JOURNAL OF
MATHEMATICAL
PHYSICS

VOLUME 8, NUMBER 4 APRIL 1967

Energy Tensor of the Null Electromagnetic Field

P. C. BARTRUM
"Longmead," Kitsbury Terrace, Berkhamsted, Hertfordshire, England

(Received 31 May 1966)

The differential conditions are obtained which must be satisfied by the energy tensor of a null, source-free electromagnetic field. At the same time a method is indicated of determining the electromagnetic field when the energy tensor satisfying the necessary conditions is given.

1. INTRODUCTION

GIVEN a symmetrical tensor field, $T_{\mu\nu}$, we ask ourselves two questions: (1) What are the conditions that this could be the energy tensor of a source-free electromagnetic field? (2) If the conditions are satisfied, how can the electromagnetic field be determined?

For the general case, when the electromagnetic field is not null, these questions have been answered by Rainich[1] and later again by Misner and Wheeler.[2] However, the case of a null electromagnetic field is substantially different, as pointed out by Witten,[3] and no answers have been given for this case. The purpose of this paper is to answer the questions for the case of a null field. In effect the two questions are considered together, the process of seeking the electromagnetic field giving the necessary conditions that it should exist.

We assume the signature of the metric to be $+, +, +, -$. $g_{\mu\nu}$ is the metric tensor, and g is the determinant $|g_{\mu\nu}|$. Ordinary differentiation with respect to x^μ is indicated by a comma and covariant differentiation by a semicolon.

$\mathsf{f}_{\mu\nu}$ is the antisymmetric tensor describing the electromagnetic field, and its dual is defined as

$$*\mathsf{f}^{\mu\nu} = \tfrac{1}{2}(-g)^{-\frac{1}{2}}\epsilon^{\mu\nu\alpha\beta}\mathsf{f}_{\alpha\beta}, \tag{1.1}$$

where $\epsilon^{\mu\nu\alpha\beta}$ is the numerical antisymmetric tensor of weight $+1$. Thus

$$*\mathsf{f}_{\mu\nu} = -\tfrac{1}{2}(-g)^{\frac{1}{2}}\eta_{\mu\nu\alpha\beta}\mathsf{f}^{\alpha\beta}, \tag{1.2}$$

where $\eta_{\mu\nu\alpha\beta}$ is the numerical antisymmetric tensor of weight -1.

Two invariants can be formed from the tensor $\mathsf{f}_{\mu\nu}$,

$$I_1 = \tfrac{1}{2}\mathsf{f}_{\alpha\beta}\mathsf{f}^{\alpha\beta} = -\tfrac{1}{2}*\mathsf{f}_{\alpha\beta}*\mathsf{f}^{\alpha\beta}, \tag{1.3}$$

$$I_2 = \tfrac{1}{2}\mathsf{f}_{\alpha\beta}*\mathsf{f}^{\alpha\beta}. \tag{1.4}$$

The electromagnetic energy tensor is

$$\begin{aligned}\mathsf{T}_{\mu\nu} &= \mathsf{f}_{\mu\alpha}\mathsf{f}^{\alpha}_{\nu} - \tfrac{1}{4}\mathsf{g}_{\mu\nu}\mathsf{f}_{\alpha\beta}\mathsf{f}^{\alpha\beta} \\ &= \tfrac{1}{2}(\mathsf{f}_{\mu\alpha}\mathsf{f}^{\alpha}_{\nu} + *\mathsf{f}_{\mu\alpha}*\mathsf{f}^{\alpha}_{\nu}).\end{aligned} \tag{1.5}$$

From (1.5) we find

$$\mathsf{T}^{\alpha}_{\alpha} = 0, \tag{1.6}$$

$$\mathsf{T}_{\mu\alpha}\mathsf{T}^{\alpha}_{\nu} = \tfrac{1}{4}\mathsf{g}_{\mu\nu}\mathsf{T}_{\alpha\beta}\mathsf{T}^{\alpha\beta} = \tfrac{1}{4}\mathsf{g}_{\mu\nu}T^2, \tag{1.7}$$

where

$$T^2 = \mathsf{T}_{\alpha\beta}\mathsf{T}^{\alpha\beta} = I_1^2 + I_2^2. \tag{1.8}$$

Equations (1.6) and (1.7) are two algebraic conditions which must be satisfied by the tensor $\mathsf{T}_{\mu\nu}$. The third

[1] G. Y. Rainich, Trans. Am. Math. Soc. **27**, 106 (1925).
[2] C. W. Misner and J. A. Wheeler, Ann. Phys. (Paris) **2**, 525 (1957).
[3] L. Witten, in *Gravitation: An Introduction to Current Research*, L. Witten, Ed., (John Wiley & Sons, Inc., New York, 1962). Chap. 9.

algebraic condition is that in local Minkowski coordinates at any point, T_{44} must be positive. In tensor form this may be written as

$$T_{\alpha\beta}v^{\alpha}v^{\beta} \geq 0, \tag{1.9}$$

where v^{μ} is any timelike vector.

2. THE NULL CASE

The electromagnetic field is said to be "null" when $T = 0$, i.e., when

$$I_1 = I_2 = 0. \tag{2.1}$$

In this case (1.7) becomes

$$T_{\mu\alpha}T^{\alpha}_{\nu} = 0 \tag{2.2}$$

and we may write

$$T_{\mu\nu} = C_{\mu}C_{\nu}. \tag{2.3}$$

Thus (1.9) is automatically satisfied.

By (1.6) it is seen that

$$C_{\alpha}C^{\alpha} = 0, \tag{2.4}$$

i.e., C_{μ} is a null vector. It is completely determined by Eq. (2.3). The sign is immaterial.

Let G_{μ} be any unit vector normal to C_{μ}. Since it is normal to a null vector it must be spacelike, i.e., $G_{\alpha}G^{\alpha} = +1$. The tensor

$$f_{\mu\nu} = C_{\mu}G_{\nu} - C_{\nu}G_{\mu} \tag{2.5}$$

is a suitable tensor to use as an electromagnetic field tensor, for if we substitute in (1.5) it gives (2.3). The dual may be written as

$$*f_{\mu\nu} = C_{\mu}H_{\nu} - C_{\nu}H_{\mu}, \tag{2.6}$$

where H_{μ} is a unit spacelike vector normal to C_{μ} and G_{μ}. Other suitable tensors can be found by the duality rotation

$$f'_{\mu\nu} = f_{\mu\nu}\cos\alpha + *f_{\mu\nu}\sin\alpha, \tag{2.7}$$

where α is an invariant, but it is easily seen that this is equivalent to choosing new unit vectors

$$\begin{aligned} G'_{\mu} &= G_{\mu}\cos\alpha + H_{\mu}\sin\alpha, \\ H'_{\mu} &= H_{\mu}\cos\alpha - G_{\mu}\sin\alpha. \end{aligned} \tag{2.8}$$

3. DIFFERENTIAL CONDITIONS

If $f_{\mu\nu}$ is of the form (2.5) the algebraic conditions (1.6) and (2.2) are satisfied. But $f_{\mu\nu}$ also has to satisfy the Maxwell equations for a source-free electromagnetic field. These are

$$f^{\alpha}_{\mu;\alpha} = 0, \quad *f^{\alpha}_{\mu;\alpha} = 0, \tag{3.1}$$

which may also be written, respectively,

$$\begin{aligned} *f_{\mu\nu;\sigma} + *f_{\nu\sigma;\mu} + *f_{\sigma\mu;\nu} &= 0, \\ f_{\mu\nu;\sigma} + f_{\nu\sigma;\mu} + f_{\sigma\mu;\nu} &= 0. \end{aligned} \tag{3.2}$$

In terms of the vectors C_{μ}, G_{μ}, H_{μ}, (3.1) gives

$$C_{\mu}G^{\alpha}_{;\alpha} - G_{\mu}C^{\alpha}_{;\alpha} + G^{\alpha}C_{\mu;\alpha} - C^{\alpha}G_{\mu;\alpha} = 0, \tag{3.3}$$

whence

$$C^{\alpha}C^{\beta}G_{\alpha;\beta} = 0, \quad G^{\alpha}G^{\beta}C_{\alpha;\beta} = C^{\alpha}_{;\alpha} \tag{3.4}$$

with similar results for H_{μ}.

From (3.2) we get

$$\begin{aligned} G_{\mu}(C_{\sigma;\nu} - C_{\nu;\sigma}) + G_{\nu}(C_{\mu;\sigma} - C_{\sigma;\mu}) + G_{\sigma}(C_{\nu;\mu} - C_{\mu;\nu}) \\ + C_{\mu}(G_{\nu;\sigma} - G_{\sigma;\nu}) + C_{\nu}(G_{\sigma;\mu} - G_{\mu;\sigma}) \\ + C_{\sigma}(G_{\mu;\nu} - G_{\nu;\mu}) = 0. \end{aligned} \tag{3.5}$$

Multiply by $G^{\sigma}C^{\nu}$ and we get, using (3.4),

$$C_{\mu;\alpha}C^{\alpha} = -C_{\mu}C^{\alpha}_{;\alpha}. \tag{3.6}$$

This may be written $(C^{\mu}C^{\alpha})_{;\alpha} = 0$, whence $T^{\mu\alpha}_{;\alpha} = 0$. This is the equation of conservation which, as is well known, derives from Maxwell's equations. Equation (3.6) also shows that the vector field C_{μ} defines a congruence of null geodesics. Equation (3.6) consists of three independent differential conditions on C_{μ}, i.e., on $T_{\mu\nu}$.

Again, multiply (3.5) by C^{σ} and use (3.6), and we get

$$\begin{aligned} C_{\mu}[C^{\alpha}(G_{\nu;\alpha} - G_{\alpha;\nu}) - G_{\nu}C^{\alpha}_{;\alpha}] \\ = C_{\nu}[C^{\alpha}(G_{\mu;\alpha} - G_{\alpha;\mu}) - G_{\mu}C^{\alpha}_{;\alpha}], \end{aligned}$$

whence

$$C^{\alpha}(G_{\mu;\alpha} - G_{\alpha;\mu}) - G_{\mu}C^{\alpha}_{;\alpha} = \lambda_g C_{\mu}, \tag{3.7}$$

where λ_g is some invariant.

Similarly, we find

$$C^{\alpha}(H_{\mu;\alpha} - H_{\alpha;\mu}) - H_{\mu}C^{\alpha}_{;\alpha} = \lambda_h C_{\mu},$$

where λ_h is some invariant.

Eliminate $C^{\alpha}G_{\mu;\alpha}$ between (3.3) and (3.7), we find

$$G^{\alpha}(C_{\mu;\alpha} + C_{\alpha;\mu} - 2g_{\mu\alpha}C^{\beta}_{;\beta}) = C_{\mu}(\lambda_g - G^{\alpha}_{;\alpha}). \tag{3.8}$$

Similarly, with H_{μ} for G_{μ} and λ_h for λ_g.

4. CONDITIONS THAT (3.8) ARE CONSISTENT

Write

$$E_{\mu\nu} \equiv C_{\mu;\nu} + C_{\nu;\mu} - 2g_{\mu\nu}C^{\alpha}_{;\alpha}, \tag{4.1}$$

and let $e^{\mu\nu}$ be the minor of $E_{\mu\nu}$ in the determinant $|E_{\mu\nu}|$, so that

$$E_{\alpha\mu}e^{\alpha\nu} = \delta^{\nu}_{\mu}|E_{\alpha\beta}|. \tag{4.2}$$

Multiply (3.8) by $e^{\mu\nu}$ and we have

$$G^{\nu}|E_{\alpha\beta}| = (\lambda_g - G^{\alpha}_{;\alpha})C_{\beta}e^{\beta\nu}.$$

Using (3.6) and (A2) of the Appendix, we find that

$$C_{\beta}e^{\beta\nu} = gJC^{\nu},$$

where J is an invariant. So that

$$G^\nu |E_{\alpha\beta}| = (\lambda_g - G^\alpha_{;\alpha}) g J C^\nu.$$

Multiply by G_ν and we get

$$|E_{\alpha\beta}| = 0. \tag{4.3}$$

This is another differential condition on C_μ, i.e., on $T_{\mu\nu}$. We show that it is, effectively, two conditions.

From (4.3) it follows that there must be a vector A^μ such that

$$A^\alpha(C_{\mu;\alpha} + C_{\alpha;\mu} - 2g_{\mu\alpha}C^\beta_{;\beta}) = 0. \tag{4.4}$$

We also have by (3.6)

$$C^\alpha(C_{\mu;\alpha} + C_{\alpha;\mu} + g_{\mu\alpha}C^\beta_{;\beta}) = 0. \tag{4.5}$$

Now, consider the process of finding the principal directions and principal invariants determined by the tensor $C_{\mu;\nu} + C_{\nu;\mu}$.

We solve the equation

$$|C_{\mu;\nu} + C_{\nu,\mu} - Ug_{\mu\nu}| = 0 \tag{4.6}$$

for U, and find four roots U_1, U_2, U_3, U_4. By (4.4) and (4.5) two of these roots are

$$U_1 = 2C^\alpha_{;\alpha}, \quad U_3 = -C^\alpha_{;\alpha}, \tag{4.7}$$

and the corresponding vectors are A^μ and C^μ. The sum of the roots, $U_1 + U_2 + U_3 + U_4$, is equal to minus the coefficient of U^3 in the expansion of (4.6). By (A1) in Appendix A this is $2C^\alpha_{;\alpha}$, so that

$$U_2 + U_4 = C^\alpha_{;\alpha}.$$

From this point we assume that $C^\alpha_{;\alpha} \neq 0$ in the domain considered. When $C^\alpha_{;\alpha} = 0$ we have the "null–null" case which is considered separately in Sec. 10.

The fact that the vector corresponding to U_3 is a null vector leads to only one possibility, namely,

$$U_1 = U_2 = 2C^\alpha_{;\alpha}, \quad U_3 = U_4 = -C^\alpha_{;\alpha}. \tag{4.8}$$

However, the elementary divisors corresponding to the pairs of roots may be simple or multiple. (See Eisenhart,[4] pp. 111–112.) In the present case there are only two practical possibilities: (1) All the elementary divisors are simple, in which case $U_1 = U_2$ correspond to a pencil of vectors all of which satisfy (4.4), and $U_3 = U_4$ correspond to a pencil of vectors, all of which satisfy (4.5). The two pencils are normal to each other. (2) The elementary divisors are $(U - U_1)$, $(U - U_2)$, and $(U - U_3)^2$. In this case there is a pencil of vectors corresponding to $U_1 = U_2$ and satisfying (4.4), and a single null vector corresponding to $U_3 = U_4$ and satisfying (4.5). This null vector is, of course, C^μ.

[4] L. P. Eisenhart, *Riemannian Geometry*, (Princeton University Press, Princeton, New Jersey, 1926).

5. DETERMINATION OF AN ORTHONORMAL TETRAD

Case 1: Since C^μ appears in the second pencil, the second pencil must consist of spacelike and timelike vectors. The first pencil must therefore consist of spacelike vectors.

Let X^μ and Y^μ be a pair of unit vectors in the second pencil, normal to each other, and suppose that X^μ is spacelike and Y^μ timelike. If we give suitable signs to X^μ and Y^μ, we can arrange that

$$X^\mu - Y^\mu = nC^\mu, \tag{5.1}$$

where n is an invariant. The other null vector in the pencil is then given by

$$X^\mu + Y^\mu = (2/n)S^\mu. \tag{5.2}$$

The choice of invariant in (5.2) is made, for convenience, so that

$$S^\alpha C_\alpha = 1. \tag{5.3}$$

Although X^μ and Y^μ are to some extent arbitrary, the vector S^μ is determined unambiguously. This must be so since there are only two null vectors in a pencil, while the magnitude and sign of S^μ are fixed by (5.3). The orientation of the pencil is determined by

$$X^\mu Y^\nu - X^\nu Y^\mu = C^\mu S^\nu - C^\nu S^\mu. \tag{5.4}$$

We also have

$$X^\mu X^\nu - Y^\mu Y^\nu = C^\mu S^\nu + C^\nu S^\mu. \tag{5.5}$$

In the first pencil we choose a pair of unit spacelike vectors normal to each other, A^μ and B^μ, say. These are to some extent arbitrary, but the quantities $A^\mu A^\nu + B^\mu B^\nu$ are definite, while the quantities $A^\mu B_\nu - A^\nu B^\mu$, which determine the orientation of the pencil, are definite except for sign. We may fix the sign of the latter by the convention

$$\begin{aligned} A_\mu B_\nu - A_\nu B_\mu &= (-g)^{\frac{1}{2}} \eta_{\mu\nu\alpha\beta} X^\alpha Y^\beta \\ &= (-g)^{\frac{1}{2}} \eta_{\mu\nu\alpha\beta} C^\alpha S^\beta. \end{aligned} \tag{5.6}$$

A_μ and B_μ are arbitrary to the extent that they may be replaced by

$$\begin{aligned} A'_\mu &= A_\mu \cos\alpha + B_\mu \sin\alpha, \\ B'_\mu &= B_\mu \cos\alpha - A_\mu \sin\alpha, \end{aligned} \tag{5.7}$$

where α is any invariant. If there is any difficulty in finding such a pair, they can be determined by choosing any spacelike vector z_μ not in the pencil of

$\mathbf{C}_\mu$ and $\mathbf{S}_\mu$, i.e., linearly independent of $\mathbf{C}_\mu$ and $\mathbf{S}_\mu$. We may then take

$$\begin{aligned}\mathbf{A}_\mu &= (-g)^{\frac{1}{2}}\eta_{\mu\alpha\beta\gamma}\mathbf{z}^\alpha\mathbf{C}^\beta\mathbf{S}^\gamma[\mathbf{z}^\epsilon\mathbf{z}_\epsilon - 2(\mathbf{z}^\epsilon\mathbf{C}_\epsilon)(\mathbf{z}^\rho\mathbf{S}_\rho)]^{-\frac{1}{2}},\\ \mathbf{B}_\mu &= -(-g)^{\frac{1}{2}}\eta_{\mu\alpha\beta\gamma}\mathbf{A}^\alpha\mathbf{C}^\beta\mathbf{S}^\gamma.\end{aligned} \tag{5.8}$$

Case 2: The pencil of vectors is normal to $\mathbf{C}^\mu$, and they are therefore spacelike. We take any pair of unit vectors $\mathbf{A}^\mu$ and $\mathbf{B}^\mu$ in this pencil, normal to each other. There is a pencil of vectors normal to the former pencil, of which $\mathbf{C}^\mu$ is one of the null vectors. Let $\mathbf{S}^\mu$ be the other null vector in this second pencil. Then we may define pairs of unit vectors $\mathbf{X}^\mu$ and $\mathbf{Y}^\mu$ normal to each other in this pencil by (5.1) and (5.2). As in Case 1 we fix the relative signs of $\mathbf{A}^\mu$ and $\mathbf{B}^\mu$ by the convention (5.6).

Cases 1 and 2: We now treat the two cases together. The unit vectors $\mathbf{A}^\mu$, $\mathbf{B}^\mu$, $\mathbf{X}^\mu$, and $\mathbf{Y}^\mu$ form an orthonormal tetrad of which $\mathbf{Y}^\mu$ is the timelike member. Hence by (B1) in Appendix B we have

$$\begin{aligned}g_{\mu\nu} &= \mathbf{A}_\mu\mathbf{A}_\nu + \mathbf{B}_\mu\mathbf{B}_\nu + \mathbf{X}_\mu\mathbf{X}_\nu - \mathbf{Y}_\mu\mathbf{Y}_\nu\\ &= \mathbf{A}_\mu\mathbf{A}_\nu + \mathbf{B}_\mu\mathbf{B}_\nu + \mathbf{C}_\mu\mathbf{S}_\nu + \mathbf{C}_\nu\mathbf{S}_\mu.\end{aligned} \tag{5.9}$$

We may also write

$$\begin{aligned}\mathbf{C}_{\mu;\nu} + \mathbf{C}_{\nu;\mu} &= 2(\mathbf{A}_\mu\mathbf{A}_\nu + \mathbf{B}_\mu\mathbf{B}_\nu)\mathbf{C}^\alpha_{;\alpha}\\ &\quad - (\mathbf{X}_\mu\mathbf{X}_\nu - \mathbf{Y}_\mu\mathbf{Y}_\nu)\mathbf{C}^\alpha_{;\alpha} + q\mathbf{C}_\mu\mathbf{C}_\nu\\ &= 2(\mathbf{A}_\mu\mathbf{A}_\nu + \mathbf{B}_\mu\mathbf{B}_\nu)\mathbf{C}^\alpha_{;\alpha}\\ &\quad - (\mathbf{C}_\mu\mathbf{S}_\nu + \mathbf{C}_\nu\mathbf{S}_\mu)\mathbf{C}^\alpha_{;\alpha} + q\mathbf{C}_\mu\mathbf{C}_\nu,\end{aligned} \tag{5.10}$$

where q is an invariant which vanishes in Case 1.

It is seen that Case 2 is slightly more general than Case 1. In fact, Case 1 corresponds to Case 2 with $q = 0$.

Eliminate $(\mathbf{A}_\mu\mathbf{A}_\nu + \mathbf{B}_\mu\mathbf{B}_\nu)$ between (5.9) and (5.10) and we get

$$\begin{aligned}\mathsf{E}_{\mu\nu} &\equiv \mathbf{C}_{\mu;\nu} + \mathbf{C}_{\nu;\mu} - 2g_{\mu\nu}\mathbf{C}^\alpha_{;\alpha}\\ &= -3(\mathbf{C}_\mu\mathbf{S}_\nu + \mathbf{C}_\nu\mathbf{S}_\mu)\mathbf{C}^\alpha_{;\alpha} + q\mathbf{C}_\mu\mathbf{C}_\nu,\end{aligned} \tag{5.11}$$

whence we deduce

$$\mathbf{C}_\sigma\mathsf{E}_{\mu\nu} - \mathbf{C}_\mu\mathsf{E}_{\nu\sigma} - \mathbf{C}_\nu\mathsf{E}_{\mu\sigma} = (6\mathbf{S}_\sigma\mathbf{C}^\alpha_{;\alpha} - q\mathbf{C}_\sigma)\mathbf{C}_\mu\mathbf{C}_\nu. \tag{5.12}$$

Finally, eliminating S_μ, we have

$$\mathbf{C}_\lambda\mathbf{C}_\sigma\mathsf{E}_{\mu\nu} + \mathbf{C}_\mu\mathbf{C}_\nu\mathsf{E}_{\lambda\sigma} = \mathbf{C}_\lambda\mathbf{C}_\mu\mathsf{E}_{\nu\sigma} + \mathbf{C}_\nu\mathbf{C}_\sigma\mathsf{E}_{\lambda\mu}. \tag{5.13}$$

These equations are in terms of $\mathbf{C}_\mu$ and its derivatives only, and therefore give necessary differential conditions which must be satisfied by $\mathbf{C}_\mu$ in order that (3.8) should be consistent. *A fortiori* they are differential equations which must be satisfied by $\mathsf{T}_{\mu\nu}$.

We also deduce from (5.11)

$$\mathsf{E}_{\mu\alpha}\mathbf{E}^\alpha_\nu = -3\mathbf{C}^\alpha_{;\alpha}(\mathsf{E}_{\mu\nu} + q\mathbf{C}_\mu\mathbf{C}_\nu). \tag{5.14}$$

$\mathbf{S}_\mu$ and q may be determined in terms of $\mathbf{C}_\mu$ as follows. Provided (5.13) holds we may determine from (5.12) the vector

$$\mathbf{R}_\sigma \equiv 6\mathbf{S}_\sigma\mathbf{C}^\alpha_{;\alpha} - q\mathbf{C}_\sigma. \tag{5.15}$$

We then have

$$\begin{aligned}q &= -\mathbf{R}^\alpha\mathbf{R}_\alpha/12\mathbf{C}^\beta_{;\beta},\\ \mathbf{S}_\mu &= (\mathbf{R}_\mu + q\mathbf{C}_\mu)/6\mathbf{C}^\alpha_{;\alpha}.\end{aligned} \tag{5.16}$$

If $q = 0$ a further condition is satisfied by $\mathbf{C}_\mu$, for by (5.14) we then have

$$\mathsf{E}_{\mu\alpha}\mathbf{E}^\alpha_\nu = -3\mathbf{C}^\alpha_{;\alpha}\mathsf{E}_{\mu\nu}. \tag{5.17}$$

It is found that (5.13) consists of 30 different equations of which 10 are identities. Of the remaining 20 only five are independent. Equation (5.14) provides no further conditions except when $q = 0$, and then provides one further condition equivalent to $q = 0$. But this is not a necessary condition. The three conditions (3.6) are deducible from (5.13), so that (5.13) effectively contains two more conditions not contained in (3.6). This at first is surprising in view of the fact that the additional two conditions must be equivalent to (4.3). The fact is that by virtue of (3.6) $|\mathsf{E}_{\mu\nu}|$ reduces to the form

$$|\mathsf{E}_{\mu\nu}| = -g(\mathbf{C}^\alpha_{;\alpha})^2(P^2 + Q^2).$$

Thus if $\mathbf{C}^\alpha_{;\alpha} \neq 0$ we must have $P = 0$, $Q = 0$, i.e., two conditions.

6. FORM OF G_μ AND H_μ

If conditions (5.13) are satisfied, then Eqs. (3.8) are satisfied by *any* vector $\mathbf{G}_\mu$ normal to $\mathbf{C}_\mu$.

Any vector normal to $\mathbf{C}_\mu$ must be of the form

$$\mathbf{G}_\mu = \mathbf{A}_\mu\cos\theta + \mathbf{B}_\mu\sin\theta + p_g\mathbf{C}_\mu, \tag{6.1}$$

where θ and p_g are invariants. p_g clearly has no significance, since it disappears in the expression $\mathbf{C}_\mu\mathbf{G}_\nu - \mathbf{C}_\nu\mathbf{G}_\mu$. Substituting (6.1) in (3.8) and using (5.11), we find that (3.8) is satisfied provided

$$\lambda_g = \mathbf{G}^\alpha_{;\alpha} - 3p_g\mathbf{C}^\alpha_{;\alpha}.$$

This in effect fixes λ_g which is otherwise undetermined. Similar results hold for $\mathbf{H}_\mu$ which may be written

$$\mathbf{H}_\mu = \mathbf{B}_\mu\cos\theta - \mathbf{A}_\mu\sin\theta + p_h\mathbf{C}_\mu. \tag{6.2}$$

It does not follow that $\mathbf{G}_\mu$ and $\mathbf{H}_\mu$ satisfy the differential equations (3.3) and its pair. We look at this further in Sec. 8 and find that, provided certain integrability conditions are satisfied, (3.3) and its pair in general determine θ to within a constant. First of all we deduce some required formulas in Sec. 7.

7. MISCELLANEOUS FORMULAS

The Invariant N

Let

$$\mathbf{Z}^\mu = (-g)^{-\frac{1}{2}}\boldsymbol{\epsilon}^{\mu\alpha\beta\gamma}\mathbf{C}_\alpha\mathbf{C}_{\beta;\gamma},$$

then we find

$$-(-g)^{\frac{1}{2}}\boldsymbol{\eta}_{\mu\nu\beta\alpha}\mathbf{C}^\beta\mathbf{Z}^\alpha = \mathbf{C}_\mu.\mathbf{C}_{\nu;\beta}\mathbf{C}^\beta - \mathbf{C}_\nu.\mathbf{C}_{\mu;\beta}\mathbf{C}^\beta = 0 \quad \text{by (3.6).}$$

Thus $\mathbf{C}^\mu\mathbf{Z}^\nu = \mathbf{C}^\nu\mathbf{Z}^\mu$ so that $\mathbf{Z}^\mu = N\mathbf{C}^\mu$, where N is an invariant. Thus we have the important result

$$(-g)^{-\frac{1}{2}}\boldsymbol{\epsilon}^{\mu\alpha\beta\gamma}\mathbf{C}_\alpha\mathbf{C}_{\beta;\gamma} = N\mathbf{C}^\mu. \tag{7.1}$$

In what follows we make considerable use of the orthonormal tetrad, $\mathbf{A}_\mu$, $\mathbf{B}_\mu$, $\mathbf{X}_\mu$, $\mathbf{Y}_\mu$ or rather the equivalent set of vectors $\mathbf{A}_\mu$, $\mathbf{B}_\mu$, $\mathbf{C}_\mu$, $\mathbf{S}_\mu$.

The Form of $\mathsf{C}_{\mu;\nu} - \mathsf{C}_{\nu;\mu}$

If we express $\mathsf{C}_{\mu;\nu} - \mathsf{C}_{\nu;\mu}$ in terms of its six tetradic components [see Appendix, (B3)], we find that by virtue of (3.6) and (7.1) two of these components vanish and we are left with

$$\mathsf{C}_{\mu;\nu} - \mathsf{C}_{\nu;\mu} = N(\mathbf{A}_\mu\mathbf{B}_\nu - \mathbf{A}_\nu\mathbf{B}_\mu) - \mathsf{C}^\alpha_{;\alpha}(\mathbf{C}_\mu\mathbf{S}_\nu - \mathbf{C}_\nu\mathbf{S}_\mu) + \rho(\mathbf{C}_\mu\mathbf{A}_\nu - \mathbf{C}_\nu\mathbf{A}_\mu) + \sigma(\mathbf{C}_\mu\mathbf{B}_\nu - \mathbf{C}_\nu\mathbf{B}_\mu), \tag{7.2}$$

or, equivalently

$$(-g)^{-\frac{1}{2}}\boldsymbol{\epsilon}^{\mu\nu\alpha\beta}\mathsf{C}_{\alpha;\beta} = N(\mathbf{C}^\mu\mathbf{S}^\nu - \mathbf{C}^\nu\mathbf{S}^\mu) + \mathsf{C}^\alpha_{;\alpha}(\mathbf{A}^\mu\mathbf{B}^\nu - \mathbf{A}^\nu\mathbf{B}^\mu) + \rho(\mathbf{C}^\mu\mathbf{B}^\nu - \mathbf{C}^\nu\mathbf{B}^\mu) - \sigma(\mathbf{C}^\mu\mathbf{A}^\nu - \mathbf{C}^\nu\mathbf{A}^\mu). \tag{7.3}$$

ρ and σ are invariants which depend on the choice of $\mathbf{A}_\mu$ and $\mathbf{B}_\mu$, but $(\rho^2 + \sigma^2)$ is independent of the choice, and depends only on $\mathbf{C}_\mu$.

By (5.11) and (7.2) we may now write $\mathsf{C}_{\mu;\nu}$ in terms of the tetrad and invariants:

$$\mathsf{C}_{\mu;\nu} = \mathsf{C}^\alpha_{;\alpha}(\mathsf{g}_{\mu\nu} - 2\mathbf{C}_\mu\mathbf{S}_\nu - \mathbf{C}_\nu\mathbf{S}_\mu) + \tfrac{1}{2}N(\mathbf{A}_\mu\mathbf{B}_\nu - \mathbf{A}_\nu\mathbf{B}_\mu) + \tfrac{1}{2}\rho(\mathbf{C}_\mu\mathbf{A}_\nu - \mathbf{C}_\nu\mathbf{A}_\mu) + \tfrac{1}{2}\sigma(\mathbf{C}_\mu\mathbf{B}_\nu - \mathbf{C}_\nu\mathbf{B}_\mu) + \tfrac{1}{2}q\mathbf{C}_\mu\mathbf{C}_\nu. \tag{7.4}$$

From (7.4) we deduce

$$\mathsf{C}^\alpha_{;\beta}\mathsf{C}^\beta_{;\alpha} = 3(\mathsf{C}^\alpha_{;\alpha})^2 - \tfrac{1}{2}N^2, \qquad \mathsf{C}^\alpha_{;\beta}\mathsf{C}^\beta_{\alpha;} = 2(\mathsf{C}^\alpha_{;\alpha})^2 + \tfrac{1}{2}N^2. \tag{7.5}$$

These give N^2 explicitly in terms of $\mathbf{C}_\mu$.

Further Properties of N

By (7.2) and (7.3) we find

$$(-g)^{-\frac{1}{2}}\boldsymbol{\epsilon}^{\mu\nu\lambda\sigma}\mathsf{C}_{\mu;\nu}\mathsf{C}_{\lambda;\sigma} = 2N\mathsf{C}^\alpha_{;\alpha}. \tag{7.6}$$

Differentiate (7.1), and we get

$$(N\mathbf{C}^\mu)_{;\mu} = (-g)^{-\frac{1}{2}}\boldsymbol{\epsilon}^{\mu\alpha\beta\gamma}\mathsf{C}_{\alpha;\mu}\mathsf{C}_{\beta;\gamma} + (-g)^{-\frac{1}{2}}\boldsymbol{\epsilon}^{\mu\alpha\beta\gamma}\mathbf{C}_\alpha\mathsf{C}_{\beta;\gamma;\mu}.$$

The second term on the right-hand side vanishes by the properties of the Riemann tensor. Using (7.6) we get

$$(N\mathbf{C}^\alpha)_{;\alpha} = -2N\mathsf{C}^\alpha_{;\alpha},$$

whence

$$N_{,\alpha}\mathbf{C}^\alpha = -3N\mathsf{C}^\alpha_{;\alpha}. \tag{7.7}$$

It follows that

$$(N^{\frac{1}{3}}\mathbf{C}^\alpha)_{;\alpha} = 0, \tag{7.8}$$

so that $N^{\frac{1}{3}}$ is proportional to the "density of rays" in the congruence of null geodesics defined by $\mathbf{C}_\mu$. We also find, by (7.7) and (3.6),

$$\mathbf{C}^\alpha(N^{-\frac{1}{3}}\mathbf{C}^\mu)_{;\alpha} = 0. \tag{7.9}$$

These are the conditions for a "geodesic congruence" in their canonical form.

Formulas Involving $\mathsf{A}_{\mu;\sigma}$, $\mathsf{B}_{\mu;\sigma}$

Differentiate (5.9) covariantly with respect to x^σ and multiply by $\mathbf{A}^\nu$, we get

$$\mathsf{A}_{\mu;\sigma} = -\mathsf{B}_{\mu}.\mathbf{A}^\alpha\mathsf{B}_{\alpha;\sigma} - \mathsf{C}_{\mu}.\mathbf{A}^\alpha\mathsf{S}_{\alpha;\sigma} - \mathsf{S}_{\mu}.\mathbf{A}^\alpha\mathsf{C}_{\alpha;\sigma}. \tag{7.10}$$

Similarly,

$$\mathsf{B}_{\mu;\sigma} = \mathsf{A}_{\mu}.\mathbf{A}^\alpha\mathsf{B}_{\alpha;\sigma} - \mathsf{C}_{\mu}.\mathbf{B}^\alpha\mathsf{S}_{\alpha;\sigma} - \mathsf{S}_{\mu}.\mathbf{B}^\alpha\mathsf{C}_{\alpha;\sigma}. \tag{7.11}$$

By (7.10) and (7.11) and using (5.6) we deduce

$$\mathsf{A}^\alpha_{;\mu}\mathsf{B}_{\alpha;\nu} - \mathsf{A}^\alpha_{;\nu}\mathsf{B}_{\alpha;\mu} = (\mathbf{A}^\alpha\mathbf{B}^\beta - \mathbf{A}^\beta\mathbf{B}^\alpha)(\mathsf{C}_{\alpha;\mu}\mathsf{S}_{\beta;\nu} - \mathsf{C}_{\alpha;\nu}\mathsf{S}_{\beta;\mu}) = -(-g)^{-\frac{1}{2}}\boldsymbol{\epsilon}^{\alpha\beta\gamma\epsilon}\mathbf{C}_\gamma\mathbf{S}_\epsilon(\mathsf{C}_{\alpha;\mu}\mathsf{S}_{\beta;\nu} - \mathsf{C}_{\alpha;\nu}\mathsf{S}_{\beta;\mu}). \tag{7.12}$$

8. THE DETERMINATION OF θ

Having chosen a specific pair of unit spacelike vectors $\mathbf{A}_\mu$ and $\mathbf{B}_\mu$ normal to each other and satisfying (4.4), we take $\mathbf{G}_\mu$ and $\mathbf{H}_\mu$ to be the expressions (6.1) and (6.2), and substitute in (3.3) and its pair. We get

$$(\mathbf{C}^\mu\mathbf{A}^\alpha - \mathbf{C}^\alpha\mathbf{A}^\mu)_{;\alpha} + (\mathbf{C}^\mu\mathbf{B}^\alpha - \mathbf{C}^\alpha\mathbf{B}^\mu)\theta_{,\alpha} = 0, \tag{8.1}$$

$$(\mathbf{C}^\mu\mathbf{B}^\alpha - \mathbf{C}^\alpha\mathbf{B}^\mu)_{;\alpha} - (\mathbf{C}^\mu\mathbf{A}^\alpha - \mathbf{C}^\alpha\mathbf{A}^\mu)\theta_{,\alpha} = 0. \tag{8.2}$$

Multiply (8.1) in turn by $\mathbf{B}_\mu$ and $\mathbf{S}_\mu$, we get

$$\mathbf{C}^\alpha\theta_{,\alpha} = \mathbf{B}_\mu(\mathbf{C}^\mu\mathbf{A}^\alpha - \mathbf{C}^\alpha\mathbf{A}^\mu)_{;\alpha} = \mathbf{A}^\alpha\mathbf{B}^\beta\mathsf{C}_{\beta;\alpha} - \mathbf{C}^\alpha\mathbf{B}^\beta\mathsf{A}_{\beta;\alpha} = -\tfrac{1}{2}N + \mathbf{C}^\alpha\cdot\mathbf{A}^\beta\mathsf{B}_{\beta;\alpha} \quad \text{by (7.4),} \tag{8.3}$$

$$\mathbf{B}^\alpha\theta_{,\alpha} = -\mathbf{S}_\mu(\mathbf{C}^\mu\mathbf{A}^\alpha - \mathbf{C}^\alpha\mathbf{A}^\mu)_{;\alpha} = -\mathsf{A}^\alpha_{;\alpha} - \mathbf{A}^\alpha\mathbf{S}^\beta\mathsf{C}_{\beta;\alpha} + \mathbf{C}^\alpha\mathbf{S}^\beta\mathsf{A}_{\beta;\alpha} = -\rho + \mathbf{B}^\alpha\cdot\mathbf{A}^\beta\mathsf{B}_{\beta;\alpha}, \tag{8.4}$$

using (7.10) and (7.4). Similarly, multiply (8.2) in

turn by $\mathbf{A}_\mu$ and $\mathbf{S}_\mu$, we get (8.3) again, and as in (8.4) we find

$$\mathbf{A}^\alpha\theta_{,\alpha} = \sigma + \mathbf{A}^\alpha \cdot \mathbf{A}^\beta\mathbf{B}_{\beta;\alpha}. \tag{8.5}$$

By Appendix (B3) we may write

$$\begin{aligned}\theta_{,\mu} &= (\mathbf{A}^\alpha\theta_{,\alpha})\mathbf{A}_\mu + (\mathbf{B}^\alpha\theta_{,\alpha})\mathbf{B}_\mu + (\mathbf{X}^\alpha\theta_{,\alpha})\mathbf{X}_\mu - (\mathbf{Y}^\alpha\theta_{,\alpha})\mathbf{Y}_\mu \\ &= (\mathbf{A}^\alpha\theta_{,\alpha})\mathbf{A}_\mu + (\mathbf{B}^\alpha\theta_{,\alpha})\mathbf{B}_\mu + (\mathbf{C}^\alpha\theta_{,\alpha})\mathbf{S}_\mu + (\mathbf{S}^\alpha\theta_{,\alpha})\mathbf{C}_\mu.\end{aligned} \tag{8.6}$$

Substituting (8.3), (8.4), and (8.5) in (8.6), we get

$$\begin{aligned}\theta_{,\mu} &= (\mathbf{A}^\beta\mathbf{A}_\mu + \mathbf{B}^\beta\mathbf{B}_\mu)\mathbf{A}^\alpha\mathbf{B}_{\alpha;\beta} + \sigma\mathbf{A}_\mu - \rho\mathbf{B}_\mu \\ &\quad - (\tfrac{1}{2}N - \mathbf{C}^\alpha\mathbf{A}^\beta\mathbf{B}_{\beta;\alpha})\mathbf{S}_\mu + (\mathbf{S}^\alpha\theta_{,\alpha})\mathbf{C}_\mu \\ &= \mathbf{A}^\alpha\mathbf{B}_{\alpha;\mu} + \sigma\mathbf{A}_\mu - \rho\mathbf{B}_\mu - \tfrac{1}{2}N\mathbf{S}_\mu \\ &\quad + \mathbf{C}_\mu(\mathbf{S}^\alpha\theta_{,\alpha} - \mathbf{S}^\alpha\mathbf{A}^\beta\mathbf{B}_{\beta;\alpha}), \quad \text{using (5.9).}\end{aligned} \tag{8.7}$$

$\mathbf{S}^\alpha\theta_{,\alpha}$ is undetermined, and we may therefore write

$$\mathbf{S}^\alpha(\theta_{,\alpha} - \mathbf{A}^\beta\mathbf{B}_{\beta;\alpha}) = w, \tag{8.8}$$

where w is an undetermined invariant. Using (7.3) we may write (8.7) in the form

$$\begin{aligned}\theta_{,\mu} = \mathbf{L}_\mu \equiv \mathbf{A}^\alpha\mathbf{B}_{\alpha;\mu} - (-g)^{\frac{1}{2}}\eta_{\mu\alpha\beta\gamma}\mathbf{S}^\alpha\mathbf{C}^{\beta;\gamma} \\ + \tfrac{1}{2}N\mathbf{S}_\mu + w\mathbf{C}_\mu.\end{aligned} \tag{8.9}$$

This gives θ to within a constant provided the right-hand side is integrable. The conditions of integrability are $\mathsf{L}_{\mu;\nu} = \mathsf{L}_{\nu;\mu}$. Carrying out this operation and using (7.12), we may write the conditions wholly in terms of $\mathbf{C}_\mu$, $\mathbf{S}_\mu$, N, and w:

$$\begin{aligned}\mathsf{K}_{\mu\nu} &\equiv (\mathsf{C}_{\alpha;\mu}\mathsf{S}_{\beta;\nu} - \mathsf{C}_{\alpha;\nu}\mathsf{S}_{\beta;\mu} + \tfrac{1}{2}\mathsf{R}_{\alpha\beta\mu\nu})(-g)^{-\frac{1}{2}}\epsilon^{\alpha\beta\gamma\epsilon}\mathbf{C}_\gamma\mathbf{S}_\epsilon \\ &\quad - (-g)^{\frac{1}{2}}\eta_{\mu\nu\alpha\beta}(\mathbf{C}^{\alpha;\beta}\mathbf{S}^\epsilon + \mathbf{C}^{\beta;\epsilon}\mathbf{S}^\alpha + \mathbf{C}^{\epsilon;\alpha}\mathbf{S}^\beta)_{;\epsilon} \\ &\quad + \tfrac{1}{2}(N\mathbf{S}_\mu)_{;\nu} - \tfrac{1}{2}(N\mathbf{S}_\nu)_{;\mu} \\ &= (w\mathbf{C}_\nu)_{;\mu} - (w\mathbf{C}_\mu)_{;\nu}.\end{aligned} \tag{8.10}$$

The invariant w in (8.10) is undetermined. In fact, the six equations (8.10) may be regarded as one to determine w and five conditions of integrability. Denote the left-hand side of (8.10) by $\mathsf{K}_{\mu\nu}$. We may eliminate $w_{,\mu}$ from (8.10) and obtain

$$\begin{aligned}\tfrac{1}{2}(-g)^{-\frac{1}{2}}\epsilon^{\mu\alpha\beta\gamma}\mathbf{C}_\alpha\mathsf{K}_{\beta\gamma} &= w(-g)^{-\frac{1}{2}}\epsilon^{\mu\alpha\beta\gamma}\mathbf{C}_\alpha\mathbf{C}_{\beta;\gamma} \\ &= wN\mathbf{C}^\mu \quad \text{by (7.1).}\end{aligned} \tag{8.11}$$

These are three independent equations. If $N \neq 0$ they may be regarded as one to determine w and two conditions of integrability. If these are satisfied and we substitute this value of w in (8.10), we then have three more conditions of integrability. Whichever way we look at it, there are five conditions of integrability. These, together with the five conditions contained in (5.13), give ten differential conditions to be satisfied by the energy tensor $\mathsf{T}_{\mu\nu}$ in order that it should represent a source-free electromagnetic field.

If these ten conditions are satisfied, $\theta_{,\mu}$ is determined from (8.9), and θ is determined to within an arbitrary constant, except in a special case considered in Sec. 9. We note that $\theta_{,\mu}$ depends on the choice of $\mathbf{A}_\mu$ and $\mathbf{B}_\mu$. This is because θ is measured against the fields of $\mathbf{A}_\mu$ and $\mathbf{B}_\mu$. A transformation of $\mathbf{A}_\mu$ and $\mathbf{B}_\mu$ of the form (5.7) simply has the effect of subtracting the invariant angle α from θ. On the other hand, the conditions of integrability are independent of the choice of $\mathbf{A}_\mu$ and $\mathbf{B}_\mu$ as they should be. By (2.5) and (6.1) the electromagnetic field is then given by

$$\mathsf{f}_{\mu\nu} = (\mathbf{C}_\mu\mathbf{A}_\nu - \mathbf{C}_\nu\mathbf{A}_\mu)\cos\theta + (\mathbf{C}_\mu\mathbf{B}_\nu - \mathbf{C}_\nu\mathbf{B}_\mu)\sin\theta. \tag{8.12}$$

9. SPECIAL CASE

It may happen that $\mathbf{C}_\mu$ has the property that an invariant a exists such that

$$a\mathbf{C}_\mu = \epsilon_{,\mu}, \tag{9.1}$$

where ϵ is an invariant. Then ka can be added to w, where k is any constant, without affecting the conditions (8.10). In this case a more general solution of (8.9) exists, namely,

$$\begin{aligned}\theta'_{,\mu} &= \theta_{,\mu} + ka\mathbf{C}_\mu = \theta_{,\mu} + k\epsilon_{,\mu}, \\ \theta' &= \theta + k\epsilon.\end{aligned} \tag{9.2}$$

By (9.1) we have

$$(a\mathbf{C}_\mu)_{;\nu} = (a\mathbf{C}_\nu)_{;\mu},$$

whence

$$\mathsf{C}_{\mu;\nu} - \mathsf{C}_{\nu;\mu} = \mathbf{C}_\nu(\log a)_{,\mu} - \mathbf{C}_\mu(\log a)_{,\nu}.$$

Comparing this with (7.2), we see that a necessary condition that $\mathbf{C}_\mu$ should have this property is that

$$N = 0. \tag{9.3}$$

In effect, this special case means that the solution of (8.10) for w, if it exists, is of the form

$$w = w_0 + ka, \tag{9.4}$$

where w_0 and a are invariants and k is any constant. As we have seen, this can only happen when $N = 0$.

10. THE NULL-NULL CASE

In this case, throughout the domain considered,

$$\mathsf{C}^\alpha_{;\alpha} = 0. \tag{10.1}$$

(3.6) now becomes

$$\mathbf{C}^\alpha\mathsf{C}_{\mu;\alpha} = 0, \tag{10.2}$$

whence

$$\mathbf{C}^\alpha(\mathsf{C}_{\mu;\alpha} + \mathsf{C}_{\alpha;\mu}) = 0. \tag{10.3}$$

(3.8) becomes

$$\mathbf{G}^\alpha(\mathsf{C}_{\mu;\alpha} + \mathsf{C}_{\alpha;\mu}) = \mathbf{C}_\mu(\lambda_g - \mathsf{G}^\alpha_{;\alpha}). \tag{10.4}$$

In order to satisfy these conditions it is necessary and sufficient that $C_{\mu;\nu} + C_{\nu;\mu}$ can be written in the form

$$E_{\mu\nu} = C_{\mu;\nu} + C_{\nu;\mu} = m(C_\mu A_\nu + C_\nu A_\mu), \quad (10.5)$$

where $\mathbf{A}_\mu$ is a unit vector normal to $\mathbf{C}_\mu$ and is therefore spacelike, and m is an invariant.

It is seen that if (10.5) holds then (5.13) is satisfied, and therefore the conditions (5.13) cover the null–null case.

As before, we introduce vectors $\mathbf{A}_\mu$, $\mathbf{B}_\mu$, $\mathbf{S}_\mu$. $\mathbf{A}_\mu$ is already to hand in (10.5) as one member. $\mathbf{B}_\mu$ and $\mathbf{S}_\mu$ may be obtained in terms of an arbitrary vector field $\mathbf{y}_\mu$ not normal to $\mathbf{C}_\mu$; thus

$$\mathbf{B}^\mu = (\mathbf{C}^\epsilon \mathbf{y}_\epsilon)^{-1}(-g)^{-\frac{1}{2}}\epsilon^{\mu\alpha\beta\gamma}\mathbf{A}_\alpha \mathbf{C}_\beta \mathbf{y}_\gamma, \quad (10.6)$$

$$\mathbf{S}^\mu = (\mathbf{C}^\epsilon \mathbf{y}_\epsilon)^{-1}[\mathbf{y}^\mu - (\mathbf{y}^\alpha \mathbf{A}_\alpha)\mathbf{A}^\mu + \tfrac{1}{2}(\mathbf{y}^\gamma \mathbf{C}_\gamma)^{-1}\{(\mathbf{y}^\alpha \mathbf{A}_\alpha)^2 - \mathbf{y}^\alpha \mathbf{y}_\alpha\}\mathbf{C}^\mu]. \quad (10.7)$$

We then have, as before,

$$\mathbf{A}^\alpha \mathbf{B}_\alpha = \mathbf{A}^\alpha \mathbf{C}_\alpha = \mathbf{B}^\alpha \mathbf{C}_\alpha = \mathbf{A}^\alpha \mathbf{S}_\alpha = \mathbf{B}^\alpha \mathbf{S}_\alpha = 0,$$
$$\mathbf{C}^\alpha \mathbf{C}_\alpha = \mathbf{S}^\alpha \mathbf{S}_\alpha = 0, \quad \mathbf{A}^\alpha \mathbf{A}_\alpha = \mathbf{B}^\alpha \mathbf{B}_\alpha = \mathbf{S}^\alpha \mathbf{C}_\alpha = 1.$$

Equations (5.6) and (5.9) hold for these vectors. $\mathbf{B}_\mu$ and $\mathbf{S}_\mu$ are arbitrary to the extent that they can be replaced by

$$\mathbf{B}'_\mu = \mathbf{B}_\mu + r\mathbf{C}_\mu, \qquad \mathbf{S}'_\mu = \mathbf{S}_\mu - r\mathbf{B}_\mu - \tfrac{1}{2}r^2\mathbf{C}_\mu, \quad (10.8)$$

where r is any invariant. Proceeding as before, we find

$$C_{\mu;\nu} - C_{\nu;\mu} = N(\mathbf{A}_\mu \mathbf{B}_\nu - \mathbf{A}_\nu \mathbf{B}_\mu) + \rho(\mathbf{C}_\mu \mathbf{A}_\nu - \mathbf{C}_\nu \mathbf{A}_\mu) + \sigma(\mathbf{C}_\mu \mathbf{B}_\nu - \mathbf{C}_\nu \mathbf{B}_\mu), \quad (10.9)$$

$$(-g)^{-\frac{1}{2}}\epsilon^{\mu\nu\alpha\beta}C_{\alpha;\beta} = N(\mathbf{C}^\mu \mathbf{S}^\nu - \mathbf{C}^\nu \mathbf{S}^\mu) + \rho(\mathbf{C}^\mu \mathbf{B}^\nu - \mathbf{C}^\nu \mathbf{B}^\mu) - \sigma(\mathbf{C}^\mu \mathbf{A}^\nu - \mathbf{C}^\nu \mathbf{A}^\mu). \quad (10.10)$$

If we make a transformation of the form (10.8), we find that σ is independent of the choice of $\mathbf{B}_\mu$, while ρ transforms according to

$$\rho' = \rho + rN. \quad (10.11)$$

As before, we write

$$\mathbf{G}_\mu = \mathbf{A}_\mu \cos\theta + \mathbf{B}_\mu \sin\theta + p_g \mathbf{C}_\mu,$$
$$\mathbf{H}_\mu = \mathbf{B}_\mu \cos\theta - \mathbf{A}_\mu \sin\theta + p_h \mathbf{C}_\mu,$$

and find that

$$\theta_{,\mu} = \mathbf{A}^\alpha \mathbf{B}_{\alpha;\mu} + \sigma\mathbf{A}_\mu - \rho\mathbf{B}_\mu - \tfrac{1}{2}N\mathbf{S}_\mu + w\mathbf{C}_\mu, \quad (10.12)$$

where w is an undetermined invariant.

We are thus led to conditions which are formally the same as in the general null case, but in the present case $\mathbf{S}_\mu$ and $\mathbf{B}_\mu$ are to some extent arbitrary, so that it would appear that the conditions of integrability, in the form (8.10), are of no real value.

However, by performing the transformations (10.8) and (10.11), it can easily be seen that $\theta_{,\mu}$ given by (10.12) is independent of our choice of $\mathbf{B}_\mu$. That this must be so can be seen from the fact that, by (8.1) and (8.2), $\theta_{,\mu}$ depends only on the quantities

$$\mathbf{C}^\mu \mathbf{A}^\nu - \mathbf{C}^\nu \mathbf{A}^\mu \quad \text{and} \quad \mathbf{C}^\mu \mathbf{B}^\nu - \mathbf{C}^\nu \mathbf{B}^\mu,$$

which in this case are quite definite.

It follows that the values of $\theta_{,\mu}$, given by (10.12) and the integrability conditions (8.10), are independent of the choice of $\mathbf{B}_\mu$, and therefore of $\mathbf{S}_\mu$. It appears necessary to introduce these quantities in order to solve the equations and to express the necessary conditions of integrability.

A specific example of the null–null case is given by Witten[3] (see p. 395), from Peres.[5]

11. GRAVITATIONAL EQUATIONS

So far we have made no use of the relativity gravitational equations which may be written

$$R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R - \lambda g_{\mu\nu} = -\gamma T_{\mu\nu}, \quad (11.1)$$

where $R_{\mu\nu}$ is the Ricci tensor derived from the Riemann tensor [Ref. 4, Eq. (8.14)]. γ is a constant depending on the units used, and λ is the cosmological constant.

Contracting, we have in the present case, since $T^\alpha_\alpha = 0$,

$$R + 4\lambda = 0, \quad (11.2)$$

therefore

$$R_{\mu\nu} + \lambda g_{\mu\nu} = -\gamma T_{\mu\nu}. \quad (11.3)$$

Thus in the null case, by (2.3),

$$R_{\mu\nu} = -\gamma C_\mu C_\nu - \lambda g_{\mu\nu}. \quad (11.4)$$

Differentiate (3.6) and we get, after re-arrangement,

$$\mathbf{C}^\alpha \mathbf{C}^\beta R_{\beta\mu\alpha\nu} + \mathbf{C}^\alpha C_{\mu;\nu;\alpha} + C_\mu C^\alpha_{;\alpha;\nu} + C_{\mu;\alpha}C^\alpha_{;\nu} + C_{\mu;\nu}C^\alpha_{;\alpha} = 0.$$

Multiply by $g^{\mu\nu}$, the first term vanishes by (11.4) leaving

$$2C^\alpha(C^\beta_{;\beta})_{,\alpha} + (C^\alpha_{;\alpha})^2 + C^\beta_{;\alpha}C^\alpha_{;\beta} = 0. \quad (11.5)$$

By (7.5) this gives

$$N^2 = 8(C^\alpha_{;\alpha})^2 + 4C^\alpha(C^\beta_{;\beta})_{,\alpha}. \quad (11.6)$$

It follows that, when $C^\alpha_{;\alpha} = 0$ throughout any domain, then $N = 0$. Hence in the null–null case, if the gravitational equations are assumed to hold, N is always zero. This makes no difference to the results already obtained except that a few terms vanish.

12. SUMMARY

We first test whether $T_{\mu\nu}$ can be the energy tensor of a null, source-free, electromagnetic field.

Algebraic Conditions

$$\mathbf{T}^\alpha_\alpha = 0, \quad (1.6)$$

$$T_{\mu\alpha}\mathbf{T}^\alpha_\nu = 0. \quad (2.2)$$

[5] A. Peres, Phys. Rev. **118**, 1105 (1960).

If these hold we may write

$$T_{\mu\nu} = C_\mu C_\nu . \tag{2.3}$$

This defines C_μ, which is a null vector.

Differential Conditions

Let

$$E_{\mu\nu} \equiv C_{\mu;\nu} + C_{\nu;\mu} - 2g_{\mu\nu}C^\alpha_{;\alpha} . \tag{4.1}$$

Then we must have

$$C_\lambda C_\sigma E_{\mu\nu} + C_\mu C_\nu E_{\lambda\sigma} = C_\lambda C_\mu E_{\nu\sigma} + C_\nu C_\sigma E_{\lambda\mu} . \tag{5.13}$$

These consist of five independent conditions. If these conditions are satisfied, we can find the invariant, N, defined by the equation

$$NC^\mu = (-g)^{-\frac{1}{2}}\epsilon^{\mu\alpha\beta\gamma}C_\alpha C_{\beta;\gamma} . \tag{7.1}$$

Case I. When $C^\alpha_{;\alpha} \neq 0$ Over the Domain is Considered

If the conditions (5.13) are satisfied, we define a vector R_σ by the equation

$$R_\sigma C_\mu C_\nu = C_\sigma E_{\mu\nu} - C_\mu E_{\nu\sigma} - C_\nu E_{\mu\sigma} . \quad (5.12), (5.15)$$

From R_σ we find

$$\begin{aligned} q &= -R_\alpha R^\alpha / 12 C^\beta_{;\beta} , \\ S_\mu &= (R_\mu + qC_\mu)/6C^\alpha_{;\alpha} . \end{aligned} \tag{5.16}$$

S_μ is a null vector, depending only on C_μ and its derivatives. $S^\alpha C_\alpha = 1$.

Case II. When $C^\alpha_{;\alpha} = 0$ Over the Domain is Considered

We may write

$$E_{\mu\nu} = m(C_\mu A_\nu + C_\nu A_\mu), \tag{10.5}$$

where A_μ is a unit spacelike vector. This determines A_μ. The sign is immaterial.

We find two other vectors B_μ and S_μ by the method of (10.6) and (10.7). These can perhaps be simplified by a transformation of the form (10.8).

Cases I and II

The following integrability conditions must now be satisfied:

$$\begin{aligned} &(C_{\alpha;\mu}S_{\beta;\nu} - C_{\alpha;\nu}S_{\beta;\mu} + \tfrac{1}{2}R_{\alpha\beta\mu\nu})(-g)^{-\frac{1}{2}}\epsilon^{\alpha\beta\gamma\epsilon}C_\gamma S_\epsilon \\ &\quad - (-g)^{\frac{1}{2}}\eta_{\mu\nu\alpha\beta}(C^{\alpha;\beta}S^\epsilon + C^{\beta;\epsilon}S^\alpha + C^{\epsilon;\alpha}S^\beta)_{;\epsilon} \\ &\quad + \tfrac{1}{2}(NS_\mu)_{;\nu} - \tfrac{1}{2}(NS_\nu)_{;\mu} = (wC_\mu)_{;\nu} - (wC_\nu)_{;\mu} . \end{aligned} \tag{8.10}$$

These are effectively five conditions and an equation to determine the invariant w.

Electromagnetic Field with Above Conditions Satisfied

Case I. When $C^\alpha_{;\alpha} \neq 0$

We choose two spacelike unit vectors A_μ and B_μ normal to each other and normal to C_μ and S_μ. These can be found by the method of (5.8) and perhaps simplified by a transformation of the form (5.7).

Case II. When $C^\alpha_{;\alpha} = 0$

In this case A_μ and B_μ have already been determined.

Cases I and II

θ is now determined to within a constant (and sometimes more generally, depending on the determination of w) from the equation

$$\theta_{,\mu} = A^\alpha B_{\alpha;\mu} - (-g)^{\frac{1}{2}}\eta_{\mu\alpha\beta\gamma}S^\alpha C^{\beta;\gamma} + \tfrac{1}{2}NS_\mu + wC_\mu , \tag{8.9}$$

which is known to be integrable. The electromagnetic field is then given by

$$f_{\mu\nu} = (C_\mu A_\nu - C_\nu A_\mu)\cos\theta + (C_\mu B_\nu - C_\nu B_\mu)\sin\theta . \tag{8.12}$$

The electromagnetic field is indeterminate to the extent that θ can always have an added constant and in some cases may be even more general.

APPENDIX A

Let $L^{\cdot\nu}_{\mu}$ be any mixed tensor in four dimensions. Then the determinant $|L^{\cdot\nu}_{\mu}|$ is given by

$$\begin{aligned} 24\,|L^{\cdot\nu}_{\mu}| &= \delta^{\alpha\beta\gamma\epsilon}_{\lambda\mu\nu\rho}L^{\cdot\lambda}_{\alpha}L^{\cdot\mu}_{\beta}L^{\cdot\nu}_{\gamma}L^{\cdot\rho}_{\epsilon} \\ &= -6L^{\cdot\beta}_{\alpha}L^{\cdot\gamma}_{\beta}L^{\cdot\delta}_{\gamma}L^{\cdot\alpha}_{\delta} + 8L^{\cdot\alpha}_{\alpha.}L^{\cdot\gamma}_{\beta}L^{\cdot\delta}_{\gamma}L^{\cdot\beta}_{\delta} \\ &\quad - 6(L^{\cdot\alpha}_{\alpha})^2L^{\cdot\gamma}_{\beta}L^{\cdot\beta}_{\gamma} + 3(L^{\cdot\beta}_{\alpha}L^{\cdot\alpha}_{\beta})^2 + (L^{\cdot\alpha}_{\alpha})^4 . \end{aligned} \tag{A1}$$

Let the minor of $L^{\cdot\nu}_{\mu}$ in the determinant be $l^{\mu}_{\cdot\nu}$. Then we find

$$\begin{aligned} 6l^{\mu}_{\cdot\nu} &= \delta^{\mu}_{\nu}[2L^{\cdot\beta}_{\alpha}L^{\cdot\gamma}_{\beta}L^{\cdot\alpha}_{\gamma} - 3L^{\cdot\alpha}_{\alpha.}L^{\cdot\gamma}_{\beta}L^{\cdot\beta}_{\gamma} + (L^{\cdot\alpha}_{\alpha})^3] \\ &\quad - 6L^{\cdot\alpha}_{\nu}L^{\cdot\beta}_{\alpha}L^{\cdot\mu}_{\beta} + 6L^{\cdot\alpha}_{\alpha.}L^{\cdot\beta}_{\nu}L^{\cdot\mu}_{\beta} \\ &\quad - 3L^{\cdot\mu}_{\nu}[L^{\cdot\beta}_{\alpha}L^{\cdot\alpha}_{\beta} - (L^{\cdot\alpha}_{\alpha})^2] . \end{aligned} \tag{A2}$$

APPENDIX B

For formulas concerning the orthonormal tetrad, see Eisenhart, Ref. 4, Chap. 3.

Denote the tetrad by $\lambda_{i/\mu}$, $i = 1, 2, 3, 4$. $\lambda_{4/\mu}$ is the timelike member of the tetrad.

$$\begin{aligned} &h_{11} = h_{22} = h_{33} = 1, \quad h_{44} = -1, \\ &h_{ij} = 0 \quad \text{if} \quad i \neq j; \\ &\lambda_i \,|\, {}^\alpha\lambda_{j/\alpha} = h_{ij}, \quad h_{ij}\lambda_{i/\mu}\lambda_{j/\nu} = g_{\mu\nu} . \end{aligned} \tag{B1}$$

The tetradic components of tensors T_μ and $T_{\mu\nu}$ are

$$t_i = \lambda_i \,|\, {}^\alpha T_\alpha , \quad t_{ij} = \lambda_i|^\alpha\lambda_j|^\beta T_{\alpha\beta} . \tag{B2}$$

Then

$$\begin{aligned} T_\mu &= h_{ia}\lambda_{a/\mu}.t_i , \\ T_{\mu\nu} &= h_{ia}h_{jb}\lambda_{a/\mu}\lambda_{b/\nu}.t_{ij} . \end{aligned} \tag{B3}$$

The summation convention holds for repeated Latin suffixes.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Relation of the $O(2, 1)$ Partial-Wave Expansion to the Regge Representation

J. F. BOYCE*
International Centre for Theoretical Physics, Trieste, Italy

(Received 16 May 1966)

The general two-particle scattering amplitude is expanded in terms of partial waves corresponding to the crossed channel little group, $O(2, 1)$. Under the assumption of square integrability over the group manifold, the invariance of the S matrix under the complex Lorentz group, which follows from the Bargmann–Hall–Wightmann theorem, enables this expansion to be identified with the Regge representation in the crossed channel, whenever no dynamical singularities occur to the right of $\mathrm{Re}\, j = -\frac{1}{2}$. The identification requires the assumption of the fixed t dispersion relation necessary for the definition of the Regge representation.

1. INTRODUCTION

DUE to the crossing symmetry of the S matrix the two-particle scattering amplitude may be expressed in terms of two-particle helicity states which correspond to one incoming and one outgoing particle. The spacelike character of the total momentum of such a state permits its expansion in terms of eigenstates of the little group, $O(2, 1)$, maintaining, however, the reality of the masses of the component single-particle states. This expansion in turn enables the amplitude to be expanded in terms of the irreducible unitary representations of $O(2, 1)$, a subset of which forms a complete set for the expansion of any function which is square integrable over the group manifold. The manner of making the expansion which is adopted below is due to Dr. J. A. Strathdee, as is also the tenor of the approach.

The invariance of the S matrix under the complex Lorentz group enables this restricted expansion to be identified with the Regge continuation of the crossed channel $O(3)$ partial-wave expansion, subject to the condition that the partial-wave amplitudes have no dynamical singularities to the right of $\mathrm{Re}\, j = -\frac{1}{2}$.

The chief results are the identification of the principal series of $O(2, 1)$ representations with the background integral of the Regge continuation, and the discrete series with the nonsense channel terms, which for $O(2, 1)$ are perfectly natural contributions.

The major assumptions are the square integrability of the amplitude over the group manifold and the absence of dynamical singularities to the right of $\mathrm{Re}\, j = -\frac{1}{2}$ in the physical region of the s channel, together with the fixed t dispersion relation necessary for the definition of the Regge continuation of the amplitude.

We have tried to alleviate the complication due to spin and the presence of exchange forces by presenting the essential framework of the paper in Sec. 2. A summary of the crossed channel $O(3)$ expansion in Sec. 3 precedes the definition of $O(2, 1)$ helicity states in Sec. 4 and the expansion of the S matrix in terms of them in Sec. 5. The analytic continuation in Sec. 6 enables its identification with the $O(3)$ expansion of Sec. 3.

This work may be regarded as the continuation of an enquiry suggested by Joos[1] and is complementary to recent works by Toller,[2] Hadjioannou,[3] and Roffman.[4] It draws heavily upon the properties of the representations of $SL(2, R)$ which have been established by Andrews and Gunson.[5]

2. CENTER-OF-MASS AND BRICK-WALL FRAMES

The invariance of the S matrix under the Poincaré group $\mathscr{P}$ enables the scattering amplitude to be expanded in terms of its unitary irreducible representations. The usual partial-wave expansion is based upon the representations of the little group $O(3)$, which corresponds to positive definite eigenvalues of the Casimir operator P^2 of $\mathscr{P}$. The general two-particle transition amplitude

$$\langle p_1\lambda_1, p_2\lambda_2|\, T(s, t)\, |p_3\lambda_3, p_4\lambda_4\rangle,$$

for the process represented in Fig. 1, is expanded by transforming to the center-of-mass frame, in which $(p_1 + p_2)$ is along the time axis; and by basing the definition of the scattering amplitude on the single particle helicity state

$$|p\lambda\rangle = U(L_p)\, |\lambda\rangle, \tag{2.1}$$

where

$$p = m(\cosh\gamma, \sinh\gamma \sin\theta\cos\varphi, \sinh\gamma \sin\theta \sin\varphi, \sinh\gamma\cos\theta),$$

$$U(L_p) = \exp(-i\varphi J_{12}) \exp(-i\theta J_{31}) \exp(-i\gamma J_{03}),$$

* Permanent address: Imperial College, London, England.

[1] H. Joos, in *Lectures in Theoretical Physics* (University of Colorado, Boulder, 1964), Vol. 7A.
[2] M. Toller, Nuovo Cimento **37**, 631 (1965); Istituto di Fisica "G. Marconi", Note Interne 76 and 84.
[3] F. T. Hadjioannou, Nuovo Cimento **44A**, 185 (1966).
[4] E. H. Roffman, Phys. Rev. Letters **16**, 210 (1966).
[5] M. Andrews and J. Gunson, J. Math. Phys. **5**, 1391 (1964).

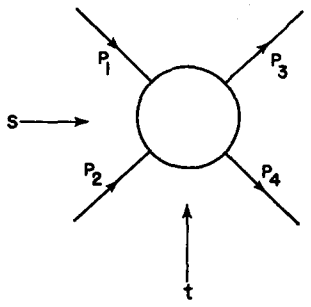


FIG. 1. General two-particle scattering.

and $|\lambda\rangle$ corresponds to a single particle of mass m at rest, having spin s with z component λ.

However, it is also possible to expand the amplitude in terms of representations which are based on the little group $O(2, 1)$,[6] which corresponds to negative definite eigenvalues of P^2.

This is achieved by transforming to the brick-wall frame, in which $(p_1 - p_3)$ is along the z axis, and by basing the amplitude upon the helicity states $|p\lambda\rangle'$, which are appropriate to the $O(2, 1)$ group, these being defined by

$$|p\lambda\rangle' = U(L_p')\,|\lambda\rangle, \tag{2.2}$$

where

$$p = m(\cosh\alpha\cosh\beta,\ \cosh\alpha\sinh\beta\cos\phi,\ \cosh\alpha\sinh\beta\sin\phi,\ \sinh\alpha)$$

and

$$U(L_p') = \exp(-i\phi J_{12})\exp(-i\beta J_{01})\exp(-i\alpha J_{03}).$$

The significance of this definition is that J_{12} and J_{01} are both generators of the relevant $O(2, 1)$ group.

Since L_p and L_p' both transform the vector $\hat{p} = (m, \underline{0})$ into the same vector p it follows that

$$(L_p')^{-1}L_p\hat{p} = \hat{p}$$

and consequently that $S_p = (L_p')^{-1}L_p$ is a pure $O(3)$ rotation.

Therefore, the $O(3)$ and $O(2, 1)$ helicity states may be simply related by

$$|p\lambda\rangle = \sum_{\mu} |p\mu\rangle'\langle\mu|\,S_p\,|\lambda\rangle, \tag{2.3}$$

where the summation is over the states of the irreducible representation of $SU(2)$ which is characterized by the spin s of the particle, and by using a specific representation for the generators it may be shown that

$$S_p = \exp(-i\Theta J_{31}),$$

where

$$\cos\Theta = \cosh\beta\cos\theta = \sinh\alpha\cosh\beta\{\sinh^2\alpha\cosh^2\beta + \sinh^2\beta\}^{-\frac{1}{2}}.$$

[6] The representations of $SU(1, 1)$ are in (2, 1) correspondence with the representation of $O(2, 1)$, which is locally isomorphic to $SL(2, R)$. See A. Moussa and R. Stora, in *Lectures in Theoretical Physics* (University of Colorado, Boulder, 1964), Vol. 7A.

Hence we may establish two equivalent representations for the general amplitude

$$\begin{aligned}&\langle p_1\lambda_1, p_2\lambda_2|\,T(s, t)\,|p_3\lambda_3, p_4\lambda_4\rangle\\ &\quad= \langle p_1\lambda_1, p_2\lambda_2|\,U^{-1}(\mathfrak{L}_1)T(s, t)U(\mathfrak{L}_1)\,|p_3\lambda_3, p_4\lambda_4\rangle\\ &\quad= \langle p_1\lambda_1, p_2\lambda_2|\,U^{-1}(\mathfrak{L}_2)T(s, t)U(\mathfrak{L}_2)\,|p_3\lambda_3, p_4\lambda_4\rangle,\end{aligned} \tag{2.4}$$

where $\mathfrak{L}_1$ is the Lorentz transformation to the center-of-mass frame, $(p_1 + p_2)$ along the time axis, and $\mathfrak{L}_2$ is the transformation to the brick-wall frame, $(p_1 - p_3)$ along the z axis. The subsequent expansions of the amplitudes are in terms of the center-of-mass angle θ, a function of the momentum transfer t, and a corresponding hyperbolic angle β, which is a function of s.

The individual vectors may be expressed as follows; for the center-of-mass frame, denoting the vectors by $\hat{q}$,

$$\begin{aligned}\hat{q}_1 &= m_1(\cosh\hat{\gamma}_1,\ \sinh\hat{\gamma}_1\sin\hat{\theta},\ 0,\ \sinh\hat{\gamma}_1\cos\hat{\theta}),\\ \hat{q}_2 &= m_2(\cosh\hat{\gamma}_2,\ \sinh\hat{\gamma}_2\sin\hat{\theta},\ 0,\ \sinh\hat{\gamma}_2\cos\hat{\theta}),\\ \hat{q}_3 &= m_3(\cosh\hat{\gamma}_3,\ 0,\ 0,\ \sinh\hat{\gamma}_3),\\ \hat{q}_4 &= m_4(\cosh\hat{\gamma}_4,\ 0,\ 0,\ \sinh\hat{\gamma}_4),\end{aligned} \tag{2.5}$$

with

$$m_1\sinh\hat{\gamma}_1 + m_2\sinh\hat{\gamma}_2 = m_3\sinh\hat{\gamma}_3 + m_4\sinh\hat{\gamma}_4 = 0, \tag{2.6}$$

while for the brick-wall frame

$$\begin{aligned}p_1 &= m_1(\cosh\alpha_1\cosh\beta,\ \cosh\alpha_1\sinh\beta,\ 0,\ \sinh\alpha_1),\\ p_2 &= m_2(\cosh\alpha_2,\ 0,\ 0,\ \sinh\alpha_2),\\ p_3 &= m_3(\cosh\alpha_3\cosh\beta,\ \cosh\alpha_3\sinh\beta,\ 0,\ \sinh\alpha_3),\\ p_4 &= m_4(\cosh\alpha_4,\ 0,\ 0,\ \sinh\alpha_4),\end{aligned} \tag{2.7}$$

with

$$m_1\cosh\alpha_1 - m_3\cosh\alpha_3 = m_4\cosh\alpha_4 - m_2\cosh\alpha_2 = 0. \tag{2.8}$$

All of the components of the vectors are uniquely determined by the values of s and t. This enables the angles to be interrelated by

$$\hat{\theta}(s, t) = i\beta(t, s)$$

and

$$\begin{aligned}&\hat{\gamma}_1(s, t) = \alpha_1(t, s) - \tfrac{1}{2}i\pi;\quad \hat{\gamma}_2(s, t) = \alpha_3(t, s) - \tfrac{1}{2}i\pi;\\ &\hat{\gamma}_3(s, t) = \alpha_2(t, s) - \tfrac{1}{2}i\pi;\quad \hat{\gamma}_4(s, t) = \alpha_4(t, s) - \tfrac{1}{2}i\pi,\end{aligned} \tag{2.9}$$

where $\alpha(t, s)$ means the value obtained from $\alpha(s, t)$ by interchanging s and t. These somewhat curious relations are indicative of the nature of the $O(2, 1)$ expansion which, although performed in the physical

region of the s channel, is essentially in terms of t channel variables, viz., t channel partial waves. Thus, although $\alpha(s, t)$ is real for the physical region of the s channel, $\alpha(t, s)$ is complex in this region. The relations indicate that the $O(2, 1)$ expansion is related to the continuation of the $O(3)$ expansion in the *crossed channel* at values of the parameters which correspond to the physical region of the direct channel. This, we see below, is simply the Regge continuation by means of the Sommerfeld–Watson transformation.

Since use has been made of the analyticity of the S matrix in order to continue from one channel to another it is not possible to connect the frames in which the expansions are to be made by a real Lorentz transformation. However, they may be related by a complex Lorentz transformation, while the amplitudes may be related by using the Bargmann–Hall–Wightmann theorem.[7]

The choice of variables appropriate to the center-of-mass frame of the t channel, viz., $(p_1 - p_3)$ along the t axis, is

$$\begin{aligned} q_1 &= -m_1(\cosh\gamma_1, \ \sinh\gamma_1 \sin\theta, \ 0, \ \sinh\gamma_1 \cos\theta), \\ q_2 &= \ \ m_2(\cosh\gamma_2, \ \ 0, \ \ 0, \ \ \sinh\gamma_2 \ \ ), \\ q_3 &= \ \ m_3(\cosh\gamma_3, \ \sinh\gamma_3 \sin\theta, \ 0, \ \sinh\gamma_3 \cos\theta), \\ q_4 &= -m_4(\cosh\gamma_4, \ \ 0, \ \ 0, \ \ \sinh\gamma_4 \ \ ), \end{aligned} \tag{2.10}$$

with

$$m_1 \sinh\gamma_1 + m_3 \sinh\gamma_3 = m_2 \sinh\gamma_2 + m_4 \sinh\gamma_4 = 0. \tag{2.11}$$

As might be anticipated, the Lorentz transformation from the t channel center-of-mass frame, $(p_1 - p_3)$ along the t axis, to the s channel brick-wall frame, $(p_1 - p_3)$ along the z axis, is that rotation which transforms a unit vector along the time axis into a unit vector along the z axis, while preserving its length. This may be verified explicitly by using the identity of s and t for both frames to interrelate the angular coordinates of the two systems, viz.,

$$\beta(s, t) = i\theta(s, t), \tag{2.12}$$

i.e.,

$$\zeta = \cosh\beta = \cos\theta = z$$

with

$$\gamma_1(s, t) = \alpha_1(s, t) + \tfrac{1}{2}i\pi; \quad \gamma_2(s, t) = \alpha_2(s, t) - \tfrac{1}{2}i\pi;$$
$$\gamma_3(s, t) = \alpha_3(s, t) - \tfrac{1}{2}i\pi; \quad \gamma_4(s, t) = \alpha_4(s, t) + \tfrac{1}{2}i\pi,$$

and in consequence

$$p = Rq, \tag{2.13}$$

where $R = \exp(\frac{1}{2}\pi J_{03})$ is the required rotation, while in addition

$$L_p = RL'_q \tag{2.14}$$

and

$$L'_p = RL_q,$$

(p, q) denoting any corresponding pair of $p_1, \cdots, p_4$; $q_1, \cdots, q_4$.

From the Bargmann–Hall–Wightmann theorem,[7] given an amplitude which satisfies:

(1) Under the real Lorentz transformation $\Lambda(A)$, corresponding to the element A of $SL(2, C)$, viz., (A, A^+) of $SL(2, C) \otimes SL(2, C)$, the amplitude transforms as

$$\begin{aligned} &\langle p_1\lambda_1, p_2\lambda_2| M(s, t) |p_3\lambda_3, p_4\lambda_4\rangle \\ &\quad = \mathfrak{D}^{(s_1 0)^\dagger}(\Lambda)^{\mu_1}_{\lambda_1} \mathfrak{D}^{(s_2 0)^\dagger}(\Lambda)^{\mu_2}_{\lambda_2} \mathfrak{D}^{(s_3 0)}(\Lambda)^{\lambda_3}_{\mu_3} \mathfrak{D}^{(s_4 0)}(\Lambda)^{\lambda_4}_{\mu_4} \\ &\qquad \times \langle \Lambda p_1\mu_1, \Lambda p_2\mu_2| M(s, t) |\Lambda p_3\mu_3, \Lambda p_4\mu_4\rangle, \end{aligned}$$

where $\mathfrak{D}^{(s0)}$ is a representation of $SL(2, C)$.

(2) Holomorphic in the tube $\eta_j \in V^+$, where $p_j = \xi_j - \eta_j$, $j = 1, \cdots, 4$, and V_+ is the set of all four vectors p which satisfy $p^2 > 0$, $p_0 > 0$. Then the amplitude has a single-valued analytic continuation which transforms as above, but where Λ is now the complex Lorentz transformation which corresponds to the element (A, B) of $SL(2, C) \otimes SL(2, C)$.

The above amplitude may be identified with the M function which corresponds to the transition amplitude under consideration. Explicitly

$$\begin{aligned} &\langle p_1\lambda_1, p_2\lambda_2| M(s, t) |p_3\lambda_3, p_4\lambda_4\rangle \\ &\quad = \mathfrak{D}^{(s_1 0)}(L_{p_1})^{\mu_1}_{\lambda_1} \mathfrak{D}^{(s_2 0)}(L_{p_2})^{\mu_2}_{\lambda_2} \mathfrak{D}^{(s_3 0)^\dagger}(L_{p_3})^{\lambda_3}_{\mu_3} \mathfrak{D}^{(s_4 0)^\dagger}(L_{p_4})^{\lambda_4}_{\mu_4} \\ &\qquad \times \langle p_1\mu_1, p_2\mu_2| T(s, t) |p_3\mu_3, p_4\mu_4\rangle. \end{aligned}$$

This enables a single-valued analytic continuation of the transition amplitude to be defined[7] which, under the complex Lorentz transformation, $\Lambda = R = \exp(\frac{1}{2}\pi J_{03})$, transforms as

$$\begin{aligned} &\langle p_1\lambda_1, p_2\lambda_2| T(s, t) |p_3\lambda_3, p_4\lambda_4\rangle \\ &\quad = \mathfrak{D}^{s_1\dagger}(W_1)^{\mu_1}_{\lambda_1} \mathfrak{D}^{s_2\dagger}(W_2)^{\mu_2}_{\lambda_2} \mathfrak{D}^{s_3}(W_3)^{\lambda_3}_{\mu_3} \mathfrak{D}^{s_4}(W_4)^{\lambda_4}_{\mu_4} \\ &\qquad \times \langle q_1\mu_1, q_2\mu_2| T(s, t) |q_3\mu_3, q_4\mu_4\rangle, \end{aligned} \tag{2.15}$$

where

$$\begin{aligned} W_p &= (L_{R^{-1}p})^{-1} R^{-1} L_p = (L_q)^{-1} R^{-1} L_p \\ &= (L'_p)^{-1} L_p = S_p, \end{aligned} \tag{2.16}$$

and therefore may be identified with the spin rotation which entered into the definition of the $O(2, 1)$ helicity states, Eq. (2.3).

Hence the general amplitude may be transformed to the s channel brick-wall system, $(p_1 - p_3)$ along the

[7] R. F. Streater and A. S. Wightmann, *PCT, Spin and Statistics and all that* (W. A. Benjamin Inc., New York, 1964), Theorem (2.11). H. P. Stapp, Phys. Rev. **125**, 2139 (1962); I. J. Muzinich, J. Math. Phys. **5**, 1481 (1964). The same analysis applies to the crossing relation between the s channel Breit frame and t channel barycentric frame as that between the s and t channel barycentric frames considered by Muzinich.

z axis, which may then either be expressed in terms of $O(2,1)$ helicity states and expanded in terms of $O(2,1)$ representations, or be subjected to the complex rotation $R = \exp(\frac{1}{2}\pi J_{03})$ which transforms it to the t channel center-of-mass system, $(p_1 - p_3)$ along the t axis, and then expanded in terms of $O(3)$ representations, i.e., the amplitude in the brick-wall frame is expressible as

$$\langle p_1\lambda_1, p_2\lambda_2|\, T(s,t)\,|p_3\lambda_3, p_4\lambda_4\rangle$$
$$= \langle\lambda_1|\, S_1^+\,|\mu_1\rangle\langle\lambda_2|\, S_2^+\,|\mu_2\rangle\backslash\langle p_1\mu_1, p_2\mu_2|\, T(s,t)\,|p_3\mu_3, p_4\mu_4\rangle/ \times \langle\mu_3|\, S_3\,|\lambda_3\rangle\langle\mu_4|\, S_4\,|\lambda_4\rangle \quad (2.17)$$
$$= \langle\lambda_1|\, S_1^+\,|\mu_1\rangle\langle\lambda_2|\, S_2^+\,|\mu_2\rangle\langle q_1\mu_1, q_2\mu_2|\, T(s,t)\,|q_3\mu_3, q_4\mu_4\rangle \times \langle\mu_3|\, S_3\,|\lambda_3\rangle\langle\mu_4|\, S_4\,|\lambda_4\rangle, \quad (2.18)$$

where, upon removing appropriate kinematical factors $\mathcal{N}$ and $\mathcal{N}'$ from the initial and final states,

$$T_{\lambda\mu}(\zeta,t) = \backslash\langle p_1\lambda_1, p_2\lambda_2|\, T(s,t)\,|p_3\lambda_3, p_4\lambda_4\rangle/\,[\mathcal{N}^*\mathcal{N}']^{-1}$$

and

$$T_{\lambda\mu}(z,t) = \langle q_1\lambda_1, q_2\lambda_2|\, T(s,t)\,|q_3\lambda_3, q_4\lambda_4\rangle[\mathcal{N}^*\mathcal{N}']^{-1},$$

are expandable in terms of $O(3)$ and $O(2,1)$ representations, respectively.

The content of this paper is the formulation of these two expansions and the proof that under certain conditions one may be analytically continued into the other; it being most convenient to continue the $O(2,1)$ into the $O(3)$. There are two consequences. Firstly, we are able to identify the principal and discrete series of the $O(2,1)$ representation with the background integral and nonsense channel terms, respectively, of the usual Regge continuation, thereby indicating that from the viewpoint of the $O(2,1)$ little group the nonsense channel terms appear to be as valid as any other contribution. Secondly, we find that we may base the analytic continuation on that subset of irreducible unitary representations of $O(2,1)$ which appear in the expansion of any function which is square integrable over the group manifold, although an extension of this set is indicated. In addition there is a suggestion that the remaining irreducible unitary representations, notably the supplementary series, which may all be located to the left of $\mathrm{Re}\,j = -\frac{1}{2}$, may also be useful in a representation of the scattering amplitude.

The main points of the argument are as follows. The definition of $O(2,1)$ helicity states enables the amplitude $\bar{T}_{\lambda\mu}(\zeta,t)$, obtained from $T_{\lambda\mu}(\zeta,t)$ by crossing symmetry, to be expanded as

$$\bar{T}_{\lambda\mu}(\zeta,t) = \sum_{k=0\text{ or }\frac{1}{2}}^{|\mu|-1}(2k+1)\bar{T}_{\lambda\mu}(k,t)d^k_{\lambda\mu}(\zeta) + \frac{1}{2i}\int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty} dj(2j+1)\bar{T}_{\lambda\mu}(j,t)d^j_{\lambda\mu}(\zeta), \quad (2.19)$$

where we have assumed $\lambda \geq |\mu| \geq 0$, since the symmetry of the representations enables the general case to be expressed in terms of these. The expansion coefficients are

$$\bar{T}_{\lambda\mu}(k,t) = \frac{1}{2}\int_1^\infty d\zeta \bar{T}_{\lambda\mu}(\zeta,t)\,d^k_{\lambda\mu}(\zeta),$$
$$\bar{T}_{\lambda\mu}(j,t) = \frac{1}{2i}\int_1^\infty d\zeta \bar{T}_{\lambda\mu}(\zeta,t)e^j_{\lambda\mu}(\zeta). \quad (2.20)$$

The application of an inverse Sommerfeld–Watson transformation to the principal series integral is prevented by the well-known effect of the exchange potential. We circumvent this by assuming the fixed t dispersion relation[8]

$$\bar{T}_{\lambda\mu}(\zeta,t) = \left(\frac{\zeta-1}{2}\right)^{\frac{1}{2}(\lambda-\mu)}\left(\frac{\zeta+1}{2}\right)^{\frac{1}{2}(\lambda+\mu)} \times \left[\int_{\zeta_R}^\infty d\zeta' \frac{\bar{\rho}^R_{\lambda\mu}(\zeta',t)}{\zeta'-\zeta} + \int_{-\infty}^{-\zeta_L} d\zeta' \frac{\bar{\rho}^L_{\lambda\mu}(\zeta',t)}{\zeta'-\zeta}\right], \quad (2.21)$$

which enables the amplitude to be expressed as the sum of two terms which may be analytically continued in j. In addition it enables the $O(2,1)$ expansion coefficients to be related to those of the $O(3)$ expansion via the lemma

$$e^j_{\lambda\mu}(\zeta)\left(\frac{\zeta-1}{2}\right)^{\frac{1}{2}(\lambda-\mu)}\left(\frac{\zeta+1}{2}\right)^{\frac{1}{2}(\lambda+\mu)} e^{i\pi(j-\lambda)} = \frac{\sin\pi(j-\lambda)}{\pi}\int_1^\infty d\zeta' \times \frac{e^j_{\lambda\mu}(\zeta')[\frac{1}{2}(\zeta'-1)]^{\frac{1}{2}(\lambda-\mu)}[\frac{1}{2}(\zeta'+1)]^{\frac{1}{2}(\lambda+\mu)}}{\zeta'-\zeta} - \frac{1}{2}\int_{-1}^1 d\zeta' \times \frac{d^j_{\lambda-\mu}(-\zeta')[\frac{1}{2}(\zeta'-1)]^{\frac{1}{2}(\lambda-\mu)}[\frac{1}{2}(\zeta'+1)]^{\frac{1}{2}(\lambda+\mu)}}{\zeta'-\zeta}, \quad (2.22)$$

which holds for $\mathrm{Re}\,j > M-1$, $M = \max(|\lambda|, |\mu|)$. Upon deforming the principal series integral to the right in the j plane it is found that kinematical singularities of the integrand exist which give rise to contributions which exactly cancel the discrete series and reproduce the familiar $O(3)$ expansion in the t channel.

3. NORMAL t CHANNEL PARTIAL-WAVE EXPANSION

From crossing symmetry,[9] the transition amplitude in the t channel center-of-mass frame may be

[8] F. Calogero, J. M. Charap, and E. J. Squires, Ann. Phys. (N.Y.) **25**, 325 (1963).
[9] A. O. Barut, Phys. Rev. **130**, 436 (1963).

expressed as

$$\langle q_1\lambda_1, q_2\lambda_2|\, T(s,t)\,|q_3\lambda_3, q_4\lambda_4\rangle$$
$$= \mathfrak{D}^{s_1}(C)^{\mu_1}_{\lambda_1}\mathfrak{D}^{s_4}(C^{-1})^{\mu_4}_{\lambda_4}\langle -q_4\mu_4, q_2\lambda_2|\, \bar{T}(s,t)\,|q_3\lambda_3, -q_1\mu_1\rangle$$
$$= \mathcal{N}^*\mathcal{N}'\mathfrak{D}^{s_1}(C)^{\mu_1}_{\lambda_1}\mathfrak{D}^{s_4}(C^{-1})^{\mu_4}_{\lambda_4}\bar{T}_{\lambda\mu}(z,t), \tag{3.1}$$

where by the normal partial-wave decomposition in terms of helicity states[10]

$$\bar{T}_{\lambda\mu}(z,t) = \sum_{j=M}^{\infty}(2j+1)\bar{T}_{\lambda\mu}(j,t)\, d^j_{\lambda\mu}(\cos\theta), \tag{3.2}$$

in which

$$\lambda = \lambda_2 - \lambda_4, \quad \mu = \lambda_3 - \lambda_1, \quad M = \max(|\lambda|, |\mu|)$$

and

$$\mathcal{N} = \{4\pi(m_1 \sinh\gamma_1^{-1})(m_3\cosh\gamma_3 + m_1\cosh\gamma_1)\}^{\frac{1}{2}}$$

while

$$\bar{T}_{\lambda\mu}(j,t) = \tfrac{1}{2}\int_{-1}^{1} dz\bar{T}_{\lambda\mu}(z,t)\, d^j_{\lambda\mu}(z). \tag{3.3}$$

The definition of the angular coordinate system, Eq. (2.10), enables the identification

$$z = \cos\theta = \frac{-2t(s - m_1^2 - m_2^2) - (t + m_1^2 - m_3^2)(t + m_2^2 - m_4^2)}{[\{t - (m_1 - m_3)^2\}\{t - (m_1 + m_3)^2\} \times \{t - (m_2 - m_4)^2\}\{t - (m_2 + m_4)^2\}]^{\frac{1}{2}}}. \tag{3.4}$$

The amplitude is expressed in a j plane analytically continuable form[8] by using the dispersion relation, Eq. (2.21) to express

$$T_{\lambda\mu}(z,t) = A_{\lambda\mu}(z,t) - e^{i\pi\lambda}B_{\lambda-\mu}(-z,t) \tag{3.5}$$

when

$$\bar{T}_{\lambda\mu}(j,t) = \bar{A}_{\lambda\mu}(j,t) - e^{i\pi j}\bar{B}_{\lambda-\mu}(j,t),$$

where

$$\bar{A}_{\lambda\mu}(j,t) = \int_{z_R}^{\infty} dz'\,\tilde{\rho}^R_{\lambda\mu}(z',t) \times \left(\frac{z'-1}{2}\right)^{\frac{1}{2}(\lambda-\mu)}\left(\frac{z'+1}{2}\right)^{\frac{1}{2}(\lambda+\mu)} e^j_{\lambda\mu}(z') \tag{3.6}$$

and

$$\bar{B}_{\lambda-\mu}(j,t) = -\int_{z_L}^{\infty} dz'\,\tilde{\rho}^L_{\lambda\mu}(-z',t) \times \left(\frac{z'-1}{2}\right)^{\frac{1}{2}(\lambda+\mu)}\left(\frac{z'+1}{2}\right)^{\frac{1}{2}(\lambda-\mu)} e^j_{\lambda-\mu}(z'). \tag{3.7}$$

The expansions of $\bar{A}_{\lambda\mu}(z,t)$ and $\bar{B}_{\lambda\mu}(z,t)$ are amenable to the Sommerfeld–Watson transformation for $z > 0$) and hence define an analytic continuation of $T_{\lambda\mu}(z,t)$ in the j plane, whose uniqueness follows from Carlson's theorem. In terms of them,

$$\bar{T}_{\lambda\mu}(z,t) = \sum_{j=M}^{\infty}(2j+1)[\bar{A}_{\lambda\mu}(j,t) - e^{i\pi j}\bar{B}_{\lambda-\mu}(j,t)]\, d^j_{\lambda\mu}(z). \tag{3.8}$$

4. THE DEFINITION OF $O(2,1)$ HELICITY STATES

We proceed by analogy with the normal $O(3)$ development.[10] The single-particle states are defined by Eq. (2.2). The general two-particle state is defined by

$$|p_1\lambda_1, -p_2\lambda_2\rangle' = |p_1\lambda_1\rangle' \otimes |-p_2 - \lambda_2\rangle', \tag{4.1}$$

where, if $(p_1 - p_2)$ is purely along the z axis

$$\varphi_1 = \varphi_2 = \varphi; \quad \beta_1 = \beta_2 = \beta;$$
$$m_1\cosh\alpha_1 - m_2\cosh\alpha_2 = 0,$$

and hence

$$|p_1\lambda_1, -p_2\lambda_2\rangle' = \exp(-i\varphi J_{12})\exp(-i\beta J_{01}) \times \{\exp(-i\alpha_1 J_{03})\,|\lambda_1\rangle \otimes \exp(-i\alpha_2 J_{03})\,|-\lambda_2\rangle\} = \mathcal{N}\exp(-i\varphi J_{12})\exp(-i\beta J_{01})\,|\alpha;\lambda_1\lambda_2\rangle', \tag{4.2}$$

where $|-\lambda_2\rangle$ means $|(-m_2, \underline{0}), -\lambda_2\rangle$ while

$$P_\mu\,|\alpha;\lambda_1\lambda_2\rangle' = 0 \qquad (\mu = 0,1,2),$$
$$P_3\,|\alpha;\lambda_1\lambda_2\rangle' = \alpha\,|\alpha;\lambda_1\lambda_2\rangle',$$

and hence may be expanded in terms of the bases of irreducible representations of $O(2,1)$,[11] as

$$|\alpha;\lambda_1\lambda_2\rangle' = \sum_{q,m} N(q,m)\,|\alpha;\lambda_1\lambda_2;qm\rangle' = \sum_q N(q)\,|\alpha;\lambda_1\lambda_2;qm = \lambda_1 - \lambda_2\rangle' \tag{4.3}$$

since

$$J_{12}\,|\alpha;\lambda_1\lambda_2\rangle' = (\lambda_1 - \lambda_2)\,|\alpha;\lambda_1\lambda_2\rangle'.$$

The representation of $O(2,1)$ may be characterized by q and the individual basis vectors by m, where

$$(J_{12}^2 - J_{01}^2 - J_{02}^2)\,|q,m\rangle' = q(q+1)\,|q,m\rangle',$$
$$J_{12}\,|q,m\rangle' = m\,|q,m\rangle'.$$

As q assumes discrete, integral or half-integral, and continuous values, $\sum$ signifies both summation and integration.

