. 10 it is required that the generalized Dirac Hamiltonian have the form (6.3) where the gamma matrices satisfy the usual Dirac rules. One further requires that the matrices γ^α and the matrices M_i and N_i constitute an irreducible set. All such irreducible sets are then found.

It is shown that in general the solutions of the generalized Dirac equation correspond to multiple spin particles. However, by choosing the representation properly, single spin theories are obtained by setting some of the components of the spinor wavefunction equal to zero.

In the present section we consider the simplest generalized Dirac equation after the original Dirac equation itself. We first reduce the wavefunction which transforms according to the rules of Ref. 10, even when the wavefunction does not satisfy the Dirac equation. It may be seen that our knowledge of the transformation properties of scalar, electromagnetic, and vector wavefunctions enables us to reduce the wavefunction immediately.

We then require that the wavefunction satisfy the wave equation and thereby eliminate some of the irreducible representations which appear.

We find that we obtain essentially the same two-spin theory that we obtained for the vector wavefunction which satisfies the second-order wave equation when the Lorentz condition is not imposed. A single-spin theory is then obtained by setting the wavefunction corresponding to a scalar equal to zero.

(a) Matrices for the Generalized Dirac Equation. Reduction of the Wavefunction

In our example the spinor wavefunction $\Psi(x)$ is a column vector with eight components. The γ matrices and the matrices M_i and N_i are all eight-by-eight matrices. We rewrite the results of Ref. 9 slightly to conform to the notation of the present paper.

---

[9] V. Bargmann and E. P. Wigner, Proc. Natl. Acad. Sci. U.S. **34**, 211 (1948).

[10] J. S. Lomont and H. E. Moses, Phys. Rev. **118**, 337 (1960).

The γ matrices are given by

$$\gamma_0 = \begin{pmatrix} 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{7.1}$$

$$\gamma_1 = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -i & 0 \\ 0 & 0 & 0 & 0 & 0 & i & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & -i & 0 & 0 & 0 & 0 & 0 \\ 0 & i & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}, \tag{7.2}$$

$$\gamma_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & i & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -i & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & i & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ -i & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}, \tag{7.3}$$

$$\gamma_3 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & -i & 0 & 0 \\ 0 & 0 & 0 & 0 & i & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -i & 0 & 0 & 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}. \tag{7.4}$$

The matrices M_i and N_i appear as completely reduced matrices in which the irreducible representations are those for the scalar, (Sec. 3), the electromagnetic field (the matrices M_i, N_i of Sec. 4), and the four-vector of Sec. 5. That is, the eight-by-eight matrices M_i and N_i for the present case are written as

$$M_i = \begin{pmatrix} \hat{S}_i & & & \\ & 0_1 & & \\ & & \hat{S}_i & \\ & & & 0_1 \end{pmatrix}, \qquad N_i = \begin{pmatrix} i\hat{S}_i & & & \\ & \hat{N}_i & & \\ & & & 0_1 \end{pmatrix}. \tag{7.5}$$

In (7.5) the symbol 0_1 is the one-by-one zero matrix.

It is now a simple matter to express the transformation properties of the eight-component spinor $\Psi(x)$. From the top three components of the spinor let us construct the vector wavefunction $\boldsymbol{\Psi}$. Let us denote the fourth component by V and form the vector $\mathbf{A}$ from the fifth, sixth, and seventh components. Finally the function φ denotes the last component. To summarize we write

$$\Psi(x) = \begin{pmatrix} \boldsymbol{\Psi} \\ V \\ \mathbf{A} \\ \varphi \end{pmatrix}. \tag{7.6}$$

Thus under the transformations of the inhomogeneous Lorentz group φ transforms as a scalar as in Sec. 3, $\boldsymbol{\Psi}$ transforms as the wavefunction treated in Sec. 4, and the functions V, $\mathbf{A}$ transform as the four-vector of Sec. 5. We expand these quantities in terms of the irreducible representations as earlier. We consider (3.2) with Ψ replaced by φ, (4.15) and (5.9a) as giving the expansion in terms of irreducible representations in the form in which we wish to use them. We see that for each mass μ there are four scalar wavefunctions, namely, $f(\mu, \mathbf{p})$, $h(\mu, \mathbf{p})$, $f^{(0)}(\mu, \mathbf{p})$, $h^{(0)}(\mu, \mathbf{p})$, and four vector wavefunctions, namely, $\mathbf{f}(\mu, \mathbf{p})$, $\mathbf{h}(\mu, \mathbf{p})$, $\mathbf{f}^{(1)}(\mu, \mathbf{p})$, and $\mathbf{h}^{(1)}(\mu, \mathbf{p})$, all of positive energy.

(b) Solution of the Generalized Dirac Equation

We now consider the spinor $\Psi(x)$ to be a solution of the generalized Dirac equation. For exactly the same

reason as for the usual Dirac equation only the mass m appears in the decomposition to the irreducible representations. It is no longer necessary to integrate over the variable μ. For simplicity we take the constants C and D, which appear in the expansions when only single masses are used, to be 1. This restriction is not severe, since these constants could always be absorbed in the wavefunction.

Having used the fact that the Dirac equation permits only representations belonging to the mass m, we substitute $\Psi(x)$ in the form (7.6) into the Dirac equation, using the γ matrices of the present section.

The Dirac equation is completely equivalent to the following set of equations.

$$\begin{aligned}(\partial/\partial t)\mathbf{A} + \nabla V + i\nabla \times \mathbf{A} + m\Psi &= 0,\\ -(\partial/\partial t)\varphi + \nabla \cdot \Psi + mV &= 0,\\ -(\partial/\partial t)\Psi + \nabla\varphi + i\nabla \times \Psi + m\mathbf{A} &= 0,\\ (\partial/\partial t)V + \nabla \cdot \mathbf{A} + m\varphi &= 0.\end{aligned} \tag{7.7}$$

On substituting the expansions for the various functions and vectors we obtain the following conditions on the representatives:

$$\begin{aligned}f(m, \mathbf{p}) &= if^{(0)}(m, \mathbf{p}),\\ h(m, \mathbf{p}) &= ih^{(0)}(m, \mathbf{p}),\\ \mathbf{f}(m, \mathbf{p}) &= i\mathbf{f}^{(1)}(m, \mathbf{p}),\\ \mathbf{h}(m, \mathbf{p}) &= i\mathbf{h}^{(1)}(m, \mathbf{p}).\end{aligned} \tag{7.8}$$

Thus the number of independent representations have been reduced considerably.

Finally, to get a spin 1 theory only we must have

$$h^{(0)} = f^{(0)} = 0. \tag{7.9}$$

It is easily seen that condition (7.9) is a necessary and sufficient condition for the function φ to be identically zero.

APPENDIX. DERIVATION OF THE ALGORITHM FOR REDUCING RELATIVISTIC WAVEFUNCTIONS

We now prove the basic algorithm of the present paper, namely Eqs. (2.23) and (2.24). We use the "recipe" for reduction given in Ref. 1 and simply identify the various quantities which appear in that reference which we assume that the reader has before him.

First of all, the variables collectively denoted by ζ in Ref. 1 are identified with the variables $\mathbf{x}$, t, and γ. The function $f(\zeta)$ is then identified with the components of the wavefunction $\Psi(\mathbf{x}, t, \gamma)$.

We seek the transformation function $\langle \zeta \mid \mu, \epsilon, \mathbf{p}, \lambda\rangle$ of Ref. 1 which we now call $\langle \mathbf{x}, t, \gamma \mid \mu, \epsilon, \mathbf{p}, \lambda\rangle$. We must also find the spin matrices denoted in Ref. 1 by $S_i(\lambda \mid \lambda')$.

To find the transformation function we turn to Eq. (1.16) of Ref. 1 and we show how to choose the function denoted there by $g(\zeta \mid \mu, \epsilon, \mathbf{p}, \lambda)$ which we now call $g(\mathbf{x}, t, \gamma \mid \mu, \epsilon, \mathbf{p}, \lambda)$. If A is an operator, we mean by A^ζ the operator as it acts on the variables $\mathbf{x}$, t, γ. These operators are given by (2.13)–(2.15) of the present paper.

The first of Eqs. (1.16) of Ref. 1 becomes

$$(\partial/\partial x_i)g(\mathbf{x}, t, \gamma \mid \mu, \epsilon, \lambda) = 0, \tag{A1}$$

from which it is clear that g is independent of $\mathbf{x}$; we thus write

$$g(\mathbf{x}, t, \gamma \mid \mu, \epsilon, \lambda) = G(t, \gamma \mid \mu, \epsilon, \lambda) \tag{A2}$$

to emphasize this independence.

The second of Eqs. (1.16) of Ref. 1 leads to

$$i(\partial/\partial t)G(t, \gamma \mid \mu, \epsilon, \lambda) = \epsilon\mu G(t, \gamma \mid \mu, \epsilon, \lambda). \tag{A3}$$

From (A2) and (A3)

$$g(\mathbf{x}, t, \gamma \mid \mu, \epsilon, \mathbf{p}, \gamma) = e^{-i\epsilon\mu t}g(\gamma \mid \mu, \epsilon, \lambda), \tag{A4}$$

where $g(\gamma \mid \mu, \epsilon, \lambda)$ is the constant of integration which comes from (A3).

We must obtain all linearly independent solutions in λ of $g(x, t, \gamma \mid \mu, \epsilon, \lambda)$. The variable λ is not unique, as explained in Ref. 1. A very convenient way of choosing λ is to make it a discrete variable which goes through the same range of values as γ and choose the constant of integration $g(\gamma \mid \mu, \epsilon, \lambda)$ to be given by

$$g(\gamma \mid \mu, \epsilon, \lambda) = C(\mu, \epsilon)\delta_{\gamma,\lambda}. \tag{A5}$$

We may take $C(\mu, \epsilon)$ to be positive but otherwise arbitrary and may even have δ functions in μ. Indeed we may take $C(\mu, \epsilon)$ to have any value including unity or depend on λ, for example. Each choice will lead to a different—but equivalent (for a given mass spectrum)—decomposition theorem (2.24). The choice we have made seems the simplest for our purposes and includes all possible spectra for μ. On using Eq. (1.18) of Ref. 1 (which is the principal formula of the recipe) we obtain [using also (2.12) and (2.12a) of the present paper] a surprisingly simple result:

$$\begin{aligned}\langle \mathbf{x}, t, \gamma \mid \mu, \epsilon, \mathbf{p}, \lambda\rangle &= C(\mu, \epsilon)\\ &\times \exp\{i[\mathbf{p} \cdot \mathbf{x} - \epsilon\omega(\mu, p)t]\}\{\exp[-i\boldsymbol{\nu} \cdot \mathbf{N}]\}_{\gamma,\lambda}.\end{aligned} \tag{A6}$$

It is easily seen using Ref. 1 that the spin matrices S_i are identical to the matrices M_i and hence the spin

matrices are in reduced form because of the assumptions which we have made in setting up the configuration space representation.

The functions $f^{(r)}(\mu, \epsilon, \mathbf{p}, \lambda)$ are formed from the representatives $F(\mu, \epsilon, \mathbf{p}, \lambda)$ by restricting λ in F to take on only values corresponding to the rth block in the decomposition of M_i.

A measure function $M(\mu, \epsilon)$ is obtained from

$$dM(\mu, \epsilon) = C(\mu, \epsilon)\, d\mu.$$

The measure function $M^{(r)}(\mu, \epsilon)$ is obtained from $M(\mu, \epsilon)$ on redefining $f^{(r)}$ somewhat. We refrain from belaboring the point.

Leading Landau Curves for a Larger Class of Feynman Diagrams

MAMATA PATTNAIK
Saha Institute of Nuclear Physics, Calcutta, India

(Received 14 September 1966)

The leading parts of Landau curves for a larger class of Feynman diagrams are found to give no singularities on the physical sheet provided certain conditions are imposed upon the external and internal masses. The number of such masses involved is found to be fixed, no matter how complicated a member of this class is considered.

1. INTRODUCTION

In a recent paper Islam[1] has found out that the leading Landau curves of certain Feynman diagrams do not give singularities on the physical sheet if the external and internal masses satisfy certain inequalities. His paper includes: (a) A class of diagrams obtained by twisting a ladder with $n+1$ rungs around one side of it which is subsequently joined to a one-particle exchange diagram; (b) some other diagrams formed out of a twisted square and double square joined in a similar fashion. In the first case, Islam found some conditions imposed upon the masses which are identically satisfied in the equal-mass case. However, in the second case he found one extra condition which remained unsatisfied in the equal-mass case.

In the present paper we want to apply his procedure to a larger class of Feynman diagrams, to which "some other diagrams" of Islam form a subclass. To get this new class of diagram one takes two ladders, one with m and another with n straight rungs, which are twisted around one side of respective ladders. Then the two free ends of one ladder are joined to the two free ends of the other ladder.

As expected, we also find here a set of mass inequalities. In the equal-mass case the validity of some of the inequalities depends upon the particular value of n and m chosen. For example, we find, if $n, m > 2$, all but two such inequalities; if $n = 2$, $m > 2$ or vice versa we find all but one such inequality; if $n = 2$, $m = 2$ all such inequalities are satisfied in the equal-mass cases. (The absence of singular acnodes[2,3] is assumed.) Nevertheless this is not very unpromising, since the number of masses involved is fixed no matter how large the values of n and m are taken. The motivation behind this type of work has been well explained in previous papers,[1,4] so we need not explain it again.

[1] J. N. Islam, J. Math. Phys. **7**, 652 (1966).
[2] R. J. Eden, P. V. Landshoff, J. C. Polkinghorne, and J. C. Taylor, J. Math. Phys. **2**, 656 (1961).
[3] J N. Islam, J. Math. Phys. **4**, 872 (1963).
[4] J. N. Islam, Nuovo Cimento **30**, 259 (1963).

2. A CLASS OF FEYNMAN DIAGRAMS

In this section we discuss the leading Landau curves of a class of Feynman diagrams, a typical member of which is shown in Fig. 1. This is formed out of two twisted ladders, one with n and another with m rungs, in a manner already explained in the Introduction. The momenta, masses, and the Feynman

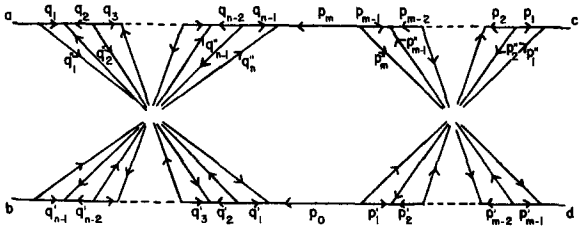


FIG. 1. A member of a class of Feynman diagrams under consideration.

parameters associated with the first twisted ladder are denoted as q_i, q_i', q_j''; $m_{q_i}, m_{q_i}', m_{q_j}''$; $\alpha_i, \alpha_i', \alpha_j''$ (where $i = 1, 2, \cdots, n-1, j = 1, 2, \cdots, n$); that of second are, p_k, p_k', p_l''; $m_{p_k}, m_{p_k}', m_{p_l}''$; $\beta_k, \beta_k', \beta_l''$ (where $k = 1, 2, \cdots, m-1$ and $l = 1, 2, \cdots, m$); that of joined ends are denoted by p_m, p_0; β_m, β_0; m_{pm} and m_{p0}, respectively. For convenience we restrict ourselves to the case for which both n and m are even. For both odd or one odd and another even, the procedure is exactly similar. In the latter case, however, the arrows at each vertex are not all incoming or outgoing.

We know that the singularity corresponding to a Feynman diagram lies on a curve which is given by a set of Landau equations[5]

$$\sum_i \alpha_i q_i = 0,$$

around each independent loop.

$$\sum_i \alpha_i (q_i^2 - m_i^2) = 0,$$

where q_i, m_i, and α_i are the momentum, mass, and Feynman parameter associated with the ith line.

[5] L. D. Landau, Nucl. Phys. **13**, 181 (1959).

Since in the diagrams under consideration there are $n+m-1$ independent loops, we get $n+m-1$ Landau equations which are given below.

$$\begin{aligned}
&\beta_1 p_1 + \beta_1' p_1' - \beta_1'' p_1'' - \beta_2'' p_2'' = 0,\\
&\beta_2 p_2 + \beta_2' p_2' - \beta_2'' p_2'' - \beta_3'' p_3'' = 0,\\
&\cdots\\
&\cdots\\
&\beta_{m-1} p_{m-1} + \beta_{m-1}' p_{m-1}' - \beta_{m-1}'' p_{m-1}'' - \beta_m'' p_m'' = 0,\\
&-\beta_1' p_1' + \beta_2' p_2' - \cdots - \beta_{m-1}' p_{m-1}' + \beta_m'' p_m''\\
&\qquad + \alpha_1 q_1 - \alpha_2 q_2 + \cdots + \alpha_{n-1} q_{n-1}\\
&\qquad\qquad - \alpha_1'' q_1'' + \beta_0 p_0 - \beta_m p_m = 0,\\
&\alpha_1 q_1 + \alpha_1' q_1' - \alpha_1'' q_1'' - \alpha_2'' q_2'' = 0,\\
&\alpha_2 q_2 + \alpha_2' q_2' - \alpha_2'' q_2'' - \alpha_3'' q_3'' = 0,\\
&\cdots\\
&\cdots\\
&\alpha_{n-1} q_{n-1} + \alpha_{n-1}' q_{n-1}' - \alpha_{n-1}'' q_{n-1}'' - \alpha_n'' q_n'' = 0.
\end{aligned} \tag{1}$$

Besides these Landau equations there are $2(n+m-2)$ conservation equations at the vertices. These are

$$\begin{aligned}
&p_0 + p_1' + p_1'' = 0,\\
&p_m + p_{m-1} + p_m'' = 0,\\
&p_{s-1} + p_s + p_s'' = 0, \quad 2 \le s \le m-1,\\
&p_{s-1}' + p_s' + p_s'' = 0, \quad 2 \le s \le m-1,
\end{aligned} \tag{2}$$

and

$$\begin{aligned}
&q_1' + q_1'' + p_0 = 0,\\
&q_{n-1} + q_n'' + p_m = 0,\\
&q_{r-1} + q_r + q_r'' = 0, \quad 2 \le r \le n-1,\\
&q_{r-1}' + q_r' + q_r'' = 0, \quad 2 \le r \le n-1.
\end{aligned}$$

On rearranging, one writes

$$\begin{aligned}
&p_m'' = -(p_{m-1} + p_m),\\
&p_s'' = -(p_{s-1} + p_s), && 2 \le s \le m-1,\\
&p_s' = (-1)^s (p_1 - p_1') + p_s, && 2 \le s \le m-1,\\
&q_n'' = -(p_m + q_{n-1}),\\
&q_r'' = -(q_{r-1} + q_r), && 2 \le r \le n-1,\\
&q_r' = (-1)^r (q_1 - q_1') + q_r, && 2 \le r \le n-1.
\end{aligned} \tag{3}$$

In order to find out the equation of the Landau curve at the vertex whose momentum conservation is given by the first of Eq. (2), i.e.,

$$p_0 + p_1' + p_1'' = 0,$$

we multiply it by p_1 and obtain

$$p_0 p_1 + p_1 p_1' + p_1 p_1'' = 0, \tag{4}$$

where $p_1 \cdot p_1''$ is a constant determined by masses m_{p_1}, m_{p_1}'', and m_c

$$p_1 \cdot p_1'' = \tfrac{1}{2}(m_c^2 - m_{p_1}'^2 - m_{p_1}''^2), \tag{5}$$

To get $p_1 \cdot p_1'$ we multiply the first of Eq. (1) by p_1' and find

$$p_1 \cdot p_1' = -(1/\beta_1)[\beta_1 p_1'^2 - \beta_1'' p_1' \cdot p_1'' - \beta_2'' p_1' \cdot p_2'']. \tag{6}$$

Next we wish to find out $p_1 \cdot p_0$. We multiply all of Eq. (1) and all of Eq. (2) except the first one [(2(n + m) − 5) vertex conservation equations] by p_0. After a little algebraic manipulation by the use of Eq. (3) we get $n+m-1$ equations, from which $n+m-1$ unknowns $p_i \cdot p_0$ ($i = 1, 2, \cdots, m$) and $q_j \cdot p_0$ ($j = 1, \cdots, n-1$) are determined. These equations are

$$\begin{aligned}
&(\beta_1 + \beta_2'') p_1 \cdot p_0 + \beta_2'' p_2 \cdot p_0 = -\beta_1' p_1' \cdot p_0 + \beta_1'' p_1'' \cdot p_0,\\
&(\beta_2' + \beta_2'') p_1 \cdot p_0 + M_2 p_2 \cdot p_0 + \beta_3'' p_3 \cdot p_0 = \beta_2' p_1' \cdot p_0,\\
&\cdots\\
&\cdots\\
&(-1)^s \beta_s' p_1 \cdot p_0 + \beta_s'' p_{s-1} \cdot p_0 + M_s p_s \cdot p_0\\
&\qquad + \beta_{s+1}'' p_{s+1} \cdot p_0 = (-1)^s \beta_s' p_0' \cdot p_0,\\
&\cdots\\
&\cdots\\
&-\beta_{m-1}' p_1 \cdot p_0 + \beta_{m-1}'' p_{m-2} \cdot p_0 + M_{m-1} p_{m-1} \cdot p_0\\
&\qquad + \beta_m'' p_m \cdot p_0 = -\beta_{m-1}' p_1' \cdot p_0,\\
&\sum_{s=2}^{m-1} \beta_s' p_1 \cdot p_0 + \beta_2' p_2 \cdot p_0 - \beta_3' p_3 \cdot p_0 + \cdots\\
&\quad + \beta_{m-2}' p_{m-2} \cdot p_0 - (\beta_{m-1}' + \beta_m'') p_{m-1} \cdot p_0\\
&\quad - (\beta_m + \beta_m'') p_m \cdot p_0 + \alpha_1 q_1 \cdot p_0 - \alpha_2 q_2 \cdot p_0 + \cdots\\
&\quad + \alpha_{n-1} q_{n-1} \cdot p_0\\
&= -\beta_0 p_0^2 + \sum_{s=1}^{m-1} \beta_s' p_1' \cdot p_0 + \alpha_1'' q_1'' \cdot p_0,
\end{aligned} \tag{7}$$

$$\begin{aligned}
&(\alpha_1 + \alpha_2'') q_1 \cdot p_0 + \alpha_2'' q_2 \cdot p_0 = -\alpha_1' q_1' \cdot p_0 + \alpha_1' q_1'' \cdot p_0,\\
&(\alpha_2' + \alpha_2'') q_1 \cdot p_0 + N_2 q_2 \cdot p_0 + \alpha_3'' q_3 \cdot p_0 = \alpha_2' q_1' \cdot p_0,\\
&\cdots\\
&\cdots\\
&(-1)^r \alpha_r' q_1 \cdot p_0 + \alpha_r'' q_{r-1} \cdot p_0 + N_r q_r \cdot p_0 + \alpha_{r+1}'' q_{r+1} \cdot p_0\\
&\qquad = (-1)^r \alpha_r' q_1' \cdot p_0,\\
&\cdots\\
&\cdots\\
&\alpha_n'' p_m \cdot p_0 - \alpha_{n-1}' q_1 \cdot p_0 + \alpha_{n-1}'' q_{n-2} \cdot p_0 + N_{n-1} q_{n-1} \cdot p_0\\
&\qquad = -\alpha_{n-1}' q_1' \cdot p_0,
\end{aligned}$$

where

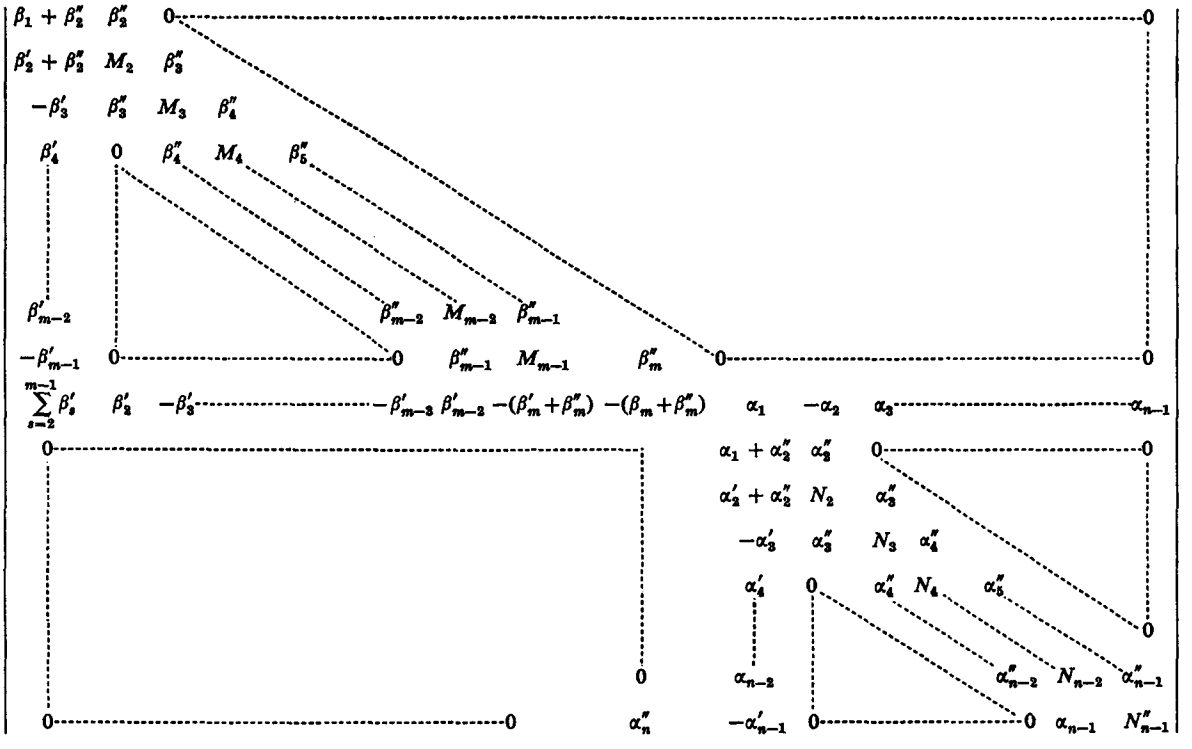
$$\begin{aligned}
M_s &= \beta_s + \beta_s' + \beta_s'' + \beta_{s+1}'',\\
N_r &= \alpha_r + \alpha_r' + \alpha_r'' + \alpha_{r+1}''.
\end{aligned}$$

Using Cramer's rule we obtain the solution for $p_i \cdot p_0$ in determinantal form

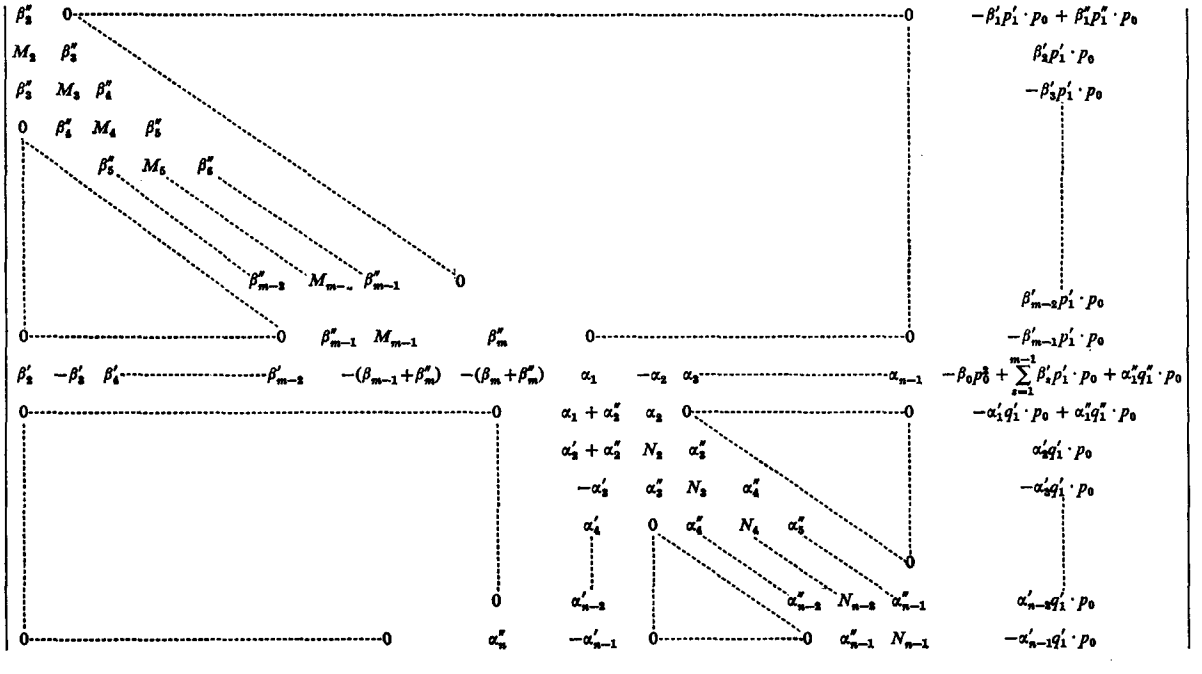
$$p_1 \cdot p_0 = B_{m,n}'/A_{m,n}', \tag{8}$$

where

$$A'_{m,n} =$$



$$B'_{m,n} =$$



Now, we can expand the determinants $A'_{m,n}$ and $B'_{m,n}$ in terms of a certain row or column as is convenient and get

$$A'_{m,n} = -A_{m,n},$$
$$B'_{m,n} = -[-C_{m,n}p'_1 \cdot p_0 + D_{m,n}p''_1 \cdot p_0 + E_{m,n}q'_1 \cdot p_0 + F_{m,n}q''_1 \cdot p_0 - G_{m,n}mp_0^2], \quad (9)$$

where

$$A_{m,n} = \Bigg[\Bigg\{\sum_{s=3}^{m-1}\beta_s\beta''_{s+1}\beta''_{s+2}\cdots\beta''_m A^{\beta}_{s-1} + \beta_m A^{\beta}_{m-1} + \beta''_m\beta''_{m-1}\cdots\beta''_3(\beta''_2\beta_2 + \beta''_2\beta_1 + \beta_2\beta_1)\Bigg\}A^{\alpha}_{n-1} + \sum_{r=1}^{n-1}\alpha_r\alpha''_{r+1}\alpha''_{r+2}\cdots\alpha''_n A^{\alpha}_{r-1}A^{\beta}_{m-1}\Bigg], \quad (10)$$

$$C_{m,n} = \Bigg[\sum_{s=2}^{m-2}\beta'_s\beta''_2\beta''_3\cdots\beta''_s I'_s + \beta'_1 I'_1 + \beta'_{m-1}\beta''_2\beta''_3\cdots\beta''_{m-1}(\beta''_m A^{\alpha}_{n-1} + \beta_m A^{\alpha}_{n-1} + \alpha''_n J_{n-1,n-1}) - \sum_{s=1}^{m-1}\beta'_s\beta''_2\beta''_3\cdots\beta''_m A^{\alpha}_{n-1}\Bigg],$$
$$D_{m,n} = \beta''_1 I'_1, \quad (11)$$
$$E_{m,n} = \beta''_2\beta''_3\cdots\beta''_m\sum_{r=1}^{n-1}\alpha'_r J_{r,n-1},$$
$$F_{m,n} = \beta''_2\beta''_3\cdots\beta''_m\alpha''_1(A^{\alpha}_{n-1} - J_{1,n-1}),$$
$$G_{m,n} = \beta_0\beta''_2\beta''_3\cdots\beta''_m A^{\alpha}_{n-1}.$$

It is shown in the Appendix that $A_{m,n}$, $C_{m,n}$, $D_{m,n}$, $E_{m,n}$, and $G_{m,n}$ are all positive, when the β's and α's are positive whereas the sign of $F_{m,n}$ is not certain. Since both A^{α}_{n-1} and $J_{1,n-1}$ are positive quantities and one does not know which one is greater than the other for any value of n and m and for positive Feynman parameters.