The elements U of $O(2,1)$ may be parametrized by (φ, β, ψ), defined analogously to the Euler angles by

$$U(\varphi,\beta,\psi) = \exp(-i\varphi J_{12})\exp(-i\beta J_{01})\exp(-i\psi J_{12})$$
$$0 \le \varphi, \psi \le 2\pi, \qquad -\infty \le \beta \le \infty.$$

[10] M. Jacob and G. C. Wick, Ann. Phys. (N.Y.) 7, 404 (1959); G. C. Wick, *ibid.* 18, 65 (1962); L. Durand, P. C. de Celles, and R. B. Marr, Phys. Rev. 126, 1882 (1962).

[11] V. Bargmann, Ann. Math. 48, 568 (1947); A. O. Barut and C. Fronsdal, Proc. Roy. Soc. (London) A287, 532 (1965); L. C. Biedenharn, J. Nuyts, and N. Straumann, CERN Preprint 65/76115-TH. 555; F. R. Halpern and E. Branscomb, UCRL (Livermore)-1235 Rev. 1, & Errata.

Under the transformation characterized by (φ, β, ψ)

$$|q, m\rangle' \to U(\varphi, \beta, \psi)\,|q, m\rangle' = \sum_{m'} |q, m'\rangle' \mathfrak{D}^q_{m'm}(\varphi, \beta, \psi),$$

where

$$\mathfrak{D}^q_{mm'}(\varphi, \beta, \psi) = \exp(-im\varphi)\, d^q_{mm'}(\cosh\beta) \exp(-im'\psi)$$

and $d^q_{mm'}(\cosh\beta)$ is the analytic continuation of $d^j_{mm'}(\cos\theta)$, this being defined by its expression in terms of hypergeometric functions.[5] Hence the two-particle state may be expanded as

$$|p_1\lambda_2, -p_2\lambda_2\rangle' = \mathcal{N} \sum_{q,m} N(q)\, |\alpha; \lambda_1\lambda_2; qm\rangle' \mathfrak{D}^q_{m\lambda_1-\lambda_2}(\varphi, \beta, 0). \tag{4.4}$$

The usual normalization of two-particle states

$$\Delta^+(p_1)\Delta^+(-p_2)\langle p_1\lambda_1, -p_2\lambda_2 \,|\, p_3\lambda_3, -p_4\lambda_4\rangle = (2\pi)^8 \delta^4(p_1 - p_3)\delta^4(p_2 - p_4)\delta_{\lambda_1\lambda_3}\delta_{\lambda_2\lambda_4},$$

where

$$\Delta^+(p) = 2\pi\delta(p^2 - m^2)\theta(p_0) = 2\pi\delta(p^2 - m^2)\theta(p_0)$$

implies

$$\begin{aligned}(2\pi)^{-8}&\Delta^+(p_1)\Delta^+(-p_2)\,{}^{\backslash}\langle p_1\lambda_1, -p_2\lambda_2 \,|\, p_3\lambda_3, -p_4\lambda_4\rangle' \\ &= \delta^4(p_1 - p_3)\delta^4(p_2 - p_4)\delta_{\lambda_1\lambda_3}\delta_{\lambda_2\lambda_4} \\ &= \nu^2 u^{-1}\delta^4(p_1 + p_2 - p_3 - p_4)\delta(p_{1_z} - u_1) \\ &\times \delta(p_{2_z} - u_2)\delta(\cosh\beta - \cosh\beta')\delta(\varphi - \varphi')\delta_{\lambda_1\lambda_3}\delta_{\lambda_2\lambda_4},\end{aligned} \tag{4.5}$$

where

$$u_1 = m_1 \sinh\alpha_1;\ u_2 = m_2 \sinh\alpha_2,$$

$$\nu = m_1 \cosh\alpha_1;$$

$$u = \frac{m_1 \cosh\alpha_1 (m_1 \sinh\alpha_1 - m_2 \sinh\alpha_2)}{m_1 \sinh\alpha_1\, m_2 \sinh\alpha_2}$$

the variables being with respect to the frame in which $(p_1 - p_2)$ is purely along the z axis, primed and unprimed indices referring to final and initial states.

Hence the normalization of Eq. (4.2) by

$$N = (4\pi\nu^{-2} u\, u_1 u_2)^{\frac{1}{2}},$$

which may be identified with the $O(3)$ normalization, yields

$$\begin{aligned}\pi\Delta^+(p_1)\Delta^+(p_2)&\delta^3(\tilde{p}_1 + \tilde{p}_2 - \tilde{p}_3 - \tilde{p}_4) \\ &\times \sum_{\substack{q,q' \\ mm'}} {}^{\backslash}\langle\alpha; \lambda_1\lambda_2; qm \,|\, \alpha'; \lambda_3\lambda_4; q'm'\rangle' \\ &\times \mathfrak{D}^{q\,\dagger}_{\lambda_1-\lambda_2 m}(\varphi, \beta, 0)\mathfrak{D}^{q'}_{m'\lambda_3-\lambda_4}(\varphi', \beta', 0) N^*(q)N(q') \\ = \delta(p_{1_z} - u_1)&\delta(p_{2_z} - u_2)\delta^4(p_1 + p_2 - p_3 - p_4) \\ \times \delta(\cosh\beta - \cosh\beta')&\delta(\varphi - \varphi')\delta_{\lambda_1\lambda_3}\delta_{\lambda_2\lambda_4}(2\pi)^8(4u_1u_2)^{-1},\end{aligned}$$

where

$$\tilde{p} = (p_0, p_1, p_2) \tag{4.6}$$

and the orthogonality relations

$${}^{\backslash}\langle\alpha; \lambda_1\lambda_2; qm \,|\, \alpha'; \lambda_3\lambda_4; q'm'\rangle' = 2\pi\delta(\alpha - \alpha')\delta_{\lambda_1\lambda_3}\delta_{\lambda_2\lambda_4}\delta_{qq'}\delta_{mm'} \tag{4.7}$$

imply

$$N(q) = (2q + 1)^{\frac{1}{2}} \tag{4.8}$$

since

$$\tfrac{1}{2}\sum_q (2q + 1)\, d^q_{\lambda\lambda'}(\cosh\beta)\, d^q_{\lambda\lambda'}(\cosh\beta') = \delta(\cosh\beta - \cosh\beta'). \tag{4.9}$$

5. $O(2, 1)$ DECOMPOSITION IN THE s CHANNEL

In the s channel brick-wall frame, $(p_1 - p_3)$ along the z axis, the specification of the angular coordinates, Eq. (2.7) permits the identification

$$\zeta = \cosh\beta = \frac{\left[\begin{matrix}-2t(s - m_1^2 - m_2^2) \\ -(t + m_1^2 - m_3^2)(t + m_2^2 - m_4^2)\end{matrix}\right]}{\left[\begin{matrix}[\{t - (m_1 - m_3)^2\}\{t - (m_1 + m_3)^2\} \\ \times \{t - (m_2 - m_4)^2\}\{t - (m_2 + m_4)^2\}]^{\frac{1}{2}}\end{matrix}\right]}, \tag{5.1}$$

which may be compared with Eq. (3.4), it following that $z = \zeta$; however, for clarity, we maintain them as distinct variables. Crossing symmetry, which takes the same form for the amplitude defined in terms of $O(2, 1)$ helicity states as for the normal definition, may be utilized to express the $O(2, 1)$ helicity amplitude in the s channel brick-wall frame as

$$\begin{aligned}{}^{\backslash}\langle p_1\lambda_1, p_2\lambda_2 |\, T(s, t)\, |p_3\lambda_3, p_4\lambda_4\rangle' &= \mathfrak{D}^{s_1}(C)^{\mu_1}_{\lambda_1}\mathfrak{D}^{s_4}(C^{-1})^{\mu_4}_{\lambda_4} \\ &\times {}^{\backslash}\langle -p_4\mu_4, p_2\lambda_2 |\, T(s, t)\, |p_3\lambda_3, -p_1\mu_1\rangle',\end{aligned} \tag{5.2}$$

where

$$\begin{aligned}|p_3\lambda_3, -p_1\lambda_1\rangle' &= \exp(-i\beta J_{01}) \\ &\times \{\exp(-i\alpha_3 J_{03})\, |\lambda_3\rangle \otimes \exp(-i\alpha_1 J_{03})\, |-\lambda_1\rangle\} \\ &= \mathcal{N}' \exp(-i\beta J_{01})\, |\alpha; \lambda_3\lambda_1\rangle',\end{aligned}$$

while

$$|-p_4\lambda_4, p_2\lambda_2\rangle' = \mathcal{N}\, |\alpha; \lambda_4\lambda_2\rangle'$$

and hence

$$\begin{aligned}{}^{\backslash}\langle -p_4\lambda_4, p_2\lambda_2 |\, T(s, t)\, |p_3\lambda_3, -p_1\lambda_1\rangle' &= \mathcal{N}^*\mathcal{N}' \sum_{\substack{q,q' \\ m,m'}} {}^{\backslash}\langle\alpha; \lambda_4\lambda_2; qm |\, T(s, t)\, |\alpha; \lambda_3\lambda_1; q'm'\rangle' \\ &\times \mathfrak{D}^{q\,\dagger}_{m\,\lambda_2-\lambda_4}(0, 0, 0)\mathfrak{D}^{q}_{m'\,\lambda_3-\lambda_1}(0, \beta, 0) N^*(q)N(q').\end{aligned}$$

But

$$[J_{12}^2 - J_{01}^2 - J_{02}^2,\, T] = [J_{12},\, T] = 0$$

enables the definition

$${}^{\backslash}\langle\alpha; \lambda_4\lambda_2; qm |\, T(s, t)\, |\alpha; \lambda_3\lambda_1; q'm'\rangle' = \delta_{qq'}\delta_{mm'}\, {}^{\backslash}\langle\lambda_4\lambda_2 |\, \bar{T}(q, t)\, |\lambda_3\lambda_1\rangle' \tag{5.3}$$

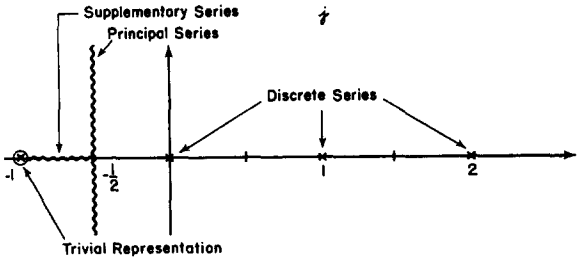


FIG. 2. Single-valued representations of $O(2, 1)$.

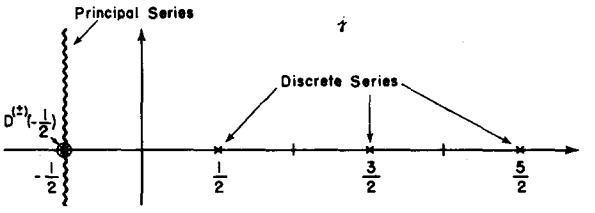


FIG. 3. Double-valued representations of $O(2, 1)$.

to be made, in terms of which

$$\langle -p_4\lambda_4, p_2\lambda_2|\, T(s,t)\,|p_3\lambda_3, -p_1-\lambda_1\rangle' = \sum_q (2q+1)\,\langle\lambda_4\lambda_2|\,\bar{T}(q,t)\,|\lambda_3\lambda_1\rangle'\, d^q_{\lambda\mu}(\cosh\beta)\mathcal{N}^*\mathcal{N}'$$

$$\lambda = \lambda_2 - \lambda_4; \quad \mu = \lambda_3 - \lambda_1;$$

i.e.,

$$\langle p_1\lambda_1, p_2\lambda_2|\, T(s,t)\,|p_3\lambda_3, p_4\lambda_4\rangle' = \mathfrak{D}^{s_1}(C)^{\mu_1}_{\lambda_1}\mathfrak{D}^{s_4}(C^{-1})^{\mu_4}_{\lambda_4}\bar{T}_{\lambda\mu}(\zeta,t)\mathcal{N}^*\mathcal{N}',$$

$$\lambda = \lambda_2 - \mu_4, \quad \mu = \lambda_3 - \mu_1, \qquad (5.4)$$

where

$$\bar{T}_{\lambda\mu}(\zeta,t) = \sum_q (2q+1)\bar{T}_{\lambda\mu}(q,t)\, d^q_{\lambda\mu}(\zeta), \qquad (5.5)$$

which may be identified as the expansion formula of the $O(2, 1)$ group, Eq. (2.19), which we now consider in more detail.

Just as in the $O(3)$ expansion, the representations of $SU(2)$ are required to form a complete set, so the representations of the corresponding spinor group $SU(1, 1)$ are needed in the $O(2, 1)$ analysis. These have been established by a number of authors.[11] We make use of the formulation of Barut and Fronsdal, replacing, however, Φ by $(-j-1)$; for, as the representations corresponding to Φ and $(-\Phi-1)$ are equivalent, we are at liberty to do this. Upon denoting each irreducible unitary representation by a point in the complex j plane, the single- and double-valued representations may be exhibited as shown in Figs. 2 and 3. They may be identified in the usual way as

Principal series

$$j = -\tfrac{1}{2} + i\sigma \qquad -\infty \leq \sigma \leq \infty.$$

Eigenvalue of $J_{12} = 0, \pm 1, \cdots$ single valued

$= \pm\tfrac{1}{2}, \pm\tfrac{3}{2}, \cdots$ double valued.

Supplementary series

$$-1 < j < -\tfrac{1}{2}.$$

Eigenvalue of $J_{12} = 0, \pm 1, \cdots$ single valued.

Positive discrete series, D^+

$j = 0, 1, 2, \cdots$ single valued

$= -\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{3}{2}, \cdots$ double valued.

Eigenvalue of $J_{12} = j + 1, j + 2, \cdots$.

Negative discrete series, D^-

$j = 0, 1, 2, \cdots$ single valued

$= -\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{3}{2}, \cdots$ double valued.

Eigenvalue of $J_{12} = -j - 1, -j - 2, \cdots$.

Trivial representation

$$j = -1.$$

Eigenvalue of $J_{12} = 0$.

The analog of the Peter–Weyl theorem, Theorem 9 of Bargmann, states that any function which is square integrable over the group manifold may be expanded in terms of the principal series and those members of the discrete series having $j > -\frac{1}{2}$. The virtue of our choice of representation parameter now becomes clear, for as far as square integrable functions are concerned we may discard the representations lying to the left of Re $j = -\frac{1}{2}$, viz., the supplementary series, the trivial representation and $D^{(\pm)}(-\frac{1}{2})$. The principal series and the remaining members of the discrete series may be identified with the background integral and nonsense channel terms, respectively, this identification being validated by the next section.

6. ANALYTIC CONTINUATION OF THE $O(2, 1)$ DECOMPOSITION

The set of irreducible unitary representations of $O(2, 1)$ which occur in the expansion of any square integrable function is just that set of kinematical terms which are obtained by the $O(3)$ channel expansion; the principal and discrete series being simply the background integral and nonsense channel contributions. It is most convenient to show the inverse, i.e., to express the amplitude in a form which permits an inverse Sommerfeld–Watson transformation to be performed, the result of which may be identified with the $O(3)$ series. In order to do this we must, in addition, continue in ζ to $|\zeta| < 1$. Two distinct difficulties arise. The principal series of the $O(2, 1)$ expansion diverges

if $|\zeta| < 1$. This can be overcome by using the dispersion relation (2.21) to consider direct and exchange forces independently; which simply corresponds to the introduction of signature. Even so, the partial-wave amplitude diverges as $\mathrm{Re}\, j \to \infty$, unless we utilize the invariance of the principal series integral under $j \to -(j+1)$ to redefine $\bar{T}_{\lambda\mu}(j,t)$ in such a manner that it is bounded as $\mathrm{Re}\, j \to \infty$. As will emerge, the natural way of doing this leads to an amplitude which has poles at integer values of $(j-\lambda)$, whose residues are identical with the crossed channel $O(3)$ partial-wave amplitudes. When these poles are crossed by shifting the principal series integral to the right, they give rise to a set of terms which cancels the discrete series and reproduces the familiar $O(3)$ expansion of the crossed channel. In this section, since we are concerned with kinematic singularities, we assume that the partial-wave amplitudes have no dynamical singularities for $\mathrm{Re}\, j > -\frac{1}{2}$; a discussion of the validity of this assumption is postponed until the next section.

We begin by using Eq. (2.21) to divide the amplitude into direct and exchange components

$$\bar{T}^D_{\lambda\mu}(\zeta,t) = \left(\frac{\zeta-1}{2}\right)^{\frac{1}{2}(\lambda-\mu)}\left(\frac{\zeta+1}{2}\right)^{\frac{1}{2}(\lambda+\mu)}\int_{\zeta_R}^{\infty} d\zeta' \frac{\bar{\rho}^R_{\lambda\mu}(\zeta',t)}{\zeta'-\zeta}, \tag{6.1}$$

$$\bar{T}^E_{\lambda-\mu}(-\zeta,t) = -e^{-i\pi\lambda}\left(\frac{\zeta-1}{2}\right)^{\frac{1}{2}(\lambda-\mu)}\left(\frac{\zeta+1}{2}\right)^{\frac{1}{2}(\lambda+\mu)} \times \int_{-\infty}^{-\zeta^L} d\zeta' \frac{\bar{\rho}^L_{\lambda\mu}(\zeta',t)}{\zeta'-\zeta}, \tag{6.2}$$

in terms of which

$$\bar{T}_{\lambda\mu}(\zeta,t) = \bar{T}^D_{\lambda\mu}(\zeta,t) - e^{i\pi\lambda}\bar{T}^E_{\lambda-\mu}(-\zeta,t). \tag{6.3}$$

The anti-symmetry of the principal series integral under $j \to -(j+1)$ enables it to be replaced by

$$\frac{1}{2i}\int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty} dj \frac{(2j+1)}{\sin\pi(j-\lambda)} \times [\bar{\mathcal{A}}_{\lambda\mu}(j,t) - \bar{\mathcal{B}}_{\lambda-\mu}(j,t)]\, d^j_{\lambda\mu}(\zeta), \tag{6.4}$$

where

$$\bar{\mathcal{A}}_{\lambda\mu}(j,t) = \pi^{-1}\sin\pi(j-\lambda)\int_1^{\infty} d\zeta e^j_{\lambda\mu}(\zeta)\bar{T}^D_{\lambda\mu}(\zeta,t) + \tfrac{1}{2}e^{-i\pi\lambda}\int_{-1}^{1} d\zeta\, d^j_{\lambda-\mu}(-\zeta)\bar{T}^D_{\lambda\mu}(\zeta,t), \tag{6.5}$$

$$\bar{\mathcal{B}}_{\lambda-\mu}(j,t) = -\pi^{-1}\sin\pi(j-\lambda)e^{i\pi\lambda}\int_1^{\infty} d\zeta e^j_{\lambda\mu}(\zeta) \times \bar{T}^E_{\lambda-\mu}(-\zeta,t) - \tfrac{1}{2}\int_{-1}^{1} d\zeta\, d^j_{\lambda-\mu}(-\zeta)\bar{T}^E_{\lambda-\mu}(-\zeta,t). \tag{6.6}$$

It follows from the asymptotic behavior[5]

$$e^j_{\lambda\mu}(\zeta) = O(|\zeta|^{-j-1}) \quad \text{as} \quad |\zeta| \to \infty \tag{6.7}$$

that the existence of $\bar{\mathcal{A}}_{\lambda\mu}(j,t)$ and $\bar{\mathcal{B}}_{\lambda-\mu}(j,t)$ for $\mathrm{Re}\, j = -\frac{1}{2}$ implies their existence for all j such that $\mathrm{Re}\, j \geq -\frac{1}{2}$. The significance of Eqs. (6.5) and (6.6) is that for $\mathrm{Re}\, j$ sufficiently large we may substitute for $\bar{T}^D_{\lambda\mu}(\zeta,t)$ and $\bar{T}^E_{\lambda-\mu}(-\zeta,t)$ using Eqs. (6.1) and (6.2), invert the orders of integration, and use Eq. (2.22), which is proven in the Appendix to show that

$$\bar{\mathcal{A}}_{\lambda\mu}(j,t) = -e^{-i\pi(j-\lambda)}\bar{A}_{\lambda\mu}(j,t), \tag{6.8}$$

$$\bar{\mathcal{B}}_{\lambda-\mu}(j,t) = -e^{i\pi\lambda}\bar{B}_{\lambda-\mu}(j,t), \tag{6.9}$$

where $\bar{A}_{\lambda\mu}(j,t)$ and $\bar{B}_{\lambda-\mu}(j,t)$ are the crossed channel $O(3)$ amplitudes defined by (3.6) and (3.7). If

$$|\bar{\rho}^L_{\lambda\mu}(\zeta,t)|, \quad |\bar{\rho}^R_{\lambda\mu}(\zeta,t)| \leq 0(|\zeta|^{-\alpha}),$$

then the above equations are valid for $\mathrm{Re}\, j > M - \alpha$, where

$$M = \max(|\lambda|, |\mu|).$$

Due to the j plane asymptotic behavior[5]

$$e^j_{\lambda\mu}(\zeta) = \kappa j^{-\frac{1}{2}}\exp[(j+\tfrac{1}{2})\log\{\zeta - (\zeta^2-1)^{\frac{1}{2}}\}], \quad \text{as} \quad |j| \to \infty. \tag{6.10}$$

It follows that $\bar{A}_{\lambda\mu}(j,t)$ and $\bar{B}_{\lambda\mu}(j,t)$ are bounded by $|j|^{-\frac{1}{2}}$ as $\mathrm{Re}\, j \to \infty$, which is not the case for the original $O(2,1)$ partial waves, unmodified by the terms which are symmetric under $j \to -(j+1)$.

The expansion is not yet in a form which may be continued in ζ to $|\zeta| < 1$, since it still contains the signature factors. However, before removing these, it is convenient to show the cancellation of the discrete series. To do this we transform the principal series integrals from $\mathrm{Re}\, j = -\frac{1}{2}$ to $\mathrm{Re}\, j = M - \frac{1}{2}$, while utilizing the symmetry of $d^j_{\lambda\mu}(\zeta)$ to restrict attention to the cases $\lambda \geq |\mu| > 0$.

The integrand consists of an entire function of j times the factor

$$\begin{aligned}\frac{\pi}{\sin\pi(j-\lambda)}\cdot\frac{\Gamma(j+\lambda+1)\Gamma(j-\mu+1)}{\Gamma(j-\lambda+1)\Gamma(j+\mu+1)} &= \frac{\pi}{\sin\pi(j-\lambda)}\cdot(j+\mu+1)\cdots(j+\lambda) \\ &\times (j-\lambda)\cdots(j-\lambda+1), \quad \lambda \geq \mu \geq 0 \\ &= \frac{\pi}{\sin\pi(j-\lambda)}(j+\lambda)\cdots(j-\lambda+1) \\ &\times (j-\mu)\cdots(j+\mu+1), \quad \lambda \geq -\mu \geq 0.\end{aligned} \tag{6.11}$$

When $\lambda \geq \mu \geq 0$ poles occur from $j = 0$ or $\frac{1}{2}$ up to $\mu - 1$ and from λ to ∞, while for $\lambda \geq -\mu \geq 0$

they occur only from λ to ∞. This is in agreement with the occurrence of discrete $O(2, 1)$ representations, as these do not contribute to amplitudes for which $\lambda\mu < 0$.

The residue of the integrand at these poles is

$$(4\pi i)^{-1}(2j+1)\Big[\int_1^\infty d\zeta\, d^j_{\lambda\mu}(\zeta)\bar{T}^D_{\lambda\mu}(\zeta, t) - e^{i\pi\lambda}\int_1^\infty d\zeta\, d^j_{\lambda\mu}(\zeta)\bar{T}^E_{\lambda-\mu}(-\zeta, t)\Big]$$

and hence their contribution to the amplitude is

$$-\theta(\mu)\sum_{k=0\,or\,\frac{1}{2}}^{\mu-1}(2k+1)\Big\{\tfrac{1}{2}\int_1^\infty d\zeta'\, d^j_{\lambda\mu}(\zeta')\bar{T}_{\lambda\mu}(\zeta', t)\Big\}\, d^j_{\lambda\mu}(\zeta), \tag{6.12}$$

which exactly cancels the contribution from the discrete series.

Assuming that $\bar{\mathcal{A}}_{\lambda\mu}(j, t)$ and $\bar{\mathcal{B}}_{\lambda\mu}(j, t)$ have no dynamical singularities to the right of $\mathrm{Re}\, j = -\frac{1}{2}$, the $O(2, 1)$ expansion becomes

$$\bar{T}_{\lambda\mu}(\zeta, t) = \tfrac{1}{2}i\int_{M-\frac{1}{2}-i\infty}^{M-\frac{1}{2}+i\infty} dj\, \frac{(2j+1)}{\sin\pi(j-\lambda)}\, e^{i\pi\lambda} \times [e^{-i\pi j}\bar{A}_{\lambda\mu}(j, t) - \bar{B}_{\lambda-\mu}(j, t)]\, d^j_{\lambda\mu}(\zeta), \tag{6.13}$$

where $\bar{A}_{\lambda\mu}(j, t)$ and $\bar{B}_{\lambda-\mu}(j, t)$ are the crossed channel $O(3)$ amplitudes. Although these are bounded by $|j|^{-\frac{1}{2}}$ as $|j| \to \infty$, since for $\zeta > 1$

$$d^j_{\lambda\mu}(\zeta) \sim |j|^{-\frac{1}{2}}\exp(\beta\,\mathrm{Re}\, j), \quad \text{as } \mathrm{Re}\, j \to \infty,\ \zeta = \cosh\beta,$$
$$\leq |j|^{-\frac{1}{2}}, \quad \text{as } \mathrm{Im}\, j \to \infty,$$

while for $|\zeta| < 1$

$$d^j_{\lambda\mu}(\zeta) \leq |j|^{-\frac{1}{2}}, \quad \text{as } \mathrm{Re}\, j \to \infty,$$
$$\sim |j|^{-\frac{1}{2}}\exp(|\theta\,\mathrm{Im}\, j|), \quad \text{as } \mathrm{Im}\, j \to \infty,\ \zeta = \cos\theta,$$

we must continue Eq. (6.13) to $|\zeta| < 1$ before allowing $\mathrm{Re}\, j$ to become infinite. This continuation is prevented by the signature factor $e^{-i\pi j}$, which may be removed by using[5]

$$e^{-i\pi(j-\lambda)}\, d^j_{\lambda\mu}(\zeta) = d^j_{\lambda-\mu}(-\zeta) + 2\pi^{-1}\sin\pi(j-\lambda)e^j_{\lambda\mu}(\zeta) \tag{6.14}$$

and noticing that in the absence of dynamical singularities in the right-hand plane

$$\int_{M-\frac{1}{2}-i\infty}^{M-\frac{1}{2}+i\infty} dj(2j+i)e^j_{\lambda\mu}(\zeta)\bar{A}_{\lambda\mu}(j, t) = 0, \tag{6.15}$$

which enables the amplitude to be expressed as

$$\bar{T}_{\lambda\mu}(\zeta, t) = \tfrac{1}{2}i\int_{M-\frac{1}{2}-i\infty}^{M-\frac{1}{2}+i\infty} dj\, \frac{(2j+1)}{\sin\pi(j-\lambda)} \times [\bar{A}_{\lambda\mu}(j, t)\, d^j_{\lambda-\mu}(-\zeta) - e^{i\pi\lambda}\bar{B}_{\lambda-\mu}(j, t)\, d^j_{\lambda\mu}(\zeta)], \tag{6.16}$$

if we now continue this amplitude in ζ to $|\zeta| < 1$ and then perform an inverse Sommerfeld–Watson transform, the resulting expression is

$$\bar{T}_{\lambda\mu}(\zeta, t) = \sum_{j=M}^{\infty}(2j+1) \times [\bar{A}_{\lambda\mu}(j, t) - e^{i\pi j}\bar{B}_{\lambda-\mu}(j, t)]\, d^j_{\lambda\mu}(\zeta),$$

which may be identified with the $O(3)$ expansion Eq. (3.8).

7. CONCLUSION

Whenever the amplitudes $\bar{\mathcal{A}}_{\lambda\mu}(j, t)$ and $\bar{\mathcal{B}}_{\lambda\mu}(j, t)$, defined by Eqs. (6.5) and (6.6) exist and have no dynamical singularities to the right of $\mathrm{Re}\, j = -\frac{1}{2}$ then the $O(2, 1)$ partial-wave expansion may be identified with the Regge continuation of the crossed channel amplitude. Their existence follows directly from the assumption of square integrability over the group manifold. However, due to the integration with respect to $\cosh\beta$ being over an infinite interval, $1 \leq \cosh\beta < \infty$, the general amplitude is not square integrable.

For the scattering of spinless particles the Martin–Froissart bound[12]

$$|T(s, t)| \leq Ks(\ln s)^2 \quad \text{as} \quad s \to \infty,\ t < 0$$

enables $\bar{\mathcal{A}}_{\lambda\mu}(j, t)$ $\bar{\mathcal{B}}_{\lambda\mu}(j, t)$, to be defined by their analytic continuations from $\mathrm{Re}\, j > 1$. The amplitude defined by Eq. (6.16) may now differ from the physical amplitude, but upon performing an inverse Sommerfeld–Watson transformation on it as in Sec. 6 the result agrees with the analytically continuable $O(3)$ partial-wave amplitude which itself may differ from the physical amplitude by just those terms $d^j_{\lambda\mu}(z)$, $j = 0, \frac{1}{2}, 1$, which correspond to the possible existence of elementary particles having spins 0, $\frac{1}{2}$, or 1.

On the basis of analogy with potential theory we may expect that all Regge poles are to the left of $\mathrm{Re}\, j = -\frac{1}{2}$ for sufficiently large negative t. In such circumstances, if we ignore the possibility of elementary particles, the representation of the amplitude may be identified with the $O(2, 1)$ expansion of a square integrable function. As t increases to positive values, singularities may penetrate the contour. We may infer that they are caused by the amplitude becoming nonsquare integrable.

ACKNOWLEDGMENTS

The author is grateful to Professor A. Salam and the IAEA for hospitality at the International

[12] M. Froissart, Phys. Rev. **123**, 1053 (1961); A. Martin, *ibid.* **129**, 1432 (1963).

Centre for Theoretical Physics, Trieste. He wishes to thank Professor A. Salam for suggesting the work and for his interest during its progress. He is also indebted to Dr. J. Strathdee, to whom the definition of $O(2, 1)$ helicity states is due, for much sage counsel and to C. Cook and A. Akyeampong for many discussions.

APPENDIX

The proof of Eq. (2.22), which is simply the generalization of the well-known relation between the Legendre polynomials of the first and second kinds, follows from the fact that

$$e^j_{\lambda\mu}(\zeta)[\tfrac{1}{2}(\zeta-1)]^{\frac{1}{2}(\lambda-\mu)}[\tfrac{1}{2}(\zeta+1)]^{\frac{1}{2}(\lambda+\mu)}$$

is an analytic function of ζ in the plane cut from 1 to $-\infty$, with discontinuity given by[5]

$$\begin{aligned}\text{Disc}\,&[e^j_{\lambda\mu}(\zeta)[\tfrac{1}{2}(\zeta-1)]^{\frac{1}{2}(\lambda-\mu)}[\tfrac{1}{2}(\zeta+1)]^{\frac{1}{2}(\lambda+\mu)}]\\ &= -i\pi\, d^j_{\lambda\mu}(\zeta)[\tfrac{1}{2}(\zeta-1)]^{\frac{1}{2}(\lambda-\mu)}[\tfrac{1}{2}(\zeta+1)]^{\frac{1}{2}(\lambda+\mu)} \qquad (-1 \le \zeta \le -1)\\ &= 2i \sin\pi(j-\lambda) e^j_{\lambda-\mu}(-\zeta)\\ &\quad\times [\tfrac{1}{2}(\zeta-1)]^{\frac{1}{2}(\lambda-\mu)}[\tfrac{1}{2}(\zeta+1)]^{\frac{1}{2}(\lambda+\mu)} \quad (\zeta < -1).\end{aligned}$$

In addition, since

$$e^j_{\lambda\mu}(\zeta) = O(|\zeta|^{-j-1}) \quad \text{as} \quad |\zeta| \to \infty,$$

$$\therefore \left| \int_{-\pi+\epsilon}^{\pi-\epsilon} d\theta\, |\zeta'|\, e^{i\theta} e^j_{\lambda\mu}(\zeta') \left(\frac{\zeta'-1}{2}\right)^{\frac{1}{2}(\lambda-\mu)} \left(\frac{\zeta'+1}{2}\right)^{\frac{1}{2}(\lambda+\mu)} \right| < \kappa \int_{-\pi+\epsilon}^{\pi-\epsilon} d\theta \frac{|\zeta'|^{-\mathrm{Re}\, j+M}}{|\zeta'-\zeta|}$$

and hence $\to 0$ as $|\zeta'| \to \infty$ for $\mathrm{Re}\, j > M-1$. Therefore, if ζ is any point not on the real axis < 1, and C is a finite circle, center ζ and radius $< |\zeta - 1|$, then by Cauchy's integral theorem

$$\begin{aligned}e^j_{\lambda\mu}(\zeta)&\left(\frac{\zeta-1}{2}\right)^{\frac{1}{2}(\lambda-\mu)}\left(\frac{\zeta+1}{2}\right)^{\frac{1}{2}(\lambda+\mu)}\\ &= \frac{1}{2\pi i}\int_c d\zeta' \frac{e^j_{\lambda\mu}(\zeta')[\tfrac{1}{2}(\zeta'-1)]^{\frac{1}{2}(\lambda-\mu)}[\tfrac{1}{2}(\zeta'+1)]^{\frac{1}{2}(\lambda+\mu)}}{\zeta'-\zeta}\\ &= -\frac{1}{2}\int_{-1}^{1} d\zeta' \frac{d^j_{\lambda\mu}(\zeta')[\tfrac{1}{2}(\zeta'-1)]^{\frac{1}{2}(\lambda-\mu)}[\tfrac{1}{2}(\zeta'+1)]^{\frac{1}{2}(\lambda+\mu)}}{\zeta'-\zeta}\\ &\quad + \frac{\sin\pi(j-\lambda)}{\pi}\\ &\quad\times \int_{-\infty}^{-1} d\zeta' \frac{e^j_{\lambda-\mu}(-\zeta')[\tfrac{1}{2}(\zeta'-1)]^{\frac{1}{2}(\lambda-\mu)}[\tfrac{1}{2}(\zeta'+1)]^{\frac{1}{2}(\lambda+\mu)}}{\zeta'-\zeta}\end{aligned}$$

for $\mathrm{Re}\, j > M-1$, from which Eq. (2.22) follows.

Tensor Operators and Mass Formula in the Minimal Extension of U_3 by Charge Conjugation*

A. O. BARUT AND H. KLEINERT
Department of Physics, University of Colorado, Boulder, Colorado

(Received 19 September 1966)

The transformation property of the mass (or mass square) M is specified in the extended group. The most general mass formula, $M = M_0 + a(\alpha)[I(I+1) - \frac{1}{4}Y^2 - K(\alpha)] + b_1(\alpha)NY + b_2(\alpha)Q_2Y$, is derived where $a(\alpha)$, $b_i(\alpha)$, $K(\alpha)$ are constants depending on a charge-even set α, specified in the text, and Q_2 is the second invariant operator of SU_3.

THE Gell-Mann–Okubo mass formula, $M = M_0 + a[I(I+1) - \frac{1}{4}Y^2] + bY$, is incomplete in the sense that it does not reflect, as it stands, the charge conjugation properties of the mass splitting. Thus, the last term should have opposite signs for baryons and antibaryons and should vanish for mesons. Therefore, the baryon number and charge conjugation parities must enter into the formula as well. In order to modify the form of the equation, we start from the invariance group $\mathcal{E} = \{U_3, \text{coset } U_3C\}$, which is the minimal extension[1] of U_3 by charge conjugation C.[2]

I. THE STATES

The states of irreducible representations of U_3 are labeled by $|N, Q_1, Q_2; I, I_3, Y\rangle$, where Q_1 and Q_2 are the two invariant (Casimir) operators of SU_3. Note that Q_1 and Q_2 are chosen to be the components of the maximal weight of the representation:

$$Q_1 = \tfrac{1}{2}(p+q), \quad Q_2 = \frac{p-q}{2\sqrt{3}}.$$

Since $CpC = q$ and $CqC = p$, we find $CQ_1C = Q_1$, $CQ_2C = -Q_2$. Also, $CNC = -N$. The quadratic Casimir operator is

$$\sum_{i=1}^{8} F_i^2 = \tfrac{1}{3}[p(p+3) + q(q+3) + pq].$$

The charge conjugation C can be adjoined to the representations of U_3 in two possible ways: either the states $C|\rangle$ are unrelated to the original states, then the extended group $\mathcal{E} = \{U_3, U_3C\}$ is represented in the Hilbert space $\begin{pmatrix} |\rangle \\ C|\rangle \end{pmatrix}$ of doubled states; or the states $C|\rangle$ are linearly related to the states $|\rangle$, then we can form the combinations $(2)^{-\frac{1}{2}}[|\rangle \pm C|\rangle]$ with definite C-parities η_c, which are denoted by

$$|\eta_c; N, Q_1, Q_2; I, I_3, Y\rangle, \quad \eta_c = \pm 1. \tag{1}$$

This second case occurs only if the invariant operators of U_3 are also invariant under C, that is, if $N = 0$, $Q_2 = 0$ in our choice of the Casimir operators (see note above) (i.e., self-adjoint representations of U_3).

II. THE INVARIANTS

The invariant operators out of which the most general invariant function is constructed is different in both cases. In the first case, where the states are doubled, the two representations of U_3 with $\pm N$, $\pm Q_2$ are the same irreducible representation of the extended group $\mathcal{E}$ characterized by N^2, Q_2^2. Thus, due to the additional requirement

$$Cf(N, Q_1, Q_2)C^{-1} = f(N, Q_1, Q_2), \tag{2}$$

all invariants are functions of only

$$N^2, Q_1, NQ_2, \quad \text{and} \quad Q_2^2. \tag{3}$$

The operator NQ_2 fixes the relative sign between N and Q_2 to distinguish, for example, between

$$\begin{pmatrix} 10, N=1 \\ \overline{10}, N=-1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \overline{10}, N=1 \\ 10, N=-1 \end{pmatrix};$$

only the first case is known to be realized for the $\frac{3}{2}^+$ baryons.

In the second case the invariants are functions of η_c and Q_1. Thus we can write in both cases the invariants as functions of the set

$$\alpha = \{\eta_c\delta_{N,0}, N^2, Q_1, NQ_2, Q_2^2\}. \tag{4}$$

* This work was supported by the United States Air Force Office of Scientific Research under Grant AF-AFOSR-30-65.

[1] The concept of minimal extension of a group by a discrete operation is discussed in detail in the articles by E. P. Wigner and L. Michel, in *Group Theoretical Concepts and Methods in Elementary Particle Physics*, F. Gürsey, Ed. (Gordon and Breach Science Publishers, Inc., New York, 1964); T. D. Lee and G. C. Wick, Phys. Rev. **148**, 1385 (1966).

[2] Mathematically the charge conjugation C defines an automorphism of U_3 and this automorphism characterizes the minimal extension. L. C. Biedenharn, J. Nuyts, and H. Ruegg, CERN preprint (1965); S. Okubo and N. Mukunda, Ann. Phys. (N.Y.) **36**, 311 (1966). Depending on whether the representation of the automorphism is inner or outer, one gets the two cases of doubling or no doubling of states discussed below.

III. THE TENSOR OPERATORS

The general tensor operator is first constructed for the group U_3 in the usual way from the infinitesimal generators F_i, $i = 0, 1, \cdots, 8$ in the form

$$T_i = t_i + t_{ij}F_j + t_{ijk}F_jF_k + \cdots, \tag{5}$$

where $t_{i_1 \cdots i_n}$ are all possible invariant symmetric tensors of the adjoint representation of U_3. This follows from

$$Ut_{ijk} \cdots F_jF_k \cdots U^{-1}$$

$$= t_{ijk} \cdots F_{j'}AdU_{j'j}F_{k'}\ AdU_{k'k} = t_{i'jk} \cdots F_jF_kAdU_{i'i}.$$

Hence

$$AdU_{j'j}AdU_{k'k}AdU_{i'i}t_{i'j'k'} = t_{ijk}.$$

Applying this to T_8, for example, one gets

$$T_8 = a[I(I+1) - \tfrac{1}{4}Y^2 - K] + bY, \tag{6}$$

where a, b, K are functions of the U_3-invariants N, Q_1, Q_2. The constant K has the effect of making the average of T_8 over a multiplet vanish and is given by

$$K = \tfrac{1}{3}\sum_{i=1}^{8} F_i^2 = \tfrac{1}{9}[p(p+3) + q(q+3) + pq]. \tag{7}$$

Then we impose the condition of charge invariance

$$CT_iC^{-1} = T_i. \tag{8}$$

IV. THE MASS FORMULA

We require that the mass splitting be invariant under C in addition to the T_8-property, i.e.,

$$CT_8C^{-1} = T_8. \tag{9}$$

(Note that $CF_8C^{-1} = -F_8$!) The condition (9) restricts the coefficient a in Eq. (6) to be a function only of the set α, Eq. (4); K remains the same because $CKC^{-1} = K$, while b has to be the most general odd function under C, hence

$$b = b_1(\alpha)N + b_2(\alpha)Q_2.$$

Therefore, the most general mass formula in $\mathfrak{S} = \{U_3, U_3C\}$ under the stated assumptions is

$$M = M_0 + a(\alpha)[I(I+1) - \tfrac{1}{4}Y^2 - K] + b_1(\alpha)NY + b_2(\alpha)Q_2Y. \tag{10}$$

V. OTHER CONCLUSIONS

(i) Because the group SU_n for $n \geq 3$ has only one outer automorphism, we do not expect any further extension of the internal symmetry group except the one discussed.

(ii) In the case of $N = 0$, only one value of η_c is known at present. Because the coefficients in the mass formula (10) depend on η_c, the states with $\eta_c = -1$ could lie higher.

(iii) Note the presence of the term b_2Q_2Y in Eq. (10) which distinguishes, for example, the $N = 1$ octet and decouplet even if the coefficients $a_1b_i(\alpha)$ are the same for these two multiplets.

Evaluation of Phase-Space Integrals*

GRAHAM H. CAMPBELL, JOSEPH V. LEPORE, AND ROBERT J. RIDDELL, JR.
Lawrence Radiation Laboratory, University of California, Berkeley, California

(Received 2 September 1966)

The relativistic phase-space integral over the submanifold defined by total momentum zero and fixed total energy is reduced to a single contour integration. The number of particles involved and their masses are arbitrary. It is shown that the contour integration may be readily approximated by the saddle-point technique and yields a result which is easily handled by a computer. In the nonrelativistic and extreme relativistic limits, this method leads to expressions for the phase space which may be obtained from the exact results for these cases by replacing Γ-function factors by the Stirling approximation.

I. INTRODUCTION

INTEREST in the evaluation of phase-space integrals involving the constraints of momentum and energy conservation arises from the study of multiple production of particles in high-energy nuclear collisions. A knowledge of the phase-space factor for a particular process allows the separation of the dynamical and kinematical features peculiar to the situation. For example, a knowledge of phase-space factors can be important for the determination of whether very-short-lived particles or "resonances" play a role in a particular reaction.

It has been shown that the general relativistic phase-space integrals are easily reduced to two integrations.[1] This paper shows how still another integration may be performed.

II. REDUCTION OF THE INTEGRAL

The integral to be evaluated is (using units in which $c = 1$)

$$S_n(E) = \int \delta\Big(E - \sum_{i=1}^{n} \omega_i\Big)\, \delta(\textstyle\sum \mathbf{p}_i) \prod_{i=1}^{n} \frac{d^3 p_i}{\omega_i}, \qquad (1)$$

where $\omega_i = (\mathbf{p}_i^2 + m_i^2)^{\frac{1}{2}}$. This is a Lorentz-invariant quantity which we evaluate in the center-of-momentum frame. If we insert a Fourier representation of the δ function we may write

$$S_n(E) = \frac{1}{(2\pi)^4} \int_{-\infty - i\epsilon}^{\infty - i\epsilon} d\alpha e^{i\alpha E} \int d^3\boldsymbol{\lambda} \prod_{i=1}^{n} J(\boldsymbol{\lambda}, \alpha, m_i), \qquad (2)$$

where

$$J(\boldsymbol{\lambda}, \alpha, m) = \int \frac{d^3 p}{\omega} \exp\,[i(\boldsymbol{\lambda} \cdot \mathbf{p} - \alpha\omega)]. \qquad (3)$$

The variable α has been given a small negative imaginary part to make the integration over momenta well defined. After the trivial angular integrations are performed we can write

$$J(\boldsymbol{\lambda}, \alpha, m) = -(2\pi/\lambda)(d/d\lambda) I(\lambda, \alpha, m), \qquad (4)$$

where

$$I(\lambda, \alpha, m) = \int_{-\infty}^{\infty} \frac{dp}{\omega} e^{i(\lambda p - \alpha\omega)}. \qquad (5)$$

If we now let

$$p = m \sinh\theta,$$
$$\omega = m \cosh\theta,$$
$$\alpha = (\alpha^2 - \lambda^2)^{\frac{1}{2}} \cosh\psi,$$
$$\lambda = (\alpha^2 - \lambda^2)^{\frac{1}{2}} \sinh\psi,$$

then[2]

$$I(\lambda, \alpha, m) = \int_{-\infty}^{\infty} d\theta \exp\,[-im(\alpha^2 - \lambda^2)^{\frac{1}{2}} \cosh(\theta - \psi)] \qquad (6)$$

$$= -i\pi H_0^{(2)}[m(\alpha^2 - \lambda^2)^{\frac{1}{2}}]. \qquad (7)$$

Therefore

$$J(\lambda, \alpha, m) = 2\pi^2 i m (\alpha^2 - \lambda^2)^{-\frac{1}{2}} H_1^{(2)}[m(\alpha^2 - \lambda^2)^{\frac{1}{2}}], \qquad (8)$$

$$S_n(E) = \frac{1}{4\pi^3} \int_{-\infty - i\epsilon}^{\infty - i\epsilon} d\alpha \int_0^{\infty} \lambda^2\, d\lambda e^{i\alpha E} f[(\alpha^2 - \lambda^2)^{\frac{1}{2}}], \qquad (9)$$

where

$$f(z) = \prod_{k=1}^{n} \left[\frac{2\pi^2 i m_k}{z} H_1^{(2)}(m_k z)\right]. \qquad (10)$$

The sole role of ϵ in this equation is to define the continuation of the square root. If we had carried the analysis to this point using an arbitrary Lorentz frame, this expression would be a Fourier transformation in a space with a timelike dimension. Thus the following steps seem to be a generalization of a theorem of Bochner's on Fourier transformations of radial functions.[3]

We now replace the α integration by an integration over $\rho = (\alpha^2 - \lambda^2)^{\frac{1}{2}}$. This necessitates a separate

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

[1] J. V. Lepore and R. N. Stuart, Phys. Rev. 94, 1724 (1954).

[2] G. N. Watson, *Theory of Bessel Functions* (The Macmillan Company, New York, 1948), 2nd ed., p. 180.

[3] S. Bochner and K. Chandrasekharan, *Fourier Transforms* (Princeton University Press, Princeton, New Jersey, 1949), p. 69.

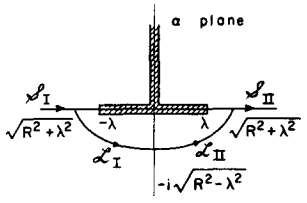


FIG. 1. Integration contour in the α plane for $\lambda < R$.

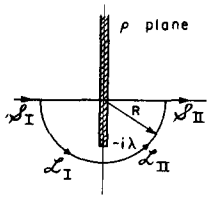


FIG. 2. Integration contour in the ρ plane for $\lambda < R$.

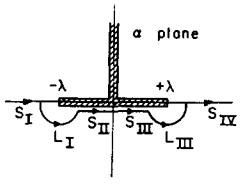


FIG. 3. Integration contour in the α plane for $\lambda > R$.

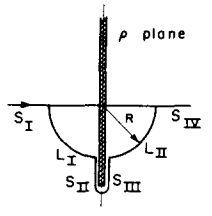


FIG. 4. Integration contour in the ρ plane for $\lambda > R$.

consideration of different domains of integration. We write

$$S_n(E) = \frac{1}{4\pi^3}\left\{\int_0^R d\lambda \int_{C_1} d\alpha + \int_R^\infty d\lambda \int_{C_2} d\alpha\right\} \times \lambda^2 e^{i\alpha E} f[(\alpha^2 - \lambda^2)^{\frac{1}{2}}]. \quad (11)$$

In the first integral in Eq. (11), $\lambda \leq R$. In this case, we first distort the original α contour so that it follows the real axis except for an arc below the singular points at $\pm\lambda$ (see Fig. 1). For the variable ρ, there is a corresponding contour traced out in the ρ plane as α follows its contour (see Fig. 2). For convenience, we choose the arc in the ρ plane as a half circle of radius R, so the α traces out an arc which is similar to an ellipse. In the second integral, $\lambda \geq R$, the α contour is distorted to follow the real axis except for separate arcs under the singularities at $\alpha = \pm\lambda$ (see Fig. 3). As before, we have a corresponding path in ρ and we choose the nonstraight portions to be circles of radius R (see Fig. 4).[4] We denote the contributions to S_n from the various pieces of the α contours simply by the labels as indicated in Figs. 1–4. The integrals can now all be expressed in terms of a real independent variable, but we must treat each one separately, taking into account the analytic continuation in the appropriately cut plane for the various functions involved.

Let us consider, for example, the calculation of S_I:

$$S_I = \frac{1}{4\pi^3}\int_R^\infty d\lambda \int_{-\infty}^{-(\lambda^2+R^2)^{\frac{1}{2}}} d\alpha \lambda^2 e^{i\alpha E} f[(\alpha^2 - \lambda^2)^{\frac{1}{2}}]. \quad (12)$$

If we now choose ρ as the integration variable, and introduce the variable r, using

$$\rho = (\alpha^2 - \lambda^2)^{\frac{1}{2}} = -r, \quad \text{for} \quad r > 0, \qquad \alpha = (\rho^2 + \lambda^2)^{\frac{1}{2}} = -(r^2 + \lambda^2)^{\frac{1}{2}}, \quad (13)$$

we have

$$S_I = -\frac{1}{4\pi^3}\int_R^\infty d\lambda \int_{-\infty}^{-R} d\rho \rho \lambda^2 (\lambda^2 + r^2)^{-\frac{1}{2}} \times \exp[-iE(\lambda^2 + r^2)^{\frac{1}{2}}] f(\rho). \quad (14)$$

The phase of ρ is $-\pi$. If we now make the same change of variable in S_I, interchange order of integration in both S_I and S_I, and add the two, we have, in terms of the variable $u = |(\lambda/r)^2 + 1|^{\frac{1}{2}}$,

$$S_I + S_I = \frac{1}{4\pi^3}\int_{-\infty}^{-R} d\rho f(\rho)\rho^3 \int_1^\infty du(u^2 - 1)^{\frac{1}{2}} e^{-iEru}. \quad (15)$$

We must now face the problem, which we ignored above, that the u integration is divergent. However, it is a limiting case of a convergent integral, namely that in which r has a negative imaginary part. This is because we are dealing with generalized functions [recall the Fourier representation for the δ function introduced into Eq. (2)]. We define this function and others to follow as the limit of the generalized function as parameters approach their final value through values which make the integral convergent. These statements applied here to the u integration actually refer to the method by which the original δ function was represented. With this intepretation we now have[5]

$$S_I + S_I = \left(\frac{1}{4\pi^3}\right)\int_{-\infty}^{-R} d\rho f(\rho)\rho^3 \left[\frac{i\pi}{2Er} H_1^{(2)}(Er)\right]. \quad (16)$$

[4] In Figs. 3 and 4 the contours are shown displaced from the cuts for clarity, although in fact they are to be taken on the cuts.

[5] Reference 2, p. 167.

An exactly analogous calculation, this time using

$$\rho = (\alpha^2 - \lambda^2)^{\frac{1}{2}} = r, \quad \text{for} \quad r > 0,$$
$$\alpha = (\rho^2 + \lambda^2)^{\frac{1}{2}} = (\lambda + r^2)^{\frac{1}{2}} \tag{17}$$

yields[5]

$$S_{IV} + \mathcal{S}_{II} = \left(\frac{1}{4\pi^3}\right)\int_R^\infty d\rho f(\rho)\rho^3\left[\frac{-i\pi}{2Er} H_1^{(1)}(Er)\right]. \tag{18}$$

Similar calculations using

$$\rho = (\alpha^2 - \lambda^2)^{\frac{1}{2}} = -ir, \quad \text{for} \quad r > 0,$$
$$\alpha = (\rho^2 + \lambda^2)^{\frac{1}{2}} = \mp(\lambda^2 - r^2)^{\frac{1}{2}} \tag{19}$$

(minus sign for S_{II}, plus for S_{III}), followed by the introduction of the variable $u = |(\lambda/r)^2 - 1|^{\frac{1}{2}}$, give

$$S_{II} = \frac{1}{4\pi^3}\int_{-iR}^{-i\infty} d\rho f(\rho)\rho^3 \int_0^\infty du(u^2 + 1)^{\frac{1}{2}} e^{-iEru},$$
$$S_{III} = \frac{1}{4\pi^3}\int_{-i\infty}^{-iR} d\rho f(\rho)\rho^3 \int_0^\infty du(u^2 + 1)^{\frac{1}{2}} e^{-iEru}. \tag{20}$$

Now, in a manner consistent with our interpretation of the u integrals, we can deform the u contours through the convergent quadrant to the imaginary axis, avoiding the branch points at $\pm i$ in the appropriate manner. Then using the variable $v = iu$ in S_{II} and $v = -iu$ in S_{III} we find that the contributions to the two integrals from the region $0 \leq v \leq 1$ cancel, while the remainder gives[6]

$$S_{II} + S_{III} = -\frac{1}{2\pi^3}\int_{-iR}^{-i\infty} d\rho f(\rho)\rho^3 \int_1^\infty dv(v^2 - 1)^{\frac{1}{2}} e^{-Erv}$$
$$= \frac{1}{2\pi^3}\int_{-iR}^{-i\infty} d\rho f(\rho)\rho^3\left[-\frac{K_1(Er)}{Er}\right]. \tag{21}$$

To evaluate the loop integrals (Figs. 1 and 2) we let

$$\alpha = -(\lambda^2 + R^2 e^{2i\theta})^{\frac{1}{2}} \tag{22}$$

in both L_I and $\mathcal{L}_I$, and take θ as the new variable. In both cases θ ranges from 0 to $+\pi/2$. We may now add L_I and $\mathcal{L}_I$, with the result that λ ranges from zero to infinity. We can now go back to $\rho = Re^{i\theta}$, and we set

$$L_I + \mathcal{L}_I = \frac{1}{4\pi^3}\int_{-R}^{-iR} d\rho\rho f(p)\int_0^\infty d\lambda\lambda^2(\lambda^2 + \rho^2)^{-\frac{1}{2}}$$
$$\times \exp\left[iE(\lambda^2 + \rho^2)^{\frac{1}{2}}\right]. \tag{23}$$

In this expression the phase of the square root is to be chosen so that it becomes a negative real quantity as ρ becomes real and negative. Thus the function represented by the λ integration is the analytic continuation of the corresponding function in $S_I + \mathcal{S}_I$ [see Eq. (14)] as ρ follows the circular arc from $-R$ to $-iR$. The corresponding calculation for $L_{II} + \mathcal{L}_{II}$ yields the corresponding continuation of $S_{IV} + \mathcal{S}_{II}$ with the net result that the R limits on the remaining ρ integration in both $S_I + \mathcal{S}_I$ and $S_{IV} + \mathcal{S}_{II}$ can be replaced by $-iR$ upon adding the contribution of the loop integrals.

The contour of the integral for $S_{II} + S_{III}$ may be deformed, by Jordan's lemma, to go to infinity along the real axis and may then be combined with the other integrals. Upon replacing r by ρ times an appropriate phase factor we have

$$S_I + \mathcal{S}_I + L_I + \mathcal{L}_I + \tfrac{1}{2}(S_{II} + S_{III})$$
$$= \frac{1}{4\pi^3}\int_{-\infty}^{-iR} d\rho\rho^2 f(\rho)\left\{\frac{i\pi}{2E}\left[H_1^{(2)}(E\rho e^{i\pi}) + H_1^{(2)}(E\rho)\right]\right\},$$
$$S_{IV} + \mathcal{S}_{II} + L_{II} + \mathcal{L}_{II} + \tfrac{1}{2}(S_{II} + S_{III}) \tag{24}$$
$$= \frac{1}{4\pi^3}\int_{-iR}^{\infty} d\rho\rho^2 f(\rho)\left\{\frac{-i\pi}{2E}\left[H_1^{(1)}(E\rho) + H_1^{(2)}(E\rho)\right]\right\}.$$

The sum of the Hankel functions yields $-2J_1(E\rho)$ in the first case and $2J_1(E\rho)$ in the second, so we have

$$S_n(E) = \frac{-i}{(2\pi)^2 E}\int d\rho\rho^2 f(\rho)J_1(E\rho), \tag{25}$$

where $f(\rho)$ is given by Eq. (10). The ρ contour runs from $-\infty$ to $+\infty$ below the origin. If this contour is chosen to be symmetric under $\rho \to e^{-i\pi}\rho^*$, then it can be shown that the contribution from the left half of the contour is the negative complex conjugate of the contribution from the right half. Thus the above expression for S_n is real, as is required.

We note that we may also write

$$S_n(E) = \frac{-i}{2(2\pi)^2 E}\int d\rho\rho^2 f(\rho)H_1^{(1)}(E\rho), \tag{26}$$

since the contour integral in which $H_1^{(2)}$ replaces $H_1^{(1)}$ vanishes because the integrand is analytic in the entire lower half-plane. This form is convenient for consideration of the nonrelativistic limit.

III. EVALUATION OF THE INTEGRAL

The integrand of the expression for S_n [Eq. (25)] is, apart from constant, real factors,

$$g(\rho) = -i\rho^2 J_1(E\rho)\prod_{k=1}^{n}\frac{H_1^{(2)}(m_k\rho)}{-i\rho}. \tag{27}$$

This function has a single saddle point on the negative imaginary axis because it takes the form

$$g(-iy) = y^2 I_1(Ey)\prod_{k=1}^{n}\left[\frac{2}{\pi y} K_1(m_k y)\right]. \tag{28}$$

[6] Reference 2, p. 172.

Paths of constant phase in $g(\rho)$ generally connect the consecutive zeros of $J_1(E\rho)$, and there are saddle points for each of these segments. Since $g(\rho)$ alternates in sign on different sections of these portions of the contour, large cancellations are expected, and, further, the integrand has its largest value at the saddle point on the negative imaginary axis. Thus the dominant contribution to the integral, taken along a path of constant phase, comes from the neighborhood of the saddle point on the negative imaginary axis. We may therefore approximate S_n by the standard saddle-point technique. Thus if we let $\rho = -iy$, we find that the saddle point is located at the solution of

$$\frac{1-2n}{y} + \frac{EI_0(Ey)}{I_1(Ey)} - \sum_{k=1}^{n} \frac{m_k K_0(m_k y)}{K_1(m_k y)} = 0. \quad (29)$$

Let y_0 be this root. Then

$$S_n(E) \approx \frac{y_0^2 I_1(Ey_0)}{(2\pi)^{\frac{3}{2}} E} \prod_{k=1}^{n} \left[\frac{4\pi m_k}{y_0} K_1(m_k y_0)\right] \times \left\{\frac{4n-2}{y_0^2} + E^2 + \sum_{k=1}^{n} m_k^2 - \left[\frac{EI_0(Ey_0)}{I_1(Ey_0)}\right]^2 - \sum_{k=1}^{n}\left[\frac{m_k K_0(m_k y_0)}{K_1(m_k y_0)}\right]^2\right\}^{-\frac{1}{2}}. \quad (30)$$

The accuracy of this approximation can be judged by a comparison with the exact results for the nonrelativistic and extreme relativistic limits.

On examination of Eq. (28) one finds that the saddle point occurs as a result of a balance between the decreasing functions in the product over k and the increasing term, $y^2 I_1(Ey)$. As the energy is reduced the minimum is reached for larger and larger y. Thus, to obtain the exact nonrelativistic limit, the predominant contribution to the integral arises from large values of ρ and we may use the asymptotic expansions for the Hankel functions in Eq. (26). After some rearrangement we find[7]

$$S_n \approx (2\pi)^{(3n-5)/2} P^{\frac{1}{2}} E^{-\frac{3}{2}} \sum_{l=0}^{\infty} \frac{(-1)^l (1,l)}{(2E)^l} \times \sum_{j_1=0}^{\infty} \cdots \sum_{j_n=0}^{\infty} \left[\prod_{k=1}^{n} \frac{(1,j_k)}{(2m_k)^{j_k}}\right] \times \left\{ i^{[(5n-5)/2]-l-\Sigma j_k} \int d\rho e^{iT\rho} \rho^{[-(3n-2)/2]-l-\Sigma j_k}\right\}. \quad (31)$$

We have used the abbreviation P for the product of the masses and will use M for their sum. The kinetic energy is $T(=E-M)$. Now the term in braces is simply

$$\frac{T^{\frac{3}{2}n-\frac{5}{2}+l+\Sigma j_k}}{\Gamma\left(\frac{3}{2}n - \frac{3}{2} + l + \sum_k j_k\right)}. \quad (32)$$

[7] Reference 2, p. 198. The symbols (v, m) are defined there.

Thus we have obtained an asymptotic expansion of S_n in powers of T (if the inverse powers of E are so expanded). The first term of this expansion is

$$S_n^{NR} = (2\pi)^{\frac{1}{2}(3n-3)} P^{\frac{1}{2}} M^{-\frac{3}{2}} \frac{T^{\frac{1}{2}(2n-5)}}{\Gamma[\frac{1}{2}(3n-3)]}. \quad (33)$$

This agrees with the nonrelativistic phase space computed by more elementary methods.

In this limit we can examine the accuracy of the saddle-point approximation. The saddle point is located at

$$\rho_0 = -\tfrac{3}{2}i(n-1)/T. \quad (34)$$

Application of the saddle-point method to the integral yields the above result with the Γ function replaced by Stirling's approximation to it. For $n = 2$ this approximation is in error by about 6%, and the error decreases as n increases. We note that this estimate applies only when E_{ρ_0} and $m_k\rho_0$ are all large, that is,

$$\tfrac{1}{2}(3n-3)(m_k/T) \gg 1; \quad (35)$$

otherwise the saddle point of the original integral would not occur in the asymptotic region of the Hankel functions.

We can also easily determine the asymptotic value of the phase space in the extreme relativistic case. In this case, we expect ρ_0 to be very small. Thus we can expand the $H_1^{(2)}(m_k\rho)$ for small values of the argument. If we assume that E is large, however, and use the asymptotic form of $H_1^{(1)}(E)$, we then deduce that

$$\rho_0 = -i(2n - \tfrac{3}{2})/E. \quad (36)$$

Thus $E\rho_0$ is large only if $2n - \frac{3}{2} \gg 1$. On the other hand, we can carry out the integration in Eq. (25) without expanding $J_1(\rho E)$, and we find

$$S_n^{ER} = \frac{2\pi^{n-1}}{(n-1)!\,(n-2)!} E^{2n-4}. \quad (37)$$

This result agrees with the saddle-point approximation again to the extent that the factorials are replaced by the Stirling approximation. In this case the Stirling approximation is not as good as before, and an 11% error is found for $n = 3$. The ratio of the approximate to the exact result approaches 1 with reasonable rapidity as n increases. This is to be expected, for as n increases the saddle point of this integral moves into the asymptotic region of J_1, where the saddle-point method was applied to find S_n.

IV. APPLICATIONS AND VARIATIONS

Extensive numerical calculations have been made with the phase-space or "statistical" model by use of this approximation.[8] Numerous comparisons with

[8] G. H. Campbell, University of California, Lawrence Radiation Laboratory Report No. UCRL-16315 (1965).

the exact two- and three-body phase space have confirmed the error estimates presented in Sec. III.

The noncovariant form of the phase-space integral may be treated by similar methods to those used for the covariant form. The noncovariant form is

$$S_n(E) = \int \delta\left(E - \sum_{i=1}^{n} \omega_i\right) \delta^{(3)}\left(\sum_{i=1}^{n} p_i\right) \prod_{i=1}^{n} d^3\mathbf{p}_i. \quad (38)$$

This has previously been reduced to[1]

$$S_n = \frac{1}{4\pi^3} \int_{-\infty - i\infty}^{\infty - i\epsilon} d\alpha \int_0^\infty d\lambda \lambda^2 \alpha^n e^{i\alpha E} \prod_{k=1}^{n} \times \left[\frac{2\pi^2 m_k}{\alpha^2 - \lambda^2} H_2^{(2)}(m_k(\alpha^2 - \lambda^2)^{\frac{1}{2}})\right]. \quad (39)$$

The only essential difference between this and Eq. (9) is the factor α^n. These integrals define generalized functions; hence we may replace this factor by n-fold differentiation with respect to the energy. Then, using the results of the covariant calculation, we may write

$$S_n(E) = \frac{-i}{(2\pi)^2} \int d\rho \rho^2 f(\rho) \left(\frac{d}{dE}\right)^n \frac{J_1(E\rho)}{\rho}, \quad (40)$$

where

$$f(\rho) = \prod_{k=1}^{n} \left[\frac{2\pi^2 m_k}{i\rho^2} H_2^{(2)}(m_k\rho)\right]. \quad (41)$$

Although the saddle-point method is in principle applicable to this integral it is not convenient for numerical approximation, because the form of the integrand depends upon n.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Canonical Definition of Wigner Coefficients in U_n

L. C. BIEDENHARN AND A. GIOVANNINI*
Department of Physics, Duke University, Durham, North Carolina

J. D. LOUCK
Los Alamos Scientific Laboratory, Los Alamos, New Mexico

(Received 21 April 1966; revised manuscript received 30 December 1966)

Two general results applicable to the problem of a canonical definition of the Wigner coefficient in U_n are demonstrated: (1) the existence of a canonical imbedding of $U_n \times U_n$ into U_{n^2} and (2) a general factorization lemma for operators defined in the boson calculus. Using these results, a resolution of the multiplicity problem for U_3 is demonstrated, in which all degenerate operators are shown to split completely upon projection into U_2.

I. INTRODUCTION

AN extension of the techniques of the Wigner–Racah angular momentum calculus (WRC) from SU_2 to an arbitrary group is a problem of evident importance in quantum mechanics, and it has been recognized as such since the problem was first formulated in the works of Racah[1] and of Wigner.[2] A very extensive literature[3] has developed, especially in recent years, and partial solutions (that is, for certain groups) have been developed.[4] The unitary groups U_n play a special role in this problem, partly for their own sake as higher symmetry groups in elementary particle and nuclear physics, but more basically as a structure sufficiently large to encompass all compact groups. Thus, for example, it appears likely, owing to the close relationship between the symmetric and unitary groups,[5] that a solution for U_n entails a corresponding solution for S_n.

The present work has as its objective to demonstrate

* On leave from Istituto di Fisica dell'Università, Torino, Italy.

[1] G. Racah, Ergeb. Exakt. Naturw. **37**, 28 (1965) (this paper is the published version of Racah's Princeton Lectures, 1951).

[2] E. P. Wigner, in *Selected Papers on the Quantum Theory of Angular Momentum*, L. C. Biedenharn and H. van Dam, Eds., (Academic Press Inc., New York, 1965) (this paper is the published version of Wigner's unpublished paper, 1940).

[3] Recent references which cite the earlier literature extensively are J. G. Nagel and M. Moshinsky, J. Math. Phys. **6**, 682 (1965), and J. D. Louck, J. Math. Phys. **6**, 1786 (1965); see also Ref. 4.

[4] This has been discussed in a series of papers by G. E. Baird and L. C. Biedenharn (a) J. Math. Phys. **4**, 436 (1963); (b) J. Math. Phys. **4**, 1499 (1963); (c) J. Math. Phys. **5**, 1723 (1964); (d) J. Math. Phys. **5**, 1730 (1964); (e) J. Math. Phys. **6**, 1847 (1965). The present paper continues this investigation.

[5] G. de B. Robinson, *Representation Theory of the Symmetric Group* (University of Toronto Press, Toronto, 1961), cf. Chap. III, especially.

the exact two- and three-body phase space have confirmed the error estimates presented in Sec. III.

The noncovariant form of the phase-space integral may be treated by similar methods to those used for the covariant form. The noncovariant form is

$$S_n(E) = \int \delta\left(E - \sum_{i=1}^{n} \omega_i\right) \delta^{(3)}\left(\sum_{i=1}^{n} p_i\right) \prod_{i=1}^{n} d^3\mathbf{p}_i. \quad (38)$$

This has previously been reduced to[1]

$$S_n = \frac{1}{4\pi^3} \int_{-\infty - i\infty}^{\infty - i\epsilon} d\alpha \int_0^\infty d\lambda \lambda^2 \alpha^n e^{i\alpha E} \prod_{k=1}^{n} \times \left[\frac{2\pi^2 m_k}{\alpha^2 - \lambda^2} H_2^{(2)}(m_k(\alpha^2 - \lambda^2)^{\frac{1}{2}})\right]. \quad (39)$$

The only essential difference between this and Eq. (9) is the factor α^n. These integrals define generalized functions; hence we may replace this factor by n-fold differentiation with respect to the energy. Then, using the results of the covariant calculation, we may write

$$S_n(E) = \frac{-i}{(2\pi)^2} \int d\rho \rho^2 f(\rho) \left(\frac{d}{dE}\right)^n \frac{J_1(E\rho)}{\rho}, \quad (40)$$

where

$$f(\rho) = \prod_{k=1}^{n} \left[\frac{2\pi^2 m_k}{i\rho^2} H_2^{(2)}(m_k\rho)\right]. \quad (41)$$

Although the saddle-point method is in principle applicable to this integral it is not convenient for numerical approximation, because the form of the integrand depends upon n.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Canonical Definition of Wigner Coefficients in U_n

L. C. BIEDENHARN AND A. GIOVANNINI*
Department of Physics, Duke University, Durham, North Carolina

J. D. LOUCK
Los Alamos Scientific Laboratory, Los Alamos, New Mexico

(Received 21 April 1966; revised manuscript received 30 December 1966)

Two general results applicable to the problem of a canonical definition of the Wigner coefficient in U_n are demonstrated: (1) the existence of a canonical imbedding of $U_n \times U_n$ into U_{n^2} and (2) a general factorization lemma for operators defined in the boson calculus. Using these results, a resolution of the multiplicity problem for U_3 is demonstrated, in which all degenerate operators are shown to split completely upon projection into U_2.

I. INTRODUCTION

AN extension of the techniques of the Wigner–Racah angular momentum calculus (WRC) from SU_2 to an arbitrary group is a problem of evident importance in quantum mechanics, and it has been recognized as such since the problem was first formulated in the works of Racah[1] and of Wigner.[2] A very extensive literature[3] has developed, especially in recent years, and partial solutions (that is, for certain groups) have been developed.[4] The unitary groups U_n play a special role in this problem, partly for their own sake as higher symmetry groups in elementary particle and nuclear physics, but more basically as a structure sufficiently large to encompass all compact groups. Thus, for example, it appears likely, owing to the close relationship between the symmetric and unitary groups,[5] that a solution for U_n entails a corresponding solution for S_n.

The present work has as its objective to demonstrate

* On leave from Istituto di Fisica dell'Università, Torino, Italy.

[1] G. Racah, Ergeb. Exakt. Naturw. **37**, 28 (1965) (this paper is the published version of Racah's Princeton Lectures, 1951).

[2] E. P. Wigner, in *Selected Papers on the Quantum Theory of Angular Momentum*, L. C. Biedenharn and H. van Dam, Eds., (Academic Press Inc., New York, 1965) (this paper is the published version of Wigner's unpublished paper, 1940).

[3] Recent references which cite the earlier literature extensively are J. G. Nagel and M. Moshinsky, J. Math. Phys. **6**, 682 (1965), and J. D. Louck, J. Math. Phys. **6**, 1786 (1965); see also Ref. 4.

[4] This has been discussed in a series of papers by G. E. Baird and L. C. Biedenharn (a) J. Math. Phys. **4**, 436 (1963); (b) J. Math. Phys. **4**, 1499 (1963); (c) J. Math. Phys. **5**, 1723 (1964); (d) J. Math. Phys. **5**, 1730 (1964); (e) J. Math. Phys. **6**, 1847 (1965). The present paper continues this investigation.

[5] G. de B. Robinson, *Representation Theory of the Symmetric Group* (University of Toronto Press, Toronto, 1961), cf. Chap. III, especially.

two basic results related to the existence of a canonical[6] solution to the problem of extending the Racah–Wigner calculus to the complete family of unitary groups U_n. Earlier [see Ref. 4(d)] it was shown that the splitting between the two independent octet operators (F and D in Gell-Mann's notation) could be extended to all operators in SU_3 and thereby defined a resolution of the multiplicity problem for this particular group. As is often the case, a more general viewpoint can yield results that are at once both simpler to understand and more far-reaching; by focusing attention on the complete family of groups U_n this turns out to be the case here. Accordingly, we do not assume any detailed knowledge of our earlier papers on this subject and emphasize instead the basically simple group-theoretic structure of our solution by developing the results directly.

From a group-theoretic point of view, there is but one type of problem that enters in the extension of the WRC to any group: this is the problem of labeling (uniquely) the irreducible representations (irreps) of a subgroup that occur in the restriction to this subgroup of an irrep of a larger group.

Put this way the problem seems rather vague and unattractive; let us illustrate the content of the problem by considering the "state-labeling problem for SU_n." The subgroup to be considered here is the Abelian group H (generated by the Cartan subalgebra) of the *generalized charges*, for example, the ordinary charge Q and the hypercharge Y for SU_3. For a given irrep of SU_n, the charges (denoting irreps of H) are well known to be insufficient labels. The solution to the labeling problem is equally well known: it is the chain of subgroups denoted by $U_n \supset U_1 \times U_{n-1} \supset \cdots \supset H$ whose representation labels, by the Weyl branching theorem, uniquely label all states belonging to a given irrep of U_n. The Gel'fand pattern [Ref. 4(b)] is just a convenient shorthand to express this labeling induced by the Weyl branching law.

It is tempting to denote this elegant solution to the U_n state labeling problem as "canonical." In the sense in which Artin[7] uses this term[8] this would be incorrect, for there are indeed many free choices involved. The situation is rather like that discussed by Artin of obtaining an affine plane from a projective plane by deleting a particular line. A canonical construction has to be explained as *an equivalence class* corresponding to the designation of a particular U_1, out of the set of all equivalent U_1 groups, at each stage of the decomposition. Viewed in this way, it is then proper to denote the Weyl decomposition as "canonical." This is important, for we regard the Weyl result as the very model of a successful labeling, and the results to follow are in fact generalizations of this idea.

The essential labeling problem to which the extension of the WRC leads is this: In considering a quantum mechanics on the manifold of states belonging to U_n, the relevant set of operators is that which carries any state of any irrep to any other state of a (possibly different) irrep. This set of operators belongs to the diagonal subgroup of the group $U_n \times U_n$ (where we have mapped the adjoint states of U_n onto U_n). The "diagonal subgroup"—denoted by $U_n \otimes U_n$ (both the term and notation are due to Mackey[9])—is that subgroup of $U_n \times U_n$ whose elements have the form (g, g) where $g \in U_n$.

One set of labels for this set of operators is induced from the group U_n; this is the *tensor operator labeling*, determined by the generators of U_n. This set of labels is sufficient whenever in the restriction of an irrep of $U_n \times U_n$ to U_n, every irrep of U_n occurs at most once. If this is true for every irrep of $G \times G$ then the group G is said to be *multiplicity free* (mf).

The problem of labeling the operators of a not mf group is the multiplicity problem to which this paper is addressed. Among the unitary groups only U_1 and U_2 are mf.

In order to resolve the multiplicity problem, one possible group-theoretic procedure is that of *imbedding*. The group (taken to be not mf) is embedded in a larger group G' such that $G' \supset G$ and G' is itself mf. It appears offhand very unlikely that such a procedure could ever be successful but two examples, are known: (a) A_4 (the alternating group on 4 elements) imbedded in S_4 (the symmetric group on 4 elements), and (b) A_5 imbedded in R_3 (the three-dimensional rotation group). Since these examples are models for the work to follow let us consider the first in more detail.

There are four irreps of A_4, denoted by Γ_0, Γ_1, $\Gamma_{1'}$, and Γ_3 of dimension 1, 1, 1, and 3, respectively. Only Γ_3 is not mf and one finds that: $\Gamma_3 \times \Gamma_3 = \Gamma_0 + \Gamma_1 + \Gamma_{1'} + 2\Gamma_3$.

For S_4 there are 5 irreps denoted by $\Gamma_{(4)}$, $\Gamma_{(31)}$, $\Gamma_{(2^2)}$, $\Gamma_{(21^2)}$, and $\Gamma_{(1^4)}$ of dimension 1, 3, 2, 3, and 1, respectively. The group S_4 is mf.

To split the multiplicity we note that both $\Gamma_{(31)}$ and $\Gamma_{(21^2)}$ yield Γ_3 upon restriction to A_4. If we carry out

[6] The precise meaning intended for this term is discussed below in this section.

[7] E. Artin, *Geometric Algebra* (Interscience Publishers, Inc., New York, 1964).

[8] "The word 'canonical,' or also 'natural,' is applied in a rather loose sense to any mathematical construction which is unique in as much as no free choices of objects are used in it." (Ref. 7, p. 3.)

[9] G. W. Mackey, "The Theory of Group Representations," lecture notes, The University of Chicago (1955) (unpublished).

the multiplication in S_4 and then restrict to A_4, we find for example

$$\begin{array}{ccccccccc} \Gamma_{(31)} & \times & \Gamma_{(31)} & = & \Gamma_{(4)} & + & \Gamma_{(2^2)} & + & \Gamma_{(31)} & + & \Gamma_{(21^2)} \\ \downarrow & & \downarrow & & \downarrow & & \swarrow\searrow & & \swarrow & & \downarrow \\ \Gamma_3 & \times & \Gamma_3 & = & \Gamma_0 & + & \Gamma_1 & + & \Gamma_{1'} & + & 2\Gamma_3. \end{array}$$

It is clear that the group S_4 allows one to define, *group-theoretically*, labels for two distinguishable (in S_4) irreps of A_4: Γ_3^+ and Γ_3^- whose "multiplication" is mf. In this way the imbedding allows one to split the degeneracy and hence affords a group theoretical resolution of the multiplicity.

Now, it should be noted that this model for the resolution really demands only that the group G' have a sufficiently large set of mf representations to label all the representations of G; it is *not* necessary that G' itself be mf. It is this fact which makes the U_n problem solvable, for generally speaking the larger the group the larger the multiplicity.

We are now in a position to categorize our proposed solution to the multiplicity problem. We demonstrate that every irreducible representation of U_n may be given a suitable set of additional labels by imbedding U_n in $U_n \times U_n$ which is itself imbedded in a mf way in a totally-symmetric representation of U_{n^2} (Proposition 1). For such representations of U_{n^2}, the multiplication is itself mf (Lemma 5). In consequence, this shows that there exists a labeling scheme sufficient to split all multiplicity in U_n.[10]

Of itself this is an interesting and suggestive result, but it does not suffice for demonstrating the uniqueness (to within phases and normalization) of the Wigner coefficients. For this, we first categorize the Wigner coefficients in U_n as projections from U_{n^2}, next we show the Wigner coefficients in U_n to be eigenvectors of a U_{n^2} Wigner coefficient, and then establish Lemma 7 which explicitly determines the eigenvalues to which these eigenvectors belong, for all U_n. To resolve the multiplicity we demonstrate that, for U_3, all operators split completely upon projection onto suitable operators in U_2. The results for U_3 are summarized as Proposition 2.

In order to make the presentation self-contained and precise, we have chosen to depart from the discursive style of physics to the more succinct "theorem-proof" style. We hope by this to gain in clarity—not to imply that the results are necessarily mathematically novel.

[10] It has been called to the attention of the authors by Professor M. Moshinsky that a similar type of idea was proposed by T. A. Brody, M. Moshinsky, and I. Renero, J. Math. Phys. **6**, 1540 (1965).

II. THE EXISTENCE OF AN ADMISSIBLE IMBEDDING

In a group which is not mf, the first ancillary task is to determine and categorize the extent of the multiplicity. For U_n this was a basic result contained in Ref. 4(d): The tensor operators in U_n may be characterized by two Gel'fand patterns each carrying the same representation labels in U_n, i.e., each Gel'fand pattern refers to the same Young frame.[11] The two Gel'fand labels (one inverted) can be placed one above the other in a diamond shaped pattern $\diamondsuit$; this extended Gel'fand pattern (operator-pattern) then characterized all tensor operators in U_n in a one-to-one fashion.

The two Gel'fand patterns, however, play quite different roles; the lower pattern (ℓ-space) is completely defined *operationally* as the tensor operator labeling under commutation with the generators of U_n. The upper Gel'fand pattern by contrast is *not* (at this stage) a true labeling scheme but rather a shorthand to enumerate the multiplicity.

Nevertheless the notation strongly suggests that one attempt to define upper pattern space as the carrier space of a U_n group. In fact, we had two realizations in mind when this notation was introduced, and if we elaborate on this it may help to motivate the work to follow.

(a) The first realization lies in the representation matrices $D^{[m_{in}]}_{(m);(m')}(\{\theta_i\})$, themselves. These matrices carry three sets of labels: the irrep labels in U_n, $[m_{in}]$, and two sets of Weyl decomposition labels [Gel'fand labels (m), (m') of U_{n-1}]. It is not difficult to see that these are precisely equivalent to an operator-pattern. But now the two spaces (u- and ℓ-space) are completely definite: they are each isomorphic to U_n and their generators commute. Hence the representation matrices may be looked upon as the carrier space of irreps of $U_n \star U_n$, where the $\star$ denotes a direct product group modulo the relations that specify that the irreps of u- and ℓ-space share the same U_n labels (Young pattern).

(This may appear complicated but it is really very familiar; recall that the D matrices of SU_2 are themselves *orbital* irreps of $R_4 \cong SU_2 \times SU_2$. The u-space here is just that of the angular momentum *in the body-fixed frame*. All these ideas generalize; see Refs. 12 and 13.)

(b) The second realization lies in the boson calculus.

[11] An independent proof of this result is contained as a corollary to Lemma 7 below, and hence the present work is self-contained.

[12] L. C. Biedenharn and P. J. Brussaard, *Coulomb Excitation* (Clarendon Press, Oxford, England, 1965), see Chap. 7, especially.

[13] G. E. Baird and L. C. Biedenharn, paper presented at the Eastern Theoretical Physics Conference (1963).

Consider the (creation) operators $\{a_i\}$ which transform according to the irrep $[1\ \dot{0}]$ in U_n (the dotted zero denotes "repeat as often as required"). If we consider n distinct copies of these operators we may denote them by the set $\{a_i^j\}$. Clearly, the space of the index j is isomorphic to that of the index i; moreover, $\{a_i^j\}$ has the same Young frame $[1\ \dot{0}]$ in both u- and ℓ-space. In other words, $\{a_i^j\}$ is a realization of the operator-pattern $\langle 1\ \dot{0}\rangle$; it follows that multinomials on the set $\{a_i^j\}$ may be completely classified by operator-patterns as states in $U_n \star U_n$.

Both realizations are essential in the work to follow. We must emphasize, however, once again that the operator-patterns introduced in Ref. 4(d) to enumerate the Wigner operators do not designate states of $U_n \star U_n$ (since u-space had no group theoretical significance there). We should also emphasize that it is the operator-patterns in U_n which are of basic interest, and not so much the special realizations by state vectors in $U_n \star U_n$. (In fact, a notation which distinguished these ideas might be better.)

Let us now consider in more detail $U_n \star U_n$. The first result we wish to assert is almost trivial, though necessary.

Lemma 1: The irreps of $U_n \star U_n$ are of dimension $(\dim [m_{in}])^2$, where m_{in} are the Young pattern labels of the two U_n groups, and $\dim [m_{in}]$ denotes the Weyl dimension formula.

We next establish formulas equivalent to the Weyl dimension formula which represent two distinct ways to regard Weyl's result, one from the point of view of *hook patterns*, the other in terms of the symmetric group and the Schur functions.

Weyl's result for the dimensionality of the representation $[m_{in}]$ of U_n (where $m_{1n}, m_{2n}, \cdots, m_{nn}$ denote the lengths of the rows of the Young frame) is given by

$$\dim [m_{in}] = \prod_{i<k} (l_i - l_k)/[1!\,2! \cdots (n-1)!] \quad (1)$$

with

$$l_i \equiv n - i + m_{in}. \quad (2)$$

Let us now introduce the idea of the *hook length* h_{ik} of the (ij)th node (box) of the Young frame $[m_{in}]$:

$$h_{ik} \equiv (m_{in} - i) + (m'_{kn} - k) + 1, \quad (3)$$

where m'_{kn} is defined to be the number of nodes (boxes) in the kth column of the Young frame $[m_{in}]$.

The product of all the hook lengths in the Young frame is denoted by

$$H^{[m_{in}]} \equiv \prod_{i,k} h_{ik}. \quad (4)$$

Similarly we introduce a function $G^{[m_{in}]}(x)$ defined on the Young frame by:

$$G^{[m_{in}]}(x) \equiv \prod_{i,k} (x + k - i), \quad (5)$$

where (i, k) runs over each node of the Young frame $[m_{in}]$.

With these definitions the Weyl dimension formula, as shown by Robinson,[14] takes the following useful and very intriguing form.

Lemma 2: $\dim [m_{in}] = G^{[m_{in}]}(n)/H^{[m_{in}]}$.

Next let us discuss the Weyl dimension formula from the point of view of Schur functions.[15] We denote first by S_r the symmetric functions of the n variables α_i $(i = 1, \cdots, n)$ which are homogeneous product sums, i.e.,

$$S_r \equiv \sum_{i=1}^{n} (\alpha_i)^r. \quad (6)$$

We define, as usual, the matrix Z_r by the form:

$$(Z_r) \equiv \begin{pmatrix} S_1 & 1 & 0 & & \cdots & 0 \\ S_2 & S_1 & 2 & 0 & \cdots & 0 \\ S_3 & S_2 & S_1 & 3 & 0 & \cdots \\ & & & \cdots & & \\ S_{r-1} & S_{r-2} & & \cdots & S_1 & r-1 \\ S_r & S_{r-1} & & \cdots & S_2 & S_1 \end{pmatrix}, \quad (7)$$

and define an *immanant* of the $n \times n$ matrix (a_{st}) corresponding to the Young frame $[\lambda]$ (having n nodes) to be:

$$|a_{st}|^{[\lambda]} \equiv \sum \chi^{[\lambda]}(S) P_S, \quad (8)$$

where

(a) the sum is over all $n!$ permutations S;
(b) $P_S \equiv a_{1i_1} a_{2i_2} \cdots a_{ni_n}$ and S is the permutation $(i_1 \cdots i_n)$ of the integers $1 \cdots n$;
(c) $\chi^{[\lambda]}(S)$ is the character of the representation $[\lambda]$ of S_n, the symmetric group on n symbols.

With these definitions *the Schur function* $\{\lambda\}$ is given by

$$r!\,\{\lambda\} \equiv |Z_r|^{[\lambda]}. \quad (9)$$

Since, as is well known, the characters of the unitary group correspond to Schur functions of the characteristic roots of the (matrix) representation, we see that the dimension formula of Weyl must be the special Schur function corresponding to the identity element. That is,

Lemma 3:

$$\dim [m_{in}] = (k!)^{-1} \sum_{\mathcal{C}} (n)^{l(\mathcal{C})} \cdot \#(\mathcal{C}) \cdot \chi^{[m_{in}]}(\mathcal{C}),$$

[14] See Ref. 5, p. 60.
[15] D. E. Littlewood, *The Theory of Group Characters* (Clarendon Press, Oxford, England, 1950), see Chap. VI.

where

(a) k denotes the number of nodes in the Young frame $[m_{in}]$;
(b) $\mathcal{C}$ denotes the *class* of the permutations in S_n written in cycle structure form, i.e., as a Young frame;
(c) $l(\mathcal{C})$ denotes the number of rows in the Young pattern of the class $\mathcal{C}$;
(d) $\#(\mathcal{C})$ denotes the number of elements in the class $\mathcal{C}$; and
(e) $\chi^{[m_{in}]}(\mathcal{C})$ is the character of the irrep $[m_{in}]$ of S_n for the class $\mathcal{C}$.

Proof: This is simply a matter of direct substitution of $\alpha_i = 1$ (corresponding to the identity element of U_n) into the definition of the Schur function given above.

Lemma 3 is relatively complicated, but it is our desired expression for the Weyl dimension formula of U_n in terms of the symmetric group S_n. It should be noted that the hook form gives the Weyl formula in *factored* form, whereas the result above gives the same formula as an explicit polynomial on n. Both are useful, but our specific application is the following.

Lemma 4: $\sum (\dim[m_{in}])^2 = \dim [k\dot{0}]_{n^2}$,

where the sum on the left extends over *all* Young frames with

$$k = \sum_{i=1}^{n} m_{in}$$

nodes each taken *once*.

Proof: Taking the square of dim $[m_{in}]$, using Lemma 3, and summing over all Young frames, one may invoke the orthogonality theorem for the characters of S_n to eliminate one of the class sums (noting the necessity to normalize correctly).

This yields

$$\sum (\dim [m_{in}])^2 = (k!)^{-1} \sum_{\mathcal{C}} (n^2)^{l(\mathcal{C})} \cdot \#(\mathcal{C}). \quad (10)$$

But by Lemma 3, the right-hand side is just the dimension of the irrep $[k\ \dot{0}]$ in U_{n^2}, and the result follows.[16]

As it stands, this result, Lemma 4, is an interesting—possibly even rather surprising—relation between certain representations of U_n and U_{n^2}. The real significance of this result, however, which we wish to emphasize is much more important and we assert it as Proposition 1.

Proposition 1: There exists a multiplicity-free imbedding of any representation $\langle m_{in} \rangle$ of $U_n \star U_n$ in a totally symmetric representation $[k\ \dot{0}]$ of U_{n^2}, where $k = \sum_{i=1}^{n} m_{in}$. (Multiplicity-free means here that $[k\ \dot{0}]_{n^2}$ contains each irrep $\langle m_{in} \rangle$, whose Young frame contains exactly k boxes, once and only once.)

We may readily appreciate the importance of this imbedding if we remark now that,

Lemma 5: The multiplication of the total symmetric representation in U_{n^2} is multiplicity-free.

To prove this assertion we use the results of Ref. 4(d), or simply observe that the fundamental Wigner operator [1 0] has this property and that an induction on k establishes the desired answer directly.

III. THE FACTORIZATION LEMMA AND APPLICATION TO THE CANONICAL DEFINITION OF THE WIGNER COEFFICIENTS IN U_3

Proposition 1 and Lemma 5 of the preceding section demonstrate that the imbedding of U_n in U_{n^2} allows one to associate with each representation of $U_n \star U_n$ a canonical set of labels given by two Weyl decompositions. (It is clear that canonical here means "to within an equivalence class" in accord with earlier comments on the Weyl decomposition.) The set of labels in upper pattern space allows one to distinguish in U_{n^2} and in $U_n \star U_n$ representations of U_n (more definitely, *states* in U_n, i.e., lower pattern labels) which are otherwise indistinguishable. Since this extra labelling as well as the multiplication is mf in U_{n^2}, we have clearly made progress toward our goal in precisely the manner indicated by the example of imbedding A_4 in S_4.

However, this is *not* yet enough to assert our desired final result; the desired final result concerns itself with defining a unique Wigner coefficient, i.e., the matrix element of a fully labeled (possibly, though not necessarily, labeled in $U_n \star U_n$) tensor operator acting between *states of* U_n. We have proved, however, only that matrix elements of states, all in $U_n \star U_n$, are uniquely fixed—and it is conceivable that in suppressing the extra labels (in some unspecified way) a multiplicity could once again be re-established. It is the purpose of this section to demonstrate that this does indeed occur, and show how the problem is resolved for U_3.

Let us begin once again with an easy result.

Lemma 6: The matrix element

$$\left\langle \begin{matrix} [k+k'\ \dot{0}]_{(n^2)} \\ [\lambda'']\ [\lambda''] \\ (\mu'')\ (m'') \end{matrix} \middle| \begin{matrix} [k'\ \dot{0}]_{(n^2)} \\ [\lambda']\ [\lambda'] \\ (\mu')\ (m') \end{matrix} \middle| \begin{matrix} [k\ \dot{0}]_{(n^2)} \\ [\lambda]\ [\lambda] \\ (\mu)\ (m) \end{matrix} \right\rangle$$

$$\equiv \left\langle \left\langle \begin{matrix} \mu'' \\ \lambda'' \\ m'' \end{matrix} \right\rangle \middle| \mathcal{O} \begin{pmatrix} \mu' \\ \lambda' \\ m' \end{pmatrix} \middle| \left\langle \begin{matrix} \mu \\ \lambda \\ m \end{matrix} \right\rangle \right\rangle \equiv f \begin{pmatrix} \mu & \mu' & \mu'' \\ \lambda & \lambda' & \lambda'' \\ m & m' & m'' \end{pmatrix}$$

[16] Louck, Ref. 3, has independently obtained this result.

is a uniquely determined real number fully specified by the $U_n \star U_n$ labels.