If we substitute Eqs. (6), (9), and (10) in Eq. (4) we get

$$A_{m,n}(\beta_1 p_1 \cdot p''_1 + \beta''_1 p'_1 \cdot p''_1 + \beta''_2 p'_1 \cdot p''_2 - \beta'_1 p'^2_1) - C_{m,n}\beta_1 p'_1 \cdot p_0 + D_{m,n}\beta_1 p''_1 \cdot p_0 + E_{m,n}\beta_1 q'_1 \cdot p_0 + F_{m,n}\beta_1 q''_1 \cdot p_0 - G_{m,n}\beta_1 m p_0^2 = 0. \quad (12)$$

Case a: When $m, n > 2$ the sign of $F_{m,n}$ is not certain so we cannot say anything about the solution for positive β and α. If we put $q''_1 \cdot p_0 = 0$, then Eq. (12) has solution for positive β and α if

$$q''_1 \cdot p_0 = 0,$$
$$p_1 \cdot p''_1,\ p'_1 \cdot p''_1,\ p''_2 \cdot p'_1,\ p'_1 \cdot p_0,\ q'_1 \cdot p_0 < 0, \quad (13)$$
$$p''_1 \cdot p_0 > 0;$$

i.e.,

$$m'^2_{q_1} = m''^2_{q_1} + m^2_{p_0},\quad m'^2_{p_2} < m''^2_{p_2} + m'^2_{p_1},$$
$$m''^2_{p_1} > m'^2_{p_1} + m^2_{p_0},\quad m'^2_{p_1} < m''^2_{p_1} + m^2_{p_0},$$
$$m^2_c < m''^2_{p_1} + m^2_{p_1},\quad m''^2_{q_1} < m'^2_{q_1} + m^2_{p_0}.$$
$$m^2_{p_0} < m''^2_p + m'^2_{p_1},$$

Here all the inequalities except the first two are satisfied in equal-mass cases.

Case b: When $m > 2$, $n = 2$, say,

$$A^{\alpha}_{n-1} = A^{\alpha}_1 = \alpha_1 + \alpha''_2,$$
$$J_{1,n-1} = J_{1,1} = \alpha_1,$$

so

$$A^{\alpha}_{n-1} - J_{1,n-1} = \alpha''_2;$$

thus $F_{m,n}$ is positive.

Thus the leading Landau curve has solutions for positive β and α if

$$p_1 \cdot p''_1,\ p'_1 \cdot p''_1,\ p'_1 \cdot p''_2,\ p''_1 \cdot p_0,\ q'_1 \cdot p_0,\ q''_1 \cdot p_0 < 0,$$
$$p'_1 \cdot p_0 > 0;$$

i.e.,

$$m''^2_{p_1} > m'^2_{p_1} + m^2_{p_0},\quad m'^2_{p_2} < m''^2_{p_2} + m'^2_{p_1},$$
$$m^2_c < m''^2_{p_1} + m^2_{p_1},\quad m'^2_{q_1} < m''^2_{q_1} + m^2_{p_0},$$
$$m^2_{p_0} < m''^2_p + m'^2_{p_1},\quad m''^2_{q_1} < m'^2_{q_1} + m^2_{p_0}.$$
$$m'^2_{p_1} < m''^2_{p_1} + m^2_{p_0}, \quad (14)$$

The third and fourth inequalities follow from first inequality of Eq. (14). Here all the inequalities except the first are satisfied in equal-mass cases.

Case c: When $n = 2$, $m = 2$.

$$A_{2,2} = (\beta_2\beta''_2\alpha_1 + \beta_2\beta''_2\alpha''_2 + \beta_1\beta''_2\alpha_1 + \beta_1\beta''_2\alpha''_2 + \beta_1\beta_2\alpha_1 + \beta_1\beta_2\alpha''_2 + \beta_1\alpha_1\alpha''_2 + \beta''_2\alpha_1\alpha''_2),$$
$$C_{2,2} = \beta'_1(\beta_2\alpha_1 + \beta_2\alpha''_2 + \alpha_1\alpha''_2),$$
$$D_{2,2} = \beta''_1(\beta_2\alpha_1 + \beta_2\alpha''_2 + \beta''_2\alpha_1 + \beta''_2\alpha''_2 + \alpha_1\alpha''_2), \quad (15)$$
$$E_{2,2} = \beta''_2\alpha_1\alpha'_1,$$
$$F_{2,2} = \beta''_2\alpha''_1\alpha''_2,$$
$$G_{2,2} = \beta_0\beta''_2(\alpha_1 + \alpha''_2).$$

Substituting these values into Eq. (12) and making use of some conservation equations at the vertices we get

$$\begin{aligned}&(\beta_2\beta''_2\alpha_1 + \beta_2\beta''_2\alpha''_2 + \beta_1\beta''_2\alpha_1 + \beta_1\beta''_2\alpha''_2 + \beta_1\beta_2\alpha_1 + \beta_1\beta_2\alpha''_2 + \beta_1\alpha_1\alpha''_2 + \beta''_2\alpha_1\alpha''_2)\\ &\times(\beta_1 p_1 \cdot p''_1 + \beta''_2 p'_1 \cdot p''_2) + (\beta_1\beta'_1 + \beta''_1\beta''_2)\\ &\times(\beta_2\alpha_1 + \beta_2\alpha''_2 + \alpha_1\alpha''_2)p'_1 \cdot p''_1\\ &+ \beta''_2\beta_1\alpha_1\alpha'_1 q'_1 \cdot p_0 + \beta_1\beta''_2\alpha''_1\alpha''_2 q''_1 \cdot p_0\\ &- \beta_1\beta''_1(\beta_2\alpha_1 + \beta_2\alpha''_2 + \beta''_2\alpha_1 + \beta''_2\alpha''_2 + \alpha_1\alpha''_2)m''^2_{p_1}\\ &- \beta'_1\beta''_2(\beta_1\alpha_1 + \beta_1\alpha''_2 + \beta_2\alpha_1 + \beta_2\alpha''_2 + \alpha_1\alpha''_2)m'^2_{p_1}\\ &- \beta_0\beta_1\beta''_2(\alpha_1 + \alpha''_2)m^2_{p_0} = 0. \quad (16)\end{aligned}$$

For positive β and α and for no solution of Eq. (16) the conditions are

$$p_1 \cdot p_1'', \quad p_1' \cdot p_2'', \quad p_1' \cdot p_1'', \quad q_1' \cdot p_0, \quad q_1'' \cdot p_0 < 0;$$

i.e.,

$$\begin{aligned} &m_a^2 < m_{p_1}^2 + m_{p_1}''^2, \quad m_{q_1}''^2 < m_{q_1}'^2 + m_{p_0}^2, \\ &m_c^2 < m_{p_1}'^2 + m_{p_2}''^2, \quad m_{q_1}'^2 < m_{q_1}''^2 + m_{p_0}^2, \\ &m_{p_0}^2 < m_{p_1}'^2 + m_{p_1}''^2. \end{aligned} \tag{17}$$

We see that Eq. (16) and the condition (17) are exactly the same as obtained in Ref. 4, Eqs. (7) and (8). Here all the conditions are satisfied in the equal-mass case.

ACKNOWLEDGMENT

It is a pleasure to thank Professor T. Pradhan for his interest in this work.

APPENDIX

In this appendix we introduce the quantities $A_{m,n}$ $C_{m,n}$, $D_{m,n}$, $E_{m,n}$, $G_{m,n}$, $F_{m,n}$ and prove that all except the last are positive for positive values of β and α and for any value of n and m greater than 1. The sign of $F_{m,n}$ is not definite except for the values $n = 2$, $m = 2$, for which it is proved to be positive.

We first expand $A'_{m,n}$ in terms of the mth row, which gives

$$\begin{aligned} -A'_{m,n} = A_{m,n} = \Bigg[\Bigg\{ &\sum_{s=2}^{m-1} \beta_s' \beta_2'' \beta_3'' \cdots \beta_m'' \\ &- \sum_{s=2}^{m-1} \beta_s' \beta_{s+1}'' \beta_{s+2}'' \cdots \beta_m'' A_{s-1}^{\beta} \\ &- \beta_m''^2 A_{m-2}^{\beta} + (\beta_m + \beta_m'') A_{m-1}^{\beta} \Bigg\} A_{n-1}^{\alpha} \\ &+ \sum_{r=1}^{n-1} \alpha_r \alpha_{r+1}'' \alpha_{r+2}'' \cdots \alpha_n'' A_{r-1}^{\alpha} A_{m-1}^{\beta} \Bigg], \end{aligned} \tag{A1}$$

where

$$A_{s-1}^{\beta} = \begin{vmatrix} \beta_1 + \beta_2'' & \beta_2'' & 0 & \cdots & 0 \\ \beta_2' + \beta_2'' & M_2 & \beta_3'' & & \\ -\beta_3' & \beta_3'' & M_3 & \beta_4'' & \\ \beta_4' & 0 & \beta_4'' & M_4 & \beta_5'' \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ (-1)^{s-2}\beta_{s-2}' & & & \beta_{s-2}'' & M_{s-2} \quad \beta_{s-1}'' \\ (-1)^{s-1}\beta_{s-1}' & 0 & \cdots & 0 & \beta_{s-1}'' \quad M_{s-1} \end{vmatrix}$$

For A_{r-1}^{α} we have a similar determinant only, β, M, s are replaced by α, N, and r, respectively. The determinants A_{s-1}^{β} and A_{r-1}^{α} are of the same form as A_r' of Ref. 1, so they are positive (we need not prove this again) and satisfy the recursion relation

$$A_{s-1}^{\beta} = M_{s-1} A_{s-2}^{\beta} - \beta_{s-1}''^2 A_{s-3}^{\beta} - \beta_{s-1}' \beta_2'' \beta_3'' \cdots \beta_{s-1}'', \tag{A2}$$

a similar relation exists for A_{r-1}^{α}.

In spite of the fact that A_{s-1}^{β}, A_{r-1}^{α} are positive, it is difficult to say at first from (A1) whether $A_{m,n}$ is positive or not, because the expression contains terms of both signs. To make further simplification we use the recursion relations for A_{m-1}^{β} and separate the last two terms from the summations, which are canceled, leaving

$$\begin{aligned} A_{m,n} = \Bigg[\Bigg\{ &\sum_{s=2}^{m-2} \beta_s' \beta_2'' \beta_3'' \cdots \beta_m'' \\ &- \sum_{s=2}^{m-2} \beta_s' \beta_{s+1}'' \beta_{s+2}'' \cdots \beta_m'' A_{s-1}^{\beta} - \beta_m'' \beta_{m-1}''^2 A_{m-3}^{\beta} \\ &+ (\beta_{m-1} + \beta_{m-1}'') \beta_m'' A_{m-2}^{\beta} + \beta_m A_{m-1}^{\beta} \Bigg\} A_{n-1}^{\alpha} \\ &+ \sum_{r=1}^{n-1} \alpha_r \alpha_{r+1}'' \alpha_{r+2}'' \cdots \alpha_n'' A_{r-1}^{\alpha} A_{m-1}^{\beta} \Bigg]. \end{aligned}$$

Using the recursion relations (A2) for A_{m-2}^{β}, $A_{m-3}^{\beta} \cdots$ etc. in succeeding steps and separating the last terms from summations we finally get

$$\begin{aligned} A_{m,n} = \Bigg[&(\beta_1 \beta_2 + \beta_1 \beta_2'' + \beta_2 \beta_2'') \beta_3'' \beta_4'' \cdots \beta_m'' A_{n-1}^{\alpha} \\ &+ \sum_{s=3}^{m-1} \beta_s \beta_{s+1}'' \beta_{s+2}'' \cdots \beta_m'' A_{s-1}^{\beta} A_{n-1}^{\alpha} + \beta_m A_{m-1}^{\beta} A_{n-1}^{\alpha} \\ &+ \sum_{r=1}^{n-1} \alpha_r \alpha_{r+1}'' \alpha_{r+2}'' \cdots \alpha_n'' A_{r-1}^{\alpha} A_{m-1}^{\beta} \Bigg]. \end{aligned} \tag{A3}$$

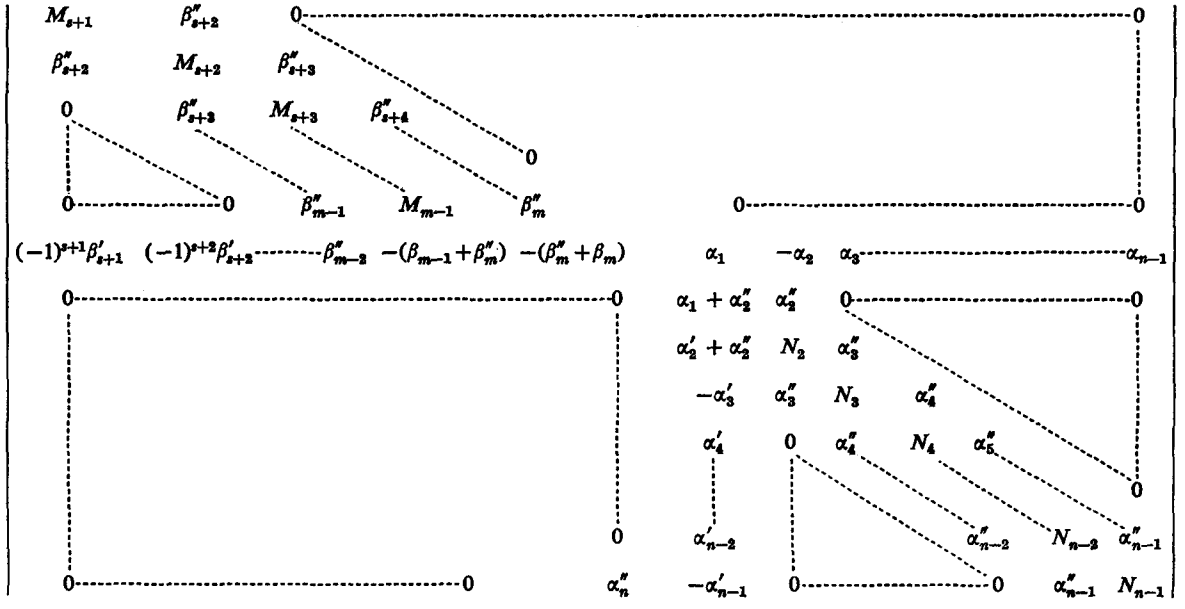
Since A_{s-1}^{β}, A_{r-1}^{α} are positive for positive β and α and any values of s and r, $A_{m,n}$ is positive.

We now expand $B'_{m,n}$ in terms of the last column and get

$$\begin{aligned} B'_{m,n} = \Bigg[\Bigg\{ &\sum_{s=2}^{m-2} \beta_s' \beta_2'' \beta_3'' \cdots \beta'' I_s' + \beta_1' I_1' \\ &- \sum_{s=1}^{m-1} \beta_s' \beta_2'' \beta_3'' \cdots \beta_m'' A_{n-1}^{\alpha} + \beta_{m-1}' \beta_2'' \beta_3'' \cdots \beta_{m-1}'' \\ &\times (\beta_m'' A_{n-1}^{\alpha} + \beta_m A_{n-1}^{\alpha} + \alpha_n'' J_{n-1,n-1}) \Bigg\} p_1' \cdot p_0 \\ &- \beta_1'' I_1' p_1'' \cdot p_0 - \beta_2'' \beta_3'' \cdots \beta_m'' \Bigg\{ \sum_{r=1}^{n-1} \alpha_r J_{r,n-1} q_1' \cdot p_0 \\ &+ \alpha_1'' (A_{n-1}^{\alpha} - J_{1,n-1}) q_1'' \cdot p_0 - \beta_0 m_{p_0}^2 A_{n-1}^{\alpha} \Bigg\} \Bigg], \end{aligned} \tag{A4}$$

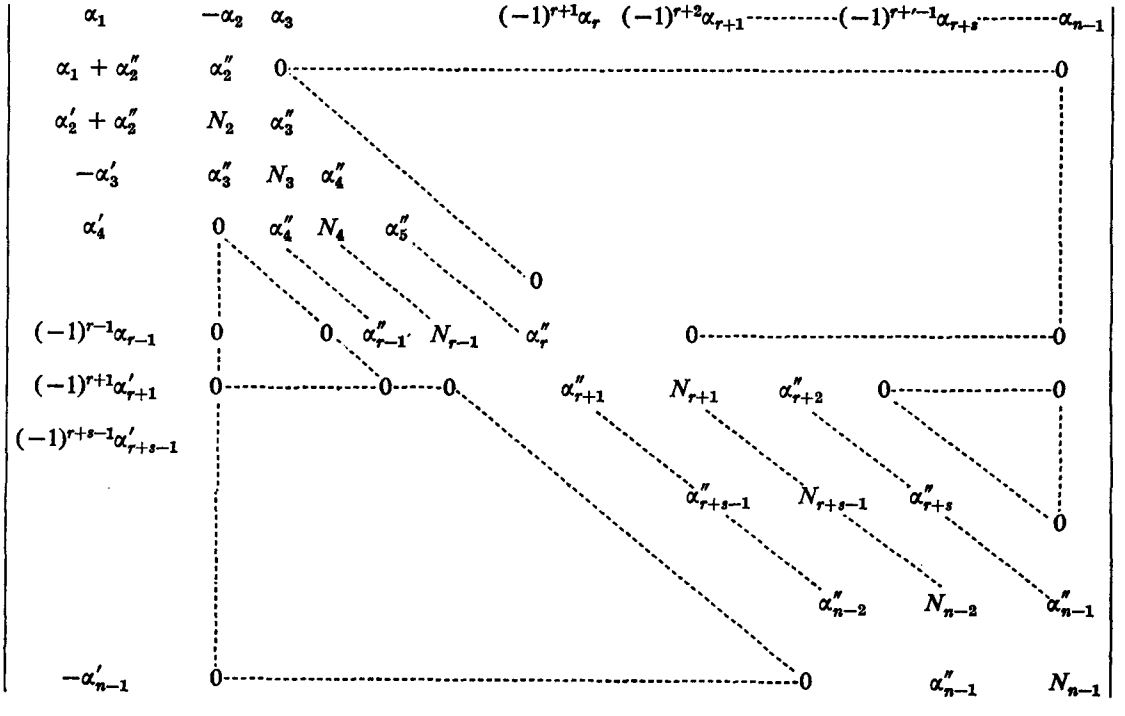
where

$$I'_s = -$$



and

$$J_{r,n-1} =$$



Expanding I'_s in terms of the $(m-s)$th row we get

$$I'_s = \Bigg[A^{\alpha}_{n-1} \Bigg\{ (\beta_m + \beta''_m)\Delta_{s+1,m-1} - \beta''^2_m \Delta_{s+1,m-2} - \sum_{l=1}^{m-s-1} \beta'_{s+l}\beta''_{s+l+1}\beta''_{s+l+2} \cdots \beta''_m \Delta_{s+1,s+l-1} \Bigg\} + \sum_{r=1}^{n-1} \alpha_r \alpha''_{r+1}\alpha''_{r+2} \cdots \alpha''_n A^{\alpha}_{r-1} \Delta_{s+1,m-1} \Bigg], \tag{A5}$$

where

$$\Delta_{s,m} = \begin{vmatrix} M_s & \beta''_{s+1} & 0 & \cdots & 0 \\ \beta''_{s+1} & M_{s+1} & \beta''_{s+2} & & \\ 0 & \beta''_{s+2} & M_{s+2} & & \\ & & & & \beta''_m \\ 0 & \cdots & 0 & \beta''_m & M_m \end{vmatrix}$$

satisfies the recursion relation

$$\Delta_{s,m} = M_m \Delta_{s,m-1} - \beta''^2_s \Delta_{s,m-2}, \tag{A6}$$

which is proved to be positive (Ref. 1). The trick, which we have made earlier for proving $A_{m,n}$ to be positive, is used here to prove that I'_s is positive. That is, in subsequent steps we use relations (A6) for $\Delta_{s+1,m-1}$, $\Delta_{s+1,m-2}$, $\cdots$ etc., and find the negative terms occurring in Eq. (A5) are gradually canceled, giving finally a positive expression for I'_s which is

$$I'_s = \Bigg[\Bigg\{ (\beta_{s+1} + \beta''_{s+1})\beta''_{s+2}\beta''_{s+3} \cdots \beta''_m + \beta_m \Delta_{s+1,m-1} + \sum_{l=1}^{m-s-2} \beta_{s+l+1}\beta''_{s+l+2}\beta''_{s+l+3} \cdots \beta''_m \Delta_{s+1,s+l} \Bigg\} A^{\alpha}_{n-1} + \sum_{r=1}^{n-1} \alpha_r \alpha''_{r+1}\alpha''_{r+2} \cdots \alpha''_n A^{\alpha}_{r-1} \Delta_{s+1,m-1} \Bigg]. \tag{A7}$$

We now go over to $J_{r,n-1}$. (Note that this determinant is different from others occurring in the problem in the respect that none of the elements of first row and first column are zero.) Expanding in terms of the first row we have

$$J_{r,n-1} = \Bigg[\Bigg\{ \sum_{v=1}^{r-1} \alpha_v \alpha''_{v+1}\alpha''_{v+2} \cdots \alpha''_r A^{\alpha}_{v-1} + \alpha_r A^{\alpha}_{r-1} \Bigg\} \Delta_{r+1,n-1} + \sum_{s=1}^{n-1} \alpha_{r+s} J'_{r,r+s,n-1} \Bigg], \tag{A8}$$

where

$J'_{r,r+s,n-1} =$

$$\begin{vmatrix} \alpha_1 + \alpha''_2 & \alpha''_2 & 0 & \cdots & & & & & & & & & 0 \\ \alpha'_2 + \alpha''_2 & N_2 & \alpha''_3 & 0 & \cdots & & & & & & & & 0 \\ -\alpha'_3 & \alpha''_3 & N_3 & \alpha''_4 & & & & & & & & & \\ & 0 & & \alpha''_{r-1} & N_{r-1} & \alpha''_r & 0 & 0 & & & & & \\ (-1)^{r-1}\alpha'_{r-1} & & & & 0 & \alpha''_{r+1} & N_{r+1} & \alpha''_{r+2} & & & & & \\ (-1)^{r+1}\alpha'_{r+1} & & & & & \alpha''_{r+2} & N_{r+2} & \alpha''_{r+3} & & & & & \\ & & & & & & & & \alpha''_{r+s-2} & N_{r+s-2} & \alpha''_{r+s-1} & & \\ (-1)^{r+s-1}\alpha'_{r+s-1} & & & & & & & & & \alpha''_{r+s-1} & N_{r+s-1} & 0 & \\ & & & & & & & & & & \alpha''_{r+s} & \alpha''_{r+s+1} & \\ (-1)^{r+s+2}\alpha'_{r+s+2} & & & & & & & & 0 & & N_{r+s+1} & \alpha''_{r+s+2} & \\ & & & & & & & & & & \alpha''_{r+s+2} & N_{r+s+3} & \alpha''_{r+s+3} \\ \alpha'_{n-2} & & & & & & & & & & & & \alpha''_{n-2} \\ -\alpha'_{n-1} & 0 & \cdots & & & & & & & & & \alpha''_{n-1} & N_{n-1} \end{vmatrix}$$

and it satisfies the recursion relations

$$J'_{r,r+s,n-1} = N_{n-1}J'_{r,r+s,n-2} - \alpha''^2_{n-1}J'_{r,r+s,n-3} + \alpha'_{n-1}\alpha''_2\alpha''_3 \cdots \alpha''_r \Delta_{r+1,r+s-1}, \tag{A9}$$

$$J'_{r,r+s,r+s+1} = N_{r+s+1}J'_{r,r+s,r+s} + \alpha'_{r+s+1}\alpha''_2\alpha''_3 \cdots \alpha''_r \Delta_{r+1,r+s-1}, \tag{A10}$$

$$J'_{r,r+s,r+s} = \alpha''_{r+s}J'_{r,r+s,r+s-1} + \alpha'_{r+s}\alpha''_2\alpha''_3 \cdots \alpha''_r \Delta_{r+1,r+s-2}, \tag{A11}$$

$$J'_{r,r+s,r+2} = \alpha''_{r+2}J'_{r,r+s,r+1} + \alpha'_{r+2}\alpha''_2\alpha''_3 \cdots \alpha''_r N_{r+1}, \tag{A12}$$

$$J'_{r,r+s,r+1} = \alpha''_{r+1}A^{\alpha}_{r-1} + \alpha'_{r+1}\alpha''_2\alpha''_3 \cdots \alpha''_r, \tag{A13}$$

$$J'_{r,r+s,1} = 1. \tag{A14}$$

Equation (A13) is valid for $r > 1$. For $r = 1$;

$$J'_{1,1+s,2} = \alpha'_2 + \alpha''_2.$$

All the J'_s occurring in between $J'_{r,r+s,n-1}$ and $J'_{r,r+s,r+s+1}$ satisfy recursion relation (A9). All J'_s in between $J'_{r,r+s,r+s}$ and $J'_{r,r+s,r+2}$ satisfy Eq. (A11) and all J'_s between $J'_{r,r+s,r+1}$ and $J'_{r,r+s,1}$ satisfy the recursion relation for A^{α}_{r-1}. We now prove by induction that $J'_{r,r+s,n-1}$ is positive. Since A^{α}_{r-1} is positive, $J'_{r,r+s,r+1}$ is positive for positive β and α which implies $J'_{r,r+s,r+s+1}$ is positive, because Δ_s are also positive. Using Eq. (A9) we find that if $J'_{r,r+s,r+s+1}$ is positive then $J'_{r,r+s,s+s+2}$ is positive. Thus by induction we can prove that $J'_{r,r+s,n-1}$ is positive, which implies that $J'_{r,n-1}$ is also positive.

Now, substituting the value of I'_s and I'_1 from Eq. (A7) in the expression for $C_{m,n}$ given by Eq. (11) we get

$$\begin{aligned} C_{m,n} = \sum_{s=2}^{m-2} \beta'_s\beta''_2\beta''_3 \cdots \beta''_s \Bigg\{ \Bigg(& \beta''_{s+2}\beta''_{s+3} \cdots \beta''_m\beta_{s+1} \\ & + \sum_{l=1}^{m-s-2} \beta_{s+l+1}\beta''_{s+l+2}\beta''_{s+l+3} \cdots \beta''_m\Delta_{s+1,s+l} \\ & + \beta_m\Delta_{s+1,m-1}\Bigg) A^{\alpha}_{n-1} \\ & + \sum_{r=1}^{m-1} \alpha_r\alpha''_{r+1}\alpha''_{r+2} \cdots \alpha''_n A^{\alpha}_{r-1}\Delta_{s+1,m-1}\Bigg\} \\ & + \beta'_1\Bigg\{\Bigg(\beta_2\beta''_3\beta''_4 \cdots \beta''_m + \sum_{l=1}^{m-2} \beta_{l+2}\beta''_{l+3}\beta''_{l+4} \cdots \beta''_m\Delta_{2,1+l} \\ & + \beta_m\Delta_{2,m-1}\Bigg) A^{\alpha}_{n-1} + \sum_{r=1}^{m-1} \alpha_r\alpha''_{r+1}\alpha''_{r+2} \cdots \alpha''_n A^{\alpha}_{r-1}\Delta_{2,m-1}\Bigg\}. \end{aligned}$$

This shows that $C_{m,n}$ is positive for positive β_s and α_s. Recalling the expressions from Eq. (11),

$$\begin{aligned} D_{m,n} &= \beta''_1 I'_1, \\ E_{m,n} &= \beta''_2\beta''_3 \cdots \beta''_m \sum_{r=1}^{n-1} \alpha'_r J_{r,n-1}, \\ G_{m,n} &= \beta_0\beta''_2\beta''_3 \cdots \beta''_m A^{\alpha}_{n-1}. \end{aligned}$$

We find that these are all positive, since I'_s and J_s, A^{α}_s are positive. However,

$$F_{m,n} = \beta''_2\beta''_3 \cdots \beta''_m(A^{\alpha}_{n-1} - J_{1,n-1})\alpha''_1$$

does not have a definite sign but for $n = 2$, $m = 2$

$$A^{\alpha}_{n-1} - J_{1,n-1} = \alpha''_2,$$

which is positive.

Weyl Coefficients in $SU(3)$†

KLAUS J. LEZUO
University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

(Received 16 November 1966)

The Weyl coefficients are by definition the matrix elements of the Weyl operators in $SU(3)$. They are found to be generalized hypergeometric series of the type ${}_4F_3$, and can be written down in a simple way from the Gel'fand patterns involved.

INTRODUCTION

THE Weyl coefficients in $SU(3)$ are of interest to the physicist who employs "U-spin" techniques in certain problems concerned with scattering cross sections in elementary particle physics. As such, use is made of the general property of the Weyl operators, namely to diagonalize a set of invariants, which is obtained from a subgroup decomposition[1,2] in terms of any other such set. In other words, a state with a definite "J spin" is transformed into a state with a definite "U spin." In addition, as a symmetry operation in the group, Weyl operations yield relations between the Wigner coefficients of $SU(3)$ for weights which are equivalent under Weyl reflections.

For the low-dimensional representations, which are used in physics, there are several ways to calculate the Weyl coefficients.[3] These methods, however, do not lend themselves to direct generalization. A simple alternative procedure makes explicit use of the boson calculus, which has been introduced by Schwinger[4] for $SU(2)$ and has been generalized to $U(n)$ by Baird and Biedenharn.[2] It requires the knowledge of the states as realized by boson quanta and the definition of the Weyl operators on the system of boson quanta. We discuss this procedure in detail below.

I. THE WEYL GROUP OF $SU(3)$. WEYL OPERATIONS AND WEYL REFLECTIONS

The Lie algebra of $SU(3)$ is defined by

$$[A_{ij}, A_{kl}] = \delta_{jk}A_{il} - \delta_{il}A_{kj}, \tag{1}$$

where the indices range from 1 to 3. The problem of the Weyl reflections is to find the group of inner automorphisms of the Lie algebra.[5] It can be seen that an operation W_p, which permutes the indices of the generators in the manner indicated by P, i.e.,

$$W_pA_{ij}W_p^{-1} \equiv A_{p(ij)}, \tag{2}$$

indeed maps the algebra onto itself. One should, however, distinguish this operation from the trivial one of permuting all indices in (1). W_p leaves the indices of the structure constants unchanged. As defined in (2), W_p is called a *Weyl operation*, with an additional phase to be established below. The group of all Weyl operations is called the Weyl group and from its definition seen to be isomorphic to the symmetric group S_3. It is of order 3!, has three classes, and is generated by two elements, the independent transpositions. The generating Weyl operations are labeled W_{12} and W_{23} and are called *simple* Weyl operations, because they are associated with simple roots.[6] The particular labeling scheme used is a consequence of the canonical subgroup decomposition. In addition, Biedenharn[5] has shown that Weyl operations are but special rotations, which can be written in the form

$$W = R_\alpha(\pi) = \exp\{i\pi(2\boldsymbol{\alpha}\boldsymbol{\alpha})^{-\frac{1}{2}}(E_\alpha + E_{-\alpha})\}, \tag{3}$$

where $\boldsymbol{\alpha}$ is the root which defines W.

Alternatively, a mapping of the Lie algebra onto itself is obtained, by mapping the root system onto itself. This is expressed in the well-known relation

$$\mathbf{m}' = \mathbf{m} - 2\boldsymbol{\alpha}(\mathbf{m}\boldsymbol{\alpha})/(\boldsymbol{\alpha}\boldsymbol{\alpha}) \quad (\boldsymbol{\alpha} \text{ fixed}). \tag{4}$$

In words, it states that if $\mathbf{m}$ is a root (or a weight) and $\boldsymbol{\alpha}$ is a root, then $\mathbf{m}'$ is also a root (or a weight), where the transformation $\mathbf{m} \to \mathbf{m}'$ can be interpreted as a reflection of the root (or weight) diagram through a plane perpendicular to the root $\boldsymbol{\alpha}$. This interpretation stems from Weyl and is hence known as a "Weyl

† Excerpted in part from a Ph.D. thesis, University of North Carolina, Chapel Hill, (1967). Supported by the U.S. Atomic Energy Commission.

[1] L. C. Biedenharn, J. Math. Phys. **4**, 436 (1963).

[2] G. E. Baird and L. C. Biedenharn, J. Math. Phys. **4**, 1449 (1963).

[3] M. E. Mayer, in *Brandeis Summer Institute in Theoretical Physics* (W. A. Benjamin Inc., New York, 1963), Vol. 1.

[4] J. Schwinger, in *Selected Papers in Quantum Theory of Angular Momentum*, L. C. Biedenharn and H. Van Dam, Eds. (Academic Press Inc., New York, 1965).

[5] L. C. Biedenharn, in *Lectures in Theoretical Physics*, W. E. Britten, B. W. Downs, and Joanne Downs, Eds. (Interscience Publishers, Inc., New York, 1963), Vol. V. (This is originally due to Weyl, Princeton Lectures, but the cited reference is more accessible.)

[6] Every positive root can be written as a linear combination of simple roots, with nonnegative integers only.

reflection." (This way the Weyl group appears to have a somewhat different structure, because the product of two reflections is a rotation.) The Weyl group can be described as isomorphic to the dihedral group D_3,[5] with its abstract definition:

$$(AB)^3 = (B^2) = E$$

and is, of course, isomorphic to S_3.

The two ways of considering the Weyl group can be conveniently distinguished by the names "Weyl operation" and "Weyl reflection." One should, in fact, reserve the term "reflection" to the latter interpretation, because it is strictly linked to the root diagram. As an operation in the group, W_p is a *rotation* and the word "reflection" in this connection is misleading.

Since Weyl operations are automorphisms of the Lie algebra it follows that all results derived therefrom must be symmetric under Weyl reflections. The study of these symmetries is then the primary aim to be pursued here. The situation, however, becomes more complicated upon introducing the subgroup decomposition, which is necessary to label the states of the irreducible representations of $SU(3)$ uniquely. This introduces a superimposed structure which is not invariant under Weyl operations. In fact, under Weyl operations a given subgroup decomposition is transformed into a different one, which is equivalent to the former one. The term "canonical" subgroup decomposition—meaning unique up to equivalence classes[1,2]—refers exactly to this kind of equivalence. It is quite clear that the simple change of the subgroup decomposition in the way indicated yields trivial results only. The interest lies in considering the relations between two possible subgroup decompositions which are related by a Weyl operation,[7] to which we now turn. Let the invariants of a $U(2)$ subgroup of $SU(3)$ be denoted by Λ_{12}, i.e.,

$$\Lambda_{12} \equiv E_{12}E_{21} + E_{21}E_{12} + \tfrac{1}{4}(N_1 - N_2)^2,$$

where N_1, N_2 are diagonal generators, namely $H_1 = \frac{1}{2}(N_1 - N_2)$. States which are eigenstates of Λ_{12} (and of the other diagonal operators in the complete set) may be denoted by $|j\rangle$ and have eigenvalues $j(j+1)$. A different subgroup $U(2)$ can be characterized by Λ_{13}, to be defined in an analogous manner, and has eigenstates $|u\rangle$. The invariants of these two subgroups do not commute with each other, of course, and are related by the Weyl operator W_{23}, i.e.,

$$W_{23}\Lambda_{12}W_{23}^{-1} = \Lambda_{13},$$
$$W_{23}\,|j\rangle = |u\rangle.$$

This is a rotation of the basis $|j\rangle$ to a new basis, which can be defined as a linear combination of states $|j\rangle$, in the form

$$W_{23}\,|j\rangle = \sum_{j'} w_{23}(jj')\,|j'\rangle.$$

The coefficients in this expansion, namely the matrix elements of W_{23}, we call *Weyl coefficients*, and are calculated explicitly in the following sections, using the boson calculus. Our basic idea is to define the Weyl operators on a system of bosons for which the states are known. Weyl operations can then be explicitly performed on these realizations of the states, and the expansion in terms of the original states yields the Weyl coefficients.

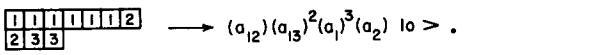
II. BASIS STATES IN THE REPRESENTATIONS OF $SU(3)$ AS REALIZED BY BOSON CREATION OPERATORS

Baird and Biedenharn[2] show how to map the Lie algebra of $SU(3)$ onto suitable boson operators, and proceed to construct all states of the irreducible representations of $SU(3)$. Their result is essentially reproduced in (5a). For the present calculation it is necessary to make a few additional remarks.

The states are labeled by means of a *Gel'fand pattern*.[2,8,9] This is a triangular array of numbers m_{ij} in the form

$$\begin{pmatrix} m_{13} & & m_{23} & & m_{33} \\ & m_{12} & & m_{22} & \\ & & m_{11} & & \end{pmatrix},$$

which reflects the subgroup decomposition (Weyl branching law). The m_{ij} are arbitrary nonnegative integers lying within the limits $m_{ij+1} \geq m_{ij} \geq m_{i+1,j+1}$ ("betweenness relation"). These states span an orthogonal basis, called the Gelfand basis.[2] The realization of these states in terms of the boson calculus yields a polynomial of boson operators. The individual "monomial" terms can be visualized by means of the Weyl pattern.[2] A Weyl pattern is a Young frame,[10] which is filled in a lexicographical way with the numbers 1, 2, 3. A mapping onto the boson calculus is established by "reading off the quanta," best demonstrated by example:



[7] There is a Weyl operation to relate any two possible subgroup decompositions. This in fact also exhausts the Weyl group.

[8] I. M. Gel'fand and M. L. Zetlin, Dokl. Akad. Nauk SSSR 71, 825 (1950).

[9] L. C. Biedenharn, "Group Theory and the Classification of the Elementary Particles," lecture notes, CERN 65-41 (1965).

[10] We reserve Young tableau to be a *Young frame* lexically filled with the integers 1 to n.

To be more specific, such a pattern is referred to as a Weyl basis tableau, and the state is said to belong to the Weyl basis. The Weyl basis states span the space of all the irreducible representations of $SU(3)$. Although this basis is generally not orthogonal, two states with different weights and/or which belong to different irreducible representations are orthogonal. Beyond the boson monomials so obtained, there are still others, such as for example $(a_{13})(a_{23})(a_1)(a_2)\,|0\rangle$, with the simultaneous occurrence of a_1 and a_{23}, which cannot be obtained from a Weyl basis tableau. Accordingly the quanta a_1 and a_{23} are "incompatible" in this sense. One of the two can always be eliminated using the identity $a_1a_{23} - a_2a_{13} + a_3a_{12} \equiv 0$. The Weyl basis is then singled out by supplying a linearly independent set of boson monomial states. This is precisely the property needed below.

Baird and Biedenharn's expression can be transformed to the Weyl basis using the identity just mentioned. In their expression it is also seen that the result (for $SU(3)$) does not directly depend on the integers m_{ij}, but *only on differences thereof.* In addition one notes that their expression has a direct meaning only for $m_{11} \geq m_{23}$. This is an expression of the fact that m_{11} *and* m_{23} *are not constrained by any ordering relation.* It will therefore be necessary to distinguish three cases: $m_{11} > m_{23}$; $m_{11} < m_{23}$; $m_{11} = m_{23}$. This is most conveniently done by defining a *path on the Gelfand pattern* which establishes a total ordering of all the m_{ij} (for each numerical occurrence). Let the line start at m_{33} (the lowest integer in the pattern). It is said to be *regular*, if the m's along the line *never decrease.* The line has of course been introduced to be regular, yet it is sometimes useful to refer to the opposite case, an irregular line.

Beyond this immediate purpose, the regular line is an extremely useful tool, which introduces a significant structure into the Gelfand pattern and makes it possible to write down the expansion of the Gelfand basis states in terms of the Weyl basis states directly from the Gelfand pattern. [This is even more transparent in $SU(n)$.[11]] This is facilitated, using a different notation, obtained as follows: The Gelfand pattern is mapped onto a pattern which looks like a Weyl pattern. (It is, of course, impossible to map the Gelfand pattern onto a single Weyl basis tableau.) This is achieved by interpreting the meaning of m_{ij}, namely as the numbers which indicate how far the indices j reach in the ith row of this pattern,[9] e.g.,

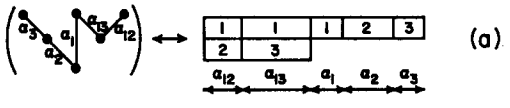
$$\begin{pmatrix} 6 & & 2 & & 0 \\ & 4 & & 2 & \\ & & 3 & & \end{pmatrix} \longleftrightarrow \begin{array}{|c|c|c|c|c|c|c|} \hline 1 & 1 & 1 & 2 & 3 & 3 \\ \hline 2 & 2 \\ \cline{1-2} \end{array}.$$

The latter pattern is completely specified, if the number of columns of a certain type is given. The subscripts of the symbols α serve to distinguish the type of the columns. As an example, the pattern drawn above is specified by

$$\alpha_{12} = 2, \quad \alpha_{13} = 0, \quad \alpha_1 = 1, \quad \alpha_2 = 1, \quad \alpha_3 = 2.$$

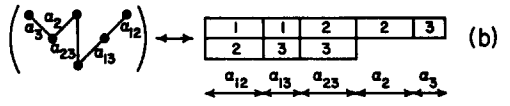
The same symbols will then be inserted into the regular line of the Gelfand pattern, replacing the m_{ij}, and by construction denote the positive differences of adjacent m's along the line.

The two possibilities in $SU(3)$, $m_{11} \geq m_{23}$ and $m_{11} \leq m_{23}$ are indicated as follows:



(a)

and



(b)

These two types of patterns are called *conjugates* of each other, because in going from a representation to the conjugate one, a pattern of type (a) goes into a pattern of type (b) and vice versa (as noted already in Ref. 2). If $m_{11} = m_{23}$, the two cases coincide, both, however, remain regular by definition. It is not necessary to make this a special case since either (a) or (b) yield identical results. This in fact has to be so, because it is but an expression of the required continuity of the orthogonalizing function, with respect to the application of E_{12} or E_{21}. The expansions in terms of the Weyl basis for the cases (a) and (b) are then as follows:

(a)
$$|(\alpha)\rangle = N(a_3)^{\alpha_3}(a_2)^{\alpha_2}(a_1)^{\alpha_1}(a_{13})^{\alpha_{13}}(a_{12})^{\alpha_{12}} \sum_i g_i^{(-)}(\alpha_2, \alpha_1, \alpha_{13}) \left(\frac{a_3a_{12}}{a_2a_{13}}\right)^i |0\rangle, \tag{5a}$$

where

$$N^{-2} = \frac{(\alpha_{12} + \alpha_{13} + \alpha_1 + \alpha_2 + 1)!\,\alpha_{12}!\,(\alpha_{13} + \alpha_1 + \alpha_2 + \alpha_3 + 1)!\,\alpha_{13}!\,\alpha_1!\,\alpha_2!\,\alpha_3!\,(\alpha_2 + \alpha_1)!\,(\alpha_1 + \alpha_{13})!}{(\alpha_{13} + \alpha_1 + \alpha_2 + 1)!(\alpha_1 + \alpha_2 + \alpha_3 + 1)!\alpha_1!\,(\alpha_2 + \alpha_1 + \alpha_{13})!},$$

(b)
$$|(\alpha)\rangle = N(a_3)^{\alpha_3}(a_2)^{\alpha_2}(a_{23})^{\alpha_{23}}(a_{13})^{\alpha_{13}}(a_{12})^{\alpha_{12}} \sum_i g_i^{(-)}(\alpha_2, \alpha_{23}, \alpha_{13}) \left(\frac{a_3a_{12}}{a_2a_{13}}\right)^i |0\rangle, \tag{5b}$$

[11] K. J. Lezuo, Ph.D. thesis, University of North Carolina (1967).

where

$$N^{-2} = \frac{(\alpha_{12}+\alpha_{13}+\alpha_{23}+\alpha_2+1)!\,\alpha_{12}!\,(\alpha_{13}+\alpha_{23}+\alpha_2+\alpha_3+1)!\alpha_{13}!\,(\alpha_{23}+\alpha_2+\alpha_3+1)!\,\alpha_{23}!\,\alpha_2!\,\alpha_3!}{(\alpha_{13}+\alpha_{23}+\alpha_2+1)!(\alpha_{23}+\alpha_2+\alpha_3+1)!(\alpha_2+\alpha_3+1)!}$$
$$\times \frac{(\alpha_2+\alpha_{23})!\,(\alpha_{23}+\alpha_{13})!}{\alpha_{23}!\,(\alpha_2+\alpha_{23}+\alpha_{13})!}.$$

The leading term in these expansions corresponds to a Weyl basis tableau, which looks identical to the tableau onto which the Gel'fand pattern has been mapped. Accordingly the latter one is called the *representative* Weyl tableau. Equation (5a) is, except for a transformation, identical to the result given in Ref. 2. Equation (5b) is obtained by specializing (5a) to $\alpha_1 = 0$ and a subsequent application of $(E_{21})^{\alpha_{23}}$. The normalization of (5b) has been calculated directly in the form $\langle(\alpha)\,|\,(\alpha)\rangle$. One should note that the normalizations in both cases can be obtained directly from the representative Weyl tableau, using Biedenharn's idea of "entanglement." The second factor is closely related to the structure of the series. (In fact it is the sum of all the coefficients in the series. A detailed discussion based on combinatorial analysis is given by Ciftan and Biedenharn.[12]) The standard convention of specifying hypergeometric series has not been used, to adjust the notation more closely to the problem at hand. In particular, the parameters have been inserted in the order in which they are read off the regular line. Note also that the arguments of the series follow from the subscripts of the parameters. The definition of $g_i^{(-)}(\alpha_2, \alpha_1, \alpha_{13})$ is

$$g_i^{(-)}(\alpha_2, \alpha_1, \alpha_{13}) \equiv (-)^i \frac{\alpha_2!\quad \alpha_{13}!\quad (\alpha_2+\alpha_1+\alpha_{13}-i)!}{i!\,(\alpha_2-i)!\,(\alpha_{13}-i)!\quad (\alpha_2+\alpha_1+\alpha_{13})!}$$
$$= (-)^i \frac{\binom{\alpha_2}{i}\binom{\alpha_{13}}{i}}{\binom{\alpha_2+\alpha_1+\alpha_{13}}{i}}. \tag{6}$$

This function may be called the characteristic function of the state $|(\alpha)\rangle$.

III. CALCULATION OF THE WEYL COEFFICIENTS

The knowledge of the states as realized by the boson calculus allows for the calculation of the Weyl coefficients, provided the Weyl operations are defined on the system of boson quanta. To find this expression use is made of the fact that Weyl operations are special rotation operations. A rotation operator R can be defined as[4]

$$Ra_i^l R^{-1} = x_i^j a_j^l, \quad (x_i^j) \text{ unitary}, \tag{7}$$

which is easily verified. The numbers (x_i^j) are the matrix elements of R in the defining representation (three-dimensional). If the matrix (x_i^j) is chosen so that in every row and column there appears a single ± 1 only, then the operation maps the boson system in a permuted fashion onto itself. Forming the adjoint of Eq. (7) in the operator space, one finds that the destruction operators transform with the same matrix (x_i^j). These properties qualify the operation as a Weyl operation, independently of previous arguments. There remains to fix an ambiguous phase. This is done in such a way that the effect of W_{13} on the state of highest weight in a self-conjugate representation coincides with the operation of conjugation.[13] In addition, this choice makes the matrix element $+1$, when the state of highest weight in any irreducible representation is involved.

Since W_{12} is in $SU(2)$ and its matrix elements known, i.e.,

$$\langle j'm'|\, W_{12}\,|jm\rangle = \delta_{jj'}\delta_{-m,m'}(-)^{j-m},$$

the calculation can be restricted to W_{23}, which is defined on the boson system as

$$W_{23}a_1^l W_{23}^{-1} = a_1^l; \quad W_{23}a_2^l W_{23}^{-1} = a_3^l; \quad W_{23}a_3^l W_{23}^{-1} = -a_2^l, \tag{8}$$

using the phase convention just mentioned.

The states which can be reached by W_{23} from a given one, $|(\alpha)\rangle$, are limited and can be expressed as $|(\alpha')\rangle \equiv |(\alpha), n\rangle$, where $\alpha_3' = \alpha_2 - n$, $\alpha_2' = \alpha_3 + n$, $\alpha_1' = \alpha_1$, $\alpha_{13}' = \alpha_{12} + n$, $\alpha_{12}' = \alpha_{13} - n$. n distinguishes the states with the same weight and is found to be

$$n = j' - \tfrac{1}{2}(\alpha_3 + \alpha_1 + \alpha_{12}),$$

where $j' = \frac{1}{2}(m_{12}' - m_{22}')$ is the label of a $U(2)$ multiplet. The requirement of α_3', α_2' etc. to be nonnegative limits n: $-\{\alpha_3, \alpha_{12}\} \leq n \leq \{\alpha_2, \alpha_{13}\}$; $\{x, y\}$ meaning the smaller one of x and y. This relation, by the way, gives the multiplicity M of the weight m, and in more standard notation reads

$$2(M-1) = p - |s+m+\lambda| - |s+m-\lambda| - |s-2m|,$$

where

$$s \equiv \tfrac{1}{3}(p-2q).$$

[12] M. Ciftan and L. C. Biedenharn, Science 154, 418 (1966). Abstract of papers presented at the Autumn Meeting of the National Academy of Science, Durham, N.C.

[13] G. E. Baird and L. C. Biedenharn, J. Math. Phys. 5, 1723 (1964).

In terms of the Gel'fand pattern there is a convenient method to obtain the states which are connected to a given one, $|(m)\rangle$, by W_{23}. Denoting the sums of the m's in a row of the Gel'fand pattern by s_i, i.e., $s_1 \equiv m_{11}$, $s_2 \equiv m_{12} + m_{22}$, $s_3 \equiv m_{13} + m_{23} + m_{33}$ the weight is obtained as[9] $m = s_1 - \frac{1}{2}s_2$; $\lambda = \frac{1}{2}s_2 - \frac{1}{3}s_3$. Equation (4) then yields the sums in the transformed pattern (m'). They are $s_1' = s_1$, $s_2' = s_3 - s_2 + s_1$, $s_3' = s_3$. For example,

$$\begin{pmatrix} 7 & 3 & 0 \\ & 4 & 2 \\ & & 4 \end{pmatrix} \xrightarrow{W_{23}} \begin{pmatrix} 7 & 3 & 0 \\ & x & y \\ & & 4 \end{pmatrix}$$

with $x + y = 10 - 6 + 4 = 8$.

The calculation of the matrix elements of W_{23} can now be carried out in the form

$$W_{23}\,|(\alpha)\rangle = \sum_n w_{23}[(\alpha), n]\,|(\alpha)'\rangle. \tag{9}$$

Using the states of the type (5a), and comparing the coefficients of the same Weyl basis state on the two sides of Eq. (9), a system of equations for the matrix elements $w_{23}[(\alpha), n]$ is obtained. It should be stressed that such a comparison is only possible on states which are linearly independent. This singles out the Weyl basis among other possible boson operator monomials. The system of equations obtained is

$$(-)^k N g_i^{(-)}(\alpha_2, \alpha_1, \alpha_{13}) = \sum_n w_{23}[(\alpha), n] N_n g_{n-i}^{(-)}(\alpha_3 + n, \alpha_1, \alpha_{12} + n), \tag{10}$$

where $k \equiv \alpha_3 + \alpha_{13} = s_3 - s_2$ (if $m_{33} = 0$). To solve the problem the inverse of the matrix $g_{n-i}^{(-)}(\alpha_3 + n, \alpha_1, \alpha_{12} + n)$ has to be found. The result is a function $g_{i-n}^{(+)}(\alpha_3 + n, \alpha_1, \alpha_{12} + n)$ to be defined as

$$g_i^{(+)}(\alpha_2, \alpha_1, \alpha_{13}) \equiv \frac{(\alpha_2 + i)!\,(\alpha_{13} + i)!(\alpha_2 + \alpha_1 + \alpha_{13} + 1)!}{i!\alpha_2!\alpha_{13}!(\alpha_2 + \alpha_1 + \alpha_{13} + i + 1)!}. \tag{11}$$

The proof consists of showing that

$$\sum_i g_{i-n}^{(+)}(\alpha_3 + n, \alpha_1, \alpha_{12} + n) \times g_{n'-i}^{(-)}(\alpha_3 + n', \alpha_1, \alpha_{12} + n') = \delta_{nn'}, \tag{12}$$

which involves the summation of a hypergeometric series, and can be done by Gauss' theorem.

These $g^{(+)}$ functions are, like the $g^{(-)}$ functions, the coefficients of the general term of a hypergeometric series. The $(+)$ and $(-)$ signs indicate that these hypergeometric series depend on positive or negative parameters only. Except for adding $+1$ to certain parameters a $g^{(+)}$ function is obtained from a $g^{(-)}$ function by changing the sign of the parameters, i.e.,

$$g_i^{(+)}(\alpha_2, \alpha_1, \alpha_{13}) = g_i^{(-)}[-(\alpha_2 + 1), -\alpha_1, -(\alpha_{13} + 1)].$$

Solving Eq. (10) by means of (12) the Weyl coefficients $w_{23}[(\alpha), n]$ are obtained as

$$w_{23}[(\alpha), n] = (-)^k \frac{N}{N_n} \sum_i g_{i-n}^{(+)}(\alpha_3 + n, \alpha_1, \alpha_{12} + n) g_i^{(-)}(\alpha_2, \alpha_1, \alpha_{13}), \tag{13}$$

where

$$k \equiv \alpha_3 + \alpha_{13}.$$

Although not made explicit in the notation, the series in (13) is a generalized hypergeometric series of type ${}_4F_3$, and is Saalschuetzian, i.e., the sum of the denominator parameters exceeds the sum of the numerator parameters by unity.[14]

In deriving Eq. (13) only states of the type (5a) have been used. It is clear that states of the type (5b) will give essentially the same result, which can be written down by inspection. This, however, is not necessary, because the expression (13) is well defined for all values of the parameters, i.e., even if the line is irregular. As shown below this is the consequence of the invariance of the Weyl coefficients under conjugation, an operation which takes states of type (5a) into states of type (5b). Independent of this argument it can be seen from the derivation. The states of type (5a) are really well defined, even if the line is irregular, were it not for the occurrence of inverse powers of a_1. In the method used, however, the boson quanta play the role of a counting device and inverse powers of a_1 serve the purpose just as well.

The method used for the calculation needs some justification too. At first glance it seems much more straightforward to calculate the scalar product $\langle(\beta)|\,W_{23}\,|(\alpha)\rangle$ directly, particularly because this involves only the two states $|(\alpha)\rangle$ and $|(\beta)\rangle$, whereas the method of expansion requires a complete set of states. The formal difficulty is that of finding the closed forms of various series which occur in the scalar product calculation, and can be traced essentially back to the normalization "integral". In addition one has to perform a rather unnatural operation, namely to expand the antisymmetric combinations, such as a_{12}, in terms of its elementary constituents, however, it is easily demonstrated that a_{12}, etc., can be considered entities by themselves, omitting the reference to the substructure. This is then seen to be the real advantage of using the expansion method.

[14] It seems worth remarking that a series of the same type occurs in the Racah coefficients of $SU(2)$. There are, in fact, certain similarities in the two problems, which may be sufficient to give rise to similar series. At a closer examination, however, the problems seem to have little else in common. Attempting to identify the result (13) with a Racah coefficient, it is already found that both the structure of the irrationals and the "substructure" of the series are quite distinct.

In the following and for explicit evaluation it may be useful to introduce the common labels for the parameters, namely p, q, λ, m, j. The expression of the α's in terms of these is as follows:

$$\begin{aligned}
\alpha_3 &= \tfrac{1}{3}(2p-q) - j - \lambda; \quad \alpha_2 = j - m; \\
\alpha_1 &= \tfrac{1}{3}(p-2q) + m + \lambda, \\
\alpha_{13} &= -\tfrac{1}{3}(p-2q) + j - \lambda; \\
\alpha_{12} &= \tfrac{1}{3}(p+q) - j + \lambda, \\
n &= -\tfrac{1}{3}(2p-q) + j + j' - \tfrac{1}{2}(\lambda + m).
\end{aligned} \tag{14}$$

The explicit substitution of these into Eq. (13) is rather elaborate and spoils the simplicity of the result. It may be most convenient to insert values for p, q, λ, m, j into (14) and the resulting α's into (13) if numerical values are desired. In terms of the labels p, q, λ, m, j the Weyl coefficients are denoted by $w_{23}(pq\lambda m, jj')$ and the selection rules are

$$\begin{aligned}
&\langle p'q'\lambda'm'j' |\, W_{23} \,| pq\lambda mj\rangle \\
&\quad = \delta(p', p)\delta(q', q)\delta(\lambda', \tfrac{1}{2}m - \tfrac{1}{2}\lambda) \\
&\qquad \times \delta(m', \tfrac{1}{2}m + \tfrac{3}{2}\lambda) w_{23}(pq\lambda m, jj').
\end{aligned}$$

IV. INTERPRETATION

It should be observed now that the actual calculation which has been carried out is more general than is suggested by Eq. (12). Without adding anything to the content of the statement, but the known orthogonality of the Weyl basis with respect to different weights and/or different representation, Eq. (12) can be rewritten in the form

$$\begin{aligned}
&\sum_{ij} g_j^{(+)}(\beta_2, \beta_1, \beta_{13}) g_i^{(-)}(\alpha_2, \alpha_1, \alpha_{13}) \\
&\qquad \times \delta(\alpha_2 - i, \beta_2 - j)\delta(\alpha_3 + i)\beta_3 + j)\delta(\alpha_1, \beta_1) \\
&\qquad \times \delta(\alpha_{13} - i, \beta_{13} - j)\delta(\alpha_{12} + i, \beta_{12} + j) \\
&\quad = \delta(\alpha_2\beta_2)\delta(\alpha_3, \beta_3)\delta(\alpha_1, \beta_1)\delta(\alpha_{13}, \beta_{13})\delta(\alpha_{12}, \beta_{12}),
\end{aligned} \tag{15}$$

where (α) and (β) specify the Gel'fand patterns of two states. The Kronecker deltas introduced in the sum specify the Weyl basis state uniquely, or in other words, make sure that the coefficients of the same Weyl basis state have been compared in the expansion problem. This is implicit in (10). The result is then sufficient to "invert a complete set of states" as may be necessary for an arbitrary operator $\mathcal{O}$. In particular, it reduces the problem of finding the matrix elements of $\mathcal{O}$ in the Gel'fand basis to that of finding the matrix elements in the Weyl basis. The latter may be easy for operators which are defined on the boson quanta in a form

$$\mathcal{O} a_i^j \mathcal{O}^{-1} = f(a_i^j).$$

Calling $g_i^{(-)}(\alpha_2, \alpha_1, \alpha_{13})$ the characteristic function of (α), Eq. (15) can be interpreted as supplying the prescription on how to "contract" characteristic functions. One would then refer to $g_i^{(+)}(\alpha_2, \alpha_1, \alpha_{13})$ as the dual function. In this more general language, the Weyl coefficients (13) can be written as follows:

$$\begin{aligned}
&\langle(\alpha)|\, W_{23} \,|(\beta)\rangle \\
&\quad = (-)^k N_\alpha / N_\beta \sum_{ij} g_j^{(+)}(\beta_2, \beta_1, \beta_{13}) g_i^{(-)}(\alpha_2, \alpha_1, \alpha_{13}) \\
&\qquad \times \delta(\alpha_3 + i, \beta_3 - j)\delta(\alpha_2 - i, \beta_2 + j)\delta(\alpha_1, \beta_1) \\
&\qquad \times \delta(\alpha_{13} - i, \beta_{13} + j)\delta(\alpha_{12} + i, \beta_{12} - j)
\end{aligned} \tag{16}$$

and contains all selection rules appropriately. Thus the Weyl coefficients appear as the simplest nontrivial structure involving the characteristic functions of two states.

From (12) or (15) it is seen that the functions $g_i^{(+)}$ are but the coefficients in the expansion of the Weyl basis in terms of the Gel'fand basis. The calculation of the Weyl coefficients is thus solved, by finding the expansion of the Gel'fand basis in terms of the Weyl basis and vice versa, problems which are independent of the structure of any operator.

V. ORTHOGONALITY AND SYMMETRY RELATIONS OF THE WEYL COEFFICIENTS

We have seen that Weyl operators transform the states of a certain subgroup decomposition into the states of a different decomposition. Since both bases are orthogonal, the matrix elements of the Weyl operators belong to an orthogonal matrix. For such operators the inverse equals the transpose, i.e., $W^{-1} = W^{\sim}$. The matrix elements of $W^{\sim}$ are denoted by $w_{23}^{\sim}(pq\lambda m, jj')$, and it should be noted that

$$w_{23}^{\sim}(pq\lambda m, jj') \neq w_{23}(pq\lambda m, j'j).$$

With these remarks in mind the orthogonality relations can be obtained from

$$W_{23} \,|pq\lambda m, j_1\rangle = \sum_{j_1'} w_{23}(pq\lambda m, j_1 j_1') \,|pq\lambda' m', j_1'\rangle$$

and

$$W_{23} \,|pq\lambda m, j_2\rangle = \sum_{j_2'} w_{23}(pq\lambda m, j_2 j_2') \,|pq\lambda' m', j_2'\rangle.$$

The results are

$$\begin{aligned}
&\sum_j w_{23}^{\sim}(pq\lambda m, jj_1) w_{23}(pq\lambda m, jj_2) = \delta_{j_1 j_2} \\
&\text{and} \\
&\sum_j w_{23}^{\sim}(pq\lambda m, j_1 j) w_{23}(pq\lambda m, j_2 j) = \delta_{j_1 j_2}.
\end{aligned} \tag{16}$$

An important symmetry to note is the invariance of the Weyl coefficients (up to a phase) under the change of the weight, namely

$$w_{23}(pq\lambda m, jj') = (-)^\rho w_{23}^{\sim}(pq\lambda' m', j'j), \tag{17}$$

where $\rho \equiv \frac{2}{3}(p+q) + \lambda + m = \alpha_3 + \alpha_{13} + \alpha_2 + \alpha_{12}$. This follows from $W_{23}^2 \,|pq\lambda mj\rangle = (-)^\rho \,|pq\lambda mj\rangle$, which

is obtained by applying W_{23}^2 on the states as realized by the boson calculus.

With the help of Eq. (17) it is easy to see that the matrices of the Weyl coefficients are either symmetric or antisymmetric, i.e.,

$$\tilde{w_{23}}(pq\lambda m, jj') = (-)^{\rho} w_{23}(pq\lambda m, jj') \tag{18}$$

resulting from $WW^{-1}|pq\lambda m,j\rangle = |pq\lambda m,j\rangle$.