Proof: The proof is immediate: one need only note that this matrix element is no more and no less than a Wigner coefficient in U_{n^2} corresponding to unique states of totally symmetric representations. [The algebraic result is in fact deduced from the fundamental Wigner coefficient in U_{n^2} as given in Ref. 4(b).] It should be noted that $f(\cdots)$ is symmetric under the interchange $\mu\ \mu'\ \mu'' \leftrightarrow m\ m'\ m''$.

A Wigner coefficient must now be defined as a *projection* on the matrix element $f(\cdots)$, and this projection must suppress *all* upper pattern labels. (Otherwise one is implicitly assuming, without justification, that upper pattern labels have a significance not demonstrated so far.) Lower pattern labels, on the other hand, have precisely the desired significance by virtue of the original group U_n. This shows, moreover, that the projection $P(\cdots)$ must be *linear*, so that we must have the general form:

$$\sum_{\mu\mu'\mu''} f\begin{pmatrix}\mu & \mu' & \mu''\\ \lambda & \lambda' & \lambda''\\ m & m' & m''\end{pmatrix} P^{(g)}\begin{pmatrix}\lambda & \lambda' & \lambda''\\ \mu & \mu' & \mu''\end{pmatrix} \equiv G\begin{pmatrix} & g & \\ \lambda & \lambda' & \lambda''\\ m & m' & m''\end{pmatrix}, \quad (11)$$

where $G(\cdots)$ is a Wigner coefficient, possibly complex, with (g) labeling the multiplicity. It is easily established that this coefficient has the transformation properties under U_n specified by the (λ, m) indices.

We must now prove an essential lemma concerning the matrix element $f(\cdots)$, viewed as a matrix in (μ), (m) space.

Lemma 7: The matrix

$$F(\lambda\lambda'\lambda'') \equiv (F_{(\mu),(m)}) \equiv \left[f\begin{pmatrix}\mu & \mu' & \mu''\\ \lambda & \lambda' & \lambda''\\ m & m' & m''\end{pmatrix}\right]$$

is real and symmetric and may be brought to diagonal form. The orthonormal eigenvectors defined by this diagonalization are Wigner coefficients in U_n. This diagonalization takes the explicit form:

$$f\begin{pmatrix}\mu & \mu' & \mu''\\ \lambda & \lambda' & \lambda''\\ m & m' & m''\end{pmatrix} = \left[\left(\sum_{i=1}^{n}\lambda_{i,n}\right)!\Big/\left(\sum_{i=1}^{n}\lambda''_{i,n}\right)!\right]^{\frac{1}{2}} \times \sum_{\Gamma}\Delta(\lambda\lambda'\lambda'')\left\langle\begin{matrix}\lambda''\\ m''\end{matrix}\middle|\left\langle\begin{matrix}\Gamma\\ \lambda'\\ m'\end{matrix}\right\rangle\middle|\begin{matrix}\lambda\\ m\end{matrix}\right\rangle \times \left\langle\begin{matrix}\lambda''\\ \mu''\end{matrix}\middle|\left\langle\begin{matrix}\Gamma\\ \lambda'\\ \mu'\end{matrix}\right\rangle\middle|\begin{matrix}\lambda\\ \mu\end{matrix}\right\rangle,$$

where

(a) $\Delta(\lambda\lambda'\lambda'') \equiv [\chi^{[\lambda']}(E)]^{\frac{1}{2}} \cdot \left[\prod_{i=1}^{n}\left(\frac{(\lambda''_{i,n}+n-i)!}{(\lambda_{i,n}+n-i)!}\right)\right]^{\frac{1}{2}} \times (\dim[\lambda]/\dim[\lambda''])^{\frac{1}{2}}$,

(b) $\chi^{[\lambda']}(E)$ is the identity character of $[\lambda']$ in the symmetric group, $S_{k'}$,

(c) U_n Wigner operators are denoted by

$$\left\langle\begin{matrix}\Gamma\\ \lambda'\\ m'\end{matrix}\right\rangle \text{ and } \left\langle\begin{matrix}\Gamma\\ \lambda'\\ \mu'\end{matrix}\right\rangle$$

in ℓ- and u-space respectively,

(d) Γ denotes the set of operator (upper pattern) labels in U_n or, equivalently, any set in 1–1 correspondence with this.

Proof: Let us first state the specific normalization of the matrix F; as a matrix element of a (unique) Wigner operator in U_{n^2}, this obeys the usual normalization conventions, i.e., the operator is orthonormal and the matrix element between maximal states with maximal change in irrep labels is defined as $+1$.

The fact that F is real and symmetric has already been established. Any such matrix may be brought to diagonal form by a real orthogonal transformation.[17] Wherever the eigenvalues (necessarily real) are distinct this defines unique eigenvectors (to within order and phases).

Next let us note that

$$f\begin{pmatrix}\mu & \mu' & \mu''\\ \lambda & \lambda' & \lambda''\\ m & m' & m''\end{pmatrix}$$

transforms under the group $U_n \times U_n$ as the identity component of the direct product $\langle\overline{\lambda''}\rangle \times \langle\lambda\rangle \times \langle\lambda'\rangle$. It follows that $f(\cdots)$ has the general form

$$f\begin{pmatrix}\mu & \mu' & \mu''\\ \lambda & \lambda' & \lambda''\\ m & m' & m''\end{pmatrix} = \sum_{g,g'} A^{\lambda\lambda'\lambda''}_{g,g'} \times G\begin{pmatrix} & g & \\ \lambda & \lambda' & \lambda''\\ m & m' & m''\end{pmatrix} G\begin{pmatrix} & g' & \\ \lambda & \lambda' & \lambda''\\ \mu & \mu' & \mu''\end{pmatrix}, \quad (12)$$

where the $G(\cdots)$ are (orthonormal) Wigner coefficients in any (canonical or noncanonical, i.e., arbitrary) breaking of the multiplicity. It follows that the matrix $A_{gg'}$ is square and may be brought to diagonal form, thereby defining Wigner coefficients as eigenvectors. The main task here concerns the explicit evaluation of the (real) eigenvalues, $\Delta(\lambda\lambda'\lambda'')$.

[17] F. Tricomi, *Lezioni di Analisi Matematica* (CEDAM, Padova, 1948), Pt. 1, p. 317.

We establish the validity of the stated result by recursion; it is most convenient for this purpose to use the language of operators. The matrix element $F(\lambda\lambda'\lambda'')$ is a specific matrix element of the unique Wigner operator in U_{n^2} which (1) transforms as the state (with $k' = \sum \lambda'_{\text{in}}$ quanta)

$$\left| \begin{matrix} [k'0]_{(n^2)} \\ [\lambda'][\lambda'] \\ (\mu)(m) \end{matrix} \right\rangle$$

and (2) induces maximal change in the irrep labels of states of U_{n^2}. We denote this operator as $\mathcal{O}\begin{pmatrix} \mu' \\ \lambda' \\ m' \end{pmatrix}$. The factorization stated by the lemma applies only to totally symmetric initial states in U_{n^2}, and we suppose throughout the following that the initial states are restricted to this class only. The plan of the proof is then the following. The result for $k' = 1$ (i.e., fundamental Wigner operators) is not trivial but can be shown directly [or equivalently from Ref. 4(b)]. The recursion proceeds by assuming the lemma valid for k' quanta and demonstrating that this implies the validity for $k' + 1$ quanta.

To relate the operators for $k' + 1$ quanta to the operators for k' quanta, we use the coupling aspect of the Wigner operator [1 0̇] in U_{n^2} to couple the operators for k' quanta and for 1 quanta to form the operator for $k' + 1$ quanta. This step shows that the normalized operator for $k' + 1$ quanta is of the form

$$\begin{aligned} \mathcal{O}\begin{pmatrix} \mu' \\ M' \\ m' \end{pmatrix} = \sum_{\substack{[M](\gamma) \\ (m)(\mu) \\ (r)(\rho)}} (k'+1)^{-\frac{1}{2}} \Delta([M], [1\,\dot{0}], [M']) \\ \times \left\langle \begin{matrix} [M'] \\ (m') \end{matrix} \middle| \left\langle \begin{matrix} (\gamma) \\ [1\,\dot{0}] \\ (r) \end{matrix} \right\rangle \middle| \begin{matrix} [M] \\ (m) \end{matrix} \right\rangle \\ \times \left\langle \begin{matrix} [M'] \\ (\mu') \end{matrix} \middle| \left\langle \begin{matrix} (\gamma) \\ [1\,\dot{0}] \\ (\rho) \end{matrix} \right\rangle \middle| \begin{matrix} [M] \\ (\mu) \end{matrix} \right\rangle \\ \times \mathcal{O}\begin{pmatrix} \rho \\ 1\,\dot{0} \\ r \end{pmatrix} \mathcal{O}\begin{pmatrix} \mu \\ M \\ m \end{pmatrix}. \end{aligned} \tag{13}$$

[The order is essential. The operator $\mathcal{O}(M)$ acts first, then $\mathcal{O}(1\,\dot{0})$; hence the coupling reads in the same (right to left) order.]

By hypothesis, we may introduce the eigenvalue expansion for $\mathcal{O}(1\,\dot{0})$ and $\mathcal{O}(M)$ as an operator relation. Note that (a) when the $\langle \cdots \rangle$ act in different spaces (ℓ- vs. u-space) the Wigner operators commute, otherwise not; and that (b) the particular structure of $\Delta(\cdots)$ shows that these operators always commute. With these simplifications in hand, the operator $\mathcal{O}(M')$ takes the form:

$$\begin{aligned} \mathcal{O}\begin{pmatrix} \mu' \\ M' \\ m' \end{pmatrix} = (k'+1)^{-\frac{1}{2}} \left[\left(\sum_{i=1}^{n} \lambda_{\text{in}}^{\text{initial}} \right)! \Big/ \left(\sum_{i=1}^{n} \lambda_{\text{in}}^{\text{final}} \right)! \right]^{\frac{1}{2}} \\ \times \sum_{\substack{[M] \\ (\gamma),(\Gamma_1),(\Gamma_k)}} \Delta(M, M') \Delta([M], [1\,\dot{0}], [M']) \\ \times \sum_{(m),(r)} \left\{ \left\langle \begin{matrix} [M'] \\ (m') \end{matrix} \middle| \left\langle \begin{matrix} (\gamma) \\ [1\,\dot{0}] \\ (r) \end{matrix} \right\rangle \middle| \begin{matrix} [M] \\ (m) \end{matrix} \right\rangle \right. \\ \left. \times \left\langle \begin{matrix} (\Gamma_1) \\ [1\,\dot{0}] \\ (r) \end{matrix} \right\rangle \left\langle \begin{matrix} (\Gamma_k) \\ [M] \\ (m) \end{matrix} \right\rangle \right\} \\ \times \sum_{(\mu),(\rho)} \left\{ \left\langle \begin{matrix} [M'] \\ (\mu') \end{matrix} \middle| \left\langle \begin{matrix} (\gamma) \\ [1\,\dot{0}] \\ (\rho) \end{matrix} \right\rangle \middle| \begin{matrix} [M] \\ (\mu) \end{matrix} \right\rangle \right. \\ \left. \times \left\langle \begin{matrix} (\Gamma_1) \\ [1\,\dot{0}] \\ (\rho) \end{matrix} \right\rangle \left\langle \begin{matrix} (\Gamma_k) \\ [M] \\ (\mu) \end{matrix} \right\rangle \right\}. \end{aligned} \tag{14}$$

[In this equation, the operator $\Delta(M, M')$ denotes the product of the Δ operators for $\mathcal{O}(1\,\dot{0})$ and $\mathcal{O}(M)$ and is specified fully by the arguments $[M]$ and $[M']$ once the initial irrep (λ^{initial}) is specified.]

Note that in the coupling $[M] \otimes [1\,\dot{0}] = [M']$, the irrep labels $[M]$ are specified by $[M']$ and (γ). Note also the structure of the terms in this result for $\mathcal{O}(M')$: each sum in brackets $\sum\{\cdots\}$ specifies, for fixed (γ), a normalized operator in U_n which transforms as $[M']$, (m'), or $[M']$ (μ') in ℓ- or u-space, respectively. Since such operators are not necessarily unique, we may rewrite these sums in terms of an *arbitrary* breaking of the multiplicity

$$\begin{aligned} \sum_{(m),(r)} \left\langle \begin{matrix} [M'] \\ (m') \end{matrix} \middle| \left\langle \begin{matrix} (\gamma) \\ 1\,\dot{0} \\ (r) \end{matrix} \right\rangle \middle| \begin{matrix} [M] \\ (m) \end{matrix} \right\rangle \left\langle \begin{matrix} (\Gamma_1) \\ [1\,\dot{0}] \\ (r) \end{matrix} \right\rangle \left\langle \begin{matrix} (\Gamma_k) \\ [M] \\ (m) \end{matrix} \right\rangle \\ \equiv \sum_{(\Gamma_t)} R(\Gamma_1 \Gamma_k \Gamma_t) \left\langle \begin{matrix} (\Gamma_t) \\ [M'] \\ (m') \end{matrix} \right\rangle. \end{aligned} \tag{15}$$

Here $R(\Gamma_1\Gamma_k\Gamma_t)$ denotes a Racah coefficient in U_n, defined, by this relation, for an arbitrary breaking of the multiplicity. (We have suppressed various indices in writing this coefficient in order to simplify the notation; this should cause no ambiguity.) A relation similar to Eq. (15) holds in u-space.

Introducing these relations into Eq. (14) and summing over all (Γ_1) and (Γ_k) [for fixed (Γ_t) and $(\Gamma_{t'})$] we find from the general orthonormality properties of the Racah coefficient[18] (*independent* of the multiplicity

[18] J. R. Derome and W. T. Sharp, J. Math. Phys. 6, 1584 (1965).

problem) that the following sum factors out and yields

$$\sum_{(\Gamma_1),(\Gamma_k)} R(\Gamma_1\Gamma_k\Gamma_t)R(\Gamma_1\Gamma_k\Gamma_{t'}) = \delta(\Gamma_t, \Gamma_{t'}). \quad (16)$$

The final step requires the use of the identity

$$\sum_{\substack{[M] \\ (\gamma)}} [\chi^{[M]}(E)]^{\frac{1}{2}}(k'+1)^{-\frac{1}{2}} \cdot \left[\prod_{i=1}^{n} \left(\frac{(M'_{in}+n-i)!}{(M_{in}+n-i)!}\right) \frac{\dim [M]}{\dim [M']}\right]^{\frac{1}{2}} = [\chi^{[M']}(E)]^{\frac{1}{2}}. \quad (17)$$

This identity can be shown directly by recursion, and although hardly an obvious result, need not be proved here.

These results suffice to show that the operator $\mathcal{O}(M')$ for $k'+1$ quanta has precisely the form specified by the lemma, if $\mathcal{O}(M)$ for k' quanta has this form. Hence the lemma is proved.

Remarks:

A. The essential point about Lemma 7 is that it asserts *a general factorization of the boson operators of* $U_n \star U_n$ *into characteristic operators of the* U_n *group itself;* we refer to Lemma 7 as the "factorization lemma for the boson calculus". This result is not only interesting from the point of view of structure, but is a very powerful tool for actual computations. Lacking such a general tool, it was necessary [in Ref. 4(b)] to define the fundamental Wigner operators, $\langle 1\ \dot{0}\rangle$, by means of the 'trick' of projecting U_{n+1} generators into the fundamental operators of U_n. It is clear (in retrospect) that an uncritical use of the elementary boson operators a_i^j is not fully satisfactory for defining $\langle 1\dot{0}\rangle$, precisely because the idea of factorization needs to be clarified first.

B. The occurrence of the character $\chi(E)$ of symmetric group in the eigenvalue expression given in the lemma is a direct consequence of requiring the operators $\left\{\mathcal{O}\begin{pmatrix}\mu\\ \lambda\\ m\end{pmatrix}\right\}$ to comprise a single Wigner operator in U_{n^2}. The character can be eliminated if we normalized each $\langle\lambda\rangle$ component of the operator individually. Note, too, that the eigenvalue [aside from $\chi(E)$] is just the ratio of the pattern measures [cf. Ref. 4(b)] of the maximal initial and final states.

The factorization lemma turns out to be basic to the systematic development of the boson calculus for U_n. Let us indicate this by stating a number of assertions that result from Lemma 7.

1. (a) The number of distinct eigenvectors of the operator $[k\ \dot{0}]$ in U_{n^2} is given by precisely

$$\sum_{[M]} \dim [M],$$

where the sum is over all lexical patterns in k quanta.

(b) There are as many eigenvectors (= Wigner operators) that transform in U_n as the irrep $[M]$ as there are states, namely dim $[M]$.

2. The dimensionality of the matrix $A_{gg'}$ [cf. Eq. (11)] is precisely $I([\overline{\lambda''}], [\lambda] \otimes [\lambda'])$, where $I(\cdots)$ is the intertwining number.[9]

3. The degeneracy structure of the eigenvalues of $\mathcal{O}([M])$ exactly parallels the degeneracy structure of the state vector of the irrep $[M]$. Put somewhat differently we may assert that

(state vectors of $[M]\leftrightarrow$ weight vectors) $\leftrightarrow$ (operator $\langle M\rangle\leftrightarrow$ changes in irrep labels).

[This equivalence was proved earlier in Ref. 4(d); it underlies the enumeration of the Wigner operators $\langle M\rangle$ by operator (upper Gel'fand) patterns which specify the changes in irrep labels.]

4. Every Wigner operator in U_n may be obtained from the reduction of direct products of fundamental Wigner coefficients. An alternative form for this result is the assertion that the set of Wigner operators $\{[k\ \dot{0}]\}$ in U_{n^2} [equivalently the set of boson (creation) operators in $U_n \star U_n$] yield Wigner operators in U_n, each operator occurring as an eigenvector once and only once.

The factorization lemma shows that the embedding of U_n into U_{n^2} does not, of itself, resolve the multiplicity problem (since the eigenvalues of Lemma 7 have precisely the degeneracy of the intertwining number). Nevertheless the factorization lemma *does* supply the key to the resolution; we demonstrate this for U_3 in detail and obtain Proposition 2 below. For the generalization to U_n we refer to a subsequent paper.[19]

The key to the resolution lies in the splitting of the multiplicity under restriction to operators belonging to a subgroup. *In other words, the multiplicity for operators is resolved in a manner analogous to the resolution of the state labeling problem—in both, the key idea is a branching law.*

It is helpful to introduce a notation. The U_3 operator $\langle M\rangle$ is designated by the upper pattern labels

$$\begin{pmatrix} & a_{11} & \\ a_{12} & & a_{22}\end{pmatrix} \equiv (a)$$

[19] The assertion of Proposition 2 in our original paper contained a serious error, and our original statement of Lemma 7 was incorrect. We discovered this error while our paper was being set in type, and have revised the present version accordingly (30 December 1966). The generalization to U_n requires additional considerations which appear best treated in a separate paper.

and the subgroup labels

$$\begin{pmatrix} m_{12} & & m_{22} \\ & m_{11} & \end{pmatrix} \equiv (m).$$

The operator is fully designated by the lexical operator pattern

$$\left\langle \begin{matrix} & & a_{11} & & \\ & a_{12} & & a_{22} & \\ m_{13} & & m_{23} & & m_{33} \\ & m_{12} & & m_{22} & \\ & & m_{11} & & \end{matrix} \right\rangle .$$

The significance of the upper labels (a) is that they specify the *changes* induced by operating with $\langle M \rangle$ on a definite initial state $|(\lambda)\rangle$

$$\Delta_{i3} \equiv \lambda_{i3}^{\text{final}} - \lambda_{i3}^{\text{initial}}$$

with

$$\Delta_{13} = a_{11}, \quad \Delta_{23} = a_{12} + a_{22} - a_{11},$$
$$\Delta_{33} = m_{13} + m_{23} + m_{33} - a_{12} - a_{22}.$$

Operators with the same $a_{12} + a_{22}$ induce the same changes, that is, the same set $\{\Delta_{i3}\}$. The resolution of this multiplicity (in other words, the significance of the a_{12}, a_{22}) is supplied by the following lemma.

Lemma 8: The Wigner operators $\langle m_{13}m_{23}m_{33} \rangle$ in U_3 under restriction to the U_2 operators $\left\langle \begin{matrix} & m_{13} & \\ m_{13} & . & m_{33} \end{matrix} \right\rangle$ split completely. Operators whose upper patterns have maximal a_{12}–a_{22} have a nonzero restriction; all other operators have zero restriction.

Proof: We use the factorization lemma and consider the U_2 Wigner operator having the $U_n \star U_n$ irrep labels $[m_{13}m_{23}m_{33}]$,

$$(m') = \begin{pmatrix} m_{13} & & m_{33} \\ & m_{13} & \end{pmatrix} \text{ and } (\mu') \text{ arbitrary}$$

taken between the initial state (in $U_n \star U_n$)

$$\left| \left\langle \begin{matrix} & & \mu_{11} & & \\ & \mu_{12} & & \mu_{22} & \\ \lambda_{13} & & \lambda_{23} & & \lambda_{33} \\ & \lambda_{12} & & \lambda_{22} & \\ & & \lambda_{12} & & \end{matrix} \right\rangle \right\rangle$$

and the final state

$$\left| \left\langle \begin{matrix} & & \mu''_{11} & & \\ & \mu''_{12} & & \mu''_{22} & \\ \lambda_{13}+\Delta_{13} & & \lambda_{23}+\Delta_{23} & & \lambda_{33}+\Delta_{33} \\ & \lambda_{12}+m_{13} & & \lambda_{22}+m_{33} & \\ & & \lambda_{12}+m_{13} & & \end{matrix} \right\rangle \right\rangle .$$

Next, we write this same operator in the form

$$\mathcal{O}^{(\mu')}_{(m')}([m_{13}m_{23}m_{33}]) = \mathcal{O}([m_{33}m_{33}m_{33}]) \otimes \mathcal{O}\left(\begin{bmatrix} m_{23}-m_{33} & m_{23}-m_{33} & 0 \\ & m_{23}-m_{33} & 0 \\ & & m_{23}-m_{33} \end{bmatrix}\right) \otimes \mathcal{O}\left(\begin{bmatrix} m_{13}-m_{23} & 0 & 0 \\ & m_{13}-m_{23} & 0 \\ & & m_{13}-m_{23} \end{bmatrix}\right),$$

where we note that the lower patterns are so chosen that they multiply by addition of their patterns [appropriate (unique, therefore known) Wigner coefficients are implied for u-space].

Now, we apply the factorization lemma to each term in the product above, introducing a complete set of intermediate (U_{n^2}) states between each operator. The scalar operator $\mathcal{O}([m_{33}m_{33}m_{33}])$ is of course trivial. The U_n operators

$$\langle m_{23}-m_{33} \quad m_{23}-m_{33} \quad 0 \rangle$$

and $\langle m_{13} - m_{23} 0\ 0 \rangle$ have no multiplicity; explicit evaluation, for the special states chosen, shows that matrix elements in ℓ-space completely factor into parts dependent on U_3 and U_2 labels *separately*. It follows that the u- and ℓ-space matrix elements themselves factor into a *single* product.

In other words, for this specific choice of ℓ-space states, there exists, by the factorization lemma, but *one* operator. We choose to designate this operator by maximal a_{12}–a_{22} labels.

Lemma 9: The Wigner operators $\langle m_{13}m_{23}m_{33} \rangle$ in U_3 under restriction to the U_2 operators $\left\langle \begin{matrix} & b_{12} & \\ m_{13} & . & m_{33} \end{matrix} \right\rangle$ split completely.

The distinction between this and the previous lemma lies in the fact that a more general U_2 operator is used. The desired result is established by iteration: For $b_{12} = m_{13}$ we determine all maximal $(a_{12}$–$a_{22})$ operators. We next apply the factorization lemma, just as before, except this time $b_{12} = m_{13} - 1$, and all maximal operators are explicitly *removed* from the relation established by the factorization lemma. By explicit evaluation of the matrix elements in ℓ-space, there now exists again but a single factor. This yields all operators $[a_{12} - 1, a_{22} + 1]$. By iteration we obtain all operators having distinct upper patterns.

Remark: The existence of the splitting demonstrated above is familiar in special cases. Thus, for

example, it is well known that the generators in U_3, when restricted to [20], have zero components on the operators

$$\left\langle \begin{matrix} 2 \\ 20 \\ \cdot \end{matrix} \right\rangle \quad \text{and} \quad \left\langle \begin{matrix} 0 \\ 20 \\ \cdot \end{matrix} \right\rangle,$$

as the lemmas above require. The net effect of Lemmas 8 and 9 is to assert how this property extends to all U_3 operators (in particular, the precise subgroup operator on which the splitting occurs). Ultimately the validity of the lemma hinges on the unique multiplication properties of the operator decomposition used, as well as the special choice of subgroups. All of these features generalize to U_n, but to carry out the proof is highly unwieldy without the introduction of simplifying techniques.

It is gratifying that the operators obtained in this resolution of the multiplicity are *precisely* the orthonormal operators discussed in our earlier work.[20] In particular, the F and D operators used in high-energy physics are exactly the canonical operators

$$\left\langle \begin{matrix} 1 \\ 1\ 1 \\ 2\ 1\ 0 \end{matrix} \right\rangle \quad \text{and} \quad \left\langle \begin{matrix} 1 \\ 2\ 0 \\ 2\ 1\ 0 \end{matrix} \right\rangle,$$

respectively.

We have hence proved our desired result for U_3, which we assert as a proposition.

Proposition 2: The matrix

$$F \equiv (F_{(m),(\mu)}) \equiv f\begin{pmatrix} \mu & \mu' & \mu'' \\ \lambda & \lambda' & \lambda'' \\ m & m' & m'' \end{pmatrix}$$

is a unique Wigner coefficient defined in the group $U_3{}^2$ labeled by two Weyl subgroup decompositions of $U_3 \star U_3$. The Wigner coefficients of U_3 are precisely the eigenvectors of the transformation which brings the matrix F to diagonal form. This diagonal form exists since the matrix F has precisely N degenerate but nonzero eigenvalues, where N is the intertwining number $I([\overline{\lambda''}], [\lambda] \otimes [\lambda'])$. The operators corresponding to a degenerate eigenvalue split completely upon restriction to appropriate operators in U_2. The Wigner coefficients so defined are orthonormal and unique to within ordering, and canonical in the sense of equivalence classes of Weyl decompositions.

The generalization of this result to U_n is to be presented in a subsequent paper.

Let us note that it is now a straightforward matter to determine explicitly not only the Wigner coefficients but also the complete family of coefficients denoted by "$3n - j$ symbols" in SU_2. For this purpose the boson calculus of the $\{a_i^j\}$ is ideal (see for example, Louck, Ref. 3), and if one combines the boson calculus with the pattern calculus of hooks, great simplifications can be made. These techniques are to be discussed systematically elsewhere.

ACKNOWLEDGMENTS

We are indebted to many colleagues for discussions on the work presented above. In particular, we would like to thank Dr. M. Danos, Dr. U. Fano, Dr. D. Kleima, and Dr. J. Macek.

This paper was supported in part by the United States Army Research Office (Durham) and the National Science Foundation. Part of this work was performed under the auspices of the U.S. Atomic Energy Commission.

[20] The techniques of the present paper are a great deal more convenient than the earlier technique involving conjugation. In particular, the operators appear correctly orthonormalized—automatically.

Null Tetrad Approach to Motions in Empty Space-Time

C. D. COLLINSON AND D. C. FRENCH
Department of Applied Mathematics, University of Hull, Hull, England

(Received 26 July 1966)

The integrability conditions of conformal motions are written in the null tetrad formalism of Newman and Penrose. The maximum order of the group of conformal motions admitted by nonflat empty space-times of given Petrov type is shown to be at most one greater than the maximum order of the group of Killing motions. The symmetries of those empty space-times which possess hypersurface orthogonal geodesic rays with nonvanishing divergence are determined. Among these space-times is one of type III which admits a group of Killing motions of order three. This provides a counter example to a result of Petrov which states that the maximum order of the group of Killing motions for such space-times is two.

1. INTRODUCTION

A TETRAD of null vectors l^μ, n^μ, m^μ, and $\bar{m}^\mu$, with l^μ, n^μ real and m^μ complex, satisfying

$$l^\mu n_\mu = -m^\mu \bar{m}_\mu = 1,$$

all other contractions being zero, have been introduced into space-time by Newman and Penrose.[1] Tetrad indices m, n, $\cdots$, are introduced by contracting tensor indices μ, ν, $\cdots$, with the array

$$Z_{m\mu} = (l_\mu, n_\mu, m_\mu, \bar{m}_\mu).$$

These indices are raised and lowered by the "metric"

$$\eta_{mn} = \eta^{mn} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix}.$$

Certain linear combinations of the Ricci rotation coefficients[2] $\gamma^m{}_{np}$ are used, namely,

$$\kappa = \gamma_{131}, \quad \pi = -\gamma_{241}, \quad \epsilon = \tfrac{1}{2}(\gamma_{121} - \gamma_{341}), \quad \rho = \gamma_{134},$$
$$\lambda = -\gamma_{244}, \quad \alpha = \tfrac{1}{2}(\gamma_{124} - \gamma_{344}), \quad \sigma = \gamma_{133},$$
$$\mu = -\gamma_{243}, \quad \beta = \tfrac{1}{2}(\gamma_{123} - \gamma_{343}), \quad \nu = -\gamma_{242},$$
$$\gamma = \tfrac{1}{2}(\gamma_{122} - \gamma_{342}), \quad \text{and} \quad \tau = \gamma_{132}.$$

These are known as the spin coefficients. If the congruence defined by l^μ is geodesic, then $\kappa = 0$. In this case $\rho + \bar{\rho}$, $|\rho - \bar{\rho}|$, and $|\sigma|$ describe the divergence, rotation, and shear of the congruence.[3]

The tetrad components of the Weyl tensor are written as linear combinations of five complex scalars $\psi_0, \cdots, \psi_4$. The null tetrad can be chosen so that the ψ's take a canonical form as in Table I.

A null geodesic is called a geodesic ray if the tangent direction at each point is a principal direction of the Weyl tensor.[3] Hence l^μ is a geodesic ray if and only if $\psi_0 = \kappa = 0$. Goldberg and Sachs have proved that a shear-free geodesic ray in empty space is a repeated principal vector and, conversely, a repeated principal vector is geodesic and shear free. Hence

$$\psi_0 = \psi_1 = 0 \quad \text{if and only if} \quad \sigma = \kappa = 0,$$
$$\psi_3 = \psi_4 = 0 \quad \text{if and only if} \quad \lambda = \nu = 0.$$

The tangent vector V^μ to the trajectory of a conformal motion of space-time satisfies the equation[4]

$$\underset{v}{\pounds} g_{\mu\nu} = 2\phi g_{\mu\nu}, \tag{1.1}$$

where $\pounds$ denotes the Lie derivative, $g_{\mu\nu}$ is the metric of the space-time, and $\phi = V^\mu{}_{;\mu}$. The first set of integrability conditions of this equation is

$$\underset{v}{\pounds} C_{\nu\mu\lambda}{}^\chi = 0, \tag{1.2}$$

$$\underset{v}{\pounds} C_{\nu\mu\lambda} = -C_{\nu\mu\lambda}{}^\chi \phi_{,\chi}, \tag{1.3}$$

where $C_{\nu\mu\lambda}{}^\chi$ is the conformal Weyl tensor of the space-time and $C_{\nu\mu\chi}$ is defined in terms of the Ricci tensor $R_{\lambda\mu}$ and scalar R by

$$C_{\nu\mu\lambda} = -R_{\lambda[\mu;\nu]} + \tfrac{1}{6} g_{\lambda[\mu} R_{;\nu]}$$

Two special types of conformal motion exist. If ϕ is a constant the motion is a homothetic motion, and if

TABLE I. Canonical forms for each Petrov type.

type	principal vectors	canonical form
I	l, n	$\psi_0 = \psi_4 = 0, \psi_1\psi_3 \neq 0, 9\psi_2^2 \neq 16\psi_1\psi_3$
II	l, l, n	$\psi_0 = \psi_1 = \psi_4 = 0, \psi_2\psi_3 \neq 0$
D	l, l, n, n	$\psi_0 = \psi_1 = \psi_3 = \psi_4 = 0, \psi_2 \neq 0$
III	l, l, l, n	$\psi_0 = \psi_1 = \psi_2 = \psi_4 = 0, \psi_3 \neq 0$
N	l, l, l, l	$\psi_0 = \psi_1 = \psi_2 = \psi_3 = 0, \psi_4 \neq 0$

[1] E. Newman and R. Penrose, J. Math. Phys. 3, 565 (1962).
[2] L. P. Eisenhart, *Riemannian Geometry* (Princeton University Press, Princeton, New Jersey, 1925).
[3] R. Sachs, Proc. Roy. Soc. (London) **A264**, 309 (1961).
[4] K. Yano, *Theory of Lie Derivatives* (North-Holland Publishing Company, Amsterdam, 1955).

ϕ is zero the motion is a Killing motion. Killing motions which generate the symmetries of the space-time[2,4] are used in the later sections of this paper.

Equations (1.1)–(1.3) in the tetrad notation become

$$V_{m;n} + V_{n;m} = \tfrac{1}{2}\phi\eta_{mn} + V^s(\gamma_{msn} + \gamma_{nsm}), \quad (1.4)$$

$$\begin{aligned} C_{mnpq;s}V^s &+ C_{snpq}V^s{}_{;m} \\ &+ C_{mspq}V^s{}_{;n} + C_{mnsq}V^s{}_{;p} + C_{mnps}V^s{}_{;q} \\ &= \tfrac{1}{2}C_{mnpq}\phi + V^s[C_{rnpq}(\gamma_m{}^r{}_s + \gamma^r{}_{sm}) \\ &\quad + C_{mrpq}(\gamma_n{}^r{}_s + \gamma^r{}_{sn}) + C_{mnrq}(\gamma_p{}^r{}_s + \gamma^r{}_{sp}) \\ &\quad + C_{mnpr}(\gamma_q{}^r{}_s + \gamma^r{}_{sq})], \end{aligned} \quad (1.5)$$

$$\begin{aligned} C_{mnp;s}V^s &+ C_{snp}V^s{}_{;m} + C_{msp}V^s{}_{;n} + C_{mns}V^s{}_{;p} \\ &= -C_{mnp}{}^q\phi_{,q} + V^s[C_{rnp}(\gamma_m{}^r{}_s + \gamma^r{}_{sm}) \\ &\quad + C_{mrp}(\gamma_n{}^r{}_s + \gamma^r_s{}_n) + C_{mnr}(\gamma_p{}^r{}_s + \gamma^r{}_{sp})]. \end{aligned} \quad (1.6)$$

These equations are used in the next section to prove the following two theorems.

Theorem 1: A conformal motion of nonflat empty space-time must be homothetic, unless the space-time is type N with hypersurface orthogonal (twist free) geodesic rays.

Theorem 2: For each Petrov type the maximum order of the group of conformal motions admitted in nonflat empty space-time is at most one greater than the maximum order of the group of Killing motions.

In the following sections the equations are solved with $\phi = 0$ for those space-times which possess diverging hypersurface orthogonal geodesic rays.

2. MAXIMUM ORDER OF THE GROUP OF CONFORMAL MOTIONS ADMITTED BY EMPTY SPACE-TIMES

Equations (1.4)–(1.6) in empty space can be written explicitly as

$$DV_1 = (\epsilon + \bar{\epsilon})V_1 - \bar{\kappa}V_3 - \kappa V_4, \quad (2.1a)$$

$$\Delta V_2 = -(\gamma + \bar{\gamma})V_2 - \nu V_3 + \bar{\nu}V_4, \quad (2.1b)$$

$$\delta V_3 = \bar{\lambda}V_1 - \sigma V_2 - (\bar{\alpha} - \beta)V_3, \quad (2.1c)$$

$$\Delta V_1 + DV_2 - \tfrac{1}{2}\phi = (\gamma + \bar{\gamma})V_1 - (\epsilon + \bar{\epsilon})V_2 + (\pi - \bar{\tau})V_3 + (\bar{\pi} - \tau)V_4, \quad (2.1d)$$

$$\delta V_1 + DV_3 = (\bar{\alpha} + \beta + \bar{\pi})V_1 - \kappa V_2 + (\epsilon - \bar{\epsilon} - \bar{\rho})V_3 - \sigma V_4, \quad (2.1e)$$

$$\delta V_2 + \Delta V_3 = \bar{\nu}V_1 - (\bar{\alpha} + \beta + \tau)V_2 + (\mu + \gamma - \bar{\gamma})V_3 + \bar{\lambda}V_4, \quad (2.1f)$$

$$\delta V_3 + \bar{\delta}V_4 + \tfrac{1}{2}\phi = (\mu + \bar{\mu})V_1 - (\rho + \bar{\rho})V_2 + (\alpha - \bar{\beta})V_3 + (\bar{\alpha} - \beta)\bar{V}_3, \quad (2.1g)$$

$$\begin{aligned} V_2D\psi_2 &+ V_1\Delta\psi_2 - V_3\bar{\delta}\psi_2 - V_4\delta\psi_2 + \tfrac{1}{2}\phi\psi_2 + (\bar{\delta}V_2 - \Delta V_4)\psi_1 + (\delta V_1 - DV_3)\psi_3 \\ &= \psi_1[\nu V_1 + (2\pi + \bar{\tau} - \alpha - \bar{\beta})V_2 - \lambda V_3 + (\gamma - \bar{\gamma} + \bar{\mu} - 2\mu)V_4] \\ &\quad + \psi_3[(\bar{\alpha} + \beta - 2\tau - \bar{\pi})V_1 - \kappa V_2 + (2\rho - \bar{\rho} - \epsilon + \bar{\epsilon})V_3 + \sigma V_4], \end{aligned} \quad (2.2a)$$

$$\begin{aligned} V_1\Delta\psi_0 &+ V_2D\psi_0 - V_3\bar{\delta}\psi_0 - V_4\delta\psi_0 - \tfrac{1}{2}\phi\psi_0 + 2\psi_0(DV_2 - \delta V_4) + 2\psi_1(\delta V_1 - DV_3) \\ &= 2\psi_0[(2\gamma - \mu)V_1 + (\epsilon - \bar{\epsilon} + \bar{\rho})V_2 + (\pi - 2\alpha)V_3 + (\bar{\pi} - \bar{\alpha} - \beta)V_4] \\ &\quad + 2\psi_1[(\bar{\alpha} + \beta - \bar{\pi} - 2\tau)V_1 - \kappa V_2 + (2\rho - \bar{\rho} + \bar{\epsilon} - \epsilon)V_3 + \sigma V_4], \end{aligned} \quad (2.2b)$$

$$\begin{aligned} V_1\Delta\psi_4 &+ V_2D\psi_4 - V_3\bar{\delta}\psi_4 - V_4\delta\psi_4 - \tfrac{1}{2}\phi\psi_4 + 2\psi_4(\Delta V_1 - \delta V_3) + 2\psi_3(\delta V_2 - \Delta V_3) \\ &= 2\psi_4[(\bar{\gamma} - \gamma - \bar{\mu})V_1 + (\rho - 2\epsilon)V_2 + (\alpha + \bar{\beta} - \bar{\tau})V_3 + (2\beta - \tau)V_4] \\ &\quad + 2\psi_3[\nu V_1 + (2\pi + \bar{\tau} - \alpha - \bar{\beta})V_2 - \lambda V_3 + (\gamma - \bar{\gamma} + \bar{\mu} - 2\mu)V_4], \end{aligned} \quad (2.2c)$$

$$\begin{aligned} V_1\Delta\psi_1 &+ V_2D\psi_1 - V_3\bar{\delta}\psi_1 - V_4\delta\psi_1 - \psi_0\Delta V_4 + \psi_1(DV_2 - \delta V_4) + \psi_2(\delta V_1 - 2DV_3) \\ &= \psi_0[(\pi + \bar{\tau})V_2 - \lambda V_3 + (\gamma - \bar{\gamma} - \mu)V_4] + \psi_1[(2\gamma - \mu)V_1 + (\epsilon - \bar{\epsilon} + \bar{\rho})V_2 + (\pi - 2\alpha)V_3 + (\bar{\pi} - \bar{\alpha} - \beta)V_4] \\ &\quad + \psi_2[(\bar{\alpha} + \beta - 2\bar{\pi} - 3\tau)V_1 - \kappa V_2 + (3\rho - \bar{\rho} - 2\epsilon + 2\bar{\epsilon})V_3 + 2\sigma V_4], \end{aligned} \quad (2.2d)$$

$$\begin{aligned} V_1\Delta\psi_3 &+ V_2D\psi_3 - V_3\bar{\delta}\psi_3 - V_4\delta\psi_4 - \psi_4DV_3 + \psi_3(\Delta V_1 - \bar{\delta}V_4) + \psi_2(\bar{\delta}V_2 - 2\Delta V_4) \\ &= \psi_4[-(\bar{\pi} + \tau)V_1 + (\bar{\epsilon} - \epsilon + \rho)V_3 + \sigma V_4] + \psi_3[(\bar{\gamma} - \gamma - \bar{\mu})V_1 + (\rho - 2\epsilon)V_2 + (\alpha + \bar{\beta} - \bar{\tau})V_3 \\ &\quad + (2\beta - \tau)V_4] + \psi_2[\nu V_1 + (3\pi + 2\bar{\tau} - \alpha - \bar{\beta})V_2 - 2\lambda V_3 + (2\gamma - 2\bar{\gamma} + \bar{\mu} - 3\mu)V_4], \end{aligned} \quad (2.2e)$$

$$\psi_2 D\phi - \psi_1 \bar{\delta}\phi = 0, \tag{2.3a}$$

$$\psi_2 \Delta\phi - \psi_3 \delta\phi = 0, \tag{2.3b}$$

$$\psi_2 \delta\phi - \psi_1 \Delta\phi = 0, \tag{2.3c}$$

$$\psi_2 \bar{\delta}\phi - \psi_3 D\phi = 0, \tag{2.3d}$$

$$\psi_1 D\phi - \psi_0 \bar{\delta}\phi = 0, \tag{2.3e}$$

$$\psi_3 \Delta\phi - \psi_4 \delta\phi = 0, \tag{2.3f}$$

$$\psi_0 \Delta\phi - \psi_1 \delta\phi = 0, \tag{2.3g}$$

$$\psi_4 D\phi - \psi_3 \bar{\delta}\phi = 0. \tag{2.3h}$$

In the above, D, Δ, δ, and $\bar{\delta}$ are the intrinsic derivatives defined by Newman and Penrose which satisfy the commutation relations

$$\Delta D - D\Delta = (\gamma + \bar{\gamma})D + (\epsilon + \bar{\epsilon})\Delta - (\tau + \bar{\pi})\bar{\delta} - (\bar{\tau} + \pi)\delta, \tag{2.4}$$

$$\delta D - D\delta = (\bar{\alpha} + \beta - \bar{\pi})D + \kappa\Delta - \sigma\bar{\delta} - (\bar{\rho} + \epsilon - \bar{\epsilon})\delta, \tag{2.5}$$

$$\delta\Delta - \Delta\delta = -\bar{\nu}D + (\tau - \bar{\alpha} - \beta)\Delta + \bar{\lambda}\bar{\delta} + (\mu - \gamma + \bar{\gamma})\delta, \tag{2.6}$$

$$\bar{\delta}\delta - \delta\bar{\delta} = (\bar{\mu} - \mu)D + (\bar{\rho} - \rho)\Delta - (\bar{\alpha} - \beta)\bar{\delta} - (\bar{\beta} - \alpha)\delta. \tag{2.7}$$

If the space-time is not of type N, Eqs. (2.3) yield $D\phi = \Delta\phi = \delta\phi = 0$. If the space-time is of type N, then the equations yield $D\phi = \delta\phi = 0$ (the canonical form for the ψ's is used in each Petrov type). Putting ϕ in Eq. (2.7) gives

$$(\rho - \bar{\rho})\Delta\phi = 0.$$

Hence if $\bar{\rho} \neq \rho$, $\Delta\phi$ must be zero. Now $D\phi$, $\Delta\phi$, and $\delta\phi$ are all zero only if ϕ is a constant and the motion homothetic. Hence Theorem 1 is proved.

The order of the group of conformal motions admitted by a space-time is equal to the number of independent solutions to (1.4)–(1.6) and subsequent integrability conditions considered as *algebraic* equations in the unknowns V_m, $V_{m;n}$, ϕ, and $\phi_{,m}$.[4,5] If ϕ is a constant, there can only be one more independent solution to the equations than in the case $\phi = 0$. This therefore proves Theorem 2 for all but type N space-times with $\rho = \bar{\rho}$.

It remains to prove Theorem 2 for type N space-times with $\rho = \bar{\rho}$. In this case Eqs. (2.1), (2.2e), and (2.2c) are used to obtain all derivatives of $\mathbf{V}_m$ in terms of ΔV_3, V_m, and ϕ. Substituting V_2 in Eq. (2.4) then gives $\Delta\phi$ in terms of ΔV_3, V_m, and ϕ. The maximum possible order of the group of conformal motions is therefore $2 + 4 + 1 = 7$. This completes the proof of Theorem 2.

[5] L. P. Eisenhart, *Continuous Groups of Transformations* (Princeton University Press, Princeton, New Jersey, 1933).

3. EMPTY SPACE-TIMES POSSESSING DIVERGING, HYPERSURFACE ORTHOGONAL, GEODESIC RAYS

The space-times of this section are characterized in the tetrad notation by

$$\psi_0 = \kappa = 0, \quad \rho = \bar{\rho} \neq 0.$$

There are three distinct classes:

Class A: $\rho^2 \neq \sigma\bar{\sigma} \neq 0$ (Newman and Tamburino: spherical class)[6];

Class B: $\rho^2 = \sigma\bar{\sigma} \neq 0$ (Newman and Tamburino: cylindrical class)[6];

Class C: $\sigma = 0$ (Robinson and Trautman).[7]

These metrics have been found by Newman and Tamburino[6,8] using the coordinate system first introduced by Robinson and Trautman.[7] In this coordinate system the tetrad vectors become

$$l^\mu = \delta^\mu_2,$$
$$n^\mu = \delta^\mu_1 + U\delta^\mu_2 + X^\alpha\delta^\mu_\alpha,$$
$$m^\mu = \omega\delta^\mu_2 + \xi^\alpha\delta^\mu_\alpha,$$

where $\alpha = 3, 4$. Equations (2.1) are now solved with $\phi = 0$. The resulting Killing motions are of physical

[6] E. Newman and L. A. Tamburino, J. Math. Phys. **3**, 902 (1962).
[7] I. Robinson and A. Trautman, Proc. Roy. Soc. (London) **A265**, 463 (1962).
[8] L. A. Tamburino, Ph.D. thesis, University of Pittsburgh (1962).

importance since they describe the symmetries of the space-times.

4. SPACE-TIMES OF CLASS A

The metric of this class is

$$g^{22} = -\frac{2r^2(\zeta\bar{\zeta})^{\frac{1}{2}}}{R^2} + \frac{2rL}{A} + \frac{2r^3A(\zeta^2+\bar{\zeta}^2)}{R^4} - \frac{4r^2A^2(\zeta\bar{\zeta})^{\frac{3}{2}}}{R^4},$$

$$g^{23} = 4A^2(\zeta\bar{\zeta})^{\frac{3}{2}}x^3\left(\frac{L}{2a^3} - \frac{(r-2a)}{2a^2R^2} - \frac{(r-a)}{R^4}\right),$$

$$g^{24} = 4A^2(\zeta\bar{\zeta})^{\frac{3}{2}}x^4\left(\frac{L}{2a^3} - \frac{(r+2a)}{2a^2R^2} - \frac{(r+a)}{R^4}\right),$$

$$g^{33} = \frac{2(\zeta\bar{\zeta})^{\frac{3}{2}}}{(r+a)^2}, \quad g^{44} = -\frac{2(\zeta\bar{\zeta})^{\frac{3}{2}}}{(r-a)^2},$$

$$g^{12} = 1, \quad g^{34} = g^{13} = g^{14} = 0,$$

where

$$x^1 \equiv u, \quad x^2 \equiv r, \quad x^3 + ix^4 \equiv \zeta, \quad R^2 = r^2 - a^2,$$

$$L = \tfrac{1}{2}\log_e\left(\frac{r+a}{r-a}\right), \quad a = A(\zeta\bar{\zeta})^{\frac{1}{2}}$$

with

$$A = Bu \quad \text{or} \quad B \quad (B \text{ is a } \textit{real} \text{ constant}).$$

The spin coefficients used in the calculations are

$$\epsilon = \kappa = \pi = 0, \quad \tau = \bar{\alpha} + \beta,$$
$$\rho = -r/R^2, \quad \sigma = a/R^2,$$

$$\alpha = \frac{\psi_1^0 Lr}{2a^2R^2} + \frac{1}{R^2}\left[r\alpha^0 + a\left(\bar{\alpha}^0 - \frac{\psi_1^0}{2a^2}\right)\right],$$

$$\beta = -\frac{\psi_1^0 L}{2aR^2} - \frac{1}{R^2}[r\bar{\alpha}^0 + a\alpha^0],$$

$$\lambda = -\frac{rL^2\psi_1^{0^2}}{4a^4R^2} - \frac{rL\psi_1^0\bar{\psi}_1^0}{2a^4R^2} + \frac{ra_{,1}}{R^2} + \frac{\psi_1^0\bar{\psi}_1^0}{2a^3R^2},$$

$$\gamma = -\frac{rL^2\psi_1^0\bar{\psi}_1^0}{4a^4R^2} + \frac{\psi_1^{0^2}L^2}{4a^3R^2} + \frac{aLr(\psi_1^0\bar{\alpha}^0 - \bar{\psi}_1^0\alpha^0)}{2a^3R^2} + \frac{L\psi_1^0\bar{\psi}_1^0}{2a^3R^2} - \frac{\psi_1^0\bar{\psi}_1^0}{4a^3}\left[\frac{L}{2a^2} - \frac{r}{2aR^2}\right] + \frac{1}{2a^2R^2}\left[\frac{\psi_1^{0^2}}{2a} - a(\psi_1^0\bar{\alpha}^0 - \bar{\psi}_1^0\alpha^0)\right],$$

where

$$\psi_1^0 = 2A^2(\zeta\bar{\zeta})^{\frac{3}{4}}\zeta, \quad \alpha^0 = \tfrac{3}{4}\bar{\zeta}^{\frac{3}{4}}/\zeta^{\frac{1}{4}}.$$

The "metric variables" used are

$$\omega = -\bar{\psi}_1^0 L/2a^2 + (r\omega^0 - a\bar{\omega}^0)/R^2, \quad \omega^0 = -A(\zeta\bar{\zeta})^{\frac{1}{4}}\bar{\zeta},$$
$$\xi^\alpha = (r\xi^{0\alpha} - a\bar{\xi}^{0\alpha})/R^2, \quad \xi^{03} = -i\xi^{04} = p = (\zeta\bar{\zeta})^{\frac{3}{4}}.$$

Since $\psi_0 = 0$ and $\psi_1 \neq 0$, Eq. (2.2b) becomes simply

$$\delta V_1 - DV_3 = -\tau V_1 + \rho V_3 + \sigma V_4. \qquad (4.1)$$

Equations (2.1a) and (2.1e) plus (4.1) yield

$$V_1 = V_1(u)$$

and then from (2.1e) minus (4.1)

$$V_3 = (\psi_1^0 V_1 rL/2a^3) + a^0 r + a\bar{a}^0 - (\psi_1^0 V_1/2a^2),$$

where a^0 is independent of r. The imaginary part of (2.1c) yields

$$\bar{\xi}^{0\alpha}a^0_{,\alpha} - \xi^{0\alpha}\bar{a}^0_{,\alpha} + 3a^0\psi_1^0/4a^2 - 3\bar{a}^0\psi_1^0/4a^2 = 0, \quad (4.2)$$

$$\bar{\xi}^{0\alpha}\bar{a}^0_{,\alpha} - \xi^{0\alpha}a^0_{,\alpha} + 3\bar{a}^0\bar{\psi}_1^0/4a^2 - 3a^0\psi_1^0/4a^2 = 0, \quad (4.3)$$

while the real part of (2.1c) yields

$$\begin{aligned} V_2 = {} & rL^2(-\psi_1^{0^2}V_1 - \bar{\psi}_1^{0^2}V_1)/4a^5 \\ & - L^2\psi_1^0\bar{\psi}_1^0V_1/4a^4 - L(\bar{a}^0\psi_1^0 + a^0\bar{\psi}_1^0)/2a \\ & + r(2aV_1a_{,1} - a^0\bar{\psi}_1^0 - \bar{a}^0\psi_1^0)/2a^2 \\ & + (-2a^3a^0\psi_1^0 - 2a^3\bar{a}^0\bar{\psi}_1^0 + V_1\psi_1^0\bar{\psi}_1^0)/4a^4 \\ & + Lr(-2a^3a^0\psi_1^0 - 2a^3\bar{a}^0\bar{\psi}_1^0 - V^1\psi_1^0\bar{\psi}_1^0)/4a^5 \\ & + R^2(-4a^2\xi^{0\alpha}a^0_{,\alpha} - 4a^2\bar{\xi}^{0\alpha}\bar{a}^0_{,\alpha} - 3a^0\psi_1^0 - 3\bar{a}^0\bar{\psi}_1^0)/8a^3. \end{aligned}$$

Equation (2.1d) gives

$$\xi^{0\alpha}a^0_{,\alpha} + \bar{\xi}^{0\alpha}\bar{a}^0_{,\alpha} + 3a^0\psi_1^0/4a^2 + 3\bar{a}^0\bar{\psi}_1^0/4a^2 = 0, \quad (4.4)$$

$$-dV_1/du - V_1a_{,1}/a + \bar{a}^0\psi_1^0/2a^2 + a^0\bar{\psi}_1^0/2a^2 = 0. \quad (4.5)$$

Equations (4.2)–(4.5) are consistent only if

$$a^0 = 0 \quad \text{and} \quad V_1 = \text{const } A^{-1}.$$

There is therefore at most one Killing motion. Furthermore, substitution into (2.1g) yields

$$V_1a_{,1} = 0,$$

and so the Killing motion exists only when $a_{,1} = 0$. In this case the coordinate x^1 is ignorable and the Killing motion generates the consequent symmetry.

5. SPACE-TIMES OF CLASS B

There are two metrics for this class

(a) $\psi_1 \neq \bar{\psi}_1$ (see Refs. 6, 8),

$$\begin{aligned} g_{11} &= -4a^2(cn^2ay)(\log_e r)^2 - [e + a^2\log_e(r^2cn^4ay)]/cn^2ay, \\ g_{12} &= 1, \quad g_{23} = g_{24} = 0, \\ g_{13} &= -4Y[r + 4a^2u(cn^2ay)\log_e r], \\ g_{14} &= -2(cn^2ay)\log_e r, \\ g_{33} &= -r^2/2 - 64a^4Y^2u^2/cn^2ay, \\ g_{34} &= -8uYcn^2(ay), \\ g_{44} &= -cn^2(ay)/a^2, \end{aligned}$$

where $x^1 \equiv u$, $x^2 \equiv r$, $x^3 \equiv y$, $cn(ay)$ is an elliptic function of modulus $\kappa = 1/\sqrt{2}$ and

$$Y = \pm\frac{a(1 - cn^4ay)^{\frac{1}{2}}}{2\sqrt{2}\,cn(ay)}, \quad \begin{matrix} + & \text{if} & y > 0, \\ - & \text{if} & y < 0, \end{matrix}$$

with a and e arbitary constants.

(b) $\psi_1 = \bar{\psi}_1$ (see Refs. 3, 6, 8),

$$g_{11} = -[\kappa + \log_e (r^2 y^4)],$$
$$g_{13} = 2r/y, \quad g_{12} = -1,$$
$$g_{33} = -r^2, \quad g_{44} = -y^2,$$
$$g_{14} = g_{23} = g_{24} = g_{34} = 0,$$

where κ is an arbitrary constant.

Equations (2.1) are solved for these two space-times in a similar manner as in the previous section. It is found that there are two Killing motions for the second metric (corresponding to the ignorable coordinates x^4 and u) but only one Killing motion for the first (corresponding to the ignorable coordinate x^4).

6. SPACE-TIMES OF CLASS C

The metric of this class is

$$g^{22} = 2U^0 - 4\gamma^0 r - 2\psi_2^0/r,$$
$$g^{12} = 1, \quad g^{33} = g^{44} = -2P^2/r^2,$$
$$g^{11} = g^{13} = g^{14} = g^{23} = g^{24} = g^{34} = 0,$$

where

$$U^0 = -P^2 \nabla^2 L, \quad 2\gamma^0 = -\partial L/\partial u, \quad L = \log P,$$
$$x^1 \equiv u, \quad x^2 \equiv r, \quad x^3 + ix^4 = \zeta, \quad \text{and} \quad \nabla \equiv 2\partial/\partial\bar{\zeta}.$$

P is a real function, independent of r, satisfying the equation

$$(\partial/\partial u + 6\gamma^0)\psi_2^0 = P^2 \nabla^2 U^0. \tag{6.1}$$

Equations (2.1) together with (2.2) readily yield the following information:

$$V_1 = g(u), \quad V_2 = -Ug - r\dot{g}, \quad V_3 = rf(\zeta)/P,$$

where

$$\psi_2^0 \dot{g} = 0, \tag{6.2}$$
$$f + \bar{f} = f\bar{\nabla}L + \bar{f}\nabla L + 2g\gamma^0 - \dot{g}, \tag{6.3}$$
$$f\bar{\nabla}U^0 + \bar{f}\nabla U^0 = g\partial U^0/\partial u + 2U^0\dot{g}, \tag{6.4}$$
$$f\bar{\nabla}\gamma^0 + \bar{f}\nabla\gamma^0 = g\partial\gamma^0/\partial u + \gamma^0\dot{g} - \tfrac{1}{2}\ddot{g}. \tag{6.5}$$

Equations (6.4) and (6.5) are integrability conditions of Eq. (6.3).

Equations (6.2)–(6.5) can be solved, but many different cases arise and, in order to make the notation concise, it becomes necessary to define several auxiliary functions. Information is more readily obtained by adapting the coordinate system to the Killing vectors. The metric is invariant under the transformations

$$u' = s(u), \quad r' = r/\dot{s}, \quad l^{\mu'} = \dot{s} l^\mu, \quad n^{\mu'} = n^\mu/\dot{s}, \tag{6.6}$$
$$\zeta' = h(\zeta), \quad m^{\mu'} = e^{ic} m^\mu, \quad \text{where} \quad e^{ic} = \dot{h}/|\dot{h}|. \tag{6.7}$$

TABLE II. Canonical forms for the function L.

case	set	canonical form	remaining freedoms
(i)	$g = 1$	$L = L(\zeta, \bar{\zeta})$	(6.7) and $u' = u +$ const.
(ii)	$f = 1$	$L = L(\zeta - \bar{\zeta}, u)$	(6.6) and $\zeta' = \zeta +$ const.
(iii)	$g = 2f = 1$	$L = L(\zeta + u, \bar{\zeta} + u)$	$u' = u +$ const. $\zeta' = \zeta +$ const.

If $\psi_2^0 \neq 0$, then using (6.5) one can set $\psi_2^0 = 1$. The remaining coordinate freedom is $u^1 = u +$ const. Under (6.6) and (6.7) $g' = \dot{s}g$ and $f' = hf$.

Suppose now that a Killing vector exists. There are three cases invariant under (6.6) and (6.7):

$$\text{(i) } g \neq 0, \quad f = 0; \quad \text{(ii) } g = 0, \quad f \neq 0;$$
$$\text{(iii) } g \neq 0, \quad f \neq 0.$$

The coordinate system can be *adapted* to each case as in Table II. The canonical form is obtained from Eq. (6.3). If $\psi_2^0 \neq 0$, Eq. (6.2) yields $g =$ const. By scaling the Killing vector, one can set $g = 1$. Hence the above adapted coordinate systems can be chosen and at the same time ψ_2^0 can be set equal to unity.

Equations (6.2)–(6.5) are now solved for each canonical form of L. The results are summarized below.

Case (i) with $\psi_2^0 = 1$

If $\nabla U^0 = 0$, the metric is of type D.S.[7] Following Robinson and Trautman one can set

$$\sqrt{2}P = 1 - \tfrac{1}{2}U^0 \zeta\bar{\zeta}.$$

Equations (6.2)–(6.5) then yield the four Killing vectors,

$$V_m = \delta^1_m - U\delta^2_m,$$
$$V_m = \delta^3_m \zeta r/P - \zeta^4_m \bar{\zeta} r/P,$$
$$V_m = (2 - U^0\zeta^2 r/P)\delta^3_m + (2 - U^0\bar{\zeta}^2 r/P)\delta^4_m,$$
$$V_m = (2 + U^0\zeta^2 r/P)\delta^3_m - (2 + U^0\bar{\zeta}^2 r/P)\delta^4_m.$$

If $\nabla U^0 \neq 0$, one can set $U^0 = \zeta + \bar{\zeta}$. When $\nabla L \neq \bar{\nabla} L$ only one Killing vector exists, but if $\nabla L = \bar{\nabla} L$, two Killing vectors exist, namely,

$$V_m = \delta^1_m - U\delta^2_m \quad \text{and} \quad V_m = \delta^3_m r/P - \zeta^4_m r/P. \tag{6.8}$$

In this last case the field equation (6.1) becomes simply

$$z = -4e^{2L}\ddot{L}, \tag{6.9}$$

where $z = \zeta + \bar{\zeta}$ and L is a function of z alone. A *singular* solution of (6.9) is

$$L = \tfrac{1}{2}\log(z^3/6). \tag{6.10}$$

Case (i) with $\psi_2^0 = 0$

If $\eth U^0 = 0$, the space is flat. If $\eth U^0 \neq 0$, one can again set $U^0 = \zeta + \bar{\zeta}$. When L is given by (6.10) with $z = \zeta + \bar{\zeta}$, the following three Killing vectors exist,

$$V_m = \delta^1_m - U\delta^2_m,$$
$$V_m = \delta^3_m r/P - \delta^4_m r/P,$$
$$V_m = u\delta^1_m - Uu\delta^2_m - r\delta^2_m + \delta^3_m r\zeta/P + \delta^4_m r\bar{\zeta}/P. \tag{6.11}$$

If L is a function of z alone satisfying (6.9), then two Killing vectors of the form (6.8) exist. It is interesting to note that it is the *singular* solution of (6.9) which gives rise to the three Killing vectors. Finally, if $\bar{\eth} L \neq \eth L$, there can exist at most two Killing vectors, and if these exist one can set

$$L = \tfrac{3}{2} \log \zeta + F(z),$$

where $z = (\zeta' + \bar{\zeta})/\zeta$ and

$$z = 4e^{2F}[(z-1)\ddot{F} + \dot{F}]. \tag{6.12}$$

The two Killing vectors are

$$V_m = \delta^1_m - U\delta^2_m,$$
$$V_m = u\delta^1_m - uU\delta^2_m - r\delta^2_m + \delta^3_m r\zeta/P + \zeta^4_m r\bar{\zeta}/P. \tag{6.13}$$

Equation (6.12) has the *singular* solution

$$F = \tfrac{1}{2} \log (z^3/6)$$

and this again gives rise to the three Killing vectors.

Case (ii) with $\psi_2^0 = 1$

If $\eth^2 U^0 = 0$, then (6.1) implies that L is independent of u, a situation covered by case (i). If $\eth^2 U^0 \neq 0$, then there exists at most two Killing vectors, and if these exist L can be set equal to one of the two expressions

$$L = \log (\zeta - \bar{\zeta}) + \log F(z), \quad \text{where} \quad z = (\zeta - \bar{\zeta})e^{au}$$

or

$$L = \log F(z), \quad \text{where} \quad z = \zeta - \bar{\zeta} + iau.$$

In both the above, a is a real constant and $i = (-1)^{\frac{1}{2}}$. The field equations in the two cases become

$$3a \frac{dF}{du} = -16F^3 z \frac{d^2}{du^2}\left[F^2 - z^2\left(\frac{d^2F}{du^2}F - \frac{dF}{du}\frac{dF}{du}\right)\right], \tag{6.14}$$

$$3ia \frac{dF}{du} = -16F^3 \frac{d^2}{du^2}\left[\frac{dF}{du}\frac{dF}{du} - F\frac{d^2F}{du^2}\right]. \tag{6.15}$$

The two Killing vectors are

$$V_m = \delta^3_m r/P + \delta^4_m r/P,$$
$$V_m = \delta^1_m - U\delta^2_m + ar\zeta\delta^3_m/2P + ar\bar{\zeta}\delta^4_m/2P$$

or

$$V_m = \delta^3_m r/K + \delta^4_m r/P,$$
$$V_m = \delta^1_m - U\delta^2_m + iar\zeta\delta^3_m/4P - iar\bar{\zeta}\delta^4_m/4P.$$

Equation (6.15) has the *singular* solution

$$F = (-iaz^3/4)^{\frac{1}{3}}. \tag{6.16}$$

Case (ii) with $\psi_2^0 = U^0 = 0$

Since $U^0 = 0$,

$$L = iG(u)(\zeta - \bar{\zeta}) + H(u).$$

If $\dot{G}(u) = 0$ the space is flat. Otherwise, one can set $G(u) = u$ and then there are at most two Killing vectors. If these exist, one can set

$$H = -2 \log u + \text{const}$$

and the Killing vectors are

$$V_m = \delta^3_m r/P + \delta^4_m r/P,$$
$$V_m = u\delta^1_m - uU\delta^2_m - r\delta^2_m + \delta^3_m r\zeta/2P + \delta^4_m r\bar{\zeta}/2P.$$

Case (ii) with $\psi_2^0 = \eth U^0 = 0$, $U^0 \neq 0$

One can set $U^0 = \epsilon$, where $\epsilon = \pm 1$ and then the *general* solution to the field equation (6.1) is

$$P = (\epsilon/4\nu^2)^{\frac{1}{2}} \sinh [\mu - \nu^2 \epsilon i(\zeta - \bar{\zeta})], \tag{6.17}$$

where μ and ν are functions of u alone. The case $\dot{\mu} = \dot{\nu} = 0$ has been covered in (i). If not both $\dot{\mu}$ and $\dot{\nu}$ are zero, then there exists at most two Killing vectors, and if these exist one can set

$$\nu = e^{au} \quad \text{and} \quad \mu = b,$$

where a and b are arbitary real constants. The two Killing vectors are

$$V_m = \delta^3_m r/P + \delta^4_m r/P,$$
$$V_m = \delta^1_m - U\delta^2_m + \delta^3_m r(a\zeta - 2i\epsilon a)/P + \zeta^4_m r(a\bar{\zeta} + 2i\epsilon a)/P.$$

Case (ii) with $\psi_2 = 0$, $\eth U^0 \neq 0$

In this case there are again at most two Killing vectors, and if these exist L can be taken in one of the following forms:

$$L = -3 \log u + F(z), \quad \text{where} \quad z = i\epsilon u^2(\zeta - \bar{\zeta}) + a \tag{6.18}$$

or

$$L = F(z), \quad \text{where} \quad z = au + (\zeta - \bar{\zeta})i\epsilon. \tag{6.19}$$

In both the above, a is a constant and $\epsilon = \pm 1$. The field equation takes the form

$$z = -4e^{2F}\ddot{F}$$

and has the singular solution

$$F = \tfrac{1}{2} \log (z^3/6).$$

The Killing vectors for (6.18) are

$$V_m = \delta^3_m r/P + \delta^4_m r/P,$$
$$V_m = u\delta^1_m - uU\delta^2_m - r\delta^2_m + \delta^3_m r\zeta/P + \delta^4_m r\bar{\zeta}/P.$$

The Killing vectors for (6.19) are

$$V_m = \delta^3_m r/P + \delta^4_m r/P,$$
$$V_m = \delta^1_m - U\delta^2_m - \delta^3_m i\epsilon ra/P + \delta^4_m i\epsilon ra/P.$$

Case (iii) with $\psi_2^0 = 1$ or $\psi_2^0 = \nabla U^0 = 0$, $U^0 \neq 0$

The only metrics which admit more than one Killing vector have already been covered by either case (i) or case (ii).

Case (iii) with $\psi_2^0 = U^0 = 0$

Now

$$L = F(\zeta + u) + \bar{F}(\bar{\zeta} + u)$$

and the condition for nonflat space is

$$\dddot{F} + 2\ddot{F}\dot{F} \neq 0.$$

There are, again, at most two Killing vectors, and if they exist $\dot{F}$ can be set equal to one of the three forms:

$$\dot{F} = [-ba + bd(a + 4)e^{2b(u+\zeta)}][a - ade^{2b(u+\zeta)}]^{-1}, \tag{6.20}$$

where a, b, and d are nonzero constants with a imaginary and b real;

$$\dot{F} = -[a(\zeta + u)]^{-1}, \tag{6.21}$$

where a is a nonzero imaginary constant;

$$\dot{F} = -b + de^{b(\zeta+u)}, \tag{6.22}$$

where b and d are nonzero constants with b real. In each case one Killing vector is

$$V_m = \delta^1_m - U\delta^2_m + \delta^3_m r/2P + \delta^4_m r/2P. \tag{6.23}$$

The second Killing vectors are

$$V_m = -2e^{2bu}\delta^1_m + 2Ue^{2bu}\delta^2_m + 4bre^{2bu}\delta^2_m - \delta^3_m re^{-2b\zeta}/Pd - \delta^4_m re^{-2b\bar{\zeta}}/P\bar{d},$$
$$V_m = u\delta^1_m - uU\delta^2_m - r\delta^2_m - \delta^3_m r\zeta/2P - \delta^4_m r\bar{\zeta}/2P,$$
$$V_m = \delta^3_m re^{-b\zeta}/Pd - \delta^4_m re^{-b\bar{\zeta}}/P\bar{d}.$$

Case (iii) with $\psi_2^0 = 0$, $\nabla U^0 \neq 0$

Excluding those metrics already covered by cases (i) and (ii), it is found that if there exist more than one Killing vector then L takes the form

$$L = b(\zeta + \bar{\zeta} + 2u)/4 + F(z), \tag{6.24}$$

with

$$z = [kb/2d + e^{-\frac{1}{2}b(\zeta+u)}][\bar{k}b/2\bar{d} + e^{-\frac{1}{2}b(\bar{\zeta}+u)}]^{-1},$$

where k, b, and d are constants with b, d nonzero and b real, or

$$L = F(z) \tag{6.25}$$

with $z = [u + \zeta][u + \bar{\zeta}]^{-1}$. Both of these metrics admit the Killing vector (6.23) together with one other, namely,

$$V_m = -2\delta^1_m e^{-\frac{1}{2}bu}/b + 2\delta^2_m Ue^{-\frac{1}{2}bu}/b - \delta^2_m re^{-\frac{1}{2}bu} + \delta^3_m rke^{\frac{1}{2}b\zeta}/2dP + \delta^4_m r\bar{k}e^{\frac{1}{2}\bar{\zeta}b}/2\bar{d}P$$

or

$$V_m = u\delta^1_m - uU\delta^2_m - r\delta^2_m - \delta^3_m r\zeta/2P - \delta^4_m r\bar{\zeta}/2P.$$

The function F in both (6.24) and (6.25) must satisfy the field equation

$$8e^{2F}[\dot{F} + z\ddot{F}] = 1/d^2z^2 + 1/\bar{d}^2. \tag{6.26}$$

In the discussion of this section emphasis has been placed on those metrics admitting more than one Killing vector. It is worth noting that Robinson and Trautman give examples of metrics possessing only one or no Killing vectors.

7. MAXIMUM ORDER OF THE GROUP OF KILLING MOTIONS ADMITTED BY EMPTY SPACE-TIMES OF TYPE III

Equations (2.1) and (2.2) have been analyzed to give the maximum order of the group of Killing motions admitted by empty space-times of each Petrov type. The results are in agreement with those of Bialas and Bialas[9] excepting for type-II and type-III space-times. The authors were inclined to believe that the discrepancy arose because only the first set of integrability conditions were considered. However, it has been shown that the metric of the last section, for which $\psi_2^0 = 0$ and $L = \frac{1}{2}\log(\zeta + \bar{\zeta})^3/6$, admits a group of Killing motions of order three. This metric is of type III and provides a counterexample to the Bialas result which states that the maximum order for type-III space-times is two. The results of Bialas and Bialas are based on the work of Petrov.[10]

8. REMARKS

It is interesting to see how the use of the tetrad calculus helps in solving the problems discussed in this paper. Although the first set of integrability conditions of conformal motions looks formidable, it becomes quite concise and quickly yields any required information when simplified by taking the canonical forms for each particular Petrov type. It seems reasonable, having obtained metrics by the tetrad method, to investigate their symmetries using the same methods. Here the integrability conditions proved extremely helpful, and, of course, they would not have been so readily available using the more usual techniques.

[9] E. Bialas and A. Bialas, Acta Phys. Polon. **24**, 515 (1963).

[10] A. Z. Petrov, *Einstein Spaces* (Pergamon Press, Inc., New York, 1964).

For space-times of type I it is found that Theorem 2 can be strengthened so that the maximum orders of the groups of conformal and Killing motions are equal.

Note added in proof: It has been suggested that the metrics obtained by Newman and Tamburino[6] do not exhaust all space-times containing diverging, hypersurface orthogonal, geodesic rays (see Sec. 5). One of the authors (C. D. C.) has made an independent check of the calculations and has found that no other metrics do in fact exist. However, the metric (b) was found directly rather than as a limiting case of the metric (a) and the g_{33} component of the metric (a) itself was found to be misprinted in the paper of Newman and Tamburino. This misprint has been corrected here.

ACKNOWLEDGMENT

One of the authors (D. C. F.) wishes to express his gratitude to the Kingston upon Hull Education Committee for financial support.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Induced Representations of Strong Coupling Groups*

T. Cook†
University of Wisconsin, Madison, Wisconsin, and *Argonne National Laboratory, Argonne, Illinois*

B. Sakita‡
Argonne National Laboratory, Argonne, Illinois

(Received 27 June 1966)

Mackey's method of induced representations is applied to the strong coupling group $G = K \cdot T$, where K is compact and T is Abelian, to obtain the general irreducible representations. The form of the meson–isobar–isobar couplings is obtained by reducing these, with the use of the Peter–Weyl theorem, to irreducible representations of the compact subgroup K. The results are applied to the cases: pseudoscalar octet mesons, $K = SU_2 \otimes SU_3$, $T = T_{24}$, and SU_6 35-plet mesons, $K = SU_6$, $T = T_{35}$. Explicit representations are obtained which are consistent with mass formulas $\mathcal{M} \propto f_2(\mu) - \frac{5}{8}j(j+1)$ and $\mathcal{M} \propto G^2$ in the respective cases.

I. INTRODUCTION

In previous papers[1,2] we have exhibited the group of the strong coupling theory. The strong coupling problem is defined by the equations

$$[A_\alpha, A_\beta] = 0, \tag{1}$$

$$\Lambda_{\beta\alpha} = [A_\beta, [\mathcal{M}, A_\alpha]], \tag{2a}$$

$$\Lambda_{\beta\alpha} = \sum_\gamma \Lambda_{\beta\gamma}\Lambda_{\gamma\alpha}. \tag{2b}$$

In the above all terms are considered as operators on the isobar states. A_α is the Hermitian source operator of meson α and hence matrix elements of A_α are proportional to the meson–isobar–isobar couplings. A factor g^2 has been removed so that A_α is finite. $\mathcal{M}$ is a diagonal operator whose matrix elements are proportional to the isobar mass differences. Here a factor of $1/g^2$ has been removed so that $\mathcal{M}$ is not zero. Matrix elements of $\Lambda_{\beta\alpha}$ are proportional to the scattering amplitude[3] for the reactions meson α + isobar → meson β + isobar. Equations (2) are called the "mass condition" because they restrict the form of the mass operator $\mathcal{M}$. However, for some allowed mass operators they also restrict the source operators A_α.

The most general strong coupling problem is then solved by finding the (matrix) representations of (1) and (2). By (1) the A_α may be thought of as generators

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

† Associated Midwest Universities–Argonne National Laboratory Predoctoral Fellow. Present address: University of Notre Dame, Notre Dame, Indiana.

‡ Present address: Universiy of Wisconsin, Madison, Wisconsin.

[1] T. Cook, C. J. Goebel, and B. Sakita, Phys. Rev. Letters 15, 35 (1965).

[2] C. J. Goebel, T. Cook, and B. Sakita, in *Proceedings of the Third Coral Gables Conference on Symmetry Principles at High Energy* (W. H. Freeman and Company, San Francisco, 1966).

[3] C. J. Goebel, in Proceedings of the 1965 Midwest Conference on Theoretical Physics (1965).

For space-times of type I it is found that Theorem 2 can be strengthened so that the maximum orders of the groups of conformal and Killing motions are equal.

Note added in proof: It has been suggested that the metrics obtained by Newman and Tamburino[6] do not exhaust all space-times containing diverging, hypersurface orthogonal, geodesic rays (see Sec. 5). One of the authors (C. D. C.) has made an independent check of the calculations and has found that no other metrics do in fact exist. However, the metric (b) was found directly rather than as a limiting case of the metric (a) and the g_{33} component of the metric (a) itself was found to be misprinted in the paper of Newman and Tamburino. This misprint has been corrected here.

ACKNOWLEDGMENT

One of the authors (D. C. F.) wishes to express his gratitude to the Kingston upon Hull Education Committee for financial support.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Induced Representations of Strong Coupling Groups*

T. Cook†
University of Wisconsin, Madison, Wisconsin, and *Argonne National Laboratory, Argonne, Illinois*

B. Sakita‡
Argonne National Laboratory, Argonne, Illinois

(Received 27 June 1966)

Mackey's method of induced representations is applied to the strong coupling group $G = K \cdot T$, where K is compact and T is Abelian, to obtain the general irreducible representations. The form of the meson–isobar–isobar couplings is obtained by reducing these, with the use of the Peter–Weyl theorem, to irreducible representations of the compact subgroup K. The results are applied to the cases: pseudoscalar octet mesons, $K = SU_2 \otimes SU_3$, $T = T_{24}$, and SU_6 35-plet mesons, $K = SU_6$, $T = T_{35}$. Explicit representations are obtained which are consistent with mass formulas $\mathcal{M} \propto f_2(\mu) - \frac{5}{8}j(j+1)$ and $\mathcal{M} \propto G^2$ in the respective cases.

I. INTRODUCTION

In previous papers[1,2] we have exhibited the group of the strong coupling theory. The strong coupling problem is defined by the equations

$$[A_\alpha, A_\beta] = 0, \tag{1}$$

$$\Lambda_{\beta\alpha} = [A_\beta, [\mathcal{M}, A_\alpha]], \tag{2a}$$

$$\Lambda_{\beta\alpha} = \sum_\gamma \Lambda_{\beta\gamma}\Lambda_{\gamma\alpha}. \tag{2b}$$

In the above all terms are considered as operators on the isobar states. A_α is the Hermitian source operator of meson α and hence matrix elements of A_α are proportional to the meson–isobar–isobar couplings. A factor g^2 has been removed so that A_α is finite. $\mathcal{M}$ is a diagonal operator whose matrix elements are proportional to the isobar mass differences. Here a factor of $1/g^2$ has been removed so that $\mathcal{M}$ is not zero. Matrix elements of $\Lambda_{\beta\alpha}$ are proportional to the scattering amplitude[3] for the reactions meson α + isobar → meson β + isobar. Equations (2) are called the "mass condition" because they restrict the form of the mass operator $\mathcal{M}$. However, for some allowed mass operators they also restrict the source operators A_α.

The most general strong coupling problem is then solved by finding the (matrix) representations of (1) and (2). By (1) the A_α may be thought of as generators

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

† Associated Midwest Universities–Argonne National Laboratory Predoctoral Fellow. Present address: University of Notre Dame, Notre Dame, Indiana.

‡ Present address: Universiy of Wisconsin, Madison, Wisconsin.

[1] T. Cook, C. J. Goebel, and B. Sakita, Phys. Rev. Letters 15, 35 (1965).

[2] C. J. Goebel, T. Cook, and B. Sakita, in *Proceedings of the Third Coral Gables Conference on Symmetry Principles at High Energy* (W. H. Freeman and Company, San Francisco, 1966).

[3] C. J. Goebel, in Proceedings of the 1965 Midwest Conference on Theoretical Physics (1965).

of an Abelian Lie group T whose representations are restricted by (2) for a given mass operator. In interesting physical cases there exists also a group K of symmetry operators on the isobar states, which we assume is a compact group. The mass operator $\mathcal{M}$ is invariant under K. In order that the interactions be invariant under the action of this group, the source operators must transform as a tensor. That is for $k \in K$

$$k^{-1}A_\alpha k = \sum_\beta D^l_{\beta\alpha}(k)A_\beta, \tag{3}$$

where $D_{\beta\alpha}$ is the $\beta\alpha$ matrix element of the operator k in the representation, l, to which the mesons belong. By virtue of equation (3) if $t \in T$ and $k \in K$ the operator

$$t' = k^{-1}\, tk \tag{4}$$

is an element of T. Mathematically we say that (3) defines a homomorphism of the group K into the group of automorphisms of T. Thus, we may define the semidirect product of K and T

$$G = K \cdot T. \tag{5}$$

Here any $g \in G$ may be decomposed uniquely

$$g = kt,$$

$k \in K$, $t \in T$, where k and t do not commute as in a direct product but satisfy (4) instead.

In this paper we concern ourselves with the case $K = SU_2 \otimes SU_3$, $T = T_{24}$ (octet pseudoscalar meson) and the case $K = SU_6$, $T = T_{35}$. In the case of $[SU_2 \otimes SU_3] \cdot T_{24}$ Goebel[4] has considered a mass operator of the form[5]

$$\mathcal{M} = aF^2 + bJ^2 + \frac{c}{\alpha^2}\sum_{i\alpha\beta} A_{i\alpha}A_{i\beta}F_\alpha F_\beta, \tag{6}$$

where J_i and F_α are the generators of SU_2 and SU_3, respectively, and

$$\alpha^2 = \tfrac{1}{3}\sum_{i\alpha} A_{i\alpha}A_{i\alpha}.$$

Such a mass operator can be made to satisfy Eqs. (2) if we require either (1) $a = 0$, $b + \frac{1}{4}\lambda^2 c = (4\alpha^2)^{-1}$ or (2) $a = \lambda/3\alpha^2$, $b + \frac{1}{4}\lambda^2 c = -\frac{1}{4}\lambda^2(\lambda/3\alpha^2)$ or (3) $a = \lambda/3\alpha^2$, $b + \frac{1}{4}\lambda^2 c = [(3/4\lambda) - \frac{1}{4}\lambda^2](\lambda/3\alpha^2)$, where λ can be either 1 or 2, and if we require the condition

$$\sum_{ij\alpha\beta} f_{\alpha\beta\gamma}\epsilon_{ijk}A_{i\alpha}A_{j\beta} = \lambda_\alpha A_{k\gamma}, \tag{7}$$

where ϵ_{ijk} and $f_{\alpha\beta\gamma}$ are the structure constants of SU_2 and SU_3.

In the $SU_6 \cdot T_{35}$ case the operator A_α is the adjoint representation of SU_6. A mass operator of the form

$$\mathcal{M} \propto G^2, \tag{8}$$

where G_α, $\alpha = 1$–35 are the generators of SU_6 requires that the source operators satisfy

$$\sum_{\alpha\beta} d_{\alpha\beta\gamma}A_\alpha A_\beta = \lambda\alpha A_\gamma, \tag{9}$$

where $d_{\alpha\beta\gamma}$ are the analogously defined symmetric couplings for SU_6,

$$\alpha^2 = \tfrac{1}{3}\sum_\alpha A_\alpha A_\alpha$$

and λ^2 equals either 4/15, 1/12, or 0.

Thus the strong coupling problem has been reduced to the mathematical problem of finding those unitary irreducible representations of $G = K \cdot T$ which are consistant with the conditions on the A's [Eq. (7) or (9)]. Fortunately this problem has been solved by mathematicians[6] by the method of induced representation, the mathematical generalization of the procedure by which Wigner constructed the representations of the Poincaré group. An advantage of using this representation is that its irreducibility can be exhibited relatively easily owing to the Mackey irreducibility criterion.[6] However, the usual basis for an induced representation does not diagonalize the mass operator. To do this we must reduce the induced representation to irreducible representations of K, which is a symmetry group of the isobar states. This reduction can be carried out by using the Peter–Weyl theorem[7] for obtaining a basis irreducible under K. The matrix elements of the source operators with respect to such a basis give the explicit form of the meson–isobar–isobar couplings. In Sec. II we carry out this program for a general case. In Secs. III and IV we apply the results to the specific cases mentioned before.

II. INDUCED REPRESENTATIONS

The method as it applies to our case is as follows. Since T is Abelian, any irreducible representation α of T is simply a character of T

$$\alpha(t)U^a = \chi^a(t)U^a, \tag{10}$$

where U^a is the base vector of the (one-dimensional)

[4] C. J. Goebel, in *Proceedings of the Conference on Non-Compact Groups in Particle Physics* (W. A. Benjamin, Inc. New York, 1967); Phys. Rev. Letters **16**, 1130 (1966).

[5] In $SU_2 \otimes SU_3$ the latin indices $i, j, k \cdots$ run from 1 to 3 and refer to the SU_2 part; and the greek indices $\alpha, \beta, \gamma \cdots$ run from 1 to 8 and refer to the SU_3 part.

[6] G. W. Mackey, "Group Representations in Hilbert Space," Appendix to *Mathematical Problems of Relativistic Physics* by Irving E. Segal (American Mathematical Society, Providence, Rhode Island, 1963). See also R. Hermann, *Lie Groups for Physicists* (W. A. Benjamin, Inc., New York, 1966), Chap. 9.

[7] See for example, L. Pontrjagin, *Topological Groups* (Princeton University Press, Princeton, New Jersey, 1939).

representation space of α; a is an index to label the character; and the character $\chi^a(t)$ is a complex-valued function with magnitude 1 which satisfies the condition $\chi^a(t_1)\chi^a(t_2) = \chi^a(t_1t_2)$. The action of K on T [Eq. (3)] defines an action of K on the characters of T given by

$$\chi^a(k^{-1}tk) = \chi^{a(k)}(t). \tag{11}$$

The set of all characters which may be obtained by applying the elements of K to a particular character is called an orbit of K in the characters. The set of all elements of K which leaves a particular character a_0 unchanged forms a subgroup of K. This subgroup is called the "little group" K_0 associated with a_0 and its form depends on the orbit of K to which a_0 belongs.

Let β be an irreducible representation of the little group K_0.

$$\beta(k_0)U^L_M = \sum_{M'} \mathfrak{D}^L_{M'M}(k_0)U^L_{M'}, \tag{12}$$

where the U^L_M form a basis for the representation space of β, and the $\mathfrak{D}^L_{M'M}(k_0)$ are the matrix elements with respect to this basis. We can form a representation of the group $H = K_0 \cdot T$ by taking the direct product of the representations α and β.

$$\sigma(h)U^{a_0L}_M = \sum_{M'} \chi^{a_0}(t)\mathfrak{D}^L_{M'M}(k_0)U^{a_0L}_{M'}, \tag{13}$$

where h is decomposed $h = k_0t$, and $U^{a_0L}_M$ is a base vector of the direct product space given by $U^{a_0L}_M = U^{a_0} \otimes U^L_M$. Call this representation space U.

Now consider the set of functions ψ on G which take values in U and satisfy the property:

$$\psi(gh) = \sigma(h^{-1})\psi(g) \tag{14}$$

for $g \in G$, $h \in H$. These functions form a vector space V with linear combination defined by

$$(a\psi_1 + b\psi_2)(g) = a\psi_1(g) + b\psi_2(g),$$

where a and b are complex numbers, since property (14) is preserved by this linear combination. The induced representation ρ of G can be given in this vector space V by

$$\rho(g_1)\psi = \psi',$$

where

$$\psi'(g) = \psi(g_1^{-1}g). \tag{15}$$

Clearly, ψ' retains the property (14), and the action of ρ is linear. Also, $\rho(g_2)\rho(g_1) = \rho(g_2g_1)$ because

$$(\rho(g_2)\rho(g_1)\psi)(g) = (\rho(g_1)\psi)(g_2^{-1}g) = \psi(g_1^{-1}g_2^{-1}g)$$
$$= \psi[(g_2g_1)^{-1}g] = [\rho(g_2g_1)\psi](g).$$

Since α and β are irreducible, by Theorem 8.1 in Mackey,[6] the induced representation ρ is irreducible and specified by the orbit determined by a_0 and the representation L of K_0.

Now let K_1 be a subset of K chosen so that one and only one element of K_1 is in each coset of K/K_0. The set K_1 does not form a subgroup since K_0 is not normal, but every $k \in K$ may be decomposed uniquely into

$$k = k_1k_0 \quad \text{with} \quad k_1 \in K_1,\ k_0 \in K_0. \tag{16}$$

Because of the restriction [Eq. (14)] on ψ, it need only be defined on K_1. We may define an inner product in V by the formula

$$(\psi_a, \psi_b) = \int_{K_1} dk_1(\psi_a(k_1), \psi_b(k_1)), \tag{17}$$

and hence, V is a Hilbert space.

By a straightforward extension of the method for finite groups[8] a basis for V can be taken as the functions $\psi^{a_0L}_{k_1M}$ defined on K_1 by

$$\psi^{a_0L}_{k_1M}(k_1') = \delta(k_1', k_1)U^{a_0L}_M, \tag{18}$$

where $\delta(k_1', k_1)$ has the properties of the Dirac δ function, that is,

$$\int_{K_1} dk_1\, \delta(k_1', k_1)f(k_1) = f(k_1').$$

Using Eq. (14), one can extend these functions to the whole group G;

$$\psi^{a_0L}_{k_1'M}(g) = \sum_{M'} \delta(k_1', k_1)\chi^{a_0}(t^{-1})\mathfrak{D}^L_{M'M}(k_0^{-1})U^{a_0L}_{M'}, \tag{19}$$

where g is decomposed $g = k_1k_0t$. The action of G on these basis vectors can then be easily shown to be given by

$$\rho(g)\psi^{a_0L}_{k_1'M} = \sum_{M'} \int dk_1'''\, \delta(k_1', k_1'')\chi^{a_0}(k_1'''^{-1}ktk^{-1}k_1''')$$
$$\times \mathfrak{D}^L_{M'M}(k_1'''^{-1}kk_1'')\psi^{a_0L}_{k_1''M'}, \tag{20}$$

where k_1'' is given by the decomposition

$$g^{-1}k_1''' = k_1''k_0''t'' \quad \text{and} \quad g = kt.$$

Useful special cases of this are:

$$\rho(k_1)\psi^{a_0L}_{1M} = \psi^{a_0L}_{k_1M}, \tag{21}$$

$$\rho(k_0)\psi^{a_0L}_{1M} = \sum_{M'} \mathfrak{D}^L_{M'M}(k_0)\psi^{a_0L}_{1M'}, \tag{22}$$

$$\rho(t)\psi^{a_0L}_{k_1M} = \chi^{a(k_1)}(t)\psi^{a_0L}_{k_1M}. \tag{23}$$

As we have discussed in the Introduction, we wish to reduce this representation with respect to the irreducible representations of K. To accomplish this we define the vectors

$$X^{a_0L}_{\mu\nu}(\nu'M) = \int dk\ D^{\mu}_{\nu\nu'}(k)^*\rho(k)\psi^{a_0L}_{1M}, \tag{24}$$

[8] M. Iyanaga and M. Sugiura, *Algebras for Applied Mathematicians* (Iwanami-shoten, Tokyo, 1960) (in Japanese).

where D^{μ} is an irreducible representation of K. Clearly

$$\rho(k)X^{a_0L}_{\mu\nu}(\nu''M) = \sum_{\nu'} D^{\mu}_{\nu'\nu}(k)X^{a_0L}_{\mu\nu'}(\nu''M). \quad (25)$$

Thus the vectors $X^{a_0L}_{\mu\nu}(\nu'M)$ transform irreducibly under the action of the subgroup K. However, they do not form a basis because they are in general not linearly independent. By the Peter–Weyl theorem,[7]

$$\rho(k)\psi^{a_0L}_{1M} = \sum_{\mu\nu\nu'} N(\mu)\, D^{\mu}_{\nu\nu'}(k)X^{a_0L}_{\mu\nu}(\nu'M),$$

where $N(\mu)$ is the dimensionality of the representation μ. Now, using Eqs. (21) and (22),

$$\rho(k)\psi^{a_0L}_{1M} = \sum_{M'} \mathfrak{D}^{L}_{M'M}(k_0)\psi^{a_0L}_{k_1M'},$$

where k is decomposed $k = k_1k_0$. If we multiply the right-hand sides of these two equations by $\mathfrak{D}^{L}_{M''M'}(k_0)$ and integrate over K_0 we obtain

$$\frac{1}{N(L)}\delta_{LL'}\delta_{MM'}\psi^{a_0L}_{k_1M''} = \sum_{\mu\nu\nu''} N(\mu)D^{\mu}_{\nu\nu''}(k_1)\sum_{\nu'} \times\left[\int dk_0 D^{\mu}_{\nu''\nu'}(k_0)\mathfrak{D}^{L}_{M''M'}(k_0)^*\right]X^{a_0L}_{\mu\nu}(\nu'M). \quad (26)$$

To evaluate the integral of Eq. (26) we must investigate the way in which K_0 is contained in K. The components of an irreducible representation μ of K have been specified by ν. Since K_0 is a subgroup of K the index ν may be written (ξLM), where L, M specifies the base of an irreducible representation of K_0 and ξ specifies irreducible representations of a chain of subgroups of K which contain K_0, i.e., $K \supset K' \supset K'' \cdots \supset K_0$. Then the matrix $D^{\mu}_{\nu'\nu}(k)$ restricted to K_0 can be written

$$D^{\mu}_{(\xi'L'M')(\xi LM)}(k_0) = \delta_{\xi'\xi}\delta_{L'L}\mathfrak{D}^{L}_{M'M}(k_0). \quad (27)$$

Thus we can do the integral of Eq. (26) and obtain

$$\psi^{a_0L}_{k_1M} = \sum_{\mu\nu\xi}\left[\frac{N(\mu)}{N(L)}\right]^{\frac{1}{2}} D^{\mu}_{\nu(\xi LM)}(k_1)\phi^{a_0L}_{\mu\nu\xi}, \quad (28)$$

where

$$\phi^{a_0L}_{\mu\nu\xi} = [N(\mu)N(L)]^{\frac{1}{2}}X^{a_0L}_{\mu\nu}[(\xi LM)M], \quad (29)$$

and that $X^{a_0L}_{\mu\nu}[(\xi LM)M]$ is independent of M. From the definition of X [Eq. (24)] we have

$$\phi^{a_0L}_{\mu\nu\xi} = \sum_{M}\int dk_1\left[\frac{N(\mu)}{N(L)}\right]^{\frac{1}{2}} D^{\mu}_{\nu(\xi LM)}(k_1)^*\psi^{a_0L}_{k_1M}. \quad (30)$$

One may derive the expression

$$\int_{K_0} dk_0 D^{\mu}_{\nu\xi LM)}(k)^* D^{\mu'}_{\nu'(\xi'L'M')}(k) = \frac{1}{N(L)}\delta_{LL'}\delta_{MM'}\sum_{M''} D^{\mu}_{\nu(\xi LM'')}(k_1)^* D^{\mu'}_{\nu'(\xi'LM'')}(k_1). \quad (31)$$

Using this it is easy to show that the $\phi^{a_0L}_{\mu\nu\xi}$ are orthonormal so that they form a basis for the irreducible representation (a_0L) of G, which exhibits explicitly the reduction into irreducible representations μ of K. Considering carefully the nonzero coefficients of Eq. (30) we see that the representation (a_0L) contains those representations, μ, which have a component (ξLM) and with multiplicity given by the number of different values that ξ may take on for the same L.

Now, to obtain the explicit form of the physical couplings, we wish to construct the matrix elements of the source operators in the above representation. The action of A_α on the ψ's is implied by Eq. (23)

$$A_\alpha\psi^{a_0L}_{k_1M} = a_\alpha(k_1)\psi^{a_0L}_{k_1M}.$$

Also $a(k_1)$ may be expressed in terms of a_0 by means of Eqs. (3) and (11)

$$a_\alpha(k_1) = \sum_{\beta} D^{l}_{\alpha\beta}(k_1)^*(a_0)_\beta.$$

Using the expansion coefficients of Eq. (30) and the above expressions we may write:

$$(\phi^{a_0L}_{\mu'\nu'\xi'}, A_\alpha\phi^{a_0L}_{\mu\nu\xi}) = \sum_{M,\beta}\int_{K_1} dk_1\left[\frac{N(\mu')}{N(L)}\right]^{\frac{1}{2}} D^{\mu'}_{\nu'(\xi'LM)}(k_1) \times D^{l}_{\alpha\beta}(k_1)^*(a_0)_\beta\left[\frac{N(\mu)}{N(L)}\right]^{\frac{1}{2}} D^{\mu}_{\nu(\xi LM)}(k_1)^*.$$

With the use of Eq. (31) and the definition of the little group the region of integration may be extended to all of K. Then using the well-known results that the integral of three D's is two "3-j" symbols we have

$$(\phi^{a_0L}_{\mu'\nu'\xi'}, A_\alpha\phi^{a_0L}_{\mu\nu\xi}) = [N(\mu')N(\mu)]^{\frac{1}{2}}\sigma[\mu'\nu'(\xi'LM)]\sum_{\gamma}\begin{pmatrix}\bar{\mu}' & l & \mu\\ \bar{\nu}' & \alpha & \nu\end{pmatrix}_\gamma \times\sum_{\beta}(a_0)_\beta\begin{pmatrix}\bar{\mu}' & l & \mu\\ (\xi'LM) & \beta & (\xi LM)\end{pmatrix}_\gamma, \quad (32)$$

where

$$\begin{pmatrix}\cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot\end{pmatrix}_\gamma$$

is a "3-j" symbol [symmetrized Wigner coupling coefficient] for K. Here γ is used to distinguish different ways of coupling the same representations μ', l, μ. The phase σ is given by

$$D^{\mu}_{\nu\nu'}(k) = \sigma(\mu\nu\nu')D^{\bar{\mu}}_{\bar{\nu}\bar{\nu}'}(k)^*.$$

The expression is independent of M. This is not obvious since it comes from the nature of $(a_0)_\beta$.

III. $[SU_2 \otimes SU_3] \cdot T_{24}$

In this case, the characters of T_{24} are specified by 24 numbers. Thus, the index a should be taken to be

a 3 × 8 matrix whose components $a_{i\alpha}$ are the values of the generators $A_{i\alpha}$ in the representation given by Eq. (10). The action of K on the characters is given by

$$a(k)_{i\alpha} = \sum_{i'\alpha'} D^{1,8}_{i\alpha,i'\alpha'}(k)^*(a_0)_{i'\alpha'}. \tag{33}$$

Here 1, 8 stands for spin 1 (p-wave), octet. If now we make the restriction of Eq. (7) on the generators, we restrict the characters by the condition

$$\sum_{ij\alpha\beta} f_{\alpha\beta\gamma}\epsilon_{ijk}a_{i\alpha}a_{j\beta} = \lambda\alpha a_{k\gamma}. \tag{34}$$

This condition restricts the character to two orbits of K which are distinguished by the two values of λ. To show this we consider the 3 × 3 matrices

$$a_i = \sum_\alpha \lambda_\alpha a_{i\alpha},$$

where the λ_α are the usual Gell-Mann matrices. Condition (34) is then equivalent to the condition

$$[a_i, a_j] = i\lambda\alpha \sum_k \epsilon_{ijk}a_k.$$

Thus the matrices $(1/\lambda\alpha)a_i$ form a three-dimensional representation of the Lie algebra of SU_2. It is well known that there are only two Hermitian, nontrivial, nonequivalent, not necessarily irreducible three-dimensional representations; i.e., $2 \oplus 1$ and 3. Equivalent representations correspond to points of the same orbit since a unitary transformation on the a_i matrices is the same as the action of the corresponding element of SU_3 on the $a_{i\alpha}$. Thus, we have two orbits and can choose a_0 to be given by either

$$(a_0)_i = \lambda\alpha\tfrac{1}{2}\lambda_i \qquad (i = 1, 2, 3)$$

or

$$(a_0)_1 = \lambda\alpha\lambda_2, \quad (a_0)_2 = \lambda\alpha\lambda_5, \quad (a_0)_3 = \lambda\alpha\lambda_7.$$

Consideration of $\sum_i \mathrm{Tr}\,(a_i a_i)$ shows that in the former case $\lambda = 2$ and in the latter $\lambda = 1$, so the two orbits may be distinguished by λ. Finally we have for $\lambda = 2$

$$(a_0)_{i\alpha} = \alpha\delta_{i\alpha} \tag{35}$$

and for $\lambda = 1$, $(a_0)_{12} = \alpha$, $(a_0)_{25} = \alpha$, $(a_0)_{37} = \alpha$, and all the rest are zero. In this paper we consider only the case $\lambda = 2$.

By careful consideration of the Lie algebras[9] involved we can show that the little group in this case is generated by

$$v_i = J_i + I_i, \quad \text{where} \quad I_i = F_i, \quad i = 1 - 3,$$

and

$$Y = (2/\sqrt{3})F_8. \tag{36}$$

Let us note it by $(SU_2)_v \otimes (U_1)_Y$, where the subscripts refer to the generators. A sufficient chain to specify ξ is

$$(SU_2)_J \otimes (SU_3)_F \supset (SU_2)_J \otimes (SU_2)_I \otimes (U_1)_Y \supset (SU_2)_v \otimes (U_1)_Y.$$

Only the invariant I_0 of $(SU_2)_I$ is not otherwise specified and it may be used for ξ.

In this notation Eq. (30) becomes

$$\phi^{\alpha v Y_0}_{j\mu m\nu I_0} = \sum_{v_3} \int dk_1 \left[\frac{N(j)N(\mu)}{N(v)}\right]^{\frac{1}{2}} \times D^{j\mu}_{m\nu(I_0 v Y_0 v_3)}(k_1)^* \psi^{\alpha v Y_0}_{k_1 v_3}. \tag{30'}$$

Here j, m refer to a representation of SU_2 and μ, ν to SU_3. If we recouple v, v_3 to j and I_0 using Eqs. (36) Eq. (32) becomes[10]

$$\begin{aligned}(\phi^{\alpha v Y_0}_{j'\mu' m'\nu' I_0'}, A_{i\alpha}\phi^{\alpha v Y_0}_{j\mu m\nu I_0}) &= [(2j'+1)(2j+1)]^{\frac{1}{2}}[N(\mu')N(\mu)]^{\frac{1}{2}} \\ &\times \sum_\gamma \begin{pmatrix} j' & 1 & j \\ -m' & i & m \end{pmatrix}\begin{pmatrix} \bar{\mu}' & 8 & \mu \\ \bar{\nu}' & \alpha & \nu \end{pmatrix}_\gamma \\ &\times (-)^{(j-j'+I_0-I_0'+v)}(-)^{(m'+I_3')} \times \sigma(I', Y', I_0', Y_0) \\ &\times \alpha\begin{bmatrix} \bar{\mu}' & 8 & \mu \\ (I_0, -Y_0) & (1,0) & (I_0, Y_0) \end{bmatrix}_\gamma \begin{Bmatrix} j' & I_0' & v \\ I_0 & j & 1 \end{Bmatrix},\end{aligned} \tag{32'}$$

where the index ν has been decomposed in the usual way into I, I_3, Y;

$$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$$

is a "3-j" symbol for either SU_2 or SU_3;

$$\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$$

is an isoscalar factor for SU_3;

$$\begin{Bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{Bmatrix}$$

is a "6-j" symbol for SU_2; and σ is a phase defined by

$$D^{\mu}_{\nu'\nu}(k) = (-)^{I_3 - I_3'}\sigma(I', Y', I, Y)D^{\bar{\mu}}_{\bar{\nu}'\bar{\nu}}(k)^*.$$

[9] If we consider the commutators of the $A_{i\alpha}$ with v_i ($i = 1, 2, 3$) and Y evaluated at the point defined by Eq. (35) we see that the algebra of the little group contains that with basis v_i and Y, since $[A_{i\alpha}, v_j]|_{A_{i\alpha}=\alpha\delta_{i\alpha}} = 0$ and $[A_{i\alpha}, Y]|_{A_{i\alpha}=\alpha\delta_{i\alpha}} = 0$. The orthogonal complement of this algebra in the algebra of $SU_2 \otimes SU_3$ reduces into three components whose bases form irreducible tensors under the action of this algebra. These are $(J_i - I_i$, $i = 1, 2, 3)$; $(F_4 + iF_5, F_6 + iF_7)$; and $(F_4 - iF_5, F_6 - iF_7)$. If each of these is adjoined to the above algebra they produced an algebra which is larger than that of the little group, since $[A_{13}, J_2]|_{A_{i\alpha}=\alpha\delta_{i\alpha}} = i\alpha$ and $[A_{17}, F_4 \pm iF_5]|_{A_{i\alpha}=\alpha\delta_{i\alpha}} = -i\frac{1}{2}\alpha$. Thus, assuming the little group is a Lie group we can conclude that the little group is generated by v_i and Y.

[10] Here we have slipped into spherical notation for the indices i, α. In this notation $(a_0)_{i\alpha} = \alpha(-)^\alpha\delta_{i-\alpha}$, where the index α refers to the I_3 value for the $I = 1$ part of the octet, $(a_0)_{i\alpha} = 0$ for the other components of the octet.

The expression for the mass, Eq. (6), becomes in this representation for case 3

$$\mathcal{M} = (2/3\alpha^2)\{f_2(\mu) - \tfrac{5}{8}[a'j(j+1) + b'I_0(I_0+1)]\}, \tag{37}$$

where $a' + b' = 1$, $a' - b'$ is arbitrary, and $f_2(\mu)$ is the value of the second-order Casimir operator of SU_3 in the representation μ.

In the case $v = 0; j = I_0$ and the mass is

$$\mathcal{M} = (2/3\alpha^2)[f_2(\mu) - \tfrac{5}{8}j(j+1)]. \tag{38}$$

This mass formula is explored in detail in the reference of Ref. 4.

IV. $SU_6 \cdot T_{35}$

The characters of T_{35} are specified by 35 numbers or a 6×6 traceless matrix a whose components a_{ij} are the values of the operators $\sum_\alpha D^6_{ij}(G_\alpha)A_\alpha$ in the representation of Eq. (10). Any such matrix a may be diagonalized by the action of K on the characters. Thus the orbits of K in the character space are specified by the eigenvalues of a. This is a great simplification over the case of $[SU_2 \otimes SU_3] \cdot T_{24}$ since we now know the most general orbit. Thus a_0 may be taken in general to be

$$(a_0)_{ij} = \delta_{ij}a_j. \tag{39}$$

The restriction (9) on the generators requires that a have only two distinct eigenvalues. There are three possible a_0 which satisfy this condition. These three cases, along with the associated little group, are given below.

Case (a)

$$\lambda^2 = 4/15,$$
$$a_1 = (5/6)^{\frac{1}{2}}\alpha,$$
$$a_2 = a_3 = a_4 = a_5 = a_6 = -\tfrac{1}{5}(5/6)^{\frac{1}{2}}\alpha.$$

The little group is the $U_1 \otimes SU_5$ subgroup generated by

$$X_{11};\ X_{ij} - \tfrac{1}{5}\delta_{ij}\sum_{k=2}^{6} X_{kk}, \quad i,j = 2-6,$$

where the generators of SU_6 are written

$$X_{ij} = \sum_\alpha D^6_{ij}(G_\alpha)G_\alpha.$$

Case (b)

$$\lambda^2 = \tfrac{1}{12},$$
$$a_1 = a_2 = (\tfrac{1}{3})^{\frac{1}{2}}\alpha,$$
$$a_3 = a_4 = a_5 = a_6 = -\tfrac{1}{2}(\tfrac{1}{3})^{\frac{1}{2}}\alpha.$$

The little group is the $U_2 \otimes SU_4$ subgroup generated by

$$X_{ij}, \quad i,j = 1,2; \quad X_{ij} - \tfrac{1}{4}\delta_j\sum_{k=3}^{6} X_{kk}, \quad i,j = 3-6.$$

Case (c)

$$\lambda = 0,$$
$$a_1 = a_2 = a_3 = (\tfrac{1}{6})^{\frac{1}{2}}\alpha,$$
$$a_4 = a_5 = a_6 = -(\tfrac{1}{6})^{\frac{1}{2}}\alpha.$$

The little group is the $U_3 \otimes SU_3$ subgroup generated by

$$X_{ij}, \quad i,j = 1-3; \quad X_{ij} - \tfrac{1}{3}\delta_{ij}\sum_{k=4}^{6} X_{kk}, \quad i,j = 4-6.$$

We consider only Case (a). The induced representation space is spanned by

$$\psi^{\alpha\rho s}_{k_1\sigma},$$

where ρ is an irreducible representation of SU_5 with components indexed by σ and s specifies the representation of U_1. In this case, the little group is a maximal subgroup, and hence the index ξ does not appear. Equations (30) and (32) become

$$\phi^{\alpha\rho s}_{\mu\nu} = \sum_\sigma \int dk_1 \left[\frac{N(\mu)}{N(\rho)}\right]^{\frac{1}{2}} D^\mu_{\nu(\rho s\sigma)}(k_1)\psi^{\alpha\rho s}_{k_1\sigma} \tag{30''}$$

and

$$(\phi^{\alpha\rho s}_{\mu'\nu'}, A_\alpha\phi^{\alpha\rho s}_{\mu\nu}) = [N(\mu')N(\mu)]^{\frac{1}{2}}\sigma[\mu'\nu'(\rho s\sigma)] \times \sum_\gamma \begin{pmatrix} \bar{\mu}' & 35 & \mu \\ \bar{\nu}' & \alpha & \nu \end{pmatrix}_\gamma \times \sqrt{3}\,\alpha \begin{pmatrix} \bar{\mu}' & 35 & \mu \\ \overline{(\rho s\sigma)} & (100) & (\rho s\sigma) \end{pmatrix}_\gamma. \tag{32''}$$

If we use the notation of Young's diagrams to specify the representations of SU_5 and SU_6 [example $(\lambda_1\lambda_2\lambda_3\lambda_4)$ and $(\lambda_1\lambda_2\lambda_3\lambda_4\lambda_5)$, where $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$] we can show that the only representations which contain the 56 representations of SU_6 are

(a)	$\rho = (0000)$,	$s = -15/2$,
(b)	(1000),	$-9/2$,
(c)	(2000),	$-3/2$,
(d)	(3000),	$3/2$.

Of these only (a) has no state of mass less than the mass of the 56. This representation gives the spectrum

μ	$N(\mu)$	$2G^2$
(30000)	56	45
(51111)	700	57
(72222)	4536	125
(93333)	20580	177

ACKNOWLEDGMENTS

The authors thank Professor C. J. Goebel for stimulating discussions. He had obtained many of the results contained in this paper in a slightly different way.

Symmetry Properties of the $3j$ Symbols for $SU(3)$

JEAN-ROBERT DEROME*†
Laboratoire de Physique Théorique et Hautes Energies, Orsay, France

(Received 10 June 1966)

For an arbitrary compact group it is in general not possible to choose the $3j$ symbol $(j_1j_2j_3)_{r,m_1m_2m_3}$ such that its absolute value is invariant under every permutation of the j's and of the corresponding m's. Still, it is commonly assumed that for $SU(3)$ such a choice is possible. In this paper it is shown that this assumption is indeed justified.

1. INTRODUCTION

FOR the group $SU(3)$ it is often stated that the $3j$ symbol $(j_1j_2j_3)_{r,m_1m_2m_3}$ can be chosen such that its absolute value is invariant under every permutation of the j's and of the corresponding m's.[1] However, this statement requires proof for every particular group since it has been shown to be false for S_6 (the symmetric group on six symbols).[2] The purpose of this paper is to show that for $SU(3)$ the statement is indeed correct.

We first recall the definition of the $3j$ symbol $(j_1j_2j_3)_{r,m_1m_2m_3}$ for an arbitrary compact group[3] G

$$j_1(R)^{m_1}{}_{m_1'}j_2(R)^{m_2}{}_{m_2'} = \sum_{j_3}[j_3]\{(j_1j_2j_3)_{r,m_1m_2m_3}\}^*j_3(R)_{m_3}{}^{m_3'}(j_1j_2j_3)_{r,m_1'm_2'm_3'}, \tag{1}$$

where a sum is implied over repeated m and r indices. Here $[j_3]$ is the dimension of j_3, $j(R)^m{}_{m'} = \{j(R)_m{}^{m'}\}^*$ are the matrix elements of the irreducible representation j of G, and $R \in G$. The index m is a generalized magnetic quantum number and r is a multiplicity index which takes on as many values as the multiplicity of j_3^* in the Kronecker product $j_1 \times j_2$. Using the orthogonality relations for the irreducible representations, one obtains[4]

$$\int j_1(R)^{m_1}{}_{m_1'}j_2(R)^{m_2}{}_{m_2'}j_3(R)^{m_3}{}_{m_3'}\,dR = \{(j_1j_2j_3)_{r,m_1m_2m_3}\}^*(j_1j_2j_3)_{r,m_1'm_2'm_3'}, \tag{2}$$

where the integral has been normalized to $\int dR = 1$. Also, from the definition of the $3j$ symbols and from their unitarity it follows that

$$(j_1j_2j_3)_{r,m_1m_2m_3}j_1(R)^{m_1}{}_{m_1'}j_2(R)^{m_2}{}_{m_2'}j_3(R)^{m_3}{}_{m_3'} = (j_1j_2j_3)_{r,m_1'm_2'm_3'}. \tag{3}$$

For the sake of brevity, we say that the $3j$ symbol $(j_1j_2j_3)_{r,m_1m_2m_3}$ can be chosen in symmetric form or can be symmetrized if

$$|(j_1j_2j_3)_{r,m_1m_2m_3}| = |[\pi(j_1j_2j_3)]_{r,\pi(m_1m_2m_3)}|, \tag{4}$$

where $\pi(j_1j_2j_3)$ can be any permutation of $j_1j_2j_3$ and $\pi(m_1m_2m_3)$ is the same permutation of $m_1m_2m_3$, that is if Eq. (4) holds for every π in S_3, the symmetric group on three symbols. In a previous paper[2] we have shown that if the $3j$ symbol $(j_1j_2j_3)_{r,m_1m_2m_3}$ cannot be chosen in symmetric form, then $j_1j_2j_3$ are equivalent representations. Thus, in this paper, only $3j$ symbols of the type $(jjj)_{r,m_1m_2m_3}$ need be considered.

2. SYMMETRIZATION CONDITION IN TERMS OF CHARACTERS

It follows from Eqs. (2) and (3) that the unitary matrix

$$M(\pi)_r^{r'} = \{(jjj)_{r',m_1m_2m_3}\}^*(jjj)_{r,\pi(m_1m_2m_3)} \tag{5}$$

satisfies the following equation:

$$(jjj)_{r,\pi(m_1m_2m_3)} = M(\pi)_r^{r'}(jjj)_{r',m_1m_2m_3}. \tag{6}$$

From Eq. (6) one sees that M is a representation of S_3 and that the $3j$ symbol $(jjj)_{r,m_1m_2m_3}$ can be chosen in symmetric form if and only if the matrices $M(\pi)$ can be diagonalized simultaneously. Now, the arbitrariness in the definition of the $3j$ symbol $(jjj)_{r,m_1m_2m_3}$ implies that the matrices $M(\pi)$ are determined only up to the similarity transformation

$$M(\pi)' = UM(\pi)U^{-1},$$

where the unitary matrix U must be the same for every $\pi \in S_3$.[2] Thus the $3j$ symbol $(jjj)_{r,m_1m_2m_3}$ can be symmetrized if and only if [21] (the only irreducible representation of S_3 of dimension larger than one) is not a constituent of M, that is, if and only if the multiplicity of [21] in M is zero. The multiplicity of [21] in M is given by

$$m(M,[21]) = \frac{1}{3!}\sum_{\pi\in S_3}\chi^{[21]}(\pi)\chi^M(\pi),$$

* Post-doctorate fellow of the National Research Council of Canada.

† Present address: Département de physique, Université de Montréal, Montréal, P. Q. Canada.

[1] J. J. de Swart, Rev. Mod. Phys. **35**, 916 (1963).

[2] J. R. Derome, J. Math. Phys. **7**, 612 (1966).

[3] J. R. Derome and W. T. Sharp, J. Math. Phys. **6**, 1584 (1965).

[4] W. T. Sharp, Atomic Energy of Canada, Ltd., Report AECL-1098 (1960).

where $\chi^M(\pi)$ = trace $M(\pi)$ and $\chi^{[21]}(\pi)$ is the character of [21]. One obtains

$$m(M, [21]) = \tfrac{1}{3}\{m(j \times j, j^*) - \chi^M[(3)]\},$$

where $m(j \times j, j^*) = \int [\chi^j(R)]^3\, dR$ is the multiplicity of j^* in $j \times j$, and where "(3)" denotes the class of S_3 consisting of the two even permutations (123) and (213). Explicitly $\chi^M[(3)]$ is given by

$$\begin{aligned}\chi^M[(3)] &= M(123)_r^r = M(213)_r^r \\ &= \{(jjj)_{r,m_1m_2m_3}\}^*(jjj)_{r,m_2m_3m_1} \\ &= \int j(R)^{m_1}{}_{m_2} j(R)^{m_2}{}_{m_3} j(R)^{m_3}{}_{m_1}\, dR \\ &= \int \chi^j(R^3)\, dR.\end{aligned}$$

Thus

$$m(M, [21]) = \tfrac{1}{3}\left\{m(j \times j, j^*) - \int \chi^j(R^3)\, dR\right\}, \quad (7)$$

and it is possible to choose the 3j symbol $(jjj)_{r,m_1m_2m_3}$ in symmetric form if and only if

$$m(j \times j, j^*) = \int \chi^j(R^3)\, dR. \quad (8)$$

We have already shown[2] that, for the irreducible representation of dimension 16 of S_6, Eq. (8) is not satisfied. We now show that Eq. (8) holds for every irreducible representation of $SU(3)$.

3. SYMMETRIZATION OF THE 3j SYMBOLS FOR $SU(3)$

Let $[\lambda\mu]$ stand for the irreducible representation of $SU(3)$ which corresponds to the Young diagram of $(\lambda + \mu)$ boxes in the first row and μ boxes in the second row so that the complex conjugate of $[\lambda\mu]$ is $[\mu\lambda]$. The multiplicity of $[\mu\lambda]$ in $[\lambda\mu] \times [\lambda\mu]$ is easily found to be $[1 + \min(\lambda, \mu)]$.[5] Thus

$$m([\lambda, \mu] \times [\lambda, \mu], [\lambda, \mu]^*) = 1 + \min(\lambda, \mu).$$

To evaluate the integral $\int \chi^{[\nu,\mu]}(R^3)\, dR$ we use the following expression for the characters of $SU(3)$

$$\chi^{[\lambda,\mu]}(\epsilon_1\epsilon_2) = \sum_{r=0}^{\lambda}\sum_{s=0}^{\mu}\sum_{t=0}^{\mu+r-s} \epsilon_1^{\lambda+\mu-r-2s-t}\epsilon_2^{\mu+r-s-2t},$$

where $\epsilon_1\epsilon_2$ are two of the eigenvalues of the group element R. If R is in the class labeled by $\epsilon_1\epsilon_2$ then R^3 is in the class labeled by $\epsilon_1^3\epsilon_2^3$ so that

$$\begin{aligned}\chi^{[\lambda,\mu]}(R^3) &= \chi^{[\lambda,\mu]}(\epsilon_1^3\epsilon_2^3) \\ &= \sum_{r=0}^{\lambda}\sum_{s=0}^{\mu}\sum_{t=0}^{\mu+r-s} \epsilon_1^{3(\lambda+\mu-r-2s-t)}\epsilon_2^{3(\mu+r-s-2t)}.\end{aligned}$$

[5] B. Preziosi, A. Simoni, and B. Vitale, Nuovo Cimento **34**, 1011 (1964).

Using the normalized invariant measure for $SU(3)$

$$dR = (1/3\pi^2)[1 - \cos(\alpha - \beta)][1 - \cos(2\alpha + \beta)] \times [1 - \cos(2\beta + \alpha)]\, d\alpha\, d\beta,$$

where $\alpha = \arg\epsilon_1$, $\beta = \arg\epsilon_2$, and $-\pi \leqslant \alpha, \beta \leqslant \pi$, one obtains

$$\begin{aligned}&\int \chi^{[\lambda,\mu]}(R^3)\, dR \\ &\quad = \frac{1}{3\pi^2}\sum_{r=0}^{\lambda}\sum_{s=0}^{\mu}\sum_{t=0}^{\mu+r-s}\int e^{3im\alpha}e^{3in\beta}[1 - \cos(\alpha - \beta)] \\ &\qquad \times [1 - \cos(2\alpha + \beta)][1 - \cos(2\beta + \alpha)]\, d\alpha\, d\beta, \end{aligned} \quad (9)$$

where $m = \lambda + \mu - r - 2s - t$ and $n = \mu + r - s - 2t$. In view of the fact that

$$\sum_{r=0}^{\lambda}\sum_{s=0}^{\mu}\sum_{t=0}^{\mu+r-s}\int e^{3im\alpha}e^{3in\beta}e^{-iP\alpha}e^{-iQ\beta}\, d\alpha\, d\beta$$

may differ from zero only when both P and Q are multiples of 3, the right-hand side of Eq. (9) becomes

$$\begin{aligned}&\frac{1}{4\pi^2}\sum_{r=0}^{\lambda}\sum_{s=0}^{\mu}\sum_{t=0}^{\mu+r-s}\int e^{3im\alpha}e^{3in\beta} \\ &\quad \times \{1 + \tfrac{2}{3}[\cos 3(\alpha + \beta) + \cos 3\alpha + \cos 3\beta]\}\, d\alpha\, d\beta \\ &\qquad = \sum_{r=0}^{\lambda}\sum_{s=0}^{\mu}\sum_{t=0}^{\mu+r-s} F(m, n),\end{aligned}$$

where

$$\begin{aligned}F(m, n) &= \delta(m)\delta(n) + \tfrac{1}{3}\{\delta(m + 1)\delta(n + 1) \\ &\quad + \delta(m)\delta(n - 1) + \delta(m - 1)\delta(n)\} + \tfrac{1}{3}\{\delta(m - 1) \\ &\quad \times \delta(n - 1) + \delta(m)\delta(n + 1) + \delta(m + 1)\delta(n)\} \end{aligned} \quad (10)$$

and

$$\begin{aligned}\delta(x) &= 0, \quad \text{if} \quad x \neq 0, \\ &= 1, \quad \text{if} \quad x = 0.\end{aligned}$$

Since $m + n = \lambda + 2\mu - 3(s + t) = (\lambda - \mu) \bmod 3$, the sum

$$G(p, q) = \sum_{r=0}^{\lambda}\sum_{s=0}^{\mu}\sum_{t=0}^{\mu+r-s}\delta(m - p)\delta(n - q)$$

is nonzero only if $(p + q) = (\lambda - \mu) \bmod 3$. If $\lambda = \mu \bmod 3$ only the first term on the right-hand side of Eq. (10) contributes to the sum. If $(\lambda - \mu) = 1 \bmod 3$ only the next three terms and if $(\lambda - \mu) = 2 \bmod 3$ only the last three terms contribute.

Suppose $(\lambda - \mu) = (p + q) \bmod 3$ and $\mu \geqslant \lambda$. $G(p, q)$ is then equal to the number of terms in the sum for which $m = p$ and $n = q$. Since our values of p and q always satisfy the inequalities

$$-1 \leqslant p, q \leqslant 1 \quad \text{and} \quad pq \geqslant 0,$$

it follows that

$$0 \leqslant s = \tfrac{1}{3}(2\lambda + \mu - 2p + q) - r \leqslant \mu,$$
$$0 \leqslant t = \tfrac{1}{3}[\mu - \lambda + p - 2q] + r \leqslant \mu + r - s$$
$$= \tfrac{1}{3}[2(\mu - \lambda) + 2p - q] + 2r$$

for every value of r in $0 \leqslant r \leqslant \lambda$. Thus in this case

$$G(p, q) = 1 + \lambda.$$

If $\lambda \geqslant \mu$, one deduces that

$$0 \leqslant r = \tfrac{1}{3}(2\lambda + \mu - 2p + q) - s \leqslant \lambda,$$
$$0 \leqslant t = \tfrac{1}{3}(\lambda + 2\mu - p - q) - s \leqslant \mu + r - s$$
$$= \mu + \tfrac{1}{3}(2\lambda + \mu - 2p + q) - 2s$$

for every value of s in $0 \leqslant s \leqslant \mu$ and

$$G(p, q) = 1 + \mu.$$

It follows that

$$\int \chi^{[\lambda,\mu]}(R^3)\, dR = \sum_{r=0}^{\lambda} \sum_{s=0}^{\mu} \sum_{t=0}^{\mu+r-s} F(m, n) = 1 + \min(\lambda, \mu);$$

that is, Eq. (8) is satisfied for every irreducible representation of $SU(3)$ and the $3j$ symbols of $SU(3)$ can all be chosen in symmetric form.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Bootstrap Prediction of Symmetry for a Soluble Static Model

ANDREW A. CUNNINGHAM
Department of Mathematical Physics, The University of Adelaide, South Australia

(Received 18 June 1966)

The Huang–Low bootstrap criterion of self-consistency, in the form of Levinson's theorem, is imposed on the exact solutions of the two-channel Low equation with an arbitrary crossing matrix. It is found that this condition, together with a number of dynamical conditions are sufficient to restrict the continuous crossing-matrix parameter to discrete values corresponding to an SU_2 symmetry.

I. INTRODUCTION

In recent years, much interest has been displayed in possible bootstrap predictions of internal symmetry. Studies have been made using an exact algebraic approach,[1] or approximate dynamical methods.[2] In work of the latter kind, a number of classified particles is given, and then, by insisting on a self-consistent solution, it is hoped that internal symmetries may be predicted. It would appear much better to avoid any approximations,[3] and to obtain an exact solution for the S-matrix elements defined by as general a set of postulates as possible, and then to impose a criterion of self-consistency. This is the approach of this paper. The postulates referred to are taken to be analyticity, unitarity, and crossing symmetry, each of which is considered quite fundamental to any description of interactions. The role of crossing symmetry in particular has been greatly stressed by Wigner.[4] It remains, of course, highly desirable that the number of dynamical assumptions should be as restricted as possible, and as general as possible. It is hoped to show how such a set of assumptions, together with the imposition of a bootstrap criterion on an exactly soluble two-channel static model will lead to a prediction of internal symmetry. The model to be considered is particularly suitable for such a study, as it incorporates a general formulation of crossing symmetry.

This model, with a crossing matrix corresponding to the scattering of a particle of isospin 1 by a baryon of isospin $\frac{1}{2}$, was exactly solved by Wanders.[5] Martin and McGlinn[6] generalized the crossing matrix for the model to include a continuous parameter, and, using an extension of the technique given by Wanders, obtained exact solutions for the two S-matrix elements. They had hoped that the existence of solutions satisfying analyticity, unitarity, and crossing symmetry would be limited to the restricted values of the crossing matrix parameter defining SU_2 symmetry. This, however, proved not to be the case, and further self-consistency and dynamical conditions are apparently needed before such a symmetry prediction can be made.

The self-consistency condition to be imposed in this

[1] R. Cutkosky, Bull. Am. Phys. Soc. **8**, 591 (1963); Phys. Rev. **131**, 1888 (1963).
[2] R. Capps, Phys. Rev. Letters **10**, 312 (1963); J. Sakurai, *ibid.* **10**, 446 (1963).
[3] B. Diu and H. Rubinstein, Nuovo Cimento **32**, 1103 (1964).
[4] E. P. Wigner, Phys. Today **17**, 34 (1964).
[5] G. Wanders, Nuovo Cimento **23**, 817 (1962).
[6] A. Martin and W. McGlinn, Phys. Rev. **136**, 1515 (1964).

it follows that

$$0 \leqslant s = \tfrac{1}{3}(2\lambda + \mu - 2p + q) - r \leqslant \mu,$$
$$0 \leqslant t = \tfrac{1}{3}[\mu - \lambda + p - 2q] + r \leqslant \mu + r - s$$
$$= \tfrac{1}{3}[2(\mu - \lambda) + 2p - q] + 2r$$

for every value of r in $0 \leqslant r \leqslant \lambda$. Thus in this case

$$G(p, q) = 1 + \lambda.$$

If $\lambda \geqslant \mu$, one deduces that

$$0 \leqslant r = \tfrac{1}{3}(2\lambda + \mu - 2p + q) - s \leqslant \lambda,$$
$$0 \leqslant t = \tfrac{1}{3}(\lambda + 2\mu - p - q) - s \leqslant \mu + r - s$$
$$= \mu + \tfrac{1}{3}(2\lambda + \mu - 2p + q) - 2s$$

for every value of s in $0 \leqslant s \leqslant \mu$ and

$$G(p, q) = 1 + \mu.$$

It follows that

$$\int \chi^{[\lambda,\mu]}(R^3)\, dR = \sum_{r=0}^{\lambda} \sum_{s=0}^{\mu} \sum_{t=0}^{\mu+r-s} F(m, n) = 1 + \min(\lambda, \mu);$$

that is, Eq. (8) is satisfied for every irreducible representation of $SU(3)$ and the $3j$ symbols of $SU(3)$ can all be chosen in symmetric form.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Bootstrap Prediction of Symmetry for a Soluble Static Model

ANDREW A. CUNNINGHAM
Department of Mathematical Physics, The University of Adelaide, South Australia

(Received 18 June 1966)

The Huang–Low bootstrap criterion of self-consistency, in the form of Levinson's theorem, is imposed on the exact solutions of the two-channel Low equation with an arbitrary crossing matrix. It is found that this condition, together with a number of dynamical conditions are sufficient to restrict the continuous crossing-matrix parameter to discrete values corresponding to an SU_2 symmetry.

I. INTRODUCTION

In recent years, much interest has been displayed in possible bootstrap predictions of internal symmetry. Studies have been made using an exact algebraic approach,[1] or approximate dynamical methods.[2] In work of the latter kind, a number of classified particles is given, and then, by insisting on a self-consistent solution, it is hoped that internal symmetries may be predicted. It would appear much better to avoid any approximations,[3] and to obtain an exact solution for the S-matrix elements defined by as general a set of postulates as possible, and then to impose a criterion of self-consistency. This is the approach of this paper. The postulates referred to are taken to be analyticity, unitarity, and crossing symmetry, each of which is considered quite fundamental to any description of interactions. The role of crossing symmetry in particular has been greatly stressed by Wigner.[4] It remains, of course, highly desirable that the number of dynamical assumptions should be as restricted as possible, and as general as possible. It is hoped to show how such a set of assumptions, together with the imposition of a bootstrap criterion on an exactly soluble two-channel static model will lead to a prediction of internal symmetry. The model to be considered is particularly suitable for such a study, as it incorporates a general formulation of crossing symmetry.

This model, with a crossing matrix corresponding to the scattering of a particle of isospin 1 by a baryon of isospin $\frac{1}{2}$, was exactly solved by Wanders.[5] Martin and McGlinn[6] generalized the crossing matrix for the model to include a continuous parameter, and, using an extension of the technique given by Wanders, obtained exact solutions for the two S-matrix elements. They had hoped that the existence of solutions satisfying analyticity, unitarity, and crossing symmetry would be limited to the restricted values of the crossing matrix parameter defining SU_2 symmetry. This, however, proved not to be the case, and further self-consistency and dynamical conditions are apparently needed before such a symmetry prediction can be made.

The self-consistency condition to be imposed in this

[1] R. Cutkosky, Bull. Am. Phys. Soc. **8**, 591 (1963); Phys. Rev. **131**, 1888 (1963).
[2] R. Capps, Phys. Rev. Letters **10**, 312 (1963); J. Sakurai, *ibid.* **10**, 446 (1963).
[3] B. Diu and H. Rubinstein, Nuovo Cimento **32**, 1103 (1964).
[4] E. P. Wigner, Phys. Today **17**, 34 (1964).
[5] G. Wanders, Nuovo Cimento **23**, 817 (1962).
[6] A. Martin and W. McGlinn, Phys. Rev. **136**, 1515 (1964).

paper is that the exact solutions, satisfying the generalized crossing symmetry, should also verify the bootstrap criterion proposed by Huang and Low.[7] This condition is essentially that of the Levinson theorem of potential scattering.[8] Huang and Low[9] have shown that this condition is also equivalent to the N/D prescription for a bootstrap solution, at least in the context of the Low equation. The condition, as a general one, has further motivation by such results as those obtained by Van Hove[10] on the possible asymptotic behavior of the S matrix.

In their second paper,[7] Huang and Low imposed their consistency condition on the Wander's solution, and obtained conclusions concerning subtractions, cutoffs, and position and location of bound states. Also, Huang and Mueller have, using the same criterion, considered the model with the generalized crossing matrix, and obtained similar quantitative conclusions without recourse to the exact solutions. In addition they were able to show that the crossing matrix parameter (defined in Sec. II) is restricted to values greater than $-\frac{1}{2}$. In this paper, it is principally sought to generalize the Huang–Low paper and so restrict the crossing matrix parameter to that set corresponding to SU_2 symmetry. To do this it is necessary to assume certain results of the Huang–Mueller paper, for certain cases.

In Sec. II the mathematical formulation of the problem is given, together with the set of dynamical assumptions to be made. In Sec. III, the consistency requirements are imposed and the solutions restricted. In Sec. IV, the results and possible extensions are discussed.

For ease of comparison, the same notation and development as that of Huang–Low has, as far as possible, been preserved.

II. FORMULATION

A. General Solution and Conditions

We denote the two S-matrix elements for the two-channel static model by $S_\alpha(\omega)$ ($\alpha = 1, 2$), where ω is the energy of the incident particle.

The momentum of the incident particle is therefore given by $q = (\omega^2 - 1)^{\frac{1}{2}}$. The branch cuts of q in the complex ω plane are taken to be from 1 to ∞ and from -1 to $-\infty$, and q is chosen to be real and positive just above the cut from 1 to ∞, so that iq is a real analytic function in the cut plane.

$S_\alpha(\omega)$ must satisfy, then, the following conditions of the model:

(i) Analyticity, so that S_α is real meromorphic in the cut ω plane.

(ii) Elastic unitarity, i.e., S_α has only 1 branch point on the positive real axis, which is of the square root type at $\omega = 1$, the threshold point. Also, the analytic continuation of S_α onto the second Riemann surface is given by

$$S_\alpha^{(2)}(\omega) = 1/S_\alpha(\omega).$$

(iii) Crossing symmetry, i.e., $S_\alpha(-\omega) = \sum_\beta A_{\alpha\beta} S_\beta(\omega)$. $(A_{\alpha\beta})$ is a real 2×2 matrix, and for the type of model under consideration has the general form

$$(A_{\alpha\beta}) = \frac{1}{(2t+1)} \begin{pmatrix} -1 & 2t+2 \\ 2t & 1 \end{pmatrix}, \tag{1}$$

where t is a real parameter.

We now write

$$S_\alpha(\omega) = 1 + 2iqV(\omega)h_\alpha(\omega), \tag{2}$$

where $V(\omega)$ is a cutoff function, taken to be of the form

$$V(\omega) = \kappa^{2c}/(q^2 + \kappa^2)^c, \quad \kappa > 1, \quad c = 0, 1, 2, \cdots. \tag{3}$$

The three requirements for S_α above are expressed by the following dispersion relation[11] for h_α:

$$h_\alpha(\omega) = P_\alpha(\omega) + \frac{1}{\pi} \int_1^\infty d\omega^1 q^1 V(\omega^1) \times \left[\frac{|h_\alpha(\omega^1)|^2}{\omega^1 - \omega} + \sum_\beta A_{\alpha\beta} \frac{|h_\beta(\omega^1)|^2}{\omega^1 + \omega} \right], \tag{4}$$

where P_α is the sum of poles located on the real axis between the branch points:

$$P_\alpha(\omega) = \sum_i \left(\frac{\lambda_{i\alpha}}{\omega_i - \omega} + \sum_\beta A_{\alpha\beta} \frac{\lambda_{i\beta}}{\omega_i + \omega} \right), \quad \lambda_{i\alpha} \geq 0, \quad |\omega_i| < 1. \tag{5}$$

The above dispersion relation may require one or more subtractions to account for high-energy effects.

The crossing relation ensures that each subtraction introduces only one further parameter.

We impose the following physical conditions on any solution:

(a) $\lambda_{i\alpha} > 0$. $\lambda_{i\alpha}$, being the squared coupling constant would, if negative, display a ghost state.

(b) $S_\alpha(\omega)$ have the correct threshold behavior,

$$S_\alpha(\omega) \underset{\omega \to 1}{\longrightarrow} 1 + 0(q), \tag{6}$$

which requires that $h_\alpha(1)$ be finite.

(c) $S_\alpha(\omega)$ have the correct high-energy behavior, which depends on the number of subtractions made in

[7] K. Huang and F. Low, J. Math. Phys. **6**, 795 (1965).
[8] N. Levinson, Kgl. Danske' Videnskab. Selskab., Mat. Fys. Medd. **25**, No. 9 (1949).
[9] K. Huang and F. Low, Phys. Rev. Letters **13**, 596 (1964).
[10] L. Van Hove, Nuovo Cimento **25**, 392 (1962).
[11] F. E. Low, Phys. Rev. **97**, 1392 (1955).

the dispersion relation. Since the unitarity condition may be written, for the unsubtracted case, as

$$\operatorname{Im} h_\alpha = qV\,|h_\alpha|^2,$$

so that

$$|h_\alpha(\omega)| \leq |qV(\omega)|^{-1} (\omega \geq 1),$$

subtractions are needed only if the cutoff factor is such that

$$\lim_{\omega\to\infty} qV(\omega) < 1. \tag{7}$$

(d) The target particle should occur as a bound state in one of the channels, in order that the incident particle–target particle coupling constant should be nonvanishing.

(e) No bound state should have a smaller mass than that of the target baryon, otherwise there would be an inelastic threshold below the elastic threshold.

(f) The S-matrix elements should not have an essential singularity at infinity, in order that the asymptotic phase shifts be determinable.

(g) There should be at most a finite number of resonances in the scattering.

The solution dictated by analyticity, unitarity, and crossing symmetry, and obtained by Rothleitner[12] may be written thus:

$$S_1 = U[(B - t - 1)/(B + t)]D, \tag{8}$$

$$S_2 = UD, \tag{9}$$

where

$$B = \tfrac{1}{2} + i[\pi^{-1} \log(\omega + q) - (\omega/q)\beta(\omega)] \tag{10}$$

and $\beta(\omega)$ is a real meromorphic and even function in the entire ω plane.

$$U(B) = \frac{\tan \tfrac{1}{2}(\pi B)}{\tan \pi[\tfrac{1}{2}(B + t)]} \frac{\Gamma[\tfrac{1}{2}(B + t + 1)]\Gamma[\tfrac{1}{2}(B - t)]}{\Gamma[\tfrac{1}{2}(B - t + 1)]\Gamma[\tfrac{1}{2}(B + t)]}$$

$$\equiv \frac{\tan \tfrac{1}{2}(\pi B)}{\tan \pi[\tfrac{1}{2}(B + t)]} U_0(B). \tag{11}$$

D is an arbitrary real symmetric S-matrix element, i.e.,

$$D(\omega) = D(-\omega), \tag{12a}$$

$$D^*(\omega) = D(\omega^*), \tag{12b}$$

$$D^{(2)}(\omega) = 1/D(\omega). \tag{12c}$$

The postulates (f) and (g) above imply that D may be represented in the form

$$D(\omega) = \prod_m \frac{1 - ir_m q}{1 + ir_m q} \prod_n \frac{(1 - a_n q)}{(1 + a_n q)} \frac{(1 + a_n^* q)}{(1 - a_n^* q)},$$

$$\operatorname{Im} r_m = 0, \quad \operatorname{Re} a_n > 0. \tag{13}$$

The normalization $D \to 1$, $q \to 0$ is chosen to comply with the threshold condition, which is then satisfied if $\beta(1) \neq 0$. The high-energy behavior is obtained by noting that

$$D(\omega) \xrightarrow[\omega\to\infty]{} 1 + \frac{d_1}{\omega} + \frac{d_2}{\omega^2} + \cdots, \tag{14}$$

$$B(\omega) \xrightarrow[\omega\to\infty]{} \left(\frac{i}{\pi}\right) \log \omega - \beta(\omega), \tag{15}$$

$$U(\omega) \xrightarrow[\omega\to\infty]{} \left[1 + \frac{t}{B(\omega)}\right]. \tag{16}$$

Hence

$$h_\alpha = \frac{S_\alpha - 1}{2iqV} \to \omega^{2c-1}\left[\frac{\mu_\alpha}{B(\omega)} + \frac{d_1}{\omega} + \frac{d_2}{\omega^2} + \cdots\right]. \tag{17}$$

Since $\beta(\omega)$ is an even meromorphic function

$$\beta(\omega) \xrightarrow[\omega\to\infty]{} k\omega^{2n} \quad (n = 0, \pm 1, \pm 2, \cdots). \tag{18}$$

Hence,

$$B(\omega) \approx \log \omega \quad \text{if} \quad n \leq 0,$$

$$B(\omega) \approx \omega^{2n} \quad \text{if} \quad n > 0.$$

The number of subtractions required in (4) is determined by the values of c, n and the numbers d_1, $d_2, \cdots$.

For the case of K subtractions we must have that:

$$\begin{aligned} &\text{if } c = \tfrac{1}{2}(1 + K)\text{: no condition on } \beta(\omega) \\ &\qquad\qquad\qquad\qquad \text{and } D(\omega); \\ &\text{if } c > \tfrac{1}{2}(1 + K)\text{: } n > c - \tfrac{1}{2}(1 + K), \\ &\qquad\qquad\qquad\qquad \text{plus a condition on } D(\omega). \end{aligned} \tag{19}$$

B. Bootstrap Conditions

For $\omega \geq 1$, we write

$$S_\alpha(\omega) = \exp[2i\delta_\alpha(\omega)], \tag{20}$$

$$D(\omega) = \exp[-2i\theta(\omega)], \tag{21}$$

$$(B(\omega) - t - 1)/(B(\omega) + t) = \exp[-2i\psi(\omega)], \tag{22}$$

$$U(\omega) = \exp[2i\varphi(\omega)], \tag{23}$$

where $\delta(\omega)$, $\theta(\omega)$, $\psi(\omega)$, and $\varphi(\omega)$ are real. The representations of Eqs. (22) and (23) are valid since $B^{(2)}(\omega) = -B(\omega) + 1$, and so

$$U^{(2)}(\omega) = +1/U(\omega),$$

$$\therefore \quad \delta_1(\omega) = \varphi(\omega) - \theta(\omega) - \psi(\omega), \tag{24a}$$

$$\delta_2(\omega) = \varphi(\omega) - \theta(\omega). \tag{24b}$$

Using the representation (13)

$$\theta(\omega) = \sum_m \tan^{-1}(r_m q) + \sum_n \tan^{-1} \frac{2qI_m a_n}{1 - q^2\,|a_n|^2}. \tag{25}$$

[12] J. Rothleitner, Z. Physik 177, 287 (1964).

ψ may be verified to be given by

$$\cot\psi = -[2/(2t+1)][\pi^{-1}\log(\omega+q) - (\omega/q)\beta(\omega)]. \tag{26}$$

We use the notation $\Delta\alpha = \alpha(\infty) - \alpha(1)$, so that, by the above, the phase shifts for each channel may be written as

$$\Delta\delta_1 = \Delta\varphi - \Delta\psi - \Delta\theta, \tag{27a}$$

$$\Delta\delta_2 = \Delta\varphi - \Delta\theta. \tag{27b}$$

The Huang–Low bootstrap criterion hence becomes a condition for $\Delta\theta$, $\Delta\varphi$, and $\Delta\psi$.

It is easily verified that

$$\Delta\theta/\pi = M_+ - M_- + \tfrac{1}{2}(m_+ - m_-), \tag{28}$$

where

$M_\pm$ = number of a_n's in (13) with $I_m a_n \gtrless 0$,

$m_\pm$ = number of r_m's in (13) with $r_m \gtrless 0$.

In Appendix B of the Huang–Low paper it is shown that

$$\Delta\psi/\pi = \tfrac{1}{2}(K_0 - K_1), \tag{29}$$

K^γ being the number of roots of the equation $B(\omega) = \gamma$. The following theorem, vital to the ensuing analysis, is also proved, but is here, again, only quoted:

(i) There is no root of $B = \gamma$ on the imaginary axis except for $\gamma = 0$.

(ii) There is no root on the real axis $|\omega| > 1$ for any γ. (30)

(iii) $K_\gamma = K_{-\gamma}$.

(iv) K_γ is independent of γ for $|\gamma| \geq \tfrac{1}{2}$.

(v) $K_1 = 1 - \nu + \max(Z_\beta, P_\beta)$

where $$\nu = \begin{cases} 1 \text{ if } \beta(\omega) \text{ has a pole at } \omega = 0 \\ 0 \text{ otherwise} \end{cases} \tag{31}$$

and Z_β, P_β are respectively the total number of zeros and poles of $\beta(\omega)$ on the physical sheet, and so are given by

$$\begin{aligned} P_\beta &= 2(N + N_+ + N_-), \\ Z_\beta &= 2(n + N + N_+ + N_-). \end{aligned} \tag{32}$$

Here, $N_\pm$ = number of poles of $\beta(\omega)$ on the real axis $\omega > 1$, with positive or negative residues. $2N$ = number of poles of $\beta(\omega)$ not on the real axis $|\omega| \geq 1$. It may also be shown by precisely similar methods that for $|\gamma| < \tfrac{1}{2}$, $K_\gamma = K_0$. This result is also used.

C. Location of Bound-State Poles

$S_\alpha(\omega)$ has the same poles as h_α, plus the poles of $V(\omega)$, which are of order c, located at $\omega = \pm i(\kappa^2 - 1)^{\frac{1}{2}}$. The poles on the real ω axis between ± 1 must conform to

$$P_\alpha(\omega) = \sum_i \left(\frac{\lambda_{i\alpha}}{\omega_i - \omega} + \sum_\beta A_{\alpha\beta} \frac{\lambda_{i\beta}}{\omega_i + \omega} \right), \qquad \lambda_{i\alpha} \geq 0, \quad |\omega_i| < 1.$$

The poles of $S_\alpha(\omega)$ can occur only at the poles of D, of U, and the roots of $B = -t$.

The poles of $U(B)$ in the physical region lie at

$$\begin{aligned} &B = +t, \\ &B = \pm(2n+1), \qquad n = 0, 1, 2\cdots, \\ &B = \pm(t - 2n - 2), \quad n = 0, 1, 2\cdots. \end{aligned} \tag{33a}$$

The zeros of $U(B)$ in the physical region lie at

$$\begin{aligned} &B = \pm 2n, \qquad n = 0, 1, 2, \cdots, \\ &B = \pm(t - 2n - 1), \quad n = 0, 1, 2, \cdots. \end{aligned} \tag{33b}$$

Unless t is an integer, there cannot be an infinite number of cancellations among these zeros and poles. For t integer, $\tan\pi(\frac{1}{2}B)/\tan\pi[\frac{1}{2}(B+t)]$ becomes a symmetric S-matrix element and can be absorbed in D, leaving $U = U_0$ with a finite number of zeros and poles, provided only that $B = \gamma$ has a finite number of roots for each γ. $B = \gamma$ has no roots on the imaginary axis except for $\gamma = 0$, and therefore unless t is an even integer greater than 2, the cutoff poles must occur in D.

In order that the pole terms of S_α have the required structure, poles of D must occur only at:

(1) $\pm i(\kappa^2 - 1)^{\frac{1}{2}}$;

(2) roots of $U = 0$;

(3) real roots of $B = t + 1$, lying between $\omega = \pm 1$, and that

(4) there must be zeros of D at all complex poles of U and

(5) there must be zeros of D at all bound-state poles of U.

The residues of bound-state poles in S_α have the opposite signs to those in $h_\alpha(\omega)$ since $iq < 0$ for $|\omega| < 1$. Denoting by Λ_α the effective coupling constant, which in terms of the actual squared coupling constant, λ_α, is given by

$$\Lambda_\alpha = 2V(\omega_\alpha)(1 - \omega_\alpha^2)^{\frac{1}{2}}\lambda_\alpha. \tag{34}$$

The connection between poles and bound states is shown in Table I.

III. NECESSARY BOOTSTRAP CONDITIONS

We now restrict the solutions by imposing the Levinson's theorem form of the bootstrap criterion, thereby obtaining necessary bootstrap conditions. By means of these conditions, the solutions for t noninteger are excluded.

We need the following lemma, which is proved in Appendix B.

TABLE I. Conditions for bound-state poles.

bound-state energy	channel	conditions	coupling constants
0	1, 2 or both	either $U = \infty$, $D = 0$ or $U = 0$, $D = \infty$ and UD has a simple pole	$\Lambda_1 - \Lambda_2 = -\frac{2t+1}{2t}\,\mathrm{Res}\,UD$
$0 < \lvert\omega_0\rvert < \lvert$	ω_0 in 1, $-\omega_0$ in 2	either (a) ω_0 is a pole of U, $D = 0$ and UD has a simple pole or (b) ω_0 is a pole of D, $U = 0$ and UD has a simple pole	$\Lambda_2 = (2t+1)\,\mathrm{Res}\,UD$ $\Lambda_1 = \mathrm{Res}\,UD\left[\frac{B(\omega_0) - t - 1}{B + t} - 2(t+1)\right]$
	ω_0 in 1	$B + t = 0$ (simple zero) $UD \neq 0$, or ∞	$\Lambda_1 = -(2t+1)\,\mathrm{Res}\,\frac{UD}{B+t}$
	ω_0 in 2	$B - t - 1 = 0$ (simple zero) $UD = \infty$ (simple pole)	$\Lambda_2 = U(B = t+1)\,\mathrm{Res}\,D$

Lemma: Writing $U(B, t) = \exp[i2\varphi_t(\omega)]$, then for $n = 0, 1, 2, \cdots$,

$\Delta\varphi_t = 0$ if $t \in (2n - s, 2n + s)$, $0 \leq s < \frac{1}{2}$,

$\Delta\varphi_t = \Delta\psi$ if $t \in (2n + 1 - s, 2n + 1 + s)$, $0 \leq s < \frac{1}{2}$,

$\Delta\varphi_t = \frac{1}{2}\Delta\psi$ if $t = \frac{1}{2}(2n + 1)$.

It is necessary to examine each of these three possible cases individually.

A. The case $\Delta\varphi = 0$

By (27) we may now write

$$\Delta\delta_1 = -\Delta\theta - \Delta\psi,$$
$$\Delta\delta_2 = -\Delta\theta.$$

The bootstrap criterion then takes the form

$$\Delta\theta/\pi = b_2, \tag{35a}$$
$$\Delta\psi/\pi = b_1 - b_2, \tag{35b}$$

where b_1, b_2 are respectively the number of bound states in channels 1 and 2.

We further write

$$b_1 = u_{10} + b_{1t} + p_1, \tag{36a}$$
$$b_2 = u_{20} + b_{21+t} + p_2, \tag{36b}$$

where u_{10} is the number of bound states in channel 1 at the roots of $U = 0$, b_{1t} is the number of bound states in channel 1 at the roots of $B = t$, p_1 is the number of bound states in channel 1 at the poles of $(B - t)U$, u_{20} is the number of bound states in channel 2 at the roots of $U = 0$, b_{21+t} is the number of bound states in channel 2 at the roots of $B = 1 + t$, p_2 is the number of bound states in channel 2 at the poles of $(B - t)U$. Now, for all values of t, U has poles at the roots of $B = t$, and for all values of t other than a positive even integer, has zeros at the roots of $B = 0$.

We therefore first consider the case when t is non-integral and hence write

$$U = [B/(B - t)]V(B). \tag{37}$$

By (28) we may write

$$\Delta\theta/\pi = \tfrac{1}{2}c + (\Delta\theta_0 + \Delta\theta_V + \Delta\theta_{-t} + \Delta\theta_{t+1})/\pi - \tfrac{1}{2}X, \tag{38}$$

where $\Delta\theta_0/\pi$ is the contribution to $\Delta\theta/\pi$ from zeros of $B = 0$, $\Delta\theta_{-t}/\pi$ is the contribution to $\Delta\theta/\pi$ from zeros of $B + t = 0$, $\Delta\theta_{1+t}/\pi$ is the contribution to $\Delta\theta/\pi$ from zeros of $B - (t + 1) = 0$, $\Delta\theta_V/\pi$ is the contribution to $\Delta\theta/\pi$ from zeros and poles of V, and $\frac{1}{2}X$ is the contribution to $\Delta\theta/\pi$ from "extra zeros" of D. X is, of course, either zero or a positive integer. Consider the roots of $B = -t$. Let there be K_{-tR} real roots and $2K_{-tc}$ complex roots. Therefore $K_{-tR} + 2K_{-tc} = K_{-t}$. Every complex root must be canceled and some or all of the real roots may be canceled by zeros of D.

Suppose K'_{-tR} real roots are canceled. Then

$$\Delta\theta_{-t}/\pi = -K_{-tc} - \tfrac{1}{2}K'_{-tR}, \tag{39}$$

as is seen from (28). Also

$$b_{1t} = K_{-tR} - K'_{-tR}. \tag{40}$$

Therefore

$$\Delta\theta_{-t}/\pi = \tfrac{1}{2}(b_{1t} - K_{-t}). \tag{41}$$

Consider the roots of $B = t + 1$ $(t \neq -1)$. Let there be K_{1+tR} real roots and $2K_{1+tc}$ complex roots. Therefore $K_{1+tR} + 2K_{1+tc} = K_{1+t}$.

Poles of D may be placed at the real roots but not at the complex roots. Suppose K'_{1+tR} of the real roots are canceled by poles of D. Then

$$\Delta\theta_{1+t}/\pi = \tfrac{1}{2}K'_{1+tR}, \tag{42}$$
$$b_{21+t} = K'_{1+tR}, \tag{43}$$
$$\Delta\theta_{1+t}/\pi = \tfrac{1}{2}b_{21+t}. \tag{44}$$

Consider the roots of $B = 0$. There is always a root at $\omega = 0$, unless $\beta(\omega)$ has a pole at the origin. We therefore designate the number of roots as follows:

at $\omega = 0$: $(1 + 2m_0)\delta_{\nu 0}$, where ν is as given in (31) and m_0 is a positive integer or zero.

at $\omega \neq 0$: $2K_{0R}$ real roots,
$2K_{0I}$ pure imaginary roots,
$4K_{0C}$ complex roots.

Therefore

$$(1 + 2m_0)\delta_{\nu 0} + 2(K_{0R} + K_{0I} + 2K_{0C}) = K_0. \quad (45)$$

Let there be a factor $[(1 - iq)/(1 + iq)]^s$ of D to cancel some of the roots at $\omega = 0$, and let there be poles of D such as to cancel $2K'_{0R}$, $2K'_{0I}$, $4K'_{0C}$ of the other roots. Hence, the contribution to $\Delta\theta/\pi$ from the B factor of U is

$$\Delta\theta_0/\pi = \tfrac{1}{2}s + \tfrac{1}{2}(K'_{0R} + K'_{0I}) + K'. \quad (46)$$

Suppose $V(B)$ has $2K_{VR}$ real roots; $4K_{Vc}$ complex roots; $2Z_{VR}$ real poles; $4Z_{Vc}$ complex poles [$V(B)$ has no pure imaginary zeros or poles]. D must have zeros at all the complex poles. Suppose D has zeros at $2Z_{VR}$ of the real poles. Suppose D has zeros at $4K'_{Vc}$ of the complex roots. Suppose D has poles at $4K'_{VR}$ of the real roots. Therefore

$$\Delta\theta_V/\pi = \tfrac{1}{2}K'_{VR} + K'_{Vc} - \tfrac{1}{2}Z'_{VR} - Z_{Vc}. \quad (47)$$

$BV(B)D$ can have at most a simple pole at the real roots and cannot have a pole at the nonreal roots. Therefore

$$K'_{0C} \leq K_{0C}, \quad (48)$$

$$K'_{0I} \leq K_{0I}, \quad (49)$$

$$K'_{Vc} \leq K_{Vc}. \quad (50)$$

Let z be the number of pairs of simple poles of $BV(B)D$ located at some or all of the nonzero roots of $B = 0$, $V = 0$, and the poles of V. Then

$$K'_{0R} - K_{0R} + K'_{VR} - K_{VR} + Z_{VR} - Z'_{VR} \leq z. \quad (51)$$

In the neighborhood of the origin $BD \sim \omega^x$ where $x = (1 + 2m_0)\delta\nu_0 - 2s$. For the target particle to be a bound state of the model, we must have $x = -1$. This is satisfied only when

$$s = 1 + m_0, \quad (52)$$

$$\nu = 0. \quad (53)$$

From Table I,

$$u_{10} + u_{20} + p_1 + p_2 = 2z + 1. \quad (54)$$

Using inequalities (48)–(51) we obtain

$$2[\Delta\theta_0/\pi + \Delta\theta_V/\pi] \leq [\tfrac{1}{2} + \tfrac{1}{2}K_0 + z] + [(K_{VR} + 2K_{Vc}) - (Z_{VR} + 2Z_{Vc})]. \quad (55)$$

But

$$\Delta\psi/\pi \equiv (u_{10} + b_{1t} + p_1) - (u_{20} + b_{21+t} + p_2) = \left(u_{10} + \frac{2\Delta\theta_{-t}}{\pi} + K_{-t} + p_1\right) - (u_{20} + 2\Delta\theta_{1+t}/\pi + p_2).$$

But

$$u_{20} + 2\Delta\theta_{1+t}/\pi + p_2 = \tfrac{1}{2}c + (\Delta\theta_0 + \Delta\theta_V + \Delta\theta_{-t} + \Delta\theta_{1+t})/\pi - \tfrac{1}{2}X,$$

so that

$$2\Delta\theta_{1+t}/\pi = c + 2(\Delta\theta_0 + \Delta\theta_V + \Delta\theta_{-t})/\pi - X - 2(u_{20} + p_2).$$

Hence

$$\Delta\psi/\pi = -c + K_{-t} - 2(\Delta\theta_0 + \Delta\theta_V)/\pi + (u_{10} + u_{20} + p_1 + p_2) + X;$$

therefore using (55) and (54)

$$-\Delta\psi/\pi \leq c - K_{-t} + [-\tfrac{1}{2} + \tfrac{1}{2}K_0 - z] + [(K_{VR} + 2K_{Vc}) - (Z_{VR} + 2Z_{Vc})] - X.$$

Therefore from (29) we obtain

$$0 \leq c + [K_0 - K_{-t} - \tfrac{1}{2}K_1] + [(K_{VR} + 2K_{Vc}) - (Z_{VR} + 2Z_{Vc})] - \tfrac{1}{2} - z - X.$$

$V(B)$ has zeros at the roots of

$$B = \pm 2n_1, \quad n_1 = 1, 2, \cdots$$

and

$$B = \pm[2(n - n_2) + s - 1], \quad n_2 = 0, 1, 2, \cdots.$$

$V(B)$ has poles at the roots and

$$B = \pm(2n_3 + 1), \quad n_3 = 0, 1, 2, \cdots,$$

and

$$B = [2(n - n_4) + s - 2], \quad n_4 = 0, 1, 2, \cdots,$$

where

$$t = 2n + s \quad \text{and} \quad 0 < |s| < \tfrac{1}{2}, \quad n = 0, 1, 2, \cdots.$$

We therefore see that

$$[(K_{VR} + 2K_{Vc}) - (Z_{VR} + 2Z_{Vc})] = K_1 - K_s = K_1 - K_0, \quad n \neq 0,$$
$$= 0 \text{ for } n = 0.$$

The above inequality hence becomes

$$0 \leq c - \tfrac{1}{2}K_1 - \tfrac{1}{2} - z - X,$$

i.e., using (32)

$$0 \leq c - (N + N_+ + N_-) - \max(0, n) - 1 - z - X. \quad (56)$$

From this inequality follow, as shown in Huang–Low's paper, the following two theorems.

Theorem 1: There exists no bootstrap solution satisfying an unsubtracted dispersion relation.

Theorem 2: For a once subtracted dispersion relation, a bootstrap solution must have $c = 1$. If further, the target baryon is required to be a bound state, then $\beta(\omega) = \beta_0$, where β_0 is a nonvanishing constant.

We now consider the once subtracted solutions, so that $c = 1, \beta(\omega) = \beta_0 \neq 0, K_1 = 1, s = 1$. From (56), also $X = z = 0$ and hence by (54) $u_{10} + u_{20} + p_1 + p_2 = 1$. But $u_{10} + u_{20} \neq 0$ since the target baryon is required to be a bound state in at least one of the channels. Therefore $p_1 = p_2 = 0$. Therefore

$$(u_{10}, u_{20}) = (1, 0) \quad \text{or} \quad (0, 1). \quad (57)$$

For $\beta_0 < 0$,

$$K_0 = K_1 = 1 \Rightarrow \Delta\psi/\pi = 0.$$

(See Appendix B of the Huang–Low paper.) Therefore

$$b_1 = b_2 = 1 \quad (58)$$

by (57). For $\beta_0 > 0$,

$$K_0 = K_1 + 2 \Rightarrow \Delta\psi/\pi = 1.$$

Therefore

$$b_1 = b_2 + 1.$$

Therefore by (57), either

$$(b_1, b_2) = (1, 0) \quad \text{or} \quad (2, 1). \quad (59)$$

But

$$\Delta\theta_V/\pi = b_2 - (\Delta\theta_0/\pi) - \tfrac{1}{2}b_{21+t} - \tfrac{1}{2}b_{1t},$$

while for

$$\beta_0 < 0, \quad \Delta\theta_0/\pi = \tfrac{1}{2} \quad \text{and} \quad b_{21+t} + b_{1t} = 1,$$

$$\beta_0 > 0, \quad \Delta\theta_0/\pi = \tfrac{1}{2} \quad \text{or} \quad 1$$

and

$$b_{21+t} + b_{1t} = 1 \quad \text{or} \quad 0.$$

Therefore in either case $\Delta\theta_V/\pi = 0, \tfrac{1}{2}, 1, 1\tfrac{1}{2}$. But $K'_{VR} \leq K_{VR}, Z'_{VR} = Z_{VR}$, and so

$$\Delta\theta_V/\pi \leq \tfrac{1}{2}(K_{VR} - Z_{VR}) = 0, \quad \text{or} \quad \tfrac{1}{2}$$

($B - \gamma = 0$ has no imaginary or complex zeros for $\gamma \neq 0$).

Therefore in all cases $K'_{VR} = K_{VR}$, $Z'_{VR} = Z_{VR}$. Denote the countable set of zeros of $V(B)$ by $\{Z_i\}$. Denote the countable set of poles of $V(B)$ by $\{P_i\}$. Write

$$Z_i = (1 - Z_i^2)^{\frac{1}{2}}, \quad P_i = (1 - P_i^2)^{\frac{1}{2}}.$$

Then $D(\omega)$ contains as a factor, the infinite product of terms

$$\prod_i \frac{(1 - iq/Z_i)}{(1 + iq/Z_i)} \prod_j \frac{(1 + iq/P_j)}{(1 - iq/P_j)}.$$

None of the terms can be canceled in the expressions for S_α. But these products are divergent in the supposed domain of analyticity of D, since

$$B(\omega) = \tfrac{1}{2} + i[\pi^{-1} \log(\omega + q) - \omega/q\beta_0].$$

Therefore

$$\frac{dB}{d\omega} = \frac{i}{q}\left[\frac{1}{\pi} - \frac{\beta_0}{1 - \omega^2}\right].$$

Therefore B is either monotonically increasing or decreasing in the neighborhood of the threshold $\omega = +1$. If ω_N is such that $B(\omega_N) = N$, then $|\omega_N| \to |$ as $N \to \infty$

$$\therefore \lim_{N\to\infty} \frac{B(\omega_N)}{N} \to 1 \quad \text{as} \quad N \to \infty, \quad S_N = (1 - \omega_N^2)^{\frac{1}{2}},$$

where

$$\therefore \quad S_{N/1/N} \xrightarrow[N\to\infty]{} -\beta_0 \neq 0. \quad (60)$$

The zeros of V occur at the roots

$$\omega_n^{(1)} \quad \text{of} \quad B - (2n + 2) = 0,$$

$$\omega_n^{(2)} \quad \text{of} \quad B - (2n + 1 - t) = 0, \quad n = 0, 1, 2, \cdots.$$

But

$$\frac{1 + Z_n^{(1)}}{1 - Z_n^{(1)}} = 1 + \frac{2Z_n^{(1)}}{1 - Z_n^{(1)}}.$$

Therefore by (60),

$$\frac{2Z_n^{(1)}}{1 - Z_n^{(1)}/(2n + 2)^{-1}} \xrightarrow[N\to\infty]{} -\beta_0,$$

$$\therefore \quad \prod_n \frac{(1 + Z_n^{(1)})}{(1 - Z_n^{(1)})}$$

is divergent. Similarly,

$$\prod_n \frac{(1 + Z_n^{(2)})}{(1 - Z_n^{(2)})}$$

is divergent. Likewise, poles of V occur at the roots

$$p_n^{(1)} \quad \text{of} \quad B - (2n + 1) = 0,$$

$$B - (2n + 2 - t) = 0, \quad n = 0, 1, 2, \cdots.$$

As above,

$$\prod_n \frac{(1 - P_n^{(1)})}{(1 + P_n^{(1)})} \quad \text{and} \quad \prod_n \frac{(1 - P_n^{(2)})}{(1 + P_n^{(2)})}$$

diverge. This shows the infinite product factors of D to be not absolutely convergent at the origin, and so contradicts the analyticity conditions for D. The solutions for noninteger t lying in the given ranges must hence be rejected.

We now obtain the solutions for $t = 2n$, $n = 1, 2, \cdots$. We may now put $U = U_0$. U now has poles at the roots of $B = 0$, and so we write $U(B) = [1/B(B-t)]V(B)$. D must have zeros at all the imaginary and complex roots of $B = 0$.

With an analogous notation to that used above,

$$\Delta\theta_0/\pi = -\tfrac{1}{2}s - \tfrac{1}{2}(K'_{0R} + K_{0I} + 2K_{0C}), \quad (61)$$

$$\Delta\theta_V/\pi = \tfrac{1}{2}(K'_{VR} - Z'_{VR}) + (K'_{Vc} - Z_{Vc}). \quad (62)$$

If z is the number of pairs of simple poles of $V(B)D/B$ located at some or all of the nonzero real roots of $B = 0$, $V = 0$, and the poles of V, then

$$K_{0R} - K'_{0R} + K'_{VR} - K_{VR} + Z_{VR} - Z'_{VR} \leq z. \quad (63)$$

Suppose $D/B \approx \omega^x$ as $\omega \sim 0$. Then

$$2s - (1 + 2m_0)\delta_{v0} = x.$$

Again we require that $x = -1$. We must have that $s = m_0$ and $v = 0$

$$2(\Delta\theta_0 + \Delta\theta_V)/\pi = -s - (K'_{0R} + K_{0I} + 2K_{0C}) + [(K'_{VR} + 2K'_{Vc}) - (Z'_{VR} + 2Z_{Vc})]; \quad (64)$$

therefore, using (63), (64), and $K'_{Vc} \leq K_{Vc}$,

$$2(\Delta\theta_0 + \Delta\theta_V)/\pi \leq [+\tfrac{1}{2} + z - \tfrac{1}{2}K_0] + [(K_{VR} + 2K_{Vc}) - (Z_{VR} + 2Z_{Vc})].$$

Now,

$$[(K_{VR} + 2K_{Vc}) - (Z_{VR} + 2Z_{Vc})] = K_1.$$

Finally,

$$2(\Delta\theta_0 + \Delta\theta_V)/\pi \leq [\tfrac{1}{2} + z - \tfrac{1}{2}K_0 + K_1]. \quad (65)$$

As before,

$$\Delta\psi/\pi = -c + K_{-t} - 2(\Delta\theta_0 + \Delta\theta_V)/\pi + (u_{10} + u_{30} + p_1 + p_3) + X,$$

and so

$$\tfrac{1}{2}(K_1 - K_0) \leq c - K_{-t} + [\tfrac{1}{2} + z - \tfrac{1}{2}K_0 + K_1] - 2Z - 1 - X,$$

i.e.,

$$0 \leq c - \tfrac{1}{2}K_1 - X - z - \tfrac{1}{2} \quad \text{(assuming } t \neq 0\text{)}.$$

This inequality is identical to (56) and so Theorems 1 and 2 follow immediately.

For the case of one subtraction, we may therefore set $c = 1$, $\beta(\omega) = \beta_0 \neq 0$, $X = z = 0$, $K_1 = 1$. For $\beta_0 < 0$, $K_0 = K_1 = 1$, therefore $s = 0$,

$$\therefore \quad \Delta\psi/\pi = 0, \quad \therefore \quad b_1 = b_2.$$

But

$$u_{10} + u_{20} + p_1 + p_2 = 1.$$

However, $p_1 + p_2 \neq 0$, since $\omega = 0$ is a bound-state pole of U. Therefore $u_{10} = u_{20} = 0$ and $(p_1, p_2) = (1, 0)$ or $(0, 1)$.

$$\therefore \quad K'_{1+tR} = 1 = K_{1+t},$$

$$K'_{-tR} = 1 = K_{-t}.$$

Also

$$K'_{0R} = K'_{0I} = K'_{0C} = 0.$$

Since

$$\Delta\theta/\pi = 1 = \tfrac{1}{2} + \Delta\theta/\pi \quad \text{then} \quad \Delta\theta_V/\pi = \tfrac{1}{2},$$

$$K'_{VR} = Z'_{VR} + 1 = Z_{VR} + 1 = K_{VR}.$$

D has zeros and poles at all the poles and zeros of V. But

$$U = \frac{B}{B - 2n}\prod_{K=0}^{n-1}\frac{[B + 2K + 1][B - 2K - 1]}{[B + 2K][B - 2K]}.$$

Denote the roots of $B = N$ by ω_N, $N = 1, 2, \cdots$, and write $S_N = (1 - \omega_N^2)^{\frac{1}{2}}$. We now have

$$D = \frac{(1 - iq/\kappa)}{(1 + iq/\kappa)}\frac{(1 - iq/S_{2n+1})}{(1 + iq/S_{2N+1})}\frac{(1 + iq/S_{2n})}{(1 - iq/S_{2n})} \times \prod_{k=0}^{m-1}\frac{(1 - iq/S_{2k+1})}{(1 + iq/S_{2k+1})}\prod_{k=0}^{n-1}\frac{(1 + iq/S_{2k})}{(1 - iq/S_{2\,k})}.$$

By Table I, the effective coupling constant squared for incident particle and target baryon is given by

$$\Lambda_1 = -\frac{(4n+1)}{4n}\operatorname{Res} UD\,|_{\omega=0} = +\frac{4n+1}{8n^2}\frac{\kappa+1}{\kappa-1}\frac{1}{1/\pi - \beta_0}\frac{1+S_{2n+1}}{1-S_{2n+1}}\frac{1-S_{2n}}{1+S_{2n}} \times \prod_{k=1}^{n-1}\left(\frac{2k+1}{2k}\right)^2\prod_{k=0}^{n-1}\frac{1+S_{2k+1}}{1-S_{2k+1}}\prod_{k=1}^{n-1}\frac{1+S_{2k}}{1-S_{2k}}; \quad (66)$$

therefore, $\Lambda_1 > 0$. Similarly for the bound state in channel 2 at the root of $B = 2n + 1$

$$\Lambda_2 = (4n+1)\frac{\kappa + S_{2n+1}}{\kappa - S_{2n+1}}\frac{S_{2n} - S_{2n+1}}{S_{2n} + S_{2n+1}}\prod_{k=1}^{2n}\frac{2k}{2k+1} \times \prod_{\kappa=0}^{n-1}\frac{S_{2k+1} + S_{2k+1}}{S_{2k+1} - S_{2n+1}}\prod_{k=1}^{n-1}\frac{S_{2k} - S_{2n+1}}{S_{2k} + S_{2n+1}}. \quad (67)$$

We have found a bootstrap solution for $t = 2n$, $n = 1, 2, \cdots$. We try now

$$(p_1, p_2) = (0, 1), \quad (b_{1t}, b_{21+t}) = (1, 0).$$

$$K'_{-tR} = K'_{1+tR} = 0,$$

$$K'_{VR} = K_{VR}, \quad Z'_{VR} = Z_{VR},$$

$$D = \frac{(1 - iq/\kappa)}{(1 + iq/\kappa)}\prod_{K=0}^{n-1}\frac{(1 - iq/S_{2k-1})}{(1 + iq/S_{2k+1})}\prod_{k=1}^{n-1}\frac{(1 + iq/S_{2k})}{(1 - iq/S_{2k})}.$$

The effective coupling constant squared for the bound state in channel 2 is

$$\Lambda_2 = \frac{(2t+1)}{2t}\,\mathrm{Res}\,UD\,|_{\omega=0}$$
$$= -\frac{4n+1}{8n^2}\prod_{k=1}^{n-1}\left(\frac{2k+1}{2k}\right)^2\frac{1}{\frac{1}{\pi}-\beta_0}\,\frac{\kappa+1}{\kappa-1}$$
$$\times\prod_{k=0}^{n-1}\frac{(1+S_{2k+1})}{(1-S_{2k+1})}\prod_{k=1}^{n-1}\frac{(1-S_{2k})}{(1+S_{2k})};$$

therefore $\Lambda_2 < 0$, and the solution is unacceptable. Suppose now that $\beta_0 > 0$, so that $K_0 = 3$. If

$$(p_1, p_2) = (1, 0), \quad (b_{1t}, b_{21+t}) = (1, 1),$$

then

$$K'_{-tR} = 0, \quad K'_{1+tR} = 1 = K_{1+t}.$$
$$\frac{\Delta\theta}{\pi} = \frac{1}{2} + \frac{1}{2} + \frac{(\Delta\theta_0 + \Delta\theta_V)}{\pi} = b_2 = 1,$$
$$\therefore \qquad (\Delta\theta_0 + \Delta\theta_V)/\pi = 0.$$

But

$$\Delta\theta_0/\pi = -\tfrac{1}{2}s - \tfrac{1}{2}(K'_{0R} + K'_{0I}) - K'_{0C},$$
$$K'_{0C} = 0.$$

For $\beta_0 < 1/\pi$, $K_{0R} = 1$ and we must have $K'_{0R} = K_{0R}$. For $\beta_0 = 1/\pi$, $S = m_0 = 1$. For $\beta_0 > 1/\pi$, $K_{0I} = 1$ and we must have $K'_{0I} = K_{0I}$. In all cases

$$\Delta\theta/\pi = -\tfrac{1}{2},$$
$$\Delta\theta_V/\pi = 1,$$
$$Z'_{VR} = Z_{VR}, \quad K'_{VR} = K_{VR} = Z_{VR} + 1.$$

Hence,

$$D(\omega) = \frac{(1-iq/\kappa)}{(1+iq/\kappa)}\frac{(1-iq/S_{2n+1})}{(1+iq/S_{2n+1})}\frac{(1+iq/S_0)}{(1-iq/S_0)}$$
$$\times\prod_{k=1}^{n}\frac{(1-iq/S_{2k-1})}{(1+iq/S_{2k-1})}\prod_{k=1}^{n-1}\frac{(1+iq/S_{2k})}{(1-iq/S_{2k})}.$$

The effective coupling constant squared for the bound state and ω_{2n} in channel 1 is

$$\Lambda_1 = -(4n+1)\,\mathrm{Res}\,\frac{UD}{B+t}\bigg|_{\omega=\omega_{2n}}$$
$$= -\frac{(4n+1)}{8n^2}\prod_{k=1}^{n}\frac{(2k-1)(4n-2k+1)}{2k(4n-2k)}$$
$$\times\frac{(1+S_{2n}/\kappa)}{(1-S_{2n}/\kappa)}\frac{(1+S_{2n}/S_{2n+1})}{(1-S_{2n}/S_{2n+1})}\frac{(1-S_{20}/S_0)}{(1+S_{20}/S_0)}$$
$$\times\prod_{k=1}^{n}\frac{(1+S_{2n}/S_{2k-1})}{(1-S_{2n}/S_{2k-1})}\prod_{k=1}^{n-1}\frac{(1-S_{2n}/S_{2k})}{(1+S_{2n}/S_{2k})}$$
$$\times\,\mathrm{Res}\,\frac{1}{B+2n}\bigg|_{\omega=\omega_{2n}}.$$

But

$$B'(\omega) = \frac{i}{q}\left[\pi^{-1} - \frac{\beta_0}{1-\omega^2}\right]$$

while

$$\mathrm{Res}\,\frac{1}{B+2n}\bigg|_{\omega=\omega_{2n}} = \frac{1}{B'(\omega_{2n})}\frac{(1+S_{2n}/S_{2n+1})}{(1-S_{2n}/S_{2n+1})}$$
$$\times\,\mathrm{Res}\,\frac{1}{B+2n}\bigg|_{\omega=\omega_{2n}} > 0.$$

If

$$\beta_0 > \frac{1}{\pi}, \quad S_0 > 1S_{2n}$$

while

$$S_{2n} > S_{2k}, \quad S_{2k-1}, \quad \Lambda_1 < 0.$$

If $\beta_0 < 1/\pi$, $S_0 < 1$ while $S_{2n} < S_{2k}$, S_{2k-1}, again $\Lambda_1 < 0$. Hence, the assumed distribution of bound-state poles gives an unacceptable solution. Try now

$$b_1 = 1, \quad b_2 = 0,$$

so that

$$(p_1, p_2) = (1, 0) \quad \text{and} \quad (b_{1t}, b_{21+t}) = (0, 0),$$
$$K'_{-tR} = 1, \quad K'_{1+tR} = 0,$$
$$\Delta\theta/\pi = \tfrac{1}{2} - \tfrac{1}{2} + \Delta\theta_0 + \frac{\Delta\theta_V}{\pi} = 0,$$

therefore, as before, $K'_{-VR} = K_{VR}$, $Z'_{VR} = Z_{VR}$. Hence,

$$D = \frac{(1-iq/\kappa)}{(1+iq/\kappa)}\frac{(1+iq/S_{2n})}{(1-iq/S_{2n})}\frac{(1+iq/S_0)}{(1-iq/S_0)}$$
$$\times\prod_{k=1}^{n}\frac{(1-iq/S_{2k-1})}{(1+iq/S_{2k-1})}\prod_{k=1}^{n-1}\frac{(1+iq/S_{2k})}{(1-iq/S_{2k})}.$$

The effective coupling constant this time is

$$\Lambda_1 = -\frac{(4n+1)}{4n}\,\mathrm{Res}\,UD\,|_{\omega=0}$$
$$= \frac{4n+1}{8n^2}\frac{\kappa+1}{\kappa-1}\frac{1-S_{2n}}{1+S_{2n}}\frac{1-S_0}{1+S_0}\frac{1}{\frac{1}{\pi}-\beta_0}$$
$$\times\prod_{k=1}^{n}\frac{(1+S_{2k-1})}{(1+S_{2k-1})}\prod_{k=1}^{n-1}\frac{(1+S_{2k})}{(1-S_{2k})}, \quad (68)$$

since,

$$\frac{1-S_0}{\frac{1}{\pi}-\beta^0} > 0, \quad \Lambda_1 > 0.$$

For this distribution of bound-state poles we have a solution. For the case $t = 0$, $U_0(B) \equiv 1$

$$S_1 = [(B-1)/B]D, \quad S_2 = D.$$

But then it is seen that if $\omega \equiv 0$ is a pole of S_2, it is at least a double pole of S_1, and so cannot be interpreted

as a bound-state pole for either channel. This concludes the discussion of the $\Delta\varphi = 0$ case.

B. The Case $\Delta\varphi = \Delta\psi$

Consider the case for

$$t = 2n + 1 + s, \quad 0 < |s| < \tfrac{1}{2}. \tag{69}$$

We now write

$$\Delta\delta_1 = -\Delta\theta, \quad \Delta\delta_2 = \Delta\psi - \Delta\theta,$$

so that the bootstrap criterion becomes

$$\Delta\theta/\pi = b_1, \quad \Delta\psi/\pi = b_1 - b_2. \tag{70}$$

Equations (36)–(55) remain valid. Again,

$$\begin{aligned}\Delta\psi/\pi &= (u_{10} + b_{1t} + p_1) - (u_{20} + b_{21+t} + p_2)\\ &= (u_{10} + 2\Delta\theta_{-t}/\pi + K_t + p_1)\\ &\qquad - (u_{20} + 2\Delta\theta_{1+t}/\pi + p_2).\end{aligned}$$

But now,

$$\begin{aligned}&u_{10} + (2\Delta\theta_{-t}/\pi) + K_{-t} + p_1\\ &\quad = \tfrac{1}{2}c + (\Delta\theta_0 + \Delta\theta_V + \Delta\theta_{-t} + \Delta\theta_{t+1})/\pi - \tfrac{1}{2}X,\\ &\Delta\theta_{-t}/\pi = u_{10} - K_{-t} - p_1 + \tfrac{1}{2}c\\ &\qquad + (\Delta\theta_0 + \Delta\theta_V + \Delta\theta_{t+1})/\pi - \tfrac{1}{2}X,\end{aligned}$$

so that

$$\begin{aligned}\Delta\psi/\pi = c - K_{-t} + \frac{2}{\pi}(\Delta\theta_0 + \Delta\theta_V)\\ - (u_{10} + u_{20} + p_1 + p_2) - X.\end{aligned}$$

Then, proceeding as before we find the following inequality:

$$\begin{aligned}0 \le c - \tfrac{1}{2} - \tfrac{1}{2}K_1 - z - X\\ + [(K_{VR} + 2K_{Vc}) - (Z_{VR} + 2Z_{Vc})].\end{aligned} \tag{71}$$

The zeros of $V(B)$ now occur at the roots of

$$B = \pm 2n_1, \quad n_1 = 1, 2, \cdots,$$
$$B = \pm(2n - 2n_2 + s), \quad n_2 = 0, 1, 2, \cdots.$$

The poles of $V(B)$ now occur at the roots of

$$B = \pm(2n_3 + 1), \quad n_3 = 0, 1, 2, \cdots,$$
$$B = (2n - 2n_4 - 1 + s), \quad n_4 = 0, \cdots,$$
$$(K_{VR} + 2K_{Vc}) - (Z_{VR} + 2Z_{Vc}) = K_0 - K_1,$$

and the above inequality becomes

$$0 \le c - \tfrac{1}{2} + K_0 - \tfrac{3}{2}K_1 - z - X \tag{72}$$

or

$$0 \le c + N_+ - 3N_- - N - \max(n_1\, 0) - z - X.$$

It is impossible to draw any conclusions from this inequality. We therefore try to incorporate the conclusions of the Huang–Mueller paper, thereby using an analysis of the pology off the real line. They show that the number of bound-state poles in channel 1 is 1 and in channel 2, either 1 or 0. Therefore assume $b_1 = 1$ and $b_2 = 0$ or 1.

$$\Delta\psi/\pi = 0 \quad \text{or} \quad 1.$$

For $\Delta\psi/\pi = 0$, $K_0 = K_1$ and we obtain from (72) the original inequality, (56), from which Theorems 1 and 2 follow. To obtain this then, we require that the target particle appears as a bound state and that each channel has the same number of bound states. The other conclusions of these are no solution without a subtraction and cutoff may then be drawn, together with the exclusion of the noninteger case.

For $\Delta\psi/\pi = 1$, $K_0 = K_1 + 2$, and so from (72),

$$0 \le (c + 2) - (N + N_+ + N_-) - \max(n, 0).$$

We may also take $c = 1$,

$$N + N_+ + N_- + \max(n, 0) \le 3.$$

Also $n + N + N_+ + N_- \ge 0$. The parameters n_1, N_1, N_+, N_- then all take a limited number of finite values. However, we now have

$$\begin{aligned}(u_{10}, b_{1t}, p_1) &= (1, 0, 0),\\ (u_{20}, b_{21+t}, p_2) &= (0, 0, 0),\\ (u_{10} + u_{20} + p_1 + p_2) &= 1 = 2Z + 1,\end{aligned}$$

$\therefore\ z = 0$, irrespective of the value of K_1.
Also,

$$\Delta\theta_{-t}/\pi = -\tfrac{1}{2}K_{-t} = -\tfrac{1}{2}K_1$$

since $t > \tfrac{1}{2}$;

$$\begin{aligned}&\Delta\theta_{1+t}/\pi = \tfrac{1}{2}b_{21+t} \Rightarrow K'_{1+tR} = 0,\\ &1 = \Delta\theta/\pi = \tfrac{1}{2} + (\Delta\theta_0 + \Delta\theta_V)/\pi - \tfrac{1}{2}K_1 - \tfrac{1}{2}V,\\ &2(\Delta\theta_0 + \Delta\theta_V)/\pi = 1 + K_1 + X.\end{aligned} \tag{73}$$

But

$$\begin{aligned}&2(\Delta\theta_0 + \Delta\theta_V)/\pi \le [\tfrac{1}{2} + \tfrac{1}{2}K_0] + K_0 - K_1 = \tfrac{7}{2} + K_1,\\ &X \le \tfrac{5}{2} - \tfrac{1}{2}K_1,\\ &\therefore \qquad 0 \le X \le 2, \quad 1 \le K_1 \le 5.\end{aligned} \tag{74}$$

Writing (73) explicitly

$$\begin{aligned}s + (K'_{0R} + K'_{0I} + 2K'_{0c}) + K'_{VR} + 2K'_{Vc}\\ - Z'_{VR} - 2Z_{Vc} = 1 + K_1 + X.\end{aligned}$$

But

$$\begin{aligned}s + (K_{0R} + K_{0I} + 2K_{0c}) + K_{VR} + 2K_{Vc}\\ - Z_{VR} + 2Z_{Vc} = \tfrac{7}{2} + \tfrac{1}{2}K_1.\end{aligned}$$

Subtracting,

$$\begin{aligned}&(K_{0R} - K'_{0R}) + (K_{0I} - K'_{0I}) + 2(K_{0c} - K'_{0c})\\ &\quad + (K_{VR} - K'_{VR}) + 2(K_{Vc} - K'_{Vc})\\ &\qquad - (Z_{VR} - Z'_{VR}) = \tfrac{5}{2} - \tfrac{1}{2}K_1 - X \le 2,\end{aligned}$$

by (74). But, in order that unwanted bound-state poles should not occur

$$K'_{0C} \leq K_{0C}, \quad K'_{0I} \leq K_{0I}, \quad K'_{Vc} \leq K_{Vc},$$
$$K'_{0R} \leq K_{0R}, \quad K'_{VR} \leq K_{VR}, \quad Z_{VR} \leq Z'_{VR},$$
$$\therefore \quad K_{VR} \leq 2 + K'_{VR}, \quad Z'_{VR} \leq 2 + Z_{VR};$$

therefore all but at most two of the infinite number of real zeros and poles of V must be canceled by D. $B(\omega)$ is a real analytic function in the cut plane and so is continuous on the interval $(-1, 1)$ in which $B(\omega) = \gamma$ has at most five roots, and at least one.