It is also expected that the Weyl coefficients be invariant under conjugation, except for a phase. Using Baird and Biedenharn's[13] results for the phase introduced under conjugation, one obtains

$$w_{23}(pq\lambda m, jj') = (-)^{k} w_{23}(p, p-q, -\lambda, -m, j, j'), \tag{19}$$

where

$$k \equiv m_{11} - (m_{13} + m_{23}) = \tfrac{2}{3}(p+q) + \lambda + m = \rho.$$

This can be checked independently from the known result (13). (Note that these invariance requirements yield transformations of Saalschuetzian ${}_4F_3$'s which are quite difficult to derive otherwise.) The transposed matrices can be obtained directly, if j and j' are referred to their minimal values, i.e., $j = m + i$, $j' = m' + i' = \frac{1}{2}m + \frac{3}{2}\lambda + i'$, namely

$$\begin{aligned} \tilde{w_{23}}(pq\lambda m, jj') &\equiv \tilde{w_{23}}(pq\lambda m,\ m+i,\ \tfrac{1}{2}m + \tfrac{3}{2}\lambda + i') \\ &= w_{23}(pq\lambda m,\ m+i',\ \tfrac{1}{2}m + \tfrac{3}{2}\lambda + i). \end{aligned}$$

Finally it is observed that the Weyl coefficients serve a somewhat different purpose too. They are also the coefficients of the expansion of states of a certain subgroup decomposition in terms of the states of a different decomposition *at the same weight*. As such they are *not* the matrix elements of the Weyl operators. Denote eigenstates of the two subgroup decompositions by $|j\rangle$ and $|u\rangle$, ("J spin" and "U spin"). The coefficients to be found are defined in

$$|pq\lambda mj\rangle = \sum_u c(pq\lambda m, ju)\,|pq\lambda m, u\rangle.$$

Using W_{23} on this state yields

$$c(pq\lambda m, ju) = w_{23}(pq\lambda m, ju). \tag{20}$$

For completeness the matrix elements of W_{13} are given. They follow from $W_{13} = W_{12}W_{23}W_{12}$ as

$$w_{13}(pq\lambda m, jj') = (-)^{j+j'-m+m'} w_{23}(pq\lambda - m, jj'). \tag{21}$$

VI. SYMMETRIES OF THE WIGNER COEFFICIENTS OF $SU(3)$ WHICH RESULT FROM WEYL OPERATIONS

From the knowledge of the Wigner coefficients of $SU(3)$ which depend on certain weights, it should be possible to calculate all those Wigner coefficients which depend on weights that are equivalent to the former one under the Weyl group. In $SU(2)$ this is expressed in the relation

$$c(j_1 j_2 j, m_1 m_2 m) = (-)^{j_1+j_2-j} \times c(j_1 j_2 j, -m_1, -m_2, -m).$$

The Wigner coefficients of $SU(3)$ are defined in

$$\begin{aligned} &|p_1q_1p_2q_2,\ p_\gamma q_\gamma\ ;\ \lambda m\rangle \\ &\quad = \sum \begin{pmatrix} p_1q_1 & p_2q_2 & | & p_\gamma q_\gamma \\ \lambda_1 m_1 j_1 & \lambda_2 m_2 j_2 & | & \lambda mj \end{pmatrix} \\ &\qquad \times |\ p_1q_1\lambda_1 m_1 j_1\rangle\ |p_2q_2\lambda_2 m_2 j_2\rangle, \end{aligned}$$

where γ serve to distinguish representations which may occur in the reduction in a multiple fashion. Application of W_{23} on this equation yields

$$\begin{aligned} &\sum w_{23}(p_\gamma q_\gamma \lambda m, jj')\,|\cdots, \lambda' m' j'\rangle \\ &\quad = \sum \begin{pmatrix} p_1q_1 & p_2q_2 & | & p_\gamma q_\gamma \\ \lambda_1 m_1 j_1 & \lambda_2 m_2 j_2 & | & \lambda mj \end{pmatrix} \\ &\qquad \times w_{23}(p_1q_1\lambda_1 m_1, j_1 j_1')\, w_{23}(p_2q_2\lambda_2 m_2, j_2 j_2') \\ &\qquad \times |p_1q_1\lambda_1' m_1' j_1'\rangle\ |p_2q_2\lambda_2' m_2' j_2'\rangle \end{aligned}$$

and can be re-expanded at the transformed weight, i.e.,

$$\begin{aligned} &\sum w_{23}(p_\gamma q_\gamma \lambda m, jj')\,|\cdots, \lambda' m' j'\rangle \\ &\quad = \sum w_{23}(p_\gamma q_\gamma \lambda m, jj') \begin{pmatrix} p_1q_1 & p_2q_2 & | & p_\gamma q_\gamma \\ \lambda_1' m_1' j_1' & \lambda_2' m_2' j_2' & | & \lambda mj \end{pmatrix} \\ &\qquad \times |p_1q_1\lambda_1' m_1' j'\rangle\ |p_2q_2\lambda_2' m_2' j_2'\rangle. \end{aligned}$$

The comparison shows that

$$\begin{aligned} &\begin{pmatrix} p_1q_1 & p_2q_2 & | & p_\gamma q_\gamma \\ \lambda_1' m_1' j_1' & \lambda_2' m_2' j_2' & | & \lambda' m' j' \end{pmatrix} \\ &\quad = \sum_{j_1 j_2 j} \tilde{w_{23}}(p_\gamma q_\gamma \lambda m, jj')\, w_{23}(p_1q_1\lambda_1 m_1, j_1 j_1') \\ &\qquad \times w_{23}(p_2q_2\lambda_2 m_2, j_2 j_2') \begin{pmatrix} p_1q_1 & p_2q_2 & | & p_\gamma q_\gamma \\ \lambda_1 m_1 j_1 & \lambda_2 m_2 j_2 & | & \lambda mj \end{pmatrix}, \end{aligned} \tag{21}$$

where all λ', m' are related to λ, m by $m' = \frac{1}{2}m + \frac{3}{2}\lambda$; $\lambda' = \frac{1}{2}m - \frac{1}{2}\lambda$. This is the desired relation, which is somewhat more complicated than the analogous $SU(2)$ relation, because in $SU(2)$, the Weyl operator W_{12} does not change the subgroup decomposition (in fact there is none at all).

One must remark that there is no simpler relation involving the isoscalar factors only. The reason is that the result of the Weyl operation on a state $|p_1q_1, p_2q_2, \lambda_1\lambda_2\lambda, m, j_1j_2j\rangle$ is not directly a Weyl coefficient, but also involves Wigner coefficients of $SU(2)$. The answer obtained would then be the same as the factorization of (21).

ACKNOWLEDGMENTS

The author would like to thank Professor E. Merzbacher for support and Professor L. C. Biedenharn for guidance of the work. He also thanks M. Ciftan and G. Basbas for help in preparing this article.

APPENDIX

As a numerical example we give the matrix elements of the Weyl operator W_{23} in the 27-plet of $SU(3)$. Let the states be denoted by $|m\lambda j\rangle$, omitting the representation labels, namely, $p = 4$, $q = 2$. Except for a few additional phases, there are only five nontrivial numbers to be calculated due to the various symmetry properties of these coefficients. The cases of multiplicity 2 are specified by $\langle 011|\, W_{23}\,|\frac{1}{2}\,\frac{1}{2}\,\frac{1}{2}\rangle = -6^{\frac{1}{2}}/6$, and $\langle \frac{1}{2}, \frac{1}{2}, \frac{1}{2}|\, W_{23}\,|-\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\rangle = -\frac{2}{3}$, those of multiplicity 3 by

$$\langle 000|\, W_{23}\,|000\rangle = \tfrac{1}{3}; \quad \langle 001|\, W_{23}\,|000\rangle = +(3)^{\frac{1}{2}}/3,$$

and $\langle 001|\, W_{23}\,|001\rangle = \frac{1}{2}$. More explicitly this is

$$W_{23}\begin{pmatrix} \left|-\frac{1}{2}\ \frac{1}{2}\ \frac{1}{2}\right\rangle \\ \left|-\frac{1}{2}\ \frac{1}{2}\ \frac{3}{2}\right\rangle \end{pmatrix} = \begin{pmatrix} -\frac{2}{3} & -\frac{5^{\frac{1}{2}}}{3} \\ -\frac{5^{\frac{1}{2}}}{3} & \frac{2}{3} \end{pmatrix} \times \begin{pmatrix} \left|\frac{1}{2}\ -\frac{1}{2}\ \frac{1}{2}\right\rangle \\ \left|\frac{1}{2}\ -\frac{1}{2}\ \frac{3}{2}\right\rangle \end{pmatrix}$$

and

$$W_{23}\begin{pmatrix} \left|\frac{1}{2}\ \frac{1}{2}\ \frac{1}{2}\right\rangle \\ \left|\frac{1}{2}\ \frac{1}{2}\ \frac{3}{2}\right\rangle \end{pmatrix} = \begin{pmatrix} -\frac{6^{\frac{1}{2}}}{6} & -\frac{(30)^{\frac{1}{2}}}{6} \\ -\frac{(30)^{\frac{1}{2}}}{6} & \frac{6^{\frac{1}{2}}}{6} \end{pmatrix} \begin{pmatrix} |011\rangle \\ |012\rangle \end{pmatrix}.$$

The other multiplicity 2 cases follow from these using the phases in Sec. IV. For the multiplicity 3 case, the result is

$$W_{23}\begin{pmatrix} |000\rangle \\ |001\rangle \\ |002\rangle \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & \frac{3^{\frac{1}{2}}}{3} & \frac{(5)^{\frac{1}{2}}}{3} \\ \frac{3^{\frac{1}{2}}}{3} & \frac{1}{2} & -\frac{(15)^{\frac{1}{2}}}{6} \\ \frac{5^{\frac{1}{2}}}{3} & -\frac{(15)^{\frac{1}{2}}}{6} & \frac{1}{6} \end{pmatrix} \begin{pmatrix} |000\rangle \\ |001\rangle \\ |002\rangle \end{pmatrix}.$$

Since Weyl operations are but special rotations, the matrices obtained have the properties of rotation matrices. The particular matrices above, however, are submatrices only. This is to stress that as such these submatrices need not be of unit determinant. (The determinants may be ± 1.)

Dynamics of Pressure-Free Matter in General Relativity

G. F. R ELLIS*
Relativity Center, University of Texas, Austin, Texas

(Received 13 August 1966)

An orthonormal tetrad system and associated coordinate system is obtained, which may be used to locally describe any dust-filled space–time. This is used to study dust-filled spaces in which there exist multiply transitive groups of motions; all such spaces are classified in detail. Spaces containing shear-free dust are also considered; it is shown that $\sigma = 0 \Rightarrow \omega\Theta = 0$. Three classes of solution with $\sigma = 0$, $\omega \neq 0$ are studied. Several new solutions of the field equations are contained in these results.

1. INTRODUCTION AND NOTATION

IN the study of gravitationally interacting matter, particularly in cosmological models, the simplest assumption about the sources of the gravitational field in Einstein's equations

$$(R_{bd} - \tfrac{1}{2}Rg_{bd}) - \Lambda g_{bd} = -T_{bd} \qquad (1.1)$$

is that the energy-momentum tensor of matter has the form

$$T_{bd} = \rho u_b u_d, \qquad (1.2)$$

where ρ is the density of matter and u^a is the 4-velocity of matter. This vector is taken to be normalized, so

$$u^a u_a = -1. \qquad (1.3)$$

If (1.2) holds with $\rho > 0$, the matter content of the universe is called "dust."

Early study[1,2] of Eqs. (1.1)–(1.3) centered on the conservation equations

$$T^{ab}_{\ ;b} = 0, \qquad (1.4)$$

which are a consequence of (1.1); later work[3–5] (and references given there) has considered general properties of the full field equations (1.1). Many exact solutions of the field equations for dust are known,[6–16] in addition to the classic cosmological models (see, e.g., Refs. 17, 18). An elegant summary of known exact results is contained in the paper by Ehlers.[5]

Recently, the use of tetrad ("vierbein") formalism has resulted in considerable progress in the study of the vacuum field equations.[19–21] Similar methods have been applied to the field equations for dust when certain symmetries exist.[12,13]

In this paper an orthonormal tetrad system and related coordinate system are set up that can be used to describe any space–time containing dust. This system is then used to study some exact properties of dust. This study provides a convenient classification of some of the exact solutions of the field equations (listed in Refs. 6–16) and of some new solutions.

Section 2 establishes notation to be used in the description of a timelike congruence of lines. Some known results concerning dust are summarized, and a self-contained summary is given of tetrad methods basic to the rest of the paper. In Sec. 3, a general space–time filled with dust is considered. A tetrad and coordinate system is considered which simplifies the Jacobi identities and field equations; the remaining coordinate and tetrad freedom is found. The field equations in this tetrad system are used, in Sec. 4, to study dust-filled space–times in which there exists a rotational symmetry in each tangent space. These are in fact, the dust-filled space–times in which there exists a multiply transitive group of motions. All such

* On leave from Peterhouse, Cambridge, England.

[1] L. P. Eisenhart, Trans. Am. Math. Soc. **26**, 205, (1924).

[2] J. L. Synge, Proc. Math. Soc. (London) **43**, 376 (1937).

[3] K. Gödel, Proc. Intern. Congr. Math. (Cambridge, Mass.) **1**, 175, Am. Math. Soc., Providence, R. I. (1952).

[4] A. Raychaudhuri, Phys. Rev. **98**, 1123 (1955).

[5] J. Ehlers, Abhandl. Math. Naturw. Kl., Mainz Akad. Wiss. Lit., **11**, (1961).

[6] J. Ehlers, Ph.D dissertation, Hamburg University (1957); in Royaumont Conference Volume (1959); in *Recent Developments in General Relativity* (Pergamon Press, Inc., New York, 1962).

[7] C. Lanczos, Z. Physik **21**, 73 (1924); W. J. van Stockum, Proc. Roy. Soc. (Edinburgh) **57**, 135 (1937); J. P. Wright, J. Math. Phys. **6**, 103 (1965).

[8] H. Bondi, Monthly Notices Roy. Astron. Soc. **107**, 410 (1947).

[9] K. Gödel, Rev. Mod. Phys. **21**, 447 (1949).

[10] E. Schücking, Naturwiss. **19**, 507 (1957).

[11] O. Heckmann and E. Schücking, in Solvay Conference Volume, (1958); A. Raychaudhuri, Proc. Phys. Soc. (London) **72**, 263 (1958); B. B. Robinson, Proc. Natl. Acad. Sci. U.S. **47**, 1852 (1961).

[12] O. Heckmann and E. Schücking, in *Gravitation: An Introduction to Current Research*, L. Witten, Ed., (John Wiley & Sons, Inc., New York, 1962).

[13] I. Oszvath and E. Schücking, Nature **193**, 1168 (1962); I. Oszvath, Abhandl. Math. Naturw. Kl., Mainz Akad. Wiss. Lit., **13**, (1962) and **1**, (1965); J. Math. Phys. **6**, 590 (1965); D. L. Farnsworth and R. P. Kerr, *ibid.* **7**, 1625 (1966).

[14] C. Behr, Zs. f. Ap. **54**, 268 (1962); Astron. Abhandl. Hamburg Sternwarte, **7**, 5 (1965); L. Shepley, Ph.D. thesis, Princeton University (1965).

[15] S. C. Maitra, J. Math. Phys. **7**, 1025 (1966).

[16] R. Kantowski and R. K. Sachs, J. Math. Phys. **7**, 443 (1966).

[17] H. P. Robertson, Rev. Mod. Phys. **5**, 62 (1933).

[18] H. Bondi, *Cosmology* (Cambridge University Press, Cambridge, England, 1960).

[19] E. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962) and following papers by E. Newman *et al.*, e.g., *ibid.* **6**, 902 (1965).

[20] R. Debever, Lecture notes, Brussels University (1964).

[21] J. Ehlers, Lecture at London Relativity Conference (1965).

solutions are classified in detail, and integrated completely, except for one class in which there remains to be solved two ordinary differential equations. In Sec. 5, properties of shear-free dust are examined. It is shown that either vorticity or expansion vanishes in such solutions; various theorems concerning shear-free solutions are proved. Section 6 contains some concluding remarks.

Notations used are as follows: the metric tensor g_{ab} has signature $(-+++)$. Covariant differentiation in the X^i direction is $\nabla_X = {}_{;k}X^k$; partial differentiation in the X^i direction is ${}_{,k}X^k$. Covariant differentiation along the velocity vector u^a is $\nabla_u = \cdot$.

A vector is regarded as a directional derivative[22,23]. A basis of vectors is $\partial/\partial x^i$, so the vector X can be written $X = X^i\,\partial/\partial x^i$, where X^i are the components of the vector with respect to the basis $\partial/\partial x^i$. Thus, $X(f) = X^i(\partial f/\partial x^i) = f_{,i}X^i$. The commutator of the vectors X, Y is $[X, Y]$, defined by $[X, Y]f := X(Yf) - Y(Xf)$. Then $[X, Y]f = (Y^j{}_{,i}X^i - X^j{}_{,i}Y^i) \times (\partial f/\partial x^j) = (L_X Y)f$, where $L_X Y$ is the Lie derivative of Y with respect to X; this is the vector field giving the difference between the vector field Y, and the vector field produced if Y is "dragged along" by X.[22–25] This gives the commutator a useful geometric interpretation, closely related to integrability conditions.

A set of vectors $\{e_a\}$ (labeled by the index a) that are orthonormal at each point is called a *tetrad*. The notation ∂_a is used to emphasize the action of these vectors as directional derivatives: $\partial_a f := e_a(f)$.

Latin indices run from 0 to 3; $i, j, k \cdots$ are coordinate indices and $a, b, c \cdots$ are tetrad indices. Greek indices run over 1, 2, 3; the vector e_0 is timelike, so $\{e_\nu\}$ are spacelike. Round brackets denote symmetrized indices, and square brackets denote skew-symmetrized indices.[24] The skew-symmetric tensor η is defined by

$$\eta^{abcd} := 4!e_0{}^{[a}e_1{}^b e_2{}^c e_3{}^{d]};$$

so $\eta^{abcd} = \eta^{[abcd]}$, has tetrad components $\eta^{0123} = 1$, $\eta_{0123} = -1$. The Ricci tensor is defined by $R_{ab} := R^c_{acb}$, the curvature scalar by $R := R^a_a$. Finally, the metric scalar product is denoted by a dot, $X.Y := X^i Y_i = g_{ij}X^i Y^j$. Note from this that

$$(\partial/\partial x^k).(\partial/\partial x^m) = g_{ij}\delta^i_k\delta^j_m = g_{km}.$$

2. KINEMATICS OF A TIMELIKE CONGRUENCE

We first summarize some results given by Ehlers.[5] For any space–time filled with dust, the fluid flow vector u^a determines the tensor

$$h_{ab} := g_{ab} + u_a u_b, \tag{2.1}$$

which projects into the instantaneous rest space of an observer moving with 4-velocity u^a. This tensor obeys

$$h^b_a h^c_b = h^c_a;\ h^a_b u^b = 0;\ h^a_a = 3.$$

The tensors Θ_{ab}, ω_{ab}, σ_{ab}, $\dot{u}_a$, and Θ are defined by

$$\dot{u}_a := u_{a;b}u^b, \tag{2.2a}$$

$$u_{a;b} =: \omega_{ab} + \Theta_{ab} - \dot{u}_a u_b, \tag{2.2b}$$

where

$$\omega_{ab} = \omega_{[ab]},\ \omega_{ab}u^b = 0,\quad \Theta_{ab} = \Theta_{(ab)},\quad \Theta_{ab}u^b = 0;$$
$$\sigma_{ab} := \Theta_{ab} - \tfrac{1}{3}\Theta h_{ab},\quad \Theta := u^a_{;a}. \tag{2.2c}$$

$\dot{u}_a$ is the acceleration vector; $(1.3) \Rightarrow \dot{u}_a u^a = 0$. If an observer on one fluid particle observes a neighboring fluid particle at distance $\delta 1$ in direction

$$e_a \quad (e_a e^a = 1,\ e_a u^a = 0),$$

then[5]

$$\frac{(\delta 1)^{\cdot}}{\delta 1} = \Theta_{ab}e^a e^b = \sigma_{ab}e^a e^b + \tfrac{1}{3}\Theta, \tag{2.3}$$

$$h^b_a(e_b)^{\cdot} = (\omega^b_a + \sigma^b_a - (\sigma_{cd}e^c e^d)h^b_a)e_b, \tag{2.4}$$

give, respectively, the rate of change of distance and direction of the neighboring particle relative to the observer. Also, if δV is the volume enclosed by a group of neighboring particles, then

$$(\delta V)^{\cdot}/\delta V = \Theta. \tag{2.5}$$

From (2.3)–(2.5), we see that Θ_{ab} is the expansion tensor, Θ the (volume) expansion, σ_{ab} the shear tensor, and ω_{ab} the vorticity tensor. The vorticity vector ω^a may be defined by

$$\omega^a := \tfrac{1}{2}\eta^{abcd}u_b\omega_{cd}. \tag{2.6}$$

Then $\omega_{tu} = \eta_{tuas}\omega^a u^s$, so $\omega^a u_a = 0 = \omega^a\omega_{ab}$; the vector ω^a is the vector in the rest space of u^a which defines the instantaneous axis of the rotation of the fluid due to vorticity.

The scalars $\dot{u} := (\dot{u}_a\dot{u}^a)^{\frac{1}{2}}$, $\omega := (\omega^a\omega_a)^{\frac{1}{2}} = (\frac{1}{2}\omega_{ab}\omega^{ab})^{\frac{1}{2}}$, and $\sigma := (\frac{1}{2}\sigma_{ab}\sigma^{ab})^{\frac{1}{2}}$ vanish if and only if the corresponding tensors vanish.

The contracted Bianchi identities (1.4) for dust are

$$\dot{u}_a = 0,\quad \dot{\rho} + \Theta\rho = 0 \tag{2.7}$$

(the world lines of dust are *geodesic*, and the mass of any portion of the fluid is conserved).

[22] G. F. R. Ellis, Ph.D. thesis, Cambridge University, (1964); lecture notes, Cambridge University (1965).

[23] P. M. Cohn, *Lie Groups* (Cambridge University Press, Cambridge, England, 1961); S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry* (John Wiley & Sons, Inc., New York, 1963); C. W. Misner, in *Relativity, Groups and Topology* (Gordon and Breach Science Publishers, Inc., New York, 1963).

[24] J. Schouten, *Ricci Calculus* (Springer-Verlag, Berlin, 1954).

[25] F. Pirani and A. Trautmann, in *Lectures on General Relativity* (Prentice-Hall., Inc., Englewood Cliffs, New Jersey, 1965), Vol. 1.

We now describe the tetrad formalism used in this paper. All the following considerations are purely *local*; they hold in some neighborhood of each point in space–time. To obtain global results would require an investigation into the patching together of these local results.

Let $\{x^i\}$ denote a local coordinate system, $\{e_a\}$ a tetrad. The equations

$$e_a = e_a^i(\partial/\partial x^i) \quad (\Rightarrow \det \|e_a^i\| \neq 0) \tag{2.8}$$

define the functions e_a^i, which are the components of the vectors e_a with respect to the basis $\partial/\partial x^i$, and are also the directional derivatives of the coordinate functions x^i as

$$e_a^i = \partial_a(x^i). \tag{2.9}$$

Tetrad transformations

$$e_{a'} = \Lambda_{a'}^a e_a, \tag{2.10}$$

where $\Lambda_{a'}^a$ is a position-dependent Lorentz matrix, and coordinate transformations

$$x^{i'} = x^{i'}(x^j) \Rightarrow \partial/\partial x^{j'} = \frac{\partial x^i}{\partial x^{j'}}\,\partial/\partial x^i \tag{2.11}$$

both induce changes of the functions e_a^i.

The components of any vector or tensor may be found with respect to the basis $\partial/\partial x^i$ or the basis e_a (see, e.g., Refs. 12, 22, 24). Thus, the vector Z can be written as

$$Z = Z^i(\partial/\partial x^i) = Z^a e_a; \tag{2.12}$$

$(2.8) \Rightarrow Z^i = Z^a e_a^i$. The tetrad components of the metric tensor are

$$\begin{aligned} g_{ab} = e_a \cdot e_b &= (e_a^i \partial/\partial x^i) \cdot (e_b^j \partial/\partial x^j) \\ &= e_a^i e_b^j (\partial/\partial x^i) \cdot (\partial/\partial xj) = e_a^i e_b^j g_{ij}. \end{aligned} \tag{2.13}$$

As the tetrad is orthonormal,

$$g_{ab} = e_a^i e_{bi} = \operatorname{diag} \|-1, 1, 1, 1\|. \tag{2.14}$$

The metric components g^{ab} (the inverse matrix to g_{ab}) are numerically equal to the g_{ab}; tetrad indices are raised and lowered by the metric components g^{ab} and g_{ab}. The components g_a^b, g_i^j of the metric tensor are δ_a^b, δ_i^j; so $(2.13) \Rightarrow e_a^i e_i^b = \delta_a^b$, i.e., e_a^i and e_j^b are inverse matrices. Therefore

$$e_j^b e_b^i = \delta_j^i, \tag{2.15}$$

and the inverse relations to (2.8), (2.13) are

$$\partial/\partial x^i = e_i^a e_a, \tag{2.16}$$

$$g_{ij} = e_i^a e_j^b g_{ab}. \tag{2.17}$$

As (2.8), (2.10), (2.11), (2.16) each simply imply a change of vector basis, tensor algebraic relations are preserved in form under each of these transformations.

The Ricci rotation coefficients, defined by

$$\Gamma_{abc} := e_a \cdot \nabla_b e_c = e_a^i e_{ci;j} e_b^j \tag{2.18}$$

can be regarded as "tetrad components" of the Christoffel symbols. As the g_{ab} are constants,

$$g_{ij;k} = 0 \Rightarrow \Gamma_{abc} + \Gamma_{cba} = 0. \tag{2.19}$$

The commutators $[e_a, e_b]$ of the basis vectors are defined by

$$[e_a, e_b]f := \partial_a(\partial_b f) - \partial_b(\partial_a f). \tag{2.20}$$

For each set of values (a, b) this commutator is a new vector field, the Lie derivative of e_b with respect to e_a, which can be described by its tetrad components γ_{bc}^a:

$$[e_a, e_b] =: \gamma_{ab}^c e_c, \quad \gamma_{ab}^c = \gamma_{[ab]}^c. \tag{2.21}$$

It follows that

$$\gamma_{ab}^c = \Gamma_{ab}^c - \Gamma_{ba}^c. \tag{2.22}$$

From (2.19), the inverse of (2.22) are the "Christoffel relations"

$$\Gamma_{abc} = \tfrac{1}{2}(\gamma_{abc} + \gamma_{cab} - \gamma_{bca}). \tag{2.23}$$

Equations (2.22), (2.23) show that the Γ_{abc} and γ_{abc} are linear combinations of each other; given all the γ_{abc}, all the Γ_{abc} are known, and vice versa.

The general procedures used in this paper may be summarized as follows: the γ_{abc} are used as auxiliary variables, obeying certain identities and field equations. As far as possible in any given situation, we determine the γ_{abc} first, then find the e_a^i [from (2.20), (2.21) and any remaining field equations] the e_j^b [from (2.15)] and finally the g_{ij} [from (2.17), (2.14)]. In carrying out this program, the simplifications to the equations possible via the coordinate and tetrad choice (2.10), (2.11) available are essential to making solution possible.[26]

At this stage, we make the first *tetrad specialization*: the timelike vector e_0 is chosen as the fluid flow vector u. Then

$$u = e_0, \quad u^a = \delta_0^a, \quad u_a = -\delta_a^0. \tag{2.24}$$

Now the unit vectors $\{e_\nu\} = \{e_1, e_2, e_3\}$ form an orthogonal triad in the rest space of u at each point.

Each of the γ_{bc}^a has a direct geometrical interpretation as a component of a Lie derivative [see (2.21)]. However, it is often useful to have alternative interpretations for these coefficients; in particular, the choice (2.24) of the vector e_0 implies some simple

[26] It should be noted that, if desired, one can easily use 1-forms instead of vectors in what follows. The basis θ^a of 1-forms dual to the basis e_b of vectors is $\theta^a = e_i^a\,dx^i$, and obeys $\langle \theta^a, e_b \rangle = \delta_b^a$. Then $(2.21) \Rightarrow d\theta^a = -\tfrac{1}{2}\gamma_{bc}^a \theta^b \wedge \theta^c$, which makes clear the role of the functions γ_{bc}^a in terms of 1-forms. However, in this paper we proceed using vectors.

relations of the γ^a_{bc} to the tensors occurring in (2.2)–(2.6). These relations may be obtained by the technique illustrated in the following derivation[27] of the tetrad components $T_{ab;c}$ of the tensor with coordinate components $T_{ij;k}$:

$$T_{ab;c} = T_{ij;k}e^i_a e^j_b e^k_c = (T_{ij}e^i_a e^j_b)_{;k}e^k_c - T_{ij}(e^i_{a;k})e^j_b e^k_c - T_{ij}e^i_a(e^j_{b;k})e^k_c = \partial_c T_{ab} - T_{fb}\Gamma^f_{ca} - T_{af}\Gamma^f_{cb}; \quad (2.25)$$

Eq. (2.23) enables us to write out this expression in terms of the γ_{abc}. Similarly, we can obtain the explicit tetrad components of any tensor equation.

Using this technique on (2.2), (2.6) with the tetrad restriction (2.24), we find

$$\gamma^0_{0\nu} = \dot{u}_\nu, \quad (2.26a)$$

$$\gamma^\nu_{0\nu} = -\Theta_{\nu\nu} =: -\Theta_\nu \quad \text{(no sum)}, \quad (2.26b)$$

$$\gamma^0_{\mu\nu} = -2\epsilon_{\mu\nu\sigma}\omega^\sigma, \quad (2.26c)$$

$$\gamma^\mu_{0\nu} = e_\mu \cdot \dot{e}_\nu - \sigma_{\mu\nu} - \epsilon_{\mu\nu\sigma}\omega^\sigma \quad (\mu \neq \nu), \quad (2.26d)$$

$$\gamma^\mu_{\nu\mu} = e_\nu \cdot \nabla_\mu e_\mu \quad \text{(no sum)}, \quad (2.26e)$$

where

$$\epsilon_{\mu\nu\sigma} := \eta_{\mu\nu\sigma a}u^a, \quad \text{so} \quad \epsilon_{\mu\nu\sigma} = \epsilon_{[\mu\nu\sigma]}, \quad \epsilon_{123} = 1;$$

and

$$\Theta = \Theta_1 + \Theta_2 + \Theta_3 = -(\gamma^1_{01} + \gamma^2_{02} + \gamma^3_{03}). \quad (2.26f)$$

In these equations, the $\gamma^0_{0\nu}$ are acceleration components, zero for dust [see (2.7)]. The terms Θ_ν are the expansions of the fluid in the directions e_ν [see (2.3)], which are directly observable[5] in terms of redshift. The terms (2.26c) give the vorticity vector components; then the terms (2.26d) determine the three shear cross terms ($\sigma_{\mu\nu} = \Theta_{\mu\nu}, \mu \neq \nu$) and the Fermi-derivatives $e_\mu \cdot \dot{e}_\nu$ of the basis vectors along the world lines. It is reasonable to regard the frame defined by the vectors e_ν as a physically nonrotating frame if and only if the 3 terms $e_\nu \cdot \dot{e}_\mu = -e_\mu \cdot \dot{e}_\nu$ vanish.[25] The terms in (2.26e) are all spatial components of the first curvatures ("accelerations") of the vectors e_ν. Three terms $\gamma^\sigma_{\mu\nu}$ ($\sigma \neq \nu \neq \mu \neq \sigma$) are not listed in (2.26).

The Ricci identities are

$$v^b_{;cd} - v^b_{;dc} = R^b_{ecd}v^e. \quad (2.27)$$

Choosing v as the basis vector e_f ($v^a = \delta^a_f$), the methods of (2.25) applied to (2.27) yield

$$R^f_{bcd} = -\partial_d\Gamma^f_{cb} + \partial_c\Gamma^f_{db} + \Gamma^f_{cg}\Gamma^g_{db} - \Gamma^f_{dg}\Gamma^g_{cb} + \Gamma^f_{gb}\gamma^g_{dc}. \quad (2.28)$$

[27] There are more direct and fundamental ways of obtaining the tetrad-coordinate relations described in this section; see, e.g., Refs. 22, 24. However, space restrictions preclude such a presentation here.

Contraction of (2.28) gives the tetrad Ricci tensor components R_{bd}; these can be combined with (1.1)–(1.3) to give

$$R_{bd} = \partial_d\Gamma^c_{cb} - \partial_c\Gamma^c_{db} - \Gamma^c_{cg}\Gamma^g_{db} + \Gamma^g_{cb}\Gamma^c_{gd} = -(\tfrac{1}{2}\rho + \Lambda)h_{bd} - (\tfrac{1}{2}\rho - \Lambda)u_b u_d. \quad (2.29)$$

With (2.23), these are the 10 *field equations*, a set of differential equations for the γ^a_{bc}; these equations are numbered by the indices (bd).

The Riemann tensor components (2.28) are not independent, but are algebraically related by the cyclic identities

$$R^f_{[bcd]} = 0. \quad (2.30)$$

From (2.23), these equations are the conditions

$$\partial_{[d}\gamma^f_{cb]} + \gamma^g_{[dc}\gamma^f_{b]g} = 0, \quad (2.31)$$

which are in fact [from (2.20), (2.21)] the components of the *Jacobi identities*

$$[e_b, [e_c, e_d]] + [e_d, [e_b, e_c]] + [e_c, [e_d, e_b]] = 0 \quad (2.32)$$

for the basis vector fields e_a. The 16 equations (2.31) are numbered by the indices $\binom{f}{bcd}$.

Certain of the field equations and Jacobi identities may be obtained as follows[5]: in the Ricci identities (2.27), the vector v^a is chosen as u^a and the decomposition (2.2) of $u_{a;b}$ is substituted into the resulting expression for $R_{abcd}u^a$. Contracting this equation, we obtain (using the fact that for dust, $\dot{u}^a = 0$)

$$R_{ab}u^a u^b = \dot{\Theta} + \tfrac{1}{3}\Theta^2 + 2(\sigma^2 - \omega^2) = -\tfrac{1}{2}\rho + \Lambda, \quad (2.33)$$

which is *Raychaudhuri's equation*,[4] the covariant form of the (00) equation; and

$$h^{ab}R_{bc}u^c = h^a_b(\omega^{bc}_{;c} - \sigma^{bc}_{;c} + \tfrac{2}{3}\Theta^{,b}) = 0, \quad (2.34)$$

the covariant form of the (0ν) equations, which state that u^a is a Ricci eigenvector. The equations

$$R_{[abc]d}u^d = 0 \Leftrightarrow u_{[a;bc]} = 0,$$

split into the identity

$$\omega^a_{;a} = 0 \quad (2.35a)$$

and the vorticity propagation equations

$$(\omega^e)^{\cdot} = \sigma^e_d\omega^d - \tfrac{2}{3}\Theta\omega^e \quad (2.35b)$$

[(2.35a) is the covariant form of $\binom{0}{123}$, while (2.35b) is the covariant form of $\binom{0}{0\mu\nu}$, $(\mu \neq \nu)$].