Hence there exists a neighborhood $(1 - \lambda, 1)$ in which $B(\omega)$ is a monotonic continuous function

$$B(\omega) \xrightarrow[\omega \to 1-]{} +\infty, \quad \text{if} \quad \beta(1) < 0,$$
$$B(\omega) \xrightarrow[\omega \to 1-]{} -\infty, \quad \text{if} \quad \beta(1) > 0.$$

Suppose, without loss of generality $B(\omega) \to +\infty$. Let ω_N be the largest real root of $B(\omega) = N$. The set of roots $\{\omega_N\}$ is therefore an infinite countable sequence. But

$$\lim_{N \to \infty} \omega_N = 1,$$
$$\lim_{N \to \infty} \frac{B(\omega_N)}{N} = 1,$$

i.e.,

$$S_N/N \to -\beta(1) \neq 0.$$

[$\beta(\omega)$ does not have a pole at $\omega = 1$, since $v = 0$.] $D(0)$ possesses the factor

$$\prod_N \frac{(P_N - 1)}{(P_N + 1)},$$

where $N = (2n + 1)$, $n = 0, 1, \cdots$. But

$$\frac{P_N - 1}{P_N + 1} = 1 - \frac{2}{P_N + 1},$$

and

$$\frac{2/P_N + 1}{1/N} \to \frac{2}{-\beta(1)} \neq 0.$$

The product diverges as do the other infinite products. The solution is therefore rejected. Suppose now that $t = (2n + 1)$, $n = 0, 1, 2, \cdots$. Inequality (71) remains valid, but since $s = 0$,

$$(K_{VR} + 2K_{Vc}) - (Z_{VR} + 2Z_{Vc}) = 0.$$

Hence, again we have (56), from which Theorems 1 and 2 follow. For one subtraction we may therefore set

$$c = 1, \quad \beta(\omega) = \beta_0 \neq 0, \quad X = z = 0, \quad K_1 = 1.$$

As before, $p_1 = p_2 = 0$ and $(u_{10}, u_{20}) = (1, 0)$ or $(0, 1)$. Suppose firstly that $\beta_0 < 0$. In this case $K_1 = K_0$ and the roots of $B(\omega) = \gamma$ are all real and simple. Since $\Delta\psi/\pi = 0$, then $b_1 = b_2$ (75), i.e., there are the same number of bound states in each channel. We now have

$$(u_{10}, u_{20}) = \begin{matrix}(1, 0)\\(0, 1)\end{matrix} \Leftrightarrow (b_{1t}, b_{21+t}) = \begin{matrix}(0, 1)\\(1, 0)\end{matrix}.$$

In either case, $b_1 = b_2 = 1$, i.e., there is only one bound state in each channel.

$$\Delta\theta/\pi = \tfrac{1}{2} + [(\Delta\theta_0 + \Delta\theta_V)/\pi + \tfrac{1}{2}(b_{1t} + b_{21+t})] - \tfrac{1}{2}$$
$$\therefore \quad (\Delta\theta_0 + \Delta\theta_V)/\pi = \tfrac{1}{2}.$$

Since

$$\Delta\theta_0/\pi = \tfrac{1}{2},$$
$$\Delta\theta_V/\pi = 0, \quad \text{or} \quad K'_{VR} = Z'_{VR}.$$

But $K'_{VR} \leq K_{VR}$ while $Z'_{VR} \geq Z_{VR}$ since $K_{VR} = Z_{VR}$, $K'_{VR} = K_{VR} = Z_{VR} = Z'_{VR}$, i.e., D must have simple poles at all the zeros of V and simple zeros at all the poles of V. Assuming that $(b_{1t}, b_{21+t}) = (1, 0)$, so that there is a bound state at $\omega = 0$ in channel 2, and a bound state at ω_{2n+1} in channel 1, then gives

$$D = \frac{(1 - iq)}{(1 + iq)} \frac{(1 - iq/\kappa)}{(1 + iq/\kappa)} \times \prod_{K=1}^{n} \frac{(1 - iq/S_{2K})}{(1 + iq/S_{2K})} \frac{(1 + iq/S_{2K-1})}{(1 - iq/S_{2K-1})}.$$

Then calculating the effective incident particle–target particle coupling constant squared given, in Table I, by

$$\Lambda_2 = [(2t + 1)/2t] \operatorname{Res} UD\big|_{\omega=0},$$

we find that $\Lambda_2 < 0$, and so this solution is unacceptable. Assuming that $(b_{1t}, b_{21+t}) = (0, 1)$ so that the target particle appears as a bound state in channel 1, while a bound state at the root of $B = 2n + 2$ appears in channel 2. We now have

$$D = \frac{(1 - iq)}{(1 + iq)} \frac{(1 - iq/\kappa)}{(1 + iq/\kappa)} \frac{(1 - iq/S_{2n+2})}{(1 + iq/S_{2n+t})} \frac{(1 + iq/S_{2n})}{(1 - iq/S_{2n})} \times \prod_{K=1}^{n} \frac{(1 - iq/S_{2K})}{(1 + iq/S_{2K})} \frac{(1 + iq/S_{2K-1})}{(1 - iq/S_{2K-1})},$$

from which we calculate the two effective coupling constants squared to be

$$\Lambda_1 = \frac{2(4n + 3)}{(2n + 1)^2} \prod_{K=1}^{n} \left(\frac{2K}{2K - 1}\right)^2 \left(\frac{1}{\pi} - \beta_0\right) \frac{\kappa + 1}{\kappa - 1} \times \frac{S_{2n+2} + 1}{S_{2n+2} - 1} \frac{S_{2n+1} - 1}{S_{2n+1} + 1} \prod_{K=1}^{n} \frac{(1 + S_{2K})}{(1 - S_{2K})} \frac{(1 - S_{2K-1})}{(1 + S_{2K-1})}, \tag{75a}$$

$$\Lambda_2 = 4(2n+1)\frac{S_{2n+2}^2}{\omega_{2n+2}}\frac{(1+S_{2n+2})}{(1-S_{2n+2})}\frac{\kappa+S_{2n+2}}{\kappa-S_{2n+2}} \times \frac{S_{2n+1}-S_{2n+2}}{S_{2n+1}+S_{2n+2}}\prod_{K=1}^{2n}\left(\frac{2K}{2K+1}\right) \times \prod_{K=1}^{n}\frac{(S_{2K}+S_{2n+2})(S_{2K-1}-S_{2n+2})}{(S_{2K}-S_{2n+2})(S_{2K-1}+S_{2n+2})}. \quad (75b)$$

Clearly $\Lambda_1 > 0$, while for $\beta_0 < 0$, $B(\omega)$ is monotonically increasing on the interval (0, 1) and so $\Lambda_2 > 0$. This solution is then an acceptable bootstrap solution. We now investigate the case when $\beta_0 > 0$, so that

$$K_1 = 1, \quad K_0 = 3, \quad \text{and} \quad \Delta\psi/\pi = 1, \qquad b_1 = b_3 + 1, \quad (76)$$

i.e., there is one more bound state in channel 1 than in channel 2. If as before, ω_0 denotes the nonzero root of $B = 0$, then for $\beta_0 < 1/\pi$, ω_0 is real between $\omega = \pm 1$. For $\beta_0 > 1/\pi$, ω_0 is pure imaginary. For $\beta_0 = 1/\pi$, all three roots coincide at $\omega = 0$. We must have $(u_{10}, u_{20}) = (1, 0)$ and either $(b_{1t}, b_{21+t}) = (0, 0)$ or (1, 1), i.e., either $b_1 = 1, b_2 = 0$ or $b_1 = 2, b_2 = 1$. However, $\Delta\theta_0/\pi + \Delta\theta_V/\pi = 1$, in both cases. But $\Delta\theta_0/\pi = \frac{1}{2}$ or 1 while $\Delta\theta_V/\pi \leq 0$. Therefore the only possibility is that $\Delta\theta_0/\pi = 1$, i.e., $K'_{0R} + K'_{0I} = 1$ and

$$\Delta\theta_V/\pi = 0,$$

$$\therefore \qquad K'_{VR} = Z'_{VR} = Z_{VR} = K_{VR}. \quad (77)$$

Again, D has simple poles at all the zeros of V, and has simple zeros at all the poles of V. For $(b_{1t}, b_{21+t}) = (1, 1)$,

$$D = \frac{(1-iq)}{(1+iq)}\frac{(1-iq/\kappa)}{(1+iq/\kappa)}\frac{(1-iq/S_0)}{(1+iq/S_0)}\frac{(1-iq/S_{2n+2})}{(1+iq/S_{2n+2})} \times \prod_{K=1}^{n}\frac{(1+S_{2K})}{(1-S_{2K})}\frac{(1-S_{2K-1})}{(1+S_{2K-1})}.$$

For the bound state at $\omega = \omega_{2n+1}$, Table I requires that

$$\Lambda_1 = -\frac{(4n+3)}{2(2n+1)}\operatorname{Res}\frac{UD}{B+2n+1}\bigg|_{\omega=\omega_{2n+1}},$$

which is easily verified to be negative. This solution is therefore rejected. Consider the case $(b_{1t}, b_{21+t}) = (0, 0)$ so that

$$D = \frac{(1-iq)}{(1+iq)}\frac{(1-iq/\kappa)}{(1+iq/\kappa)}\frac{(1+iq/S_{2n+1})}{(1-iq/S_{2n+1})}\frac{(1-iq/S_0)}{(1+iq/S_0)} \times \prod_{K=1}^{n}\frac{(1+iq/S_{2K})}{(1+iq/S_{2K})}\frac{(1+iq/S_{2K-1})}{(1-iq/S_{2K-1})}.$$

The effective incident particle–target particle coupling constant is then given by

$$\Lambda_1 = \frac{4n+3}{2(2n+1)^2}V(0)\operatorname{Res} BD\bigg|_{\omega=0} = \frac{2(4n+3)}{(2n+1)^2}\left(\frac{1}{\pi}-\beta_0\right)\frac{\kappa+1}{\kappa-1}\prod_{K=1}^{n}\left(\frac{2K}{2K-1}\right)^2 \times \frac{S_{2n+1}-1}{S_{2n+1}+1}\frac{S_0+1}{S_0-1}\prod_{K=1}^{n}\frac{(1+S_{2K})}{(1-S_{2K})}\left(\frac{1-S_{2K-1}}{1+S_{2K-1}}\right). \quad (78)$$

This is positive since

$$\frac{(1/\pi-\beta_0)}{S_0-1} \geq 0 \quad \text{for} \quad \beta_0 > 0.$$

We therefore have an acceptable bootstrap solution.

C. The Case $\Delta\varphi = \frac{1}{2}\Delta\psi$

It remains to consider the solutions for $t = \frac{1}{2}(2n+1)$, $n = 0, 1, 2, \cdots$. Now

$$\frac{\Delta\delta_1}{\pi} = -\frac{\Delta\psi}{2\pi} - \frac{\Delta\theta}{\pi} = -b_1,$$

$$\frac{\Delta\delta_2}{\pi} = \frac{\Delta\psi}{2\pi} - \frac{\Delta\theta}{\pi} = -b_2.$$

The bootstrap conditions are now that

$$\Delta\psi/\pi = b_1 - b_2, \quad (79)$$

$$\Delta\theta/\pi = \tfrac{1}{2}(b_1 + b_2), \quad (80)$$

Equations (36)–(55) are still valid, but now

$$[(K_{VR} + 2K_{Vc}) - (Z_{VR} + 2Z_{Vc})] = K_{\frac{1}{2}} - K_1, \quad \text{for } n \text{ even}, \quad = 0, \text{ for } n \text{ odd}.$$

Suppose ω is such that

$$\pi^{-1}\log(\omega+q) = (\omega/q)\beta(\omega),$$

$$\pi^{-1}\log(-\omega+q) = -(\omega/q)\beta(\omega),$$

$$\pi^{-1}\log(-1) - \pi^{-1}\log(\omega+q) = -(\omega/q)\beta(\omega).$$

The only roots of $B(\omega) - \frac{1}{2} = 0$ are at $\omega = 1$, therefore $K_{\frac{1}{2}} = 0$. Now,

$$\left(u_{10} + \frac{2\Delta\theta_{-t}}{\pi} + K_{-t} + p_1\right) + \left(u_{20} + \frac{2\Delta\theta_{1+t}}{\pi} + p_2\right) = c + \frac{2(\Delta\theta_0 + \Delta\theta_V)}{\pi} + \frac{2\Delta\theta_{-t}}{\pi} + \frac{2\Delta\theta_{1+t}}{\pi} - X,$$

$$\therefore \quad u_{10} + u_{20} + p_1 + p_2 + K_{-t} = c + \frac{2(\Delta\theta_0 + \Delta\theta_V)}{\pi} - X,$$

$$\therefore \quad 2z + 1 + K_{-t} - c + X \leq [\tfrac{1}{2} + \tfrac{1}{2}K_0 + z] + [(K_{VR} + 2K_{Vc}) - (Z_{VR} + 2Z_{Vc})],$$

i.e.,

$$0 \leq c + \tfrac{1}{2}K_0 - K_{-t} - X - z - \tfrac{1}{2} + [(K_{VR} + 2K_{Vc}) - (Z_{VR} + 2Z_{Vc})].$$

This inequality becomes

$$0 \leq c + \tfrac{1}{2}K_0 - K_1 - X - z - \tfrac{1}{2}, \quad n = 0, \qquad (81a)$$

$$0 \leq c + K_0 - K_1 - X - z - \tfrac{1}{2}, \quad n \text{ odd}, \qquad (81b)$$

$$0 \leq c + \tfrac{1}{2}K_0 - 2K_1 - X - z - \tfrac{1}{2}, \quad n \text{ even}. \qquad (81c)$$

From the last inequality, for n even, we are able to derive Theorems 1 and 2, and so exclude the solutions, since there is no infinite number of cancellations among the zero and poles of V for this case.

For the first two inequalities it is impossible to draw any conclusions without recourse to the results of the Huang–Mueller paper.

This we do by assuming that either

$$b_1 = b_2 \quad \text{or} \quad b_1 = b_2 + 1.$$

Then, from (79), either $\Delta\psi/\pi = 0$ or 1; i.e., either $K_1 = K_0$ or $K_0 = K_1 + 2$. For both (81a) and (81b) we have

$$0 \leq c - \tfrac{1}{2}K_1 + \tfrac{1}{2} - X - z.$$

From this, with the single assumption stated above, Theorems 1 and 2 follow, and the solutions are excluded by the nonanalyticity of $D(\omega)$.

IV. DISCUSSION

We have applied the bootstrap criterion of Huang and Low to the two-channel static model, and found that only for values of the parameter t, corresponding to an internal SU_2 symmetry, can acceptable solutions be found. In addition, the possible distributions of bound-state poles have been found to be limited to one in channel 1 and either 1 or 0 bound-state poles in channel 2, the target baryon occurring as a bound state in channel 1.

It is of considerable interest to note that a similar solution to that obtained in Sec. III for $t = 1$ and for $\beta < 0$ was obtained by McGlinn and Albright[13], using the condition that there be a minimum total number of zeros of both S-matrix elements.

Even after the imposition of the consistency condition, there remain two undetermined constants, namely κ, which arose from the assumed cutoff function, and β, which may be thought of as an effective subtraction constant. As discussed in the Huang–Low paper, this nonuniqueness of the final solutions may be a manifestation of insufficient account being taken of high-energy effects and other channels, only considered by means of the cutoff. The necessity of a cutoff indicates, in fact, that the solutions depend sensitively on the high-energy behavior. It would be interesting to investigate the extent to which the assumed dynamical properties, for example, the form of the cutoff, are necessary to obtain the conclusion that t must be a positive integer. In this connection, it would seem that the model is not appropriate to the description of the scattering of two isofermions, e.g., $N - K$ scattering for which $t = \frac{1}{2}$. Nevertheless, the dynamical assumptions made are, for the most part, quite general, and it is hoped that the results obtained may be generalized to more physical models than that considered in this paper, for which similar conditions would be required to hold.

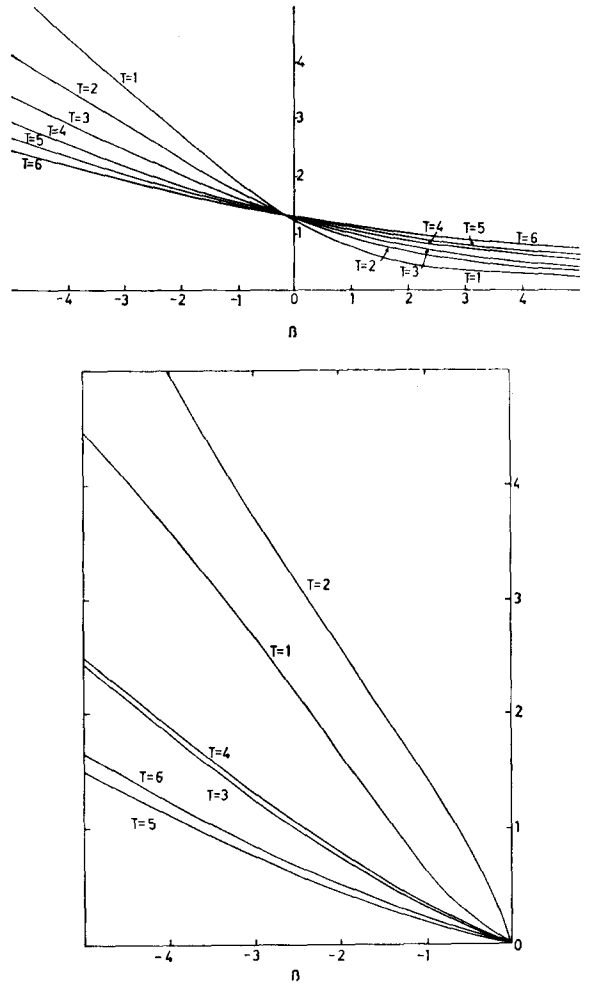


FIG. 1.(a) The coupling constant squared for the target baryon as a function of the parameter β_0, with incident isospin 1, 2, 3, 4, 5, 6. (b) The coupling constant squared for the bound state in channel 2 for incident isospin 1, 2, 3, 4, 5, 6.

[13] W. McGlinn and C. Albright, Nuovo Cimento 27, 834 (1963).

It would be very valuable to apply similar techniques to the Low equation with a 4×4 crossing matrix. However, although work is being done in this direction, and is discussed quantitatively by Huang and Mueller, exact solutions to this problem have not yet been found. An extension to the relativistic case is very difficult to formulate, and when achieved will lead naturally, it is hoped, to a prediction of internal symmetry.

A numerical computation was made of the coupling constants derived in Sec. III, for $t = 1, 2, 3, 4, 5, 6$ and with $\kappa = 7.5$ (the approximate nucleon–mesonic mass ratio) for a range of values of β from -5 to $+5$. The quantitative results are shown in Figs. 1(a) and 1(b).

ACKNOWLEDGMENTS

The author thanks Professor C. A. Hurst and Dr. J. Smith for many helpful discussions during the course of this work.

This work was undertaken in Australia under the auspices of the Commonwealth Scholarship and Fellowship Plan.

APPENDIX. PROOF OF LEMMA

We have

$$U_0(B, t) = \frac{\Gamma[\frac{1}{2}(B + t + 1)]\Gamma[\frac{1}{2}(B - t)]}{\Gamma[\frac{1}{2}(B - t + 1)]\Gamma[\frac{1}{2}(B + t)]}. \quad \text{(A1)}$$

But we may use the representation

$$\frac{\Gamma(Z_1)\Gamma(Z_2)}{\Gamma(Z_1 + Z_3)\Gamma(Z_2 - Z_3)} = \prod_{n=0}^{\infty}\left[1 + \frac{Z_3}{Z_1 + n}\right]\left[1 - \frac{Z_3}{Z_2 + n}\right] \quad \text{(A2)}$$

to give

$$U_0(B, t) = \prod_{n=0}^{\infty}\left[1 + \frac{\frac{1}{2}}{\frac{1}{2}(B - t) + n}\right] \times \left[1 - \frac{\frac{1}{2}}{\frac{1}{2}(B + t + 1 + n)}\right] = \prod_{n=0}^{\infty}\left[\frac{B - t + 2n + 1}{B - t + 2n}\right]\left[\frac{B + t + 2n}{B + t + 2n + 1}\right]. \quad \text{(A3)}$$

We examine the phase shift contribution from each bracketed term. Writing

$$\arg\left[\frac{B - t + 2n + 1}{B - t + 2n}\right] = 2\varphi_n^+(t), \quad \text{(A4)}$$

$$\arg\left[\frac{B + t + 2n}{B + t + 2n + 1}\right] = 2\varphi_n^-(t). \quad \text{(A5)}$$

It is easily verified that

$$i \tan 2\varphi_n^-(t) = \frac{2(2B - 1)}{(2B - 1)^2 - [2t - 4n - 3][2t - 4n - 1]}.$$

But

$$(2B - 1) = -i(2t + 1)\cot\psi, \quad \text{(A6)}$$

$$\tan 2\varphi_n^+(t) = \frac{2(2t + 1)\cot\psi}{(2t + 1)^2\cot^2\psi + [2t - 4n - 3][2t - 4n - 1]}. \quad \text{(A7)}$$

Similarly

$$\tan 2\varphi_n^-(t) = \frac{2(2t - 1)\cot\psi}{(2t - 1)^2\cot^2\psi + [2t + 4n + 3][2t + 4n + 1]}. \quad \text{(A8)}$$

From (A7), we see that

$\Delta\varphi_n^+(t) = 0$, if $(2t - 4n - 3)(2t - 4n - 1) > 0$,

i.e., if

$$t < 2n + \tfrac{1}{2},$$

or if

$$t > (2n + 1) + \tfrac{1}{2}.$$

$$\Delta\varphi_n^+(t) = \Delta\psi,$$

if

$$(2t - 4n - 3)(2t - 4n - 1) < 0,$$

i.e., if

$$(2n + 1) - \tfrac{1}{2} < t < (2n + 1) + \tfrac{1}{2}.$$

$$\Delta\varphi_n^+(t) = \tfrac{1}{2}\Delta\psi,$$

if

$$(2t - 4n - 3)(2t - 4n - 1) = 0,$$

i.e., if

$$t = 2n + \tfrac{1}{2} \quad \text{or} \quad (2n + 1) + \tfrac{1}{2}.$$

From (8) we see that the denominator is $+Ve$ definite for $t > -\frac{1}{2}$ and so

$$\Delta\varphi_n^-(t) = 0.$$

If $2\varphi_t = \arg U_0(B, t)$,

$\Delta\varphi_t = 0$ if $t\epsilon(2n - s, 2n + s)$, $0 < s < \frac{1}{2}$,

$\Delta\varphi_t = \Delta\psi$ if $t\epsilon(2n + 1 - s, 2n + 1 + s)$, $0 < s < \frac{1}{2}$,

$\Delta\varphi_t = \frac{1}{2}\Delta\psi$ if $t =$ half-odd integer.

For the same result to be true for $\arg U = \arg U_0 + \arg E$ we must show that $\Delta \arg E = 0$. We calculate that

$$\tan \arg E = \frac{\tan^2(\frac{1}{2}\pi t)[\tan^4(\frac{1}{2}\pi B) - 1]}{\tan^4(\frac{1}{2}\pi B)\tan^2(\frac{1}{2}\pi t) - 2\tan^2(\frac{1}{2}\pi B) + \tan^2(\frac{1}{2}\pi t)}.$$

The denominator is either $+Ve$ definite or has four real roots. But all real roots of $B = \gamma$ lie between the branch points. The lemma then follows.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Group Embedding of Space-Time and Internal Symmetry Groups

H. D. DOEBNER AND G. C. HEGERFELDT
Institut für Theoretische Physik (I), Philipps Universität, Marburg, Germany

(Received 30 May 1966)

An embedding of $SU(2)$ and an internal symmetry group G into a larger group $\tilde{G}$ containing $SU(2)$ and G as subgroups is constructed for all G possessing a generalized spin-$\frac{1}{2}$ quark model. The starting point is a set of three mathematical conditions for the embedding group $\tilde{G}$ which are derived from physically plausible assumptions. By group theoretical techniques due to Dynkin and Malcev, it is shown that the embedding group is already uniquely determined by the proposed conditions, with only one set of groups G for which two solutions are obtained. The results for $\tilde{G}$ are given explicitly. Identifying $SU(2)$ as covering of the rotation group the spin extension $\tilde{G}$ is enlarged to an embedding $\bar{G}_h$ of the homogeneous Lorentz group L_h and G. It is shown that $\bar{G}_h$ can also be obtained without using the spin extension. The minimal translation group which can be attached to $\bar{G}_h$ is calculated. The results are also taken over to the Budini–Fronsdal identification of $SU(2)$.

INTRODUCTION

IF two different and independent parts of a physical symmetry are known, given by a set of representations of internal and kinematical symmetry groups G_1 and G_2, the problem of constructing a larger group $\tilde{G}$ which combines the two different symmetries in a nontrivial way, not in the form of a pure direct product $\tilde{G} = G_1 \otimes G_2$, has by itself no unique mathematical solution. There are different methods to describe what is meant by "$\tilde{G}$ combines G_1 and G_2." For example, one can follow the idea of Michel[1] by demanding that G_1 is invariant subgroup of $\tilde{G}$ and that $G_2 \simeq \tilde{G}/G_1$, which for the Lorentz group yields a homomorphic image of the direct product; or one can consider the multitude of different groups $\tilde{G}$ in which G_1 and G_2 can be embedded, i.e., are contained as subgroups, and which are limited by the McGlinn-type theorems[2] and with respect to mass splitting by the O'Raifeartaigh–Jost theorem.[3] Further physical principles are needed to restrict $\tilde{G}$, especially to select one of the possible ways of constructing it and to assure that $\tilde{G}$ is unique with respect to this construction. Presumably such physical principles will be rather stringent and not very transparent. However, for the embedding of a wide class of internal symmetry groups, as for instance $SU(3)$, G_2, and $Sp(6)$, with $SU(2)$ describing the spin for particles at rest, a unique construction of $\tilde{G}$ is formulated in the sequel, characterized by three simple and physically plausible conditions. From this unique spin extension $\tilde{G}$, a corresponding embedding of the covering of the homogeneous Lorentz group $SL(2, C)$ in a group $\bar{G}_h$ can be constructed. If the relation between $SU(2)$ and $SL(2, C)$ is taken over to $\tilde{G}$ and $\bar{G}_h$, the result is also unique, but still open to physical interpretation, leading to different results if the translation group is attached.

[1] L. Michel, Phys. Rev., **137**, B405 (1965).
[2] W. D. McGlinn, Phys. Rev. Letters **12**, 467 (1964); L. O'Raifeartaigh, Phys. Rev. **139**, B1052 (1965).
[3] L. O'Raifeartaigh, Phys. Rev. Letters **14**, 575 (1965); M. Flato and D. Sternheimer, Phys. Rev. Letters **15**, 934 (1965); R. Jost, Helv. Phys. Acta **39**, 369 (1966).

1. FORMULATION OF THE EMBEDDING CONDITIONS

We look for an embedding of $SU(2)$ and a given connected Lie group G, describing the internal symmetry, in $\tilde{G}$ with

$$SU(2), G \subset \tilde{G},$$

where $\subset$ denotes proper inclusion.

(a) Suppose all particles of a G multiplet have equal spin, then an irreducible representation of G and $SU(2)$, characterized by a set of quantum numbers d and one quantum number s, is realized in the same representation space and the physical multiplets correspond to a subset of irreducible representations of $G \otimes SU(2)$, i.e., $\tilde{G}$ should contain G and $SU(2)$ in the form of a direct product.

(b) Suppose the irreducible unitary representations of G and $\tilde{G}$ are finite dimensional, i.e., G and $\tilde{G}$ are compact. The Lie algebra G^0 of a compact Lie group G has the form

$$G^0 = H_1^0 \oplus \cdots \oplus H_n^0 \oplus K_1^0 \oplus \cdots \oplus K_m^0. \quad (1.1)$$

H_i^0 are simple compact Lie algebras, and K_j^0 are one-dimensional algebras. The corresponding additive quantum numbers are completely independent of the rest of the group. In the spin extension this independence should be retained so that K_j^0 can be dropped. Then G becomes semisimple.

(c) Suppose G possesses a sort of quark model by which we mean that all physically realized representations can be obtained from Kronecker products (denoted by $\times$) of an irreducible representation (d_Q) and its complex conjugate $(\bar{d}_Q)$ associated with G-quarks. The dimension N of (d_Q) is as low as possible. The

image of G under the mapping by the quark representation is denoted by G_Q which is a unimodular matrix group for G semisimple. To attach the spin to the physical particles in a quark model, it is plausible to start from spin-$\frac{1}{2}$ quarks and to restrict the discussion to such symmetry groups which possess a spin-$\frac{1}{2}$ or a generalized quark model, i.e., all physical representations can be obtained from Kronecker products of the $2N$-dimensional irreducible spin-$\frac{1}{2}$ quark representations $(d_Q, \frac{1}{2})$ and $(\bar{d}_Q, \frac{1}{2})$ of $G \otimes SU(2)$.

To translate this model into some properties of $\tilde{G}$, a representation (d_Q) of $\tilde{G}$ should exist which, restricted to $G \otimes SU(2)$, is just the quark representation with no further particles. Hence the dimension of $(\tilde{d}_Q)$ is given by $2N$. This condition assures that all physical representations of $G \otimes SU(2)$ appear in the reduction of Kronecker products $(\tilde{d}_Q) \times \cdots \times (\bar{\tilde{d}}_Q) \times \cdots \times (\bar{\tilde{d}}_Q)$ to irreducible representations of $\tilde{G}$ if these are restricted to $G_Q \times SU(2)$. We can therefore identify $\tilde{G}$ with $(\tilde{d}_Q)$ and take $\tilde{G}$ as a linear group of $2N \times 2N$ matrices containing $G_Q \times SU(2)$:

$$\tilde{G} \supset G_Q \times SU(2). \tag{1.2}$$

With G not admitting a quark model, the reduction does in general not lead to all particles. Take for instance G as isospin group $SU(2)$, then $SU(4)$ contains $SU(2) \times SU(2)$. One gets all irreducible representations of $SU(4)$ by reduction of $SU(4) \times \cdots \times SU(4)$, the $SU(2) \times SU(2)$ content of which is calculated from $[SU(2) \times SU(2)] \times \cdots \times [SU(2) \times SU(2)]$, i.e., from

$$(\tfrac{1}{2}, \tfrac{1}{2}) \times \cdots \times (\tfrac{1}{2}, \tfrac{1}{2}). \tag{1.3}$$

In (1.3) there are only representations with spin and isospin being either both integer or half integer. Hence representations like $(\frac{1}{2}, 0)$ are not obtained.

(d) Suppose the embedding produces a mixing between the symmetry groups, thus excluding the already discussed direct product and a physically equivalent modification of the direct product. The trivial case enters if G and $SU(2)$ are embedded in separate groups H_1 and H_2 with $\tilde{G}$ as direct product $\tilde{G} = H_1 \otimes H_2$—an example is $\tilde{G} = U(3) \otimes SU(2)$ for $G = SU(3)$. This type of embedding is forbidden by demanding $\tilde{G}$ not to split into the product of two groups H_1, H_2 with $H_1 \supset G$, $H_2 \supset SU(2)$, that is if

$$G \subset H_1, \quad SU(2) \subset H_2, \quad \text{then} \quad \tilde{G} \neq H_1 \otimes H_2, \tag{1.4}$$

which is satisfied if $G_Q \times SU(2)$ is contained in a simple subgroup of $\tilde{G}$. For the case of simple symmetry groups G, this requirement can be proved also to be necessary.

Theorem 1.1: Let the compact Lie group $\tilde{G}$ satisfy $(\tilde{d}_Q) \supset G_Q \times SU(2)$ and be a nontrivial spin extension—defined by (4)—of the simple Lie group G with the quark representation G_Q. Then $G_Q \times SU(2)$ is contained in a simple subgroup of $\tilde{G} = (\tilde{d}_Q)$.

Proof: An irreducible (see Appendix) compact Lie group of $2N \times 2N$ matrices can be written in the form

$$\tilde{G} = H_1 \times H_2 \times \cdots \times H_n \times K_\varphi, \tag{1.5}$$

where H_1 are irreducible simple compact groups of $n_i \times n_i$ matrices and $K_\varphi \simeq \{e^{i\varphi}\}$.

Since $H = G_Q \times SU(2)$ is an irreducible unimodular subgroup of $\tilde{G}$, H can be represented by Theorem A1 (in the Appendix) which is easily carried over to compact groups (cf., beginning of Sec. 2) in the form

$$H = H_1' \times H_2' \times \cdots \times H_n' \quad \text{with} \quad H_i' \subset H_n \tag{1.6}$$

($H_1', \cdots, H_n'$ irreducible, not necessarily simple).

As both G_Q and $SU(2)$ are simple and as H is a representation of the direct product of two simple groups, at most two nontrivial factors can appear in (1.6). If there actually are two nontrivial ones, they must be equal to G_Q and $SU(2)$, respectively, and condition (1.4) is violated. Hence there is only one factor H_1' so that $G_Q \times SU(2) \subset H_1$ which proves the statement. We note that this implies that there is only one nontrivial H_i in (1.5) because of $n_1 = 2N$.

(e) The above requirements must be completed by a reasonable prescription to choose a "smallest" group under all possible groups $\tilde{G}$. This can be done in various ways. We propose to demand $\tilde{G}$ to be minimal with respect to a set of conditions, which means that it should have no proper subgroup also satisfying these conditions. Then for an embedding of a simple group G by Theorem 1.1 $\tilde{G}$ becomes also simple since H_1 fulfills all requirements for $\tilde{G}$. For semisimple G Theorem 1.1 fails, but we take $\tilde{G}$ as simple also in this case. Note that the interesting proposed symmetries $SU(3)$, G_2, $Sp(6)$, and B_3 are simple.

The plausible idea that the number of generators and (or) the rank of $\tilde{G} \supset G \otimes SU(2)$ should be as small as possible does not work in general, for example, in the spin extension $SU(6)$ of $SU(3)$. This is shown by noting that $SU(3) \otimes SU(2)$ is not only contained in $SU(6)$ but also in $SU(5)$—evidently "smaller" than $SU(6)$—in a different form though,

$$\begin{pmatrix} SU(3) & 0 \\ 0 & SU(2) \end{pmatrix}$$

and in direct contradiction to the quark model. Because of the success of the quark model one should not abandon it, although it raises the rank and the number of generators of $\tilde{G}$.

(f) The results of this discussion, the requirements on the symmetry group G and the embedding conditions for the determination of the spin extension $\tilde{G}$ can be summarized as follows. Let the compact Lie group G be an internal symmetry group which admits of a generalized quark model [cf., in (c)] with the irreducible group $(d_Q) = G_Q$ of $N \times N$ matrices [according to the remarks in (b) and (d) we can take G to be semisimple and G_Q to be unimodular].

Conditions for a spin extension $\tilde{G}$ of G:

(α) $\tilde{G}$ is simple and compact;

(β) $\tilde{G}$ is a group of $2N \times 2N$ matrices which contains $G_Q \times SU(2)$;

(γ) $\tilde{G}$ is minimal among the groups satisfying (α), (β).

It is easy to check that the spin extension $SU(6)$ of $SU(3)$ fulfills the above conditions. It was derived first by Gürsey and Radicati[4] by using implicitly similar arguments as above in the discussion of a special Lagrange function, but a systematic generalization of this field theoretic method to other symmetry groups seems to be complicated and more involved than a pure group theoretical treatment.

In the following sections, we give a full account of all spin extensions satisfying the conditions (α), (β), and (γ). One would expect a class of solutions $\tilde{G}$ for every given symmetry G and quark representation G_Q. However, this is not the case: we show that the above defined spin extension is already unique, with a few exceptions, for which two solutions are obtained.

2. A UNIQUENESS THEOREM

The evaluation of the postulates is straightforward. We note as a first simple consequence that $\tilde{G}$ is irreducible and unimodular, which follows from the corresponding property of $\tilde{G}_Q \times SU(2)$. Now all simple compact irreducible unimodular groups of $2N \times 2N$ matrices containing $G_Q \times SU(2)$ and satisfying the minimality condition are possible candidates for $\tilde{G}$. A full account of all these groups can be derived using the theory of Dynkin[5] and Malcev.[6] However, the Dynkin theory gives a classification for complex Lie groups with respect to a certain subgroup content and one has to make sure that results taken over from the Dynkin theory are valid also for the compact real forms of the complex Lie groups which occur in $\tilde{G}$. This is possible.

There exists a one-to-one correspondence between the irreducible representations of a compact Lie group and the irreducible analytic representations of its complex extension (cf., e.g., Ref. 7), and since the complex extensions of two different compact semisimple Lie groups do not coincide, which follows from the theorem of Weyl (cf., Ref. 8) stating that a complex Lie algebra cannot have more than one real compact form, it suffices to prove the theorem for the corresponding complex extension of the groups.

As for the notation, we use A_n, $B_n, \cdots, G_2$ interchangeably for the compact and complex groups as well as for the corresponding associated compact real and complex Lie algebras. The complex extension of the compact unimodular orthogonal group $O(n)$ is the group $O(n, C)$ of all complex unimodular orthogonal $n \times n$ matrices, for the compact symplectic group $Sp(n)$, one gets the group $Sp(n, C)$ of all complex symplectic $n \times n$ matrices and for $SU(n)$ the group $SL(n, C)$ of all complex unimodular $n \times n$ matrices.

Now we are prepared to prove the following uniqueness theorem, stating essentially that the spin extension is already fixed by the physically motivated conditions (α), (β), and (γ). A collection of some results of the Dynkin theory used in the proof of the theorem can be found in the Appendix.

Main Theorem:

(a) Let the irreducible group G_Q of unimodular $N \times N$ matrices be different from the lowest spinor representation of $B_n \simeq O(2n + 1)$, with $n \geq 3$, i.e., not given by the Dynkin diagram

$$\simeq B_n, \quad n \geq 3. \tag{2.1}$$

Then the spin extension $\tilde{G}$ is uniquely determined by the postulates (α), (β), and (γ), specifically one has for $N \geq 3$:

$$\tilde{G} = \begin{cases} SU(2N) \text{ if } G_Q \text{ has no bilinear invariant,} \\ Sp(2N) \text{ if } G_Q \text{ has a symmetric bilinear invariant,} \\ O(2N) \text{ if } G_Q \text{ has a skew-symmetric bilinear invariant.} \end{cases} \tag{2.2}$$

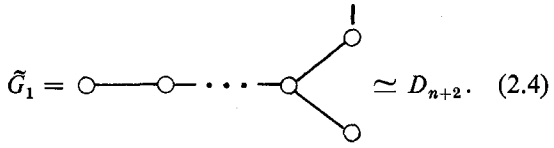
For $N = 2$ one has:

$$\tilde{G} = SU(4). \tag{2.3}$$

(b) Let G_Q be given by the lowest spinor representation of B_n, $n \geq 3$, i.e., by (2.1).

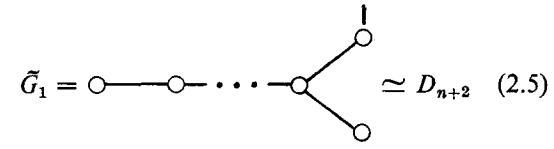
[4] F. Gürsey and L. A. Radicati, Phys. Rev. Letters **14**, 57 (1964).
[5] E. B. Dynkin, Am. Math. Soc. Transl. Ser. 2, **6**, 245 (1957).
[6] A. I. Malcev, Am. Math. Soc. Transl. Ser. 1, **9**, 172 (1950).
[7] G. Krafft, Mitt. Math. Sem. der Universität Giessen, Bd. 32 (1955).
[8] L. S. Pontrjagin, *Topologische Gruppen* (B. G. Teubner, Leipzig, 1958), Bd. II, p. 226.

(1) n even: $\tilde{G}$ is uniquely determined by (α), (β), and (γ), and is given by the lowest spinor representation of $D_{n+2} \simeq O(2n+4)$,



$$\simeq D_{n+2}. \quad (2.4)$$

(2) n odd: There are two solutions $\tilde{G}_1$ and $\tilde{G}_2$ satisfying (α), (β), and (γ), namely,

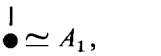


$$\simeq D_{n+2} \quad (2.5)$$

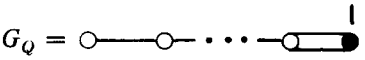
and

$$\tilde{G}_2 = \begin{cases} Sp(2^{n+1}) & \text{for } \tfrac{1}{2}n(n+1) \text{ even}, \\ O(2^{n+1}) & \text{for } \tfrac{1}{2}n(n+1) \text{ odd}. \end{cases} \quad (2.6)$$

Proof of Main Theorem: Let G_Q be an arbitrary irreducible group of unimodular $N \times N$ matrices. Because of the above correspondence between compact (real) and complex groups, Theorem A2 of the Appendix imposes severe restrictions on those irreducible real groups of $2N \times 2N$ matrices which contain $H = G_Q \times SU(2)$. Since $B_1 \simeq A_1 \simeq SU(2)$ and since the matrix group $SU(2)$ is determined by the Dynkin diagram

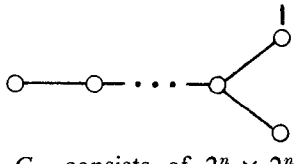


Theorem A2 implies that from the compact unimodular groups only $O(2N)$, $Sp(2N)$, and $SU(2N)$ can contain $H = G_Q \times SU(2)$, the only exception being (cf., Nr. IV_2 of Table 5 of Ref. 5) the case if G_Q is given by the lowest spinor representation of B_n for $n \geq 3$, i.e.,



$$\simeq B_n \quad (n \geq 3), \quad (2.7)$$

for in this case H is also contained in the lowest spinor representation of $D_{n+2} \simeq O(2n+4)$, i.e., in



$$\simeq D_{n+2}. \quad (2.8)$$

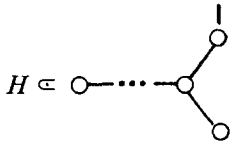
In (2.7), G_Q consists of $2^n \times 2^n$ matrices so that $N = 2^n$.

(*1*) *Proof of Part (b):*

With G_Q given by (2.1), H is a subgroup of $\tilde{G}_1$ specified by (2.4), (2.5), as is shown in (2.8). By Theorem A4, H possesses a symmetric or skew-symmetric bilinear invariant according to whether the invariant of G_Q is skew-symmetric or symmetric, because $SU(2) = Sp(2)$ has a skew-symmetric invariant. From Theorems A3 and A4 follows

$$H \subset \begin{cases} Sp(2^{n+1}) & \text{for } \tfrac{1}{2}n(n+1) \text{ even}, \\ O(2^{n+1}) & \text{for } \tfrac{1}{2}n(n+1) \text{ odd}. \end{cases} \quad (2.9)$$

(i) n even: The group given by (2.8) has a bilinear invariant which is symmetric for $\frac{1}{2}(n+2)(n+1)$ even and skew-symmetric for $\frac{1}{2}(n+2)(n+1)$ odd by Theorem A4. In the first case, $\frac{1}{2}n(n+1)$ is odd, in the second one $\frac{1}{2}n(n+1)$ is even so that by (2.9) the following proper inclusions hold:



$$\subset \begin{cases} Sp(2^{n+1}) & \text{for } \tfrac{1}{2}n(n+1) \text{ even}, \\ O(2^{n+1}) & \text{for } \tfrac{1}{2}(n(n+1) \text{ odd}. \end{cases}$$

By the minimality condition (γ), $\tilde{G}$ is realized by (2.4).

(ii) n odd: There is no bilinear invariant for (2.8) so that it is contained neither in $Sp(2^{n+1})$ nor in $O(2^{n+1})$. Conversely, it follows from Theorem A2 that neither $Sp(2^{n+1})$ nor $O(2^{n+1})$ are subgroups of (2.8). Hence one obtains the two solutions (2.5) and (2.6) for $\tilde{G}$.

(*2*) *Proof of Part (a):*

(i) $N \geq 3$: With G_Q not being given by (2.1), Theorem A2 states that the only simple compact groups which can contain $H = G_Q \times SU(2)$ are $O(2N)$, $Sp(2N)$, and $SU(2N)$. Because of the irreducibility of H, the cases $O(2N)$ and $Sp(2N)$ depend on the type of the bilinear invariant of G_Q and are mutually exclusive. If G_Q has a symmetric (skew-symmetric) invariant, H has, by Theorem A4, a skew-symmetric (symmetric) invariant. The inclusion $H \subset SU(2N)$ is always true since $SU(2N)$ is the group of all unitary unimodular matrices. The minimality condition (γ) and Theorem A3 then imply Eq. (2.2).

(ii) $N = 2$: In order to prove (2.3), it suffices to show that $SU(2) = Sp(2)$ is the only irreducible compact group of unitary unimodular 2×2 matrices. For then $G_Q = SU(2) = Sp(2)$, and (2.3) follows from Theorem A2 noting that although $SU(2) \times SU(2) = Sp(2) \times Sp(2)$ has a symmetric bilinear invariant, the inclusion $Sp(2) \times Sp(2) \subset O(4)$ is not proper since $Sp(2) \times Sp(2) = O(4)$. Now let G_Q be an irreducible compact group of unitary unimodular 2×2 matrices. G_Q can be written as a Kronecker product of irreducible compact simple matrix groups. Since there can only be one factor consisting of 2×2 matrices and since a one-dimensional representation of simple compact Lie groups is the trivial one, there remains one factor only. Hence G_Q is simple. The only simple

compact group of 2×2 matrices, however, is $SU(2) = Sp(2)$. This completes the proof.

3. DISCUSSION OF THE EMBEDDING CONDITIONS

A. The Case $G = SU(N)$

Let us consider the special case that $SU(N)$ is intrinsic symmetry group and that $SU(N)$ can be used for the construction of a generalized quark model, i.e., $G_Q = SU(N)$. Since $SU(N)$ has no bilinear invariant for $N > 2$, one obtains as spin extension of $SU(N)$ the group

$$\tilde{G} = SU(2N). \tag{3.1}$$

By (2.3), Eq. (3.1) also holds for $N = 2$. It is shown in Theorem 3.2 that $SU(N) \times SU(2)$ is even maximal subgroup of $SU(2N)$. From this it follows that for $G_Q = SU(N)$ one may drop the minimality condition (γ), since $SU(2N)$ is the largest group allowed by postulates (α), (β), and because of the maximality of $SU(N) \times SU(2)$ in $SU(2N)$ it is also the smallest. Since $SU(N)$ is simple, one may by Theorem 1.1 alternatively drop the simplicity requirement (α) for the spin extension $\tilde{G}$.

The well-known spin extension $SU(6)$ of $SU(3)$ is then uniquely determined by the postulates (α), (β), and (γ). The fact that $SU(3) \times SU(2)$ is the maximal subgroup of $SU(6)$ has also been noted in Ref. 9.

B. Dropping the Simplicity Condition for $\tilde{G}$

For a simple intrinsic symmetry group G, the simplicity of a nontrivial spin extension $\tilde{G}$ follows from Theorem 1.1 and the minimality condition (γ). It seems therefore worthwhile to investigate the consequences of dropping the simplicity condition for $\tilde{G}$; that is, to look what happens in the general case of a semisimple G if one replaces the postulate (α) by: (α') $\tilde{G}$ is semisimple and compact.

First we note that the irreducible matrix group $\tilde{G}$ can be written as

$$\tilde{G} = \tilde{G}_1 \times \cdots \times \tilde{G}_s, \quad \tilde{G}_i \text{ simple, irreducible.} \tag{3.2}$$

The group $H = G_Q \times SU(2)$ can be decomposed as

$$H = G_1 \times \cdots \times G_r \times SU(2), \qquad G_i \text{ simple, irreducible.} \tag{3.3}$$

Since $H \subset \tilde{G}$, one has by Theorem A1

$$H = H_1 \times \cdots \times H_s, \quad H_i \subset \tilde{G}_i, \quad H_i \text{ irreducible.} \tag{3.4}$$

With a suitable numbering of the $\tilde{G}_i$ and G_j, the H_i can be written as follows:

$$H_1 = G_1 \times \cdots \times G_{i_1} \times SU(2), \tag{3.5}$$

$$H_{K+1} = G_{i_{K+1}} \times \cdots \times G_{i_{K+1}}, \quad K \geq 1. \tag{3.6}$$

In (3.5), there appears at least one G_i since the spin extension $\tilde{G}$ is assumed to be nontrivial.

Now, we note that one has the following inclusions:

$$G_Q \times SU(2) \subset \tilde{G}_1 \times H_2 \times \cdots \times H_s \subset \tilde{G}_1 \times \tilde{G}_2 \times \cdots \times \tilde{G}_s = \tilde{G}. \tag{3.7}$$

The group $\tilde{G}_1 \times H_2 \times \cdots \times H_s$ fulfills postulates (α') and (β); hence the minimality condition (γ) implies

$$\tilde{G} = \tilde{G}_1 \times H_2 \times \cdots \times H_s, \tag{3.8}$$

where

$$\tilde{G}_1 \supset G_1 \times \cdots \times G_{i_1} \times SU(2).$$

Therefore, one actually has a spin extension of the group

$$G' = G_1 \times \cdots \times G_{i_1} \tag{3.9}$$

by the simple group $\tilde{G}_1$. The remaining factors $G_{i_1+1}, \cdots, G_s$ are not altered. We can formulate this result as follows.

Theorem 3.1: For an internal symmetry group G with a decomposition of $G_Q \times SU(2)$ given by (3.3), one obtains the set of all spin extensions which satisfy the postulates (α'), (β), and (γ) by first considering the spin extensions of all products $G_{i_1} \times \cdots \times G_{i_K}$ by simple groups $\tilde{G}_i$ which are determined by the main theorem, and then forming the Kronecker product of $\tilde{G}_i$ and the remaining G_j's. If the symmetry group G is not simple, there are in general several spin extensions satisfying (α'), (β), and (γ).

C. Sharpening of the Minimality Condition (γ)

We now turn to the question in which cases the minimality condition can be sharpened in a plausible way. For a spin extension $\tilde{G}$ of G constructed by the postulates (α), (β), and (γ), there may in general be groups K which lie between $\tilde{G}$ and $G_Q \times SU(2)$, that is, one has the proper inclusions

$$\tilde{G} \supset K \supset G_Q \times SU(2). \tag{3.10}$$

A compact group K which satisfies (3.10) cannot be simple by postulate (γ). From the physical point of view it would be most satisfactory, if the number of generators of the constructed spin extension is as low as possible, i.e., if there exists no such group. In this case $G_Q \times SU(2)$ is maximal subgroup of $\tilde{G}$, and it is plausible to ask if the minimality condition (γ) can be replaced by: (γ') $G_Q \times SU(2)$ is maximal subgroup of $\tilde{G}$.

Obviously condition (γ') implies the minimality condition (γ), since if (γ') holds there is no group lying between $\tilde{G}$ and $G_Q \times SU(2)$. Conversely, the postulates (α), (β), and (γ) determine the spin extension $\tilde{G}$ with one exception uniquely. Hence it may happen that there is no group satisfying the stronger conditions (α), (β), and (γ') which have a solution for a given $G_Q \times SU(2) = H$ if and only if H is a maximal subgroup of the group $\tilde{G}$ constructed from

[9] R. Hermann, *Lie Groups for Physicists* (W. A. Benjamin, Inc., New York, 1966).

the postulates (α), (β), and (γ). The following theorem completely answers this question.

Theorem 3.2:

(a) Let the group G_Q be different from the lowest spinor representation of B_n, $n \geq 3$, i.e., not given by (2.1), and let $\tilde{G}$ be the group uniquely determined by the postulates (α), (β), and (γ). Then $G_Q \times SU(2)$ is maximal subgroup of $\tilde{G}$ if and only if $G_Q = SU(N)$, $O(N)$, $Sp(N)$.

(b) Let G_Q be given by the lowest spinor representation of B_n, i.e., by (2.1). Then $G_Q \times SU(2)$ is maximal subgroup of $\tilde{G}_1$ in (2.4) and (2.5), but not of $\tilde{G}_2$ in (2.6).

Proof:

(a) Sufficiency: If $G_Q = SU(N)$, $O(N)$, and $Sp(N)$, then by the Main Theorem, $\tilde{G} = SU(2N)$, $Sp(2N)$, $O(2N)$ respectively. Assume that $G_Q \times SU(2)$ is not maximal in $\tilde{G}$. Then there is a group K lying between $G_Q \times SU(2)$ and $\tilde{G}$. Going over to the complex extensions one gets for $SU(N)$

$$SL(N, C) \times SL(2, C) \subset K_C \subset SL(2N, C), \quad (3.11)$$

where K_C denotes the complex extension of K. If K is different from $G_Q \times SU(2)$ and $\tilde{G}$, so are their complex extensions. But by Theorem A5, $SL(N, C) \times SL(2, C)$ is maximal subgroup of $SL(2N, C)$, in contradiction to the assumption. For $O(N)$ and $Sp(N)$ one proceeds on similar lines, applying Theorem A6.

Necessity: Assume G_Q different from $SU(N)$, $O(N)$, $Sp(N)$ [and (2.1)]. If G_Q has no bilinear invariant, $\tilde{G} = SU(2N)$ and one has the proper inclusions

$$G_Q \times SU(2) \subset SU(N) \times SU(2) \subset SU(2N). \quad (3.12)$$

If G_Q has a symmetric (or skew-symmetric) bilinear invariant, one has $\tilde{G} = Sp(2N)$ [or $O(2N)$, respectively] and the proper inclusions

$$G_Q \times SU(2) \subset O(N) \times SU(2) \subset Sp(2N) \quad \text{(symmetric)}, \quad (3.13)$$

$$G_Q \times SU(2) \subset Sp(N) \times SU(2) \subset O(2N) \quad \text{(skew-symmetric)}. \quad (3.14)$$

Hence, in this case $G_Q \times SU(2)$ is not a maximal subgroup of $\tilde{G}$.

(b) Let G_Q be given by (2.1), the lowest spinor representation of B_n, $n \geq 3$.

(i) Let $\tilde{G}_1$ be given by (2.4) or (2.5), the lowest spinor representation of D_{n+2}. As $\tilde{G}_1$ is simple and different from $Sp(2N)$ and $O(2N)$, one can, after complexification, apply Theorem A2. Any group lying between $G_Q \times SU(2)$ and $\tilde{G}_1$ is irreducible and unimodular. These subgroups of $\tilde{G}_1 \simeq D_{n+2}$ are given by Theorem A2. Inspection of Table AI shows there is none which has $G_Q \times SU(2) \simeq B_n \otimes A_1$ as proper subgroup. Hence in this case $G_Q \times SU(2)$ is maximal subgroup of $\tilde{G}_1$.

(ii) Let $\tilde{G}_2$ be given by (2.6), i.e., $\tilde{G}_2 = Sp(2^{n+1})$ or $O(2^{n+1})$. Then G_Q possesses a bilinear invariant, and one has the proper inclusions

$$G_Q \times SU(2) \subset O(2^n) \times SU(2) \subset Sp(2^{n+1})$$

or

$$G_Q \times SU(2) \subset Sp(2^n) \times SU(2) \subset O(2^{n+1}).$$

Hence in this case $G_Q \times SU(2)$ is not maximal subgroup.

D. On the Choice of the Generalized Quark Representation G_Q

The representation of the symmetry group G used in the formulation of the postulates (α), (β), and (γ) is essential for the determination of the spin extension $\tilde{G}$. Demanding the representation G_Q to be as low dimensional as possible does not, in general, determine G_Q uniquely. It can happen that there exist several inequivalent representations of G with equal dimensionality which could be used for a generalized quark model. One could, for instance, take the complex conjugate representation $\bar{G}_Q$ instead of G_Q. So the question arises how a different choice of G_Q affects the result of the spin extension.

A representation $D(g)$ of a Lie group G has a nondegenerate bilinear invariant if and only if D is equivalent to the contragredient representation $D(g^{-1})^T$ (Ref. 5 Theorem 0.19). For compact groups, the contragredient and complex conjugate representations are equivalent. Since in the determination of the spin extension $\tilde{G}$ in the Main Theorem only the nature of the bilinear invariant of G_Q enters and since either G_Q and $\bar{G}_Q$ have the same kind of invariant or both G_Q and $\bar{G}_Q$ have no invariant at all, one notes that the result $\tilde{G}$ of the Main Theorem is independent of whether one uses G_Q or $\bar{G}_Q$.

We now discuss the more general case that G_{Q_1} and G_{Q_2} are two inequivalent quark representations of equal dimensionality where furthermore $G_{Q_2} \neq \bar{G}_{Q_1}$. The corresponding spin extensions are denoted by $\tilde{G}_1$ and $\tilde{G}_2$. By the same argument as used before, it follows that $\tilde{G}_1 = \tilde{G}_2$ if G_{Q_1} and G_{Q_2} have the same kind of bilinear invariant or if both matrix groups have no bilinear invariant. $\tilde{G}_1 \neq \tilde{G}_2$ if and only if one of the groups has a bilinear invariant while the other has none or if one has a symmetric, the other one a skew-symmetric invariant.

This case cannot occur for G simple and G_Q the lowest nontrivial representation, since simple groups have at most two inequivalent lowest nontrivial representations which, however, are contragredient to each other. This follows from the formulas for the dimensions of the irreducible representations (cf. Ref. 5).

In the case $G = B_n$ or D_n one could try to take the lowest spinor representation for the generalized quark

model. For B_n there is only one such representation while for D_n there exist two inequivalent ones. For D_{2n+1} these two lowest spinor representations are contragredient, since according to Theorem A4 there are no bilinear invariants, and consequently, the representations are inequivalent to their respective contragredient ones. Hence they must be contragredient to each other because there are only two such lowest spinor representations. So for D_{2n+1} the spin extension is not affected. For D_{2n} both of the lowest spinor representations have the same kind of invariant, according to Theorem A4, so that $\tilde{G}_1 = \tilde{G}_2$.

For nonsimple symmetry groups the situation becomes complex. In any case, for given symmetry groups one has to choose the quark representations according to their physical usefulness.

4. APPLICATIONS

The explicit construction of the spin extension $\tilde{G}$ for special internal symmetry groups G consists of two steps: (1) A discussion of the existence of a generalized quark model and the determination of G_Q. (2) Calculation of the type of bilinear invariant of G_Q using Table AII.

For the recently discussed groups $SU(3)$,[10] G_2,[11] and $Sp(6)$,[12] a quark model holds. For B_3 (Ref. 13) a quark representation which includes hadrons as well as leptons is incompatible with the identification given in Ref. 13 and does not exist. This can be shown[14] in a similar way as in the discussion of $SU(2) \times SU(2)$ in Sec. 1. However, in this case a quark representation involving hadrons only is possible. With the invariants from Table AII, we obtain by the Main Theorem the results of Table I.

To justify this extension procedure, one has to determine the G_Q subgroup content of a given irreducible representation of $\tilde{G}$, i.e., the branching rules $\tilde{G} \downarrow G_Q \times SU(2)$. For the mathematical methods of this reduction we refer to Ref. 15. A detailed discussion of the groups $Sp(6)$, G_2, and B_3 is given elsewhere.[14,16] It turns out that the spin extension of the van Hove triplet model $Sp(6)$, given by $O(12)$, yields the quite satisfactory result that, similar to the spin extension $SU(6)$ of $SU(3)$, the scalar meson octet and the nine vector mesons as well as the baryon octet and decuplet can be fitted in one irreducible $O(12)$ representation with no more additional particles necessary than already in the $Sp(6)$-scheme. A comparison with a spin extension of $Sp(6)$ obtained by a different method[17] can be found in Ref. 18.

TABLE I. Spin extension of some symmetry groups.

Symmetry group	Spin extension
$SU(3)$	$SU(6)$
$Sp(6)$	$O(12)$
G_2	$Sp(14)$
$B_3 \simeq O(7)$	$Sp(14)$

For G_2, which has been discussed as an alternative for $SU(3)$, the spin extension $Sp(14)$ leads to a large number of additional particles if one tries to classify the known G_2-multiplets according to $Sp(14)$. The same is true for the spin extension of B_3 for which one obtains also $Sp(14)$, with a different physical interpretation though, since here one has to determine the $B_3 \otimes SU(2)$-content of $Sp(14)$-representations.

5. LORENTZ EXTENSION

Looking for a group $\tilde{G}$ comprising the covering $\tilde{L}$ of the inhomogeneous Lorentz group and an internal symmetry group G, it suggests itself to start from the spin extension $\tilde{G}$ of G and enlarge it such that it contains $\tilde{L}$. There are many ways of doing this. In the following, we discuss two possible approaches which essentially differ in the interpretation of the group $SU(2)$ which was used in constructing $\tilde{G}$. (A) $SU(2)$ is identified as the covering of the rotation group, originally suggested in the case of $SU(6) \supset SU(3) \times SU(2)$ by Fulton and Wess and Rühl.[19] This (FWR) approach leads to a large number of additional translations and to possible difficulties with unitarity. (B) $SU(2)$ here also transforms like the rotation group under homogeneous Lorentz transformations, but it is not identified with physical rotations. Furthermore, $SU(2)$ is assumed to commute with the space–time translations. For $SU(6)$ this idea was proposed by Budini and Fronsdal,[20] it leads to difficulties in the field-theoretic formulation which are connected with locality.

A generalization of the interesting $\tilde{U}(12)$-theory of

[10] M. Gell-Mann and Y. Ne'eman, *The Eightfold Way* (W. A. Benjamin, Inc., New York, 1964).

[11] R. E. Behrends and L. F. Landowitz, Phys. Rev. Letters **11**, 296 (1963); Y. Dothan and H. Harari, Nuovo Cimento **32**, 498 (1964); R. E. Behrends, Phys. Rev. **142**, 1101 (1966).

[12] H. Bacry, J. Nuyts, and L. van Hove, Phys. Letters **9**, 279 (1964); L. van Hove, in *Preludes in Theoretical Physics*, A. de-Shalit, H. Feshbach, and L. van Hove, Eds. (North-Holland Publishing Company, Amsterdam, 1966).

[13] C. Lovelace, Nuovo Cimento **37**, 180 (1965).

[14] G. C. Hegerfeldt, Ph.D. thesis, University of Marburg (1965).

[15] H. D. Doebner and G. C. Hegerfeldt, to appear in the series *Forschungberichte*, zentralstelle für Atomenergie dokumentation, Frankfurt/Main, Germany.

[16] H. D. Doebner and G. C. Hegerfeldt, Z. Phys. **199**, 369 (1967),

[17] V. G. Kadyshevsky, A. N. Tavkhelidze, and I. T. Todorov, Phys. Letters **15**, 180 (1965).

[18] H. D. Doebner and G. C. Hegerfeldt, in *Elementary Particles*, P. Urban, Ed., (Julius Springer-Verlag, Berlin, 1966), p. 369.

[19] T. Fulton and J. Wess, Phys. Letters **14**, 57 (1965); W. Rühl, *ibid.* **13**, 349 (1964).

[20] P. Budini and C. Fronsdal, Phys. Rev. Letters **14**, 968 (1965); C. Fronsdal, Trieste Lectures (1965); V. Schladming, Lectures (1966).

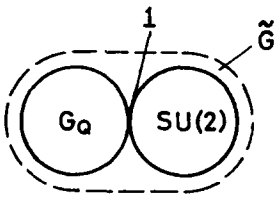


FIG. 1. Embedding of $SU(2)$.

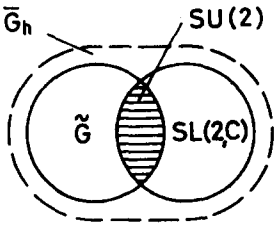


FIG. 2. Embedding of $SL(2, C)$.

Delbourgo *et al.*[21] seems to be not so straightforward and is not discussed here. We note that $\tilde{U}(12)$ does not obey our minimality condition which shows that one might lose physical information by this restriction.

A. Generalization of the FWR Approach

To construct $\tilde{G}$ starting from the identification of $SU(2)$ as covering of the physical rotation group, we carry over to $\tilde{G}$ a procedure[22] already used for $SU(6)$ by considering the groups H of $2N \times 2N$ matrices which contain $\tilde{G}$ and $SL(2, C)$ in such a way that their intersection corresponds to $SU(2)$.

By (1.2), $SU(2)$ was embedded in $\tilde{G}$ in the form $1_N \times SU(2)$, the Kronecker product of the $N \times N$ unit matrix 1_N with $SU(2)$, so that $SL(2, C)$ should lie in H as $1_N \times SL(2, C)$. We therefore demand

$$\tilde{G} \subset H, \quad 1_N \times SL(2, C) \subset H, \tag{5.1a}$$

$$\tilde{G} \cap 1_N \times SL(2, C) = 1_N \times SU(2), \tag{5.1b}$$

and take for $\tilde{G}_h$ a group minimal with respect to (5.1). If two groups H_1, H_2 fulfill (5.1), so does their intersection. Therefore, $\tilde{G}_h$ is uniquely determined and is equal to the intersection of all groups H satisfying (5.1). The difference to the problem of constructing a spin extension is contained in (5.1b), expressed graphically by Figs. 1 and 2.

Theorem 5.1: The smallest group of $2N \times 2N$ matrices satisfying Eqs. (5.1) is the complex extension $\tilde{G}_C$ of $\tilde{G}$,

$$\tilde{G}_h = \tilde{G}_C. \tag{5.2}$$

Proof: Letting the parameters of the compact group $\tilde{G}$ become complex, one gets the complex extension $\tilde{G}_C$ of $\tilde{G}$ which corresponds to the complex extension of the Lie algebra of $\tilde{G}$ (cf., the proof of the Main Theorem). The linear group $H = \tilde{G}_C$ satisfies (5.1) since it contains the complex extension $SL(2, C)$ of $SU(2)$. On the other hand $\tilde{G}$ is, as a real form of the simple complex group $\tilde{G}_C$, a maximal subgroup of $\tilde{G}_C$ (cf., Ref. 5, Theorem 1.6) so that $\tilde{G}_C$ is the smallest group H satisfying (5.1) since $\tilde{G}$ itself does not fulfill (5.1).

For the case of $\tilde{G} = SU(6)$, Theorem 5.1 gives the familiar result $\tilde{G}_h = SL(6, C)$.

In order to take the translation group T_4 into account, physical interpretation suggests the requirement that T_4 should be invariant under $\tilde{G}_h$ and transform under Lorentz transformations in the usual way. In general, however, this is impossible as is well known for $SL(6, C)$ (Ref. 19). For let g be an element of $\tilde{G}_h$ and t be an element of an arbitrary translation group T. Invariance of T then means that gtg^{-1} is again in T. The mapping

$$\alpha(g): t \to gtg^{-1} \tag{5.3}$$

of T onto T is an automorphism of T, and the mapping $g \to \alpha(g)$ is a homomorphism of $\tilde{G}_h$ into the group of automorphisms of T. This group is just given by the set of nonsingular linear transformations of T, bearing in mind that T is also real vector space.

Hence $g \to \alpha(g)$ is a representation of $\tilde{G}_h$ by linear transformations of this real vector space T. Now, $\tilde{G}_h$ in general does not have a real 4-dimensional representation in which the representation matrices of $1_N \times SL(2, C)$ act as Lorentz transformations. A possible way out of this difficulty is to embed T_4 in a larger group T and to demand the invariance of T under $\tilde{G}_h$, naturally choosing T as small as possible. The dimension of T is uniquely determined by $\tilde{G}_h$. The familiar 36-dimensional translation group T constructed for $SL(6, C)$ is a special case of the following result (parity is not taken into account).

Theorem 5.2: For $\tilde{G}_h = \tilde{G}_C = SL(2N, C), O(2N, C)$, $Sp(2N, C)$, the smallest translation group T which is invariant under $\tilde{G}_h$ and which contains a 4-dimensional subgroup T_4 on which the homogeneous Lorentz transformations embedded in $\tilde{G}_h$ act as on space–time translations, is $(2N)^2$ dimensional.

Proof: The dimension of T cannot be smaller than that of the lowest nontrivial real representation of $\tilde{G}_h$. For $\tilde{G}_h = SL(2N, C)$, $O(2N, C)$, $Sp(2N, C)$, this representation is $(2N)^2$ dimensional, and there is one and only one real representation of this dimension, as is easily deduced from Ref. 23. So one only has to show that in T the space–time translations can be embedded in the required way. To do this, we explicitly construct the representation of $\tilde{G}_h$.

[21] R. Delbourgo, A. Salam, and J. Strathdee, Nuovo Cimento **36**, 689 (1965); R. Delbourgo, A. Salam, and J. Strathdee, Proc. Roy. Soc. (London) **A284**, 146 (1965).

[22] L. Michel and B. Sakita, Ann. Inst. Henri Poincaré **2**, 167 (1965).

[23] I. Schur, Sitzber. Preuss. Akad. Wiss. Berlin, 100 (1928); R. Brauer, *ibid.*, 626 (1929); E. Mohr, Ph.D. thesis, Göttingen (1933).

The set $\{h_{2N}\}$ of all Hermitian $2N \times 2N$ matrices forms a real $(2N)^2$ dimensional linear space. In $\{h_{2N}\}$ we embed T_4 as

$$T_4 \sim 1_N \times \{h_2\}. \tag{5.4}$$

The action of $\bar{G}_h = \tilde{G}_C$ on $\{h_{2N}\}$ we define by

$$h_{2N} \xrightarrow[y_{2N}]{} y_{2N} \cdot h_{2N} \cdot y_{2N}^*, \tag{5.5}$$

where the $2N \times 2N$ matrix y_{2N} is an element of $\bar{G}_h$, y_{2N}^* being its Hermitian conjugate. The matrix on the right-hand side of (5.5) is obviously also Hermitian. One easily checks that (5.5) determines a nontrivial representation of $\bar{G}_h$ by linear transformations of the real linear space $\{h_{2N}\}$. For y_{2N} in $1_N \times \{y_2\}$,

$$y_{2N} = 1_N \times y_{2N} \tag{5.6}$$

with y_2 in $SL(2, C)$, and for h_{2N} in $1_N \times \{h_2\} \simeq T_4$ one finds from (5.5),

$$1_N \times h_2 \xrightarrow[1_N \times y_2]{} 1_N \times (y_2 h_2 y_2^*). \tag{5.7}$$

This is the action of the Lorentz transformation y_2 if one realizes the group of space-time translations by $\{h_2\}$. Q.E.D.

By combination of Theorems 5.1 and 5.2, we obtain as Lorentz extension

$$\bar{G} = \begin{cases} T_{(2N)^2} \underset{s}{\otimes} SL(2N, C) & \text{for } \tilde{G} = SU(2N), \\ T_{(2N)^2} \underset{s}{\otimes} O(2N, C) & \text{for } \tilde{G} = O(2N), \\ T_{(2N)^2} \underset{s}{\otimes} Sp(2N, C) & \text{for } \tilde{G} = Sp(2N). \end{cases} \tag{5.8}$$

The semidirect product is specified by stating that in the realization of $T_{(2N)^2}$ as the real linear space $\{h_{2N}\}$ of Hermitian $2N \times 2N$ matrices the action of the second factor on $T_{(2N)^2}$ is given by (5.5).

The homogeneous Lorentz extension $\bar{G}_h$ of G can also be obtained without using the spin extension $\tilde{G}$ as intermediate step if one adapts the postulates (α), (β), and (γ) by a slight alteration to the covering $SL(2, C)$ of the homogeneous Lorentz group.

In analogy to the relationship between the complex Lie group $SL(2, C)$ and $SU(2)$ we demand that:

$(\bar{\alpha})$ $\bar{G}_h$ is simple and complex,

$(\bar{\beta})$ $\bar{G}_h$ is a group of $2N \times 2N$ matrices which contains $G_Q \times SL(2, C)$,

$(\bar{\gamma})$ $\bar{G}_h$ is minimal among the groups which satisfy $(\bar{\alpha})$ and $(\bar{\gamma})$.

$(\bar{\alpha})$ and $(\bar{\beta})$ imply that $\bar{G}_h$ contains $G_{QC} \times SL(2, C)$ with G_{QC} the complex extension of G_Q. So the theorems of the Appendix can be used to determine $\bar{G}_h$. A practically literal repetition of the arguments leading to the Main Theorem for spin extensions renders its generalization to homogeneous Lorentz extensions. The only change is that instead of $\tilde{G}$ its complex extension $\tilde{G}_C$ appears. Therefore one obtains by $(\bar{\alpha})$, $(\bar{\beta})$, and $(\bar{\gamma})$ for $\bar{G}_h$ the same solutions [Eq. (5.2)]

$$\bar{G}_h = \tilde{G}_C$$

as in Theorem 5.1.

The postulates $(\bar{\alpha})$, $(\bar{\beta})$, and $(\bar{\gamma})$ do not possess the same direct physical interpretation as (α), (β), and (γ). This is a disadvantage and indicates that the Lorentz extensions constructed above might be unphysical.

B. Generalization of the Budini–Fronsdal Approach

In terms of Lie algebras, one formulates the assumption that the subgroup $SU(2)$ of the spin extension $\tilde{G}$ transforms under homogeneous Lorentz transformations like the covering of the rotation group by associating a set of generators $\{I_i', i = 1, 2, 3\}$ with $SU(2)$, a corresponding set I_i with rotations and a set $L_{\nu\mu}$ with the homogeneous Lorentz group and demanding that the commutators of $L_{\nu\mu}$ with I_i and $L_{\nu\mu}$ with I_i' are described by the same structure constants. Then it can be shown[20] that $\tilde{G}$ must be enlarged to a group S which contains not only $SU(2)$ but also the group $SL(2, C)$ isomorphic, though physically not identical to the homogeneous Lorentz group.

For this group S, there is a multitude of possible candidates; we note, however, that the situation is similar to that in A. So, repeating the argument for the construction of $\bar{G}_h$, one finds

$$S = \tilde{G}_C, \tag{5.9}$$

the complex extension of $\tilde{G}$.

To combine S with the covering $\bar{L}$ of the inhomogeneous Lorentz group L, one now takes a semidirect product

$$\bar{G} = S \underset{s}{\otimes} \bar{L}, \tag{5.10}$$

which is specified in general by a homomorphic mapping of

$$\bar{L} \simeq T_4 \underset{s}{\otimes} SL(2, C)$$

into the group Aut S of automorphisms of S, and in this case by the following mapping into the group int (Aut S) of inner automorphisms of S.

Let y be a homogeneuos Lorentz transformation, denote by y' the corresponding element of $SL(2, C) \subset S$, and let int (y') be the element of Aut S generated by y' and acting in S as int $(y')\, g = y' g y'^{-1}$.

Then the mapping $\bar{L} \to$ Aut S is given by

$$y \to \text{int}\,(y') \quad \text{and} \quad T_4 \to 1. \tag{5.11}$$

This is a formulation of the statement that $SU(2) \subset S$ and $SL(2, C) \subset S$ transform like the rotation and Lorentz group respectively and commute with the translations.

ACKNOWLEDGMENT

This paper was supported in part by the Bundesministerium für wissenschaftliche Forschung, Bad Godesberg, Germany.

APPENDIX

In this appendix we present some results on irreducible matrix groups due to Dynkin[5] and Malcev[6] which were used in the proofs above.

A linear or matrix group is called irreducible if there is no proper nontrivial subspace of the space in which the linear transformations operate which is left invariant by each group element. In the following, all groups are connected complex Lie groups of matrices.

Theorem A1 (Theorem 2.2 of Ref. 5): Let

$$G^* = G_1^* \times \cdots \times G_k^*, \tag{A1}$$

where $G_1^*, \cdots, G_k^*$ are irreducible simple groups of unimodular matrices. Every irreducible subgroup G of G^* can be represented in the form

$$G = G_1 \times \cdots \times G_k, \tag{A2}$$

where $G_1, \cdots, G_k$ are irreducible groups and $G_i \subset G_i^*$.

Two inclusions $G \subset G^*$ and $G_1 \subset G_1^*$ are said to belong to the same "type" if G is equivalent to G_1 and G^* equivalent to G_1^*. The inclusions $G \subset G^*$ and $G_1 \subset G_1^*$ are called "equivalent" if there exists an isomorphic mapping of the space on which the first pair of groups acts onto the space on which the second pair acts such that the first pair of groups is carried by the mapping into the second. In general, a class of inclusions of the same type can split into several classes of equivalent inclusions.

Theorem A2 (Theorem 2.3 of Ref. 5):

(a) A complete classification of all inclusion types $G \subset G^*$ where G and G^* are irreducible groups of unimodular linear transformations and where G^* is simple and distinct from $SL(N)$, $Sp(N)$, and $O(N)$, is given by Table 5 of Ref. 5. All cases of nonsimple G—only these are needed in the preceding proofs—are listed in Table AI.

(b) Every inclusion type among irreducible groups of unimodular linear transformations of Table AI and with $SL(N, C)$, $Sp(N, C)$, and $O(N, C)$ as containing group G^* consists of equivalent inclusions.

There remains to classify the irreducible complex connected subgroups of $SL(N, C)$, $O(N, C)$, and $Sp(N, C)$. The case of $SL(N, C)$ is trivial since every group of unimodular $N \times N$ matrices is contained in $SL(N, C)$.

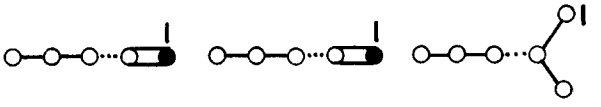
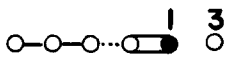
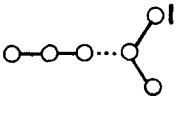
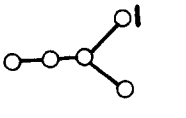
$O(N, C)$ and $Sp(N, C)$, respectively, consist of all unimodular linear transformations A of the N-dimensional complex vector space R^N which leave invariant a symmetric or, respectively, a skew-symmetric nondegenerate bilinear form $Q(x, y)$,

$$\begin{aligned} Q(Ax, Ay) &= Q(x, y) \quad \text{for all} \quad x, y \in R^N, \\ Q(x_0, y) &= 0 \qquad\quad \text{for all} \quad y \in R^N \rightharpoonup x_0 = 0. \end{aligned} \tag{A3}$$

The specific choice of the form $Q(x, y)$ is inessential because two such forms, which are either both symmetric or both skew-symmetric, can be transformed into each other.

Theorem A3 (Ref. 5): A group G of unimodular linear transformations is contained in $O(N, C)$ or $Sp(N, C)$ if and only if it has an invariant nondegenerate symmetric or, respectively, a skew-symmetric bilinear form. If G is irreducible, these two possibilities are mutually exclusive (since an irreducible

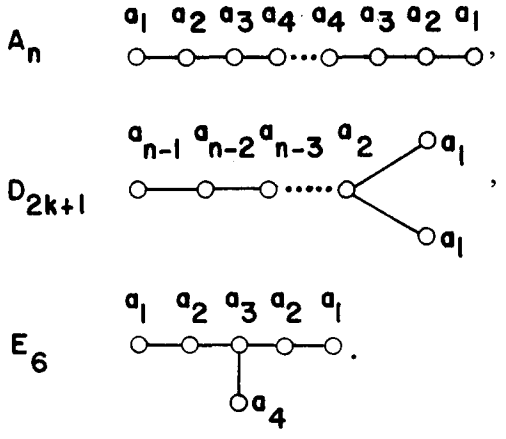
TABLE AI. Inclusion types for nonsimple irreducible G and simple $G^* \neq SL(N, C)$, $O(N, C)$, and $Sp(N, C)$.

Nr	Inclusion type $G \subset G^*$	Dynkin diagram of G	Dynkin diagram of G*
IV_2	$(B_{n_1} \cdot B_{n_2} \subset D_{n_1+n_2+1})$ $n_1 \geq 1$, $n_2 \geq 1$, $n_1+n_2 \geq 4$		
IV_3	$(B_n \cdot A_1 \subset D_{n+3})$ $n \geq 2$		
IV_5	$(A_1 \cdot A_1 \subset D_5)$		

group cannot have both a nondegenerate symmetric and a nondegenerate skew-symmetric form).

We note that if a compact group G of unitary matrices has an invariant form $Q(x, y)$ then the complex extension of G also leaves $Q(x, y)$ invariant since the invariance of $Q(x, y)$ is an algebraic identity for the matrix elements which are analytic functions of the group parameters. On the other hand, if the complex extension of G leaves $Q(x, y)$ invariant, so does G.

Theorem A4 (Ref. 6): All simple irreducible groups of the type B_n, C_n, D_{2k}, G_2, F_4, E_7, E_8 have bilinear invariants. Irreducible groups of type A_n, D_{2k+1}, E_6 have bilinear invariants if and only if the numbers in the associated Dynkin diagrams are symmetrically distributed as in (A4).



(A 4)

A semisimple irreducible group has a bilinear invariant if and only if the connected parts of the associated Dynkin diagram correspond to simple groups with bilinear invariants. In order to see if the invariant is symmetric or skew-symmetric, one multiplies each Dynkin number of a Dynkin diagram corresponding to a simple group by the respective number given in Table AII and adds the products. The invariant is symmetric or skew-symmetric according to whether the sum of these products is even or odd.

For the discussion in Sec. 3 some theorems on maximal subgroups of $SL(N, C)$, $O(N, C)$, and $Sp(N, C)$ are needed.

Theorem A5 (Theorem 1.3, Ref. 5): The set of matrices

$$SL(s, C) \times SL(t, C) \quad (st = N, 2 \leq s \leq t) \tag{A5}$$

is a maximal subgroup of $SL(N, C)$. The irreducible nonsimple maximal subgroups of $SL(N, C)$ are exhausted by the subgroups of this type (to within conjugacy).

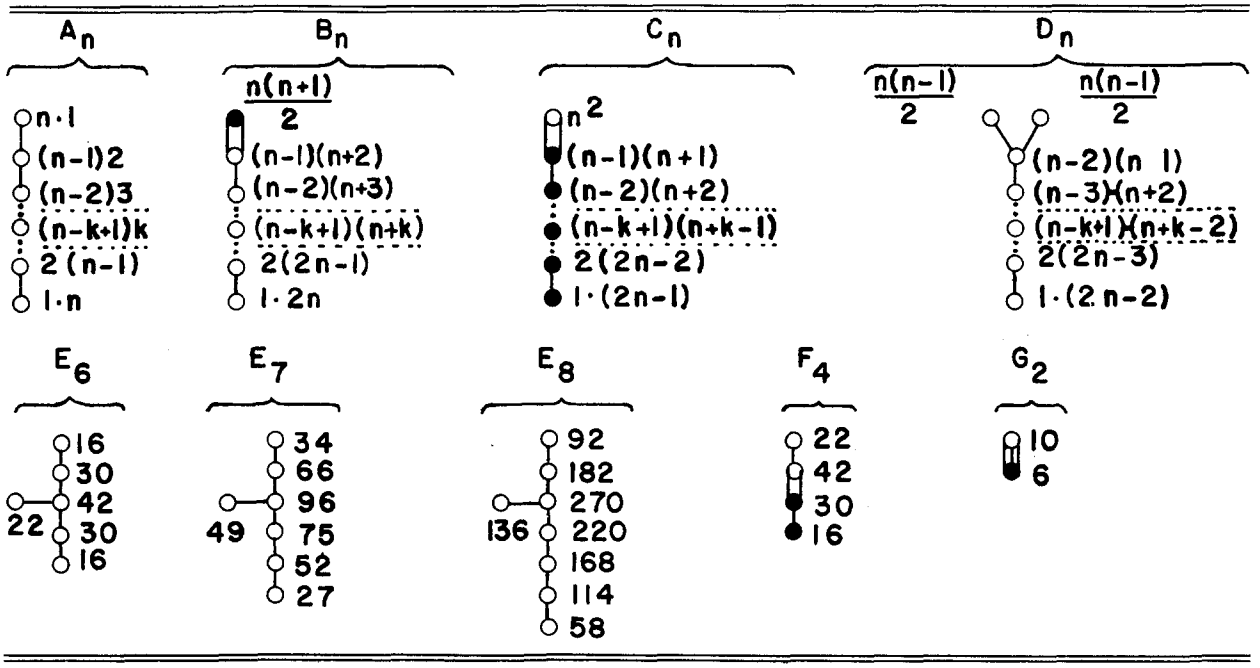
Theorem A6 (Theorem 1.4, Ref. 5): The set of matrices

$$Sp(s, C) \times O(t, C) \quad (st = N; s \geq 2, t \geq 3, t \neq 4; \text{ or } s = 2, t = 4) \tag{A6}$$

is maximal subgroup of $Sp(N, C)$. Every irreducible nonsimple subgroup of $Sp(N, C)$ is conjugate in $Sp(N, C)$ to one of the groups of the form (A6). The set of matrices

$$Sp(s, C) \times Sp(t, C) \quad (st = N, 2 \leq s \leq t) \tag{A7}$$

TABLE AII. Calculation of the bilinear invariant (by Theorem A4).



and

$$O(s, C) \times O(t, C) \quad (st = N, 3 \leq s \leq t, \quad s, t \neq 4) \tag{A8}$$

are maximal subgroups of $O(N, C)$. Every irreducible nonsimple subgroup of $O(N, C)$ is conjugate in the group of all orthogonal matrices (proper and improper) to one of the groups of the form (A7) or (A8).

We note that $O(2, C)$ is Abelian and hence reducible and that

$$O(4, C) = Sp(2, C) \times Sp(2, C). \tag{A9}$$

From this one has for $s \neq 2$ the following proper inclusions:

$$Sp(s, C) \times O(4, C) = Sp(s, C) \times Sp(2, C) \times Sp(2, C) \subset O(2s, C) \times Sp(2, C) \subset Sp(4s, C).$$

Hence $Sp(s, C) \times O(4, C)$ is not maximal in $Sp(4s, C)$ [cf., (6)]. In the same way one has for $s \geq 3$

$$O(s, C) \times O(4, C) = O(s, C) \times Sp(2, C) \times Sp(2, C) \subset Sp(2s, C) \times Sp(2, C) \subset O(4s, C),$$

therefore, the restriction $s, t \neq 4$ in (A8). In (A7) one has for $s = t = 2$:

$$Sp(2, C) \times Sp(2, C) = O(4, C),$$

so that $Sp(2, C) \times Sp(2, C)$ is no proper subgroup of $O(4, C)$.

Summary of the Appendix: Let U be an irreducible connected complex Lie group of unimodular $N \times N$ matrices. A simple complex Lie group of $N \times N$ matrices which contains U is in general equal to $SL(N, C)$, $O(N, C)$, or $Sp(N, C)$, the exceptions being specified by Theorem A2. U is subgroup of $O(N, C)$ or $Sp(N, C)$ if and only if it has a symmetric or, respectively, a skew-symmetric bilinear invariant, the form of which is determined by Theorem A4. U is a maximal connected subgroup (up to conjugacy)

of $SL(N, C)$ if and only if

$$U = SL(s, C) \times SL(t, C), \quad st = N, \quad 2 \leq s \leq t;$$

of $Sp(N, C)$ if and only if

$$U = Sp(s, C) \times O(t, C), \quad st = N, \quad s \geq 2, \quad t \geq 3, \quad t \neq 4; \quad \text{or} \quad s = 2, \quad t = 4;$$

of $O(N, C)$ if and only if

$$U = O(s, C) \times O(t, C), \quad st = N, \quad 3 \leq s \leq t, \quad s, t \neq 4$$

or

$$U = Sp(s, C) \times Sp(t, C), \quad st = N, \quad 2 \leq s \leq t.$$

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Determination of Weight Factors in Linked-Cluster Expansions for Lattice Systems

JOHN W. ESSAM
Westfield College, London, England

(Received 4 May 1966)

In a previous paper, it was shown that any function $\Phi(G)$, defined for a general linear graph G and having the extensive property, can be expanded in terms of the lattice constants of connected subgraphs of G. In this paper, a graphical interpretation of the weight factors occurring in this expansion is given. The usefulness of the expansion in deriving series expansions for properties associated with crystal lattices is discussed with particular reference to percolation problems, dilute ferromagnets, and lattice gases. A result in the theory of linear graphs, recently proved by Rushbrooke in a paper concerned with dilute ferromagnets, is rederived.

1. INTRODUCTION

THE general theory of linked-cluster expansions for the many-body problem has received much attention in the literature. Because of its generality, it is not always clear what conditions are necessary for its application to problems involving crystal lattices. For such problems, the theory takes on a purely algebraic form and has been developed from first principles in a recent paper,[1] hereinafter referred to as I.

The use of lattice constants[1,2] has been found to be of great practical importance in this field, and high-order tables of these are now available. Consequently, it is very useful to formulate a theory in terms of these. Precise definitions of the various types of lattice constant have been given in I in terms of the theory of linear graphs. The most important result of I, in the present context, was that any function $\Phi(G)$, defined for a general linear graph G and having the extensive property that

$$\Phi(G \cup G') = \Phi(G) + \Phi(G'), \tag{1.1}$$

[1] M. F. Sykes, J. W. Essam, B. R. Heap, and B. J. Hiley, J. Math. Phys. 7, 1557 (1966).

[2] C. Domb, Phil. Mag. Suppl. 9, 149 (1960).

and

$$O(s, C) \times O(t, C) \quad (st = N, 3 \leq s \leq t, \quad s, t \neq 4) \tag{A8}$$

are maximal subgroups of $O(N, C)$. Every irreducible nonsimple subgroup of $O(N, C)$ is conjugate in the group of all orthogonal matrices (proper and improper) to one of the groups of the form (A7) or (A8).

We note that $O(2, C)$ is Abelian and hence reducible and that

$$O(4, C) = Sp(2, C) \times Sp(2, C). \tag{A9}$$

From this one has for $s \neq 2$ the following proper inclusions:

$$Sp(s, C) \times O(4, C) = Sp(s, C) \times Sp(2, C) \times Sp(2, C) \subset O(2s, C) \times Sp(2, C) \subset Sp(4s, C).$$

Hence $Sp(s, C) \times O(4, C)$ is not maximal in $Sp(4s, C)$ [cf., (6)]. In the same way one has for $s \geq 3$

$$O(s, C) \times O(4, C) = O(s, C) \times Sp(2, C) \times Sp(2, C) \subset Sp(2s, C) \times Sp(2, C) \subset O(4s, C),$$

therefore, the restriction $s, t \neq 4$ in (A8). In (A7) one has for $s = t = 2$:

$$Sp(2, C) \times Sp(2, C) = O(4, C),$$

so that $Sp(2, C) \times Sp(2, C)$ is no proper subgroup of $O(4, C)$.

Summary of the Appendix: Let U be an irreducible connected complex Lie group of unimodular $N \times N$ matrices. A simple complex Lie group of $N \times N$ matrices which contains U is in general equal to $SL(N, C)$, $O(N, C)$, or $Sp(N, C)$, the exceptions being specified by Theorem A2. U is subgroup of $O(N, C)$ or $Sp(N, C)$ if and only if it has a symmetric or, respectively, a skew-symmetric bilinear invariant, the form of which is determined by Theorem A4. U is a maximal connected subgroup (up to conjugacy)

of $SL(N, C)$ if and only if

$$U = SL(s, C) \times SL(t, C), \quad st = N, \quad 2 \leq s \leq t;$$

of $Sp(N, C)$ if and only if

$$U = Sp(s, C) \times O(t, C), \quad st = N, \quad s \geq 2, \quad t \geq 3, \quad t \neq 4; \quad \text{or} \quad s = 2, \quad t = 4;$$

of $O(N, C)$ if and only if

$$U = O(s, C) \times O(t, C), \quad st = N, \quad 3 \leq s \leq t, \quad s, t \neq 4$$

or

$$U = Sp(s, C) \times Sp(t, C), \quad st = N, \quad 2 \leq s \leq t.$$

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Determination of Weight Factors in Linked-Cluster Expansions for Lattice Systems

JOHN W. ESSAM
Westfield College, London, England

(Received 4 May 1966)

In a previous paper, it was shown that any function $\Phi(G)$, defined for a general linear graph G and having the extensive property, can be expanded in terms of the lattice constants of connected subgraphs of G. In this paper, a graphical interpretation of the weight factors occurring in this expansion is given. The usefulness of the expansion in deriving series expansions for properties associated with crystal lattices is discussed with particular reference to percolation problems, dilute ferromagnets, and lattice gases. A result in the theory of linear graphs, recently proved by Rushbrooke in a paper concerned with dilute ferromagnets, is rederived.

1. INTRODUCTION

THE general theory of linked-cluster expansions for the many-body problem has received much attention in the literature. Because of its generality, it is not always clear what conditions are necessary for its application to problems involving crystal lattices. For such problems, the theory takes on a purely algebraic form and has been developed from first principles in a recent paper,[1] hereinafter referred to as I.

The use of lattice constants[1,2] has been found to be of great practical importance in this field, and high-order tables of these are now available. Consequently, it is very useful to formulate a theory in terms of these. Precise definitions of the various types of lattice constant have been given in I in terms of the theory of linear graphs. The most important result of I, in the present context, was that any function $\Phi(G)$, defined for a general linear graph G and having the extensive property that

$$\Phi(G \cup G') = \Phi(G) + \Phi(G'), \tag{1.1}$$

[1] M. F. Sykes, J. W. Essam, B. R. Heap, and B. J. Hiley, J. Math. Phys. 7, 1557 (1966).

[2] C. Domb, Phil. Mag. Suppl. 9, 149 (1960).

can be expressed in the form

$$\Phi(G) = \sum_{i=1}^{k} f(c_i)\pi_i(G), \tag{1.2}$$

where G is any graph in which each component is isomorphic with one of the connected graphs $c_1 \cdots c_k$, $\pi_i(G)$ is the weak or strong lattice constant of c_i in G, and $f(c_i)$ is a weight function independent of G, which is a linear combination of the Φ's for the connected subgraphs of c_i [see, for example, Eq. (5.1)]. The derivation of (1.2) is summarized later in Sec. 2 in a form due to Sykes.[3] Equation (1.2) is the lattice problem equivalent of the linked-cluster expansion, the right-hand side being a sum over connected graphs. Its application to problems involving crystal lattices becomes clear if the sites of the lattice are taken to be the vertices of G, and if, as is usual but not necessary, the nearest neighbor bonds are taken to be the edges of G. Thus, a lattice may be thought of as an infinite connected linear graph. For mathematical purposes, it is convenient to consider first the case when G is finite and to approach the infinite lattice by a sequence of finite graphs.

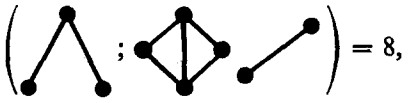
In I an algebraic expression for $f(c_i)$ was given. It is the object of this paper to provide a simple graphical interpretation of the $f(c_i)$. As a result of this, it is possible to write down the weight of any particular graph without knowing the weights of its subgraphs. This is important in developing expansions to high order, since it enables spot checks.

Because of the simplicity of our result, it is possible to see under what conditions the $f(c_i)$ are independent of the list of connected graphs $c_1 \cdots c_k$. In I the complete list of connected graphs was taken, but this, however, is by no means necessary. This aspect of the theory is discussed in Sec. 5. Also in this section, the application to infinite lattices is discussed with the object of deriving power series expansions which is then illustrated by applying the theory to percolation problems, dilute ferromagnets, and the hard square-lattice gas.

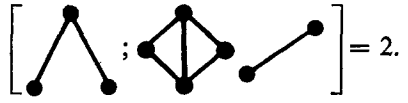
As a by-product of this work, an important result in a recent paper of Rushbrooke[4] is rederived [see Eq. (4.7)].

2. SUMMARY OF CLUSTER EXPANSION THEORY

In this section, the cluster expansion theory given in I is summarized. The weak lattice constant $(G'; G)$ of a linear graph G' in a linear graph G is the number of subgraphs (see I) of G isomorphic with G'. The strong lattice constant $[G'; G]$ is the number of section graphs (see I) of G isomorphic with G'. For example,



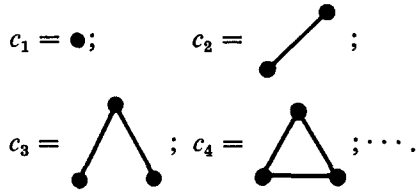
but



Denote by c_i, the ith graph in a complete list of connected graphs with no multi-edges or loops, and suppose that they have been ordered in such a way that

$$(c_i; c_j) = [c_i; c_j] = 0 \quad \text{for} \quad i > j.$$

Such a list is called a graph dictionary, and the graphs are referred to as being in dictionary order when it is wished to stress the order. For instance, the list may begin as follows:



The subsequent theory applies to any function $\Phi(G)$ defined for a graph G and which has the extensive property,

$$\Phi(G) = \sum_{j=1}^{k} \Phi(c_j)\pi_j^*(G), \tag{2.1}$$

where $\pi_j^*(G)$ is the number of components of G isomorphic with c_j, and k is some arbitrarily chosen integer independent of G, such that $\pi_j^*(G) = 0$ for $j > k$. In practice, this imposes no restriction on G, since k can be chosen as large as we please. Equation (2.1) is derived by successive applications of (1.1), using the fact that G is the union of its components. If G is connected, the sum reduces to a single term and is trivially satisfied. An example of the class of function considered is the free energy of a Heisenberg or Ising ferromagnet which may even be diluted with non-magnetic atoms. In this case, the vertices of G represent spins, and the edges of G represent possible interactions (see Sec. 7).

To obtain a cluster expansion for $\Phi(G)$, it is necessary to express the $\pi_j^*(G)$ in terms of the lattice constants of G corresponding to connected graphs. The theory is valid in either weak or strong lattice constants, and we use $\pi_j(G)$ to denote either $(c_j; G)$ or $[c_j; G]$, depending on whether weak or strong lattice constants are being used. Since only connected subgraphs or section graphs of G are being considered,

$$\pi_i(G) = \sum_{j=1}^{k} \Delta_{ij}\pi_j^*(G), \tag{2.2}$$

[3] M. F. Sykes, private communication.
[4] G. S. Rushbrooke, J. Math. Phys. 5, 1106 (1964).

where Δ_{ij} is either $(c_i; c_j)$ or $[c_i; c_j]$, depending on which set of lattice constants are used. By our choice of ordering,

$$\begin{aligned} \Delta_{ij} &= 0, \quad i > j, \\ &= 1, \quad i = j, \end{aligned} \tag{2.3}$$

so that $\det\{\Delta_{ij}\} = 1$, and Eq. (2.2) may be inverted to yield

$$\pi_j^*(G) = \sum_{i=1}^{k} \Gamma_{ji}\pi_i(G), \tag{2.4}$$

where $\Gamma_{ij} = (\Delta^{-1})_{ij}$. Substituting (2.4) into (2.1) and interchanging the order of summation, we obtain the required cluster expansion,

$$\Phi(G) = \sum_{i=1}^{k} f(c_i)\pi_i(G), \tag{2.5}$$

where

$$f(c_i) = \sum_{j=1}^{k} \Phi(c_j)\Gamma_{ji}. \tag{2.6}$$

It is thus a direct consequence of Eq. (2.1) that $\Phi(G)$ may be written as a linear combination of lattice constants of the connected section graphs or subgraphs of G only, the *weight functions* $f(c_i)$ being determined in terms of the Φ's for connected graphs and the Γ_{ij}. The weight functions are thus independent of G, and the Γ_{ij} are the same for all functions Φ. The main purpose of this paper is to provide a graphical interpretation for the Γ_{ij}, an algebraic expression for which has already been given in I.

3. WEIGHT FACTOR DETERMINATION

In this section, we first introduce the concept of a full perimeter lattice constant. The *strong full perimeter lattice constant* $[G'; G]^F$ of G' in G is defined to be the number of *section graphs* of G isomorphic with G', with the restriction that all the *vertices* of G not in G' (if any) must be adjacent to vertices of G'. For example, $[c_1; c_3]^F = 1$, whereas $[c_1; c_3] = 3$. The *weak full perimeter lattice constant* $(G'; G)^F$ of G' in G is defined to be the number of *subgraphs* of G isomorphic with G', with the restriction that all the *edges* of G not in G' (if any) are incident with vertices in G'. For example, $(c_1; c_4)^F = 0$, whereas $(c_1; c_4) = 3$.

Our results may now be stated as

$$\Gamma_{ij}^w = (-1)^{l_j - l_i}(c_i; c_j)^F, \tag{3.1}$$

$$\Gamma_{ij}^s = (-1)^{v_j - v_i}[c_i; c_j]^F, \tag{3.2}$$

where v_i and l_i are, respectively, the numbers of vertices and edges in c_i, and the superscripts on Γ_{ij} distinguish weak and strong cases. To illustrate our result, the matrices relevant to the case $k = 4$ are

$$\{\Delta_{ij}^w\} = \begin{array}{c} \\ 1 \\ 0 \\ 0 \\ 0 \end{array} \begin{array}{c} \\ 2 \\ 1 \\ 0 \\ 0 \end{array} \begin{array}{c} \\ 3 \\ 2 \\ 1 \\ 0 \end{array} \begin{array}{c} \\ 3 \\ 3 \\ 3 \\ 1 \end{array},$$

$$\{\Gamma_{ij}^w\} = \begin{array}{c} \\ 1 \\ 0 \\ 0 \\ 0 \end{array} \begin{array}{c} \\ -2 \\ 1 \\ 0 \\ 0 \end{array} \begin{array}{c} \\ 1 \\ -2 \\ 1 \\ 0 \end{array} \begin{array}{c} \\ 0 \\ 3 \\ -3 \\ 1 \end{array},$$

$$\{\Delta_{ij}^s\} = \begin{array}{c} \\ 1 \\ 0 \\ 0 \\ 0 \end{array} \begin{array}{c} \\ 2 \\ 1 \\ 0 \\ 0 \end{array} \begin{array}{c} \\ 3 \\ 2 \\ 1 \\ 0 \end{array} \begin{array}{c} \\ 3 \\ 3 \\ 0 \\ 1 \end{array},$$

$$\{\Gamma_{ij}^s\} = \begin{array}{c} \\ 1 \\ 0 \\ 0 \\ 0 \end{array} \begin{array}{c} \\ -2 \\ 1 \\ 0 \\ 0 \end{array} \begin{array}{c} \\ 1 \\ -2 \\ 1 \\ 0 \end{array} \begin{array}{c} \\ 3 \\ -3 \\ 0 \\ 1 \end{array}.$$

To prove (3.1), it is sufficient to show that

$$\sum_m (-1)^{l_m-l_i}(c_i; c_m)^F(c_m; c_j) = \delta_{ij}. \qquad (3.3)$$

The proof is based on a well-known theorem in the theory of probability. Suppose that $A_1, A_2, \cdots, A_N$ are N events which are not necessarily mutually exclusive. Denote by A_i^r the event that the ith set of r events chosen from $A_1, \cdots, A_N$ all occur (the order of choice being irrelevant), and by $A_1 \cup A_2 \cup \cdots \cup A_N$ the event that at least one of the events $A_1, A_2, \cdots, A_N$ occurs; then, of $P(A)$ denotes the probability of occurrence of event A,

$$P\{A_1 \cup A_2 \cup \cdots \cup A_N\} = \sum P\{A_i^1\} - \sum P\{A_i^2\} + \sum P\{A_i^3\} - \sum P\{A_i^4\} + \cdots \pm P\{A_i^N\}, \qquad (3.4)$$

where the rth summation contains $\binom{N}{r}$ terms. Equation (3.4) is proved in Ref. 5 and is related to the principle of inclusion and exclusion.

Suppose now that the edges of a graph G are colored at random (colored black with probability p, and colored white with probability $1-p$). Any realization of the probability distribution defines a partial graph P_B of G consisting of all the vertices of G and the black edges.

Theorem I: The probability $\bar{P}(p; c_i; G)$ that P_B has a component isomorphic with the graph c_i is

$$\bar{P}(p; c_i; G) = \sum_m (-1)^{l_m-l_i}(c_i; c_m)^F(c_m; G)p^{l_m}. \qquad (3.5)$$

Proof: Let $\bar{Q}(p; c_i; G)$ be the probability that P_B has a subgraph isomorphic with c_i, which is not a component of P_B; then

$$\bar{P}(p; c_i; G) = (c_i; G)p^{l_i} - \bar{Q}(p; c_i; G). \qquad (3.6)$$

To determine $\bar{Q}(p; c_i; G)$, we use Eq. (3.4). Let C be a particular subgraph of G isomorphic with c_i, and suppose that c_i is not isomorphic with G (otherwise the theorem is trivial). Number all the edges of G not in C but adjacent to C from 1 to N. Let A_r be the event that C is a subgraph of P_B and that the rth numbered edge of G is also black. $P\{A_1 \cup A_2 \cup \cdots \cup A_N\}$ is then the probability that C is a subgraph of P_B but not a component, and so $\bar{Q}(p; c_i; G)$ may be obtained by summing this over all C isomorphic with c_i. Suppose now that C_i^r is the subgraph of G, the edge set of which is the union of the edge set of C with the r numbered edges corresponding to the event A_i^r; then

$$P\{A_i^r\} = p^{l(C_i^r)}, \qquad (3.7)$$

where $l(C_i^r)$ is the number of edges in C_i^r. Thus, the summations in (3.4) may be replaced by summations over the C_i^r. If then we sum over all possible subgraphs C, isomorphic with c_i, and collect together contributions from the C_i^r, which are isomorphic with c_m, the contribution to $\bar{Q}(p; c_i; G)$, using (3.4), is

$$(c_i; c_m)^F(c_m; G)p^{l_m}(-1)^{l_m-l_i-1}. \qquad (3.8)$$

This follows since there are $(c_m; G)$ subgraphs of G isomorphic with c_m, and each one contains $(c_i; c_m)^F$ subgraphs isomorphic with c_i such that the edges of c_m not in c are adjacent to c_i. Thus,

$$\bar{Q}(p; c_i; G) = \sum_m{}' (c_i; c_m)^F(c_m; G)p^{l_m}(-1)^{l_m-l_i-1}, \qquad (3.9)$$

where the prime indicates that the term $m = i$ is to be omitted. Equation (3.5) follows directly from (3.6) and (3.9), since $-(c_i; G)p^{l_i}$ corresponds to the term $m = i$ which is missing from (3.9). This completes the proof of Theorem I.

If now G is the connected graph c_j and $p = 1$, then the only subgraph of c_j, which can be a component of P_B, is c_j itself, since all the edges are black; thus

$$\bar{P}(1; c_i; c_j) = \delta_{ij}, \qquad (3.10)$$

and Eq. (3.3) follows by combining (3.5) and (3.10).

The proof of Eq. (3.2) proceeds along similar lines. Suppose now that it is the vertices of G which are colored black with probability p and colored white with probability $1-p$. In any realization of the probability distribution, let R_B be the section graph of G corresponding to the black vertices.

Theorem II: The probability $P(p; c_i; G)$ that R_B has a component isomorphic with c_i is

$$P(p; c_i; G) = \sum_m [c_i; c_m]^F[c_m; G]p^{v_m}(-1)^{v_m-v_i}. \qquad (3.11)$$

Proof: The proof is the same as for Theorem I, if, for $\bar{Q}(p; c_i; G)$ read $Q(p; c_i; G)$, for subgraph read section graph, for P_B read R_B, for edge read vertex, for l_i read v_i, and for weak lattice constants read strong lattice constants.

Again

$$P(1; c_i; c_j) = \delta_{ij}, \qquad (3.12)$$

since when all the vertices are black c_i is the only possible component of R_B, and combining this with (3.11)

$$\sum_m (-1)^{v_m-v_i}[c_i; c_m]^F[c_m; c_j] = \delta_{ij}, \qquad (3.13)$$

which is equivalent to (3.2).

Since $\{\Gamma_{ij}\}$ is the inverse of $\{\Delta_{ij}\}$,

$$\sum_m \Gamma_{im}\Delta_{mj} = \sum_m \Delta_{im}\Gamma_{mj} = \delta_{ij}, \qquad (3.14)$$

so that two further results equivalent to (3.1) and (3.2) are

$$\sum_m (c_i; c_m)(-1)^{l_j-l_m}(c_m; c_j)^F = \delta_{ij}, \qquad (3.15)$$

$$\sum_m [c_i; c_m](-1)^{v_j-v_m}[c_m; c_j]^F = \delta_{ij}, \qquad (3.16)$$

which are not obvious from a graphical point of view.

[5] W. Feller, *An Introduction to Probability Theory and Its Applications* (John Wiley & Sons, Inc., New York, 1961) 2nd ed, Vol. I, Chap. 4, p. 88.

4. CONVERSION MATRICES

Further relations between the lattice constants may be obtained by using the conversion matrices of I. The $(m\text{-}n)$th element $a_{mn}(r)$ of the rth-order conversion matrix was defined to be $(g_m; g_n)$, where m and n run over the graphs with r vertices in a dictionary of all graphs with no multi-edges and loops. It was also pointed out that similar matrices could be defined for the dictionaries of connected and multiply connected graphs. Here, we denote by A_{ij} the $(i\text{-}j)$th element of the $k \times k$ super matrix, the diagonal blocks of which are the conversion matrices corresponding to connected graphs, and the off diagonal blocks of which are zero, that is

$$A_{ij} = (c_i; c_j)\delta_{v_i v_j}. \tag{4.1}$$

For any graph G, the matrix $\{A_{ij}\}$ determines the set of weak connected lattice constants of G in terms of the strong connected lattice constants,

$$(c_i; G) = \sum_j A_{ij}[c_j; G]. \tag{4.2}$$

This follows since every connected subgraph of G may be associated with a section graph of G having the same vertex set and which is also connected, and, if $v_i = v_j$, (c_i, c_j) is the number of subgraphs isomorphic with c_i having the associated section graph c_j.

If $\{B_{ij}\}$ is the inverse of $\{A_{ij}\}$, we show that

$$B_{ij} = (-1)^{l_j - l_i} A_{ij} = (-1)^{l_j - l_i}(c_i; c_j)\delta_{v_i, v_j}, \tag{4.3}$$

a result which is analogous to Eq. (3.8) in I, which applies to the dictionary of *all graphs* with no multi-edges or loops. To prove Eq. (4.3), we notice that, from (3.15),

$$\sum_m (c_i; c_m)(-1)^{l_j - l_m}(c_m; c_j)\delta_{v_i, v_j} = \delta_{ij}, \tag{4.4}$$

since if $v_i \neq v_j$, $i \neq j$ and both sides are zero, but if $v_i = v_j$, then the only nonzero terms on the left of (3.15) have $v_m = v_j$, and in this case, $(c_m; c_j)^F = (c_m; c_j)$. Equation (4.4) may therefore be written as

$$\sum_m (c_i; c_m)\delta_{v_i, v_m}(-1)^{l_j - l_m}(c_m; c_j)\delta_{v_m, v_j} = \delta_{ij}, \tag{4.5}$$

which is equivalent to (4.3).

We are now in a position to derive further results which relate the weak lattice constants to the strong full perimeter lattice constants and vice versa. From (3.16),

$$\sum_{i'} \sum_m A_{ii'}[c_{i'}; c_m](-1)^{v_j - v_m}[c_m; c_j]^F = \sum_{i'} A_{ii'}\delta_{i'j}, \tag{4.6}$$

which using (4.2) becomes

$$\sum_m (c_i; c_m)(-1)^{v_j - v_m}[c_m; c_j]^F = A_{ij}, \tag{4.7}$$

which is the same as Eq. (24) of Ref. 4. From (3.15) and (4.2),

$$\sum_m \sum_{i'} A_{ii'}[c_{i'}; c_m](-1)^{l_j - l_m}(c_m; c_j)^F = \delta_{ij}, \tag{4.8}$$

so that

$$\sum_m \sum_{i'} \sum_{i''} B_{ii''}A_{i''i'}[c_{i'}; c_m](-1)^{l_j - l_m}(c_m; c_j)^F = \sum_{i''} B_{ii''}\delta_{i''j}, \tag{4.9}$$

or, since $\{B_{ij}\}$ is the inverse of $\{A_{ij}\}$,

$$\sum_m [c_i; c_m](-1)^{l_j - l_m}(c_m; c_j)^F = B_{ij}. \tag{4.10}$$

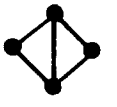
5. CHOICE OF GRAPH DICTIONARY

In the previous sections, we chose for simplicity to use the dictionary of *all* connected graphs with no multi-edges or loops. This was by no means necessary for the validity of the theory, and it is, in fact, restrictive in two important ways.

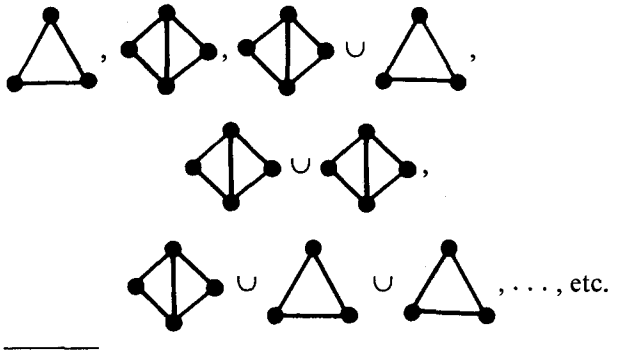
First, if G is a physical lattice L (such as the diamond lattice), then by no means all of the c_j occur as subgraphs of G. Numerical data illustrating this point were given in I. In this case, it is convenient to use the dictionary of connected graphs which are subgraphs of L.

Second, if we wish to calculate $\Phi(G)$ for a graph with multi-edges or loops, it is necessary to include connected graphs of this type in the dictionary. Such a case might be a bond percolation problem[6] in which it is desired to make the probability of two sites being connected depend on the pair of sites chosen. This may be achieved by having more than one bond joining a given pair of sites.

Let us now examine the necessary restrictions on the choice of a graph dictionary. It is to be observed that (2.1) to (2.6) are generally valid for any list of connected graphs. Having chosen such a list, G is then restricted to being the union of any number of components isomorphic with these graphs. For example, supposing the list is

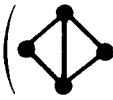
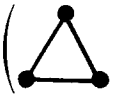
$c_1 =$  , $c_2 =$ 

consisting of just two graphs, then G may be, for example,



[6] J. W. Essam and M. F. Sykes, J. Math. Phys. 7, 1573 (1966).

In this case, Eq. (2.5) reads

$$\Phi(G) = \Phi\left(\text{  }\right)\pi_1(G) + \left\{\Phi\left(\text{  }\right) - 2\Phi\left(\text{  }\right)\right\}\pi_2(G). \quad (5.1)$$

In calculating the weight functions $f(c_i)$ defined by (2.6), it is desirable that Eqs. (3.1) and (3.2) are usable. The validity of Theorem I, from which (3.1) follows, depends on the graph dictionary containing all connected subgraphs of G which themselves have subgraphs isomorphic with c_i. Thus, the restriction imposed by the use of (3.1) is that the graph dictionary must be such that, for any graph c_j contained in the dictionary, all the connected subgraphs of c_j must also be contained in the dictionary. Similarly, for (3.2) to be valid, the presence of c_j in the dictionary must imply the presence of all its connected section graphs. The above restriction is referred to as restriction A. The usefulness of Eqs. (3.1) and (3.2) lies not only in their simplicity but also in the fact that the graph dictionary may be varied within the above limits without any change in the weight functions $f(c_i)$. For example, suppose the weight functions which are necessary for the calculation of some $\Phi(G)$ have been computed and it is then required to calculate $\Phi(G')$. A large number of the weight functions will be common to G and G', and it is therefore only necessary to calculate the new ones introduced by G' even though the graph dictionary may have been changed.

A second restriction, which is not so easily formulated, arises when we consider the usefulness of Eq. (2.5). First of all, we note that, if G is connected, $\Phi(G)$ appears on both sides of (2.5), and if we wish to use only the right-hand side to calculate $\Phi(G)$, nothing will be gained unless $f(G) = 0$. In practice, we wish to make use of (2.5) in calculating Φ for an infinite lattice L. In this case, the vertices of G are the sites of L. Actually, Φ does not exist for an infinite lattice, but if it is a bulk property and G_n $(n = 1 \cdots \infty)$ is a member of a sequence of finite graphs which tend to L as n tends to ∞, then

$$\phi(L) = \lim_{n\to\infty} \Phi(G_n)/v(G_n) \quad (5.2)$$

will normally exist. This is merely a definition of a bulk property. The lattice constants $p_i(L)$ are calculated per site and are defined by

$$p_i(L) = \lim_{n\to\infty} \pi_i(G_n)/v(G_n). \quad (5.3)$$

Thus, (2.5) becomes

$$\phi(L) = \sum_{i=1}^{\infty} f(c_i)p_i(L). \quad (5.4)$$

The restriction imposed on the useful choice of graph dictionary is the requirement that a good approximation to $\phi(L)$ may be obtained by calculating only a finite number of weight functions.

One way of calculating $\phi(L)$ is to truncate (5.4), after a finite number of terms, and then increase this number and observe the change to obtain an estimate of the terms which have been omitted. This is the so-called cluster expansion method. The success of this method requires $f(c_i)$ to become small as i increases, and this is usually valid in some region of the physical parameters. It is possible that, by varying the graph dictionary in such a way that the $f(c_i)$ are changed, better approximations may be obtained.

A second method is that of using (5.4) to obtain power series expansions in some parameter x. For the method to work, contributions to the coefficient of x^n, where n is finite, must come from a finite set of the f's. This is usually achieved by applying physical reasoning in the choice of x. It is possible, although no examples are known, that changing the graph dictionary may also achieve the same effect by changing the f's. We now consider some examples which illustrate the above points.

6. PERCOLATION PROBLEMS

Suppose that, as in the proof of Theorem I, the edges of a graph G are colored at random, black with probability p and white with probability $1 - p$. If now we take $\Phi(G)$ to be the mean number of components in P_B (the mean number of black clusters using the null cluster convention[6]), then

$$\Phi(G) \equiv \bar{K}_0(p; G) \equiv \sum_i \bar{P}(p; c_i; G). \quad (6.1)$$

The expression of $\bar{K}_0(p; G)$ as a polynomial in p has been discussed in a previous paper[6] and follows directly from (3.5). Here

$$\bar{K}_0(p; G) = \sum_m (c_m; G)\bar{k}_0(c_m)p^{t_m}, \quad (6.2)$$

where

$$\bar{k}_0(c_m) = \sum_i (-1)^{l_m - l_i}(c_i; c_m)^F. \quad (6.3)$$

The dictionary used is that of all connected subgraphs of G which automatically satisfy restriction A, so that if, instead of calculating $\bar{K}_0(p; G)$, we wish to calculate $\bar{K}_0(p; G')$, $G' \not\equiv G$; then, if c_m is a subgraph common to G and G', $\bar{k}_0(c_m)$ will be the same in both cases. Notice that if G has no multi-edges, then there are no subgraphs c_m of G having multi-edges and the weights of these graphs are then not needed. If G has multi-edges, the weights of subgraphs having no multi-edges will be the same as in the case when G has no multi-edges. For an infinite physical lattice L, the mean number of clusters is a bulk property, and

we define the mean number per site by

$$\bar{k}_0(p; L) = \lim_{n\to\infty} \bar{K}_0(p; G_n)/v(G_n) \qquad (6.4)$$

in accordance with the previous section. Equation (6.2) is in the best possible form for the derivation of a power series expansion, since, to obtain the coefficient of p^n, it is merely necessary to sum over the finite collection of subgraphs with n edges. The series will converge for p less than a certain value p_c, the critical probability which lies in general between 0 and 1. If the function $\bar{k}_0(p; L)$ is required for values of p greater than p_c, the power series in p is of no use. One would naively think that, if p is written as $1 - q$ in (6.2) and the coefficients of q^n are collected, a series expansion for $\bar{k}_0(p; L)$ could be obtained, which would be valid for p close to unity. This would be the case except for the fact that, to all orders in q, an infinity of terms contributes to each coefficient. Such an expansion does, however, exist, but this will be discussed in a subsequent paper. For the plane square lattice, for example, $\bar{k}_0(p; \mathrm{sq}) = O(q^4)$.

If now it is the vertices of G which are colored at random, as in the proof of Theorem II, a polynomial expression for the mean number of black clusters (components of R_B), $K(p; G)$ may be obtained from (3.11),

$$K(p; G) \equiv \sum_i P(p; c_i; G) = \sum_m [c_m; G]K(c_m)p^{v_m}, \qquad (6.5)$$

where

$$K(c_m) = \sum_i (-1)^{v_m - v_i}[c_i; c_m]^F. \qquad (6.6)$$

In this case, graphs G with multi-edges are not normally considered, since the replacement of an edge by a multi-edge has no effect on $K(p; G)$. Again, (6.5) is in an ideal form for obtaining power series expansions in the variable p, the coefficient of p^n coming from all connected sections graphs of G with n vertices. The properties of both $\bar{k}_0(c_m)$ and $K(c_m)$ have been discussed in a previous paper.[6]

7. DILUTE ISING AND HEISENBERG FERROMAGNETS

The problem of dilute Ising and Heisenberg ferromagnets has been considered by Rushbrooke[4]; some of the results therein are rederived. Suppose that the vertices of a finite linear graph G, with no multi-edges or loops, are occupied at random by either magnetic or nonmagnetic elements, the probability of occurrence of a magnetic element being p. Suppose also that there is a spin–spin interaction (Heisenberg or Ising) between two magnetic elements if and only if the vertices which they occupy are connected by an edge. Our theory as formulated only applies when the interaction energy is independent of the edge connecting the magnetic elements, but may easily be generalized. The function $\Phi(G)$ in this problem is taken as $-\beta F(p, T, H; G)$, where $\beta = 1/k_B T$ and F is the free energy which will be a function of the concentration of magnetic elements, the temperature and the magnetic field. It is clear that $\Phi(G)$ so defined has the extensive property (2.1) independent of whether the interaction is of the Heisenberg or the Ising type. We choose to take the graph dictionary of all connected graphs so that, by choosing k in Sec. 2 to be large enough, any finite graph G is included in the theory. To obtain the free energy of the dilute magnet, it is necessary (see Brout[7]) to calculate the free energy corresponding to a given allocation of the magnetic elements to the vertices of G and then to average over all such allocations, weighting with their respective probabilities. If the magnetic elements are identified with the black vertices in the proof of Theorem II, it follows that

$$F(p, T, H; G) = \sum_i P(p; c_i; G)F(1, T, H; c_i), \qquad (7.1)$$

since in the allocation corresponding to the section graph R_B, the free energy is the sum of the free energies calculated for each component of R_B with $p = 1$. The latter simply results, since there are no edges of G connecting magnetic elements which are in different components of R_B. Using (3.11) and defining $\Theta(G)$ by

$$\Theta(G) = -\beta F(1, T, H; G), \qquad (7.2)$$

Eq. (7.1) becomes

$$-\beta F(p, T, H; G) \equiv \Phi(G) = \sum_m [c_m; G]W(c_m)p^{v_m}, \qquad (7.3)$$

where

$$W(c_m) = \sum_i (-1)^{v_m - v_i}[c_i; c_m]^F \Theta(c_i). \qquad (7.4)$$

Now consider the situation which arises in the case of an infinite crystal lattice L. The vertices of G are now the sites of L and the edges of G represent pairs of sites which are separated by the nearest-neighbor distance. It is necessary to restrict the argument to nearest neighbors so that the interaction energy, being a function of the distance between sites, is the same for all edges of G. If second- and higher-neighbor interactions are to be included, the edges of G must be labeled accordingly, and a more general formulation of the theory used which allows the interaction energy to depend on the type of edge. The property of L which remains finite is the free energy per lattice site, and Eq. (7.3) is in the ideal form for obtaining a power series expansion in terms of p, contributions to the coefficient of p^n coming from all connected section graphs of L with n vertices. The coefficients are, of course, functions of T and H, and so the radius of convergence may depend on these parameters. Similar expansions for the zero-field

[7] R. Brout, Phys. Rev. 115, 824 (1959).

susceptibility have been used by a number of authors[8–10] to determine the radius of convergence as a function of temperature. For high enough temperatures, the series converges over the whole physical range of p; this is why a high-temperature expansion for $p = 1$ may be obtained by regrouping the terms of (7.3). The method fails to yield a low-temperature expansion for any value of p for the same reason that a high-density expansion for the mean number could not be obtained in the previous section.

8. THE HARD SQUARE LATTICE GAS

For a recent discussion of the hard square lattice gas, see Ref. 11. Suppose that the vertices of a graph G are occupied by molecules which interact in such a way that vertices connected by an edge cannot be simultaneously occupied and also that there must be no multioccupation of a vertex. If there are no other restrictions, the grand partition function is

$$\Xi = 1 + \sum_{l=1} \xi_l(G) z^l, \tag{8.1}$$

where z is the activity and $\xi_l(G)$ is the number of ways of placing l molecules on the graph in accordance with the above restrictions. In this problem,

$$\Phi(G) \equiv \Gamma(z, G) = \ln \Xi, \tag{8.2}$$

which clearly has the extensive property (2.1). Using (2.5), we may write

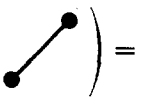
$$\Gamma(z, G) = \sum_{i=1}^{k} f(z, c_i)[c_i; G], \tag{8.3}$$

where

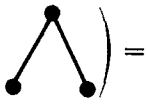
$$f(z, c_i) = \sum_{j=1}^{k} \Gamma(z, c_j)(-1)^{v_i - v_j}[c_j; c_i]^F, \tag{8.4}$$

and we have used strong lattice constants. We now attempt to derive a power series expansion for $\Gamma(z, G)$ in terms of the variable z. The dictionary of all connected graphs which are section graphs of the square lattice is used, and the expansion developed through z^4. It turns out, although it is not obvious from our formulation, that a section graph with n vertices makes no contribution to powers of z less than the nth. This, in fact, is the point we are making in this section. If it were not so, the method would break down. The first task is to calculate the Γ's which are

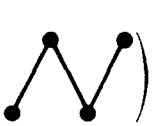
$$\Gamma(z, \bullet) = \ln(1 + z), \quad \Gamma\left(z, \; \right) = \ln(1 + 2z),$$



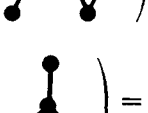
$$\Gamma\left(z, \; \right) = \ln(1 + 3z + z^2),$$



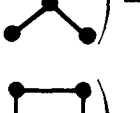
$$\Gamma\left(z, \; \right) = \ln(1 + 4z + 3z^2), \tag{8.5}$$



$$\Gamma\left(z, \; \right) = \ln(1 + 4z + 3z^2 + z^3),$$



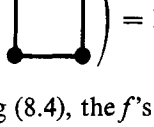
$$\Gamma\left(z, \; \right) = \ln(1 + 4z + 2z^2).$$



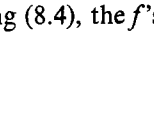
Using (8.4), the f's are

$$f(z, \bullet) = \Gamma(z, \bullet) = z - \tfrac{1}{2}z^2 + \tfrac{1}{3}z^3 - \tfrac{1}{4}z^4 + O(z^5),$$

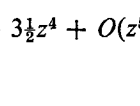
$$f\left(z, \; \right) = \Gamma\left(z, \; \right) - 2\Gamma(z, \bullet) = -z^2 + 2z^3 - 3\tfrac{1}{2}z^4 + O(z^5),$$



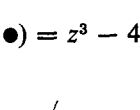
$$f\left(z, \; \right) = \Gamma\left(z, \; \right) - 2\Gamma\left(z, \; \right) + \Gamma(z, \bullet) = z^3 - 4z^4 + O(z^5),$$



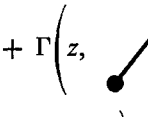
$$f\left(z, \; \right) = \Gamma\left(z, \; \right) - 2\Gamma\left(z, \; \right) + \Gamma\left(z, \; \right) = -z^4 + O(z^5), \tag{8.6}$$



$$f\left(z, \; \right) = \Gamma\left(z, \; \right) - 3\Gamma\left(z, \; \right) + 3\Gamma\left(z, \; \right) - \Gamma(z, \bullet) = -z^4 + O(z^5),$$



$$f\left(z, \; \right) = \Gamma\left(z, \; \right) - 4\Gamma\left(z, \; \right) + 4\Gamma\left(z, \; \right) = -3z^4 + O(z^5).$$



[8] G. S. Rushbrooke and D. J. Morgan, Mol. Phys. **4**, 1 (1961).
[9] R. J. Elliot and B. R. Heap, Proc. Roy. Soc. (London) **A265**, 264 (1962).
[10] B. R. Heap, Proc. Phys. Soc. (London) **82**, 252 (1963).
[11] D. S. Gaunt and M. E. Fisher, J. Chem. Phys. **43**, 2840 (1965).

For a lattice L, define

$$\gamma(z, L) = \lim_{n\to\infty} \Gamma(z, G_n)/v(G_n), \tag{8.7}$$

so that, using the strong lattice constants per site for the plane square lattice, we obtain from (8.3) the expansion for the hard square lattice gas in terms of the activity,

$$\gamma(z, \text{sq}) = z - 2\tfrac{1}{2}z^2 + 10\tfrac{1}{3}z^3 - 52\tfrac{1}{4}z^4 + O(z^5). \tag{8.8}$$

Here, we are not interested in deriving higher terms which are available,[11] but only wish to emphasize the point made in Sec. 5 that, by making a physical choice of expansion variable, only a finite number of terms contributes to the expansion coefficients. The reason for the cancellation in this case may be seen by using the method of Rushbrooke and Scoins,[12] but the general considerations governing a choice of variable are not obvious.

9. SUMMARY AND CONCLUSION

A graphical interpretation has been given to the coefficients which appear in the weight functions associated with the expansion of an extensive property $\Phi(G)$ in terms of connected lattice constants. The result may be summarized by

$$\Phi(G) = \tilde{f}\pi(G), \tag{9.1}$$

where

$$f = \tilde{\Gamma}\Phi \tag{9.2}$$

and $\tilde{A}$ denotes the transpose of a matrix A. The column matrices $\pi(G)$, Φ, and f have ith elements $\pi_i(G)$, $\Phi(c_i)$, and $f(c_i)$, respectively. Subject to quite broad restrictions, given in Sec. 5, the i–jth element of the square matrix Γ is given by (3.1), when weak lattice constants are being used, and by (3.2), when strong lattice constants are being used, and apart from a possible factor of -1, Γ_{ij} is the full perimeter lattice constant (weak or strong) of c_i in c_j.

Γ is the inverse of a matrix Δ, the i–j the element of which is $(c_i; c_j)$ in the weak case and $[c_i; c_j]$ in the strong case, thus

$$\Gamma^W\Delta^W = \Gamma^S\Delta^S = I, \tag{9.3}$$

$$\Delta^W\Gamma^W = \Delta^S\Gamma^S = I. \tag{9.4}$$

In Sec. 4, we defined the conversion matrix A, and Eq. (4.2) implies that

$$\Delta^W = A\Delta^S \tag{9.5}$$

and relates weak lattice constants to strong lattice constants. The inverse matrix B of A is given by (4.3)

$$\Delta^S = B\Delta^W, \tag{9.6}$$

and using Eqs. (9.5) and (9.6) in conjunction with (9.3) and (9.4), we obtain Eqs. (4.7) and (4.10), namely,

$$\Delta^W\Gamma^S = A, \tag{9.7}$$

$$\Delta^S\Gamma^W = B. \tag{9.8}$$

An independent proof of (9.7) has been given by Rushbrooke.[4]

The derivation of power series expansions in the case when G is an infinite lattice has been discussed and illustrated by percolation problems, dilute ferromagnets, and lattice gases.

Theorems I and II, which were used in deriving (9.3) and (9.4), are especially useful in connection with the cluster size distribution for percolation problems. This application will be discussed in a subsequent paper.

ACKNOWLEDGMENTS

The author is indebted to Professor M. E. Fisher and Dr. M. F. Sykes for useful criticism of this paper.

[18] G. S. Rushbrooke and H. J. Scoins, Proc. Roy. Soc. (London) **A230**, 74 (1955).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

CTP Invariance of the S-Matrix in a Theory of Local Observables

HENRI EPSTEIN*
The Institute for Advanced Study, Princeton, New Jersey

(Received 14 July 1966)

In a theory of local observables as proposed by Haag and Araki, the assumptions which make possible a collision theory also guarantee the CTP invariance of the S-matrix.

1. INTRODUCTION

THE purpose of this paper is to prove the CTP invariance of the S-matrix in a theory of the type proposed by Haag and Araki.[1-5] Such a theory is defined by giving a Hilbert space $\mathcal{H}$, a continuous unitary representation $\{a, \Lambda\} \to U(a, \Lambda)$ of the Poincaré group $\mathfrak{P}_+^\uparrow$ acting in $\mathcal{H}$ and, for every domain B in the real Minkowski space ($\mathbf{R}^4$), a von Neumann algebra $R(B)$ of bounded operators acting in $\mathcal{H}$ with the following properties:

(1) If $B_1 \subset B_2$ then $R(B_1) \subset R(B_2)$;

(2) If $B = \bigcup_{n=1}^{\infty} B_n$ then $R(B) = \left\{ \bigcup_{n=1}^{\infty} R(B_n) \right\}''$;

(3) If $(x_1 - x_2)^2 < 0$ for every $x_1 \in B_1$ and $x_2 \in B_2$, then $R(B_1)$ and $R(B_2)$ commute [i.e., $R(B_1) \subset R(B_2)'$];

(4) $U(a, \Lambda)R(B)U(a, \Lambda)^{-1} = R(a + \Lambda B)$;

(5) $U(a, 1) = \exp i(a, P)$; the spectrum of P (the energy-momentum operator) is in $\bar{V}^+$;

(6) There is a vector Ω (vacuum) in $\mathcal{H}$, unique up to a scalar factor, such that $U(a, \Lambda)\Omega = \Omega$;

(7) Ω is a cyclic vector for the union of all $R(B)$ and, as a consequence of the Reeh–Schlieder theorem, a cyclic vector for each $R(B)$.

The above assumptions are appropriate to the case when the theory contains no spin $-\frac{1}{2}(2n + 1)$ states. It may however contain different charge superselection sectors; in this case $R(B)$ contains not only local observables, but also operators (similar to charged fields) carrying charge. The case when there are states with spin $\frac{1}{2}(2n + 1)$ can be treated by a slight modification of the assumptions to which we return later.[4,5]

We call HA-field an operator-valued function $x \to A(x)$ of the form $A(x) = U(x, 1)A(0)U(x, 1)^{-1}$, where $A(0) \in R(B)$ for some *bounded* domain B, and $(\Omega, A(0)\Omega) = 0$. Note that $\|A(x)\| = \|A(0)\|$ and that $x \to A(x)$ is weakly continuous. We say that A is $\mathcal{C}^\infty$ if $A(0) = \int \varphi(x)A'(x)\,dx$, $A'(0) \in R(B')$, B' bounded, $\varphi \in \mathcal{D}$. In this case $A(0) \in R(\text{supp } \varphi + B')$.

Let $A_0, \cdots, A_n$ be $(n + 1)$ $\mathcal{C}^\infty$ HA-fields. The permuted Wightman functions of these fields,

$$(\Omega, A_{\pi 0}(x_{\pi 0})A_{\pi 1}(x_{\pi 1}) \cdots A_{\pi n}(x_{\pi n})\Omega),$$

are bounded $\mathcal{C}^\infty$ functions of the difference variables $x_1 - x_0, \cdots, x_n - x_{n-1}$. Generalized retarded functions (grf) can be defined for these fields by the formulas which define such functions in an LSZ theory.[2,6-8] Note that these formulas contain expressions of the form $\theta(x_i^0 - x_j^0)$. We therefore assume in the following that a fixed Lorentz frame has been chosen once and for all for the purpose of writing down these formulas. Multiplication by θ functions is here legitimate and leads to grf which are, in x-space, piecewise $\mathcal{C}^\infty$. In momentum space, the grf are boundary values (in the sense of tempered distributions) of the same analytic function $H(k)$ ($k = p + iq$). The domain of analyticity of $H(k)$ is the same as in an LSZ theory.[2,6,7,9] However, the supports of the grf in x-space are not exactly those of an LSZ theory but are contained in cones obtained by translation from those of an LSZ theory.[9] It follows that $H(k)$ is not bounded by polynomials at infinity but grows exponentially in imaginary directions. This is seen in more detail in the next section.

2. A PROPERTY OF THE p-SPACE ANALYTIC FUNCTION

Although this section is reasonably self-contained, it is the logical sequel of Ref. 10, and uses the same techniques. The notations are those of Ref. 9.

* On leave from Centre National de la Recherche Scientifique, France.

[1] R. Haag, Nuovo Cimento Suppl. **14**, 131 (1959).

[2] H. Araki, "Einfuehrung in die Axiomatische Quantenfeldtheorie," Lecture Notes, Zurich (1961–62).

[3] H. Araki, Progr. Theoret. Phys. (Kyoto) **32**, 844 (1964).

[4] H. J. Borchers, Commun. Math. Phys. **1**, 57 (1965).

[5] H. J. Borchers, Commun. Math. Phys. **1**, 281 (1965).

[6] O. Steinmann, Helv. Phys. Acta **33**, 257 (1960); **33**, 347 (1960).

[7] D. Ruelle, Nuovo Cimento **19**, 356 (1961).

[8] H. Araki and N. Burgoyne, Nuovo Cimento **8**, 342 (1960).

[9] H. Epstein, in *Brandeis University Summer Institute of Theoretical Physics, Vol. 1, Axiomatic Field Theory*, M. Chrétien and S. Deser, Eds., (Gordon and Breach Science Publishers, Inc., New York, 1967).

[10] J. Bros, H. Epstein, and V. Glaser (to be published).

A. Retarded and Advanced Functions

We start by considering the retarded and advanced functions

$$\tilde{a}_0(x) = \sum_{\substack{\pi\in\gamma_{n+1}\\ \pi 0=0}} \theta(x^0_{\pi n} - x^0_{\pi(n-1)}) \cdots \theta(x^0_{\pi 1} - x^0_0) \times (\Omega, [A_{\pi n}(x_{\pi n}), [\cdots, [A_{\pi 1}(x_{\pi 1}), A_0(x_0)] \cdots]]\Omega),$$

$$\tilde{r}_0(x) = \sum_{\substack{\pi\in\gamma_{n+1}\\ \pi 0=0}} \theta(x^0_0 - x^0_{\pi 1}) \cdots \theta(x^0_{\pi(n-1)} - x^0_{\pi n}) \times (\Omega, [[\cdots [A_0(x_0), A_{\pi 1}(x_{\pi 1})], \cdots], A_{\pi n}(x_{\pi n})]\Omega)$$

(π runs over the permutations of 0, 1, 2, $\cdots n$ which do not move 0).

It can be shown[9] that there is a 4-vector $\alpha = (a, 0, 0, 0)$, $a > 0$ such that

$$\text{supp } \tilde{a}_0(x) \subset \{x: x_j - x_0 \in \bar{V}^+ - \alpha,\ 1 \le j \le n\},$$
$$\text{supp } \tilde{r}_0(x) \subset \{x: x_0 - x_j \in \bar{V}^+ - \alpha,\ 1 < j \le n\}.$$

The Laplace transforms of $\tilde{a}_0$ and $\tilde{r}_0$ are given by

$$a_0(k) = \int \exp\left[i\sum_{j=1}^n k_j(x_j - x_0)\right] \tilde{a}_0(x)\, d^4x_1 \cdots d^4x_n = \exp\left[-i\sum_{j=1}^n k_j\alpha\right] \int \exp\left[i\sum_{j=1}^n k_j\xi_j\right] \times \tilde{a}_0(x)\, d^4\xi_1 \cdots d^4\xi_n, \quad (1)$$

where $\xi_j = x_j - x_0 + \alpha$, $(1 \le j \le n)$; $k_j = p_j + iq_j$, $q_j \in V^+$;

$$r_0(k) = \exp\left[i\sum_{j=1}^n k_j\alpha\right] \int \exp\left[-i\sum_{j=1}^n k_j\xi'_j\right] \times \tilde{r}_0(x)\, d^4\xi'_1 \cdots d^4\xi'_n, \quad (2)$$

where $\xi'_j = x_0 - x_j + \alpha$; $k_j = p_j + iq_j$, $q_j \in V^-$; $1 \le j \le n$. These functions are respectively holomorphic in $\mathcal{T}^{S_0}$ and $-\mathcal{T}^{S_0}$, where

$$\mathcal{T}^{S_0} = \{k = p + iq: q_j \in V^+,\ 1 \le j \le n\}.$$

When $q \to 0$, $a_0(p + iq)$ [resp $r_0(p + iq)$] tends, in the sense of S′, to the Fourier transform $a_0(p)$ [resp $r_0(p)$] of $\tilde{a}_0$ (resp $\tilde{r}_0$). The tempered distributions a_0 and r_0 coincide in the real region defined by

$$\left[p: \left(\sum_{j\in J} p_j\right)^2 < 0 \text{ for every subset } J \text{ of } \{1, 2, \cdots, n\}\right],$$

in particular at Jost points. It therefore follows from the theorem of Glaser and Streater[10–12] that $a_0(k)$ and $r_0(k)$ have a common analytic continuation [the restriction of $H(k)$] in the "extended tube"

$$\bigcup_{\Lambda\in L(C)} \Lambda\mathcal{T}^{S_0} = \mathcal{T}'^{S_0}.$$

For every nonzero complex number λ we denote $[\lambda]$ the complex Lorentz transformation defined by

$$[\lambda] = \begin{bmatrix} \frac{1}{2}\left(\lambda + \frac{1}{\lambda}\right) & \frac{1}{2}\left(\lambda - \frac{1}{\lambda}\right) & 0 & 0 \\ \frac{1}{2}\left(\lambda - \frac{1}{\lambda}\right) & \frac{1}{2}\left(\lambda + \frac{1}{\lambda}\right) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

(we are still in our original privileged Lorentz frame). $H([\lambda]k)$ defines a function of λ and k holomorphic in $[\mathbf{C} - \{0\}] \times \mathcal{T}'^{S_0}$, in particular in $[\mathbf{C} - \{0\}] \times \mathcal{T}^{S_0}$, for which we find estimates.

(1) Estimates for $a_0(k)$ in $\mathcal{T}^{S_0}$ and $r_0(k)$ in $-\mathcal{T}^{S_0}$

Using the formulas (1) and (2), it is easy[10] to find estimates for $a_0(k)$ and $r_0(k)$. Defining

$$C_0 = \sup_x \{|\tilde{r}_0(x)|, |\tilde{a}_0(x)|\},$$

we find, for $k_j = p_j + iq_j$:

$$\begin{aligned} |a_0(k)| &\le C_0 \left|\exp\left(-i\alpha\sum_{j=1}^n k_j\right)\right| \times \prod_{j=1}^n \int_{\xi_j\in\bar{V}^+} \exp(-q_j^0\xi_j^0 + \mathbf{q}_j\cdot\boldsymbol{\xi}_j)\, d^4\xi_j \\ &\le C_0 \left|\exp\left(-i\alpha\sum_{j=1}^n k_j\right)\right| \times \prod_{j=1}^n \int_0^\infty \exp[-(q_j^0 - |\mathbf{q}_j|)\xi_j^0]\tfrac{1}{3}\pi(\xi_j^0)^3\, d\xi_j^0 \\ &= 8\pi C_0 \left|\exp\left(-i\alpha\sum_{j=1}^n k_j\right)\right| \prod_{j=1}^n (q_j^0 - |\mathbf{q}_j|)^{-4}; \end{aligned}$$

here $q_j \in V^+$ $(1 \le j \le n)$.

A similar estimate holds for $r_0(k)$ when $q_j \in V^-$, so that:

$$|H(k)| \le C \exp\left|\sum_{j=1}^n \alpha q_j\right| \prod_{j=1}^n (|q_j^0| - |\mathbf{q}_j|)^{-4} \quad (3)$$

for $k_j = p_j + iq_j$, $|q_j^0| > |\mathbf{q}_j|$ $(1 \le j \le n)$. $(C = 8\pi C_0)$.

(2) Use of Certain Jost Points

We introduce the new coordinates

$$u_j = k_j^0 + k_j^1, \quad v_j = k_j^0 - k_j^1,$$

and consider (as in Ref. 10) a closed ball Σ in the real momentum space, $(\mathbf{R}^{4n})$ such that, for every $p = (p_1, \cdots, p_n) \in \Sigma$, the following inequality is satisfied:

$$\min\{u_1, u_2, \cdots, u_n, -v_1, -v_2, \cdots, -v_n\} \ge u > 0$$

(u depending only on Σ). Then, for every $p \in \Sigma$,

$$[\lambda]p \in \mathcal{T}^{S_0} \quad \text{for} \quad \text{Im}\,\lambda > 0,$$
$$[\lambda]p \in -\mathcal{T}^{S_0} \quad \text{for} \quad \text{Im}\,\lambda < 0,$$
$$[\lambda]p \text{ is a Jost point for Im } \lambda = 0, \quad \lambda \neq 0.$$

[11] R. F. Streater, J. Math. Phys. 3, 256 (1962).
[12] R. Jost, *The General Theory of Quantized Fields* (American Mathematical Society, Providence, Rhode Island, 1965).

The inequality (3) yields, for every $p \in \Sigma$, $\operatorname{Im} \lambda \neq 0$,

$$|H([\lambda]p)| \leq C\left\{\exp \tfrac{1}{2}a \sum_{j=1}^{n} \left|\operatorname{Im} \lambda u_j + v_j \operatorname{Im} \frac{1}{\lambda}\right|\right\} \times \left[\frac{1+|\lambda|^2}{u\,|\operatorname{Im}\lambda|}\right]$$

On the other hand, every Jost point being a point of analyticity of $H(k)$ (by the edge-of-the-wedge theorem), we know that $H([\lambda]p)$ is also analytic at real values of $\lambda \neq 0$. If $\operatorname{Im} \rho = 0$, $\rho \neq 0$, we apply the maximum principle to the function (of λ)

$$(\lambda - \tfrac{1}{2}\rho)^{4n}(\lambda - \tfrac{3}{2}\rho)^{4n} H([\lambda]p)$$

in the square

$$\{\lambda : |\operatorname{Re}\lambda - \rho| \leq \tfrac{1}{2}|\rho|,\ |\operatorname{Im}\lambda| \leq \tfrac{1}{2}|\rho|\}$$

to obtain an upper bound for $H([\rho]p)$. Finally, we find that there exist constants $A(\Sigma) > 0$, $B(\Sigma) > 0$ such that, for all $\lambda \neq 0$ and all $p \in \Sigma$

$$|H([\lambda]p)| \leq A(\Sigma) \exp\left[B(\Sigma)\left(|\lambda| + \frac{1}{|\lambda|}\right)\right]. \quad (4)$$

The set $[i]\Sigma$ is compact and contained in $\mathcal{T}^{S_0}$. The function $F(\lambda, k) = H([\lambda]k)$ is analytic in a neighborhood of the set

$$E = \{\lambda \neq 0,\ k : [\lambda]k \in \mathcal{T}^{S_0},\ k \in \mathcal{T}^{S_0}\} \cup \{\lambda,\ k : \lambda \neq 0,\ k \in [i]\,\Sigma\}.$$

This set is invariant under the transformation: $(\lambda, k) \to (\rho\lambda, k)$ for any $\rho > 0$. Using the inequalities (3) and (4) it is easy to show that, for every compact $K \subset E$, there exist constants $A(K) > 0$, $B(K) > 0$ such that, for all $(\lambda, k) \in K$ and $\rho > 0$,

$$F(\rho\lambda, k) < A(K) \exp\left[B(K)\left(\rho + \frac{1}{\rho}\right)\right].$$

It is possible to show (see Appendix I) that the envelope of holomorphy of any neighborhood of E contains the topological product

$$\{\lambda, k : \lambda \neq 0,\ k \in \mathcal{T}^{S_0}\},$$

and that, for every compact K contained in this topological product, there exists a compact $K' \subset E$, such that K is contained in the envelope of holomorphy of every neighborhood of K'; hence, for every function $G(\lambda, k)$ holomorphic in $\{\lambda, k : \lambda \neq 0,\ k \in \mathcal{T}^{S_0}\}$,

$$\max_{(\lambda,k)\in K} |G(\lambda, k)| \leq \max_{(\lambda,k)\in K'} |G(\lambda, k)|.$$

Applying this to $G(\lambda, k) = F(\rho\lambda, k)$, we find that for $(\lambda, k) \in K$ and $\rho > 0$

$$|F(\rho\lambda, k)| < A(K') \exp\left[B(K')\left(\rho + \frac{1}{\rho}\right)\right].$$

In particular, for every compact $K_1 \subset \mathcal{T}^{S_0}$, the set $\{\lambda, k : |\lambda| = 1,\ k \in K_1\}$ is compact. Therefore there exist constants $A(K_1) > 0$, $B(K_1) > 0$ (by abuse of notation) such that, for every $k \in K_1$ and $\lambda \neq 0$, one has

$$|F(\lambda, k)| < A(K_1) \exp\left[B(K_1)\left(|\lambda| + \frac{1}{|\lambda|}\right)\right]. \quad (5)$$

(3) Estimates for $F(\lambda, k)$ in $\{\lambda, k : \lambda \neq 0,\ k \in \mathcal{T}^{S_0}\}$

The estimate provided by (5) is insufficient for our purposes and is improved by the use of the Phragmén–Lindelöf theorem.

Let $k = \{k_1, \cdots, k_n\}$, $k_j = p_j + iq_j$, $u_j = k_j^0 + k_j^1$, $v_j = k_j^0 - k_j^1$, $(q_j^2)^2 + (q_j^3)^2 = r_j^2$ $(1 \leq j \leq n)$. The condition for k to be in $\mathcal{T}^{S_0}$ is: $q_j^0 > |\mathbf{q}_j|$ $(1 \leq j \leq n)$, i.e.,

$$\operatorname{Im} u_j \operatorname{Im} v_j > r_j^2, \quad \operatorname{Im} u_j > 0, \quad \operatorname{Im} v_j > 0.$$

Denote:

$$u = \min_{1\leq j\leq n} \{\operatorname{Im} u_j, \operatorname{Im} v_j\},$$

$$M = \frac{1}{u} \max_{1\leq j\leq n} \{\operatorname{Im} u_j, \operatorname{Im} v_j\},$$

$$M' = \frac{1}{u} \max_{1\leq j\leq n} \{|\operatorname{Re} u_j|, |\operatorname{Re} v_j|\} = \frac{U}{u},$$

$$\kappa = 1 - \max_{1\leq j\leq n} \{r_j^2 (\operatorname{Im} u_j \operatorname{Im} v_j)^{-1}\};\quad 0 < \kappa \leq 1.$$

Let $\lambda = \rho e^{i\theta} \neq 0$; denote $k_j' = p_j' + iq_j' = [\lambda]k_j$. The condition for $[\lambda]k$ to be in $\pm\mathcal{T}^{S_0}$ is

$$(q_j', q_j') = (q_j'^0)^2 - |\mathbf{q}_j'|^2 > 0,$$

i.e.,

$$(\cos\theta \operatorname{Im} u_j + \sin\theta \operatorname{Re} u_j) \times (\cos\theta \operatorname{Im} v_j - \sin\theta \operatorname{Re} v_j) - r_j^2 > 0.$$

Noting that $|\operatorname{Re} u_j| \leq M' \operatorname{Im} u_j$, $|\operatorname{Re} v_j| \leq M' \operatorname{Im} v_j$, we see that, if $|\cos\theta| - M'|\sin\theta| > 0$, we have

$$(q_j', q_j') \geq [|\cos\theta| - M'|\sin\theta|]^2 \operatorname{Im} u_j \operatorname{Im} v_j - r_j^2 \geq \operatorname{Im} u_j \operatorname{Im} v_j [\kappa - (1 + 2M')|\sin\theta|].$$

Let us choose

$$|\sin\theta| = \gamma\kappa/(1 + 2M') \quad \text{with} \quad 0 \leq \gamma < 1$$

(then $|\tan\theta| < [(1 + 2M')^2 - 1]^{-\frac{1}{2}} \leq M'^{-1}$). We find

$$(q_j', q_j') \geq \kappa(1 - \gamma) \operatorname{Im} u_j \operatorname{Im} v_j,$$

$$(|q_j'^0| - |\mathbf{q}_j'|)^{-1} \leq \frac{2\,|q_j'^0|}{\kappa(1-\gamma) \operatorname{Im} u_j \operatorname{Im} v_j}$$

$$\leq \kappa^{-1}(1-\gamma)^{-1}\left[\left(\frac{\rho}{\operatorname{Im} v_j} + \frac{1}{\rho \operatorname{Im} u_j}\right)|\cos\theta| + \frac{\rho + (1/\rho)}{\operatorname{Im} v_j \operatorname{Im} u_j (1 + 2M')} \frac{\gamma\kappa}{} U\right]$$

$$\leq \kappa^{-1}(1-\gamma)^{-1}(\rho + 1/\rho)(2/u). \quad (6)$$

On the other hand,

$$\exp|\alpha q_j'| = \exp a\,|q_j'^0| \leq \exp \tfrac{1}{2}a[\rho + (1/\rho)][Mu + |\sin\theta|\,U]. \quad (7)$$

For $\gamma \leq \gamma_0 = \frac{1}{2}$, $|\sin\theta| \leq \sin\theta_0 = \kappa/(2+4M') < \frac{1}{2}$, we have, using (3), (6), and (7)

$$|F(\lambda,k)| \leq C\left[\frac{4}{\kappa u}\left(|\lambda| + \left|\frac{1}{\lambda}\right|\right)\right]^{4n} \times \exp \tfrac{1}{2}na\left(|\lambda| + \frac{1}{|\lambda|}\right)[Mu + \sin\theta_0 U] \quad (8)$$

(note that $U\sin\theta_0 = [\kappa/(2+4M')]M'u < u$).

To obtain an estimate for $|\sin\theta| > \sin\theta_0$, we first rewrite (8) in the form:

$$|F(\lambda,k)| \leq C\left[\frac{4}{\kappa u}\left(|\lambda| + \frac{1}{|\lambda|}\right)\right]^{4n} \exp \tfrac{1}{2}na\left(|\lambda| + \frac{1}{|\lambda|}\right) \times \left[\frac{2Mu}{\kappa}(1+2M') + U\right]\sin\theta_0$$

or

$$|F(\lambda,k)] \leq C\left[\frac{4}{\kappa u}\left(|\lambda| + \frac{1}{|\lambda|}\right)\right]^{4n} \exp \tfrac{1}{2}na\left(|\lambda| + \frac{1}{|\lambda|}\right) \times \left[\frac{2Mu}{\kappa} + \left(\frac{4M}{\kappa} + 1\right)U\right]\sin\theta_0 \quad (9)$$

(valid for $\lambda = \rho e^{i\theta}$, $|\sin\theta| \leq \sin\theta_0$).

We now apply the following form of the Phragmén–Lindelöf theorem.

Lemma 1: Let f be a function of one complex variable, holomorphic in a neighborhood of the set D:

$$D = \{\lambda = \rho e^{i\varphi},\ \rho > 0,\ \theta_0 \leq \varphi \leq \pi - \theta_0\},$$

where $0 < \theta_0 < \frac{1}{2}\pi$. Assume that there are constants $A_0 > 0$, $B_0 > 0$, $A > 0$, and $B > 0$, such that

(1) for $\lambda = \rho e^{i\theta_0}$ or $\lambda = -\rho e^{-i\theta_0}$, $\rho > 0$,

$$|f(\lambda)| \leq A_0 \exp B_0[|\lambda| + (1/|\lambda|)],$$

(2) for $\lambda \in D$

$$|f(\lambda)| \leq A \exp B[|\lambda| + (1/|\lambda|)].$$

Then, for all $\lambda \in D$,

$$|f(\lambda)| \leq A_0 \exp(B_0/\sin\theta_0)\,\mathrm{Im}\,[\lambda - (1/\lambda)].$$

Proof: (Given here for completeness; it is a paraphrase of the well-known proof of the Phragmén–Lindelöf theorem and can be omitted.) Denote h the function,

$$h(\lambda) = \{\exp i(B_0/\sin\theta_0)[\lambda - (1/\lambda)]\}f(\lambda).$$

For $\lambda = \rho e^{i\varphi}$, $\rho > 0$, $0 \leq \varphi \leq \pi$, we have

$$|h(\lambda)| \leq \{\exp(-B_0/\sin\theta_0)\sin\varphi[\rho + (1/\rho)]\}\,|f(\lambda)|,$$

so that

$$|h(\lambda)| \leq A_0 \quad \text{for} \quad \lambda = \rho e^{i\theta_0} \quad \text{or} \quad -\rho e^{-i\theta_0}, \quad (\rho > 0),$$

$$|h(\lambda)| \leq A \exp B\left(|\lambda| + \frac{1}{|\lambda|}\right) \quad \text{for} \quad \lambda \in D.$$

The function h_ϵ defined by

$$h_\epsilon(\lambda) = \exp[-\epsilon(\lambda/i)^{(\pi-\theta_0)/(\pi-2\theta_0)} - \epsilon(i/\lambda)^{(\pi-\theta_0)/(\pi-2\theta_0)}]h(\lambda), \quad (\epsilon > 0),$$

satisfies, for $\lambda \in D$,

$$|h_\epsilon(\lambda)| \leq A \exp\left\{B\left(|\lambda| + \frac{1}{|\lambda|}\right) - \epsilon\sin\tfrac{1}{2}\theta_0(|\lambda|^\sigma + |\lambda|^{-\sigma})\right\}$$

while, for $\lambda = \rho e^{i\theta_0}$ or $-\rho e^{i\theta_0}$, $(\rho > 0)$,

$$|h_\epsilon(\lambda)| \leq A_0.$$

Here $\sigma = (\pi - \theta_0)/(\pi - 2\theta_0) > 1$. By applying the maximum principle to h_ϵ in sets of the form

$$\{\lambda : \lambda = \rho e^{i\varphi},\ L^{-1} \leq \rho \leq L,\ \theta_0 \leq \varphi \leq \pi - \theta_0\}$$

for all sufficiently large $L > 0$, we find that $|h_\epsilon(\lambda)| \leq A_0$ for all $\lambda \in D$. Letting ϵ tend to zero yields the same inequality for $|h(\lambda)|$. Hence,

$$|f(\lambda)| \leq A_0\left|\exp -i\frac{B_0}{\sin\theta_0}\left(\lambda - \frac{1}{\lambda}\right)\right| = A_0 \exp\frac{B_0}{\sin\theta_0}\,\mathrm{Im}\left(\lambda - \frac{1}{\lambda}\right)$$

in D.

Applying this result to the function of λ

$$F(\lambda,k)/[2i + \lambda - (1/\lambda)]^{4n},$$

and using the inequalities

$$|\lambda| + \left|\frac{1}{\lambda}\right| \leq \left|2i + \lambda - \frac{1}{\lambda}\right| \leq \left|2 + |\lambda| + \frac{1}{|\lambda|}\right|$$

in the upper half-plane, we obtain

$$|F(\lambda,k)| \leq C\left[\frac{4}{\kappa u}\left(2 + |\lambda| + \frac{1}{|\lambda|}\right)\right]^{4n} \times \exp\tfrac{1}{2}na\left(|\lambda| + \frac{1}{|\lambda|}\right)\left[\frac{2Mu}{\kappa} + \left(\frac{4M}{\kappa} + 1\right)U\right]; \quad (10)$$

this inequality is valid for: $|\sin\theta| > \sin\theta_0$. But since the right-hand side also majorizes the right-hand side of (9), it is valid for all $\lambda \neq 0$. For each fixed $\lambda \neq 0$, the inequality (10) shows that $F(\lambda, p + iq)$ tends to a distribution in p when q tends to 0 in any cone where M/κ is bounded. To obtain estimates on this boundary

value, we restrict our attention to points k of the form

$$k = p + i\epsilon, \quad \epsilon_j = (u, 0, 0, 0), \quad u > 0, \quad 1 \leq j \leq n.$$

Let $\varphi \in \mathcal{D}(\mathbf{R}^{4n})$ with support in $\{p : \|p_j\| < R,\ 0 \leq j \leq n\}$ (recall that $p_0 + p_1 + \cdots + p_n = 0$;

$$\|p_j\|^2 = \sum_{\mu=0}^{3} |p_j^\mu|^2\Big).$$

Denote

$$G(\lambda; u; \varphi) = \int H\{[\lambda](p + i\epsilon)\}\varphi(p)\, dp$$

($dp = d^4p_1 \cdots d^4p_n$). This is a $\mathcal{C}^\infty$ function of λ and u ($\lambda \neq 0$, $u > 0$) with

$$(\partial^m/\partial u^m)G(\lambda; u; \varphi) = G(\lambda; u; A^m\varphi),$$

where A is the differential operator,

$$A\varphi(p) = -i\sum_{j=1}^{n} \frac{\partial \varphi}{\partial p_j^0}(p).$$

We can take $M = 1 = \kappa$ so that, by (10),

$$\left|\frac{\partial^m}{\partial u^m} G(\lambda; u; \varphi)\right| \leq C\left[\frac{4}{u}\left(2 + |\lambda| + \frac{1}{|\lambda|}\right)\right]^{4n} \times \left[\exp \tfrac{1}{2}na\left(|\lambda| + \frac{1}{|\lambda|}\right)(2u + 5R)\right] \int |A^m\varphi(p)|\, dp.$$

By standard procedures,[9,13] one finds that $(\partial^m/\partial u^m) \times G(\lambda; u; \varphi)$ is continuous at $u = 0$, and that, for $0 \leq u \leq u_0$,

$$\left|\frac{\partial^m}{\partial u^m} G(\lambda; u; \varphi)\right| \leq C[4(2 + |\lambda| + 1/|\lambda|)]^{4n} \times \left[\exp \tfrac{1}{2}na\left(|\lambda| + \frac{1}{|\lambda|}\right)(2u_0 + 5R)\right] \times \sum_{r=0}^{4n+1} u_0^{r-4n} \int |A^{m+r}\varphi(p)|\, dp \qquad (11)$$

for any $u_0 > 0$. In particular, taking $u_0 = \frac{1}{2}R$, we obtain

$$|G(\lambda; 0; \varphi)| \leq \Gamma_\varphi[2 + |\lambda| + (1/|\lambda|)]^{4n} \times \exp 3naR[|\lambda| + (1/|\lambda|)], \qquad (12)$$

where Γ_φ is a constant depending on φ. The convergence of $G(\lambda; u; \varphi)$ to $G(\lambda; 0; \varphi)$ is uniform when λ remains in a compact of $\mathbf{C} - \{0\}$, as shown by (11), so that $G(\lambda; 0; \varphi)$ is holomorphic in λ for a fixed φ. Moreover, (12) and the use of Cauchy's formula prove that $G(\lambda; 0; \varphi)$ defines a holomorphic vector-valued function of λ with values in the space of distributions of order $\leq 4n + 1$. Thus, there is a distribution $a_0(\lambda; p)$ (distribution in p depending holomorphically on $\lambda \neq 0$) such that, for real $\lambda > 0$,

$$a_0(\lambda; p) = a_0([\lambda]p), \quad a_0(-\lambda; p) = r_0([-\lambda]p),$$

so that we write, by definition,

$$r_0(\lambda; p) = a_0(-\lambda; -p), \quad \lambda \in \mathbf{C} - \{0\}.$$

Using (3), we obtain, for real $\lambda \neq 0$

$$|G(\lambda; 0; \varphi)| \leq \Gamma_\varphi[2 + |\lambda| + (1/|\lambda|)]^{4n}. \qquad (13)$$

[This simply reflects the temperedness of $a_0(p)$ and $r_0(p)$.] To exploit these inequalities, we define

$$K_s\{\tfrac{1}{2}[\lambda - (1/\lambda)]; \varphi\} = \tfrac{1}{2}\{G(\lambda; 0; \varphi) + G[-(1/\lambda); 0; \varphi]\},$$

$$K_a\{\tfrac{1}{2}[\lambda - (1/\lambda)]; \varphi\} = \tfrac{1}{2}[\lambda + (1/\lambda)]^{-1} [G(\lambda; 0; \varphi) - G(-1/\lambda; 0; \varphi)].$$

The functions $K_s(\zeta; \varphi)$ and $K_a(\zeta; \varphi)$ are entire functions of the variable ζ. To see this, consider the holomorphic function of two variables z_1, z_2, defined, for $z_1 + z_2 \neq 0$ and $z_1 - z_2 \neq 0$, by

$$G(z_1 + z_2; 0; \varphi) - G(z_1 - z_2; 0; \varphi).$$

Because it is antisymmetric in z_2, it can be written $z_2 g(z_1, z_2^2)$. Setting $z_1 = \frac{1}{2}[\lambda - (1/\lambda)]$, $z_2 = \frac{1}{2}[\lambda + (1/\lambda)]$ we obtain

$$\tfrac{1}{2}[\lambda + (1/\lambda)]g\{\tfrac{1}{2}[\lambda - (1/\lambda)],\ \tfrac{1}{4}[\lambda - (1/\lambda)]^2 + 1\}.$$

To obtain bounds for these functions one can, for example, notice that

$$(\partial/\partial\lambda)\, G(\lambda; 0; \varphi) = G(\lambda; 0; -M^{01}\varphi),$$

$$M^{01}\varphi(p) = \sum_{j=1}^{n}\left(p_j^1 \frac{\partial}{\partial p_j^0} - p_j^0 \frac{\partial}{\partial p_j^1}\right)\varphi(p) = \frac{\partial}{\partial\lambda}\, \varphi([\lambda^{-1}]p)\big|_{\lambda=1}.$$

Hence

$$G(\lambda; 0; \psi) - G\left(-\frac{1}{\lambda}; 0; \varphi\right) = \int_{-1/\lambda}^{\lambda} G(\mu; 0; -M^{01}\varphi)\, d\mu,$$

the integral being taken on any contour avoiding 0. There always exists an arc $(\lambda, -1/\lambda)$ of the circle passing through λ, $-1/\lambda$, $\pm i$, $\bar{\lambda}$, $-1/\bar{\lambda}$, which is contained in $\{\mu : (|\mu| - |1/\lambda|)(|\lambda| - |\mu|) \geq 0\}$. We choose this arc as our contour of integration; its length is always $\leq \pi\,|\lambda + (1/\lambda)|$; every point μ on the contour is such that $|\mu| + (1/|\mu|) < 2[|\lambda| + (1/|\lambda|)]$. Finally, using (12), we find

$$|K_a(\zeta)| \leq \tfrac{1}{2}\pi\Gamma_{M^{01}\varphi}(4\,|\zeta| + 6)^{4n} \exp 12naR(|\zeta| + 1). \qquad (14)$$

[18] M. Zerner, "Les fonctions holomorphes a valeurs vectorielles et leurs valeurs aux bords," Lecture Notes, Orsay (1961).

It is even simpler to obtain a bound for $K_s(\zeta)$

$$|K_s(\zeta)| \leq \Gamma_\varphi(2|\zeta| + 4)^{4n} \exp 6naR(|\zeta| + 1). \quad (15)$$

For real λ, $|\lambda + (1/\lambda)| \geq 1$ so that

$$|K_a(\zeta)|, |K_s(\zeta)| \leq \Gamma_\varphi(2|\zeta| + 4)^{4n}, \quad \zeta \text{ real}. \quad (16)$$

(For real ζ, λ and $1/\lambda$ are necessarily real.) These inequalities show that for real ζ, $K_s(\zeta)$ and $K_a(\zeta)$ are Fourier transforms of distributions with support in $\{\tau:|\tau| \leq 12naR\}$. In other words

$$\int K_a(\xi;\varphi)\psi(\xi)\,d\xi = \int K_s(\xi;\varphi)\psi(\xi)\,d\xi = 0$$

for every $\psi \in \mathcal{S}(\mathbf{R})$ such that the Fourier transform $\tilde{\psi}$ of ψ has its support in $\{\tau:|\tau| > 12naR\}$. This can be rewritten

$$0 = \int_0^\infty \psi\left[\frac{1}{2}\left(\lambda - \frac{1}{\lambda}\right)\right]\left(1 + \frac{1}{\lambda^2}\right) d\lambda \times \left\{\int \frac{1}{2}\left[a_0([\lambda]p) + r_0\left(\left[-\frac{1}{\lambda}\right]p\right)\right]\varphi(p)\,dp\right\},$$

$$0 = \int_0^\infty \psi\left[\frac{1}{2}\left(\lambda - \frac{1}{\lambda}\right)\right]\frac{d\lambda}{\lambda} \times \left\{\int \frac{1}{2}\left[a_0([\lambda]p) - r_0\left(\left[-\frac{1}{\lambda}\right]p\right)\right]\varphi(p)\,dp\right\}.$$

Note that these integrals are known to converge because of the temperedness of a_0 and r_0.

B. Other Generalized Retarded Functions; Time-Ordered Function

We recall that the grf are labeled as follows[7]: consider the subspace Σ_{n+1} of $\mathbf{R}^{n+1}$ defined by

$$\left\{s = (s_0, s_1, \cdots, s_n): \sum_{j=0}^{n} s_j = 0\right\};$$

Σ_{n+1} is divided into cells by the hyperplanes of the form $\{s:s_I = 0\}$, s_I denoting $\sum_{j\in I} s_j$ for every proper nonempty subset I of $\{0, 1, \cdots, n\}$. Each cell S the subset of Σ_{n+1} defined by attributing a certain definite sign to each partial sum s_I. For example,

$$S_0 = \left\{s \in \Sigma_{n+1}: s_I > 0 \quad \text{if} \quad 0 \notin I,\ s_I < 0 \quad \text{if} \quad 0 \in I\right\}.$$

To each cell S is associated a grf denoted r^S; this is a tempered distribution in momentum space, the boundary value of a function $r^S(p + iq)$ holomorphic in the tube $\mathcal{T}^S$.

$$\mathcal{T}^S = \{k = p + iq: q_I \in V^+ \quad \text{if} \quad s_I > 0 \quad \text{in} \quad S\}.$$

Obviously $\mathcal{T}^{-S} = -\mathcal{T}^S$. Actually all functions $r^s(p + iq)$ are branches of $H(p + iq)$, i.e., $r^S(p + iq) = H(p + iq) \mid \mathcal{T}^S$. For more details we refer the reader to Refs. 2, 6, 7, 9, and 14.

Considerations analogous to those of the preceding subsection lead to the following result.

There exists, for each cell S, a distribution $r^S(\lambda; p)$ on $\mathbf{R}^{4n}$ depending holomorphically on $\lambda \in \mathbf{C} - \{0\}$ with the following properties:

(1) $r^S(\lambda; p) = r^{-S}(-\lambda; -p)$;

(2) $r^S(\lambda; p) = r^S([\lambda]p)$ for real $\lambda > 0$; by (1) this implies $r^S(\lambda; p) = r^{-S}([\lambda]p)$ for real $\lambda < 0$;

(3) there exist constants $A > 0$, $B > 0$, such that, if $\varphi \in \mathcal{D}(\mathbf{R}^{4n})$ with support in

$$\{p: \|p_j\| < R,\ 0 \leq j \leq n\},$$

one has

$$\left|\int r^S(\lambda; p)\varphi(p)\,dp\right| \leq A\left(1 + |\lambda| + \frac{1}{|\lambda|}\right)^{4n} \times \left[\exp BR\left(|\lambda| + \frac{1}{|\lambda|}\right)\right] N(\varphi), \quad (17)$$

$\varphi \to N(\varphi)$ being a certain norm in $\mathcal{D}(\mathbf{R}^{4n})$;

(4) if $\psi \in \mathcal{S}(\mathbf{R})$ is a function of a real variable whose Fourier transform vanishes in $\{\tau; |\tau| \leq 4\,BR\}$, then

$$\int_0^\infty d\lambda\psi\left[\frac{1}{2}\left(\lambda - \frac{1}{\lambda}\right)\right]\left(1 + \frac{1}{\lambda^2}\right) \times \left\{\int \frac{1}{2}\left[r^S([\lambda]p) + r^{-S}\left(\left[-\frac{1}{\lambda}\right]p\right)\right]\varphi(p)\,dp\right\} = 0, \quad (18)$$

$$\int_0^\infty \frac{d\lambda}{\lambda}\psi\left[\frac{1}{2}\left(\lambda - \frac{1}{\lambda}\right)\right] \times \left\{\int \frac{1}{2}\left[r^S([\lambda]p) - r^{-S}\left(\left[-\frac{1}{\lambda}\right]p\right)\right]\varphi(p)\,dp\right\} = 0 \quad (19)$$

for all $\varphi \in \mathcal{D}(\mathbf{R}^{4n})$ with support in $\{p: \|p_j\| \leq R,\ 0 \leq j \leq n\}$. We omit the detailed proof of this statement which uses no other ideas than those of A.

We need the analog of (18) and (19) for the "T-product" or "time-ordered function." This distribution $\tau(p)$ in momentum space is the Fourier transform of

$$\tilde{\tau}(x) = \sum_{\pi\in\gamma} {}_{n+1}\,\theta(x^0_{\pi 0} - x^0_{\pi 1}) \cdots \theta(x^0_{\pi(n-1)} - x^0_{\pi n}) \times [\Omega, A_{\pi 0}(x_{\pi 0}) \cdots A_{\pi n}(x_{\pi n})\Omega].$$

The "truncated time-ordered function" τ^T is the Fourier transform of $\tilde{\tau}^T(x)$, obtained by replacing, in the above formulas, all Wightman functions by the corresponding truncated Wightman functions.[2] It is well known[7] that the grf defined from truncated Wightman functions coincide with those defined from the ordinary vev. Using this fact and Ruelle's method,[7]

[14] J. Bros, *High-Energy Physics and Elementary Particles* (IAEA, Vienna, 1965).

assuming a strictly positive minimum mass μ for all states orthogonal to the vacuum, it is easy to see that $r^S(p)$ coincides with $\tau^T(p)$ in the domain defined by

$$\{p: p_I \notin \bar{V}^- \text{ or } p_I^2 < \mu^2 \text{ for all } I \text{ such that } s_I > 0 \text{ in } S\},$$

$$\left(p_I = \sum_{j \in I} p_j\right),$$

which contains a neighborhood of

$$\{p: p^0 \in \bar{S}\}.$$

The family of these open sets covers the whole real momentum space; to summarize: *each point p in real momentum space is the center of a real open ball in which τ^T coincides with at least one r^S*. Obviously, in a neighborhood of $-p$, τ^T coincides with r^{-S}. Let φ be a function in $\mathcal{D}(\mathbf{R}^{4n})$ with support in $\{p: \|p_j\| < R, 0 \leq j \leq n\}$. Using a partition of the unit we deduce from (18) and (19):

$$\int_0^\infty d\lambda \psi\left[\frac{1}{2}\left(\lambda - \frac{1}{\lambda}\right)\right](1 + \lambda^{-2}) \times \left\{\int \tfrac{1}{2}[\tau^T([\lambda]p) + \tau^T\left(\left[-\frac{1}{\lambda}\right]p\right)\varphi(p)\, dp\right\} = 0, \tag{20}$$

$$\int_0^\infty \frac{d\lambda}{\lambda} \psi\left[\frac{1}{2}\left(\lambda - \frac{1}{\lambda}\right)\right] \times \left\{\int \frac{1}{2}\left[\tau^T([\lambda]p) - \tau^T\left(\left[-\frac{1}{\lambda}\right]p\right)\right]\varphi(p)\, dp = 0, \right. \tag{21}$$

for all ψ in $\mathcal{S}(\mathbf{R})$ whose Fourier transform vanishes in $\{\tau: |\tau| \leq 4BR\}$. τ differs from τ^T by sums of (tensor) products of delta-functions and truncated time-ordered functions of fewer variables; thus (20) and (21) continue to hold if τ^T is replaced by τ in these formulas. (The constant B is a nondecreasing function of n.)

3. ADDITIONAL ASSUMPTIONS AND SOME CONSEQUENCES (SPINLESS CASE)

A. First additional assumption

We assume that the Hilbert space $\mathcal{H}$ decomposes into a direct sum of (charge) superselection sectors,

$$\mathcal{H} = \left\{\bigoplus_{j \in G} \mathcal{H}_j\right\},$$

where G is an Abelian group with additive notation. Each sector $\mathcal{H}_j$ is itself assumed to be a direct sum,

$$\mathcal{H}_0 = \mathbf{C}\Omega + \bigoplus_{k=1}^{N_0} \mathcal{H}_{0k} \oplus \mathcal{H}_0^c,$$

$$\mathcal{H}_j = \bigoplus_{k=1}^{N_j} \mathcal{H}_{jk} \oplus \mathcal{H}_j^c \quad (j \neq 0),$$

where the subspaces $\mathcal{H}_{jk}$, $\mathcal{H}_j^c$ are all invariant under $U(a, \Lambda)$; $P^\mu P_\mu \,|\, \mathcal{H}_j^c \geq M_j^2 > 0$; N_j is at most countably infinite; the restriction of the representation $\{a, \Lambda\} \to U(a, \Lambda)$ to the subspace $\mathcal{H}_{jk}$ is irreducible and unitarily equivalent to $[m_{jk}, 0]$, and

$$m_{j1} < m_{j2} < \cdots < m_{jr} < \cdots < M_j \text{ for all } j.$$

B. Second additional assumption

For each $j \in G$ and each domain $B \subset \mathbf{R}^4$, let

$$R_j(B) = \{A \in R(B), A\Omega \in \mathcal{H}_j\}.$$

We assume that

$$\bigcup_{B \text{ bounded}} R_j(B)\Omega \text{ is dense in } \mathcal{H}_j \quad \text{and} \quad [R_j(B)]^* = R_{-j}(B).$$

It then follows, by the Reeh–Schlieder theorem, that $\overline{R_j(B)\Omega} = \mathcal{H}_j$ for every domain B in $\mathbf{R}^4$. A partial justification of these assumptions can be found in the work of Borchers,[5] where it is also shown that they imply $m_{jk} = m_{-jk}$ for all j and k.

Let E_{jk} be the projector onto $\mathcal{H}_{jk}$, E_j^c the projector onto $\mathcal{H}_j^c$, E_Ω the projector onto $\mathbf{C}\Omega$. Define

$$R_{jk}(B) = \{A \in R_j(B), (\Omega, A\Omega) = 0 \quad (\mathcal{H}_{jr}, A\Omega) = 0 \text{ for all } r, < k\}.$$

Let $A(0) \in R_j(B)$, $A(x) = U(x, 1)A(0)U(x, 1)^{-1}$. Then

$$E_{jl} \prod_{r<k} (\Box_x + m_{jr}^2)(-m_{jk}^2 + m_{jr}^2)^{-1} A(x)\Omega = \begin{cases} E_{jk}A(x)\Omega & \text{if } l = k, \\ 0 & \text{if } l < k. \end{cases}$$

Therefore $E_{jk}\overline{R_{jk}(B)\Omega} = \mathcal{H}_{jk}$. We denote $R^\infty(B)$, $R_j^\infty(B)$, and $R_{jk}^\infty(B)$ as the subsets of $R(B)$, $R_j(B)$, and $R_{jk}(B)$, respectively, consisting of elements A such that there exists a $\mathcal{C}^\infty$ function φ with compact support K on the Poincaré group and an operator $A_1 \in R(B_1)$, $R_j(B_1)$, or $R_{jk}(B_1)$ (respectively), such that

$$A = \int \varphi(a, \Lambda)U(a, \Lambda)A_1U(a, \Lambda)^{-1}\, da\, d\Lambda, \quad \text{and} \quad K \cdot B_1 \subset B.$$

(Here $da\, d\Lambda$ stands for the invariant measure on the Poincaré group; the integral is meant in the weak sense.) The Reeh–Schlieder theorem shows that $R^\infty(B)\Omega$, $R_j^\infty(B)\Omega$, and $E_{jk}R_{jk}^\infty(B)\Omega$ are respectively dense in $\mathcal{H}$, $\mathcal{H}_j$, and $\mathcal{H}_{jk}$.

Since the restriction of the representation U to the subspace $\mathcal{H}_{jk}$ is unitarily equivalent to $[m_{jk}, 0]$, there

exists[2,15–17] a linear unitary map V_{jk} from $\mathcal{H}_{jk}$ onto the space of complex functions f on the submanifold of $\mathbf{R}^4$,

$$\{p \in \mathbf{R}^4 : (p, p) = m_{jk}^2,\ p^0 > 0\}$$

(upper sheet of mass hyperboloid), equipped with the scalar product

$$(f, g) = \int f^*(p) g(p)\, d\Omega(p),$$

$$d\Omega(\omega(\mathbf{p}), \mathbf{p}) = d^3\mathbf{p}/2\omega(\mathbf{p}); \quad \omega(\mathbf{p}) = (\mathbf{p}^2 + m_{jk}^2)^{\frac{1}{2}}.$$

Moreover,

$$V_{jk}\, U(a, \Lambda)\Psi(p) = e^{ipa}(V_{jk}\Psi)(\Lambda^{-1}p)$$

for all $\Psi \in \mathcal{H}_{jk}$.

V_{jk} is uniquely defined up to multiplication by a constant of modulus one.

Let $A_1(0) \in R_{jk}(B_1)$, $A_2(0) \in R_{jk}(B_2)$ [so that $A_1^*(0) \in R_{-jk}(B_1)$, $A_2^*(0) \in R_{-jk}(B_2)$], where B_1 and B_2 are bounded domains; denote

$$f_1(p) = [V_{jk}E_{jk}A_1(0)\Omega](p),$$
$$f_2(p) = [V_{jk}E_{jk}A_2(0)\Omega](p),$$
$$g_1^*(p) = [V_{-jk}E_{-jk}A_1^*(0)\Omega](p),$$
$$g_2^*(p) = [V_{jk}E_{-jk}A_2^*(0)\Omega](p).$$

We have

$$\int f_2^*(p) f_1(\Lambda p) e^{ipa}\, d\Omega(p) = (\Omega, A_2^*(0)U(a, \Lambda)E_{jk}A_1(0)\Omega).$$

In particular,

$$f_2^*(p) f_1(p)\delta(p^2 - m_{jk}^2)\theta(p^0) = \frac{1}{(2\pi)^4}\int (\Omega, A_2^*(0)E_{jk}A_1(x)\Omega)e^{-ipx}\, dx$$

(see Ref. 2).

Let us denote

$$\tilde{W}^+(x) = (\Omega, A_2^*(0)A_1(x)\Omega),$$
$$\tilde{W}^-(x) = (\Omega, A_1(x)A_2^*(0)\Omega),$$
$$\tilde{C}(x) = (\Omega, [A_2^*(0), A_1(x)]\Omega),$$
$$W^{\pm}(p) = (2\pi)^{-4}\int e^{-ipx}\tilde{W}^{\pm}(x)\, d^4x,$$
$$C(p) = (2\pi)^{-4}\int e^{-ipx}\tilde{C}(x)\, dx$$

(in the sense of distributions). $W^{\pm}$ has its support contained in $V^{\pm}$; it coincides with $\pm C(p)$ in $V^{\pm}$.

Moreover, the formula

$$W^+(p) = (2\pi)^{-4}\int e^{-ipx}(A_2(0)\Omega, U(x, 1)A_1(0)\Omega)\, dx$$

shows that $W^+(p)$ is a spectral *measure* of the energy-momentum operator; its support is therefore contained in the set

$$\{p : (p, p) = m_{jk}^2,\ p^0 > 0\} \cup \{p : (p, p) \geq m_{jk+1}^2,\ p^0 > 0\}.$$

Let α be a C^∞ function of a real variable with support contained in the open interval $]0, m_{j,k+1}^2[$ and such that $\alpha(m_{jk}^2) = 1$. Then

$$(2\pi)^{-4}\int e^{-ipx}(\Omega, A_2^*(0)E_{jk}U(x, 1)A_1(0)\Omega)\, dx = \alpha(p^2)W^+(p),$$

so that $f_2^*(p)f_1(p)\delta(p^2 - m_{jk}^2)\theta(p^0)$ is the restriction of $W^+(p)$, to any neighborhood of the hyperboloid $\{p : (p, p) = m_{jk}^2,\ p^0 > 0\}$ not intersecting $\{p : (p, p) \geq m_{j,k+1}^2\}$. Similarly,

$$g_1(-p)g_2^*(-p)\delta(p^2 - m_{-j,k}^2)\theta(-p^0)$$

is the restriction of $W^-(p)$ to any neighborhood of $\{p : (p, p) = m_{-jk}^2, p^0 < 0\}$ not intersecting

$$\{p : (p, p) \geq m_{-j,k+1}^2\}.$$

On the other hand, the support of $\tilde{C}(x)$ is contained in $\bar{B}_2 - \bar{B}_1 + \bar{V}^+ \cup \bar{V}^-$. Denote

$$\tilde{a}(x) = \theta(-x^0)\tilde{C}(x), \quad a(p) = (2\pi)^{-4}\int e^{-ipx}\tilde{a}(x)\, dx,$$

$$\tilde{r}(x) = -\theta(x^0)\tilde{C}(x), \quad r(p) = (2\pi)^{-4}\int e^{-ipx}\tilde{r}(x)\, dx,$$

$$a'(p) = (p^2 - m_{jk}^2)a(p), \quad r'(p) = (p^2 - m_{jk}^2)r(p);$$

$a'(p)$ [respectively, $r'(p)$] is the boundary value of a function $H^+(k)$ [respectively, $H^-(k)$] holomorphic in the forward tube (respectively, the backward tube). $a'(p)$ and $r'(p)$ coincide in $\{p : (p, p)^2 < m_{j,k+1}^2\}$ so that (by the Jost–Lehmann–Dyson representation or by the theorem of Glaser and Streater) $H^+(k)$ and $H^-(k)$ are restrictions of a function H holomorphic in

$$\{k = p + iq;\ (k, k) \notin m_{j,k+1}^2 + \mathbf{R}^+\},$$

$$C(p) = a(p) - r(p) = \lim_{\substack{\epsilon \to 0 \\ \epsilon > 0}} \left(\frac{H(p + i\epsilon q)}{(p + i\epsilon q)^2 - m_{jk}^2} - \frac{H(p - i\epsilon q)}{(p - i\epsilon q)^2 - m_{jk}^2} \right)$$

with $q \in V^+$. This coincides with

$$2i\pi H(p)\delta(p^2 - m_{jk}^2)\epsilon(p^0)$$

in any neighborhood of $\{p : p^2 = m_{jk}^2\}$ not intersecting

[15] E. P. Wigner, Ann. Math. **40**, 149 (1939); reprinted in F. J. Dyson, *Symmetry Groups in Nuclear and Particle Physics* (W. A. Benjamin, Inc., New York, 1966).

[16] R. F. Streater and A. S. Wightman, *PCT, Spin and Statistics and All That* (W. A. Benjamin, Inc., New York, 1964).

[17] A. S. Wightman, in *Relations de dispersion et particules elementaires*, C. deWitt and R. Omnes, Eds (Hermann & Cie, Paris, 1960).

$\{p: p^2 \geq m_{jk+1}^2\}$. It follows that $f_2^*(p)f_1(p)$ and $g_1(-p)g_2^*(-p)$ are restrictions of the same function holomorphic on the whole complex hyperboloid $\{k:(k,k) = m_{jk}^2\}$. This is equivalent to the following property.

For each p such that $(p,p) = m_{jk}^2$, $p^0 > 0$, and for each $\Lambda_0 \in L_+^\uparrow$ the two functions

$$\Lambda \to f_2^*(\Lambda_0^{-1}\Lambda^{-1}p)f_1(\Lambda^{-1}p), \quad \Lambda \in L_+^\uparrow,$$
$$-\Lambda \to g_2^*(\Lambda_0^{-1}\Lambda^{-1}p)g_1(\Lambda^{-1}p), \quad -\Lambda \in L_+^\downarrow$$

are the restrictions to $L_+^\uparrow$ and $L_+^\downarrow$ (respectively) of a function holomorphic on the complex Lorentz group $L_+(\mathbf{C})$. [Note that if $\Lambda_0 f_2$ denotes the function $p \to f_2(\Lambda_0^{-1}p)$, we have

$$\Lambda_0 f_2 = V_{jk}E_{jk}A_2'\Omega,$$

where

$$A_2' = U(0,\Lambda_0)A_2(0)U(0,\Lambda_0)^{-1} \in R_{jk}(\Lambda_0 B_2);$$

if $A_2(0) \in R_{jk}^\infty(B_2)$, then $A_2' \in R_{jk}^\infty(\Lambda_0 B_2)$]. (Note: A refinement of the above argument leads to the result of Borchers[5]: $m_{jk}^2 = m_{-jk}^2$; an equivalent proof uses the Jost–Lehmann–Dyson integral formula). We also notice that, if $A_1(0) \in R_{jk}^\infty(B_1)$, then f_1 is in $\mathcal{S}$ on the hyperboloid. This means that $[\mathbf{p} \to f_1(\omega(\mathbf{p}), \mathbf{p})] \in \mathcal{S}(\mathbf{R}^3)$ or, equivalently, that $[\Lambda \to f_1(\Lambda p)] \in \mathcal{S}(L_+^\uparrow)$.

We need the following result.

Lemma 2: One can normalize V_{jk} and V_{-jk} so that, for every bounded domain B_1 and every $A_1(0) \in R_{jk}^\infty(B_1)$ for which $f_1 = V_{jk}E_{jk}A_1(0)\Omega$, $g_1^* = V_{-jk}E_{-jk}A_1^*(0)\Omega$, the two functions

$$\Lambda \to f_1(\Lambda p), \quad (\Lambda \in L_+^\uparrow),$$
$$-\Lambda \to g_1(\Lambda p), \quad (-\Lambda \in L_+^\downarrow)$$

[where p is real and satisfies $(p,p) = m_{jk}^2, p^0 > 0$] are the restrictions to $L_+^\uparrow$ and $L_+^\downarrow$ (respectively) of a function holomorphic on the whole complex Lorentz group $L_+(\mathbf{C})$, and $\mathcal{C}^\infty$ in Λ and p in

$$L_+(\mathbf{C}) \times \{p: p^2 = m_{jk}^2, p^0 > 0\}.$$

The proof of this lemma is given in Appendix 2.

4. USE OF THE ASYMPTOTIC CONDITION

We now consider, as in Sec. 2, a finite sequence of $(n+1)$ HA-fields, $A_0, \cdots, A_n$ with the following properties.

(1) $A_j(0) \in R_{d_j k_j}^\infty(B_j)$ where B_j is a bounded domain in $\mathbf{R}^4$, $d_j \in G$, $0 \leq j \leq n$.

(2) $V_{d_j k_j}E_{d_j k_j}A_j(0)\Omega = f_j \neq 0 \quad (0 \leq j \leq n)$,

$V_{-d_j k_j}E_{-d_j k_j}A_j^*(0)\Omega = \bar{g}_j \neq 0$ (as a consequence).

f_j and g_j are functions (in $\mathcal{S}$) on the hyperboloid $\{p: p^2 = m_j^2, p^0 > 0\}$ (m_j is used instead of $m_{d_j k_j}$ in the rest of this section).

The fields $A_0, \cdots, A_n$ then describe the particles $(0, k_0), \cdots, (n, k_n)$ and their antiparticles. It has been shown by Araki[2] (see also the papers of Haag,[1] Ruelle,[18] Araki, Hepp, and Ruelle,[19] and Hepp[20,21]) that asymptotic outgoing and ingoing free fields can be defined for each of these particles and that the corresponding S-matrix is Lorentz invariant.

Let S_1 and S_2 be the two distributions (kernels of the S-matrix) defined, on the manifold

$$\left\{p = (p_0, \cdots, p_n): \sum_{j=0}^{r} p_j = \sum_{l=r+1}^{n} p_l; \right.$$
$$\left. p_h^2 = m_h^2, p_h^0 > 0 \quad (0 \leq h \leq n)\right\},$$

by

$$S_1(p_0, \cdots, p_r; p_{r+1}, \cdots, p_n)\delta\left(\sum_{j=0}^{n} p_j - \sum_{l=r+1}^{n} p_l\right)$$
$$= (\Omega, a_{0\,\mathrm{out}}(p_0)\cdots a_{r\,\mathrm{out}}(p_r)b_{r+1\,\mathrm{in}}^*(p_{r+1})\cdots b_{n\,\mathrm{in}}^*(p_n)\Omega),$$

$$S_2(p_n, \cdots, p_{r+1}; p_r, \cdots, p_0)\delta\left(\sum_{j=0}^{r} p_j - \sum_{l=r+1}^{n} p_l\right)$$
$$= (\Omega, a_{n\,\mathrm{out}}(p_n)\cdots a_{r+1\,\mathrm{out}}(p_{r+1})b_{r\,\mathrm{in}}^*(p_r)\cdots b_{0\,\mathrm{in}}^*(p_0)\Omega).$$

The creation and annihilation operators are those of the asymptotic fields $\phi_{h\,\mathrm{out}}$ and $\phi_{h\,\mathrm{in}}$ associated with the HA-fields A_k,

$$\phi_{h\,\mathrm{ex}}(x) = (2\pi)^{-\frac{3}{2}}\int d^4p\,\delta(p^2 - m_h^2)\theta(p^0)$$
$$\times [a_{h\,\mathrm{ex}}(p)e^{-ipx} + b_{h\,\mathrm{ex}}^*(p)e^{ipx}].$$

We have, for $\Lambda \in L_+^\uparrow$, $S_1(\Lambda p) = S_1(p)$; $S_r(\Lambda p) = S_2(p)$.