3. A TETRAD AND COORDINATE SYSTEM FOR DUST

The coordinate system thus far is completely general, while the tetrad system has been restricted only by the choice (2.24): $u = e_0$. The metric form [from (2.1), (2.14), and (2.17)] is

$$Q := g_{ij}\,dx^i\,dx^j = q - (u_j\,dx^j)^2, \tag{3.1a}$$

where

$$q := h_{ij}\,dx^i\,dx^j = (e_{1i}\,dx^i)^2 + (e_{2j}\,dx^j)^2 + (e_{3k}\,dx^k)^2. \tag{3.1b}$$

Theorem 3.1: Given any space–time containing dust, a tetrad system can be locally found with $u^i = e_0^i$, $\omega^i = \omega e_1^i$, such that

$$\gamma^0_{0\nu} = \gamma^0_{12} = \gamma^0_{13} = \gamma^3_{12} = 0, \quad \gamma^\mu_{0\nu} = 0 \quad (\mu > \nu),$$
$$\gamma^\mu_{0\nu} = -2\sigma_{\mu\nu} \quad (\mu < \nu). \tag{3.2a}$$

Associated coordinates can be locally found, for which

$$e_0^\nu = e_1^0 = e_2^0 = 0, \quad e_3^0 = y(x^2, x^3)e_3^3,$$
$$e_\mu^\nu = 0 \quad (\mu > \nu), \quad e_0^0 = 1. \tag{3.2b}$$

The metric has the form

$$Q = q - (dx^0 - y(x^2, x^3)\,dx^3)^2 \tag{3.2c}$$

in these coordinates; y can be chosen zero if and only if $\omega = 0$.

In fact, this result holds for any geodesic timelike congruence with tangent vector u^i. The coordinates here are the coordinates found by Ehlers,[6] except that the coordinate x^2 used here is slightly more general: Ehlers uses $x^{2'} = -y(x^2, x^3)$.

Proof: We prove Theorem 3.1 by a series of specializations of a general coordinate and tetrad system. As all considerations are local, "everywhere" means "everywhere in a suitable neighborhood."

The first coordinate specialization fits the coordinates to the vector e_0. A transformation $x^{0'} = x^0$, $x^{\nu'} = x^{\nu'}(x^i)$ is used to set

$$e_0^\nu = \partial_0(x^\nu) = 0. \tag{3.3a}$$

The coordinates are then comoving coordinates: the world lines (integral curves of e_0) lie everywhere in the intersections of the surfaces $x^\nu = \text{const}$, and so are lines on which only x^0 varies (integral curves of $\partial/\partial x^0$). A transformation $x^{0'} = x^{0'}(x^i)$, $x^{\nu'} = x^\nu$ is next used to set

$$e_0^0 = \partial_0(x^0) = 1 \tag{3.3b}$$

everywhere. Then $u = \partial/\partial x^0$, $u^i = \delta^i_0$, and (1.3) $\Rightarrow$ $g_{00} = -1$; x^0 measures proper time along each world line. The coordinate freedom preserving (3.3) is: $x^{0'} = x^0 + f(x^\nu)$ (a "gauge transformation", representing freedom of choice of an initial surface $x^0 = \text{const}$) and $x^{\nu'} = x^{\nu'}(x^\sigma)$ (freedom to initially label the set of particles arbitrarily).

The flow lines are geodesic, so $\gamma^0_{0\nu} = 0$. In the second tetrad specialization, we first consider the case $\omega \neq 0$. Then the vector e_1 can be chosen parallel to the vorticity vector, so

$$\omega_2 = \omega_3 = 0 \Leftrightarrow \gamma^0_{31} = \gamma^0_{12} = 0 \tag{3.4a}$$

and $\omega = \omega_1 = -\frac{1}{2}\gamma^0_{23}$ is the only nonzero vorticity component. The equations $\begin{pmatrix}0\\012\end{pmatrix}$, $\begin{pmatrix}0\\013\end{pmatrix}$ now show

$$\gamma^2_{01} = \gamma^3_{01} = 0 \Leftrightarrow \sigma_{12} = e_2 \cdot \dot{e}_1, \quad \sigma_{13} = e_3 \cdot \dot{e}_1. \tag{3.4b}$$

$\begin{pmatrix}0\\023\end{pmatrix}$ is now

$$\partial_0\omega = -\omega(\Theta_2 + \Theta_3). \tag{3.5}$$

Equations (3.4b), (3.5) are the vorticity propagation equations (2.35b) (see Refs. 2, 3, 5, 11 for geometrical interpretation). Now suppose $\omega = 0$. Then (3.4a) is identically satisfied. The vector e_1 can be chosen arbitrarily on an initial surface $x^0 = \text{const}$; and can then be propagated along the world lines of matter in such a way that (3.4b) holds. Thus, whether ω is zero or not, the tetrad can be chosen so that (3.4a, b) holds. With these restrictions, the commutation relation of u and e_1 is

$$[u, e_1] = \gamma^1_{01}e_1, \tag{3.6}$$

which implies[5] that the flow vector and vorticity vector are 2-surface forming.

The vectors e_2, e_3 are still arbitrary by a rotation

$$e_{2'} = \cos\theta e_2 + \sin\theta e_3,$$
$$e_{3'} = -\sin\theta e_2 + \cos\theta e_3, \tag{3.7}$$

where θ is an arbitrary function $\theta(x^i)$. We perform such a rotation, determining $\partial\theta/\partial x^0$ by the condition that finally

$$\gamma^3_{02} = 0 \Leftrightarrow e_2 \cdot \dot{e}_3 = \omega_1 - \sigma_{23}. \tag{3.4c}$$

This leaves e_2, e_3 still arbitrary by an initial rotation in a surface $x^0 = \text{const}$.

The second coordinate specialization fits the coordinates to the vector e_1. We choose an initial surface $x^0 = c^0$, and within this surface relabel the particles by the transformation $x^{0'} = x^0$, $x^{1'} = x^1$, $x^{2'} = x^{2'}(x^\sigma)$, $x^{3'} = x^{3'}(x^\sigma)$; $x^{2'}$ and $x^{3'}$ are chosen so that $e_1^2 = e_1^3 = 0$ holds within the hypersurface $x^0 = c^0$. Applying the commutator (3.6) to the coordinate function x^2 and using (3.3a) shows: $\partial_0(e_1^2) = e_1^2\gamma^1_{01}$; so $e_1^2 = 0$ at all later times, having been chosen zero initially. Similar results hold for x^3, so we obtain

$$e_1^2 = \partial_1(x^2) = 0, \quad e_1^3 = \partial_1(x^3) = 0 \tag{3.8a}$$

at all points; the vector e_1 lies in the surfaces $x^2 =$ const, $x^3 =$ const. Since x^2, x^3 are comoving coordinates, this implies the result[11]: vortex lines consist at all times of the same particles of matter. Next, a transformation $x^{0'} = x^0 + f(x^\nu)$, $x^{\sigma'} = x^\sigma$ is used to set $e_1^0 = 0$ in some surface $x^0 =$ const: applying (3.6) to x^0 and using (3.3b) shows that

$$e_1^0 = \partial_1(x^0) = 0 \tag{3.8b}$$

holds at all points; the vector e_1 lies everywhere in the surface $x^0 =$ const.

The coordinate freedom remaining, that preserves (3.3), (3.8) is

$$x^{0'} = x^0 + f(x^2, x^3), \quad x^{1'} = x^{1'}(x^1, x^2, x^3),$$
$$x^{2'} = x^{2'}(x^2, x^3), \quad x^{3'} = x^{3'}(x^2, x^3). \tag{3.9}$$

The inverse transformations $x^i = x^i(x^{i'})$ have the same form as (3.9).

The third tetrad step is a rotation (3.7) which preserves (3.4c), where the value of $\partial\theta/\partial x^1$ is chosen by the condition: γ^3_{12} is set zero in a surface $x^0 =$ const. The Jacobi identity $\begin{pmatrix} 3 \\ 012 \end{pmatrix}$ is

$$\partial_0(\gamma^3_{12}) = \gamma^3_{12}(\gamma^1_{01} + \gamma^2_{02} - \gamma^3_{03}),$$

so we obtain everywhere

$$\gamma^3_{12} = 0. \tag{3.10}$$

This choice makes e_3 hypersurface-orthogonal, as $\gamma^3_{01} = \gamma^3_{12} = \gamma^3_{02} = 0 \Rightarrow$ the vectors (e_0, e_1, e_2) are hypersurface-forming.

The third coordinate step is to fit the coordinates to e_2, e_3. We choose the surfaces orthogonal to e_3 as the surfaces $x^3 =$ const; this defines x^3 up to transformations $x^{3'} = x^{3'}(x^3)$. Then, in addition to (3.3a), (3.8a) we have

$$\partial_2(x^3) = e_2^3 = 0. \tag{3.11a}$$

A gauge transformation $x^{0'} = x^0 + f(x^2, x^3)$, $x^{\sigma'} = x^\sigma$ can be used to set $e_2^0 = 0$ in an initial 2-surface $x^0 =$ const, $x^1 =$ const; then the commutators $[e_0, e_2]$ and $[e_1, e_2]$ applied to x^0 show that this holds at all points, so that

$$e_2^0 = \partial_2(x^0) = 0. \tag{3.11b}$$

The vector e_3 is then the vector at each point giving the spacelike direction in which the surfaces $x^0 =$ const are *not* orthogonal to u.

Using all the above restrictions on the e^i_a and γ^a_{bc}, (2.20), (2.26) show that

$$\dot{u}_3 = \omega_2 = 0 \Rightarrow e_3^0 = y(x^2, x^3)e_3^3$$

is the form of e_3^0. If and only if $\omega = 0$, $\gamma^0_{23} = \gamma^0_{31} = \gamma^0_{12} = 0$, and (e_1, e_2, e_3) form hypersurfaces orthogonal to u. Then a gauge transformation choosing these surfaces as the surfaces $x^0 =$ const determines x^0 up to an additive constant, and puts $y(x^2, x^3) = 0$.

This completes the specialization of tetrad and coordinates. The tetrad-coordinate relations (2.8) with restrictions (3.2b) are given explicitly in the Appendix [Eqs. (A1)]. The metric components g_{ij} [found from (2.17), (2.15)] are given in the Appendix [Eqs. (A2)]. As these have the form (3.2c), Theorem 3.1 is proved.

Substituting (A1) into (2.20), (2.21) gives explicitly the γ^a_{bc} in terms of the functions e^i_a; these are listed in the Appendix [Eq. (A3)]. When the restrictions (3.2a) on the γ^a_{bc} are taken into account, there remain 13 nontrivial Jacobi identities (2.31); these and the 10 field equations (2.29) are also given in the Appendix [Eqs. (A4)].

In using these equations, we need to know the remaining coordinate and the tetrad freedom that preserves the conditions (3.2a, b).

We consider first the remaining tetrad freedom. u is defined uniquely as the timelike eigenvector of the Ricci tensor. If $\omega \neq 0$, e_1 is uniquely defined; if $\omega = 0$, there is complete freedom of choice of e_1 in an initial surface $x^0 =$ const, the propagation of e_1 along the world lines being determined by (3.4b). In either case, e_2 and e_3 are free by a rotation (3.7) that preserves (3.4c), (3.10). From (2.20), (2.21) these conditions are preserved if $\theta(x^i)$ obeys

$$\partial\theta/\partial x^0 = (\gamma^2_{02} - \gamma^3_{03}) \sin\theta \cos\theta + \gamma^2_{03} \sin^2\theta,$$
$$e_1^1(\partial\theta/\partial x^1) = (\gamma^2_{12} - \gamma^3_{13}) \sin\theta \cos\theta + \gamma^2_{13} \sin^2\theta, \tag{3.12}$$

so we choose arbitrarily

$$\theta_0(x^2, x^3) := \theta(c^0, c^1, x^2, x^3)$$

on some 2-surface $x^0 = c^0$, $x^1 = c^1$ and then find $\theta(x^i)$ from (3.12).

The transformation properties of the γ^a_{bc} under this rotation are listed in the Appendix [Eqs. (A5)]. It is clear that the scalar invariants,

$$\omega = -\tfrac{1}{2}\gamma^0_{23}, \quad \Theta = -\gamma^1_{01} - \gamma^2_{02} - \gamma^3_{03},$$
$$\sigma = [\tfrac{1}{3}((\gamma^1_{01})^2 + (\gamma^2_{02})^2 + (\gamma^3_{03})^2$$
$$- \gamma^1_{01}\gamma^2_{02} - \gamma^2_{02}\gamma^3_{03} - \gamma^3_{03}\gamma^1_{01})$$
$$+ \tfrac{1}{4}((\gamma^1_{02})^2 + (\gamma^2_{03})^2 + (\gamma^1_{03})^2)]^{\frac{1}{2}}, \tag{3.13}$$

remain unchanged during this rotation. In addition, the following quantities remain invariant:

(i) the expansion of the fluid in the e_1 direction,

$$-\gamma^1_{01} = \Theta_1; \tag{3.14a}$$

(ii) the area expansion of the fluid in the plane perpendicular to e_1,

$$-\gamma^2_{02} - \gamma^3_{03} = \Theta_2 + \Theta_3; \qquad (3.14b)$$

(iii) a measure of the shear in the plane perpendicular to e_1,

$$(\gamma^2_{02} - \gamma^3_{03})^2 + (\gamma^2_{03})^2 = (\Theta_2 - \Theta_3)^2 + 4(\sigma_{23})^2; \qquad (3.14c)$$

(iv) a measure of how much e_1 differs from being a shear eigenvector,

$$\tfrac{1}{4}((\gamma^1_{02})^2 + (\gamma^1_{03})^2) = (\sigma_{12})^2 + (\sigma_{13})^2; \qquad (3.14d)$$

(v) the magnitude of the spatial projection of the first curvature of e_1,

$$(\gamma^1_{21})^2 + (\gamma^1_{31})^2 = (e_2 \cdot \nabla_1 e_1)^2 + (e_3 \cdot \nabla_1 e_1)^2; \qquad (3.14e)$$

(vi) the invariant $\omega^{-2}\eta^{abcd}u_a\omega_b\omega_{c;d}$ if $\omega \neq 0$, (if $\omega = 0$, the corresponding expression for e_1) which is, up to a factor,

$$\gamma^1_{23}; \qquad (3.14f)$$

(vii) the expansion in the rest space of u of the e_1 lines as one travels along them:

$$-(\gamma^2_{12} + \gamma^3_{13}) = e_2 \cdot \nabla_2 e_1 + e_3 \cdot \nabla_3 e_1; \qquad (3.14g)$$

(viii) the shear in the rest space of u of the e_1 lines as one travels along them:

$$(\gamma^2_{31})^2 + (\gamma^2_{12} - \gamma^3_{13})^2. \qquad (3.14h)$$

In fact, (3.14a–d) $\Rightarrow \Theta$ and σ are invariant; so in (3.13), only the invariance of ω gives a condition not in (3.14). If the e_1 vector is covariantly defined (e.g., if $\omega \neq 0$) then all of the above quantities are scalar invariants of the curvature tensor.

Under this rotation, the tetrad-coordinate relations (A1) are to be preserved; so the tetrad rotation must be accompanied by a *coordinate transformation* (3.9). As (3.11a, b) must be preserved, the following relations hold:

$$e^2_2 \frac{\partial x^{3'}}{\partial x^2} \cos\theta + \left(e^2_3 \frac{\partial x^{3'}}{\partial x^2} + e^3_3 \frac{\partial x^{3'}}{\partial x^3}\right) \sin\theta = 0, \qquad (3.15a)$$

$$e^2_2 \frac{\partial f}{\partial x^2} \cos\theta + \left(e^2_3 \frac{\partial f}{\partial x^2} + e^3_3\left(y + \frac{\partial f}{\partial x^3}\right)\right) \sin\theta = 0, \qquad (3.15b)$$

which imply

$$\frac{\partial x^{3'}}{\partial x^2}\left(y + \frac{\partial f}{\partial x^2}\right) = \frac{\partial f}{\partial x^2}\frac{\partial x^{3'}}{\partial x^3}. \qquad (3.15c)$$

From (3.7), (3.9), and (A1), the way in which the e^i_a transform can be found; this is given in the Appendix [Eqs. (A6)]. It is convenient to consider this freedom of coordinates and tetrad in two parts.

Lemma 3.2: The tetrad freedom preserving the conditions of Theorem 3.1 is a rotation (3.7) obeying (3.12). This must be accompanied by a coordinate transformation (3.9) satisfying the conditions (3.15); this transformation can be chosen to have the form $x^{0'} = x^0 + f(x^2, x^3)$, $x^{1'} = x^1$, $x^{2'} = x^2$, $x^{3'} = x^{3'}(x^2, x^3)$. The quantities (3.13), (3.14) are invariant under this rotation.

In fact, given any one of θ_0, $x^{3'}$ or f as an arbitrary function of x^2, x^3, the other two are determined (up to some initial conditions) by (3.15). Generally, one chooses θ_0 so as to simplify the γ^a_{bc} in a convenient way, and then determines f and $x^{3'}$ from (3.15).

Lemma 3.3: The coordinate freedom preserving the conditions of Theorem 3.1 for a given tetrad is

$$x^{0'} = x^0 + f(x^3), \quad x^{1'} = x^{1'}(x^1, x^2, x^3),$$
$$x^{2'} = x^{2'}(x^2, x^3), \quad x^{3'} = x^{3'}(x^3). \qquad (3.16)$$

If $\omega = 0$ and y is chosen zero, then $f(x^3)$ is a constant.

This follows from (3.9, 15). These transformations can be used [according to (A6) with $\theta = 0$] to simplify the e^i_a in some convenient way, and hence [via (A2)] to simplify the g_{ij}.

The transformation (3.16) can be choosen so that at one point (e.g., the origin of the coordinates)

$$e^i_a = \delta^i_a, \qquad (3.17)$$

which corresponds to the fact that, at a point, the metric can always be written in the special relativistic form. The field equations and Jacobi identities then determine over how large a submanifold this remains true.

Many other coordinate and tetrad specializations different from those given here are clearly possible. However, the system described here is convenient in many applications.

4. LOCALLY ROTATIONALLY SYMMETRIC DUST

The conditions (A_r), (B), and (C) defined below are possible symmetries of space–time:

(A_r) At each point p in an open neighborhood U of a point p_0, there exists a nondiscrete subgroup g of the Lorentz group in the tangent space T_p which leaves invariant the curvature tensor and all its covariant derivatives to the rth order.

(B) At each point p in an open neighborhood V of a point p_0, there exists a nondiscrete isotropy group.
(C) There exists a local group of motions G_r in an open neighborhood W of a point p_0 which is multiply transitive on some q surface through each point p of W (so $r > q$).

Clearly $(A_r) \Rightarrow (A_{r-1})$; it is known (see Refs. 22, 28) that

$$(C) \Rightarrow (B) \Rightarrow (A_\infty). \tag{4.1}$$

In this section, we consider a dust-filled space–time in which (A_3) holds on some open set U; we say that the dust is *locally rotationally symmetric*[29] in U. By (4.1), consideration of all cases of locally rotationally symmetric dust in U includes consideration of all dust-filled universes in which either (B) or (C) (for arbitrary r, q) hold in U.

The vector u^a and the density ρ are defined algebraically by R_{ab}. R_{ab} and its covariant derivatives (up to the 3rd) are invariant under g so u^a, ρ, and their covariant derivatives (up to the 3rd) are also invariant[30] under g. Therefore, at each point p, g operates in the subspace of T_p orthogonal to u^a; g can thus be either a 1-dimensional or a 3-dimensional group of rotations in T_p. We assume continuity of dimension of g in U.

If g is a 3-dimensional group of rotations, the situation is well understood (see Ref. 5). Then

$$(A_1) \Rightarrow u_{a;b} \text{ is invariant under } g, \text{ so: } \omega = \sigma = 0. \tag{4.2a}$$

Now

$$(4.2a) \Rightarrow [\text{via } (2.34, 33)]\ h_a^b \Theta_{,b} = h_a^b \rho_{,b} = 0. \tag{4.2b}$$

The 3-spaces orthogonal to u^a have constant curvature[4,5] and the space–time in U is *locally the same as a Friedmann world model.* [This result follows already from the condition (A_1); (4.2b) is also a direct consequence of the stronger condition (A_2).]

We assume now that at least one of ω, σ is nonzero; then g is 1-dimensional.

Lemma 4.1: If an open set U contains locally rotationally symmetric dust with either ω or σ nonzero, then a coordinate and tetrad system satisfying the conditions of Theorem 3.1 can be found in an open neighborhood of any point in U. The tetrad vectors e_0, e_1 and the quantities (3.13), (3.14) are uniquely determined algebraically by the curvature tensor and its covariant derivatives (up to the 2nd order).

Proof: Introduce an orthonormal tetrad with the timelike vector e_0 as the velocity vector u, and the vector e_1 at each point as the axis of the rotation g; then g has the form (3.7). Rotational symmetry $(A_1) \Rightarrow \omega_2 = \omega_3 = \sigma_{12} = \sigma_{13} = 0$. The vector e_1 is then uniquely defined throughout space–time as the direction of ω^i, as a unique shear eigenvector, or both; we have to show it satisfies the propagation equations (3.4b). Since u and e_1 are covariantly defined, $\dot{e}_1 := \nabla_u e_1$ is covariantly defined; this vector [by (A_2)] must be invariant under g, so $e_2 \cdot \dot{e}_1 = e_3 \cdot \dot{e}_1 = 0$. Thus (3.4b) is satisfied, and the tetrad and coordinates may be chosen as in Sec. 3.

Using this tetrad system, the local rotational symmetry implies strong restrictions on the γ^a_{bc}. Rotational symmetry (A_1) of the expansion quadric shows

$$\sigma_{12} = \sigma_{23} = \sigma_{31} = 0, \quad \Theta_2 = \Theta_3 =: \beta. \tag{4.3a}$$

Defining $\alpha := \Theta_1$, α and β are the expansions of the fluid in the e_1 and perpendicular directions respectively. The vector $\mathbf{\nabla}_1 e_1$ is covariantly defined, so

$$(A_2) \Rightarrow e_2 \cdot \mathbf{\nabla}_1 e_1 = e_3 \cdot \mathbf{\nabla}_1 e_1 = 0 \Leftrightarrow \gamma^1_{21} = \gamma^1_{31} = 0. \tag{4.3b}$$

The integral lines of e_1 must have no "shear" ($e_{1a;b}$ must be invariant under g), so

$$\gamma^2_{13} = 0, \quad \gamma^2_{12} = \gamma^3_{13} =: a. \tag{4.3c}$$

Finally, covariantly defined scalars must not determine any direction in the tangent space T_p perpendicular to u and e_1; so

$$(A_1) \Rightarrow \partial_2 \rho = \partial_3 \rho, \tag{4.3d}$$

$$(A_2) \Rightarrow \partial_2 \alpha = \partial_3 \alpha, \quad \partial_2 \beta = \partial_3 \beta, \quad \partial_2 \omega = \partial_3 \omega, \tag{4.3e}$$

$$(A_3) \Rightarrow \partial_2 a = \partial_3 a, \quad \partial_2 \gamma^1_{32} = \partial_3 \gamma^1_{32}, \tag{4.3f}$$

where we use the definitions $\omega := \frac{1}{2}\gamma^0_{32}$. The conditions (4.3) may be stated as follows: the invariants (3.14c, d, e, h) vanish; the invariants (3.14a, b, f, g) and (3.13) have equal derivatives in the e_2, e_3 directions.

These restrictions on the γ^a_{bc} enable us to separate out three classes of solutions of the Jacobi identities and field equations. Restrictions (4.3) in equations $\binom{1}{012}$, $\binom{1}{013}$, (02), (03), (12), (13), $\binom{2}{123}$ show:

$$\partial_2 \alpha = \partial_3 \alpha = \partial_2 \beta = \partial_3 \beta = \partial_2 \omega = \partial_3 \omega = \partial_2 a = \partial_3 a = \partial_2 \gamma^1_{32} = \partial_3 \gamma^1_{32} = 0; \tag{4.4a}$$

[28] R. P. Kerr, J. Math. Mech. **12**, 33 (1963). For the theory of groups of motions in Riemannian spaces, see, e.g., L. P. Eisenhart, *Continuous Groups of Transformations* (Dover Publications, Inc., New York, 1961); and A. Petrov, *Einstein-Raume* (Akadamie-Verlag, Berlin, 1964).

[29] As there is rotational symmetry in the tangent space at *each* point in U, this is *not* the general case of axial symmetry about an axis.

[30] See, e.g., L. P. Eisenhart, Math. Ann. **36**, 823 (1935).

then the (00) equation shows

$$\partial_2\rho = \partial_3\rho = 0. \tag{4.4b}$$

The commutator of the vectors e_2, e_3 is

$$[e_2, e_3] = -2\omega e_0 - \gamma^1_{32}e_1 - \gamma^2_{32}e_2 + \gamma^3_{23}e_3. \tag{4.5}$$

Applying this to ω,

$$\partial_2(\partial_3\omega) - \partial_3(\partial_2\omega) = 0 = -2\omega\partial_0\omega - \gamma^1_{32}\partial_1\omega.$$

Substituting in $\begin{pmatrix}0\\023\end{pmatrix}$, $\begin{pmatrix}0\\123\end{pmatrix}$ which are: $\partial_0\omega = -2\beta\omega$, $\partial_1\omega = 2a\omega$, we find that $\omega \neq 0 \Rightarrow$

$$\gamma^1_{32}a = 2\beta\omega. \tag{4.6}$$

If $\omega = 0$, then $[e_2, e_3]\gamma^1_{32} = 0 = -\gamma^1_{32}\partial_1\gamma^1_{32}$; so either $\gamma^1_{32} = 0$ and (4.6) still holds, or $\gamma^1_{32} \neq 0$, $\partial_1\gamma^1_{32} = 0$. But $\begin{pmatrix}1\\123\end{pmatrix}$ is: $-\partial_1\gamma^1_{32} = 2\omega\alpha - 2\gamma^1_{32}a$; so the second case $\Rightarrow a = 0$, and (4.6) holds in all cases.

To simplify later work, the coordinates and tetrad are specialized at this stage. Since $\gamma^2_{03} = \gamma^2_{13} = \gamma^2_{01} = 0$, e_2 is hypersurface-orthogonal. Choosing these surfaces as the surfaces $x^2 = \text{const}$,

$$\partial_3(x^2) = e^2_3 = 0. \tag{4.7a}$$

Since $\gamma^1_{02} = 0$, we can choose x^1 so that

$$e^1_2 = \partial_2(x^1) = 0. \tag{4.7b}$$

The remaining coordinate freedom (if the tetrad is left invariant) is

$$x^{0'} = x^0 + f(x^3), \quad x^{1'} = x^{1'}(x^1, x^3),$$
$$x^{2'} = x^{2'}(x^2), \quad x^{3'} = x^{3'}(x^3). \tag{4.8}$$

The tetrad freedom can be used to set $\gamma^2_{32} = 0$ on a 2-surface $x^0 = c^0$, $x^1 = c^1$; and to set $\partial_3(\gamma^3_{23}) = 0$ on a line $x^0 = c^0$, $x^1 = c^1$, $x^2 = c^2$ in the 2-surface. Since equations $\begin{pmatrix}2\\023\end{pmatrix}$, (31) are: $\partial_0(\gamma^2_{32}) = -\beta\gamma^2_{32}$, $\partial_1(\gamma^2_{32}) = a\gamma^2_{32}$,

$$\text{this choice} \Rightarrow \gamma^2_{32} = 0 \tag{4.9a}$$

holds everywhere. $\begin{pmatrix}3\\032\end{pmatrix}$ and $\begin{pmatrix}2\\032\end{pmatrix}$ are: $\partial_0(\gamma^3_{23}) = -\beta\gamma^3_{23}$, $\partial_1(\gamma^3_{23}) = a\gamma^3_{23}$. The (22) equation $\Rightarrow \partial_3(\partial_2\gamma^3_{23}) = 2\gamma^3_{23}\partial_3\gamma^3_{23}$. These equations and the commutation relations, together with (4.6), show:

$$\partial_0(\partial_3\gamma^3_{23}) = -2\beta\partial_3\gamma^3_{23}, \quad \partial_1(\partial_3\gamma^3_{23}) = 2a\partial_3\gamma^3_{23},$$
$$\partial_2(\partial_3\gamma^3_{23}) = 3\gamma^3_{23}\partial_3\gamma^3_{23}.$$

Therefore,

$$\text{this tetrad choice also} \Rightarrow \partial_3\gamma^3_{23} = 0 \tag{4.9b}$$

holds everywhere. The tetrad is now almost completely determined; the remaining freedom (we may choose θ and γ^3_{23} arbitrarily at one point) is left open, to be determined later by the choice of coordinates. The commutator (4.5) is now applied to β; $\partial_0\beta$ is found from combination of equations (00) + (11) − (22) − (33); and $\partial_1\beta$ is found from (01). Using (4.6), we obtain

$$\omega[-\Lambda + \partial_2\gamma^3_{23} - (\gamma^3_{23})^2 + \omega^2 - a^2 - \tfrac{1}{4}(\gamma^1_{32})^2 + \beta^2 + 2\alpha\beta] = 0. \tag{4.10a}$$

The commutator (4.5) can also be applied to a; where $\partial_0 a$ is found from $\begin{pmatrix}2\\012\end{pmatrix}$ and (01), while $\partial_1 a$ is found from the combination of equations (00) + (11) + (22) + (33). Using (4.6), we obtain

$$\gamma^1_{32}[\rho + \Lambda - \partial_2\gamma^3_{23} + (\gamma^3_{23})^2 + a^2 - \omega^2 + \tfrac{1}{4}(\gamma^1_{32})^2 - \beta^2 - 2\alpha\beta] = 0. \tag{4.10b}$$

From (4.10a, b),

$$\rho \neq 0 \Rightarrow \omega\gamma^1_{32} = 0. \tag{4.11}$$

Also

$$\begin{pmatrix}1\\123\end{pmatrix}, (4.6) \Rightarrow \partial_1\gamma^1_{32} = 2\omega(2\beta - \alpha), \tag{4.12}$$

and

$$(01), (4.11) \Rightarrow \partial_1\beta = a(\beta - \alpha), \tag{4.13}$$

Equations (4.6), (4.11), (4.12), and (4.13) show that three types of solutions are possible:

Case I: $\omega \neq 0 \Rightarrow \gamma^1_{32} = \beta = \alpha = 0$.

Case II: $\omega = \gamma^1_{32} = 0$.

Case III: $\omega = 0$, $\gamma^1_{32} \neq 0 \Rightarrow a = 0 = \partial_1\alpha = \partial_1\beta = \partial_1\gamma^1_{32}$.

[The commutator (4.5) applied to $\alpha \Rightarrow \partial_1\alpha = 0$ in *Case III.*]

Theorem 4.2: If dust ($\rho \neq 0$) is locally rotationally symmetric in an open set U with $\omega \neq 0$ in U, then $\Theta = \sigma = 0$ in U, and u^i is a Killing vector.

This is similar to a result of Gödel's,[3] proved under the assumption of homogeneous spatial 3-sections.

The coordinates are now specialized further in each of these cases, using the freedom (4.8), and then the field equations are examined in each case.

Case I: a, ω and γ^3_{23} are the only nonzero γ^a_{bc}. Then

$$\begin{pmatrix}0\\031\end{pmatrix}, \begin{pmatrix}2\\021\end{pmatrix}, \begin{pmatrix}3\\032\end{pmatrix} \Rightarrow \partial_0 a = \partial_0\omega = \partial_0\gamma^3_{23} = 0.$$

Surfaces $x^1 = \text{const}$ can be chosen (as $\gamma^1_{32} = 0 \Leftrightarrow \omega$ is hypersurface orthogonal) orthogonal to e_1, leaving

x^1 free by $x^{1'} = x^{1'}(x^1)$; then

$$e_3^1 = 0. \tag{4.14a}$$

Now

$$\gamma_{01}^1 = \gamma_{21}^1 = \gamma_{31}^1 = 0 \Rightarrow e_1^1 = e_1^1(x^1);$$

the remaining freedom of x^1 can be used to set

$$e_1^1 = 1. \tag{4.14b}$$

$$\gamma_{02}^2 = \gamma_{32}^2 = \partial_2(\gamma_{12}^2) = 0 \Rightarrow e_2^2 = B(x^1)f_2^2(x^2);$$

the freedom of x^2 can be used to set $f_2^2 = 1$, so

$$e_2^2 = B(x^1). \tag{4.14c}$$

$$\gamma_{03}^3 = 0, \quad \gamma_{12}^2 = \gamma_{13}^3,$$

$$\partial_3\gamma_{23}^3 = 0 \Rightarrow e_3^3 = B(x^1)f(x^2)f_3^3(x^3),$$

so suitable choice of $x^3 \Rightarrow$

$$e_3^3 = B(x^1)f(x^2). \tag{4.14d}$$

Finally,

$$\omega = \omega(x^1) \Rightarrow \omega = cB^2(x^1),$$

$$\frac{\partial y(x^2, x^3)}{\partial x^2} = -\frac{2c}{f(x^2)}, \tag{4.14e}$$

where c is a constant.

Case II: α, β, a, and γ_{23}^3 are nonzero in this case. As $\omega = \gamma_{32}^1 = 0$, x^0 and x^1 can be chosen so that

$$y = e_3^1 = 0, \tag{4.15a}$$

leaving x^1 free by $x^{1'} = x^{1'}(x^1)$; x^0 determined up to an additive constant.

$$\partial_2\gamma_{01}^1 = \partial_3\gamma_{01}^1 = \gamma_{21}^1 = \gamma_{31}^1 = 0 \Rightarrow e_1^1 = A(x^0, x^1). \tag{4.15b}$$

$$\partial_2\gamma_{02}^2 = \partial_3\gamma_{02}^2 = \gamma_{32}^2 = \partial_2\gamma_{12}^2 = 0 \Rightarrow$$
$$e_2^2 = B(x^0, x^1)f_2^2(x^2),$$

and we choose x^2 so that

$$e_2^2 = B(x^0, x^1). \tag{4.15c}$$

$$\gamma_{02}^2 = \gamma_{03}^3, \quad \gamma_{12}^2 = \gamma_{13}^3 \Rightarrow e_3^3 = B(x^0, x^1)f(x^2)f_3^3(x^3);$$

the freedom of x^3 is used to set

$$e_3^3 = B(x^0, x^1)f(x^2). \tag{4.15d}$$

Case III: α, β, γ_{32}^1, and γ_{23}^3 are nonzero.

$$\omega = 0 \Rightarrow \text{we can choose } y = 0, \tag{4.16a}$$

which determines x^0 (up to a constant); then $\alpha = \alpha(x^0)$, $\beta = \beta(x^0)$, $\gamma_{32}^1 = \gamma_{32}^1(x^0)$.

$$\gamma_{01}^1 = \gamma_{01}^1(x^0), \quad \gamma_{21}^1 = 0 \Rightarrow e_1^1 = A(x^0)f_1^1(x^1, x^3).$$

The freedom of x^1 can be used to set $f_1^1 = 1$; then

$$e_1^1 = A(x^0). \tag{4.16b}$$

$$\gamma_{02}^2 = \gamma_{02}^2(x^1),$$

$$\gamma_{12}^2 = \gamma_{32}^2 = 0 \Rightarrow e_2^2 = B(x^0)f_2^2(x^2);$$

we choose x^2 so that

$$e_2^2 = B(x^0). \tag{4.16c}$$

$$\gamma_{02}^2 = \gamma_{03}^3, \quad \gamma_{13}^3 = \partial_3\gamma_{23}^3 = 0 \Rightarrow e_3^3 = B(x^0)f(x^2)f_3^3(x^3);$$

we choose x^3 to set

$$e_3^3 = B(x^0)f(x^2). \tag{4.16d}$$

$$\gamma_{03}^1 = \gamma_{31}^1 = 0,$$

$$\gamma_{32}^1 = \gamma_{32}^1(x^0) \Rightarrow e_3^1 = B(x^0)f(x^2)h(x^2, x^3), \tag{4.16e}$$

where

$$\gamma_{32}^1 = 2C\frac{B^2(x^0)}{A(x^0)}, \quad \frac{\partial h(x^2, x^3)}{\partial x^2} = -2\frac{C}{f(x^2)} \tag{4.16f}$$

(C is a constant).

In all three cases, we obtain the coordinate tetrad relations in the form

$$e_0 = \partial/\partial x^0, \quad e_1 = A(x^0, x^1)(\partial/\partial x^1),$$
$$e_2 = B(x^0, x^1)(\partial/\partial x^2),$$
$$e_3 = B(x^0, x^1)f(x^2)[y(x^2, x^3)(\partial/\partial x^0) + h(x^2, x^3)(\partial/\partial x^1) + \partial/\partial x^3]. \tag{4.17}$$

It is possible now to determine f, y, and h for all cases.

Equation (4.17) $\Rightarrow$

$$\gamma_{23}^3 = B(x^0, x^1)\partial/\partial x^2[\log f(x^2)]. \tag{4.18}$$

Since (22), (33) show [on using (4.4a, b)] that the left-hand side is only a function of x^0, x^1 in all cases, it follows from (4.18) that

$$\partial_2\gamma_{23}^3 - (\gamma_{23}^3)^2 = B^2(x^0, x^1)K, \tag{4.19}$$

where $K = \text{const}$,

$$[d^2t/d(x^2)^2] + Kt = 0, \quad t(x^2) := 1/f(x^2). \tag{4.20}$$

The value of K determines $t(x^2)$, and so $f(x^2)$; however, there is considerable freedom of choice as to the exact form of $t(x^2)$. Part of this freedom is the final freedom of choice of the tetrad (γ_{23}^3 may be chosen arbitrarily at one point) and part corresponds to the different ways of fitting together the coordinates and tetrad. The biggest dichotomy here is whether at the origin the coordinates are regular [e.g., (3.17) holds] or singular [e.g., $f = 0$ or $t = 0$]. Having chosen t, then (4.14e), (4.16f) determine y, h from the constants c, C.

Typical choices[31] are:

Regular coordinates:

$$K>0: \quad t=\cos(K^{\frac{1}{2}}x^2), \quad y=\frac{-2c}{K^{\frac{1}{2}}}\sin(K^{\frac{1}{2}}x^2),$$

$$h=\frac{-2C}{K^{\frac{1}{2}}}\sin(K^{\frac{1}{2}}x^2),$$

$$K=0: \quad t=1, \quad y=-2cx^2, \quad h=-2Cx^2,$$

$$K<0: t=\cosh[(-K)^{\frac{1}{2}}x^2], y=\frac{-2c}{(-K)^{\frac{1}{2}}}\sinh[(-K)^{\frac{1}{2}}x^2], \tag{4.21a}$$

$$h=\frac{-2C}{(-K)^{\frac{1}{2}}}\sinh[(-K)^{\frac{1}{2}}x^2],$$

or

$$t=\exp[(-K)^{\frac{1}{2}}x^2], \quad y=\frac{-2c}{(-K)^{\frac{1}{2}}}\exp[(-K)^{\frac{1}{2}}x^2],$$

$$h=\frac{-2C}{(-K)^{\frac{1}{2}}}\exp[(-K)^{\frac{1}{2}}x^2].$$

Singular coordinates:

$$K>0: \quad t=\sin(K^{\frac{1}{2}}x^2), \quad y=\frac{2c}{K^{\frac{1}{2}}}\cos(K^{\frac{1}{2}}x^2),$$

$$h=\frac{2C}{K^{\frac{1}{2}}}\cos(K^{\frac{1}{2}}x^2), \tag{4.21b}$$

$$K=0: \quad t=x^2, \quad y=-c(x^2)^2, \quad h=-C(x^2)^2,$$

$$K<0: \quad t=\sinh[(-K)^{\frac{1}{2}}x^2],$$

$$y=\frac{-2c}{(-K)^{\frac{1}{2}}}\cosh[(-K)^{\frac{1}{2}}x^2], h=\frac{-2C}{(-K)^{\frac{1}{2}}}\cosh[(-K)^{\frac{1}{2}}x^2].$$

Theorem 4.3: If dust in an open set U is locally rotationally symmetric, a coordinate system may be found in an open neighborhood of any point in U for which the metric form is

$$\begin{aligned}Q=&-(dx^0)^2+X^2(x^0,x^1)(dx^1)^2+Y^2(x^0,x^1)\\&\times[(dx^2)^2+t^2(x^2)(dx^3)^2]+y(x^2)\\&\times[2\,dx^0-y(x^2)\,dx^3]\,dx^3-X^2(x^0,x^1)h(x^2)\\&\times[2\,dx^1-h(x^2)\,dx^3]\,dx^3.\end{aligned} \tag{4.22a}$$

The functions t, y, and h are determined by the constants K, c, and C, respectively, typical choices being given in (4.21). Coordinates may be chosen as follows in the three types of solution that occur: *Case I*, $X=1$, $Y=Y(x^1)$, $C=0$; *Case II*, $c=C=0$; *Case III*, $X=X(x^0)$, $Y=Y(x^0)$, $c=0$.

[31] Somewhat different coordinates could have been found by very similar procedures if we had set γ^3_{22}, $\partial_2\gamma^3_{12}$ zero instead of γ^3_{12}, $\partial_2\gamma^3_{22}$.

Proof: If either ω or σ is nonzero, then the coordinate and tetrad relations (4.17) show [via (A2)] that the metric has the form (4.22), on using the definitions

$$t(x^2):=[f(x^2)]^{-1}, \quad X(x^0,x^1):=[A(x^0,x^1)]^{-1},$$
$$Y(x^0,x^1):=[B(x^0,x^1)]^{-1}. \tag{4.22b}$$

If $\omega=\sigma=0$, we have a local Friedmann world model and coordinates can be chosen as in *Case II* [see (4.36)].

We now examine the field equations in each of the three cases.

Case I: Using the coordinate and tetrad (4.17), the Jacobi identities are all satisfied. The remaining field equations are

$$\tfrac{1}{2}\rho=\Lambda+2\omega^2, \tag{4.23a}$$

$$\partial_1 a=a^2+\omega^2+\Lambda, \tag{4.23b}$$

$$0=\Lambda-\partial_2\gamma^3_{23}+(\gamma^3_{23})^2+a^2-\omega^2, \tag{4.23c}$$

where $a=-\partial/\partial x^1[\log Y(x^1)]$. Equations (4.23b, c) are differential equations for B, while (4.23a) defines ρ. Two sets of solutions arise:

Case Ia: This is the exceptional case in which $a=0\Leftrightarrow\partial Y/\partial x^1=0$; then we choose $Y=1$. This is the only spatially homogeneous solution of type *II*, and is *Gödel's solution.*[9] The field equations are

$$\tfrac{1}{2}\rho=\omega^2=-\Lambda=c^2=-\tfrac{1}{2}K>0 \tag{4.24}$$

relating the constants Λ, c, and K to ρ and ω (also constants now). The space-time is invariant under a multiply-transitive G_5 of motions.

Case Ib: is the case $a\neq 0\Leftrightarrow\partial Y/\partial x^1\neq 0$. Then (4.23c) is a first integral of (4.23b). We can solve (4.23c) to get:

$$\Lambda=0, \quad K=0: Y^2=1+2cx^1, \tag{4.25a}$$

$$\Lambda=0, \quad K\neq 0: Y^2=\frac{1}{|K|}\{[(c^2+K)^{\frac{1}{2}}+Kx^1]^2-c^2\}, \tag{4.25b}$$

$\Lambda>0$:

$$Y^2=\frac{1}{2\Lambda}[(K^2+4c^2\Lambda)^{\frac{1}{2}}\sin(2(\Lambda)^{\frac{1}{2}}x^1)+K], \tag{4.25c}$$

$\Lambda<0$:

$$Y^2=\tfrac{1}{2}\exp(-(-\Lambda)^{\frac{1}{2}}x^1)\times\left\{\left[\exp(2(-\Lambda)^{\frac{1}{2}}x^1)+\frac{K}{2\Lambda}\right]^2+\frac{c^2}{\Lambda}\right\}. \tag{4.25d}$$

ω and ρ [from (4.14e), (4.23a)] are both infinite if Y is zero. In (4.25a, b) ω and $\rho\to\infty$ for a finite value

of x^1. In (4.25c), having chosen a possible range for x^1, ω and ρ become infinite at both ends of the range of x^1. In (4.25d), ρ and ω are finite for all x^1 if $K^2/4(-\Lambda) > c^2$. However, as $x^1 \to \pm\infty$, $Y^2 \to \infty$. So [from (4.23a)] there exist finite values of x^1 for which $\rho = 0$; for values of x^1 outside these bounds, $\rho < 0$ (as $\Lambda < 0$, $|\Lambda| > 2\omega^2$ for these values). Thus, of all solutions with $\omega \neq 0$, only Gödel's solution has finite positive density of matter for all coordinate values. It is also the only solution in which ρ is constant along the vortex lines.

Each of the "space-times" in *Case Ib* is invariant under a G_4 of motions multiply transitive in the 3-surfaces $x^1 = \text{const}$.

Case II: This is a generalization of the cases examined by Bondi[8] and Kantowski and Sachs.[16] With coordinates (4.17), (4.22)

$$\alpha = \partial/\partial x^0 \log X, \quad \beta = \partial/\partial x^0 \log Y,$$

and
$$a = (-1/X)\partial/\partial x^1 (\log Y). \tag{4.26}$$

The conservation equations and $\begin{pmatrix} 2 \\ 021 \end{pmatrix}$ are

$$\partial_0 a = -a\beta, \quad \partial_0 \rho = -\rho(\alpha + 2\beta), \tag{4.27}$$

which may be integrated, using (4.4a, b) to give

$$a = Y^{-1}(x^0, x^1)a_0(x^1), \quad \rho = X^{-1}(x^0, x^1)\, Y^{-2}(x^0, x^1)\rho_0(x^1). \tag{4.28}$$

The field equations still to be satisfied are[32]

$$2\partial_0\beta = \Lambda - \partial_2\gamma^3_{23} + (\gamma^3_{23})^2 + a^2 - 3\beta^2, \tag{4.29a}$$

$$2\partial_1 a = \Lambda + \rho - \partial_2\gamma^3_{23} + (\gamma^3_{23})^2 - 2\beta\alpha - \beta^2 + 3a^2, \tag{4.29b}$$

$$\partial_0\alpha + 2\partial_0\beta = \Lambda - \tfrac{1}{2}\rho - 2\beta^2 - \alpha^2. \tag{4.29c}$$

Integration now splits into two cases, depending on whether $a = 0$ or $a \neq 0$.

Case IIa: $a \neq 0 \Leftrightarrow (\partial Y/\partial x^1) \neq 0$. From (4.26), (4.28),

$$X(x^0, x^1) = -\frac{1}{a_0(x^1)}\frac{\partial Y(x^0, x^1)}{\partial x^1}. \tag{4.30}$$

Case IIai: This exceptional case occurs if $\partial Y/\partial x^0 = 0 \Leftrightarrow \beta = 0$. Then (4.30) $\Rightarrow X = X(x^1) \Rightarrow \alpha = 0$, so we have a Friedmann universe with no expansion; X can be renormalized to 1. Now (4.29a, c) can be solved to give

$$\Lambda = \tfrac{1}{2}\rho, \quad X = 1, \quad Y = \cos[(\tfrac{1}{2}\rho)^{\frac{1}{2}}x^1], \quad K = \tfrac{1}{2}\rho > 0, \tag{4.31}$$

and (4.29b) is satisfied. This is the *Einstein static universe*; space-time is invariant under a multiply transitive G_7 of motions.

Case IIaii: In this case $\partial B/\partial x^0 \neq 0 \Leftrightarrow \beta \neq 0$. Equation (4.29c) is identically satisfied as a consequence of the other equations. Bondi[8] integrated the case in which $K = 1$. Following Bondi, (4.29a) may be integrated, giving

$$\dot{Y}^2(x^0, x^1) = a_0^2(x^1) - K + \tfrac{1}{3}\Lambda Y^2 - \frac{k(x^1)}{Y}; \tag{4.32a}$$

then

$$(4.29b) \Rightarrow \rho_0(x^1) = -\frac{1}{a_0(x^1)}\frac{\partial k(x^1)}{\partial x^1}, \tag{4.32b}$$

a_0 and k are arbitrary functions of x^1 such that $\rho_0 > 0$. On an initial surface $x^0 = c^0 = \text{const}$, we may specify the expansions α, β and the density ρ as functions of x^1, and then determine $Y(c^0, x^1)$, $\dot{Y}(c^0, x^1)$, and $k(x^1)$ from the values of Y and k at a point on the surface. The values α, β, ρ are arbitrary except for two conditions: (i) $\alpha(c^0, x^1) = \beta(c^0, x^1) \Leftrightarrow \beta(c^0, x^1) = \text{const}$, by (4.13); this is the case of a Friedmann solution, see below; (ii) $a_0(x^1)$ is found from (4.32a); the values chosen must make a_0 real.

With such values, (4.32a) determines the time-development of Y—this is Friedmann's equation, cf. (4.34a)–(4.30) determines X, and (4.28), (4.32b) determine a, ρ.

Generally, the space-time is invariant under a group of motions G_3 multiply transitive on the 2-surfaces $x^0 = \text{const}$, $x^1 = \text{const}$. Farnsworth[33] has found a dust-filled space-time invariant under a G_4 multiply transitive on 3-surfaces *not* orthogonal to u^a, in which $\omega = 0$, $\sigma \neq 0$, $\Theta \neq 0$; so that although the solution may have homogeneous space sections, it does not appear homogeneous to observers moving with 4-velocity u^a. This solution is a solution of type *IIaii*, with greater symmetry than the general case.

To obtain this solution, *Case IIaiii*, we assume there is a Killing vector $Z = Z^i\partial/\partial x^i$, where at least one of Z^0, Z^1 is not zero. We can then choose x^1 so that $Z^0 = \lambda^0 = \text{const}$, $Z^1 = \lambda^1 = \text{const}$. If $\lambda^0 = 0$, we have the Friedmann world models, and if $\lambda^1 = 0$, we have *Case IIai*, so we now assume $\lambda^0\lambda^1 \neq 0$. We find that such a solution is possible, if $Y = \exp(\gamma x^0)Z(s)$, $s := \lambda^1 x^0 - \lambda^0 x^1$; $a_0(x^1) = \lambda \exp[(\lambda^0/\lambda^1)x^1]$, $k(x^1) = -(\lambda\mu\lambda^1/3\gamma\lambda^0)\exp[3\gamma(\lambda^0/\lambda^1)]$; γ, λ, μ are nonzero constants; and $K = 0$. The function $Z(s)$ is then determined by (4.32a). Note that by the arbitrariness in the choice of λ^0, λ^1 the surface of homogeneity can

[32] Bondi gives these equations in Ref. 8 [see his Appendix, Eq. (5)] for the case $K = 1$. The full set (4.29) are obtained from his equations on replacing $(1 + Y^2)$ by $(K + Y^2)$.

[33] D. Farnsworth (private communication).

initially be chosen spacelike, null or timelike; the development of these surfaces is determined by α.

It is interesting to see how the solutions *IIaii* relate to the *Friedmann (Robertson–Walker) solutions*. With the extra condition $\sigma = 0 \Leftrightarrow \alpha = \beta$, (4.30) shows that x^1 can be chosen to set

$$Y = X(x^0)Z(x^1), \quad X = X(x^0), \tag{4.33a}$$

where

$$a_0 = -\partial Z/\partial x^1, \quad \rho_0 = (\partial k/\partial x^1)(\partial Z/\partial x^1)^{-1}. \tag{4.33b}$$

The metric then has the form

$$Q = -(dx^0)^2 + X^2(x^0) \times \{(dx^1)^2 + Z^2(x^1)[(dx^2)^2 + t^2(x^2)(dx^3)^2]\}.$$

Substituting (4.33a) into (4.32a) and separating out equations in x^0 and x^1, we find

$$\dot{X}^2 - \tfrac{1}{3}\Lambda X^2 - \tfrac{1}{3}(M/X) - E = 0, \tag{4.34a}$$

$$k(x^1) = \tfrac{1}{3}MZ^3, \quad (\partial Z/\partial x^1)^2 - K = EZ^2, \tag{4.34b}$$

where M, E are constants; (4.33b), (4.28) $\Rightarrow$ $\rho = M/X^3$, so $\rho = \rho(x^0)$, $M > 0$, X obeys *Friedmann's equation* (4.34a),[12,17,18] while Z is obtained from (4.34b). The general solutions for Z are:

$$E > 0: \quad Z = A_1 \exp(E^{\frac{1}{2}}x^1) + A_2 \exp(-E^{\frac{1}{2}}x^1), \quad K = -4A_1A_2E, \tag{4.35a}$$

$$E = 0: \quad Z = A_1x^1 + A_2 \quad (A_1 \neq 0), \quad K = (A_1)^2, \tag{4.35b}$$

$$E < 0: \quad Z = A_1 \sin[(-E)^{\frac{1}{2}}x^1] + A_2 \cos[(-E)^{\frac{1}{2}}x^1], \quad K = -E\{(A_1)^2 + (A_2)^2\}, \tag{4.35c}$$

where A_1, A_2 are constants. The ratio A_1/A_2 can be altered by changing the origin of x^1, while $(A_1)^2 + (A_2)^2$ can be altered by changing the normalization of Z. We can choose A_1, A_2 to obtain solutions

$$E > 0: \quad Z = \sinh(E^{\frac{1}{2}}x^1), \quad K = E, \quad t = \sin(E^{\frac{1}{2}}x^2), \tag{4.36a}$$

or

$$Z = \cosh(E^{\frac{1}{2}}x^1), \quad K = -E, \quad t = \sinh(E^{\frac{1}{2}}x^2), \tag{4.36b}$$

$$E = 0: \quad Z = x^1 \quad K = 1, \quad t = \sin x^2, \tag{4.36c}$$

$$E < 0: \quad Z = \sin[(-E)^{\frac{1}{2}}x^1], \quad K = -E, \quad t = \sin[(-E)^{\frac{1}{2}}x^2], \tag{4.36d}$$

or

$$Z = \cos[(-E)^{\frac{1}{2}}x^1], \quad K = -E, \quad t = \sin[(-E)^{\frac{1}{2}}x^2], \tag{4.36e}$$

where the contrast between the pairs (4.36a, b) and (4.36d, e) is interesting. All Friedmann solutions are obtained, as we have solutions for $E > 0$, $E = 0$, and $E < 0$. Each of these solutions is invariant under a G_6 of motions multiply transitive in the surfaces $x^0 = \text{const}$.

Case IIb: $a = 0 \Leftrightarrow B = B(x^0)$. Now $\beta = \beta(x^0)$; however, α and ρ may still be functions of x^1 as well as x^0. Equation (4.29a) may be integrated as before; (4.32a) holds with $a_0 = 0$ and $k(x^1) = k$, a constant. Equation (4.29b) can be integrated, and then (4.29c) is identically satisfied.

Case IIbi: $K \neq 0$. (We consider here only $\Lambda = 0$.) The equations have been integrated by Kantowski and Sachs[16] for the case $X = X(x^0)$. When $X = (x^0, x^1)$, the same integrations may be used; the only difference[34] now is that the constant occurring in the integration of X (b in Ref. 16) is now an arbitrary function of x^1. The general solution admits a multiply transitive G_3 acting on the 2-surfaces $x^0 = \text{const}$, $x^1 = \text{const}$; when $X = X(x^0)$, we have the cases investigated by Kantowski and Sachs, admitting a group G_4 multiply transitive on the 3-surfaces $x^0 = \text{const}$.

Case IIbii: When $K = 0$ (we consider only $\Lambda = 0$) there occurs an exceptional case

$$X = Y = (x^0)^{\frac{2}{3}}. \tag{4.37}$$

This is the Einstein–de Sitter world model invariant under a G_6 of motions multiply transitive on the surfaces $x^0 = \text{const}$.

Case IIbiii: The general case $K = 0$, $\Lambda = 0$ has solution

$$X = [1 + C(x^1)](x^0)^{-\frac{1}{3}}, \quad Y = (x^0)^{\frac{2}{3}}, \tag{4.38}$$

where $C(x^1)$ is an arbitrary function of x^1; in this case,

$$\beta = \frac{2}{3x^0}, \quad \alpha = \frac{1}{3x^0}\frac{2C(x^1)x^0 - 1}{1 + C(x^1)x^0}.$$

In general, this is invariant under a group G_3 multiply transitive on the 2-surfaces $x^0 = c^0$, $x^1 = c^1$. If $C(x^1)$ is constant, the space is invariant under a G_4 multiply transitive in the 3-surfaces orthogonal to u. This is a special case of the Bianchi world model studied by Heckmann and Schücking[11,12] and by Raychaudhuri[11] invariant under a simply transitive subgroup G_3 of type I.

Considering *Case IIb* with Λ possibly nonzero, there are no solutions with $\alpha = 0$ or $\beta = 0$; if $\sigma = 0$, we

[34] This is clear from the equations (see Ref. 32): if $Y' = 0$, then neither X' nor X'' occur in the field equations, which are then Eqs. (4.39) with $C = 0$.

obtain Friedmann models obeying (4.34a) with $X(x^0) = Y(x^0)$, $K = E = 0$. The only solutions of type II with $\Theta = 0$ are the solutions of *Case IIai*. If we consider solutions of *Case IIb* with larger groups of motions, three cases arise. We can choose the coordinate x^1 so that there is a Killing vector $Z = Z^i\partial/\partial x^i$, $Z^0 = \lambda^0, Z^1 = \lambda^1$; λ^i constant, not both zero. If $\lambda^1 = 0$, there is no solution. If $\lambda^0 = 0$, we obtain the Kantowski–Sachs type models if $K \neq 0$, and Bianchi I models if $K = 0$ (in both cases, with Λ possibly nonzero). Finally, if $\lambda^0\lambda^1 \neq 0$, we get the following.

Case IIbiv: a solution with $\beta = \text{const}$, $\Lambda = 3\beta^2$, $K = 0$, $Y = \exp(\beta x^0)$, $X = K_1 \exp(\beta x^0 + \gamma x^1) + K_2 \exp\{-(\beta x^0 + \gamma x^1)\}$, $\gamma = -\beta\lambda^0/\lambda^1$, where γ, K_i are nonzero constants, $K_1K_2 > 0$. Then $\rho = 2\beta(\alpha - \beta)$, which becomes infinite for finite values of x^1. The ratio β/γ determines (at a point) the nature of the surface of transitivity of the group G_4; this surface may be timelike, spacelike, or null.

Case III: The space–time has homogeneous space-sections; it is invariant under a G_4 multiply transitive in the surface $x^0 = \text{const}$. There exists at least one simply transitive subgroup G_3 in these surfaces, and so these spaces are particular models of the type investigated by Heckmann and Schücking.[11,12] The remaining equations are

$$\frac{2\dot{Y}}{Y} + \frac{\dot{Y}^2 + K}{Y^2} - \frac{3C^2X^2}{Y^4} = \Lambda, \tag{4.39a}$$

$$\frac{2\dot{X}\dot{Y}}{XY} + \frac{\dot{Y}^2 + K}{Y^2} - \frac{C^2X^2}{Y^4} = \Lambda + \rho, \tag{4.39b}$$

$$\frac{\ddot{X}}{X} + \frac{\dot{X}\dot{Y}}{XY} + \frac{\ddot{Y}}{Y} + C^2\frac{X^2}{Y^4} = \Lambda, \tag{4.39c}$$

where $\rho = \rho_0/XY^2$; ρ_0, K, and C are constants; and $X = X(x^0)$, $Y = Y(x^0)$. If C were zero, the solution would be of type *II*, these then being the equations integrated by Kantowski and Sachs ($K \neq 0$) and, e.g., by Heckmann and Schucking ($K = 0$); analytic solutions of these equations for *Case III* (i.e., $C \neq 0$) have not yet been found. The case $K > 0$ has been investigated extensively by Behr[14]; the G_3 is of type IX.The case $K = 0$ has been examined by Farnsworth[33]; the G_3 is of type *II*. If $K < 0$, the G_3 is of type VIII.

If $\sigma = 0$ we obtain $X(x^0) = Y(x^0)$ by suitable choice of coordinates. These Friedmann models obey (4.34a) with $E = -C^2$ and $K = 4C^2$, and are given in a new coordinate system. If $\Theta = 0$, then $\sigma = 0$ and we have solution *IIai* in new coordinates. The coordinates can be chosen so that $X = Y = 1$; the field equations are $\rho = 2\Lambda = 2C^2 = \frac{1}{2}K$. This is the only solution *III* in which either α or β is zero.

On examining these classes of solutions, we see that in each case the space–time obeying (A_3) is invariant under a multiply transitive group of motions: so $(A_3) \Rightarrow (C)$.

Theorem 4.4: For dust, the symmetry conditions (A_3), (B), (C) are equivalent.

The set of dust solutions obeying these conditions are the solutions of *Cases I, II, III* above. In fact, for the general solutions the vectors e_0, e_1 are uniquely defined. Thus a Killing vector Z (since it preserves the tetrad orthogonality conditions) must obey

$$[Z, e_0] = 0, \quad [Z, e_1] = 0, \quad [Z, e_2] = b(x^i)e_3,$$
$$[Z, e_3] = -b(x^i)e_2, \tag{4.40}$$

which are equivalent to Killing's equations. The solution of these equations for the tetrad (4.17) with the coordinates chosen according to (4.21b) is given in the Appendix, Eq. (A7). The multiply transitive groups contain subgroups simply transitive on their surfaces of transitivity in all cases except the general solutions of *Cases IIai* and *IIbi*, $K > 0$, and the special solutions integrated by Kantowski and Sachs, of *IIbi*, $K > 0$.

5. SHEAR-FREE DUST

Shear-free motion of dust, which we study in this section, is of interest for several reasons. First, a theorem of Gödel states: if a dust-filled world model has homogeneous space sections and $\Theta \neq 0$, then $\sigma = 0 \Rightarrow \omega = 0$. (A proof of this theorem has been given by Schücking[10].) One would like to know: under what more general conditions does such a result hold?

Second, the equations governing the behavior of a timelike congruence of geodesics are very similar to the equations of a congruence of null geodesics[35]: in each case, the vorticity and expansion propagation equations contain, apart from terms in the expansion, shear and vorticity, only terms in the Ricci tensor; but the shear propagation equations contain components of the Weyl tensor. Roughly speaking, "news" comes into the congruence via the shear induced by the Weyl tensor; the shear drives the other equations.

The Goldberg–Sachs theorem[36] shows that the set of empty-space solutions containing shear-free congruences of null geodesics is severely restricted: the Weyl tensor has to be algebraically special. To what extent does the existence of a shear-free congruence

[35] Compare Ref. 5 and J. Ehlers and R. Sachs, Abhandl. Math. Naturw. Kl., Mainz Akad. Wiss. Lit. **1**, (1961). In the case of a timelike congruence, equations relating δ_{ab} to E_{ab} may be obtained by the methods used in Ref. 5.

[36] J. Goldberg and R. Sachs, Acta. Phys. Polon. **22**, 13 (1962); also Refs. 19, 25.

of world-lines of dust imply restrictions on the space-time? In each case, no "news" comes into the congruence.