Let $\varphi_0, \cdots, \varphi_n$ be $\mathcal{C}^\infty$ functions with compact support in $\mathbf{R}^4$ such that: (1) The support of φ_h is contained in $\{p \in \mathbf{R}^4: 0 < (p,p) < m_{d_h,k_h+1}, p^0 > 0\}$ for $0 \leq h \leq n$. (2) For $0 \leq j < j' \leq r$, $r+1 \leq l < l' \leq n$, φ_j and $\varphi_{j'}$ (respectively, φ_l and $\varphi_{l'}$) are "non-overlapping,"[20,21] that is, if $p_j \in \mathrm{supp}\ \varphi_j$ and $p_{j'} \in \mathrm{supp}\ \varphi_{j'}$, then

$$\omega_j(\mathbf{p}_j)\mathbf{p}_{j'} \neq \omega_{j'}(\mathbf{p}_{j'})\mathbf{p}_j,$$

where $\omega_j(\mathbf{p}_j) = [\mathbf{p}_j^2 + m_j^2]^{\frac{1}{2}}$; this quantity will be denoted ω_j in the sequel.

Using the asymptotic theory,[2,20] it can be shown

[18] D. Ruelle, Helv. Phys. Acta **35**, 147 (1962).
[19] H. Araki, K. Hepp, and D. Ruelle, Helv. Phys. Acta **35**, 164 (1962).
[20] K. Hepp, Commun. Math. Phys. **1**, 95 (1965).
[21] K. Hepp, in *Brandeis University Summer Institute of Theoretical Physics, Vol. 1, Axiomatic Field Theory*, M. Chrétien and S. Deser, Eds. (Gordon and Breach Science Publishers, Inc., New York, 1967).

that (for $\Lambda \in L_+^\uparrow$)

$$F_1(\Lambda) = \int S_1(p_0, \cdots, p_r; p_{r+1}, \cdots, p_n) \times \delta\left(\sum_{j=0}^{r} p_j - \sum_{l=r+1}^{n} p_l\right) \times \prod_{0\le j\le r} g_j(\Lambda p_j)\varphi_j(p_j)\delta(p_j^2 - m_j^2)\theta(p_j^0) \times \prod_{r+1\le l\le n} f_l(\Lambda p_l)\varphi_l(p_l)\delta(p_l^2 - m_l^2)\theta(p_l^0)\, dp_0 \cdots dp_n$$

is the limit, when $t \to \infty$ of

$$G_1(\Lambda; t) = \int \tau(\Lambda p_0, \cdots, \Lambda p_n)\delta\Big(\sum_{0\le j\le n} p_j\Big) \prod_{0\le j\le r} \varphi_j(-p_j) \times \exp\,[i(p_j^0 + \omega_j)t] \prod_{r+1\le l\le n} \varphi_l(p_l) \times \exp\,[-i(p_l^0 - \omega_l)t]\, dp_0 \cdots dp_n.$$

Furthermore, setting $\Lambda = [\lambda] (\lambda > 0)$, one finds that

$$|F_1([\lambda]) - G_1([\lambda]; t)| < M_1(\lambda)K_1(t),$$

where $K_1(t) \to 0$ and $M_1(\lambda)$ does not increase faster than a power of $[\lambda + (1/\lambda)]$ when $\lambda \to \infty$ or $\lambda \to 0$.

It follows that, if ψ is a function in $\mathcal{S}(\mathbf{R})$ we have

$$\lim_{t\to+\infty} \int_0^\infty d\lambda \psi\left[\frac{1}{2}\left(\lambda - \frac{1}{\lambda}\right)\right](1 + \lambda^{-2}) \times \{F_1([\lambda]) - G_1([\lambda]; t)\} = 0,$$

$$\lim_{t\to+\infty} \int_0^\infty \frac{d\lambda}{\lambda} \psi\left[\frac{1}{2}\left(\lambda - \frac{1}{\lambda}\right)\right]\{F_1([\lambda]) - G_1([\lambda]; t)\} = 0.$$

Similarly, denoting

$$F_2(\Lambda) = \int S_2(p_n, \cdots, p_{r+1}; p_r, \cdots, p_0) \times \delta\left(\sum_{j=0}^{r} p_j - \sum_{l=r+1}^{n} p_l\right) \times \prod_{0\le j\le r} f_j(\Lambda p_j)\varphi_j(p_j)\delta(p_j^2 - m_j^2)\theta(p_j^0) \times \prod_{r+1\le l\le n} g_l(\Lambda p_l)\varphi_l(p_l)\delta(p_l^2 - m_l^2)\theta(p_l^2)\, dp_0 \cdots dp_n$$

and

$$G_2(\Lambda; t) = \int \tau(-\Lambda p_0, \cdots, -\Lambda p_n)\delta\Big(\sum_{0\le j\le n} p_j\Big) \times \prod_{0\le j\le r} \varphi_j(-p_j) \exp\,[i(p_j^0 + \omega_j)t] \times \prod_{r+1\le l\le n} \varphi_l(p_l) \exp\,[-i(p_l^0 - \omega_l)t],$$

one finds (setting $\Lambda = [\lambda^{-1}]$), for $\psi \in \mathcal{S}(\mathbf{R})$,

$$\lim_{t\to+\infty} \int_0^\infty d\lambda \psi\left[\frac{1}{2}\left(\lambda - \frac{1}{\lambda}\right)\right](1 + \lambda^{-2}) \times \left\{F_2\left(\left[\frac{1}{\lambda}\right]\right) - G_2\left(\left[\frac{1}{\lambda}\right]; t\right)\right\} = 0,$$

$$\lim_{t\to-\infty} \int_0^\infty \frac{d\lambda}{\lambda} \psi\left[\frac{1}{2}\left(\lambda - \frac{1}{\lambda}\right)\right] \times \left\{F_2\left(\left[\frac{1}{\lambda}\right]\right) - G_2\left(\left[\frac{1}{\lambda}\right]; t\right)\right\} = 0.$$

Using the results of Sec. 2, we see that there exists a number $R > 0$ (depending on the supports of the φ_j) such that, if supp $\tilde{\psi}$ (the support of the Fourier transform of ψ) is contained in $\{\tau \in \mathbf{R}: |\tau| > R\}$, then

$$\int_0^\infty d\lambda \psi\left[\frac{1}{2}\left(\lambda - \frac{1}{\lambda}\right)\right](1 + \lambda^{-2}) \times \left\{G_1([\lambda]; t) + G_2\left(\left[\frac{1}{\lambda}\right]; t\right)\right\} = 0,$$

$$\int_0^\infty \frac{d\lambda}{\lambda} \psi\left[\frac{1}{2}\left(\lambda - \frac{1}{\lambda}\right)\right]\left\{G_1([\lambda]; t) - G_2\left(\left[\frac{1}{\lambda}\right]; t\right)\right\} = 0.$$

So that

$$\int_0^\infty d\lambda \psi\left[\frac{1}{2}\left(\lambda - \frac{1}{\lambda}\right)\right](1 + \lambda^{-2}) \times \left\{F_1([\lambda]) + F_2\left(\left[\frac{1}{\lambda}\right]\right)\right\} = 0,$$

$$\int_0^\infty \frac{d\lambda}{\lambda} \psi\left[\frac{1}{2}\left(\lambda - \frac{1}{\lambda}\right)\right]\left\{F_1([\lambda]) - F_2\left(\left[\frac{1}{\lambda}\right]\right)\right\} = 0.$$

Defining the new variable $\xi = \frac{1}{2}[\lambda - (1/\lambda)]$ and

$$F_s(\xi) = \frac{1}{2}\left\{F_1([\lambda]) + F_2\left(\left[\frac{1}{\lambda}\right]\right)\right\}, \quad \lambda = \xi + (\xi^2 + 1)^{\frac{1}{2}},$$

$$F_a(\xi) = \frac{1}{2}\left(\lambda + \frac{1}{\lambda}\right)^{-1}\left\{F_1([\lambda]) - F_2\left(\left[\frac{1}{\lambda}\right]\right)\right\},$$

we find that $F_s(\xi)$ and $F_a(\xi)$ are functions in $\mathcal{S}(\mathbf{R})$ which satisfy

$$\int_0^\infty d\xi \psi(\xi)F_s(\xi)\, d\xi = 0 = \int_0^\infty d\xi \psi(\xi)F_a(\xi)\, d\xi.$$

F_s and F_a are Fourier transforms of functions with support in $\{\tau: |\tau| \le R\}$ and are therefore restrictions (to the real axis) of entire functions. Hence,

$$F_1([\lambda]) = F_s\left[\frac{1}{2}\left(\lambda - \frac{1}{\lambda}\right)\right] + \left(\lambda + \frac{1}{\lambda}\right)F_a\left[\frac{1}{2}\left(\lambda - \frac{1}{\lambda}\right)\right]$$

and

$$F_2([-\lambda]) = F_s\left[\frac{1}{2}\left(-\frac{1}{\lambda} + \lambda\right)\right] + \left(\lambda + \frac{1}{\lambda}\right)F_a\left[\frac{1}{2}\left(-\frac{1}{\lambda} + \lambda\right)\right]$$

are restrictions to $\{\lambda > 0\}$ and to $\{\lambda < 0\}$, respectively, of a function holomorphic in $\{\lambda \in \mathbf{C}, \lambda \neq 0\}$. In other words, continuing $F_1([\lambda])$ holomorphically to a negative value λ_0 of λ, one obtains $F_2([-\lambda_0])$.

On the other hand, applying Lemma 2 [and the fact that the topological space of functions of Λ and p, $\mathcal{C}^\infty$ in Λ and p and holomorphic in Λ coincides with the space of (vector-valued) $\mathcal{C}^\infty$ functions of p with values in the space of holomorphic functions of Λ],

we find that $F_1([\lambda])$ has a holomorphic continuation in $\{\lambda \in \mathbf{C}, \lambda \neq 0\}$ which, for real $\lambda < 0$, is given by

$$\int S_1(p_0, \cdots, p_r; p_{r+1}, \cdots, p_n) \times \prod_{0 \leq j \leq r} f_j(-[\lambda]p_j)\varphi_j(p_j)\delta(p_j^2 - m_j^2)\theta(p_j^0) \times \prod_{r+1 \leq l \leq n} g_l(-[\lambda]p_l)\varphi_l(p_l)\delta(p_l^2 - m_l^2)\theta(p_l^0) \times \delta\left(\sum_{j=0}^{r} p_j - \sum_{l=r+1}^{n} p_l\right) dp_0 \cdots dp_n.$$

This must coincide with $F_2([-\lambda])$. In particular, setting $\lambda = -1$ we obtain

$$\int [S_1(p_0, \cdots, p_r; p_{r+1}, \cdots, p_n) - S_2(p_n, \cdots, p_{r+1}; p_r, \cdots, p_0)]\delta\left(\sum_{j=0}^{r} p_j - \sum_{l=r+1}^{n} p_l\right) \times \prod_{0 \leq j \leq r} f_j(p_j)\varphi_j(p_j)\delta(p_j^2 - m_j^2)\theta(p_j^0) \times \prod_{r+1 \leq l \leq n} g_l(p_l)\varphi_l(p_l)\delta(p_l^2 - m_l^2)\theta(p_l^0)\, dp_0 \cdots dp_n = 0.$$

This being true for any choice of the HA-fields $A_0, \cdots, A_n$, the wavefunctions of the type

$$\prod_j f_j\varphi_j \quad \text{and} \quad \prod_l g_l\varphi_l$$

are sufficiently numerous to ensure the equality of $S_1(p_0, \cdots, p_r; p_{r+1}, \cdots, p_n)$ and $S_2(p_n, \cdots, p_{r+1}; p_r, \cdots, p_0)$, which expresses the CTP invariance of the S-matrix.

5. GENERALIZATION: ARBITRARY SPINS

The arguments of the preceding sections can be generalized to theories describing particles of arbitrary spins. If half odd-integer spins occur, the assumptions of Secs. 1 and 3 must be generalized as follows.

(1) One must postulate the existence of a continuous unitary representation $\{a, M\} \to U(a, M)$ of the *covering group* $\overline{\mathfrak{P}}_+^\uparrow$ of the Poincaré group, the existence of a unique vacuum, and the positivity of energy. We identify the covering group $\overline{L_+(\mathbf{C})}$ of $L_+(\mathbf{C})$ with $SL(2, \mathbf{C}) \times SL(2, \mathbf{C})$ (see Ref. 11) and the covering group $\overline{L_+^\uparrow}$ of $L_+^\uparrow$ with the subgroup of $SL(2, \mathbf{C}) \times SL(2, \mathbf{C})$ formed by the pairs $(A, \bar{A})$ ($\bar{A}$ denoting the complex conjugate of the matrix A). If $M = (A, B) \in \bar{L}_+(\mathbf{C})$ and $x \in \mathbf{C}^4$, we denote Mx the vector of $\mathbf{C}^4$ such that

$$(Mx)^\mu \tau_\mu = A(x^\mu \tau_\mu)B^T$$

$\tau_0 = 1$; τ_1, τ_2, and τ_3 are the Pauli matrices.

(2) $\mathcal{H} = \bigoplus_{j \in G} \mathcal{H}_j$, G being an Abelian group as in Sec. 3.

$$\mathcal{H}_j = \mathcal{H}_j^c \oplus \bigoplus_{k=1}^{N_j} \mathcal{H}_{jk} \quad (j \neq 0),$$

$$\mathcal{H}_0 = \mathbf{C}\Omega \oplus \mathcal{H}_0^c \oplus \bigoplus_{k=1}^{N_0} \mathcal{H}_{0k},$$

N_j may be infinite ($0 \leq j$).

$$(P^\mu P_\mu) \mid \mathcal{H}_j^c \geq M_j^2,$$

$$(P^\mu P_\mu) \mid \mathcal{H}_{jk} = m_{jk}^2;$$

$$0 < m_{j1} < m_{j2} < \cdots < M_j \quad (0 \leq j).$$

(3) The restriction of the representation U of $\mathfrak{P}_+^\uparrow$ to $\mathcal{H}_{jk}$ is unitarily equivalent to the irreducible representation $[m_{jk}, s_{jk}]$. This means that there exists a unitary operator V_{jk} (unique up to a phase factor) mapping $\mathcal{H}_{jk}$ onto the space of square-integrable (vector-valued) functions on the (half) hyperboloid $\{p : (p, p) = m_{jk}^2, p^0 > 0\}$ with values in $\mathbf{C}^{2s_{jk}+1}$. Such a function has $2s_{jk} + 1$ components f_α, where $\alpha = -s_{jk}, -s_{jk} + 1, \cdots, s_{jk} - 1, s_{jk}$; $\overline{\mathfrak{P}}_+^\uparrow$ is represented in this space by

$$(U'(a, M)f)(p) = \hat{\mathfrak{D}}^s(M)f(M^{-1}p),$$

$$U'(a, M) = V_{jk}U(a, M)V_{jk}^{-1},$$

the scalar product being given by

$$(f', f) = \int \sum_{\alpha,\beta} \overline{f'_\alpha(p)}\mathfrak{D}^s_{\alpha\beta}(\tilde{p})f_\beta(p)\, d\Omega_{m_{jk}}(p)$$

with

$$s = s_{jk}.$$

Here, $L \to \mathfrak{D}^s(L)$ is a well-known representation of the full linear group $GL(2, \mathbf{C})$ in $\mathbf{C}^{2s+1}$ satisfying

$$\mathfrak{D}^s(L^*) = \mathfrak{D}^s(L)^*, \quad \mathfrak{D}^s(L^T) = \mathfrak{D}^s(L)^T; \quad M \to \hat{\mathfrak{D}}^s(M)$$

is the mapping of $\overline{L_+(\mathbf{C})}$ defined by $\hat{\mathfrak{D}}^s[(A, B)] = \mathfrak{D}^s(A)$. If $p \in \mathbf{R}^4$ (or $\mathbf{C}^4$), $\tilde{p}$ and p are the 2×2 matrices $\tilde{p} = p^0 1 - \mathbf{p} \cdot \boldsymbol{\tau}$, $p = p^0 1 + \mathbf{p} \cdot \boldsymbol{\tau} = p^\mu \tau_\mu$. If $(p, p)^2 = m^2 > 0$ and $p^0 \gtrless 0$, $m^{-1}\tilde{p}$ and $m^{-1}p$ are two inverse, positive definite matrices with determinant 1.

(4) The existence of algebras $R(B)$ with the same properties as in the preceding sections, except cyclicity of the vacuum, is postulated. Moreover, we assume that, for every domain B of $\mathbf{R}^4$, there exists a set of bounded operators $I(B)$ such that: $I(B)$ is a weakly closed vector subspace of $\mathfrak{L}(\mathcal{H})$; $I(B)^* = I(B)$; if $B_1 \subset B_2$, $I(B_1) \subset I(B_2)$;

$$U(a, M)I(B)U(a, M)^{-1} = I(a + MB).$$

If B_1 and B_2 are spacelike separated and $\psi_1 \in I(B_1)$, $\psi_2 \in I(B_2)$, $A_1 \in R(B_1)$, then

$$\psi_1\psi_2 + \psi_2\psi_1 = 0, \quad \psi_2 A_1 - \psi_1 A_2 = 0;$$

if ψ_1 and $\psi_2 \in I(B)$, then $\psi_1\psi_2 \in R(B)$. For every bounded B we define $I_j^\infty(B)$, $I_{jk}^\infty(B)$ analogously to $R_j^\infty(B)$, $R_{jk}^\infty(B)$, (see Sec. 3). We assume that $I_j^\infty(B) = I_{-j}^\infty(B)^*$ and that

$$[I_j^\infty(B) \cup R_j^\infty(B)]\Omega$$

is a total system of vectors in $\mathcal{H}_j$. We call a fermion (respectively, boson) HA-field an operator-valued function $x \to \psi(x)$ [respectively, $x \to A(x)$] such that

$$\psi(x) = U(x, 1)\psi(0)U(x, 1)^{-1}, \quad \psi(0) \in I_j(B),$$
$$A(x) = U(x, 1)A(0)U(x, 1)^{-1}, \quad A(0) \in I_j(B),$$

for some j and some bounded B.

Borchers' argument again shows that $m_{j,k} = m_{-j,k}$. Let $\psi_1(0) \in I_j^{\infty}(B_1)$, $\psi_2(0) \in I_j^{\infty}(B_2)$, B_1, and B_2 being bounded domains in $\mathbf{R}^4$. Denote

$$f_1 = V_{jk}E_{jk}\psi_1(0)\Omega, \qquad f_2 = V_{jk}E_{jk}\psi_2(0)\Omega,$$
$$\bar{g}_1 = V_{-jk}E_{-jk}\psi_1(0)^*\Omega, \quad \bar{g}_2 = V_{-jk}E_{-jk}\psi_2(0)^*\Omega$$

[where $\bar{g}_1$ means: $\bar{g}_{1\alpha}(p) = \overline{g_{1\alpha}(p)}$; E_{jk} = projector onto $\mathcal{H}_{jk}$].

It is shown in Appendix 2 (Sec. A, 2.2) that, if f_1 (or f_2) $\neq 0$, one must have $s'_{jk} = s_{jk}$ and that $2s_{jk}$ must be odd. Furthermore, with a suitable normalization of V_{-jk}, there exists a function $k \to v_1(k)$ holomorphic on the complex hyperboloid $\{k:(k, k) = m_{jk}^2\}$ with values in $\mathbf{C}^{2s_{jk}+1}$ such that, for p real, $(p, p) = m_{jk}^2$, $p^0 > 0$,

$$v_1(p) = f_1(p),$$
$$v_1(-p) = \mathfrak{D}^s(\underset{\sim}{m_{jk}^{-1}}p)\mathfrak{D}^s(\tau_2)g_1(p) \quad (s = s_{jk}).$$

A similar discussion can be made for boson HA-fields. This property allows us to generalize the argument of Sec. 4; the S-matrix kernels are no longer Lorentz-invariant, but (holomorphically) covariant. This introduces only insignificant changes in the argument, which we omit.

6. CONCLUSION

The proof presented here is provisional in two respects: (1) It only proves the CTP invariance of the S-matrix. A natural conjecture is that, in any theory satisfying the assumptions of Sec. 1, there exists an anti-unitary operator θ such that $\theta^2 = 1$, $\theta\Omega = \Omega$, $\theta U(a, \Lambda)\theta = U(-a, \Lambda)$, and $\theta[R(B)]\theta = R(-B)$. The method of the present paper sheds no light on this problem. (2) The CTP invariance of the S-matrix (with the assumptions made here) would follow trivially if it were shown that the holomorphy domain of the p-space analytic function [known to be schlicht and invariant under $L_+(\mathbf{C})$] contains points on the mass-shell arbitrarily close to all physical points; the presence of such points of analyticity is known at present (for the general n-point function) only near certain physical points, but there is good hope to prove it in general.

Finally, it hardly needs to be remarked that the result is not expected to strengthen the evidence for the CTP invariance of nature. It rather serves to confirm the close relationship of theories of local observables with Wightman field theories and the fact that experimental evidence of violation of the CTP invariance would force the abandonment of the present notions of locality.

ACKNOWLEDGMENTS

The author is very much indebted to Dr. K. Hepp for his patience in explaining to him the asymptotic condition and to Dr. J. J. Loeffel for a very useful suggestion. He wishes to thank Professor A. S. Wightman, Dr. H. J. Borchers, and Dr. A. Jaffe for very valuable discussions. He also wishes to thank Professor J. R. Oppenheimer and the members of the School of Natural Sciences of the Institute for Advanced Study for their kind hospitality.

This research was sponsored by the Air Force Office of Scientific Research, Office of Aerospace Research, United States Air Force, under Contract No. AFOSR 42-65. A Fulbright Travel Grant is also acknowledged.

APPENDIX I

This appendix is devoted to an exercise in analytic completion which is included for the reader's convenience; the results are immediate consequences of the ideas of Refs. 9, 10, 14, 22, and others.

Lemma AII: Let D be a domain in $\mathbf{C}^N$, P a subdomain of D containing the origin O, and Δ a domain in $\mathbf{C}^{N+1}$ of the form

$$\Delta = \mathcal{N} \cap \Omega,$$

where

$$\Omega = \{\lambda, z:\lambda \in \mathbf{C}, \quad \text{Im } \lambda > 0; \ z = (z_1, \cdots, z_N) \in D\},$$

and $\mathcal{N}$ is an open connected neighborhood of the set E

$$E = \{\lambda, z:\text{Im } \lambda = 0, \ \lambda \neq 0, \ z \in D\}$$
$$\cup \{\lambda, z:\text{Im } \lambda > 0, \ \text{Im } z = 0, \ z \in P\}.$$

Then (1) the envelope of holomorphy of Δ is Ω; (2) if K is a compact subset of Ω, there is a compact subset K' of E such that every function defined, and continuous in $\Omega \cup E$, holomorphic in Ω, obeys the inequality

$$\sup_{(\lambda,z)\in K} |f(\lambda, z)| \leq \sup_{(\lambda',z')\in K'} |f(\lambda', z')|.$$

Proof: We can assume without loss of generality that P contains the closure of a domain of the form

$$P_1 = \{z:|z_j| < 1, \quad \text{Im } z_j > 0, \ 1 \leq j \leq N\}.$$

[22] J. Bros and V. Glaser, *Enveloppe d'holomorphie de deux polycercles* (Centre d'Etudes Nucléaires de Saclay, Gif-sur-Yvette, 1961).

We replace the variables λ, z by

$$\mu = \log(\lambda - b)/(\lambda - a) + \log(\lambda + a)/(\lambda + b);$$
$$\zeta_j = \log(1 + z_j)/(1 - z_j); \quad z_j = \tanh \tfrac{1}{2}\zeta_j, \quad \text{(AI1)}$$

where the log is defined as holomorphic in the upper half-plane and equal to $+\infty$ at $+\infty$ on the real axis, and $0 < a < b$, $ab = 1$.

By the Heine Borel lemma, for every θ, $(0 < \theta < \pi)$, there exists $\epsilon > 0$ such that Δ contains the set

$$\Delta(\theta, b, \epsilon)$$
$$= \{\lambda, z: \pi - \epsilon < \operatorname{Im} \mu < \pi; 0 < \operatorname{Im} \zeta_j < \tfrac{1}{2}\pi, \quad (1 \le j \le N)\}$$
$$\cup \{\lambda, z: \pi - \theta < \operatorname{Im} \mu < \pi; 0 < \operatorname{Im} \zeta_j < \epsilon, \quad (1 \le j \le N)\},$$

where μ and ζ_j stand for the functions of λ and z_j defined above. $\Delta(\theta, b, \epsilon)$ is a domain if θ is sufficiently close to π [the condition is $\tan \frac{1}{2}\theta > 2b/(b^2 - 1)$]. Let f be a function holomorphic in the union of these two open sets, which is a *domain* invariant under the transformation $\lambda \to -1/\lambda$. We can always write

$$f(\lambda, z) = f_s(\lambda, z) + [\lambda + (1/\lambda)] f_a(\lambda, z),$$

where

$$f_s(\lambda, z) = \tfrac{1}{2}[f(\lambda, z) + f(-1/\lambda, z)],$$
$$f_a(\lambda, z) = \tfrac{1}{2}[f(\lambda, z) - f(-1/\lambda, z)]$$

are holomorphic in the same domain (and, if f is continuous at E, so are f_a and f_s); moreover,

$$f_{s,a} = \Phi_{s,a}(\mu, \zeta),$$

where μ and ζ_j are the functions defined by (AI1) and the function Φ_s (resp Φ_a) is holomorphic in the tube

$$\{\mu, \zeta: \pi - \epsilon < \operatorname{Im} \mu < \pi; 0 < \operatorname{Im} \zeta_j < \tfrac{1}{2}\pi, \quad (0 \le j \le N)\}$$
$$\cup \{\mu, \zeta: \theta < \operatorname{Im} \mu < \pi; 0 < \operatorname{Im} \zeta_j < \epsilon, \quad (0 \le j \le N)\}.$$

The envelope of holomorphy of this tube (its convex envelope) contains the union, when α varies between 0 and $\frac{1}{2}\pi$, of the domains

$$\Omega'(\alpha, \theta) = \{\mu, \zeta: \pi - (2\alpha\theta/\pi) < \operatorname{Im} \mu < \pi; \quad 0 < \operatorname{Im} \zeta_j < \tfrac{1}{2}\pi - \alpha (1 \le j \le N)\}.$$

It follows that $f(\lambda, z)$ is analytic in the inverse image $\Omega(\alpha, \theta, b)$ of $\Omega'(\alpha, \theta)$. For $b - a > 2 \cot(\alpha\theta/\pi)$, $\Omega(\alpha, \theta, b)$ is connected and has the shape indicated by Fig. 1. It is easy to verify that the union of the domains $\Omega(\alpha, \theta, b)$ is

$$\{\lambda, z: \operatorname{Im} \lambda > 0, z \in P_1\}.$$

Moreover, $\Omega(\alpha, \theta, b)$ is contained in the envelope of

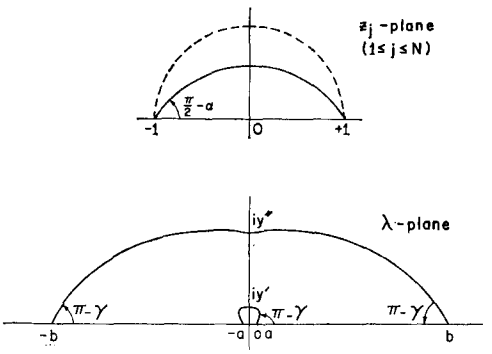


FIG. 1. The domain $\Omega(\alpha, \theta, b)$ is the topological product of the domains represented. $\gamma = 2\alpha\theta/\pi$; $(b - a)^2 \ge 4t^2$; $y' = \{(b - a) - [(b - a)^2 - 4t^2]^{\frac{1}{2}}\}/2t$; $y'' = \{(b - a) + [(b - a)^2 - 4t^2]^{\frac{1}{2}}\}$; $t = \cot(\alpha\theta/\pi)$.

holomorphy of $\Delta(\theta, b, \epsilon)$ so that, when f is continuous at E,

$$\sup_{(\lambda,z)\in\Omega(\alpha,\theta,b)} |f(\lambda, z)| \le \sup_{(\lambda',z')\in\Delta(\theta,b,\epsilon)} |f(\lambda', z')|$$

and, letting ϵ tend to 0, we obtain, when f is continuous at E,

$$\sup_{(\lambda,z)\in\Omega(\alpha,\theta,b)} |f(\lambda, z)| \le \sup_{(\lambda',z')\in E(\theta,b)} |f(\lambda', z')|,$$

where $E(\theta, b)$ is the compact subset of E given by

$$E(\theta, b) = \bigcap_{\epsilon>0} \overline{\Delta(\epsilon, \theta, b)}.$$

The lemma now follows from the following lemma.

Lemma AI2: Let D be a domain in $\mathbf{C}^N$, P a subdomain of D, and Δ a domain in $\mathbf{C}^{N+1}$ of the form

$$\Delta = \mathcal{N} \cap \Omega,$$
$$\Omega = \{\lambda, z, \lambda \in \mathbf{C}, \operatorname{Im} \lambda > 0; z = (z_1, \cdots, z_n) \in D\},$$

and $\mathcal{N}$ is an open connected neighborhood of the set F

$$F = \{\lambda, z, \operatorname{Im} \lambda \ne 0, \lambda = 0, z \in D\} \cup \{\lambda, z; \operatorname{Im} \lambda > 0, z \in P\}.$$

Then, (1) the envelope of holomorphy of Δ is Ω; (2) if K is a compact subset of Ω, there is a compact subset K' of F such that every function f defined and continuous in $\Omega \cup F$, holomorphic in Ω, obeys the inequality,

$$\sup_{z\in K} |f(z)| \le \sup_{z'\in K'} |f(z')|.$$

Proof: (1) We first prove the theorem in a special case,

$$P = \{z: |z_j - i| < 1, 1 \le j \le N\},$$
$$D = \{z: |z_j - \kappa i| < \kappa, 1 \le j \le N\},$$

where $1 < \kappa$. We make the change of variables

$$\zeta_j' = -(1/z_j)(1 \leq j \leq N),$$
$$\mu = \log(\lambda - b)/(\lambda - a) + \log(\lambda + a)/(\lambda + b)$$

and apply the tube theorem as in the first part of the proof of Lemma AI1.

(2) From the special case discussed above, it is easy to deduce that the lemma is true when

$$P = \{z : |z_j - \omega_j| < \kappa r,\ 1 \leq j \leq N\},$$
$$D = \{z : |z_j - \omega_j| < r,\ 1 \leq j \leq N\},$$

where $0 < \kappa < 1$, by taking the union of subdomain in D to which (1) applies after a linear change of coordinates.

(3) *General case:* Let $D_1 = P$, $D_2, \cdots, D_m$ be a finite sequence of polydisks

$$D_\nu = \{z : |z_j - \omega_j^{(\nu)}| < r_\nu\}$$

such that $\omega^{(\nu)} \in D_{\nu-1}$, $r_\nu > 0$ $(2 \leq \nu \leq m)$. Then D_ν contains a polydisk

$$P_\nu = \{z : |z_j - \omega_j^{(\nu)}| < \kappa_\nu r_\nu\} \quad (2 \leq \nu \leq m)$$

contained in $D_{\nu-1}$; applying (2) inductively we obtain the lemma for

$$\bigcup_{\nu=1}^{m} D_\nu$$

and, since every point in D can be reached by a finite chain of polydisks of the above type, the lemma follows. (Note: suppose f is holomorphic in Δ, $\{\Omega_j\}_{j\in I}$ is a family of subdomains of D, and $\{g_j\}_{j\in I}$ is a family of functions such that g_j is holomorphic in $\{\lambda, z : \operatorname{Im}\lambda > 0, z \in \Omega_j\}$ and coincides with f in the intersection of this domain with Δ. Then whenever $\Omega_k \cap \Omega_j \neq \phi$, g_k and g_j coincide in $\{\lambda, z : \operatorname{Im}\lambda > 0, z \in \Omega_k \cap \Omega_j\}$; hence there is a function g holomorphic in

$$\{\lambda, z : \operatorname{Im}\lambda > 0, z \in \bigcup_{j\in I} \Omega_j\}$$

which coincides with g_j in Ω_j for every $j \in I$.)

Lemma AI1 is just what is needed in Sec. 2; we leave the details to the reader.

APPENDIX II: PROOF OF LEMMA 2 AND GENERALIZATION

AII1. Zero-Spin

We use the notations of the end of Sec. 3, replacing m_{jk} by m for simplicity. Let

$$A_1(0) \in R_{jk}^{\infty}(B_1), \quad A_2(0) \in R_{jk}^{\infty}(B_2)$$

(B_1 and B_2 being bounded domains in $\mathbf{R}^4$),

$$f_\alpha = V_{jk}E_{jk}A_\alpha(0)\Omega, \quad g_\alpha^* = V_{-jk}E_{-jk}A_\alpha(0)^*\Omega \quad (\alpha = 1, 2).$$

f_α and (consequently) g_α are supposed not identically 0. We have

$$A_\alpha(0) = \int \varphi_\alpha(a, \Lambda)U(a, \Lambda)A_\alpha' U(a, \Lambda)^{-1}\, da\, d\Lambda, \quad (\alpha = 1, 2),$$

where $A_\alpha' \in R_{jk}(B_\alpha')$ and φ_α is $\mathcal{C}^\infty$ with compact support on $\mathfrak{P}_+^\uparrow$ (Poincaré group); the formulas

$$f_\alpha(p) = \int \tilde{\varphi}_\alpha(p, \Lambda)f_\alpha'(\Lambda^{-1}p)\, d\Lambda \quad (\alpha = 1, 2),$$

$$\Lambda_1 f_\alpha(p) = \int \tilde{\varphi}_\alpha(\Lambda_1^{-1}p, \Lambda)f_\alpha'(\Lambda^{-1}\Lambda_1^{-1}p)\, d\Lambda$$
$$= \int \tilde{\varphi}_\alpha(\Lambda_1^{-1}p, \Lambda_1^{-1}\Lambda)f_\alpha'(\Lambda^{-1}p)\, d\Lambda \quad (\alpha = 1, 2),$$

show that, for real Λ_1 and p, $\Lambda_1 f_\alpha(p)$ is a $\mathcal{C}^\infty$ function of these variables and for fixed Λ_1, is in $\mathcal{S}$ as a function on the hyperboloid. On the other hand, the weak continuity of $U(a, \Lambda)$ implies that

$$(\Omega, A_2'^* U(a, \Lambda_1)A_1'\Omega) \quad \text{and} \quad (\Omega, A_1' U(a, \Lambda_1)^{-1}A_2'^*\Omega)$$

are continuous functions of Λ_1 with values in the tempered distributions (in fact, continuous bounded functions) in a. By performing estimates similar to those of Sec. 2 (or by using the Jost–Lehmann–Dyson representation), it is possible to show that the p-space analytic function $H_{\Lambda_1}'(k)$, holomorphic in $\{k : (k, k) \notin m_{j,k+1}^2 + \mathbf{R}^+\}$, which reduces in the tubes $\mathcal{C}^\pm$ to the Laplace transforms of

$$(-\Box - m_{jk}^2)\theta(-a^0)(\Omega, [A_2'^*, U(a, \Lambda_1)A_1 U(a, \Lambda_1)^{-1}]\Omega)$$

and

$$(+\Box + m_{jk}^2)\theta(a^0)(\Omega, [A_2'^*, U(a, \Lambda)A_1 U(a, \Lambda_1)^{-1}]\Omega)$$

depends continuously on Λ_1 (it is a continuous function of k and Λ_1 in every compact). Therefore,

$$f_2'^*(p)f_1'(\Lambda_1^{-1}p)$$

is the restriction (to real p) of a function of Λ_1 and p defined and continuous for real Λ, and complex p (on the hyperboloid), and holomorphic in p; more precisely, for every real p ($p^2 = m^2$, $p^0 > 0$), there is a continuous function of $\Lambda_1 \in L_+^\uparrow$ and $\Lambda \in L_+(\mathbf{C})$, holomorphic in Λ, which reduces to

$$f_2'^*(\Lambda^{-1}p)f_1'(\Lambda_1^{-1}\Lambda^{-1}p).$$

As a consequence,

$$f_2^*(\Lambda^{-1}p)f_1(\Lambda_1^{-1}\Lambda^{-1}p)$$

extends to a $\mathcal{C}^\infty$ function of $(\Lambda_1, \Lambda) \in L_+^\uparrow \times L_+(\mathbf{C})$, holomorphic in Λ. It follows that, if Λ_0, $\Lambda_1 \in L_+^\uparrow$, $\zeta \in \mathbf{R}$, p is real, the function

$$\Phi_{21}(\Lambda_0; \Lambda_1; \zeta; p) = f_2^*(\Lambda_0^{-1}[e^{-\zeta}]p)f_1(\Lambda_1^{-1}[e^{-\zeta}]p)$$

can be extended to a function defined for complex values of ζ, holomorphic in ζ and $\mathcal{C}^\infty$ in all variables simultaneously. The same holds if we replace f_1 by f_2 or $\Lambda_1 f_1$ by

$$(d/d\zeta)[e^\zeta]\Lambda_1 f_1\big|_{\zeta=0},$$

which shows the existence of functions

$$\Phi_{22}(\Lambda_0;\Lambda_0;\zeta;p),\quad \Phi'_{21}(\Lambda_0;\Lambda_1;\zeta;p).$$

$\mathcal{C}^\infty$ in all variables, holomorphic in ζ, which, for real ζ, reduce to

$$f_2^*(\Lambda_0^{-1}[e^{-\zeta}]p)f_2(\Lambda_0^{-1}[e^{-\zeta}]p)$$

and

$$f_2^*(\Lambda_0^{-1}[e^{-\zeta}]p)f_1'(\Lambda_1^{-1}[e^{-\zeta}]p),$$

respectively; $f_1'(\Lambda_1^{-1}[e^{-\zeta}]p)$ stands for

$$(d/d\zeta)f_1(\Lambda_1^{-1}[e^{-\zeta}]p).$$

Define

$$\Psi_{1/2}(\Lambda_0;\Lambda_1;\zeta;p)=\frac{\Phi_{21}(\Lambda_0;\Lambda_1;\zeta;p)}{\Phi_{22}(\Lambda_0;\Lambda_0;\zeta;p)},$$

$$\Psi_{1/1}(\Lambda_1;\zeta;p)=\frac{\Phi'_{21}(\Lambda_0;\Lambda_1;\zeta;p)}{\Phi_{22}(\Lambda_0;\Lambda_1;\zeta;p)},$$

$$\Psi_{2/2}(\Lambda_0;\zeta;p)=\frac{\Phi'_{22}(\Lambda_0;\Lambda_0;\zeta;p)}{\Phi_{22}(\Lambda_0;\Lambda_0;\zeta;p)},$$

Φ'_{22} being defined analogously to Φ'_{21}. $\Psi_{1/2}$, $\Psi_{1/1}$, $\Psi_{2/2}$ are meromorphic functions of ζ depending differentiably on Λ_1, Λ_0, p; they reduce, for real ζ, to

$$\frac{f_1(\Lambda_1^{-1}[e^{-\zeta}]p)}{f_2(\Lambda_0^{-1}[e^{-\zeta}]p)},\quad \frac{f_1'(\Lambda_1^{-1}[e^{-\zeta}]p)}{f_1(\Lambda_1^{-1}[e^{-\zeta}]p)},\quad \frac{f_2'(\Lambda_0^{-1}[e^{-\zeta}]p)}{f_2(\Lambda_0^{-1}[e^{-\zeta}]p)}.$$

This implies that

$$\frac{(\partial/\partial\zeta)\Psi_{1/2}}{\Psi_{1/2}}=\Psi_{1/1}-\Psi_{2/2}.$$

As the logarithmic derivative of a meromorphic function, this expression is a meromorphic function of ζ having only simple poles. These occur (1) at the zeros of $\Phi_{21}(\Lambda_0;\Lambda_1;\zeta;p)$: the corresponding residue is then a positive integer; (2) at the zeros of $\Phi_{22}(\Lambda_0;\Lambda_1;\zeta;p)$: the corresponding residue is then a negative integer.

Let μ be a singular point of $\Psi_{1/1}(\Lambda_1;\zeta;p)$. We can adjust Λ_0 so that $\Phi_{22}(\Lambda_0;\Lambda_0;\zeta;p)$ and $\Psi_{2/2}(\Lambda_0;\Lambda_1;\zeta;p)$ have no zeros or singularities for $\zeta=\mu$. Thus μ must be a zero of $\Psi_{1/2}(\Lambda_0;\Lambda_1;\zeta;p)$, whose logarithmic derivative has a simple pole with positive integral residue at μ. Since $\Psi_{2/2}$ is regular at μ, $\Psi_{1/1}$ must have a simple pole with positive integral residue at μ. Let us define

$$\Psi_1(\Lambda_1;\zeta;p)=\exp\left[\int_{\zeta_0}^{\zeta}\Psi_{1/1}(\Lambda_1;\zeta';p)\,d\zeta'\right]\times f_1(\Lambda_1^{-1}[e^{-\zeta_0}]p)$$

the integral being taken along a smooth path skirting the singularities of the integrand; $\zeta_0\in\mathbf{R}$. We obtain a holomorphic function of ζ which, except possibly at its zeros, depends differentiably on Λ_1 and p; the same is true for

$$\Psi_2(\Lambda_0;\zeta;p)=\left[\exp\int_{\zeta_0}^{\zeta}\Psi_{2/2}(\Lambda_0;\zeta';p)\,d\zeta'\right]\times f_2(\Lambda_0^{-1}[e^{-\zeta_0}]p)$$

[ζ_0 being real and not a zero of $f_2(\Lambda_0^{-1}[e^{-\zeta}]p)$]. We have

$$(d/d\zeta)\log\Psi_1(\Lambda_1;\zeta;p)=(d/d\zeta)\log f_1(\Lambda_1^{-1}[e^{-\zeta}]p)\quad\text{for real }\zeta,$$

$\Psi_1(\Lambda_1;\zeta;p)$ and $f_1'(\Lambda_1^{-1}[e^{-\zeta}]p)$ are (for real ζ) $\mathcal{C}^\infty$ functions which are proportional to each other in any interval separating their (common) zeros; these zeros always have finite order; the two functions coincide near ζ_0. Therefore, they must coincide for all real ζ. The same holds for Ψ_2 and f_2. If $\mu\in\mathbf{C}$ we can adjust Λ_0 so that $\Psi_2(\Lambda_0;\zeta;p)$ has no zero in a neighborhood of μ; this implies that Ψ_2 depends differentiably on ζ and p near (μ,p), and $\Psi_{1/2}(\Lambda_0;\Lambda_1';\zeta;p')$ is $\mathcal{C}^\infty$ in Λ_1', ζ, p' near $\zeta=\mu$, $\Lambda_1'=\Lambda_1$, $p'=p$. We also note that, for real ξ and complex ζ

$$\Psi_1(\Lambda_1;\zeta+\xi;p)=\Psi_1([e^\xi]\Lambda_1;\zeta;p),$$

a consequence of the corresponding identity for real ζ.

We call "timelike subgroup" of $L_+(\mathbf{C})$ a mapping of $\mathbf{C}$ into $L_+(\mathbf{C})$ of the form

$$\chi(\zeta)=\Lambda_0[e^\zeta]\Lambda_0^{-1},$$

where $\Lambda_0\in L_+^\uparrow$ is fixed. Let $\chi_1,\cdots,\chi_N$ be timelike subgroups, p be real, $(p,p)=m^2$, $p^0>0$. The expression

$$f_1[\Lambda_1^{-1}\chi_N(-\zeta_N)\cdots\chi_1(-\zeta_1)p]$$

is a $\mathcal{C}^\infty$ function of Λ_1, $\zeta_1,\cdots,\zeta_N$, p for real Λ_1, $\zeta_1,\cdots,\zeta_N,p$. Furthermore, it is the restriction of the function

$$\Psi_1[\chi_{k+1}(\zeta_{k+1})\cdots\chi_N(\zeta_N)\times\Lambda_1;\zeta_k;\chi_{k-1}(-\zeta_{k-1})\cdots\chi_1(-\zeta_1)p]$$

defined, $\mathcal{C}^\infty$ in $\Lambda_1,\zeta_1,\cdots,\zeta_N,p$, and holomorphic in ζ_k for ζ_k complex, $\Lambda_1,\zeta_1,\cdots,\zeta_{k-1},\zeta_{k+1},\cdots,\zeta_N,p$, real. By the Malgrange–Zerner theorem,[9,23,24] there exists a function

$$F_{\chi_1,\cdots,\chi_N}(\Lambda_1;\zeta;p)\quad(\text{where }\zeta=\{\zeta_1,\cdots,\zeta_N\})$$

defined and $\mathcal{C}^\infty$ for real Λ_1, p, complex ζ, and holomorphic in ζ. The argument now follows that of

[23] B. Malgrange, unpublished (1961).
[24] M. Zerner, Seminar Notes, Marseilles (1961).

Ref. 10 (Sec. II.3) and is only sketched here. The functions of the type

$$F_{\chi_1,\cdots,\chi_N}(\Lambda_1;\zeta;p)$$

are locally functions of $\chi_1(\zeta_1)\cdots\chi_N(\zeta_N)\Lambda_1$, and p. For fixed p, these functions yield a set $\mathfrak{G}_p$ of germs of holomorphic functions on the complex Lorentz group. $\mathfrak{G}_p$ is a connected open subset of the sheaf of germs of analytic functions on $L_+(\mathbf{C})$. Furthermore, for every continuous curve $t\to\gamma(t)\{[0,1]\to L_+(\mathbf{C})\}$ in $L_+(\mathbf{C})$, and every $\tilde{\gamma}_0\in\mathfrak{G}_p$ projecting onto $\gamma(0)$, there exists a continuous map $t\to\tilde{\gamma}(t)$ of $[0,1]$ into $\mathfrak{G}_p$ such that $\tilde{\gamma}(0)=\tilde{\gamma}_0$ and, for each t, $\tilde{\gamma}(t)$ projects onto $\gamma(t)$. By the monodromy theorem, $\mathfrak{G}_p$ is the set of germs yielded by a function $F(A;p)$ defined and holomorphic on the covering group $\overline{L_+(\mathbf{C})}$ of $L_+(\mathbf{C})$, such that, for $A\in\overline{L_+^\uparrow}\subset\overline{L_+(\mathbf{C})}$, $F(A;p)=f_1(\Lambda(A)^{-1}p)$. Here, $A\to\Lambda(A)$ denotes the canonical homomorphism of $\overline{L_+(\mathbf{C})}$ onto $L_+(\mathbf{C})$. Let N be the kernel of this homomorphism; for $A_0\in N$ and $A\in\overline{L_+^\uparrow}$ we have $F(A;p)=F(AA_0;p)$. This identity extends to all complex $A\in\overline{L_+(\mathbf{C})}$ by analytic continuation so that $F(A;p)=F_1[\Lambda(A);p]$, where $F_1(\Lambda;p)$ is a holomorphic function of $\Lambda\in L_+(\mathbf{C})$. It is easy to verify that F_1 is $\mathcal{C}^\infty$ in both Λ and p. $F_1(\Lambda;p)=f_1(\Lambda^{-1}p)$ for real Λ. It is only necessary to verify that, for $-\Lambda\in L_+^\downarrow$, $F_1(-\Lambda;p)$ coincides with a multiple of $g_1(\Lambda^{-1}p)$. For this purpose we construct analogously a function $G_1(\Lambda;p)$ holomorphic for $\Lambda\in L_+(\mathbf{C})$ and coinciding with $g_1(-\Lambda^{-1}p)$ for $\Lambda\in L_+^\downarrow$. It is easy to verify that each germ of $G_1(\Lambda;p)$ is proportional to the germ of $F_1(\Lambda;p)$ at the same point. Hence $G_1(\Lambda;p)=\omega_1F_1(\Lambda;p)$. A similar constant ω_2 can be defined for f_2 and g_2. Because $\Lambda\to g_1(-\Lambda^{-1}p)/g_2(-\Lambda^{-1}p)$ and $\Lambda\to f_1(\Lambda^{-1}p)/f_2(\Lambda^{-1}p)$ are restrictions (to $L_+^\downarrow$ and $L_+^\uparrow$) of the same meromorphic function, we must have $\omega_1=\omega_2=\omega$. Moreover, $\bar{\omega}\omega=1$. Redefining V_{-jk} by replacing it by $\omega^{-1}V_{-jk}$ we obtain the result of Lemma 2.

AII2. Generalization (Case of Arbitrary Spin)

We use the notation of Sec. 5, replacing m_{jk} by m, s_{jk} by s and s_{-jk} by s' for simplicity. Let $h_1,\cdots,h_{2s+1}$, $f_1,\cdots,f_{2s+1}$ be (vector-valued) functions of the form

$$h_l=V_{jk}E_{jk}A_l'(0)\Omega,\quad [\text{resp } V_{jk}E_{jk}\psi_l'(0)\Omega]\qquad(1\le l\le 2s+1),$$

$$f_l=V_{jk}E_{jk}A_l(0)\Omega,\quad [\text{resp } V_{jk}E_{jk}\psi_l(0)\Omega]\qquad(1\le l\le 2s+1),$$

where $A_l(0)\in R_{jk}^\infty(B_l)$, $A_l'(0)\in R_{jk}^\infty(B_l')$, $\psi_l(0)\in I_{jk}^\infty(B_l)$, $\psi_l'(0)\in I_{jk}^\infty(B_l')$, B_l and B_l' being bounded domains $(1\le l\le 2s+1)$. For each $l(1\le l\le 2s+1)$, h_l and f_l are functions on $\{p:(pp)=m^2,\ p^0>0\}$ with values in $\mathbf{C}^{2s+1}$; their components will be denoted:

$$h_{\alpha l},\ f_{\alpha l},\quad\text{where}\quad \alpha=-s,\ -s+1,\cdots,s-1,\ s.$$

Let H and F denote the matrix-valued functions[25] (on the hyperboloid) with matrix-elements $(H)_{\alpha l}(p)=h_{\alpha l}(p)$, $(F)_{\alpha l}(p)=f_{\alpha l}(p)$. We choose H and F so that $\det H\not\equiv 0$, $\det F\not\equiv 0$. There exists a matrix-valued function Φ_{HG} of M_0, M_1, ζ, p, defined and $\mathcal{C}^\infty$ in the variables M_0, $M_1\in\overline{L_+^\uparrow}$, $\zeta\in\mathbf{C}$, $p\in\{p':(p',p')=m^2,\ p'^0>0\}$ which coincides, for real ζ, with

$$\{\hat{\mathfrak{D}}^s(e^{\zeta J^{01}}M_0)H(M_0^{-1}[e^{-\zeta}]p)\}^*\mathfrak{D}^s(\tilde{p})\times\{\hat{\mathfrak{D}}^s(e^{\zeta J^{01}}M_1)F(M_1^{-1}[e^{-\zeta}]p)\}$$

(the proof is the same as in the spinless case; Sec. AII1). Here, $\overline{L_+^\uparrow}$ has been identified with the subgroup of $SL(2,\mathbf{C})\times SL(2,\mathbf{C})=\overline{L_+(\mathbf{C})}$ consisting of pairs of the form $(A,\bar{A})$. This subgroup is isomorphic to $SL(2,\mathbf{C})$. The function $M\to\hat{\mathfrak{D}}^s(M)$ is defined on $\overline{L_+(\mathbf{C})}$ by

$$\hat{\mathfrak{D}}^s[(A,B)]=\mathfrak{D}^s(A)$$

and defines a holomorphic representation of $\overline{L_+(\mathbf{C})}$. If $M\in\overline{L_+(\mathbf{C})}$, $M=(A,B)$, and $p\in\mathbf{R}^4$ (or $p\in\mathbf{C}^4$), Mp is the vector such that

$$(Mp)^\mu\tau_\mu=A(p^\mu\tau_\mu)B^T,$$
$$(\tau_0=\mathbf{1},\tau_1,\tau_2,\tau_3\text{ Pauli matrices}),$$

J^{01} is the element of the Lie algebra of $\overline{L_+(\mathbf{C})}$ given by $(\frac{1}{2}\tau_1,\frac{1}{2}\tau_1)$; one has: $e^{\zeta J^{01}}p=[e^\zeta]p$. The above properties of $\Phi_{H,F}(\Lambda_0;\Lambda_1;\zeta;p)$ are obtained just as in the scalar case. Similarly, there exists a function $\Phi_{H\dot{F}}$, of the same variables, with the same properties, which coincides, for real ζ, with

$$\{\hat{\mathfrak{D}}^s(e^{\zeta J^{01}}M_0)H(M_0^{-1}[e^{-\zeta}]p)\}^*\mathfrak{D}^s(\tilde{p})(d/d\zeta)\times\{\hat{\mathfrak{D}}^s(e^{\zeta J^{01}}M_1)F(M_1^{-1}[e^{-\zeta}]p)\}$$

and functions, meromorphic in ζ, defined by

$$\Psi_{H^{-1}F}(M_0;M_1;\zeta;p)=[\Phi_{HH}(M_0;M_0;\zeta;p)]^{-1}\Phi_{HF}(M_0;M_1;\zeta;p),$$

$$\Psi_{F^{-1}\dot{F}}(M_1;\zeta;p)=[\Phi_{HF}(M_0;M_1;\zeta;p)]^{-1}\Phi_{H\dot{F}}(M_0;M_1\zeta;p),$$

$$\Psi_{H^{-1}\dot{H}}(M_0;\zeta;p)=[\Phi_{HH}(M_0;M_0;\zeta;p)]^{-1}\Phi_{H\dot{H}}(M_0;M_0\zeta;p),$$

note that one also has

$$\Phi_{F^{-1}\dot{F}}(M_1;\zeta;p)=[\Phi_{FF}(M_1;M_1;\zeta;p)]^{-1}\Phi_{F\dot{F}}(M_1;M_1;\zeta;p)$$

[25] The idea of using these matrices was kindly suggested to the author by Dr. J. J. Loeffel.

since the two expressions given for this meromorphic function of ζ coincide, for real ζ, with

$$\{\hat{\mathfrak{D}}^s(e^{\zeta J^{01}}M_1)F(M_1^{-1}[e^{-\zeta}]p)\}^{-1}(d/d\zeta) \times \{\hat{\mathfrak{D}}^s(e^{\zeta J^{01}}M_1)F(M_1^{-1}[e^{-\zeta}]p)\}.$$

D being a positive real number, we restrict M_0, M_1, p to take their values in small open subsets of $\overline{L_+^\uparrow}$ and of the hyperboloid chosen so that the functions $[\Phi_{HH}(M_0; M_0; \zeta; p)]^{-1}$ and $\Psi_{F^{-1}\dot{F}}(M_1; \zeta; p)$ have no common pole in $\{\zeta:|\zeta| < D\}$. Let $[\xi_0 - \tau, \xi_0 + \tau]$ be an interval of the real axis contained in $\{\zeta:|\zeta| < D\}$, where $\Psi_{F^{-1}\dot{F}}(M_1; \zeta; p)$ and $\Psi_{H^{-1}\dot{H}}(M_0; \zeta; p)$ are regular. The linear differential equations

$$\frac{d}{d\zeta}X(\zeta) = X(\zeta)\Psi_{F^{-1}\dot{F}}(M_1; \zeta; p),$$

$$\frac{d}{d\zeta}Y(\zeta) = Y(\zeta)\Psi_{H^{-1}\dot{H}}(M_0; \zeta; p)$$

[where the unknowns $X(\zeta)$ and $Y(\zeta)$ are matrix-valued functions] possess in suitable complex neighborhoods of the real segment $[\xi_0 - \tau, \xi_0 + \tau]$, unique solutions, denoted $X(M_1; \zeta; p)$ and $Y(M_0; \zeta; p)$ coinciding with

$$\hat{\mathfrak{D}}^s(e^{\zeta J^{01}}M_1)F(M_1^{-1}[e^{-\zeta}]p),$$
$$\hat{\mathfrak{D}}^s(e^{\zeta J^{01}}M_0)H(M_0^{-1}[e^{-\zeta}]p),$$

respectively, for real $\zeta \in [\xi_0 - \tau, \xi_0 + \tau]$. These solutions are holomorphic in ζ and $\mathcal{C}^\infty$ in M_1, ζ, p or M_0, ζ, p, respectively. As it is well known from the theory of linear differential equations (see Refs. 26 and 27) X (respectively, Y) can be analytically continued along any path, in the complex ζ plane, which does not pass through a pole of $\Psi_{F^{-1}\dot{F}}$ (respectively, $\Psi_{H^{-1}\dot{H}}$). Let $\mathcal{O}$ be as imply connected domain of the complex plane containing ξ_0 (e.g., a neighborhood of a path), containing no pole of $[\Phi_{HH}(M_0; M_0; \zeta; p)]^{-1}$ but possibly containing poles of $\Psi_{F^{-1}\dot{F}}$. In this domain $Y(M_0; \zeta; p)$ admits of a single-valued analytic continuation which we continue to denote by the same symbol. Since $y = \det Y$ satisfies the equation

$$y^{-1}(\zeta)(dy/d\zeta)(\zeta) = \operatorname{tr}\Psi_{H^{-1}\dot{H}}(\zeta)$$

in $\mathcal{O}$ (see Ref. 27), Y^{-1} is regular in $\mathcal{O}$. On the other hand, $\Psi_{H^{-1}F}(M_0; M_1; \zeta; p)$ is also regular in $\mathcal{O}$ and coincides near ξ_0 with $Y^{-1}X$ (since this equality holds for $\zeta \in [\xi_0 - \tau, \xi_0 + \tau]$) so that $X = Y\Psi_{H^{-1}F}$ has a single-valued continuation in $\mathcal{O}$. Using the monodromy theorem, one finds that X has a single-valued continuation in the disk $\{\zeta:|\zeta| < D\}$.

On the real axis

$$Y^{-1}X = \Psi_{H^{-1}F} = \{\hat{\mathfrak{D}}^s(e^{\zeta J^{01}}M_0)H(M_0^{-1}[e^{-\zeta}]p)\}^{-1} \times \{\hat{\mathfrak{D}}^s(e^{\zeta J^{01}}M_1)F(M_1^{-1}[e^{-\zeta}]p)\};$$

because Y and X are not singular at the same points in the disk $\{\zeta:|\zeta| < D\}$, we have

$$X = \hat{\mathfrak{D}}^s(e^{\zeta J^{01}}M_1)F(M_1^{-1}[e^{-\zeta}]p),$$
$$Y = \hat{\mathfrak{D}}^s(e^{\zeta J^{01}}M_0) \times H(M_0^{-1}[e^{-\zeta}]p)$$

for all real ζ such that $|\zeta| < D$. For given values of M_1, p, D can be chosen arbitrarily, and X is therefore the restriction of an entire function in ζ, $\mathcal{C}^\infty$ in M_1, ζ, p. The same holds for Y. Continuing the argument along the line of AII1, we find that there exist matrices $X(M; p)$, $Y(M; p)$ defined and $\mathcal{C}^\infty$ for $M \in \overline{L_+(\mathbf{C})}$, $(p, p) = m^2$, $p^0 > 0$, holomorphic in M, which, for $M \in \overline{L_+^\uparrow}$, coincide with $\hat{\mathfrak{D}}_s(M)F(M^{-1}p)$ and $\hat{\mathfrak{D}}^s(M)H(M^{-1}p)$, respectively.

Setting $X_0(M; p) = \hat{\mathfrak{D}}^s(M)^{-1}X(M; p)$, $Y_0(M, p) = \hat{\mathfrak{D}}^s(M)^{-1}Y(M; p)$, we see that $X_0[(A_0M); p] = X_0(M; p)$ and $Y_0[(A_0M); p] = Y_0(M; p)$, i.e., that $X_0(M; p)$ and $Y_0(M; p)$ are actually (for fixed p) functions on $L_+(\mathbf{C})$ [here $A_0 = (-1, -1)$].

We have proved that the vector function f_l and h_l (the columns of the matrices F and H) are restrictions of (vector-valued) functions, holomorphic over the whole complex hyperboloid $\{k:(k, k) = m^2\}$.

Denote now

$$\tilde{g}_l' = V_{-jk}E_{-jk}A_l(0)^*\Omega \text{ [resp } V_{-jk}E_{-jk}\psi_l(0)^*\Omega],$$
$$\tilde{e}_l' = V_{-jk}E_{-jk}A_l(0)'^*\Omega \text{ [resp } V_{-jk}E_{-jk}\psi_l(0)'^*\Omega],$$

and g_l, e_l the vector-valued functions defined by

$$g_l(p) = \{g_{\alpha l}(p)\} = \mathfrak{D}^{s'}(m^{-1}\underset{\sim}{p})\mathfrak{D}^{s'}(\tau_2)g_l'(p),$$
$$e_l(p) = \mathfrak{D}^{s'}(m^{-1}\underset{\sim}{p})\mathfrak{D}^{s'}(\tau_2)e_l'(p).$$

We can prove in a similar way that they are also restrictions of functions holomorphic on the whole complex hyperboloid. Forming the (possibly rectangular) matrices $G(p) = \{g_{\alpha l}(p)\}$ [whose columns are the vectors $g_l(p)$] $E(p) = \{e_{\alpha l}(p)\}$ we find that

$$\epsilon E^*(-p)\mathfrak{D}^{s'}(-\tilde{p})G(-p) \quad \text{and} \quad H(p)^*\mathfrak{D}^s(\tilde{p})F(p)$$

[where $\epsilon = +1$ (respectively, -1) if we deal with boson (respectively, fermion) HA-fields] are restrictions to the two real sheets of the hyperboloid $\{k:(k, k) = m^2\}$ of a matrix-valued function holomorphic on the whole complex hyperboloid. Since the determinant of the second matrix is not identically 0, the same holds for the determinant of the first. This implies $s' \geq s$; by symmetry, $s' = s$.

[26] E. Goursat, *Cours d'analyse mathematique* (Gauthier-Villars, Paris, 1929) Vol. II.

[27] S. Lefschetz, *Differential Equations: Geometric Theory* (Interscience Publishers, Inc., New York, 1963) 2nd ed., pp. 55 ff.

Let $(M, p) \to X(M; p)$, $Y(M; p)$, $Z(M; p)$, $T(M; p)$ denote the functions of $M \in \overline{L_+(\mathbf{C})}$ and $p \in \{(p', p') = m^2, p'^0 > 0\}$, $\mathcal{C}^\infty$ in these variables and holomorphic in M such that

$$X(M; p) = \hat{\mathfrak{D}}^s(M)F(M^{-1}p),$$
$$Y(M; p) = \hat{\mathfrak{D}}^s(M)H(M^{-1}p) \quad \text{for} \quad M \in \overline{L_+^\uparrow},$$
$$Z(M; p) = \hat{\mathfrak{D}}^s(M)G(-M^{-1}p),$$
$$T(M; p) = \hat{\mathfrak{D}}^s(M)E(-M^{-1}p) \quad \text{for} \quad M \in \overline{L_+^\downarrow},$$

and

$$X_0(M; p) = \hat{\mathfrak{D}}^s(M)^{-1}X(M; p),$$
$$Y_0(M; p) = \hat{\mathfrak{D}}^s(M)^{-1}Y(M; p),$$
$$Z_0(M; p) = \hat{\mathfrak{D}}^s(M)^{-1}Z(M; p),$$
$$T_0(M; p) = \hat{\mathfrak{D}}^s(M)^{-1}T(M; p).$$

For $M \in \overline{L_+^\uparrow}$ we know that

$$\epsilon E(-M^{-1}p)^*\mathfrak{D}^s((-M^{-1}p)^\sim)G(-M^{-1}p)$$

is the value obtained by continuing analytically the function of $M' \in \overline{L_+^\uparrow}$ given by $H(M'^{-1}p)^*\mathfrak{D}^s((M'^{-1}p)^\sim)$ $F(M'^{-1}p)$ to $M' = M$, i.e., $Y_0(\check{M}; p)^*\mathfrak{D}(^s(M^{-1}p)^\sim)$ $X_0(M; p)$, where if $M = (A, B)$, $\check{M} = (\bar{B}, \bar{A})$. Taking $M = (A, -\bar{A})$ and recalling that $Y_0(A_0M; p) = Y_0(M; p)$ for $A_0 = (-1, -1)$, we obtain

$$\epsilon Y_0(\check{M}; p)^*\mathfrak{D}^s((M^{-1}p)^\sim)X_0(M; p) = E(-M^{-1}p)^*\mathfrak{D}^s((-M^{-1}p)^\sim)G(-M^{-1}p),$$

$(M \in \overline{L_+^\uparrow})$; in particular (for $M \in \overline{L_+^\uparrow}$),

$$\epsilon X_0(M; p)^*\mathfrak{D}^s((M^{-1}p)^\sim)X_0(M; p) = G(-M^{-1}p)^*\mathfrak{D}^s((-M^{-1}p)^\sim)G(-M^{-1}p).$$

The right-hand side is a nonsingular positive definite matrix for almost all M; the left-hand side can have the same property only if $\epsilon = (-1)^{2s}$. We thus verify the well-known connection between spin and statistics.[5,12,16,28,29] It follows that, for $M = (A, -\bar{A}) \in \overline{L_+^\uparrow}$

$$Y_0(\check{M}; p)^*\mathfrak{D}^s(A)^*\mathfrak{D}^s(\tilde{p})\mathfrak{D}^s(A)X_0(M; p) = T_0(\check{M}; p)^*\mathfrak{D}^s(A)^*\mathfrak{D}^s(\tilde{p})\mathfrak{D}^s(A)\,Z_0(M; p)$$

or, since $\mathfrak{D}^s(A) = \hat{\mathfrak{D}}_s(M)$,

$$Y(\check{M}; p)^*\mathfrak{D}^s(\tilde{p})X(M; p) = T(\check{M}; p)^*\mathfrak{D}^s(\tilde{p})\,Z(M; p).$$

This identity, valid for $M \in \overline{L_+^\uparrow}$ extends to all $M \in \overline{L_+(\mathbf{C})}$ by analytic continuation. A similar identity holds if we replace $X(M; p)$ by $X(MM_0; p)$ with $M_0 \in \overline{L_+^\uparrow}$. Hence,

$$X(MM_0; p)^{-1}X(M; p) = Z(MM_0; p)^{-1}\,Z(M; p)$$

holds as an identity between meromorphic functions of M for $M_0 \in \overline{L_+^\uparrow}$ and extends, by analytic continuation, to an identity between meromorphic functions of M and M_0 in $\overline{L_+(\mathbf{C})} \times \overline{L_+(\mathbf{C})}$. Let M be a point such that $M_0 \to X(MM_0; p)^{-1}$ and $M_0 \to Z(MM_0; p)^{-1}$ are holomorphic in a neighborhood V of the identity of $\overline{L_+(\mathbf{C})}$. Let ω_p denote the invertible matrix

$$\omega_p = Z(M; p)X(M; p)^{-1},$$

then the identity

$$Z(MM_0; p) = \omega_p X(MM_0; p)$$

holds for $M_0 \in V$, hence for all $M_0 \in \overline{L_+(\mathbf{C})}$. The identity

$$X(M; M_1p) = \hat{\mathfrak{D}}^s(M_1)X(M_1^{-1}M; p), \quad M_1 \in \overline{L_+^\uparrow},$$

and the corresponding identity for Z yield

$$\omega_{M_1p} = \hat{\mathfrak{D}}^s(M_1)\omega_p\hat{\mathfrak{D}}^s(M_1)^{-1}, \quad M_1 \in \overline{L_+^\uparrow}.$$

Taking $p = P = (m, 0, 0, 0)$, $M_1 = (U, \bar{U})$, $U \in SU_2$, we find

$$\omega_P = \mathfrak{D}^s(U)\omega_P\mathfrak{D}^s(U)^{-1},$$

so that ω_P must be a multiple of the unit matrix; the preceding equation gives $\omega_p = \omega_P = \omega$ for all p. Moreover,

$$T(M; p) = \omega\, Y(M; p)$$

and $\omega^*\omega = \bar{\omega}\omega = 1$, i.e., ω is a phase factor. Changing V_{-jk} into $\omega^{-1}V_{-jk}$ we obtain the desired result, i.e., that, for each l and α, $f_{\alpha l}(p)$ and $g_{\alpha l}(-p)$ are the restrictions to $\{p:(p, p) = m^2, p^0 > 0\}$ and to $\{p:(p, p) = m^2, p^0 < 0\}$, respectively, of the same function defined and holomorphic on the whole complex hyperboloid $\{k:(k, k) = m^2\}$.

[28] N. Burgoyne, Nuovo Cimento **8**, 807 (1958).
[29] G. Lüders and B. Zumino, Phys. Rev. **110**, 1450 (1958).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Lie Algebra Extensions of the Poincaré Algebra

A. GALINDO
Facultad de Ciencias, Universidad de Zaragoza, Zaragoza, Spain

(Received 10 June 1966)

The "linear" counterpart of the problem of analytic group extensions of the Poincaré group is presented in terms of the considerably simpler (but less general) analysis of Lie algebra extensions of the Poincaré algebra P. After easily proving with this technique that every C kernel (P, θ) has an extension and that every such extension is inessential, the problem of analyzing the central extensions of P is carried out with the well-expected result that every such extension is trivial. But contrary to some claims, we exhibit an example which explicitly shows an essential noncentral extension of P.

INTRODUCTION

THE problem of group extensions of the Poincaré group $\mathfrak{P}$ has been thoroughly investigated and brilliantly exposed by Michel.[1-4] The relevance of such analysis to the question of relating the internal symmetries and the relativistic invariance is widely recognized. And, as physicists are apparently more familiarized in general with Lie algebra methods than with global group techniques, we thought that it might be worthwhile to bridge the existing gap by rederiving some results of Michel in the evidently more restricted but much simpler context of Lie algebra extensions of the Poincaré algebra P. Even though this "linearization" of some already well-established facts on group extensions of $\mathfrak{P}$ presents no special difficulty, as it will be seen, the formalism of Lie algebra extensions may prove quite useful in the actual construction of *essential noncentral* extensions of P, *contrary* to a claim by Michel,[3,4] who asserts that no such extension exists. The possibilities that this existence of essential noncentral extensions of the Poincaré algebra may afford as to the question of finding a physically interesting mixing of the Poincaré group $\mathfrak{P}$ and some internal analytic group $\mathfrak{K}$ (which must necessarily be[4] nonsemisimple nor compact, for otherwise the extension would be central) is left completely unexplored in this paper.

Section 1 contains some remarks concerning the correspondence between Lie algebra extensions and analytic group extensions. Some examples are presented to illustrate how the natural relationship between both techniques is not so complete as expected.

In Sec. 2 the extensions "by" the Poincaré algebra P are examined and it is proved that only the trivial extensions by P are to be taken into account if a physically admissible mass spectrum is to be obtained.

Next, the central extensions of P are briefly analyzed in Sec. 3 and are shown to be trivial, as expected from the much more general results by Michel[2] on abstract group central extensions of $\mathfrak{P}$.

Section 4 deals with the question of noncentral extensions of P. Contrary to a statement formulated by Michel,[3,4] an explicit example of an essential noncentral extension of P is exhibited therein.

Finally, for the reader's convenience, Appendixes A and B summarize some basic material about cohomology spaces and extensions of Lie algebras.

Notation and Definitions

All Lie algebras and modules considered in Secs. 1–4 (respectively, Appendixes A, B) are assumed to be defined over the real field R (respectively, over an arbitrary field Φ of characteristic 0) and *finite dimensional.*

If A is a Lie algebra, $Z(A)$, $D(A)$, $I(A)$ stand, respectively, for its center, its derivation Lie algebra, and the ideal of $D(A)$ consisting of the inner derivations of A. The quotient Lie algebra $D(A)/I(A)$ is denoted by $\Delta(A)$.

Let A be a Lie algebra, and let A_1 and A_2 be, respectively, an ideal and a subalgebra of A. If $A_1 \cap A_2 = 0$ and if any $a \in A$ can be expressed as $a = a_1 + a_2$ with $a_1 \in A_1$, $a_2 \in A_2$, we write $A = A_1 + A_2$ and call A the semidirect sum of A_1 and A_2. If both $A = B + C$ and $A = C + B$, we say that A is the direct sum of B and C and write $A = B \oplus C$.

P stands for the Poincaré algebra, i.e., the *real* Lie algebra of the Poincaré (inhomogeneous connected Lorentz) group $\mathfrak{P}$; clearly, $P = T + L$, T being the Abelian ideal of P spanned by the generators of the translation group $\mathfrak{T}$, and L stands for the Lie algebra of the connected Lorentz group $\mathfrak{L}$.

[1] L. Michel, *Invariance in Quantum Mechanics* and *Group Extension* (Gordon and Breach Science Publishers, Inc., New York, 1965).
[2] L. Michel, Nucl. Phys. **57**, 356 (1964).
[3] L. Michel, Phys. Rev. **137**, B405 (1965).
[4] L. Michel, in *Symmetry Principles at High Energy* (W. H. Freeman & Co., San Francisco, 1965), pp. 331–350.

1. LIE ALGEBRA EXTENSIONS AND ANALYTIC GROUP EXTENSIONS

By introducing suitable continuity-like requirements into the abstract theory[1] of group kernels and extensions, a theory of *analytic* group kernels and extensions has been developed[5,6] which is closely linked (through the linearization process) to the analogous theory for Lie algebras (see Appendix B).

We present here a few examples to show that the natural relationship existing between that analytic theory and the corresponding one for Lie algebras cannot be as simple and complete as naively expected. The obvious reason for this is based on the well-known fact that a local morphism of an analytic group into another does not always admit a global extension.

Given two analytic groups $\mathcal{B}$, $\mathcal{C}$, with respective Lie algebras B, C, we denote by $\mathcal{K}(\mathcal{C}, \mathcal{B})$ the set of all analytic (or coverable) $\mathcal{C}$-kernels $(\mathcal{B}, g)$. [Let us briefly recall that by an analytic kernel $(\mathcal{C}, \mathcal{B}, g)$ or an analytic $\mathcal{C}$-kernel $(\mathcal{B}, g)$ it is meant[5] a couple of analytic groups $\mathcal{C}$, $\mathcal{B}$ together with an algebraic morphism $g: \mathcal{C} \to \text{Out } \mathcal{B} = \text{Aut}_c\, \mathcal{B}/\text{Int } \mathcal{B}$. $\text{Aut}_c\, \mathcal{B}$ stands for the Lie group of continuous automorphisms of $\mathcal{B}$, such that g can be analytically lifted, with domain a simply connected covering group $\bar{\mathcal{C}}$ of $\mathcal{C}$, i.e., such that there exists an analytic map $\bar{g}: \bar{\mathcal{C}} \to \text{Aut}_c\, \mathcal{B}$ with the property that the following diagram

$$\begin{array}{ccc} \bar{\mathcal{C}} & \xrightarrow{\bar{g}} & \text{Aut}_c\, \mathcal{B} \\ {\scriptstyle r}\downarrow & & \downarrow{\scriptstyle s} \\ \mathcal{C} & \xrightarrow{g} & \text{Out } \mathcal{B} \end{array} \qquad (1.1)$$

is commutative, where r (respectively, s) is a covering (respectively, the natural) morphism.] Similarly, let $K(C, B)$ denote the set of all kernels (C, B, θ) (see Appendix B), C and B being kept fixed. It can be shown[5] that there exists a naturally induced map τ of $\mathcal{K}(\mathcal{C}, \mathcal{B})$ into $K(C, B)$ which is not surjective, in general, as illustrated by Example (1)

Example (1): Let $\mathcal{B} \approx \mathcal{C} \approx U(1)$; as $\text{Out } U(1) \approx \text{Aut}_c\, U(1) \approx Z_2$, $\mathcal{K}(\mathcal{C}, \mathcal{B})$ consists of a single element, namely, the trivial kernel $(\mathcal{C}, \mathcal{B}, g = 0)$. This is not the case, though, for $K(C, B)$, since the Lie algebras B, C are one-dimensional, $\Delta(B) \approx B \approx C$, and C obviously has nontrivial one-dimensional representations. Q.E.D.

Now let $\text{Ext}_g(\mathcal{C}, \mathcal{B})$ denote the set of (analytically) equivalent classes of analytic extensions of an extendible[5] element $(\mathcal{C}, \mathcal{B}, g)$ of $\mathcal{K}(\mathcal{C}, \mathcal{B})$. And let $(C, B, \theta) = \tau(\mathcal{C}, \mathcal{B}, g)$. It can be proved[6] in the standard way that there exists a natural map σ of $\text{Ext}_g(\mathcal{C}, \mathcal{B})$ into $\text{Ext}_\theta(C, B)$. Example (2) [respectively, (3)] below shows that σ is not necessarily injective (respectively, surjective).

Example (2): Let $\mathcal{B} \approx U(1)$, $\mathcal{C} \approx SO(3)$, $g = 0$ (and consequently $\theta = 0$). As C is simple, $H^2(C, B) = 0$ (see Appendix A), B being a trivial C-module. Therefore $\text{Ext}_0(C, B)$ consists of a single element: the class of trivial extensions. However, $\text{Ext}_0(\mathcal{C}, \mathcal{B})$ has at least two elements: the class of the trivial extensions, and that class containing the essential central extension $U(2)$ of $SO(3)$ by $U(1)$. Q.E.D.

Example (3): Let $\mathcal{B} \approx U(1)$, $\mathcal{C} \approx T_2$ (two-dimensional toral group), and $g = 0$ (therefore $\theta = 0$). If the short exact sequence

$$1 \to U(1) \to \mathcal{A} \to T_2 \to 1 \qquad (1.2)$$

defines an analytic extension $\mathcal{A}$ of T_2 by $U(1)$, $\mathcal{A}$ must be a three-dimensional compact analytic group, and therefore (see Ref. 7, p. 190) analytically isomorphic either to T_3, $SO(3)$ or $SU(2)$. The last two possibilities must be ruled out, for $SO(3)$ and $SU(2)$ are simple and have thus no invariant one-parameter subgroup, and (1.2) requires that $\mathcal{A}$ has one such subgroup. Therefore $\text{Ext}_0(\mathcal{C}, \mathcal{B})$ consists only of the equivalence class of trivial extensions while this is not the case for $\text{Ext}_0(C, B)$, since one may simply see that the second cohomology space $H^2(C, B)$, with B a trivial C-module, is one-dimensional. Q.E.D.

2. EXTENSIONS "BY" THE POINCARÉ ALGEBRA

Although this case is physically uninteresting, it is very simple from the mathematical viewpoint, even if formulated with the greatest generality in terms of abstract groups.[3,4] The reason for its extreme simplicity when Lie algebra extensions are considered lies upon the fact that, just like Aut $\mathcal{I}$ is a semidirect product, the extension

$$0 \longrightarrow I(P) \xrightarrow{i} D(P) \xrightarrow{e} \Delta(P) \longrightarrow 0 \qquad (2.1)$$

is also inessential, i being the identity injection, and e the canonical epimorphism.

In fact, let $d \in D(P)$; the restriction $d \mid L$ is a derivation of L into P, and the semisimplicity of L implies[8] the existence of $p_d \in P$ such that

$$dl = [p_d, l] \quad \text{for all} \quad l \in L. \qquad (2.2)$$

As the ideal T is absolutely irreducible as a left L-module, such p_d is unique. Let $e': P \to I(P)$ be the canonical epimorphism which assigns to each $p_0 \in P$ the inner derivation $e'p_0: p \to [p_0, p]$ [actually e' is an isomorphism because $Z(P) = 0$]. Writing $d' = d - e'p_d$,

[5] R. A. Macaulay, Trans. Am. Math. Soc. **95**, 530 (1960).
[6] G. Hochschild, Ann. Math. **54**, 96, 537 (1951).
[7] D. Montgomery and L. Zippin, *Topological Transformation Groups* (Interscience Publishers, Inc., New York, 1955).
[8] N. Jacobson, *Lie Algebras* (Interscience Publishers, Inc., New York, 1962).

it is clear that $d'L = 0$, so that the restriction $d' \mid T$ (note that T is the radical of P and therefore a characteristic ideal[9]) is just a multiple of the identity because of the absolute irreducibility of T under L. Consequently, $\Delta(P)$ is at most one-dimensional and so there exists a monomorphism $e_1 : \Delta(P) \to D(P)$ such that $ee_1 = 1_{\Delta(P)}$ (it suffices obviously to take as e_1 any linear injective map satisfying $ee_1 = 1_{\Delta(P)}$). Q.E.D.

That $\Delta(P)$ is actually one-dimensional can be seen by noting that any linear endomorphism of P of the form $t + l \to \lambda t$, with $t \in T$, $l \in L$, $\lambda \in R$, defines a (noninner if $\lambda \neq 0$) derivation of P. [Finally, it is clear that a completely similar argument proves that the extension

$$0 \to I(P^c) \to D(P^c) \to \Delta(P^c) \to 0$$

is also inessential, where P^c stands for the complexification of P.]

Now, we can establish the expected result: "Every C-kernel (P, θ) has an extension, and all such associated extensions are equivalent and inessential."

In fact, since $Z(P) = 0$, (k) in Appendix B applies which shows that every C-kernel (P, θ) has an extension and all extensions inducing (C, P, θ) are equivalent. It suffices therefore to prove that (C, P, θ) has an inessential extension: introduce in $C \times P$ the bracket operation

$$[(c, p), (c', p')] = ([c, c'], [p, p'] + \gamma(c)p' - \gamma(c')p), \tag{2.3}$$

where $\gamma = e_1\theta$. Under (2.3) $C \times P$ becomes a Lie algebra A, and

$$0 \longrightarrow P \xrightarrow{\beta} A \xrightarrow{\alpha} C \longrightarrow 0, \tag{2.4}$$

where $\alpha(c, p) = c$, $\beta p = (p, 0)$, is a short exact sequence which defines an inessential extension A of C by P, and induces the C-kernel (P, θ). Q.E.D.

Furthermore, due to the aforementioned freedom of choice of e_1 as any linear injective map of $\Delta(P)$ into $D(P)$, subject only to the condition $ee_1 = 1_{\Delta(P)}$, and the existence of the particular cross section [in the fiber bundle[1] $D(P)$ with base $\Delta(P)$] consisting of the (noninner for $\lambda \neq 0$) derivations of the type $t + l \to \lambda t$ ($\lambda \in R$), it is plain that γ in (2.3) can be chosen so that $\gamma(c)(t + l) = \lambda(c)t$, $c \to \lambda(c)$ being a linear map of C into R. Consequently, given any extension A of C by P inducing (C, P, θ), it is always possible to inject C (by means of some α_1) homomorphically into A in such a way that

$$A = T + (L \oplus C'), \quad \text{with} \quad C' = \alpha_1 C, \tag{2.5}$$

where, if $c' = \alpha_1 c$, $[c', t] = \lambda(c)t$.

Therefore, the extension is trivial if and only if $\theta = 0$, i.e., if and only if such extension is central. This is necessarily the case, of course, if C has no nontrivial one-dimensional representation, for instance, if $[C, C] = C$.

Finally, without entering the discussion [which incidentally turns out to be unnecessary from the physical viewpoint) as to whether an analytic group extension $\mathcal{A}$ of $\mathcal{C}$ by $\mathcal{P}$ inducing (2.5) exists or not ($\mathcal{C}$ being an analytic group with Lie algebra C], we just remark that, assuming this to be the case, Mackey's theory[10] of induced representations allows us immediately to draw the conclusion that an irreducible unitary continuous representation U of $\mathcal{A}$ in a separable Hilbert space, locally faithful for the Poincaré group $\mathcal{P}$ and containing some real nonzero mass, will have a continuous mass spectrum filling the positive real axis unless $\theta = 0$ [note that U must necessarily be locally faithful for $\mathcal{A}$ if $\theta \neq 0$ as a consequence of (2.5) and its assumed local faithfulness for $\mathcal{P}$]. Physically, thus, only the locally *trivial* extensions

$$\mathcal{A} \overset{\text{loc}}{\approx} \mathcal{P} \otimes \mathcal{C}$$

of $\mathcal{C}$ by $\mathcal{P}$ are relevant.

3. CENTRAL EXTENSIONS OF P

The problem of classifying the abstract group central extensions of the Poincaré group $\mathcal{P}$ by an abstract group $\mathcal{K}$ has been exhaustively dealt with by Michel[2] when the center $\mathfrak{Z}(\mathcal{K})$ of $\mathcal{K}$ has no divisible subgroup, with the result that any such extension is of the form $\mathcal{K} \otimes \bar{\mathcal{P}}/Z_2(\alpha)$, $\bar{\mathcal{P}}$ being the universal covering group of $\mathcal{P}$ and $Z_2(\alpha)$ a two-element group generated by $(\mathcal{k}_0, \bar{\mathcal{p}}_0)$, where $\mathcal{k}_0 \in \mathfrak{Z}(\mathcal{K})$ and $\mathcal{k}_0^2 = 1$, and $\bar{\mathcal{p}}_0 \in \bar{\mathcal{P}}$ is the 2π rotation. In the remaining cases, if there exists any solution other than the direct product, Michel[2] has shown them to be quite pathological and useless for the physicists.

These general results of Michel strongly suggest what will be the obvious conclusion as to the central extensions of P: "Any central extension of P by a Lie algebra B is trivial." For the mere purpose of illustration, let us see how this comes out.

It will be sufficient to prove (see Appendix B) that $H^2(P, Z) = 0$, $Z = Z(B)$ being the center of B considered as a trivial left P-module. But a theorem of Hochschild and Serre[11] (frequently quoted by Michel) allows us to write

$$H^2(P, Z) \approx \sum_{i+j=2} H^i(L, R) \otimes H^j(P, L, Z) \tag{3.1}$$

due to the semisimplicity of L.

[9] N. Bourbaki, *Groupes et algèbres de Lie* (Hermann & Cie., Paris, 1960), Chap. 1, ASI 1285.

[10] G. Mackey, Acta Math. **99**, 265 (1958).

[11] G. Hochschild and J. P. Serre, Ann. Math. **57**, 591 (1953).

As $H^1(L, R) = H^2(L, R) = 0$ and $H^0(L, R) \approx R$, we have

$$H^2(P, Z) \approx R \otimes H^2(P, L, Z). \tag{3.2}$$

Consequently, we have just to prove that the 2nd relative cohomology space of P mod L is null, i.e., in our specific case this is equivalent to prove that there exists no nonzero $(2; T, Z)$-cochain f_2 such that

$$f_2([l, t_1], t_2) + f_2(t_1, [l, t_2]) = 0 \tag{3.3}$$

for all $l \in L$, $t_1, t_2 \in T$.

Let $\{t_0, t_1, t_2, t_3\}$ be the standard basis for T. Equation (3.3) implies

$$f_2(t_\mu, [l, t]) = 0 \tag{3.4}$$

for every $l \in L_\mu = \{l \in L : [l, t_\mu] = 0\}$, μ being fixed and denoting one of the indices 0 to 3. The subalgebra T_μ spanned by $\{t_0, t_1, t_2, t_3\} - \{t_\mu\}$ is stable and irreducible under L_μ; therefore (3.4) implies $f_2(t_\mu, t) = 0$ for all $t \in T_\mu$, and as $f_2(t_\mu, t_\mu) = 0$, then $f_2 = 0$.

Q.E.D.

4. NONCENTRAL EXTENSIONS: EXISTENCE OF ESSENTIAL EXTENSIONS OF P WITH $\theta \neq 0$

It has been claimed by Michel[3,4] that the problem of analytic noncentral group extensions of the Poincaré group is completely solved in the mathematical literature, because when this question is formulated in terms of Lie algebra extensions the (previously referred to) theorem of Hochschild and Serre leads to the simple conclusion that every such Lie algebra extension of P is inessential, just as in the case of central extensions.

That this claim is completely unjustified is next shown with an explicit counterexample. In view of the result (j) in Appendix B, it is plain that Michel's assertion is right if and only if $H^2(P, M) = 0$ for every left P-module M, and since (3.1) holds for any (not necessarily trivial) left P-module Z, we should just need to prove that $H^2(P, L, M) \neq 0$ for some left P-module M in order to invalidate that claim. An example of an essential noncentral extension of P is exhibited instead which indirectly provides an instance of $H^2(P, L, M) \neq 0$.

Let $\bar{M}$ be a non-null vector space and let $f_2 : P \times P \to \bar{M}$ be an arbitrary nonzero bilinear alternating map such that $f_2(p_1, p_2) = 0$ whenever $p_1 \in L$. Let M be the image domain of f_2. By assumption, thus, $M \neq 0$; obviously M is also an Abelian Lie algebra.

Let us define

$$(t)_M f_2(p_1, p_2) = 0 \text{ for every } t \in T,$$
$$(l)_M f_2(p_1, p_2) = f_2([l, p_1], p_2) + f_2(p_1, [l, p_2]) \quad \text{for all } l \in L. \tag{4.1}$$

It is easy to check that (4.1) allows us to endow M with a structure of left P-module. And finally, let us introduce in $P \times M$ the bracket composition

$$[(p, m), (p', m')] = ([p, p'], (p)_M m' - (p')_M m - f_2(p, p')), \tag{4.2}$$

which can be very simply verified to induce in $P \times M$ a Lie algebra structure. If A denotes such Lie algebra, a short exact sequence

$$0 \longrightarrow M \xrightarrow{\beta} A \xrightarrow{\alpha} P \longrightarrow 0 \tag{4.3}$$

can be defined by means of the following morphisms:

$$\beta m = (0, m) \quad \text{for} \quad m \in M,$$
$$\alpha(p, m) = p \quad \text{for} \quad (p, m) \in P \times M. \tag{4.4}$$

The morphism $\theta : P \to \Delta(M) \approx D(M)$ (for M is Abelian) associated to (4.3) is in this case defined [previous identification of $\Delta(M)$ and $D(M)$] by

$$\theta p : m \to (p)_M m \tag{4.5}$$

and from (4.1), the construction of M and the fact that $[L_{t_1}, t_2] \approx T/Rt_1$ whenever t_1 is a non-null element of T such that the one-dimensional subspace Rt_1 spanned by t_1 does not contain $t_2 (\in T)$, and $L_{t_1} = \{l \in L : [l, t_1] = 0\}$, it easily follows that $\theta \neq 0$.

The extension A of P by M defined by (4.3) is thus noncentral. Furthermore, it is an essential extension: in fact, suppose otherwise that there is a monomorphism $\alpha_1 : P \to A$ such that $\alpha\alpha_1 = 1_P$, and let $\alpha_1 p = (p', m(p))$, in particular $\alpha_1 t = (t', m(t))$. As T is the unique proper ideal of P, $\alpha_1 T$ will also be the unique proper ideal of $\alpha_1 P$, and (4.2) leads therefore to $t' \in T$. Since finally $\alpha_1 T$ must be Abelian, the result clearly follows that $f_2(t_1, t_2) = 0$ for all $t_1, t_2 \in T$ and consequently $f_2 = 0$, i.e., $M = 0$. This contradicts our initial assumption $M \neq 0$. Q.E.D.

It is very simple to verify that $f_2 \in Z^2(P, L, M)$ and defines a nonzero element in $H^2(P, L, M)$, whence $H^2(P, L, M) \neq 0$, as expected.

ACKNOWLEDGMENTS

The author wishes to thank Dr. F. J. Yndurain and Dr. G. Madurga for the benefit of several discussions. He is specially indebted to Professor L. Michel for having kindly pointed out to him an error in the preprint concerning the original statement (c) in Appendix B.

This paper was partially supported by the Ayuda a la Investigación en la Universidad and the Junta de Energía Nuclear.

APPENDIX A. COHOMOLOGY SPACES OF LIE ALGEBRAS

Let A be a Lie algebra, and let M be a left A-module (let us recall that in this appendix and in Appendix B the Lie algebras and modules are supposed to be finite dimensional and defined over an arbitrary base field Φ of characteristic 0). By $C^n(A, M)$ $(n \geq 0)$ we denote[8,12] the linear space of the multilinear alternating maps of $A \times A \times \cdots \times A$ (n factors) into M. By definition, $C^0(A, M) = M$. The elements of $C^n(A, M)$ are called $(n; A, M)$-cochains.

Let $\delta_n: C^n(A, M) \to C^{n+1}(A, M)$ be the linear map defined by

$$(\delta_n f_n)(a_1, \cdots, a_{n+1}) = \sum_{1 \leq i \leq n+1} (-1)^{i+1} (a_i)_M f_n(a_1, \cdots, \hat{a}_i, \cdots, a_{n+1}) + \sum_{1 \leq i < j \leq n+1} (-1)^{i+j} \times f_n([a_i, a_j], a_1, \cdots, \hat{a}_i, \cdots, \hat{a}_j, \cdots, a_{n+1}),$$

where $(a)_M$ denotes the linear endomorphism in M associated to $a \in A$, and the caret over an argument means that this argument is omitted.

It can be shown[8,12] that $\delta_{n+1}\delta_n = 0$ $(n \geq 0)$. Let $Z^n(A, M) = \operatorname{Ker} \delta_n$, $B^n(A, M) = \operatorname{Im} \delta_{n-1}$ [by definition $B^0(A, M) = 0$]. The elements of $Z^n(A, M)$ are called $(n; A, M)$-cocyles and those of $B^n(A, M)$ $(n; A, M)$-coboundaries. The relation $\delta_{n+1}\delta_n = 0$ implies $B^n(A, M) \subset Z^n(A, M)$.

The quotient space $H^n(A, M) = Z^n(A, M)/B^n(A, M)$ is called the n-dimensional (or nth) cohomology space or group of A relative to the A-module M.

Clearly $H^0(A, M) \approx Z^0(A, M) = \operatorname{Ker} \delta_0$; therefore, $H^0(A, M)$ is isomorphic to the linear subspace of M consisting of the "invariant" elements of the A-module M, i.e., those $m \in M$ such that $(a)_M m = 0$ for every $a \in A$. Finally, it is obvious that $C^n(A, M) = 0$, and hence $H^n(A, M) = 0$, for $n > \dim A$.

If A is semisimple, the following statements hold[8,12]:

(i) $H^1(A, M) = H^2(A, M) = 0$;

(ii) if the A-module M is irreducible and (A) $M \neq 0$, then $H^n(A, M) = 0$ for every $n \geq 0$;

(iii) $H^3(A, \Phi) \neq 0$, $H^4(A, \Phi) = 0$.