Suppose that dust moves without shear in some open set W within which coordinates and a tetrad may be chosen as in Theorem 3.1. Then the γ^a_{bc} are restricted by

$$\Theta_1 = \Theta_2 = \Theta_3 =: \alpha; \quad \sigma_{12} = \sigma_{23} = \sigma_{31} = 0. \tag{5.1}$$

Using these conditions, the Jacobi identities and field equations give a set of time-derivative equations:

$$\binom{0}{023} \Rightarrow \partial_0\omega = 2\omega\alpha, \tag{5.2a}$$

$$\binom{1}{012} \Rightarrow \partial_0\gamma^1_{21} = -\alpha\gamma^1_{21} + \tfrac{1}{2}\partial_3\omega - \tfrac{1}{2}\omega\gamma^1_{31}, \tag{5.2b}$$

$$\binom{2}{023} \Rightarrow \partial_0\gamma^2_{32} = -\alpha\gamma^2_{32} - \tfrac{1}{2}\partial_2\omega + \tfrac{1}{2}\omega\gamma^1_{21}, \tag{5.2c}$$

$$\binom{3}{031} \Rightarrow \partial_0\gamma^3_{13} = -\alpha\gamma^3_{13} + \tfrac{1}{2}\omega\gamma^1_{32}, \tag{5.2d}$$

$$\binom{2}{021} \Rightarrow \partial_0\gamma^2_{12} = -\alpha\gamma^2_{12} + \tfrac{1}{2}\omega\gamma^1_{32}, \tag{5.2e}$$

$$\binom{3}{032} \Rightarrow \partial_0\gamma^3_{23} = -\alpha\gamma^3_{23} + \tfrac{1}{2}\partial_3\omega - \tfrac{1}{2}\omega\gamma^1_{31}, \tag{5.2f}$$

$$\binom{1}{013} \Rightarrow \partial_0\gamma^1_{31} = -\alpha\gamma^1_{31} - \tfrac{1}{2}\partial_2\omega + \tfrac{1}{2}\omega\gamma^1_{21}, \tag{5.2g}$$

$$\binom{1}{023}, \binom{2}{031} \Rightarrow \partial_0\gamma^1_{32} = -\alpha\gamma^1_{32}, \quad \partial_0\gamma^2_{13} = -\alpha\gamma^2_{13}, \tag{5.2h}$$

$$(00) \Rightarrow \partial_0\alpha = -\alpha^2 + \tfrac{2}{3}\omega^2 + \frac{\Lambda}{3} - \tfrac{1}{6}\rho, \tag{5.2i}$$

$$(2.7) \Rightarrow \partial_0\rho = -3\alpha\rho. \tag{5.2j}$$

Further, we obtain equations which are constraint equations in a surface $x^0 = \text{const}$. First,

$$(01) \Rightarrow \partial_1\alpha = -\tfrac{1}{2}\omega\gamma^1_{32}, \tag{5.3a}$$

$$(02) \Rightarrow \partial_2\alpha = -\tfrac{1}{2}\partial_3\omega + \tfrac{1}{2}\omega\gamma^1_{31}, \tag{5.3b}$$

$$(03) \Rightarrow \partial_3\alpha = -\tfrac{1}{2}\partial_2\omega - \tfrac{1}{2}\omega\gamma^1_{21}, \tag{5.3c}$$

$$\binom{0}{012} \Rightarrow \partial_1\omega = \omega(\gamma^2_{12} + \gamma^3_{13}). \tag{5.3d}$$

Second, (11), (22), (33) each gives a value of $\partial_0\alpha$ which may be used with (00) to obtain the following constraint equations:

$$\Phi := 2\alpha^2 + \tfrac{2}{3}\omega^2 - \tfrac{2}{3}\Lambda - \tfrac{2}{3}\rho, \tag{5.4a}$$

$$\begin{aligned}\Phi = {} & -\partial_1(\gamma^2_{12}) - \partial_1(\gamma^3_{13}) - \partial_2(\gamma^1_{21}) - \partial_3(\gamma^1_{31}) \\ & + (\gamma^2_{12})^2 + (\gamma^3_{13})^2 + (\gamma^1_{21})^2 + (\gamma^1_{31})^2 + \gamma^1_{21}\gamma^3_{23} \\ & + \gamma^1_{31}\gamma^2_{32} - \tfrac{1}{2}((\gamma^1_{32})^2 - (\gamma^2_{13})^2),\end{aligned} \tag{5.4b}$$

$$\begin{aligned}\Phi = {} & -2\partial_2(\gamma^3_{23}) - 2\partial_3(\gamma^2_{32}) + 2(\gamma^3_{23})^2 + 2(\gamma^2_{32})^2 \\ & - 4\omega^2 + (\gamma^1_{32} - \gamma^2_{13})(\tfrac{3}{2}\gamma^1_{32} + \tfrac{1}{2}\gamma^2_{13}) + 2\gamma^2_{12}\gamma^3_{13},\end{aligned} \tag{5.4c}$$

$$\begin{aligned}0 = {} & \partial_2(\gamma^1_{21}) - \partial_3(\gamma^1_{31}) + \partial_1(\gamma^2_{12}) - \partial_1(\gamma^3_{13}) - (\gamma^1_{21})^2 \\ & + (\gamma^1_{31})^2 - (\gamma^2_{12})^2 + (\gamma^3_{13})^2 - \gamma^2_{32}\gamma^1_{31} + \gamma^3_{23}\gamma^1_{21} \\ & + \gamma^2_{13}(\gamma^2_{13} - \gamma^1_{32}).\end{aligned} \tag{5.4d}$$

There remain 6 further constraint equations.

By Eq. (5.1),

$$[e_0, e_\nu] = -\alpha e_\nu. \tag{5.5}$$

Using these commutation relations, we can find the time derivatives of $\partial_2\omega$ and $\partial_3\omega$. Thus,

$$\begin{aligned}\partial_0(\partial_2\omega) &= \partial_2(\partial_0\omega) - \alpha\partial_2\omega \\ &= -3\alpha\partial_2\omega + \omega\partial_3\omega - \omega^2\gamma^1_{31}\end{aligned} \tag{5.6a}$$

on using (5.2a), (5.3b); similarly,

$$\partial_0(\partial_3\omega) = -3\alpha\partial_3\omega - \omega\partial_2\omega + \omega^2\gamma^1_{21}. \tag{5.6b}$$

From (5.2, 6) it is clear that, given the γ^a_{bc}, $\partial_2\omega$, $\partial_3\omega$, and ρ at one point on a world line, we can find all these quantities at any other point on the world line. Similarly, we can find the propagation along the world line of any derivatives $\partial_e \cdots \partial_g\gamma^a_{bc}$ of the γ^a_{bc}. This makes explicit the sense in which no "news" can come into the congruence if $\sigma = 0$.

The constraint equations must be preserved during the time development of the system. Is this condition automatically satisfied, or does it imply further constraints? Consider first Eq. (5.3a). Taking the time derivative of this equation, using methods similar to those used in deriving (5.6), we find

$$\partial_1\rho = 8\omega^2(\gamma^2_{12} + \gamma^3_{13}), \tag{5.7a}$$

which is a new constraint equation. Continuing in this way we obtain a series of further constraints on ω, α, ρ, γ^1_{32}, and $(\gamma^2_{12} + \gamma^3_{13})$. These constraints are

$$\omega[\gamma^1_{32}(\rho - \tfrac{16}{3}\omega^2) + \tfrac{16}{3}\alpha\omega(\gamma^2_{12} + \gamma^3_{13})] = 0, \tag{5.7b}$$

$$\omega^2[12\alpha\gamma^1_{32}\omega + (4\omega^2 + 2\Lambda - \rho)(\gamma^2_{12} + \gamma^3_{13})] = 0, \tag{5.7c}$$

$$\omega^2[\omega\gamma^1_{32}(\Lambda - \tfrac{2}{3}\omega^2) + \alpha(2\omega^2 + \Lambda)(\gamma^2_{12} + \gamma^3_{13})] = 0, \tag{5.7d}$$

$$\omega^2\alpha^2(\gamma^2_{12} + \gamma^3_{13})(\Lambda - 2\omega^2) = 0, \tag{5.7e}$$

$$\omega^3\alpha^2(8\alpha\omega(\gamma^2_{12} + \gamma^3_{13}) + \gamma^1_{32}(\Lambda - 2\omega^2)) = 0, \tag{5.7f}$$

where the earlier equations have been used in simplifying the later ones. Now suppose $\alpha\omega \neq 0$ at a point p. Then $\alpha\omega \neq 0$ on an open neighborhood U of p, ($U \subset W$): so (5.7e, f) $\Rightarrow$ in U,

$$\gamma^2_{12} + \gamma^3_{13} = 0 \tag{5.8a}$$

and (5.2d, e) $\Rightarrow$ in U,

$$\gamma^1_{32} = 0. \tag{5.8b}$$

Next, consider the pair of equations (5.3b, c). These are conserved if

$$\partial_2\rho = \tfrac{3}{2}\omega^2\gamma^1_{21} + \tfrac{1}{2}\omega\partial_2\omega, \tag{5.9a}$$

$$\partial_3\rho = \tfrac{3}{2}\omega^2\gamma^1_{31} + \tfrac{1}{2}\omega\partial_3\omega. \tag{5.9b}$$

As in the last case, we obtain now a series of algebraic constraints; (5.9a,b) are preserved if

$$(\partial_3\omega - \omega\gamma^1_{31})(6\rho - 29\omega^2) + 2\alpha\omega(3\omega\gamma^1_{21} + 13\partial_2\omega) = 0, \tag{5.9c}$$

$$(\partial_2\omega - \omega\gamma^1_{21})(6\rho - 29\omega^2) - 2\alpha\omega(3\omega\gamma^1_{31} + 13\partial_3\omega) = 0, \tag{5.9d}$$

and (5.9c, d) are preserved if

$$\alpha\omega^2(45\partial_3\omega - 61\omega\gamma^1_{31}) + 2\omega\left(\frac{2}{3}\omega^2 + \frac{\Lambda}{3} - \frac{1}{6}\rho\right) \times (3\omega\gamma^1_{21} + 13\partial_2\omega) = 0, \tag{5.9e}$$

$$\alpha\omega^2(45\partial_2\omega - 61\omega\gamma^1_{21}) - 2\omega\left(\frac{2}{3}\omega^2 + \frac{\Lambda}{3} - \frac{1}{6}\rho\right) \times (3\omega\gamma^1_{31} + 13\partial_3\omega) = 0. \tag{5.9f}$$

From the two pairs of equations (5.9c–f),

$$\alpha\omega \neq 0 \quad \text{at} \quad p \Rightarrow \gamma^1_{31}\partial_2\omega - \gamma^1_{21}\partial_3\omega = 0. \tag{5.10}$$

Finally, consider the equations (5.4a, b, c). We can calculate $\partial_0\Phi$ from each of these: (5.4a) $\Rightarrow$

$$\partial_0\Phi = -2\alpha\Phi + \tfrac{4}{3}\alpha\omega^2, \tag{5.11a}$$

(5.4c) $\Rightarrow$

$$\partial_0\Phi = -2\alpha\Phi + 6\alpha\omega^2 + \omega\partial_1\gamma^1_{23} + \partial_2\omega\gamma^1_{31} - \partial_3\omega\gamma^1_{21} + \omega\gamma^1_{32}(\gamma^1_{21} + \gamma^1_{31}), \tag{5.11b}$$

while (5.4b) gives a value of $\partial_0\Phi$ consistent with these results. Combining the above results, $\alpha\omega \neq 0$ at $p \Rightarrow \alpha\omega \neq 0$ on an open neighborhood U of $p \Rightarrow$ (5.8), (5.10) hold in $U \Rightarrow \alpha\omega = 0$ in U by (5.11), a contradiction.

Theorem 5.1: In a dust-filled space–time, $\sigma = 0$ on any open set $U \Rightarrow \omega\Theta = 0$ on U.

The proof of this theorem does not depend on the condition $\rho \neq 0$; so this result holds also for a timelike congruence of geodesics in empty space–time.

Theorem 5.1 shows that two cases arise for shear-free dust. The first case is when $\sigma = 0$, $\omega = 0$, and the model is locally a Friedmann world model (see Sec. 4). The second case is when $\sigma = 0$, $\omega \neq 0 \Rightarrow \Theta = 0$, u is a Killing vector, and space–time is stationary. This second case has been examined by various authors; in particular, Ehlers[6] showed that all dust solutions of this type can be obtained by suitable conformal transformations from static empty-space solutions. In the present formalism, (5.1), (5.2) $\Rightarrow$

$$\partial_0\gamma^a_{bc} = 0, \quad \partial_0\rho = 0, \quad [e_0, e_\nu] = 0. \tag{5.12}$$

Also, (5.3a) $\Rightarrow$

$$\gamma^1_{32} = 0, \tag{5.13}$$

which shows that ω^i is hypersurface orthogonal (cf. Ref. 6). Therefore x^1 can be chosen so that

$$e^1_2 = e^1_3 = 0, \tag{5.14a}$$

while $\Theta = \sigma = 0 \Rightarrow$

$$e^i_a = e^i_a(x^\nu). \tag{5.14b}$$

The coordinate freedom which preserves these conditions (with the tetrad invariant) is

$$x^{0'} = x^0 + f(x^3), \quad x^{1'} = x^{1'}(x^1), \quad x^{2'} = x^{2'}(x^2, x^3), \quad x^{3'} = x^{3'}(x^3). \tag{5.14c}$$

The field equations and the Jacobi identities yet to be satisfied can be split into equations determining propagation along the vortex lines (i.e., involving ∂_1) and initial constraints in a 3-surface $x^1 = c^1$. Four of these constraint equations are

$$\partial_2\omega = \omega\gamma^1_{21}, \tag{5.15a}$$

$$\partial_3\omega = \omega\gamma^1_{31}, \tag{5.15b}$$

$$\rho = 4\omega^2 + 2\Lambda, \tag{5.15c}$$

$$\Phi = -2(\omega^2 + \Lambda). \tag{5.15d}$$

The initial constraints must be preserved under propagation in the e_1 direction. Applying this condition to (5.15), there result the propagation equations

$$\partial_1(\gamma^1_{31}) = \partial_3(\gamma^2_{12} + \gamma^3_{13}) + \gamma^2_{13}\gamma^1_{21} - \gamma^2_{12}\gamma^1_{31}, \tag{5.16a}$$

$$\partial_1(\gamma^1_{21}) = \partial_2(\gamma^2_{12} + \gamma^3_{13}) - \gamma^3_{13}\gamma^1_{21}, \tag{5.16b}$$

$$\partial_1\rho = 8\omega^2(\gamma^2_{12} + \gamma^3_{13}), \tag{5.16c}$$

$$\partial_1\Phi = -4\omega^2(\gamma^2_{12} + \gamma^3_{13}). \tag{5.16d}$$

Now, the e_1 propagation of all the γ^a_{bc}'s is known, all the remaining constraints are conserved under propagation along the vortex lines. We characterize some of the simplest solutions of these equations.

Case A: ρ is constant in the hypersurfaces orthogonal to $\omega^i \Leftrightarrow \partial_2\omega = 0 = \partial_3\omega$. Then (5.15) $\Rightarrow$

$$\gamma^1_{21} = \gamma^1_{31} = 0, \tag{5.17a}$$

so that ω^i is geodesic. Now, Eq. (23) is

$$\partial_1(\gamma^2_{31}) = 2\gamma^2_{31}\gamma^2_{12};$$

the tetrad freedom may be used to set $\gamma^2_{31} = 0$ on a surface $x^0 = c^0$, $x^1 = c^1$; and then

$$\gamma^2_{31} = 0 \tag{5.17b}$$

holds everywhere. The coordinate x^2 can then be chosen so that

$$e^2_3 = 0. \tag{5.17c}$$

Equations (12), (13), $\begin{pmatrix} 2 \\ 123 \end{pmatrix} \Rightarrow$

$$\partial_3\gamma^2_{12} = \gamma^2_{32}(\gamma^2_{12} - \gamma^3_{13}), \tag{5.18a}$$

$$\partial_2\gamma^3_{13} = \gamma^3_{23}(\gamma^3_{13} - \gamma^2_{12}) \tag{5.18b}$$

so the case

$$\gamma^2_{12} = \gamma^3_{13} =: a \Rightarrow \partial_2 a = \partial_3 a = 0$$

and all the conditions (4.3), (4.4) on the γ^a_{bc} are satisfied. Thus, either the tetrad is uniquely determined by (5.17b) or we have *Case I* of Sec. 4. Now we consider if (5.17a, b) are preserved under propagation in the e_1 direction. Using (5.16a, b), (5.4b), and $\partial_2\Phi = \partial_3\Phi = 0$ [which follows from (5.15d)] $\Rightarrow$

$$(\gamma^2_{12} - \gamma^3_{13})\partial_3\gamma^2_{12} = 0 = (\gamma^2_{12} - \gamma^3_{13})\partial_2\gamma^2_{12}.$$

With (5.16), (5.18) this shows there are two types of solution:

Case Ai: $\gamma^2_{12} = \gamma^3_{13}$. This is *Case I* of Sec. 4 where the flow is locally rotationally symmetric, and the vortex tubes do not "shear".

Case Aii:

$$\gamma^2_{12} \neq \gamma^3_{13} \Rightarrow \gamma^2_{32} = \gamma^3_{23} = 0; \tag{5.19}$$

and the vortex tubes "shear" along themselves. With (5.17), (5.19) the equations yet to be satisfied are (on eliminating Φ) (5.15c), (5.3d), together with

$$\Lambda - \omega^2 + \gamma^2_{12}\gamma^3_{13} = 0, \tag{5.20a}$$

$$\partial_1(\gamma^2_{12} + \gamma^3_{13}) = 4\omega^2 + (\gamma^2_{12} - \gamma^3_{13})^2, \tag{5.20b}$$

$$\partial_1(\gamma^2_{12} - \gamma^3_{13}) = (\gamma^2_{12} + \gamma^3_{13})(\gamma^2_{12} - \gamma^3_{13}), \tag{5.20c}$$

where (5.20a) is preserved under e_1 propagation. From (5.20b),

$$\partial_1\omega = 0 \Leftrightarrow \gamma^2_{12} + \gamma^3_{13} = 0 \Rightarrow \omega = 0$$

so that, in the solutions of *Case Aii* , we must have $\gamma^2_{12} + \gamma^3_{13} \neq 0$, i.e., $\partial_1\omega \neq 0$.

Coordinate specializations (5.14), (5.17c) leave freedom

$$x^{0'} = x^0 + f(x^3), \quad x^{1'} = x^{1'}(x^1), \quad x^{2'} = x^{2'}(x^2), \quad x^{3'} = x^{3'}(x^3).$$

Then (5.17a), (5.19) and

$$\gamma^2_{12} = \gamma^2_{12}(x^1), \quad \gamma^3_{13} = \gamma^3_{13}(x^1)$$

[which follow from (5.16), (5.18)] show that the coordinates can be chosen so that

$$e^1_1 = 1, \quad e^2_2 = e^2_2(x^1), \quad e^3_3 = e^3_3(x^1). \tag{5.21}$$

We write $e^2_2 = B(x^1)F(x^1)$, $e^3_3 = B(x^1)/F(x^1)$. Then $\omega = \omega(x^1) \Rightarrow$

$$\omega = C_1B^2(x^1), \quad y = -2C_1x^2 \tag{5.22a}$$

is a possible choice of y, and (5.20c) $\Rightarrow$

$$\gamma^2_{12} - \gamma^3_{13} = 2C_2B^2(x^1), \tag{5.22b}$$

where C_1, C_2 are constants. Now (5.20a) is a differential equation for B, (5.22b) is a differential equation for F, and (5.20b) is identically satisfied in consequence of the other equations.

On comparing these equations with those of *Case Ib*, Sec. 4, the differential equations for B are *precisely the same* on putting

$$K = 0, \quad c = [(C_1)^2 + (C_2)^2]^{\frac{1}{2}}. \tag{5.22c}$$

The solutions for B are therefore (4.25) with K, c determined by (5.22c), where $B = 1/Y$. Then (5.22a) determines ω and (5.15c) determines ρ; thus, the behavior of ω, ρ is the same as in *Case Ib*, Sec. 4 (where $K = 0$); both become infinite for finite x^1 values in all cases. Having found B, (5.22b) $\Rightarrow$

$$F = \exp\,[C_2 \int B^2\,(x^1)\,dx^1].$$

On using coordinates (5.21), (5.22a), the Killing vectors may easily be found for *Case Aii* from the condition that, for any Killing vector Z, $[Z, e_a] = 0$ (as the e_a are uniquely defined). A basis of Killing vectors is

$$\partial/\partial x^0, \quad \partial/\partial x^2 - 2C_1x^3(\partial/\partial x^0), \quad \partial/\partial x^3,$$

so the space is invariant under a G_3 simple transitive in the surfaces $x^1 = \text{const}$.

Theorem 5.2: If dust has locally $\sigma = 0$, $\omega \neq 0$, then $\rho = \text{const}$ on the hypersurfaces normal to $\omega^i \Leftrightarrow$ space–time is locally homogeneous on these surfaces. If in addition ρ is constant along the vortex lines, then space–time is locally homogeneous.

The first part follows since it holds for both *Cases Ai* and *Aii*; since $\partial_1\omega \neq 0$, in *Case Aii*, the only case in which ρ, ${}_i\omega^i = 0$ is the homogeneous case of *Case Ai*, which (from Sec. 4) is Gödel's world model. Combining this theorem with Theorem 5.1, we get Theorem 5.3.

Theorem 5.3: If there exist spacelike hypersurfaces, $\rho = \text{const}$ in an open set U in which dust moves without shear, then space-time in U is locally either (i) a Friedmann universe or (ii) a Gödel universe.

This strengthens a theorem of Gödel,[9] since it shows that, if locally $\sigma = \Theta = 0$ and ρ is spatially constant, then space-time is locally either (i) Einstein's static universe or (ii) Gödel's universe.

Case B: There exists a Killing vector independent of u in the 2-surfaces spanned by u^i and ω^i.

This Killing vector has the form

$$Z = Z^0(\partial/\partial x^0) + Z^1(\partial/\partial x^1), \quad Z^1 \neq 0. \tag{5.23}$$

As ω is a scalar invariant, $\omega, {}_iZ^i = 0 \Rightarrow$ [via (5 12), (5.3d)]

$$\partial_1\omega = 0 \Leftrightarrow \gamma^2_{12} + \gamma^3_{13} = 0. \tag{5.24a}$$

Then (5.16b) reads

$$\partial_1\gamma^1_{21} = -\gamma^3_{13}\gamma^1_{21},$$

so we can use the freedom of rotation of e_2, e_3 to set $\gamma^1_{21} = 0$ in a surface $x^0 = c^0$, $x^1 = c^1$; then

$$\gamma^1_{21} = 0 \Leftrightarrow \partial_2\omega = 0 \tag{5.24b}$$

holds everywhere, and $\omega = \omega(x^3)$.

Case Bi: $\partial_3\omega = 0 \Leftrightarrow \gamma^1_{31} = 0 \Rightarrow \omega$ is spatially constant; this is Gödel's world model, by Theorem 5.3.

Case Bii: $\partial_3\omega \neq 0 \Leftrightarrow \gamma^1_{31} \neq 0 \Rightarrow$ now the tetrad is uniquely covariantly defined at each point, e_3 is the direction orthogonal to the surfaces $\omega = \text{const}$.

The coordinate transformations of the group of motions (produced by dragging along coordinates with the points) leave invariant the metric, and so preserve the *form* of the metric; these transformations therefore have the form (5.14c). The Killing vector (5.23) must then have the form[37] of a vector field generating transformations (5.14c), so $Z^0 = Z^0(x^3)$, $Z^1 = Z^1(x^1)$. The coordinate freedom of x^1 can be used to set $Z^1 = 1$; then

$$Z = Z^0(x^3)\partial/\partial x^0 + \partial/\partial x^1. \tag{5.25}$$

As the tetrad is now uniquely covariantly defined, Killing's equations may now be written as

$$[Z, e_a] = 0, \tag{5.26a}$$

which $\Rightarrow$

$$e^i_a = e^i_a(x^2, x^3). \tag{5.26b}$$

Now (5.26b), (5.23) $\Rightarrow$

$$\gamma^2_{13} = 0 = \gamma^3_{23}. \tag{5.27a}$$

As γ^1_{31} is an invariant [by choice (5.24b) of the tetrad], $(\gamma^1_{31}), {}_iZ^i = 0 \Rightarrow$ [on using (5.16a), (5.24a), (5.27a)]

$$\gamma^2_{12} = 0 = \gamma^3_{13}. \tag{5.27b}$$

The nonzero γ^a_{bc} are ω, γ^1_{31} and γ^2_{32}. By (5.24), (5.26), and (5.27) we can choose coordinates x^2, x^3 so that

$$e^1_1 = e^1_1(x^3), \quad e^2_3 = 0, \quad e^3_3 = 1. \tag{5.28a}$$

Also, $\begin{pmatrix} 1 \\ 123 \end{pmatrix}$, [(11) − (22) + (33)], and (5.15d) show

$$\partial_2\gamma^1_{31} = 0 = \partial_2\gamma^2_{32}, \tag{5.29}$$

so that the remaining freedom $x^{2'} = x^{2'}(x^2)$ of x^2 can be used to set

$$e^2_2 = e^2_2(x^3). \tag{5.28b}$$

The final coordinate freedom preserving the metric form after the specializations (5.28) is $x^{0'} = x^0 + f(x^3)$, $x^{\nu'} = A^\nu x^\nu + c^\nu$ (no sum; A^ν, c^ν are constants, $A^3 = 1$); with the fact $\omega = \omega(x^3)$, $\partial_3\omega \neq 0$, and (5.26a) this shows that a general Killing vector has the form

$$Z = Z^0(x^3)(\partial/\partial x^0) + \lambda^1(\partial/\partial x^1) + \lambda^2(\partial/\partial x^2)$$
$$(\lambda^1, \lambda^2 \text{ constants}).$$

From (5.26a), a basis of Killing vectors is $\partial/\partial x^0$, $\partial/\partial x^1$, and $Z^0(x^3)\partial/\partial x^0 + \partial/\partial x^2$, where Z^0 is found from $\omega(x^3)$ and (5.26a).

Theorem 5.4: If dust moves locally with $\sigma = 0$, $\omega \neq 0$ and there is a Killing vector independent of u in the 2-surfaces spanned by u^i, ω^i then space-time is locally invariant under either (i) an abelian G_3 of motions simply transitive in timelike hypersurfaces, or (ii) a G_5 of motions multiply transitive on space-time.

Case Bii is in fact the special solution of van Stockum,[7] studied also by Lanczos and Wright; we would have found the same coordinates as Wright (up to the numbering) if we had chosen a tetrad with $\gamma^1_{31} = 0$ (and ω, γ^1_{21}, $\gamma^3_{32} \neq 0$) instead of $\gamma^1_{21} = 0$.

Case C: There exists a Killing vector, independent of u, in each 3-surface orthogonal to ω^i.

With coordinates (5.14), the coordinate freedom [the tetrad still being free by a rotation $\theta(x^2, x^3)$] is

$$x^{0'} = x^0 + f(x^2, x^3), \quad x^{1'} = x^{1'}(x^1), \quad x^{2'} = x^{2'}(x^2, x^3),$$
$$x^{3'} = x^{3'}(x^2, x^3).$$

The general form of a Killing vector is therefore

$$Z = Z^0(x^2, x^3)(\partial/\partial x^0) + Z^1(x^1)(\partial x^2)$$
$$+ Z^2(x^2, x^3)(\partial/\partial x^2) + Z^3(x^2, x^3)(\partial/\partial x^3).$$

[37] We follow methods used by R. P. Kerr.

For the Killing vector considered, $Z^1 = 0$.

We may choose a 2-surface $x^0 = c^0$, $x^1 = c^1$ and rotate e_2, e_3 until at each point in this surface,

$$Z \cdot e_3 = 0 \Leftrightarrow Z^3 = 0;$$

but then

$$Z^3 = Z^3(x^2, x^3) \Rightarrow Z^3 = 0$$

everywhere. The tetrad is now fixed, so the coordinate freedom is (5.14c). We may set $Z^2 = 1$ by suitable choice of x^2; the Killing vector is then

$$Z = Z^0(x^3)(\partial/\partial x^0) + (\partial/\partial x^2). \tag{5.30}$$

As ω is a scalar invariant,

$$\omega, {}_i Z^i = 0 \Rightarrow \partial_2 \omega = 0 = \gamma^1_{21}. \tag{5.31}$$

Case Ci: $\partial_2\omega = 0 = \partial_3\omega \Leftrightarrow \gamma^1_{21} = \gamma^1_{31} = 0$; this is the *Case A* above.

Case Cii: $\partial_3\omega \neq 0 \Leftrightarrow \gamma^1_{31} \neq 0$. The tetrad is now uniquely defined; Killing's equations are therefore (5.26a), which $\Rightarrow$

$$e^i_a = e^i_a(x^1, x^3) \tag{5.32a}$$

and so

$$\gamma^3_{23} = 0, \quad \partial_2 \gamma^a_{bc} = 0. \tag{5.32b}$$

From (5.15b) and $\partial_3\omega \neq 0$,

$$\omega(x^1, x^3) = f(x^1)e^1_1(x^1, x^3);$$

we can use the freedom of x^1, x^3 to set

$$\omega = e^1_1 = e^2_2 e^3_3, \tag{5.33a}$$

and can then choose y to be

$$y = -2x^2. \tag{5.33b}$$

Equation (5.33a) can be used to eliminate e^1_1. Then (23) and (12) $\Rightarrow$

$$\gamma^2_{13} = c(e^2_2)^2, \tag{5.34a}$$

where c is a constant; (5.34a) $\Rightarrow$ we can choose

$$e^2_3 = ce^3_3\left(\int \frac{(e^2_2)^2}{(e^3_3)^2}\, dx^1\right). \tag{5.34b}$$

On eliminating Φ four equations remain to be satisfied, namely, equations (5.4) and (31). We can regard two of these as propagation equations, and the other two as constraint equations; these are conserved under the propagation (at least in the analytic case). We can specify e^3_3 and $(\partial/\partial x^1)(e^3_3)$ as arbitrary functions of x^3, and specify e^2_2, $(\partial/\partial x^1)(e^2_2)$, and $(\partial/\partial x^3)(e^2_2)$ arbitrarily at a point; then e^2_2 and e^3_3 are determined completely for any given values of the constants c and Λ. ρ is determined by (5.15c).

If the constant c is zero, e_3 and e_1 are 2-surface forming and e_0, e_1, e_3 are 3-surface forming. In this case we have van Stockum's general solution[7] in a different coordinate system. We could obtain van Stockum's type of coordinates [see Ref. 7, Eq. (3.2)] on choosing γ^1_{31} zero instead of γ^1_{21}.

Special cases of integration occur in *Case Ci*, and in *Case Ciii:*

$$\gamma^2_{12} = \gamma^3_{13} = 0. \tag{5.35}$$

This can be seen as follows: apart from *Case Ci* $[(\partial/\partial x^3)(e^1_1) = 0 \Leftrightarrow \gamma^1_{31} = 0]$ the general integration procedure may break down if $(\partial/\partial x^3)(e^2_2) = 0$, $(\partial/\partial x^1)(e^2_2) = 0$ or $(\partial/\partial x^1)(e^3_3) = 0$. If $(\partial/\partial x^3)(e^2_2) = 0$, i.e. $\gamma^2_{32} = 0$, then (31) $\Rightarrow \gamma^2_{12} = 0$. But with (5.13), (5.31), and (5.32b) $\gamma^2_{12} = \gamma^2_{32} = 0 \Rightarrow \omega = 0$ by (5.4), which is not allowed. If $(\partial/\partial x^1)(e^2_2) = 0$, i.e. $\gamma^2_{12} = 0$, then $\gamma^3_{13} = 0$ by (31), $\begin{pmatrix} 2 \\ 123 \end{pmatrix}$. Finally, $(\partial/\partial x^1)(e^3_3) = 0 \Leftrightarrow \gamma^3_{13} = 0$; conservation of this condition gives a new constraint equation [by (5.4)]; taking the ∂_1 derivative of this equation three times, $\omega \neq 0 \Rightarrow \gamma^2_{12} = 0$ also.

With condition (5.35), we can choose coordinates as before; now

$$e^2_2 = e^2_2(x^3), \quad e^3_3 = e^3_3(x^3)$$

and

$$e^2_3 = cx^1[(e^2_2)^2/e^3_3].$$

One of the four equations is now identically satisfied; using the definitions $A(x^3) = (e^2_2)^{-1}$,

$$B(x^3) = (e^3_3)^{-1}, \quad \partial/\partial x^3 = {}',$$

the remaining equations are

$$\frac{A''}{A} + \frac{A'^2}{A^2} - \frac{1}{2}c^2\frac{B^2}{A^4} = -2\Lambda B^2, \tag{5.36a}$$

$$\frac{A'B'}{AB} + \frac{A'^2 + 1}{A^2} - \frac{1}{4}c^2\frac{B^2}{A^4} = -\Lambda B^2, \tag{5.36b}$$

$$\frac{A''}{A} + \frac{A'}{A}\frac{B'}{B} + \frac{B''}{B} - \frac{B'^2}{B^2} + \frac{1}{A^2} + \frac{3}{4}c^2\frac{B^2}{A^4} = -\Lambda B^2, \tag{5.36c}$$

where (5.36b) is a first integral of the other two equations. [It is interesting to compare these equations with (4.39).] In *Case Cii*, there exist the Killing vectors $\partial/\partial x^0$ and (5.30); in *Case Ciii*, there exists a third Killing vector

$$\partial/\partial x^1 + Z^0(x^3)\partial/\partial x^0 + Z^2(x^3)\partial/\partial x^2,$$

where Z^0, Z^2 are found from (5.26a). $Z^0 = Z^2 = 0$ if and only if $c = 0$; i.e., this Killing vector lies in the 2-surface spanned by u^i and ω^i (and so we have *Case B*) if and only if $\gamma^2_{13} = c = 0$.

Cases Ci and *Ciii* are thus special cases of integration of *Cii*. A convenient way of visualizing this is as follows: excluding the case in which ρ (or ω) is

spatially homogeneous, the vectors e_0 and e_2 lie at each point in the hypersurfaces $\rho = \text{const}$. In general (*Case Cii*) neither e_1 nor e_3 lie in these surfaces; if e_3 lies in these surfaces at each point ($\partial_3\omega = 0$) then we have *Case Ci* (i.e., *Case A*); if e_1 lies in these surfaces at each point ($\partial_1\omega = 0 \Leftrightarrow \rho_{,i}\omega^i = 0$) we have (see proof below) *Case Ciii*, which contains *Case B* as the subclass when $\gamma^2_{13} = 0$. *Case I* (rotational symmetry) is contained in *Case A* (i.e., *Case Ci*); and both *Ci* and *Ciii* are special cases of van Stockum's general metric, since in both these cases $\gamma^2_{13} = 0$.

To prove that $\partial_1\omega = 0 \Rightarrow$ *Case Ciii*, we proceed as follows. We have $\gamma^2_{12} + \gamma^3_{13} = 0$. Conservation of this condition gives a new constraint equation [see (5.4b)]; this constraint equation is preserved if

$$\gamma^2_{12}(2\partial_3\gamma^1_{31} - 3(\gamma^1_{31})^2 + 2\gamma^1_{31}\gamma^2_{32}) = 0.$$

If $\gamma^2_{12} \neq 0$, this gives an expression for $\partial_3\gamma^1_{31}$ which can be substituted into (5.4b) to give

$$\Phi = 2\gamma^1_{31}\gamma^2_{32} - \tfrac{1}{2}(\gamma^1_{31})^2 + 2(\gamma^2_{12})^2 + \tfrac{1}{2}(\gamma^2_{13})^2.$$

Differentiating this twice in the e_3 direction, we find that this is inconsistent with $\gamma^2_{12} \neq 0$; so

$$\partial_1\omega = 0 \Rightarrow \gamma^2_{12} = 0 = \gamma^3_{13}.$$

6. CONCLUSION

We have chosen a tetrad and coordinate system to describe any space–time containing dust, and then used this system to study the special cases of locally rotationally symmetric dust and shear-free dust.

Solutions which are locally rotationally symmetric fall into the following cases: *Case I:* $\omega \neq 0$, $\sigma = \Theta = 0$; space is homogeneous in the 3-surfaces orthogonal to ω^i. Gödel's world model is a special case of integration. *Case IIa:* Bondi's spherically symmetric solution and closely related solutions, $\Theta \neq 0$, $\sigma \neq 0$, $\omega = 0$, (with Einstein's static space–time as a special case). The Friedmann world models are contained in this set (when $\sigma = 0$). *Case IIb:* inhomogeneous generalizations of the Kantowski–Sachs models and the Bianchi model invariant under a group G_3 of type I; $\Theta \neq 0$, $\sigma \neq 0$, $\omega = 0$. The Einstein–de Sitter space–time is a special case. *Case III:* spatially homogeneous (Bianchi) models, invariant under groups G_3 of types II, VIII, IX, with $\Theta \neq 0$, $\sigma \neq 0$, $\omega = 0$. In this case, two differential equations remain to be integrated.

The condition of local rotational symmetry is a strong condition, as it implies the existence of a multiply transitive group of motions in every case. A similar strongly restrictive condition is that of isotropy about a point. Dust-filled spaces satisfying this condition (examined by Bondi[8]) are those of type *IIa*, $K > 0$. The existence of shear-free dust flows puts fairly strong restrictions on space–time: space is either conformally flat ($\sigma = \omega = 0$) or stationary ($\sigma = \Theta = 0$). In the latter case there exist solutions[38] in which the conformal tensor is of Type I, in contrast to the case of a vacuum null congruence.

Three simple classes of solution ($\sigma = 0$, $\omega \neq 0$) are examined. *Case A* solutions are those in which ρ is constant on the 3-surfaces orthogonal to ω^i; *Case Ai* solutions are rotationally symmetric solutions (*Case I*), while *Case Aii* are a family of different solutions in which ρ, ω have similar properties. In all these solutions, there is a simply transitive group of motions in the surfaces of constant density. *Case B* solutions are those in which there is a Killing vector in the 2-surfaces spanned by u^i, ω^i (as well as u^i); these are the special solutions of van Stockum plus Gödel's world model. *Case C* solutions are those in which there is a Killing vector (other than u) in the 3-surfaces orthogonal to ω^i. Four differential equations remain to be integrated (two propagation equations and two constraint equations). These spaces are generalizations of the general solution of van Stockum; *Cases A* and *B* are special integrations of van Stockum's general solution, and so are special cases of *Case C*.

Of the spaces listed above, *Cases Ib*, *IIb* and *Aii* are (as far as the author is aware) new examples of analytic exact solutions. None of the spaces described here have both nonzero expansion and rotation, so they are all too restricted to give information about the general behavior of dust. Known analytic exact solutions with nonzero shear and rotation (invariant under simply-transitive G_4's[13] and G_3's[15]) have no expansion, and so are also rather restricted. The simplest space–times found so far in which dust has general motion are the spatially homogeneous Bianchi solutions.[3,11,12,14] (It is interesting to note that surprisingly many of the properties found for these models still hold if the restriction of homogeneity is removed.) However, no analytic solutions have yet been found with Θ and ω nonzero. Without such solutions, the interaction of expansion, shear, and rotation in the exact field equations remain obscure. Table 1 summarizes the known dust solutions admitting groups of motions G_r, $r \geq 2$.

The methods used in this paper are fairly standard tetrad methods; these have the usual advantage that, having introduced the rotation coefficients as auxiliary variables, the equations are first-order equations; and the accompanying disadvantage that a new set of

[38] J. Ehlers (private communication).

TABLE I. Dust solutions ($\rho \neq 0$) admitting groups G_r ($r \geq 2$).[a]

		Dimension s of minimum invariant varieties			
		$s = 2$	$s = 3$	$s = 4$	
Dimension q of isotropy group		Group G_2	Group G_3	Group G_4	
	$q = 0$	Metrics all known (see Petrov, Ref. 28, pp. 195–200); some models studied; for models of type T_2 see van Stockum, Ref. 7 and *Case Cii*.	In principle, all known (e.g., Petrov, Ref. 28, pp. 206–222) Type S_3 (Bianchi models) have been extensively investigated (see Refs. 9, 11, 12, 14); for Type T_3 examples see Ref. 6, 7, 15 and *Cases Aii, Bii, Ciii*.	All explicitly known (see Oszvath *et al.*, Ref. 13); ($\sigma \neq 0$, $\omega \neq 0$).	Group is simply transitive; Weyl tensor is not zero or Type N (Trumper, Szekeres Ref. 39); either $\sigma \neq 0$ or $\omega \neq 0$, $\sigma = \Theta = 0$.
		Group G_3	Group G_4	Group G_5	
	$q = 1$	Type S_2: *Case IIai* (Bondi-type); *Case IIbi*, $\Lambda \gtrless 0$ (generalized Kantowski–Sachs) and *Case IIbiii*, $\Lambda \lesseqgtr 0$ (generalized Bianchi I).	Type T_3: *Case Ib* ($\omega \neq 0$). Type S_3: *Case III; Case IIbi*, $\Lambda \gtrless 0$ (Kantowski–Sachs) and *Case IIbiii*, $\Lambda \gtrless 0$ (Bianchi I). T_3, S_3, or N_3: *Case IIaiii* (Farnsworth) and *Case IIbiv*.	Gödel world model (Ref. 9) (*Case Ia*).	Group is multiply transitive, Weyl tensor is Type D, either $\sigma \neq 0$, $\Theta \neq 0$, $\omega = 0$ or $\omega \neq 0$, $\sigma = \Theta = 0$.
			Group G_6	Group G_7	
	$q = 3$	none	Type S_3: Friedmann models, $\Theta \neq 0$ (*Case IIaii*).	Einstein static universe (*Case IIai*).	Group is multiply transitive, Weyl tensor is zero, $\sigma = \omega = 0$.

[a] *Dust solutions* ($\rho \neq 0$) *admitting groups* G_r, $r \geq 2$. Dimension of minimum invariant varieties is s, isotropy group is of dimension q, so $r = q + s$. As the vector u^a is invariant, $q \neq 2$, $q \leq 3$. If $r = 1$ or 0, then $q = 0$. For $s = 4$, space–time is homogeneous, so $\Theta = 0$. For $s = 2$ or 3, the nature of the 2- (or 3-) surfaces of transitivity is denoted by: S(spacelike), N(null), or T(timelike); e.g., S_2 = spacelike 2-surface of transitivity. The table is complete except for $q = 0$, $s = 2$ and $q = 0$, $s = 3$. The solutions admitting multiply transitive groups containing no simply transitive subgroups are Types IIai and IIbi ($s = 2$, $q = 1$) with $K > 0$, and Kantowksj–Sachs models IIbi ($s = 3$, $q = 1$) with $K > 0$. If $\Lambda = 0$, then: $\sigma \neq 0 \Rightarrow \Theta \neq 0$ or $\omega \neq 0$ (see Ehlers, Refs. 5, 6); there are no solutions for $s = 4$; and *Case IIbiv* cannot occur. When $\Lambda = 0$, then (for any r) $q = 0 \Rightarrow$ either (i) $\sigma\omega \neq 0$, (ii) $\sigma\Theta \neq 0$, $\omega = 0$, or (iii) $\sigma = \Theta = 0$, $\omega \neq 0$.

equations (integrability conditions for the new variables) have to be introduced. The tetrad methods used here differ from tetrad methods used in much recent work (e.g., Ref. 19) in two related aspects: (i) we use an orthonormal rather than a pseudo-orthonormal or null tetrad; and (ii) the Weyl tensor components and associated integrability conditions (the Bianchi identities) are not explicitly introduced.

The choice of an orthonormal tetrad is suggested by the existence of a preferred timelike congruence (the flow lines) and spacelike congruence (the vorticity lines) in space–time. Use of a null tetrad is certainly appropriate for investigation of empty space–times, in which the Weyl tensor propagates along null geodesics; in a dust-filled space–time, either the null or timelike structure might be dominant. In this paper, the timelike congruence is regarded as the principal feature; the Bianchi identities are not introduced, equivalent integrability conditions being obtained when needed by use of the vector commutators. No difficulty arises in finding the Weyl tensor components from the γ^a_{bc} when desired. This procedure contrasts with several interesting papers[39] in which the Bianchi identities are used to relate properties of dust to properties of the Weyl tensor.

ACKNOWLEDGMENTS

The author owes much to discussions with the members of the Relativity Groups at Cambridge University and The University of Texas. In particular, he would like to thank Dr. D. W. Sciama for advice and encouragement during the early part of this work, and Dr. R. K. Sachs, Dr. R. P. Kerr, Dr. E. Schücking, and Dr. J. Ehlers for much helpful advice. He would also like to thank the members of the Relativity Center at The University of Texas for their extensive hospitality.

This work was supported in part by the office of the Aerospace Research of the U.S. Air Force under grant AF-AFOSR-454-65.

[39] W. Kundt and M. Trumper, Abhandl. Math. Naturw. Kl., Mainz Akad. Wiss. Lit. 12, (1962); M. Trumper, preprint, Hamburg (1963) and Seminar, Cambridge University (1965); P. Szekeres, J. Math. Phys. 7, 751 (1966); M. Cahen (unpublished).

APPENDIX

With the coordinates and tetrad of Theorem 3.1, $u^i = e_0^i$, $\omega^i = e_1^i$, and

$$\begin{aligned} e_0 &= \partial/\partial x^0, \\ e_1 &= e_1^1(\partial/\partial x^1), \\ e_2 &= e_2^1(\partial/\partial x^1) + e_2^2(\partial/\partial x^2), \\ e_3 &= ye_3^3(\partial/\partial x^0) + e_3^1(\partial/\partial x^1) + e_3^2(\partial/\partial x^2) + e_3^3(\partial/\partial x^3), \end{aligned} \tag{A1}$$

where $e_a^i = e_a^i(x^j)$, $y = y(x^2, x^3)$, and at each point $e_1^1 e_2^2 e_3^3 \neq 0$.

The metric tensor components g_{ij} are then

$$\begin{aligned} &g_{00} = -1, \quad g_{01} = g_{02} = 0, \quad g_{03} = y(x^2, x^3), \\ &g_{11} = \frac{1}{(e_1^1)^2}, \quad g_{12} = \frac{-e_2^1}{(e_1^1)^2 e_2^2}, \\ &g_{13} = \frac{1}{(e_1^1)^2 e_3^3}\left[-e_3^1 + \frac{e_2^1 e_3^2}{e_2^2}\right], \\ &g_{22} = \frac{1}{(e_2^2)^2}\left[1 + \frac{(e_2^1)^2}{(e_1^1)^2}\right], \\ &g_{23} = \frac{1}{e_2^2 e_3^3}\left[\frac{-e_3^2}{e_2^2} + \frac{e_2^1 e_3^1}{(e_1^1)^2} - \frac{(e_2^1)^2 e_3^2}{(e_1^1)^2 e_2^2}\right], \\ &g_{33} = \frac{1}{(e_3^3)^2}\left[1 + \frac{(e_3^2)^2}{(e_2^2)^2} + \left(\frac{e_3^2 e_2^1}{e_1^1 e_2^2} - \frac{e_3^1}{e_1^1}\right)^2\right] - y^2. \end{aligned} \tag{A2}$$

The coefficients γ_{bc}^a [defined by (2.21)] are

$$\begin{aligned} &\gamma_{01}^0 = \gamma_{02}^0 = \gamma_{03}^0 = 0, \quad \gamma_{01}^2 = \gamma_{01}^3 = \gamma_{02}^3 = 0, \quad \gamma_{13}^0 = \gamma_{12}^3 = \gamma_{12}^0 = 0. \\ &\gamma_{01}^1 = \frac{\partial}{\partial x^0}(\log e_1^1) = -\Theta_1, \quad \gamma_{02}^2 = \frac{\partial}{\partial x^0}(\log e_2^2) = -\Theta_2, \quad \gamma_{03}^3 = \frac{\partial}{\partial x^0}(\log e_3^3) = -\Theta_3. \\ &\gamma_{02}^1 = \frac{e_2^1}{e_1^1}\frac{\partial}{\partial x^0}\left(\log\frac{e_2^1}{e_2^2}\right) = -2\sigma_{12}, \quad \gamma_{03}^2 = \frac{e_3^2}{e_2^2}\frac{\partial}{\partial x^0}\left(\log\frac{e_3^2}{e_3^3}\right) = -2\sigma_{23}, \quad \gamma_{03}^1 = \frac{e_3^1}{e_1^1}\frac{\partial}{\partial x^0}\left(\log\frac{e_3^1}{e_3^3}\right) - \frac{e_2^1}{e_1^1}\gamma_{03}^2 = -2\sigma_{13} \\ &\gamma_{12}^2 = e_1^1\frac{\partial}{\partial x^1}(\log e_2^2) = e_1 \cdot \nabla_2 e_2, \quad \gamma_{13}^3 = e_1^1\frac{\partial}{\partial x^1}(\log e_3^3) = e_1 \cdot \nabla_3 e_3, \\ &\gamma_{21}^1 = e_2^2\frac{\partial}{\partial x^2}(\log e_1^1) - e_2^1\frac{\partial}{\partial x^1}\left(\log\frac{e_2^1}{e_1^1 e_2^2}\right) = e_2 \cdot \nabla_1 e_1, \\ &\gamma_{31}^1 = ye_3^3\gamma_{01}^1 + e_3^3\frac{\partial}{\partial x^3}(\log e_1^1) + e_3^2\frac{\partial}{\partial x^2}(\log e_1^1) - e_3^1\frac{\partial}{\partial x^1}\left(\log\frac{e_3^1}{e_1^1 e_3^3}\right) + \frac{e_2^1}{e_1^1}\gamma_{13}^2 = e_3 \cdot \nabla_1 e_1, \\ &\gamma_{23}^3 = e_2^2\frac{\partial}{\partial x^2}(\log e_3^3) + e_2^1\frac{\partial}{\partial x^1}(\log e_3^3) = e_2 \cdot \nabla_3 e_3, \\ &\gamma_{32}^2 = ye_3^3\gamma_{02}^2 + e_3^3\frac{\partial}{\partial x^3}(\log e_2^2) + e_3^1\frac{\partial}{\partial x^1}(\log e_2^2) - e_3^2\frac{\partial}{\partial x^2}\left(\log\frac{e_3^2}{e_2^2 e_3^3}\right) - \frac{e_2^1}{e_1^1}\gamma_{13}^2 = e_3 \cdot \nabla_2 e_2, \\ &\gamma_{23}^0 = e_2^2 e_3^3\frac{\partial y}{\partial x^2} = -2\omega, \quad \gamma_{13}^2 = \frac{e_1^1 e_3^2}{e_2^2}\frac{\partial}{\partial x^1}\left(\log\frac{e_3^2}{e_3^3}\right), \\ &\gamma_{32}^1 = ye_3^3\gamma_{02}^1 + \frac{e_2^1}{e_1^1}\left[e_3^1\frac{\partial}{\partial x^1}\left(\log\frac{e_2^1 e_3^3}{e_3^1 e_2^2}\right) + e_3^2\frac{\partial}{\partial x^2}\left(\log\frac{e_2^1 e_3^2}{e_2^2 e_3^3}\right) + e_3^3\frac{\partial}{\partial x^3}\left(\log\frac{e_2^1}{e_2^2}\right) + \frac{e_2^1}{e_1^1}\gamma_{13}^2\right] - \frac{e_3^1 e_2^2}{e_1^1}\frac{\partial}{\partial x^2}\left(\log\frac{e_3^1}{e_3^3}\right). \end{aligned} \tag{A3}$$

The Jacobi identities and field equations are

$$\begin{pmatrix} 0 \\ 123 \end{pmatrix}\partial_1\omega = \omega(\gamma_{12}^2 + \gamma_{13}^3), \quad \begin{pmatrix} 0 \\ 023 \end{pmatrix}\partial_0\omega = -\omega(\Theta_2 + \Theta_3),$$

$$\binom{1}{123}\partial_2\gamma^1_{31} - \partial_3\gamma^1_{21} - \partial_1\gamma^1_{32} = 2\omega\Theta_1 - \gamma^1_{32}(\gamma^2_{12} + \gamma^3_{13}) - \gamma^2_{32}\gamma^1_{21} + \gamma^3_{23}\gamma^1_{31},$$

$$\binom{2}{123}\partial_3\gamma^2_{12} - \partial_1\gamma^2_{32} - \partial_2\gamma^2_{13} = -\gamma^2_{13}(\gamma^3_{23} + \gamma^1_{21}) - \gamma^3_{13}\gamma^2_{32} + \gamma^1_{31}\gamma^2_{12},$$

$$\binom{3}{123}\partial_1\gamma^3_{23} - \partial_2\gamma^3_{13} = -\gamma^1_{21}\gamma^3_{13} + \gamma^2_{12}\gamma^3_{23},$$

$$\binom{1}{012}2\partial_1\sigma_{12} - \partial_2\Theta_1 - \partial_0\gamma^1_{21} = \Theta_2\gamma^1_{21} + 2\sigma_{12}\gamma^2_{12},$$

$$\binom{1}{013}-2\partial_1\sigma_{13} + \partial_3\Theta_1 + \partial_0\gamma^1_{31} = -\Theta_3\gamma^1_{31} - 2\sigma_{31}\gamma^3_{13} - 2\sigma_{23}\gamma^1_{21} - 2\sigma_{12}\gamma^2_{13}, \tag{A4}$$

$$\binom{3}{032}\partial_2\Theta_3 + \partial_0\gamma^3_{23} = -\Theta_2\gamma^3_{23} - 2\sigma_{12}\gamma^3_{13},$$

$$\binom{2}{023}2\partial_2\sigma_{23} - \partial_3\Theta_2 - \partial_0\gamma^2_{32} = \Theta_3\gamma^2_{32} + 2\sigma_{23}\gamma^3_{23} + 2\sigma_{31}\gamma^2_{12} - 2\sigma_{12}\gamma^2_{13},$$

$$\binom{3}{031}\partial_1\Theta_3 + \partial_0\gamma^3_{13} = -\Theta_1\gamma^3_{13}, \quad \binom{2}{021}\partial_1\Theta_2 + \partial_0\gamma^2_{12} = -\Theta_1\gamma^2_{12},$$

$$\binom{1}{023}2\partial_2\sigma_{31} - 2\partial_3\sigma_{12} - \partial_0\gamma^1_{32} = -2\sigma_{12}(\gamma^2_{32} - \gamma^1_{31}) - 2\sigma_{31}(\gamma^1_{21} - \gamma^3_{23}) + \gamma^1_{32}(\Theta_2 + \Theta_3 - \Theta_1),$$

$$\binom{2}{031}2\partial_1\sigma_{23} + \partial_0\gamma^2_{13} = 2\sigma_{23}(\gamma^3_{13} - \gamma^2_{12}) - \gamma^2_{13}(\Theta_3 + \Theta_1 - \Theta_2).$$

(The remaining three Jacobi identities are identically satisfied.)

$$(00)\ -\Lambda + \tfrac{1}{2}\rho = -\partial_0\Theta_1 - \partial_0\Theta_2 - \partial_0\Theta_3 - (\Theta_1)^2 - (\Theta_2)^2 - (\Theta_3)^2 - 2(\sigma_{12})^2 - 2(\sigma_{23})^2 - 2(\sigma_{31})^2 + 2\omega^2,$$

$$(01)\ 0 = -\partial_1(\Theta_2 + \Theta_3) + \partial_2\sigma_{12} + \partial_3\sigma_{13} + \gamma^2_{12}(\Theta_2 - \Theta_1) + \gamma^3_{13}(\Theta_3 - \Theta_1) - \omega\gamma^1_{32} + \sigma_{23}\gamma^2_{13} - \sigma_{12}(\gamma^3_{23} + 2\gamma^1_{21}) - \sigma_{13}(\gamma^2_{32} + 2\gamma^1_{31}),$$

$$(02)\ 0 = -\partial_3\omega - \partial_2(\Theta_3 + \Theta_1) + \partial_3\sigma_{23} + \partial_1\sigma_{21} + \gamma^3_{23}(\Theta_3 - \Theta_2) + \gamma^1_{21}(\Theta_1 - \Theta_2) + \omega\gamma^1_{31} - \sigma_{31}\gamma^1_{32} - \sigma_{23}(\gamma^1_{31} + 2\gamma^2_{32}) - \sigma_{21}(\gamma^3_{13} + 2\gamma^2_{12}),$$

$$(03)\ 0 = \partial_2\omega - \partial_3(\Theta_1 + \Theta_2) + \partial_1\sigma_{31} + \partial_2\sigma_{32} + \gamma^1_{31}(\Theta_1 - \Theta_3) + \gamma^2_{32}(\Theta_2 - \Theta_3) - \omega\gamma^1_{21} + \sigma_{12}(\gamma^1_{32} - \gamma^2_{13}) - \sigma_{31}(\gamma^2_{12} + 2\gamma^3_{13}) - \sigma_{32}(\gamma^1_{21} + 2\gamma^3_{23}),$$

$$(11)\ \Lambda + \tfrac{1}{2}\rho = \partial_0\Theta_1 + \partial_1\gamma^2_{12} + \partial_1\gamma^3_{13} + \partial_2\gamma^1_{21} + \partial_3\gamma^1_{31} + \Theta_1(\Theta_1 + \Theta_2 + \Theta_3) - (\gamma^2_{12})^2 - (\gamma^3_{13})^2 - 2(\sigma_{31})^2 - 2(\sigma_{12})^2 + \tfrac{1}{2}(\gamma^1_{32})^2 - \tfrac{1}{2}(\gamma^2_{13})^2 - \gamma^1_{31}(\gamma^1_{31} + \gamma^2_{32}) - \gamma^1_{21}(\gamma^1_{21} + \gamma^3_{23}),$$

$$(22)\ \Lambda + \tfrac{1}{2}\rho = \partial_0\Theta_2 + \partial_2\gamma^3_{23} + \partial_2\gamma^1_{21} + \partial_3\gamma^2_{32} + \partial_1\gamma^2_{12} + \Theta_2(\Theta_1 + \Theta_2 + \Theta_3) - (\gamma^3_{23})^2 - (\gamma^1_{21})^2 + 2(\sigma_{12})^2 - 2(\sigma_{23})^2 + 2\omega^2 + \tfrac{1}{2}(\gamma^2_{13})^2 - \tfrac{1}{2}(\gamma^1_{32})^2 - \gamma^2_{12}(\gamma^2_{12} + \gamma^3_{13}) - \gamma^2_{32}(\gamma^2_{32} + \gamma^1_{31}),$$

$$(33)\ \Lambda + \tfrac{1}{2}\rho = \partial_0\Theta_3 + \partial_3\gamma^1_{31} + \partial_3\gamma^2_{32} + \partial_1\gamma^3_{13} + \partial_2\gamma^3_{23} + \Theta_3(\Theta_1 + \Theta_2 + \Theta_3) - (\gamma^1_{31})^2 - (\gamma^2_{32})^2 + 2(\sigma_{23} - \omega)^2 + 2(\sigma_{31})^2 - \tfrac{1}{2}(\gamma^1_{32} - \gamma^2_{13})^2 - \gamma^3_{23}(\gamma^3_{23} + \gamma^1_{21}) - \gamma^3_{13}(\gamma^3_{13} + \gamma^2_{12}),$$

$$(12)\ 0 = \partial_0\sigma_{12} + \partial_2\gamma^3_{13} + \tfrac{1}{2}\partial_3(\gamma^1_{32} - \gamma^2_{13}) + \sigma_{12}(\Theta_3 + 2\Theta_1) - 2\sigma_{13}\sigma_{23} - \gamma^1_{31}\gamma^1_{32} + \gamma^2_{32}\gamma^2_{13} + \gamma^3_{23}(\gamma^2_{12} - \gamma^3_{13}),$$

$$(23)\ 0 = \partial_0\sigma_{23} + \partial_3\gamma^1_{21} + \tfrac{1}{2}\partial_1(\gamma^2_{13} + \gamma^1_{32}) + \sigma_{23}(\Theta_1 + 2\Theta_2) + \omega(\Theta_1 - \Theta_2 + \Theta_3) + 2\sigma_{12}\sigma_{31} - \gamma^2_{12}\gamma^2_{13} - \gamma^3_{13}\gamma^1_{32} + \gamma^1_{31}(\gamma^3_{23} - \gamma^1_{21}),$$

$$(31)\ 0 = \partial_0\sigma_{31} + \partial_1\gamma^2_{32} + \tfrac{1}{2}\partial_2(\gamma^2_{13} - \gamma^1_{32}) + \sigma_{31}(\Theta_2 + 2\Theta_1) + \gamma^1_{21}(\gamma^1_{32} - \gamma^2_{13}) + \gamma^2_{12}(\gamma^1_{31} - \gamma^2_{32}).$$

The density conservation equation is

$$\partial_0\rho = -\rho(\Theta_1 + \Theta_2 + \Theta_3).$$

The tetrad freedom of a rotation (3.7) obeying (3.12) transforms the γ^a_{bc} as follows:

$$
\begin{aligned}
&(\gamma^1_{01})' = \gamma^1_{01}, \quad (\gamma^0_{23})' = \gamma^0_{23}, \quad (\gamma^1_{23})' = \gamma^1_{23},\\
&(\gamma^2_{02})' = \gamma^2_{02} + (\gamma^3_{03} - \gamma^2_{02})\sin^2\theta + \gamma^2_{03}\sin\theta\cos\theta,\\
&(\gamma^3_{03})' = \gamma^3_{03} + (\gamma^2_{02} - \gamma^3_{03})\sin^2\theta - \gamma^2_{03}\sin\theta\cos\theta,\\
&(\gamma^2_{03})' = \gamma^2_{03}(\cos^2\theta - \sin^2\theta) + 2(\gamma^3_{03} - \gamma^2_{02})\sin\theta\cos\theta,\\
&(\gamma^1_{02})' = \gamma^1_{02}\cos\theta + \gamma^1_{03}\sin\theta,\\
&(\gamma^1_{03})' = -\gamma^1_{02}\sin\theta + \gamma^1_{03}\cos\theta,\\
&(\gamma^1_{21})' = \gamma^1_{21}\cos\theta + \gamma^1_{31}\sin\theta,\\
&(\gamma^1_{31})' = -\gamma^1_{21}\sin\theta + \gamma^1_{31}\cos\theta,\\
&(\gamma^2_{12})' = \gamma^2_{12} + (\gamma^3_{13} - \gamma^2_{12})\sin^2\theta + \gamma^2_{13}\sin\theta\cos\theta,\\
&(\gamma^3_{13})' = \gamma^3_{13} + (\gamma^2_{12} - \gamma^3_{13})\sin^2\theta - \gamma^2_{13}\sin\theta\cos\theta,\\
&(\gamma^2_{13})' = \gamma^2_{13}(\cos^2\theta - \sin^2\theta) + 2(\gamma^3_{13} - \gamma^2_{12})\sin\theta\cos\theta,\\
&(\gamma^2_{32})' = (\partial_2\theta + \gamma^2_{32})\cos\theta - (-\partial_3\theta + \gamma^3_{23})\sin\theta,\\
&(\gamma^3_{23})' = (\partial_2\theta + \gamma^2_{32})\sin\theta + (-\partial_3\theta + \gamma^3_{23})\cos\theta.
\end{aligned}
\tag{A5}
$$

Simultaneously with this rotation, a coordinate transformation (3.9), obeying (3.15), is performed; then

$$
\begin{aligned}
&(e^1_1)' = e^1_1\frac{\partial x^{1'}}{\partial x^1},\\
&(e^1_2)' = \left(e^1_2\frac{\partial x^{1'}}{\partial x^1} + e^2_2\frac{\partial x^{1'}}{\partial x^2}\right)\cos\theta + \left(e^1_3\frac{\partial x^{1'}}{\partial x^1} + e^2_3\frac{\partial x^{1'}}{\partial x^2} + e^3_3\frac{\partial x^{1'}}{\partial x^3}\right)\sin\theta,\\
&(e^1_3)' = -\left(e^1_2\frac{\partial x^{1'}}{\partial x^1} + e^2_2\frac{\partial x^{1'}}{\partial x^2}\right)\sin\theta + \left(e^1_3\frac{\partial x^{1'}}{\partial x^1} + e^2_3\frac{\partial x^{1'}}{\partial x^2} + e^3_3\frac{\partial x^{1'}}{\partial x^3}\right)\cos\theta,\\
&(e^2_2)' = e^2_2\frac{\partial x^{2'}}{\partial x^2}\cos\theta + \left(e^2_3\frac{\partial x^{2'}}{\partial x^2} + e^3_3\frac{\partial x^{2'}}{\partial x^3}\right)\sin\theta,\\
&(e^2_3)' = -e^2_2\frac{\partial x^{2'}}{\partial x^2}\sin\theta + \left(e^2_3\frac{\partial x^{2'}}{\partial x^2} + e^3_3\frac{\partial x^{2'}}{\partial x^3}\right)\cos\theta,\\
&(e^3_3)' = -e^2_2\frac{\partial x^{3'}}{\partial x^2}\sin\theta + \left(e^2_3\frac{\partial x^{3'}}{\partial x^2} + e^3_3\frac{\partial x^{3'}}{\partial x^3}\right)\cos\theta.
\end{aligned}
\tag{A6}
$$

Both $y' = (\partial f/\partial x^2)/(\partial x^{3'}/\partial x^2)$ and $y' = [y + (\partial f/\partial x^3)]/(\partial x^{3'}/\partial x^3)$ hold, if well defined.

In locally rotationally symmetric dust-filled spaces, the Killing vector $Z = Z^i\partial/\partial x^i$ of the metric (4.22) obeys Eqs. (4.40) with the tetrad (4.17). General solutions are

$$
\begin{aligned}
Z^0 &= cB_0 - \frac{dZ^2}{dx^3}\left(\int f^2(x^2)y(x^2)\,dx^2\right),\\
Z^1 &= CB_1 - \frac{dZ^2}{dx^3}\left(\int f^2(x^2)h(x^2)\,dx^2\right),\\
Z^2 &= A\cos(\alpha^{\frac{1}{2}}x^3) + B_2\sin(\alpha^{\frac{1}{2}}x^3),\\
Z^3 &= B_3 - \frac{dZ^2}{dx^3}\frac{1}{\alpha}\frac{1}{f(x^2)}\frac{df(x^2)}{dx^2},
\end{aligned}
\tag{A7}
$$

where $\alpha = |K|$ if $K \neq 0$, and $\alpha = 1$ if $K = 0$; B_i, A are constants; $f(x^2)$, $y(x^2)$ and $h(x^2)$ are given in (4.21b).