Similarly, it can also be proved[12] that:

(iv) A is semisimple if and only if $H^1(A, M) = 0$ for every A-module M.

Finally, we have the useful concept[12] of relative cohomology spaces $H^n(A, A', M)$ of A mod A', A' being a subalgebra of A. An element $f_n \in C^n(A, M)$ is called orthogonal to A' if

$$(a)_M f_n(a_1, \cdots, a_n) = \sum_{1 \leq i \leq n} f_n(a_1, \cdots, [a, a_i], \cdots, a_n) \quad \text{for all} \quad a \in A'$$

and

$$f_n(a_1, \cdots, a_n) = 0 \quad \text{whenever} \quad a_1 \in A'.$$

It can be proved[12] that, if f_n is orthogonal to A', then $\delta_n f_n$ is also. Let $C^n(A, A', M)$ denote the linear subspace of $C^n(A, M)$ consisting of the $(n; A, M)$-cochains orthogonal to A', and let

$$Z^n(A, A', M) = Z^n(A, M) \cap C^n(A, A', M),$$

$$B^n(A, A', M) = \delta_{n-1} C^{n-1}(A, A', M) \quad \text{with} \quad B^0(A, A', M) = 0.$$

The relative cohomology space $H^n(A, A', M)$ of A mod A' is then defined as

$$H^n(A, A', M) = Z^n(A, A', M)/B^n(A, A', M),$$

and the following statements[12] are true:

(v) if A' is an ideal of A and the A'-module M is trivial, then

$$H^n(A, A', M) \approx H^n(A/A', M) \quad (n \geq 0);$$

(vi) if A is semisimple and the A-module M is irreducible with (A) $M \neq 0$, then

$$H^n(A, A', M) = 0 \quad \text{for all} \quad n \geq 0.$$

APPENDIX B. EXTENSIONS OF LIE ALGEBRAS

The problem of group extensions has been brilliantly exposed by Michel in his Istanbul lecture notes.[1] The general procedure carries entirely over to the case of extensions of Lie algebras, with the natural replacements, and here we just recall some basic results, extracted from the mathematical papers and treatises referred to below. As to the exact sequences and commutative diagrams language we keep to the notations of Ref. 1, with the single difference that the composition of maps α and β will be simply written as $\alpha\beta$ instead of $\alpha \circ \beta$. The objects in the diagrams will be Lie algebras and the morphisms will be Lie algebra homomorphisms.

By an *extension* A of a Lie algebra C by B, it is meant[9,12,13,14] a short exact sequence

$$0 \longrightarrow B \xrightarrow{\beta} A \xrightarrow{\alpha} C \longrightarrow 0. \qquad \text{(B1)}$$

Clearly, $C \approx A/\operatorname{Ker} \alpha$, $B \approx \operatorname{Ker} \alpha$. (Some authors,[14] however, call A an extension of B by C; although this denomination seems more natural, we keep to that given above, for it is in accordance with that of Michel for group extensions.)

For instance, if $D(B)$ and $I(B)$ denote, respectively, the derivation algebra and inner derivation algebra of

[12] C. Chevalley and S. Eilenberg, Trans. Am. Math. Soc. 63, 85 (1948).

[13] H. Cartan and S. Eilenberg, *Homological Algebra* (Princeton University Press, Princeton, New Jersey, 1956).

[14] S. MacLane, *Homology* (Springer-Verlag, Berlin, 1963).

B, and $\Delta(B) = D(B)/I(B)$, a short exact sequence is the following:

$$0 \longrightarrow I(B) \xrightarrow{i} D(B) \xrightarrow{e} \Delta(B) \longrightarrow 0, \qquad \text{(B2)}$$

where i (respectively, e) is the identity injection of $I(B)$ into $D(B)$ [respectively, the canonical map of $D(B)$ onto $\Delta(B)$].

Similarly, if $Z(B)$ is the center of B, one has the exact sequence

$$0 \longrightarrow Z(B) \xrightarrow{i'} B \xrightarrow{e'} I(B) \longrightarrow 0, \qquad \text{(B3)}$$

where i' is the identity injection of $Z(B)$ into B, and e' the natural epimorphism $B \to I(B)$.

Two extensions A, A' of C by B are called *equivalent* if there exists a morphism $\gamma: A \to A'$ such that the following diagram:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & B & \xrightarrow{\beta} & A & \xrightarrow{\alpha} & C & \longrightarrow & 0 \\ & & \| & & \downarrow{\scriptstyle \gamma} & & \| & & \\ 0 & \longrightarrow & B & \xrightarrow{\beta'} & A' & \xrightarrow{\alpha'} & C & \longrightarrow & 0 \end{array} \qquad \text{(B4)}$$

is commutative. γ is an isomorphism and this relation defined between the extensions A and A' is an *equivalence relation.*

An extension (B1) is called

(i) *inessential*, when there is an exact sequence $0 \longrightarrow C \xrightarrow{\alpha_1} A$ such that $\alpha\alpha_1 = 1_C$ (identity on C);

(ii) *trivial*, when there is an extension A of B by C

$$0 \longrightarrow C \xrightarrow{\alpha_1} A \xrightarrow{\beta_1} B \longrightarrow 0$$

such that $\alpha\alpha_1 = 1_C$, $\beta_1\beta = 1_B$.

(A non-inessential extension is called, of course, *essential.*)

It is very simple to prove that

(a) if (B1) is inessential (respectively, trivial) then

$$A = \operatorname{Ker}\alpha + \operatorname{Im}\alpha_1, \quad \text{with} \quad \operatorname{Im}\alpha_1 \approx \operatorname{Coker}\beta$$

(respectively, $A = \operatorname{Ker}\alpha \oplus \operatorname{Im}\alpha_1$);

(b) if an extension is inessential, or trivial, the same holds for an equivalent extension.

By a C-kernel (B, θ) [also denoted by (C, B, θ)] it is understood[14,15] as a pair of Lie algebras B, C, together with a morphism $\theta: C \to \Delta(B)$. Given a C-kernel (B, θ) and a linear injective map $\lambda: \Delta(B) \dashrightarrow D(B)$ such that $e\lambda = 1_{\Delta(B)}$, the map θ': $C \dashrightarrow D[Z(B)]$, which assigns to each $c \in C$ the derivation $z \to (\lambda\theta c)z$, $z \in Z(B)$, is a morphism that does not depend on the choice of λ. The C-kernel $(Z(B), \theta')$ is called the *central* kernel of (C, B, θ).

Just as for abstract groups,[1] each extension (B1) of C by B has an associated C-kernel (B, θ) defined by the following commutative diagram:

$$\begin{array}{ccccccccc} & & 0 & & & & & & \\ & & \downarrow & & & & & & \\ & & Z(B) & & & & & & \\ & & {\scriptstyle i'}\downarrow & & & & & & \\ 0 & \longrightarrow & B & \xrightarrow{\beta} & A & \xrightarrow{\alpha} & C & \longrightarrow & 0 \\ & & {\scriptstyle e'}\downarrow & & \phi\wr\downarrow & & \theta\wr\downarrow & & \\ 0 & \longrightarrow & I(B) & \xrightarrow{i} & D(B) & \xrightarrow{e} & \Delta(B) & \longrightarrow & 0. \\ & & \downarrow & & & & & & \\ & & 0 & & & & & & \end{array} \qquad \text{(B5)}$$

Those C-kernels (B, θ) induced this way by an extension of C by B are said to have an extension. Equivalent extensions (B4) clearly define the same C-kernel (B, θ).

The extension A of C by B is called[2] *central* if $\theta = 0$, or equivalently, if $\operatorname{Im}\phi \subset \operatorname{Ker} e$. Clearly, the property of being central remains under equivalence. It is easy to check that

(c) given a fixed B, all central inessential extensions by B are trivial if and only if the extension (B3) is inessential.

In the family $\mathfrak{E}(C, Z, \psi)$ of all C-kernels (B, θ), which have a common identical central C-kernel (Z, ψ) [i.e., $Z(B) = Z$, $\theta' = \psi$], a composition law and an equivalence relation can be introduced[15] in such a way that the set of equivalence classes naturally inherits a structure of additive group, which in turn becomes a linear space $E(C, Z, \psi)$ by suitably defining the composition of a scalar and an equivalence class. The null vector of this linear space represents the class of those kernels in $\mathfrak{E}(C, Z, \psi)$ having an extension. For the case of abstract groups, the corresponding procedure leads[1] to an Abelian group which is shown to be isomorphic to the third cohomology group. In our case, however, $E(C, Z, \psi)$ is in general isomorphic to a linear *subspace* of $H^3(C, Z_\psi)$ [where Z_ψ denotes the Abelian Lie algebra Z considered as the C-module defined by $(c)_Z z = (\psi c)z$, $z \in Z$] and although in some cases $E(C, Z, \psi) \approx H^3(C, Z_\psi)$, only by allowing *infinite* dimensional Lie algebras B the isomorphism of $E(C, Z, \psi)$ with $H^3(C, Z_\psi)$ can be retrieved in all instances.[16] Let us briefly indicate how this monomorphism

$$\pi: E(C, Z, \psi) \to H^3(C, Z_\psi)$$

arises.

Given a C-kernel (B, θ) in $\mathfrak{E}(C, Z, \psi)$, let $\epsilon: C \dashrightarrow D(B)$ be a linear injective map such that $e\epsilon c = \theta c$ for every $c \in C$. Clearly,

$$\epsilon[c_1, c_2] - [\epsilon c_1, \epsilon c_2] \in I(B),$$

[15] G. Hochschild, Am. J. Math. **76**, 698 (1954).

[16] G. Hochschild, Am. J. Math. **76**, 763 (1954).

so that a bilinear alternating map $f_2: C \times C \to B$ can be chosen such that

$$\epsilon[c_1, c_2] = [\epsilon c_1, \epsilon c_2] + e'(f_2(c_1, c_2)).$$

The Jacobi identity leads to the relation

$$e'\left(\sum_{\text{cycl}} f_2(c_1, [c_2, c_3])\right) + e'\left(\sum_{\text{cycl}} (\epsilon c_1) f_2(c_2, c_3)\right) = 0,$$

whence

$$g_3(c_1, c_2, c_3) = \sum_{\text{cycl}} \{(\epsilon c_1) f_2(c_2, c_3) + f_2(c_1, [c_2, c_3])\} \quad \text{(B6)}$$

defines a trilinear alternating map $g_3: C \times C \times C \to Z$, i.e., $g_3 \in C^3(C, Z_\psi)$. A straightforward computation shows that $\delta_3 g_3 = 0$, i.e., $g_3 \in Z^3(C, Z_\psi)$. Moreover, a different choice of ϵ and/or f_2 may obviously change g_3, but all the $(3; C, Z_\psi)$-cocycles so obtained run over and completely exhaust a cohomology class; any such $(3; C, Z_\psi)$-cocycle g_3 so obtained is called[15,16] an *obstruction* of the C-kernel (B, θ) and the element of $H^3(C, Z_\psi)$ thereby defined is denoted by Obs (C, B, θ). The reason for the name of obstruction[14–16] lies in the fact that (C, B, θ) has an extension if and only if Obs $(C, B, \theta) = 0$. The part "if" is shown by a direct (lengthy but simple) computation. As to the "only if", note that if Obs $(C, B, \theta) = 0$ we can always choose, because of the previous remarks, the aforementioned maps ϵ and f_2 so that $g_3 = 0$. With such a choice, a Lie algebra structure can be introduced in $C \times B$ by means of

$$[(c, b), (c', b')] = ([c, c'], [b, b'] + (\epsilon c)b' - (\epsilon c')b - f_2(c, c')).$$

If A denotes this Lie algebra, it can be very simply verified that

$$0 \longrightarrow B \xrightarrow{\beta} A \xrightarrow{\alpha} C \longrightarrow 0$$

is the required extension inducing θ, wherein $\beta b = (0, b)$, $\alpha(c, b) = c$.

It is clear that if B is Abelian then $f_2 \in C^2(C, Z_\psi)$ and (B6) tells us in this case that $g_3 = \delta_2 f_2$, and consequently, every C-kernel (B, θ), with B Abelian, has an extension.

C-kernels (B, θ) integrating the same element of $E(C, Z, \psi)$ can be shown to define the same Obs $(C, ., .)$, and conversely; and the correspondence which assigns to such element of $E(C, Z, \psi)$ that common cohomology class Obs $(C, ., .)$ of obstructions can be shown to be linear and injective and is the previously introduced monomorphism π. Hochschild[16] has given a characterization of the image domain of π which implies that

(d) if C is semisimple then Im $\pi = 0$, so that in this case every element of $\delta(C, Z, \psi)$ has an extension;

(e) if C is solvable, Im $\pi = H^3(C, Z_\psi)$.

To end with this recollection of basic results, we briefly summarize some important known facts concerning Lie algebra extensions.

Let $\text{Ext}_\psi(C, Z)$ [respectively, $\text{Ext}_\theta(C, B)$] denote the set of equivalence classes of extensions of C by an Abelian Lie algebra Z with a fixed C-kernel (Z, ψ) [respectively, of C by the Lie algebra B with a fixed C-kernel (B, θ) having an extension]. Then it can be proved[15,17] with a similar technique of factor systems as for groups that:

(f) $\text{Ext}_\psi(C, Z)$ can be given a natural structure of linear space;

(g) $\text{Ext}_\psi(C, Z) \approx H^2(C, Z_\psi)$, the equivalence class of inessential extensions of C by Z with fixed kernel (C, Z, ψ) corresponding to the null element of $H^2(C, Z_\psi)$;

(h) if (C, B, θ) has (C, Z, ψ) as central kernel, given an extension A of C by B and an extension N of C by Z inducing respectively (C, B, θ) and (C, Z, ψ), a composition of these two extensions can be defined leading to an extension of C by B which also induces (C, B, θ). This composition is such that, when A is kept arbitrarily fixed while N is let to vary [so that *one* representative N is taken for each equivalence class, i.e., for each element of $\text{Ext}_\psi(C, Z)$], the equivalence classes of the extensions of C by B resulting from these compositions are all different and exhaust $\text{Ext}_\theta(C, B)$. $\text{Ext}_\theta(C, B)$ and $H^2(C, Z_\psi)$ are therefore coordinable.

It can also be proved[12] that:

(j) every extension of C is inessential if and only if $H^2(C, M) = 0$ for every C-module M.

Finally, the following statements are now obvious:

(k) if $Z(B) = 0$ then every C-kernel (B, θ) has an extension and all extensions inducing (C, B, θ) are equivalent;

(l) if $\theta = 0$, then the kernel (C, B, θ) admits an extension (consider the direct sum of C and B);

(m) if B is complete [i.e., if $Z(B) = 0$ and $\Delta(B) = 0$] a combination of (k) and (l) leads to the conclusion that all C-kernels (B, θ) necessarily have $\theta = 0$, have an extension, and all extensions of C by B are central, equivalent, and trivial.

[17] A. Galindo, Actas Reunión Fisica Teórica, Santander (1965) JEN Report.

Peculiarities of the Eight-Dimensional Space

A. GAMBA
Istituto di Fisica dell'Università, Genoa, Italy

(Received 19 July 1966)

A mathematical review of the peculiar properties of the space of eight dimensions is presented with the view of possible applications in the study of symmetries of elementary particles. This paper, written for physicists, is self-contained in that it does not require any previous knowledge of the subject nor any advanced mathematics.

1. INTRODUCTION

THE space of eight dimensions possesses peculiar features not to be found in any other dimensionality. Since the group of rotations in eight dimensions has been recently advocated as one of the possible extensions for the symmetries of elementary particles,[1] it seems worthwhile to investigate more closely the mathematics of eight-dimensional space. The present paper gives a short account of these properties for the benefit of the physicist. It does not require any previous knowledge of the subject nor any advanced mathematics.

Although most results of this paper can already be found in the literature—either written in an easily understandable[2] or in a rather difficult[3] mathematical language—some of the results, given in Sec. 3, are believed to be new.

2. VECTORS AND SPINORS

To point out analogies and differences, let us review briefly some aspects of the ordinary four dimensional (Dirac's) theory.

A. Four-Dimensional Theory

a. Vectors

Define four basic matrices $\gamma_1, \gamma_2, \gamma_3, \gamma_4$ satisfying

$$\gamma_k\gamma_l + \gamma_l\gamma_k = 2\delta_{kl} \quad (k, l = 1, 2, 3, 4). \tag{1}$$

To be specific, take the representation

$$\gamma_1 = \begin{pmatrix} & & & -i \\ & & -i & \\ & i & & \\ i & & & \end{pmatrix}, \quad \gamma_2 = \begin{pmatrix} & & & -1 \\ & & 1 & \\ & -1 & & \\ -1 & & & \end{pmatrix},$$

$$\gamma_3 = \begin{pmatrix} & & -i & \\ & & & i \\ i & & & \\ & -i & & \end{pmatrix}, \quad \gamma_4 = \begin{pmatrix} & & 1 & \\ & & & 1 \\ 1 & & & \\ & 1 & & \end{pmatrix}, \tag{2}$$

where all matrix elements not explicitly shown are equal to zero. All the γ's are Hermitian ($\bar{\gamma}_k = \gamma_k$).

A vector $\mathbf{X}$ is then represented by the matrix

$$\mathbf{X} = \sum_{k=1}^{4} x_k\gamma_k, \tag{3}$$

where x_1, x_2, x_3, x_4 are the "components" of the vector. One can verify easily that the square of the matrix $\mathbf{X}$ is equal to $x_1^2 + x_2^2 + x_3^2 + x_4^2$ times the unit vector. We write for short

$$\mathbf{F} = \mathbf{X}^2 = x_1^2 + x_2^2 + x_3^2 + x_4^2. \tag{4}$$

An important matrix in our representation is the following

$$\mathbf{B} = \begin{pmatrix} & & -i & \\ & & & \\ i & & & \\ & & & \\ \hline & & & -i \\ & & & \\ & i & & \end{pmatrix}; \quad \mathbf{B}^2 = 1, \tag{5}$$

where $\mathbf{B}$ transforms any vector $\mathbf{X}$ into its transpose $\mathbf{X}^T$ according to the rule

$$\mathbf{X}^T = \mathbf{BXB}. \tag{6}$$

An infinitesimal rotation by an angle θ in the plane k, l is accomplished by the operator

$$\mathbf{R}_{kl} = 1 + \tfrac{1}{2}\theta\gamma_k\gamma_l \tag{7}$$

acting on a vector $\mathbf{X}$ as follows:

$$\mathbf{X}' = \mathbf{R}_{kl}\mathbf{X}\mathbf{R}_{kl}^{-1}, \tag{8}$$

where $\mathbf{X}'$ being the new (rotated) vector. In fact, since to the first order

$$\mathbf{R}_{kl}^{-1} = 1 - \tfrac{1}{2}\theta\gamma_k\gamma_l, \tag{9}$$

one has

$$\begin{aligned} x'_m\gamma_m &= (1 + \tfrac{1}{2}\theta\gamma_k\gamma_l)x_m\gamma_m(1 - \tfrac{1}{2}\theta\gamma_k\gamma_l) \\ &= x_m\{\gamma_m + \tfrac{1}{2}\theta[\gamma_k\gamma_l\gamma_m - \gamma_m\gamma_k\gamma_l]\}, \end{aligned} \tag{10}$$

whence

$$\begin{aligned} x'_m &= x_m, \quad m \neq k, l, \\ x'_k &= x_k + \theta x_l, \\ x_l &= -\theta x_k + x_l. \end{aligned} \tag{11}$$

[1] Y. Ne'eman and I. Ozsvath, Phys. Rev. **138**, B1475 (1965); see also D. C. Peaslee, J. Math. Phys. 4, 910 (1963).
[2] E. Cartan, *Leçons sur la théorie des spineurs* (Hermann & Cie, Paris, 1938), Vols. I and II.
[3] C. C. Chevalley, *The Algebraic Theory of Spinors* (Columbia University Press, New York, 1954).

b. Spinors

A general spinor ξ in four dimensions is written as a four-row, one-column matrix

$$\xi = \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \psi_1 \\ \psi_2 \end{pmatrix}. \tag{12}$$

We also use the transpose spinor

$$\xi^T = (\varphi_1, \varphi_2, \psi_1, \psi_2). \tag{13}$$

The rotation (7) transforms the spinor as follows:

$$\xi' = \mathbf{R}_{kl}\xi. \tag{14}$$

It is easily seen that in our representation (2) a spinor of the type

$$\varphi = \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ 0 \\ 0 \end{pmatrix}, \tag{15}$$

i.e., with the last two components vanishing, remains of the same type under rotations (spinor of the first kind). The same is true of the spinor of the second kind

$$\psi = \begin{pmatrix} 0 \\ 0 \\ \psi_1 \\ \psi_2 \end{pmatrix}. \tag{16}$$

The quantity

$$\xi^T\mathbf{B}\xi \tag{17}$$

is invariant under rotations. In fact, using (6)

$$\begin{aligned}\xi'^T\mathbf{B}\xi' &= \xi^T(1 + \tfrac{1}{2}\theta\gamma_l^T\gamma_k^T)\mathbf{B}(1 + \tfrac{1}{2}\theta\gamma_k\gamma_l)\xi \\ &= \xi^T\mathbf{B}(1 + \tfrac{1}{2}\theta\gamma_l\gamma_k)(1 + \tfrac{1}{2}\theta\gamma_k\gamma_l)\xi = \xi^T\mathbf{B}\xi.\end{aligned} \tag{18}$$

Decomposition of ξ into spinors of first and of second kind shows that (17) can be written as a sum of two separate rotation invariants, i.e.,

$$\Phi = \varphi^T\mathbf{B}\varphi, \tag{19}$$

$$\Psi = \psi^T\mathbf{B}\psi. \tag{20}$$

B. Eight-Dimensional Theory

a. Vectors

Define eight basic matrices $A_1, A_2, \cdots, A_8$ satisfying

$$A_kA_l + A_lA_k = 2\delta_{kl} \quad (k, l = 1, 2, \cdots, 8). \tag{21}$$

Like the smallest nontrivial representation of Dirac's γ matrices has to be four-dimensional, similarly, in this case, it turns out that the A's have to be sixteen-dimensional. We choose an Hermitian representation of the form

$$A_k = \begin{pmatrix} 0 & \alpha_k \\ \beta_k & 0 \end{pmatrix} \quad (k = 1, 2, \cdots, 8), \tag{22}$$

where all the elements in (22) are 8×8 matrices. To specify a representation it is then sufficient to give the α_k—the β_k being then immediately determined from Hermiticity. Let us choose the representation

$$\alpha_1 = \begin{pmatrix} i & & & & & & & \\ & i & & & & & & \\ & & i & & & & & \\ & & & i & & & & \\ & & & & i & & & \\ & & & & & i & & \\ & & & & & & i & \\ & & & & & & & i \end{pmatrix}, \quad \alpha_2 = \begin{pmatrix} & & & & 1 & & & \\ & & & & & 1 & & \\ & & & & & & -1 & \\ & & & & & & & -1 \\ 1 & & & & & & & \\ & 1 & & & & & & \\ & & -1 & & & & & \\ & & & -1 & & & & \end{pmatrix},$$

$$\alpha_3 = \begin{pmatrix} 1 & & & & & & & \\ & -1 & & & & & & \\ & & -1 & & & & & \\ & & & -1 & & & & \\ & & & & 1 & & & \\ & & & & & 1 & & \\ & & & & & & 1 & \\ & & & & & & & -1 \end{pmatrix}, \quad \alpha_4 = \begin{pmatrix} & & & & i & & & \\ & & & & & i & & \\ & & & & & & i & \\ & & & & & & & i \\ -i & & & & & & & \\ & -i & & & & & & \\ & & -i & & & & & \\ & & & -i & & & & \end{pmatrix},$$

$$\alpha_5 = \begin{pmatrix} & & & & & & & -i \\ & & & & & & i & \\ & & & & -i & & & \\ & & & -i & & & & \\ & & i & & & & & \\ & -i & & & & & & \\ i & & & & & & & \\ & & & & & -i & & \end{pmatrix}, \quad \alpha_6 = \begin{pmatrix} & & & & & & 1 & \\ & & & & & 1 & & \\ & & & & 1 & & & \\ & & & -1 & & & & \\ & & 1 & & & & & \\ & -1 & & & & & & \\ 1 & & & & & & & \\ & & & & & -1 & & \end{pmatrix},$$

$$\alpha_7 = \begin{pmatrix} & & & & -i & & & \\ & & & -i & & & & \\ & & & & & & -i & \\ & & & & & & & -i \\ & & & & & -i & & \\ & & & & i & & & \\ & & i & & & & & \\ & i & & & & & & \end{pmatrix}, \quad \alpha_8 = \begin{pmatrix} & & & & 1 & & & \\ & & & & & 1 & & \\ & & & & & & 1 & \\ & & & & & & & 1 \\ & & & & & 1 & & \\ & & & & 1 & & & \\ & & 1 & & & & & \\ & 1 & & & & & & \end{pmatrix}. \tag{23}$$

A vector $\mathbf{X}$ is then represented by the matrix

$$\mathbf{X} = \sum_{k=1}^{8} x_k A_k, \tag{24}$$

where $x_1, x_2, \cdots, x_8$ are the "components" of the vector. One can verify easily that the square of the matrix $\mathbf{X}$ is equal to $x_1^2 + x_2^2 + \cdots + x_8^2$ times the unit matrix. We write for short

$$\mathbf{F} = \mathbf{X}^2 = x_1^2 + x_2^2 + \cdots + x_8^2. \tag{25}$$

An important matrix in our representation is the following:

$$\mathbf{B} = \begin{pmatrix} & & & & & & & 1 \\ & & & & & & -1 & \\ & & & & & 1 & & \\ & & & & -1 & & & \\ & & & -1 & & & & \\ & & 1 & & & & & \\ & -1 & & & & & & \\ 1 & & & & & & & \end{pmatrix}; \quad \begin{pmatrix} & & & & & & & -1 \\ & & & & & & 1 & \\ & & & & & -1 & & \\ & & & & 1 & & & \\ & & & 1 & & & & \\ & & -1 & & & & & \\ & 1 & & & & & & \\ -1 & & & & & & & \end{pmatrix}$$

$$\mathbf{B}^2 = 1. \tag{26}$$

$\mathbf{B}$ transforms any vector $\mathbf{X}$ into its transpose $\mathbf{X}^T$ according to the rule

$$\mathbf{X}^T = \mathbf{BXB}. \tag{27}$$

An infinitesimal rotation by an angle θ in the plane k, l is accomplished by the operator

$$\mathbf{R}_{kl} = 1 + \tfrac{1}{2}\theta A_k A_l \tag{28}$$

acting on a vector $\mathbf{X}$ as follows:

$$\mathbf{X}' = \mathbf{R}_{kl}\mathbf{X}\mathbf{R}_{kl}^{-1}, \tag{29}$$

$\mathbf{X}'$ being the new (rotated) vector. The proof is the same as for the four-dimensional case above.

b. Spinors

A general spinor ξ in eight dimensions has sixteen components. We could write it

$$\xi = \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \cdot \\ \cdot \\ \cdot \\ \varphi_8 \\ \psi_1 \\ \psi_2 \\ \cdot \\ \cdot \\ \cdot \\ \psi_8 \end{pmatrix} \tag{30}$$

in analogy with (12). However, for later convenience, we rename the components in a somewhat peculiar manner and write it instead

$$\xi = \frac{1}{\sqrt{2}} \begin{vmatrix} -\varphi_4 + i\varphi_6 \\ \varphi_7 + i\varphi_8 \\ \varphi_1 - i\varphi_3 \\ -\varphi_5 - i\varphi_2 \\ \varphi_5 - i\varphi_2 \\ \varphi_1 + i\varphi_3 \\ -\varphi_7 + i\varphi_8 \\ -\varphi_4 - i\varphi_6 \\ \psi_6 + i\psi_4 \\ \psi_8 - i\psi_7 \\ -\psi_3 + i\psi_1 \\ -\psi_2 + i\psi_5 \\ -\psi_2 - i\psi_5 \\ \psi_3 + i\psi_1 \\ \psi_8 + i\psi_7 \\ -\psi_6 + i\psi_4 \end{vmatrix} . \tag{31}$$

A rotation (28) transforms the spinor as

$$\xi' = \mathbf{R}_{kl}\xi. \tag{32}$$

In perfect analogy with the four-dimensional theory, a spinor having only the eight φ-components different from zero (spinor of the first kind) remains of the same type under rotations. The same is true of the spinor of the second kind, i.e., having only the eight ψ-components different from zero.

The quantity

$$\xi^T \mathbf{B} \xi \tag{33}$$

is invariant under rotations. Decomposing the spinor in the two parts, φ and ψ respectively, one obtains from it two separate invariants

$$\Phi = \varphi^T \mathbf{B} \varphi = \varphi_1^2 + \varphi_2^2 + \cdots + \varphi_8^2, \tag{34}$$

$$\Psi = \psi^T \mathbf{B} \psi = \psi_1^2 + \psi_2^2 + \cdots + \psi_8^2. \tag{35}$$

Thus far, the eight-dimensional theory is perfectly analogous to the four-dimensional theory, or to the theory in any even number of dimensions. What renders the eight-dimensional theory a very special case is the fact that, in eight dimensions—and in eight dimensions only—the vector **X**, the spinor φ, and the spinor ψ all have an equal number of components (namely, eight), so that the invariant forms **F**, Φ, and Ψ [formulas (25), (34), and (35)] all look the same. Let us examine this point more carefully.

c. Peculiarities of the Eight-Dimensional Theory

Consider for simplicity a specific rotation (28), say the rotation $\mathbf{R}_{12}$ transforming the component of a vector **X** as follows:

$$\begin{cases} x_1' = x_1 + \theta x_2, \\ x_2' = -\theta x_1 + x_2, \\ x_k' = x_k \qquad (k = 3, 4, 5, 6, 7, 8). \end{cases} \tag{36}$$

The corresponding transformations for the spinors φ and ψ are

$$\begin{cases} \varphi_1' = \varphi_1 - \tfrac{1}{2}\theta\varphi_2, \\ \varphi_2' = \tfrac{1}{2}\theta\varphi_1 + \varphi_2, \end{cases} \quad \begin{cases} \varphi_3' = \varphi_3 + \tfrac{1}{2}\theta\varphi_5, \\ \varphi_5' = -\tfrac{1}{2}\theta\varphi_3 + \varphi_5, \end{cases}$$
$$\begin{cases} \varphi_4' = \varphi_4 + \tfrac{1}{2}\theta\varphi_8, \\ \varphi_8' = -\tfrac{1}{2}\theta\varphi_4 + \varphi_8, \end{cases} \quad \begin{cases} \varphi_6' = \varphi_6 + \tfrac{1}{2}\theta\varphi_7, \\ \varphi_7' = -\tfrac{1}{2}\theta\varphi_6 + \varphi_7; \end{cases} \tag{37}$$

$$\begin{cases} \psi_1' = \psi_1 - \tfrac{1}{2}\theta\psi_2, \\ \psi_2' = \tfrac{1}{2}\theta\psi_1 + \psi_2, \end{cases} \quad \begin{cases} \psi_3' = \psi_3 - \tfrac{1}{2}\theta\psi_5, \\ \psi_5' = \tfrac{1}{2}\theta\psi_3 + \psi_5, \end{cases}$$
$$\begin{cases} \psi_4' = \psi_4 - \tfrac{1}{2}\theta\psi_8, \\ \psi_8' = \tfrac{1}{2}\theta\psi_4 + \psi_8, \end{cases} \quad \begin{cases} \psi_6' = \psi_6 - \tfrac{1}{2}\theta\psi_7, \\ \psi_7' = \tfrac{1}{2}\theta\psi_6 + \psi_7. \end{cases} \tag{38}$$

Similar expressions are obtained with other rotations of the type $\mathbf{R}_{1k}$ $(k = 3, 4, 5, 6, 7, 8)$. With rotations $\mathbf{R}_{kl}$ $(k, l = 2, 3, \cdots, 8; k \neq l)$, one obtains expressions that differ from the ones above by some changes in corresponding signs of vector and spinors. For example, for the $\mathbf{R}_{23}$ rotations, one has

$$\begin{cases} x_2' = x_2 + \theta x_3, \\ x_3' = -\theta x_2 + x_3, \\ x_k' = x_k \qquad (k = 1, 4, 5, 6, 7, 8), \end{cases} \tag{36'}$$

$$\begin{cases} \varphi_2' = \varphi_2 + \tfrac{1}{2}\theta\varphi_3, \\ \varphi_3' = -\tfrac{1}{2}\theta\varphi_2 + \varphi_3, \end{cases} \quad \begin{cases} \varphi_4' = \varphi_4 + \tfrac{1}{2}\theta\varphi_7, \\ \varphi_7' = -\tfrac{1}{2}\theta\varphi_4 + \varphi_7, \end{cases}$$
$$\begin{cases} \varphi_8' = \varphi_8 + \tfrac{1}{2}\theta\varphi_6, \\ \varphi_6' = -\tfrac{1}{2}\theta\varphi_8 + \varphi_6, \end{cases} \quad \begin{cases} \varphi_5' = \varphi_5 + \tfrac{1}{2}\theta\varphi_1, \\ \varphi_1' = -\tfrac{1}{2}\theta\varphi_5 + \varphi_1; \end{cases} \tag{37'}$$

$$\begin{cases} \psi_2' = \psi_2 + \tfrac{1}{2}\theta\psi_3, \\ \psi_3' = -\tfrac{1}{2}\theta\psi_2 + \psi_3, \end{cases} \quad \begin{cases} \psi_4' = \psi_4 + \tfrac{1}{2}\theta\psi_7, \\ \psi_7' = -\tfrac{1}{2}\theta\psi_4 + \psi_7, \end{cases}$$
$$\begin{cases} \psi_8' = \psi_8 + \tfrac{1}{2}\theta\psi_6, \\ \psi_6' = -\tfrac{1}{2}\theta\psi_8 + \psi_6, \end{cases} \quad \begin{cases} \psi_5' = \psi_5 - \tfrac{1}{2}\theta\psi_1, \\ \psi_1' = \tfrac{1}{2}\theta\psi_5 + \psi_1. \end{cases} \tag{38'}$$

In the above formulas, one considers the vector rotation (36) or (36′) as the *primary* rotation that *induces* the transformations (37) and (38) or (37′)–(38′) on the spinors [the correspondence being a one-to-two correspondence because of the angle $\frac{1}{2}\theta$ that appears in (37) and (38)]. However, (37) and (38) or (37′)

and (38′) are also rotations, and, since the quadratic forms $\mathbf{F}$, $\mathbf{\Phi}$, and $\mathbf{\Psi}$ are similar, one could think, for example, that $\boldsymbol{\varphi}$ is the vector and $\mathbf{X}$ and $\mathbf{\Psi}$ the spinors. A rotation by an angle θ of $\boldsymbol{\varphi}$ then induce corresponding rotations by angles $\frac{1}{2}\theta$ in $\mathbf{X}$ and $\mathbf{\Psi}$. In other words, it is only a matter of convention to define which is the vector and which the spinors of first and second kind; the three quantities $\mathbf{X}$, $\boldsymbol{\varphi}$, and $\boldsymbol{\psi}$ being perfectly equivalent. This property is referred to in the literature as "principle of triality."[4] [It has nothing to do, however, with the concept of triality[5] introduced recently in the literature of elementary particles in connection with the representations of the group $SU(3)$.] When $\mathbf{X}$, $\boldsymbol{\varphi}$, and $\boldsymbol{\psi}$ transform as vector and spinor of the first kind and spinor of the second kind, respectively, the quantity

$$\mathcal{F} = \boldsymbol{\varphi}^T \mathbf{B} \mathbf{X} \boldsymbol{\psi} \tag{39}$$

turns out to be invariant under rotations. In fact,

$$\begin{aligned}\mathcal{F}' &= \boldsymbol{\varphi}'^T \mathbf{B} \mathbf{X}' \boldsymbol{\psi}' \\ &= \boldsymbol{\varphi}^T (1 + \tfrac{1}{2}\theta A_l^T A_k^T) \mathbf{B} (1 + \tfrac{1}{2}\theta A_k A_l) \\ &\qquad \times \mathbf{X} (1 - \tfrac{1}{2}\theta A_k A_l)(1 + \tfrac{1}{2}\theta A_k A_l) \boldsymbol{\psi} \\ &= \boldsymbol{\varphi}^T \mathbf{B} (1 + \tfrac{1}{2}\theta A_l A_k)(1 + \tfrac{1}{2}\theta A_k A_l) \mathbf{X} \boldsymbol{\psi} \\ &= \boldsymbol{\varphi}^T \mathbf{B} \mathbf{X} \boldsymbol{\psi}. \end{aligned} \tag{40}$$

Written out explicitly, the invariant (39) is equal to

$$\begin{aligned}\mathcal{F} = \sum \{ &\varphi_k x_l \psi_m + \varphi_l x_m \psi_k + \varphi_m x_k \psi_l \\ &- \varphi_l x_k \psi_m - \varphi_m x_l \psi_k - \varphi_k x_m \psi_l \} \\ &+ \sum_{k=2}^{8} \{ \varphi_1 x_k \psi_k + \varphi_k x_1 \psi_k + \varphi_k x_k \psi_1 \} - \varphi_1 x_1 \psi_1, \end{aligned} \tag{41}$$

where the first sum is over the following seven values of the triple k, l, m,

$$\begin{aligned}(k\, l\, m) = &(2\,3\,5), (3\,4\,6), (4\,5\,7), (5\,6\,8), \\ &(6\,7\,2), (7\,8\,3), (8\,2\,4). \end{aligned} \tag{42}$$

Another definition of $\mathcal{F}$ will be found later.

3. OCTONIONS AND JORDAN ALGEBRA M_3^8

A. Octonions

Octonions (also called Cayley's numbers and quasi-quaternions) are based on eight basic units $e_1, e_2, \cdots, e_8$ with the properties

$$e_1^2 = e_1, \quad e_k^2 = -e_1, \quad e_1 e_k = +e_k e_1 \qquad (k = 2, 3, \cdots, 8), \tag{43}$$

$$e_k e_l = -e_l e_k \quad (k, l = 2, 3, \cdots, 8;\ k \neq l), \tag{44}$$

$$\begin{aligned}&e_2 e_3 = e_5, \quad e_3 e_4 = e_6, \quad e_4 e_5 = e_7, \quad e_5 e_6 = e_8, \\ &e_6 e_7 = e_2, \quad e_7 e_8 = e_3, \quad e_8 e_2 = e_4. \end{aligned} \tag{45}$$

In terms of the basic units, an octonion $\mathbf{A}$ can be expressed as follows:

$$\mathbf{A} = a_1 e_1 + a_2 e_2 + a_3 e_3 + \cdots + a_8 e_8. \tag{46}$$

The unit e_1 is not always written out explicitly (like the unit "1" in the ordinary algebra of complex numbers): sometimes we write simply α instead of αe_1.

Octonion multiplication is noncommutative and nonassociative. The associator of three octonions $\mathbf{A}, \mathbf{B}, \mathbf{C}$

$$\{\mathbf{A}, \mathbf{B}, \mathbf{C}\} = \mathbf{A}(\mathbf{BC}) - (\mathbf{AB})\mathbf{C} \tag{47}$$

is antisymmetric with respect to the permutations of $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$, as can be easily verified from definitions (43)–(45). The octonion

$$\bar{\mathbf{A}} = a_1 e_1 - a_2 e_2 - a_3 e_3 - \cdots - a_8 e_8 \tag{48}$$

is the conjugate of the octonion $\mathbf{A}$. The quantity

$$n(\mathbf{A}) = \mathbf{A}\bar{\mathbf{A}} = \bar{\mathbf{A}}\mathbf{A} = a_1^2 + a_2^2 + a_3^2 + \cdots + a_8^2 \tag{49}$$

is the "norm" of the octonion $\mathbf{A}$. It can be easily verified that

$$n(\mathbf{AB}) = n(\mathbf{BA}) = n(\mathbf{A}) n(\mathbf{B}), \tag{50}$$

a property shared only by real numbers, complex numbers, quaternions, and octonions.

Let us define as "inner" product of two octonions $\mathbf{A}$ and $\mathbf{B}$ the quantity

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A} = \tfrac{1}{2}(\bar{\mathbf{A}}\mathbf{B} + \bar{\mathbf{B}}\mathbf{A}). \tag{51}$$

The inner product is nothing but the e_1 component of the ordinary product of $\mathbf{A}$ and $\mathbf{B}$. In terms of it, the norm of $\mathbf{A}$ can also be defined as

$$n(\mathbf{A}) = \mathbf{A} \cdot \mathbf{A}. \tag{52}$$

Following are a few easily proved identities, very convenient in handling octonion expressions.

$$\mathbf{A}(\bar{\mathbf{A}}\mathbf{B}) = (\mathbf{A}\bar{\mathbf{A}})\mathbf{B} = n(\mathbf{A})\mathbf{B}, \tag{53}$$

$$\overline{(\mathbf{AB})} = \bar{\mathbf{B}}\bar{\mathbf{A}}, \tag{54}$$

$$\bar{\mathbf{A}} \cdot \bar{\mathbf{B}} = \mathbf{A} \cdot \mathbf{B}, \tag{55}$$

$$\bar{\mathbf{A}} \cdot (\mathbf{BC}) = (\mathbf{AB}) \cdot \bar{\mathbf{C}}, \tag{56}$$

$$\bar{\mathbf{A}} \cdot (\mathbf{BC}) = \bar{\mathbf{B}} \cdot (\mathbf{CA}) = \bar{\mathbf{C}} \cdot (\mathbf{AB}), \tag{57}$$

$$\mathbf{A} \cdot (\mathbf{BC}) = \mathbf{B} \cdot (\mathbf{A}\bar{\mathbf{C}}). \tag{58}$$

Writing the vector $\mathbf{X}$, the spinor of the first kind $\boldsymbol{\varphi}$ and of the second kind $\boldsymbol{\psi}$, as octonions

$$\begin{aligned}\mathbf{X} &= x_1 e_1 + x_2 e_2 + x_3 e_3 + \cdots + x_8 e_8, \\ \boldsymbol{\varphi} &= \varphi_1 e_1 + \varphi_2 e_2 + \varphi_3 e_3 + \cdots + \varphi_8 e_8, \\ \boldsymbol{\psi} &= \psi_1 e_1 + \psi_2 e_2 + \psi_3 e_3 + \cdots + \psi_8 e_8, \end{aligned} \tag{59}$$

[4] See Refs. 2 and 3.

[5] C. R. Hagen and A. J. Macfarlane, Phys. Rev. 135, B342 (1964).

one can write the invariants found in the preceding section as follows:

$$\mathbf{F} = n(\mathbf{X}), \quad \mathbf{\Phi} = n(\boldsymbol{\varphi}), \quad \mathbf{\Psi} = n(\boldsymbol{\psi}), \tag{60}$$

$$-\mathcal{F} = \bar{\boldsymbol{\varphi}} \cdot (\mathbf{X}\boldsymbol{\psi}), \tag{61}$$

B. Jordan Algebras M_3^8

The set of all "octonion Hermitian" matrices $\mathbf{M}$ of the type

$$\mathbf{M} = \begin{pmatrix} p & \mathbf{A} & \bar{\mathbf{B}} \\ \bar{\mathbf{A}} & q & \mathbf{C} \\ \mathbf{B} & \bar{\mathbf{C}} & r \end{pmatrix}, \tag{62}$$

where p, q, and r are real numbers, $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are octonions ($\bar{\mathbf{A}}$, $\bar{\mathbf{B}}$, $\bar{\mathbf{C}}$ being their conjugate octonions), form an exceptional Jordan algebra. Multiplication of two elements of this algebra is defined as

$$\mathbf{MN} = \mathbf{NM} = \tfrac{1}{2}(\mathbf{M} \times \mathbf{N} + \mathbf{N} \times \mathbf{M}), \tag{63}$$

where $\mathbf{M} \times \mathbf{N}$ is the usual row-by-column product of matrices. Multiplication is therefore commutative, although not associative.

Any element $\mathbf{M}$ of this algebra satisfies a characteristic equation of the third order,

$$\mathbf{M}^3 + \alpha\mathbf{M}^2 + \beta\mathbf{M} + \gamma = 0, \tag{64}$$

where

$$\begin{aligned} \alpha &= -(p + q + r), \\ \beta &= pq + qr + pr - n(\mathbf{A}) - n(\mathbf{B}) - n(\mathbf{C}), \\ \gamma &= -pqr + pn(\mathbf{C}) + qn(\mathbf{B}) + rn(\mathbf{A}) - 2\bar{\mathbf{B}} \cdot (\mathbf{AC}). \end{aligned} \tag{65}$$

The roots of the characteristic equation (64)—considered as an algebraic equation—are the eigenvalues of $\mathbf{M}$, like in the ordinary matrix theory. Two octonion Hermitian matrices $\mathbf{M}$ and $\mathbf{M}'$ with the same eigenvalues are considered to be essentially the same matrix in different representations.

In order for $\mathbf{M}$ and $\mathbf{M}'$ to have the same characteristic equation, the octonions $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ in (62) must be transformed in such a way that α, β, and γ in (65) remain invariant. This can be obtained by means of an eight-dimensional rotation in octonion space that leaves $n(\mathbf{A})$, $n(\mathbf{B})$, $n(\mathbf{C})$, and $\bar{\mathbf{B}} \cdot (\mathbf{AC})$ invariant. The latter quantities are but the three quadratic and the trilinear invariants (60) and (61). It follows that the octonions $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ should not be transformed in the same way under an eight-dimensional rotation; one octonion should be transformed like an eight-dimensional vector and the other two like eight-dimensional spinors of the first and second kind, respectively. Which octonion is the vector and which the spinors of each kind is only a matter of convention. The principle of triality for Jordan algebra M_3^8 is but a statement on this possibility.

In ordinary matrix theory, when a matrix $\mathbf{M}'$ is equivalent to a matrix $\mathbf{M}$ (i.e., $\mathbf{M}'$ and $\mathbf{M}$ have the same eigenvalues) a relation

$$\mathbf{M}' = \mathbf{UM}\tilde{\mathbf{U}} \tag{66}$$

exists, with $\mathbf{U}$ a unitary matrix. The ordinary representation theory is essentially a study of transformations $\mathbf{U}$. When $\mathbf{U}$ differs only by an infinitesimal quantity from unity

$$\mathbf{U} = 1 + \tfrac{1}{2}\theta\mathbf{R}, \tag{67}$$

θ small, $\mathbf{R}$ anti-Hermitian ($\tilde{\mathbf{R}} = -\mathbf{R}$), Eq. (66) becomes

$$\mathbf{M}' = \mathbf{M} + \tfrac{1}{2}\theta[\mathbf{R}, \mathbf{M}]. \tag{68}$$

The commutator

$$[\mathbf{R}, \mathbf{M}] = \mathbf{RM} - \mathbf{MR} \tag{69}$$

then defines the infinitesimal transformations on which the whole (Lie's) theory of representations is based.

Can we define something analogous to Eq. (68) for the commutative, nonassociative algebra M_3^8? The answer is yes, with the associator taking the role of the commutator that is zero in a commutative algebra. The associator $\{\mathbf{L}, \mathbf{M}, \mathbf{N}\}$ of three elements $\mathbf{L}$, $\mathbf{M}$, $\mathbf{N}$ is defined as follows:

$$\{\mathbf{L}, \mathbf{M}, \mathbf{N}\} = \mathbf{L}(\mathbf{MN}) - (\mathbf{LM})\mathbf{N}. \tag{70}$$

That the associator of a Jordan algebra is indeed the analog of the commutator of the ordinary matrix algebra can be easily seen in the following special case.

Transform an ordinary matrix algebra into a Jordan algebra by defining a symmetrized product[6]

$$\mathbf{AB} = \tfrac{1}{2}(\mathbf{A} \times \mathbf{B} + \mathbf{B} \times \mathbf{A}), \tag{71}$$

$\mathbf{A}$ and $\mathbf{B}$ being ordinary matrix and the "cross" product indicating the ordinary matrix product. Then the associator is related to the commutators as follows:

$$\{\mathbf{A}, \mathbf{B}, \mathbf{C}\} = \mathbf{A}(\mathbf{BC}) - (\mathbf{AB})\mathbf{C} = \tfrac{1}{4}\{[\mathbf{A}, \mathbf{C}], \mathbf{B}\}. \tag{72}$$

Going back to the algebra M_3^8, let $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ in (62) be the octonions that transform like the vector $\mathbf{X}$, the spinor $\boldsymbol{\varphi}$, and the spinor $\boldsymbol{\psi}$, respectively. Then

[6] These are the original considerations that led Jordan to his algebra; see P. Jordan, Z. Physik **80**, 285 (1933); P. Jordan, J. Von Neumann, and E. P. Wigner, Ann. Math. **35**, 29 (1934); A. A. Albert, *Studies in Modern Algebra* (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1963), Vol. 2; A. Gamba, *High Energy Physics and Elementary Particles* (International Atomic Energy Agency, Vienna, 1965), p. 641.

the eight-dimensional rotations (28) can be written as follows:

$$\mathbf{M}' = \mathbf{M} + 0\{\mathbf{A}_k, \mathbf{M}, \mathbf{A}_l\}, \tag{73}$$

where

$$\mathbf{A}_1 = \begin{pmatrix} 0 & +e_1 & 0 \\ +e_1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$
$$\mathbf{A}_k = \begin{pmatrix} 0 & +e_k & 0 \\ -e_k & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad k = 2, 3, \cdots, 8, \tag{74}$$

as can be easily found by direct calculation.

One could wonder why the transformation matrix $\mathbf{R}$ in (68) is anti-Hermitian, whereas $\mathbf{A}_k, \mathbf{A}_l$ in (73), (74) are (octonion) Hermitian. The difference is easily explained: in ordinary matrix calculus [see formula (72)], $\mathbf{R}$ is the anti-Hermitian commutator $[A_k, A_l]$ of two Hermitian operator A_k, A_l in strict analogy with our result in M_3^8. By the way, this same consideration also explains the missing factor $\frac{1}{2}$ in (73) as compared with formula (68). A redefinition of the associator (70) could have put the formulas in more symmetrical form; however, we did not want to alter the standard definition of associator to be found in the literature only for the sake of an irrelevant factor $\frac{1}{2}$.

C. The Group G_2

We have seen that in general the octonions $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ in (62)—or equivalently the 8-dimensional vector and spinors of first and second kind—transform differently under an 8-dimensional rotation. Are they special rotations under which the above three quantities transform in the same way? The answer is yes. Consider any three units e_k of one of the triples (45). To be specific, let us choose as example the following:

$$e_4, e_5, e_7. \tag{75}$$

Consider now the three units

$$e_6 = e_3e_4, \quad e_2 = e_3e_5, \quad e_8 = e_3e_7. \tag{76}$$

With regard to (75) and (76), we can then establish the correspondence

$$4 \leftrightarrow 6, \quad 5 \leftrightarrow 2, \quad 7 \leftrightarrow 8. \tag{77}$$

We can then prove the following theorem.

Theorem: A rotation in the space (4, 5, 7) followed by an equal rotation in the space (6, 2, 8)—equality being defined by means of (77)—transforms eight-dimensional vectors and spinors in the same way.

This can easily be seen by inspection of formulas like (36′)–(38′) or by working out explicitly the corresponding associators in (73). Since there are seven triples in (45), with the above process we obtain a total of 21 rotations. Indicating the rotation in the plane (kl), followed by an equal rotation in the plane (mn), by $(kl)(mn)$, they are the following:

(34)(58)	(36)(57)	(46)(87)
(45)(62)	(47)(68)	(57)(28)
(56)(73)	(58)(72)	(68)(32)
(67)(84)	(62)(83)	(72)(43) .
(78)(25)	(73)(24)	(83)(54) ,
(82)(36)	(84)(35)	(24)(65)
(23)(47)	(25)(46)	(35)(76)

(78)

Only 14 of the above rotations are linearly independent, since the rotations in the first column of (78) are related to those of the other two columns as indicated. These 14 rotations form a group, that by inspection can be seen to be the group G_2, the lowest of the five exceptional groups of classical group theory. As a matter of fact, one could define G_2 as the subgroup of rotations that transform eight-dimensional vectors and spinors in the same way. Since the index 1 does not appear in (78), one sees that G_2 is also a subgroup of rotations in seven dimensions, a well-known result.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Topology in General Relativity

Robert P. GEROCH*
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey

(Received 19 July 1966)

A number of theorems and definitions which are useful in the global analysis of relativistic world models are presented. It is shown in particular that, under certain conditions, changes in the topology of spacelike sections can occur if and only if the model is acausal. Two new covering manifolds, embodying certain properties of the universal covering manifold, are defined, and their application to general relativity is discussed.

INTRODUCTION

A NUMBER of theorems are proved which are useful in the analysis of general relativistic world models, particularly with regard to the problem of singularities. These are called "topological" results because they do not involve the Einstein equations directly, but rather deal with the global properties of a Lorentz signature metric on a 4-manifold. We state here the various theorems and definitions in detail, but content ourselves with a survey of possible applications.

In Sec. I, we discuss some implications of the requirement that causality be maintained, i.e., that a continuous choice of the forward light cone can be made and that no closed timelike curves exist. In Sec. II, we define two new types of covering manifolds, these adapted to the Lorentz signature metric. The applications of these and of the conventional universal covering manifold to general relativity are discussed briefly.

It is convenient to define a geometry as a 4-manifold carrying a metric of signature $(-, +, +, +)$. Modifiers will then refer to the appropriate structure. (For example, a compact geometry means the manifold is compact.) Following conventional practice, a manifold is always assumed without boundary[1] unless otherwise explicitly stated.

I. CAUSALITY

We observe[2] in our local region of space–time that from any one event P we are able to influence only those events which lie within a single cone from P, the forward light cone. Suppose we assume that this represents a universal property of space–time. Then this assumption places a global restriction on those geometries which are of physical interest. This requirement of causality may be stated mathematically: (a) no closed timelike or null curves exist, and (b) a continuous choice of the forward light cone can be made. [Following the terminology of Calabi and Markus,[3] a geometry satisfying property (b) is called isochronous.] Here, we relate these two mathematical restrictions to other, more useful, properties of the geometry.

We first require a result due to Misner.[4]

Theorem 1: Let S and S' be two compact 3-manifolds. Then there exists a compact geometry M whose boundary is the disjoint union of S and S', and in which S and S' are both spacelike.

Proof: That there exists, for any compact S and S', a compact 4-manifold M whose boundary is the disjoint union of S and S' was first proved by Rohklin[5] using cobordism theory.[6] The problem of placing a Lorentz signature metric on M is equivalent[7] to that of placing a continuous vector field on M, nowhere zero and nowhere tangent to S or S'. The Poincaré–Hopf theorem[8] states that a necessary and sufficient condition that such a vector field can be found is that the Euler characteristic,[9] χ, of M vanish.

The construction reduces, therefore, to suitably modifying M so that its Euler characteristic vanishes. Let V be some other compact 4-manifold (without boundary). Suppose we remove a small 4-ball from

* National Science Foundation Predoctoral Fellow (1966).

[1] See, for example, J. R. Munkres, *Elementary Differential Topology* (Princeton University Press, Princeton, New Jersey, 1963).

[2] See, for example, H. Reichenbach, *The Direction of Time* (University of California Press, Berkeley, California, 1956); H. Reichenbach, *The Philosophy of Space and Time* (Dover Publications, Inc., New York, 1958), Sec. 21.

[3] E. Calabi and L. Markus, Ann. Math. **75**, 63 (1962).

[4] The statement and an outline of the proof of this theorem were discovered by C. W. Misner. An important step in the argument was provided by L. Markus. *Added in proof:* Essentially the same result has been obtained by B. L. Reinhart, Topology, **2**, 173 (1963).

[5] V. A. Rohklin, Dokl. Akad. Nauk USSR **81**, 355 (1951).

[6] See, for example, C. T. C. Wall, Ann. Math. **72**, 292 (1960); R. Thom, Commun. Math. Helv. **28**, 17 (1954).

[7] See, for example, N. Steenrod, *The Topology of Fibre Bundles* (Princeton University Press, Princeton, New Jersey, 1951), p. 207.

[8] H. Hopf, Math. Ann. **96**, 225 (1926). See also J. Milnor, *Topology from the Differential Viewpoint* (University Press of Virginia, Charlottesville, Virginia, 1965), p. 35.

[9] See P. J. Hilton and S. Wylie, *Homology Theory* (Cambridge University Press, Cambridge, England, 1960), p. 167.

V and from the interior of M, and identify the boundaries (3-spheres) thus created. The result is a new compact 4-manifold, M', whose boundary is the same as that of M, but whose Euler characteristic differs from that of M by[10]

$$\chi(M') - \chi(M) = \chi(V) - 2.$$

Now, the 4-torus,[11] $S^1 \times S^1 \times S^1 \times S^1$, has an Euler characteristic of zero. Complex projective 2-space, a real 4-manifold, has an Euler characteristic of three. Thus, if V is chosen to be the 4-torus in the above construction, we decrease the Euler characteristic of M by two, and if V is complex projective 2-space, we increase χ by one. Therefore, a sequence of such operations can be found which reduces the Euler characteristic of M to zero.

Theorem 1 states that, physically, whatever the topology of a compact spacelike section of our space-time (assuming one exists) at the present epoch, the mere presence of a Lorentz signature metric does not preclude any given compact spacelike section's appearing at a later epoch. Theorem 2 below shows that, if such "changes in the topology" are to occur, then causality cannot be maintained.

Theorem 2: Let M be a compact geometry whose boundary is the disjoint union of two compact spacelike 3-manifolds, S and S'. Suppose M is isochronous, and has no closed timelike curve. Then S and S' are diffeomorphic, and further M is topologically $S \times [0, 1]$.

Proof: Since M is isochronous, we may construct[12] on M a timelike vector field, ξ^μ, which is nowhere zero and nowhere tangent to S or S'. Let γ be a curve in M, beginning on S, and everywhere tangent to ξ^μ. Suppose first that γ has no future endpoint. Then γ may be parametrized by a continuous variable t with range zero to infinity, $\gamma(t)$. Consider the sequence of points $P_i = \gamma(i)$, $i = 1, 2, 3, \cdots$, on the curve γ. This is an infinite sequence on the compact set M, and therefore[13] it has a limit point, P. Let N be a sufficiently small open neighborhood of P having the property that every curve tangent to ξ^μ passing through N can be distorted in N so that, while remaining timelike, it passes through P itself. Then for any positive number s, there must be a $t > s$ with $\gamma(t)$ in the neighborhood N (since P is a limit point of the P_i), and a $t' > s$ with $\gamma(t')$ not in N (since γ has no future endpoint). That is, γ must pass into and then out of the neighborhood N an infinite number of times. By distorting γ on two of its "passages" through N so that it intersects P, we obtain a closed timelike curve, beginning and ending at P.

Since this possibility has been excluded by hypothesis, we conclude that every such curve γ must have a future endpoint. This endpoint cannot occur in the interior of M, for ξ^μ vanishes nowhere, nor on S, since M is isochronous. Thus, the future endpoint of γ lies on S'. To summarize, through every point P of M, there exists one and only one curve γ everywhere tangent to ξ^μ, and this curve has one endpoint on S, and the other on S'.

Through each point P of M draw the curve γ. Let the curve γ be written in coordinate form, $x^\mu(t)$, where the parameter t is now defined by the equations

$$dx^\mu/dt = \xi^\mu, \quad x^\mu(t) \quad \text{lies on} \quad S.$$

Suppose t assumes the values t_1 and t_2 at P and S', respectively. Now, define a scalar field, φ, on M by its value at each point P:

$$\varphi \equiv t_1/t_2.$$

The field φ now plays the role of the Morse function.[14] The surface S is given by $\varphi = 0$, and the surface S' by $\varphi = 1$. The one parameter family of surfaces, $\varphi = \text{const}$, pass through every point of M. Finally, the congruence ξ^μ provides a one-to-one correspondence between any two surfaces of this family, Hence, S and S' are diffeomorphic, and $M = S \times [0, 1]$.

An almost identical argument suffices to demonstrate the following theorem.

Theorem 3: Every compact geometry (without boundary) has a closed timelike curve.

Note added in proof: This theorem has been proved by R. W. Bass and L. Witten, Rev. Mod. Phys., **29**, 452 (1957), and by E. H. Kronheimer and R. Penrose, in a preprint, "On the Structure of Causal Spaces."

We now discuss two applications of these results. Suppose first that it has been established that at the present epoch the universe has a compact spacelike surface topologically a 3-sphere (i.e., the closed Friedmann model is a reasonable approximation). One could imagine that, during the highly contracted phase, the inhomogeneities have grown[15] to the extent

[10] See Ref. 9, p. 93.

[11] Here and elsewhere, S^n denotes the n-sphere, and an $\times$ denotes the topological product.

[12] See Ref. 7.

[13] See, for example, E. J. McShane and T. Botts, *Real Analysis* (D. Van Nostrand Company, Inc., Princeton, New Jersey, 1959), p. 51.

[14] J. Milnor, *Morse Theory* (Princeton University Press, Princeton, New Jersey, 1963).

[15] See E. Lifschitz and I. M. Khalatnikov, Adv. Phys. **12**, 185 (1963).

that some topological "twisting" takes place. The universe than emerges from the contracted phase with expanding spacelike sections (the expanding phase of the Friedmann model). In this way, the universe might hope to "bounce" while avoiding a singularity. But according to our theorem, any such "topological twisting" can occur only if the universe is acausal.

As a second application, consider the following proposal for the construction of a quantized gravitational theory.[16] One envisions a wave functional of positive definite metrics on a 3-manifold ("3-geometries"), evaluated by a (as yet to be spelled out in full detail) Feynman sum over "kinematically possible histories" (i.e., 4-geometries of the correct signature). That is,

$$\Psi({}^3\mathfrak{G}) = \sum_{\text{4-geometries}} \exp\left[-\frac{i}{\hbar}\int R(-g)^{\frac{1}{2}}\, d^4x\right].$$

Wheeler[17] has expressed the hope that in this way there will emerge a "foam-like structure" for space, i.e., resonating states between 3-geometries of different topologies. Our two theorems state that one may always find a kinematically possible history which contributes to the Feynman sum for an arbitrary configuration of the 3-geometry, but that if any variation in the topology of that geometry (the foam-like structure) is to occur, acausal histories must be admitted into the Feynman sum. Thus, it is hard to see how the "foam structure" can be consistent with causality on the macroscopic level.

II. COVERING MANIFOLDS

The universal covering manifold has been a useful concept in the global analysis of relativistic world models.[18] Intuitively, a covering manifold is a second, larger, manifold, obtained from a given manifold, in which local properties are retained, but some of the topological features have been lost. The universal covering manifold is the "largest" covering manifold. Covering manifolds are defined in general without reference to other structures (i.e., a Lorentz metric) on the manifold in question. Our goal is to define a class of covering manifolds, these definitions utilizing the metric, in such a way that certain selected topological features are retained.

Let M be a 4-manifold. A covering manifold[19] of M is defined as a second 4-manifold, M', along with a map, φ, of M' onto M, such that for any point P of M, there exists an open neighborhood N of P such that $\varphi^{-1}(N)$ is a collection of open sets of M', each mapped diffeomorphically onto N by φ. Choose some fixed point P_0 of M, and let C be the set

$$C \equiv \{P, \gamma / P \text{ is a point of } M, \quad \gamma \text{ is a curve from } P \text{ to } P_0\}.$$

If (P, γ) and (P', γ') are elements of C, we write $(P, \gamma) \approx (P', \gamma')$ if $P = P'$, and γ and γ' are homotopic.[20] Then $\approx$ is an equivalence relation in C, and the equivalence classes define the points of the universal covering manifold[19] of M.

As an example of these two concepts, consider the 2-torus (Fig. 1). Points of the torus are labeled by the two angular coordinates θ and ϕ. Three examples of covering manifolds for the torus are shown in the figure, the last being the universal covering manifold.

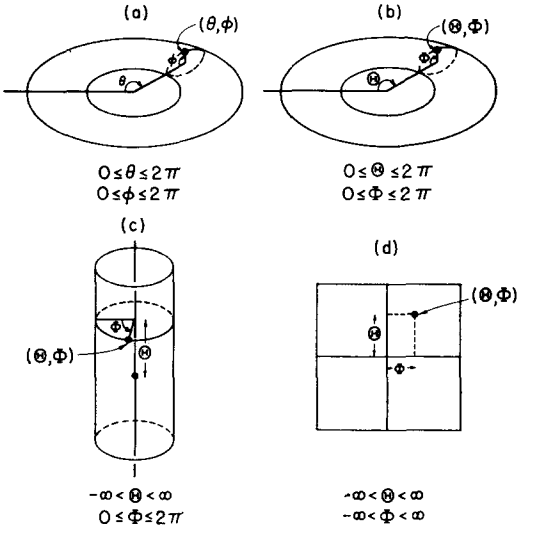


FIG. 1. An example of the torus (a), and three of its covering manifolds (b), (c), and (d). The covering manifolds are mapped onto the torus by the following equations: (b) $\theta = \Theta$, $\phi = 2\Phi$ (mod 2π); (c) $\theta = 2\Theta$ (mod 2π), $\phi = \Phi$; (d) $\theta = \Theta$ (mod 2π), $\phi = \Phi$ (mod 2π). Note that (d), the universal covering manifold, is simply connected.

From the definition, it is obvious that any tensor field residing on M determines uniquely a corresponding tensor field on any covering manifold M'. In particular, if M has a metric, then M' is assigned a unique metric.

Four elementary properties of the universal covering manifold should be mentioned.

(1) The universal covering manifolds resulting from two different choices of the initial point P_0 are identical (i.e., diffeomorphic).[19]

[16] J. A. Wheeler, Ann. Phys. (N.Y.) **2**, 604 (1957); C. W. Misner, Rev. Mod. Phys. **29**, 497 (1957); H. Leutwyler, Phys. Rev. **134**, B1155 (1964).

[17] See Wheeler, Ref. 16.

[18] See Ref. 3, also S. Hawking, Phys. Rev. Letters **15**, 689 (1965).

[19] See, for example, J. G. Hocking and G. S. Young, *Topology* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1961), p. 188.

[20] See, for example, A. H. Wallace, *Algebraic Topology* (Pergamon Press, Inc., New York, 1963), p. 63.

(2) The universal covering manifold is simply connected.[19]

(3) Every covering manifold of M is covered by the universal covering manifold of M.[19] (That is, the universal covering manifold is the "largest" covering manifold.)

(4) Let M be a geometry, and M' be its universal covering manifold. We obtain from the covering a unique metric on M'. Then M' is isochronous.[21]

Before discussing other covering manifolds, it is convenient to point out certain additional properties of the universal covering manifold in the presence of other structures on the manifold. We mention one property that the universal covering manifold does, and one property that it does not, have.

Let M be a geometry. Suppose in M we have a spacelike 3-surface S which is complete, i.e., S does not have any boundary points in M. In the universal covering manifold, M', of M, we may identify a corresponding surface S'.[22] Now, it is possible in general for some timelike curve in M to intersect S more than once, but in M' no timelike curve can intersect S' more than once. The proof is not difficult. Suppose that, for purposes of contradiction, there were in M' a timelike curve γ whose endpoints are the points P and P' of S'. We may assume without loss of generality that γ intersects S' only at P and P'. S' is two sided (i.e., a continuous choice of the forward light cone can be made on S'), since M' is isochronous. Consider a second curve $\bar{\gamma}$ which also begins at P, ends at P', and in addition coincides with γ in a finite neighborhood of P and P'. We construct $\bar{\gamma}$ as follows: $\bar{\gamma}$ begins at P, coincides with γ along a finite stretch, and thereafter remains within a neighborhood of S', staying always on the same side of S'. Then, at some point, $\bar{\gamma}$ crosses S', and again stays within a neighborhood of S', but now remaining on the other side. Finally, $\bar{\gamma}$ joins γ near P', and coincides with γ along the final stretch to P'. Now, $\bar{\gamma}$ crosses S' just once. Since M' is simply connected, γ can be continuously distorted into $\bar{\gamma}$, while always keeping fixed the finite stretches near P and P'. The impossibility of this follows from the fact that on distortion the number of intersections of γ with the complete surface S' must increase or decrease always by two, while $\bar{\gamma}$ intersects S' but once.

One might imagine that in general no closed timelike curve can occur on the universal covering manifold of a geometry M. (This is true in two dimensions.) That this is not the case in four dimensions is shown by an example. It is well known[21] that a continuous nonzero vector field can be placed on the 3-sphere, S^3. Hence, a signature $(-, +, +)$ metric can be placed thereon. Define a geometry M by

$$M \equiv S^3 \text{ (above metric)} \times (-\infty, +\infty) \text{ (usual metric)}.$$

M has a closed timelike curve (since by Theorem 3 of Sec. I, there is one on S^3). But M is simply connected, and hence is its own universal covering manifold. In particular, this universal covering manifold has a closed timelike curve.

We now try to isolate certain of the above properties of the universal covering manifold by means of suitably selected covering manifolds.

Let M be a given geometry. Choose any point P_0 of M, and define, as before,

$$C \equiv \{P, \gamma / P \text{ a point of } M, \quad \gamma \text{ is any curve from } P \text{ to } P_0\}.$$

Now, consider the equivalence relation: $(P, \gamma) \approx (P', \gamma')$ if $P = P'$, and if a timelike vector, when continuously transferred from P_0 to P along γ and back to P_0 along γ', does not reverse its time direction. The equivalence classes define a new covering manifold, which we call the Lorentz covering manifold of M. It is easy to verify that it has the following properties.

(1) The Lorentz covering manifold is isochronous.

(2) If M is isochronous, its Lorentz covering manifold is M itself.

(3) Any covering manifold of M which is isochronous covers the Lorentz covering manifold of M. (Thus, the Lorentz covering manifold is the "smallest" isochronous covering manifold.)

Next, suppose we have an isochronous geometry M, and that in M we have a complete spacelike surface S. Choose a point P_0 of M and define the set C as above. We now assign an integer, the index, to each element of C as follows. Each time γ crosses or meets S going forward in time, assign the integer $+1$, and each time going backward, assign a -1. Add the $+1$'s and -1's to obtain the index of the curve γ. Consider now the equivalence relation in C: $(P, \gamma) \approx (P', \gamma')$ if $P = P'$ and γ and γ' have the same index. The equivalence classes define the covering manifold of M relative to S. It has the properties:

(1) If S is a complete surface in the geometry M, and S' is a corresponding[22] surface in the relative covering manifold M', then no timelike curve in M' intersects S' more than once.

(2) Let M'' be another covering manifold of M having the property that no curve in M'' meets S'' exactly twice, once from each side of S''. Then M''

[21] L. Markus, Ann. Math. 62, 411 (1955).

[22] The image of S in M' may consist of several disjoint parts. Here and elsewhere, we choose any one connected part to serve as S'.

covers the relative covering manifold of M. (That is, the relative covering manifold is the "smallest" covering manifold embodying the other important property of the universal covering manifold.)

The relative covering manifold may also be characterized by the property that it is the smallest covering manifold which is divided into two disjoint parts by the surface S'.

The importance of covering manifolds in the study of singularities stems from the fact that a geometry M has a singularity[23] if and only if there is a singularity in every covering manifold of M. The technique,[24] then, is to apply the various theorems about singularities not to the geometry M, but rather to one of its covering manifolds. One is thus able to weaken the hypotheses of many of the singularity theorems.

More specifically, consider those theorems[25] which contain in their hypotheses, among other conditions, the requirements that the geometry M:

(1) be isochronous,

(2) has a complete spacelike surface S,

(3) has the property that no timelike curve intersects S more than once.

As we have seen, the universal covering manifold of a geometry M automatically satisfies conditions (1) and (3), even if M does not. There remains, therefore, the investigation of the conditions under which the other hypotheses [besides (1) and (3)] of each singularity theorem are passed from the geometry M to its universal covering manifold. This question has been investigated by Hawking[24] for the case in which S is not required to be compact.

But when[26] S is required to be compact (the "closed universe" theorems), a difficulty arises. This is the difficulty that the compactness of S is not in general passed to the universal covering manifold of M. However, if S is two-sided, its compactness is preserved under (a) the passage to the Lorentz covering manifold, and then (b) the passage to the relative covering manifold. In short, the two covering manifolds defined here can be used to weaken the hypotheses of even the theorems about singularities in closed universes. A case by case analysis of these theorems is not attempted here.

ACKNOWLEDGMENT

The author wishes to thank Professor J. A. Wheeler for many valuable discussions during the course of this work.

This research was supported in part by the United States Air Force Office of Research, Air Research and Development Command, under Contract No. AF 49-638-1545.

[23] C. W. Misner, J. Math. Phys. **4**, 924 (1963).

[24] S. Hawking, Phys. Rev. Letters **15**, 689 (1965).

[25] R. Penrose, Phys. Rev. Letters **14**, 57 (1965), S. Hawking, Proc. Roy. Soc. (London) **A294**, 511 (1966); **A295**, 490 (1966).

[26] S. Hawking, Phys. Rev. Letters **17**, 444 (1966); R. Geroch, *ibid.* **17**, 445 (1966).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Double Representations of Space Groups

M. GLÜCK, Y. GUR, AND J. ZAK
Department of Physics, Technion–Israel Institute of Technology, Haifa, Israel

(Received 26 July 1966)

It is shown that, for calculating characters of double representations of space groups, the same methods can be used as for single representations. Two methods are reviewed: the ray representation method and the induction method. Examples are presented for both methods.

I. INTRODUCTION

GROUP-THEORETICAL treatment of problems in solids requires the knowledge of irreducible representations of space groups. When the spin of the electron is not taken into account it is sufficient to know single representations of the space groups.[1] However, when problems involve spin (like spin-orbit coupling), the representations, according to which the classification of states and energy levels in crystals is carried out, are the double representations of space groups.[2]

A very simple method exists for finding irreducible representations[3] (both single and double) of symmorphic space groups.[1] According to this method, the matrix $D^{\mathbf{k}}(\{\alpha \mid \mathbf{a}\})$ of an irreducible representation for a space group element $\{\alpha \mid \mathbf{a}\}$ is given as

$$D^{\mathbf{k}}(\{\alpha \mid \mathbf{a}\}) = \exp(i\mathbf{k}\cdot\mathbf{a})D(\alpha), \tag{1}$$

where $D(\alpha)$ is the matrix of an irreducible representation of the point group of the group of $\mathbf{k}$. Relation (1) gives all the irreducible representations $D^{\mathbf{k}}(\{\alpha \mid \mathbf{a}\})$ (both single and double) of the group of $\mathbf{k}$ from all the irreducible representations $D(\alpha)$ (single and double correspondingly) of the point group of the group of $\mathbf{k}$. The same relation (1) holds also in the case of nonsymmorphic space groups for symmetry points $\mathbf{k}$ inside the Brillouin zone.[1] This simple prescription does not hold for points $\mathbf{k}$ on the surface of the Brillouin zone in nonsymmorphic space groups. Many methods have been suggested[4–8] to find the irreducible representations of single nonsymmorphic space groups for symmetry points $\mathbf{k}$ on the surface of the Brillouin zone. We describe two which seem to be very general. The one due to Zak[7,9] employs Seitz's fundamental observation about the solvability of space groups and the corresponding methods developed by Schur.[10] The other method due to Döring[5] is based on the observation that the representations of the group of the vector $\mathbf{k}$ are simply connected to ray representations of the associated point group. Döring finds the ray representations of all point groups and from there on the work is quite simple and easy. In this paper we prove that both of these methods are applicable to double representations as well. It seemed appropriate to do this since much confusion (Rudra,[11] Kitz[12]) and inaccuracy (Kovalev[8]) prevail in recent publications.

In the paper by Kitz,[12] the author finds the ray representations of double point groups and shows how to use them to construct representations of double space groups. Apart from the fact that this was done before,[13] such an approach is very confusing from the physical point of view. It is well known that a double point group is a ray representation of a single point group, and therefore a ray representation of a double point group means a ray representation of a ray representation of a single point group. Therefore, the mentioned method[12,13] does not reflect the physical nature of double groups as ray representations of single groups, and in this paper we show that there is no need for ray representations of double point groups in constructing irreducible representations of double space groups. In Kovalev's book[8] all the irreducible representations (both single and double) of space groups are listed. It turns out, however,[14] that there are many mistakes in Kovalev's tables.

In Sec. II we give a review of both Döring's[5] and

[1] G. F. Koster, in *Solid State Physics*, F. Seitz and D. Turnbull, Eds. (Academic Press, Inc., New York, 1957), Vol. 5.
[2] R. J. Elliott, Phys. Rev. **96**, 280 (1954).
[3] All through this work we have in mind representations of groups of $\mathbf{k}$ for symmetry points in the Brillouin zone (See Ref. 1).
[4] C. Herring, J. Franklin Inst. **233**, 525 (1942).
[5] W. Döring, Z. Naturforsch. **14a**, 343 (1959).
[6] T. Sugita and E. Yamaka, Nippon Telegraph and Telephone, ECL Rept. **2**, 24 (1954).
[7] J. Zak, J. Math. Phys. **1**, 165 (1960).
[8] O. V. Kovalev, *Irreducible Representations of Space Groups* (Izdatel'stuo Akademii Nauk Ukrainiskoj SSR, Kiev, 1961).
[9] J. Zak, Ph.D. thesis, Technion. Israel Institute of Technology (1960).
[10] I. Schur Sitz. Preuss. Akad. 164 (1906).
[11] P. Rudra, J. Math. Phys. **6**, 1273 (1965).
[12] A. Kitz, Phys. Stat. Sol. **8**, 813 (1965).
[13] I. B. Levinson, Tr. Akad. Nauk Lit. SSR **B4(27)**, 3 (1961).
[14] There is a book on character tables of space groups under preparation by A. Casher, M. Glück, Y. Gur, and J. Zak.

Zak's[7,9] methods and show how they are used to find irreducible representations of double space groups. In Sec. III examples are given for the groups T^5 and T_h^2.

II. METHODS FOR FINDING IRREDUCIBLE REPRESENTATIONS OF DOUBLE SPACE GROUPS

1. Use of Ray Representations

For single space groups, this method was developed by Döring.[5] Let $G_{\mathbf{k}}$, the group of the vector $\mathbf{k}$, be given by

$$G_{\mathbf{k}} = \{\epsilon \mid 0\}H + \{R_2 \mid \mathbf{a}_2\}H + \cdots + \{R_n \mid \mathbf{a}_n\}H, \tag{1a}$$

where H is the pure translation group and $\{R_i \mid \mathbf{a}_i\}$ are the "representative elements"[1,7] of $G_{\mathbf{k}}$—ϵ denotes the unit elements of the point group $O_{\mathbf{k}} = \{\epsilon, R_1, \cdots, R_n\}$. The matrices representing the translations $\{\epsilon \mid t_n\}$ are scalar and given by $D\{\epsilon \mid \mathbf{t}_n\} = \exp(i\mathbf{k}\cdot\mathbf{t}_n)I$. The problem is to find the matrices representing the elements $\{R_i \mid \mathbf{a}_i\}$. These elements satisfy

$$\{R_i \mid \mathbf{a}_i\}\{R_j \mid \mathbf{a}_j\} = \{R_k \mid \mathbf{a}_k + \mathbf{t}_{kij}\} = \{\epsilon \mid \mathbf{t}_{kij}\}\{R_k \mid \mathbf{a}_k\}; \tag{2}$$

hence,

$$D\{R_i \mid \mathbf{a}_i\}D\{R_j \mid \mathbf{a}_j\} = D\{\epsilon \mid \mathbf{t}_{kij}\}D\{R_k \mid \mathbf{a}_k\} = \exp(i\mathbf{k}\cdot\mathbf{t}_{kij})D\{R_k \mid \mathbf{a}_k\}. \tag{3}$$

There is a one-to-one correspondence between the representative elements $\{R_i \mid \mathbf{a}_i\}$ and their point parts R_i. We can therefore write (3) in the abbreviated form

$$D(R_i)D(R_j) = \exp(i\mathbf{k}\cdot\mathbf{t}_{kij})D(R_k), \tag{3a}$$

where

$$R_k = R_iR_j \tag{4}$$

and

$$\mathbf{t}_{kij} = R_i\mathbf{a}_j + \mathbf{a}_i - \mathbf{a}_k. \tag{5}$$

The matrices $D\{R_i \mid \mathbf{a}_i\} = D(R_i)$ therefore constitute a ray[5] representation of O_k with phases ω_{R_i,R_j} given by

$$\omega_{R_i,R_j} = \exp(i\mathbf{k}\cdot t_{kij}). \tag{6}$$

As is well known (Ref. 15, Chap. 12), a ray representation is defined by equations of the form

$$D^{(\lambda,\mu)}(R_i)D^{(\lambda,\mu)}(R_j) = \omega^{\lambda}_{R_i,R_j}D^{(\lambda,\mu)}(R_i\cdot R_j), \tag{7}$$

where λ denotes the factor system and μ the corresponding ray representations. The possible $\omega^{(\lambda)}$'s are restricted by the following conditions:

$$\omega^{(\lambda)}_{R_i,R_j}\omega^{(\lambda)}_{R_iR_j,R_k} = \omega^{(\lambda)}_{R_i,R_jR_k}\omega^{(\lambda)}_{R_j,R_k}. \tag{8}$$

It is obvious that the ω's defined by (6) obey these restrictions. Hence they belong to *some* factor system λ. To find this λ one must find a set of coefficients C_{R_i} satisfying

$$\omega^{(\lambda)}_{R_i,R_j} = \omega_{R_i,R_j}C_{R_i}C_{R_j}/C_{R_iR_j}. \tag{9}$$

In this set the ω_{R_i,R_j} are given by Eq. (6) and the $\omega^{(\lambda)}_{R_i,R_j}$ in tables giving the ray representations of the point groups.[5] These equations may be solved for only one definite λ (they are n^2 equations for only n C's), which must be found by a (finite) number of trials. Having found λ and the C's, one can now write down the matrices $D\{R_i \mid a_i\} \equiv D(R_i)$;

$$D^{(\mu)}\{R_i \mid \mathbf{a}_i\} = (C_{R_i})^{-1}D^{(\lambda,\mu)}(R_i) \tag{10}$$

or the characters $\chi\{R_i \mid a_i\} \equiv \chi(R_i)$,

$$\chi^{(\mu)}\{R_i \mid \mathbf{a}_i\} = (C_{R_i})^{-1}\chi^{(\lambda,\mu)}(R_i). \tag{11}$$

Reliable tables of $\chi^{(\lambda,\mu)}(R_i)$ may be found in Ref. 5. In physical application, one is often interested only in characters. However, when the matrices themselves are desired one should consult a table[16] of $D^{(\lambda,\mu)}$'s.

Let us now treat double representations. The method[17] is the same as before, but Eqs. (3) and (3a) should be replaced by

$$D(R_i)D(R_j) = \pm\exp(i\mathbf{k}\cdot\mathbf{t}_{kij})D(R_k). \tag{12}$$

The sign is determined by

$$U(R_i)U(R_j) = \pm U(R_k). \tag{13}$$

Here

$$U(R_i) = \cos\tfrac{1}{2}\theta_i - \sin(\tfrac{1}{2}\theta_i)\boldsymbol{\sigma}\cdot\mathbf{n}_i, \tag{14}$$

where $\mathbf{n}_i$ is a standard unit vector along the axis of rotation of R_i, θ_i is the corresponding angle of rotation (counterclockwise), $0 \leq \theta_i \leq 2\pi$, and $\boldsymbol{\sigma}$ stands for the Pauli matrices,

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

It should be noted that each point element of $O_{\mathbf{k}}$ is either a pure rotation or consists of a rotation followed by inversion. As inversion does *not* affect the spin degrees of freedom (Ref. 18, p. 238), our prescription for the signs [Eqs. (13) and (18)] depends only on the rotational part. The rest of the analysis is exactly as before. The method of calculation is also exactly as before, and we demonstrate it in Sec. III.

2. Induction Method

An alternative method for constructing irreducible representations of nonsymmorphic space groups was

[15] M. Hamermesh, *Group Theory and Its Application to Physical Problems* (Addison-Wesley Publishing Company, Inc., New York, 1962).

[16] A. C. Hurley, Phil. Trans. Roy. Soc. London **A260**, 1 (1966).

[17] M. Glück, M.Sc. thesis, Israel Institute of Technology (1963).

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

described by Zak.[7] We show here that this method is applicable to double space groups as well. To do this we give a summary of this method,[7] whose applicability to double groups is obvious. Let G_k be the group of vector $\mathbf{k}$, and assume it is nonsymmorphic. Since each space group has an invariant subgroup of index 2 or 3, this can be said also about G_k. Denote by H_k an invariant subgroup of index 2 or 3 of the group G_k. H_k can be either symmorphic or nonsymmorphic. If it is symmorphic we know how to find its irreducible representations (both single and double), and the method which we describe below enables us to construct the irreducible representations of G_k from those of H_k. If H_k is nonsymmorphic then we look for its invariant subgroup of index 2 or 3. We repeat this process until we finally come to an invariant subgroup of index 2 or 3 which is symmorphic. We know how to find its representations, and then step by step, using the method described below, we find the representations of G_k. Assume now that H_k is a symmorphic invariant subgroup of G_k of index 2 or 3, and let $\gamma_1, \gamma_2, \cdots, \gamma_l$ be all the irreducible representations of H_k (both single and double). G_k can be divided into cosets (1a):

$$G_k = \sum_i \{R_i \mid \mathbf{a}_i\} H_k,$$

where the summation is a direct one and i takes the values 0, 1 or 0, 1, 2, depending on whether H_k is of index 2 or 3. Take a representation γ of H_k, where a matrix $D(\{R_h \mid \mathbf{a}_h\})$ corresponds to the element $\{R_h \mid \mathbf{a}_h\}$, and define a conjugate representation $\bar{\gamma}$, where the matrix $D(\{R_i \mid \mathbf{a}_i\}^{-1}\{R_h \mid \mathbf{a}_h\}\{R_i \mid \mathbf{a}_i\})$ corresponds to the element $\{R_h \mid \mathbf{a}_h\}$. All the representations of H_k can be divided into a set of pairs (or triads) of conjugate nonequivalent representations, and a set of self-conjugate representations[7] (γ is called self-conjugate if γ and $\bar{\gamma}$ are equivalent). Each set of pairs (or triads) of conjugate representation of H_k induces one irreducible representation of G_k. Each self-conjugate representation of H_k induces 2 or 3 (depending whether H_k is invariant of index 2 or 3) irreducible representations of the group G_k. This procedure is true for any finite group containing a subgroup of index 2 or 3, and it is therefore true for both single and double space groups. We can therefore use the formulas (7), (13), (14), and (15) of Ref. 7 to derive characters of irreducible representations of double space groups.

It is worth mentioning that formulas (13) and (15) of Ref. 7 hold when it is possible to associate with an element of G, which does not belong to H, a scalar matrix. Otherwise one has to find the matrix that corresponds to the element of G (that does not belong to H) and use the multiplication rules of the group G.[17] This case is usually more complicated, but it is just a matter of simple algebra to find the characters of the representations of G. We demonstrate it by an example in Sec. III.

III. EXAMPLES

In this section we calculate characters of representations of double groups T^5 and T_h^2. For the first of these, we calculate the characters of the representations for the point $H[0(2\pi/a)0]$ by using ray representations. For group T_h^2 we calculate characters of the representations at the point $X[0(\pi/a)0]$ by the method of induction.

Let us start with point H for the group T^5. The symmetry element for this point is

$$(e \mid 0\,0\,0), \quad (C_3^{xyz} \mid 0\,0\,0), \quad (C_3^{2xyz} \mid 0\,0\,0),$$
$$\left(C_3^{\bar{x}yz} \,\middle|\, 0\,\frac{a}{2}\,\frac{a}{2}\right), \quad \left(C_3^{x\bar{y}z} \,\middle|\, \frac{a}{2}\,0\,\frac{a}{2}\right), \quad \left(C_3^{xy\bar{z}} \,\middle|\, \frac{a}{2}\,\frac{a}{2}\,0\right),$$
$$\left(C_3^{2\bar{x}yz} \,\middle|\, \frac{a}{2}\,0\,\frac{a}{2}\right), \quad \left(C_3^{2x\bar{y}z} \,\middle|\, \frac{a}{2}\,\frac{a}{2}\,0\right), \quad \left(C_3^{2xy\bar{z}} \,\middle|\, \frac{a}{2}\,\frac{a}{2}\,0\right),$$
$$\left(C_2^{x} \,\middle|\, \frac{a}{2}\,\frac{a}{2}\,0\right), \quad \left(C_2^{y} \,\middle|\, 0\,\frac{a}{2}\,\frac{a}{2}\right), \quad \left(C_2^{z} \,\middle|\, \frac{a}{2}\,0\,\frac{a}{2}\right).$$

Let us choose as generating elements:

$$a = \left(C_2^z \,\middle|\, \frac{a}{2}\,0\,\frac{a}{2}\right), \quad b = \left(C_2^x \,\middle|\, \frac{a}{2}\,\frac{a}{2}\,0\right),$$
$$c = (C_3^{xyz} \mid 0\,0\,0).$$

They satisfy the following relations:

$$a^2 = (e \mid 0\,0\,a),$$
$$b^2 = (e \mid a\,0\,0),$$
$$c^3 = (e \mid 0\,0\,0),$$
$$ba = (e \mid a\,a\,-a)ab,$$
$$ca = bc,$$
$$abc = (e \mid 0\,-a\,0)cb.$$

The matrices A, B, C corresponding to the generating elements satisfy the relations

$$A^2 = -E, \quad B^2 = -E, \quad C^3 = -E,$$
$$BA = -AB, \quad CA = BC, \quad ABC = CB.$$

By comparing with Döring's tables[5] for the point group T we find that $\alpha = -1$. We therefore get the following table for the representations of the double

group T^5 at the point $H[0(2\pi/a)0]$.

	e	$3U$	$4C_3$	$4C_3^2$
	1	1	1	1
	1	1	ω	ω^2
	1	1	ω^2	ω
	3	-1	0	0
$\omega = \exp \frac{2}{3}\pi i$				
	2	0	1	-1
	2	0	ω	$-\omega^2$
	2	0	ω^2	$-\omega$

Let us now calculate the characters for point $X[0(\pi/a)0]$ of the group T_h^2 by the induction method.[7] The symmetry elements are

$$e \qquad C_2^x \qquad C_2^y \qquad C_2^z$$

$$\left(I \left| \frac{a}{2}\frac{a}{2}\frac{a}{2}\right.\right) \quad \left(\sigma^x \left| \frac{a}{2}\frac{a}{2}\frac{a}{2}\right.\right) \quad \left(\sigma^y \left| \frac{a}{2}\frac{a}{2}\frac{a}{2}\right.\right) \quad \left(\sigma^z \left| \frac{a}{2}\frac{a}{2}\frac{a}{2}\right.\right).$$

We can write

$$G_X = H + \left(I \left| \frac{a}{2}\frac{a}{2}\frac{a}{2}\right.\right)H,$$

where H consists of the first four elements. We know the representations of H. There is only one two-dimensional representation of the double group (in addition to the single representations).

e	C_2^x	C_2^y	C_2^z
1	1	1	1
1	-1	1	-1
1	1	-1	-1
1	-1	-1	1
2	0	0	0

The one-dimensional representations are single ones. To the elements

$$\bar{e}, \quad \overline{C_2^x}, \quad \overline{C_2^y}, \quad \overline{C_2^z}$$

matrices with a minus sign correspond in the two-dimensional representation. To find the representations of G_X, we construct conjugate (or self-conjugate) representations by means of the elements

$$\left(I \left| \frac{a}{2}\frac{a}{2}\frac{a}{2}\right.\right):$$

$$\left(I \left| \frac{a}{2}\frac{a}{2}\frac{a}{2}\right.\right)^{-1} C_2^x \left(I \left| \frac{a}{2}\frac{a}{2}\frac{a}{2}\right.\right) = (C_2^x \mid 0\ a\ a),$$

$$\left(I \left| \frac{a}{2}\frac{a}{2}\frac{a}{2}\right.\right)^{-1} C_2^y \left(I \left| \frac{a}{2}\frac{a}{2}\frac{a}{2}\right.\right) = (C_2^y \mid a\ 0\ a),$$

$$\left(I \left| \frac{a}{2}\frac{a}{2}\frac{a}{2}\right.\right)^{-1} C_2^z \left(I \left| \frac{a}{2}\frac{a}{2}\frac{a}{2}\right.\right) = (C_2^z \mid a\ a\ 0). \quad (15)$$

As is easily seen, the first and second, and also the third and fourth representations are conjugate in pairs, and the two-dimensional representation is self-conjugate. The matrix that corresponds to

$$\left(I \left| \frac{a}{2}\frac{a}{2}\frac{a}{2}\right.\right)$$

in the two-dimensional representation is [as follows from (15)]

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \quad (16)$$

The characters of the representations of T_h^2 for point $X[0(\pi/a)0]$ are therefore the following.

e	C_2^y	σ^y	C_2^x	σ^x	C_2^z	σ^z	I
2	2						
2	-2			0			
2	0	$2i$					
2	0	$-2i$		0			

The last two representations are additional ones for the double group.

Note added in Proof: The basic idea behind our Sec. II is also implicitly stated in a paper by M. V. Murphy, J. Math. Phys. 7, 853 (1966), Sec. 5.

Expansion Theorem for the Linearized Fokker–Planck Equation

JORDAN D. LEWIS*
Department of Nuclear Engineering, University of Michigan, Ann Arbor, Michigan

(Received 22 December 1965)

The linearized Fokker–Planck kinetic equation for each component of a homogeneous, nondegenerate, fully ionized plasma is separated by means of a spherical harmonic expansion into an infinite set of singular integro–differential equations. Each equation is shown to generate a continuous set of eigenfunctions, for which asymptotic high-speed forms are found. By extending the theory of singular differential equations an expansion formula is developed, which is shown to be complete with respect to functions square integrable in velocity space.

INTRODUCTION

THE Fokker–Planck kinetic equation may be used to describe the temporal behavior of a homogeneous, nondegenerate, fully ionized plasma.[1–3] The Fokker–Planck equation has been derived by several authors (see e.g., the discussion in Robinson and Bernstein[3]), starting with various levels of sophistication in kinetic theory. Whatever the starting point, the results are essentially the same.[2,3] For our purposes, it is convenient to write the Fokker–Planck equation in the form[2,3]

$$\frac{\partial F_a}{\partial t} = \sum_b \frac{\partial}{\partial \mathbf{v}} \cdot \int d^3v' \left(F_b' \frac{\partial F_a}{\partial \mathbf{v}} - \frac{m_a}{m_b} F_a \frac{\partial F_b'}{\partial \mathbf{v}'} \right) \cdot \mathbf{Q}_{ab}(\mathbf{v}, \mathbf{v}'). \tag{1}$$

Here $F_a(\mathbf{v}, t)\, d^3v$ is the expected number of particles of kind a in the velocity space volume element d^3v about $\mathbf{v}$ at time t,

$$\mathbf{Q}_{ab} = \Gamma_{ab} g^{-3} (\mathbf{I} - \mathbf{gg}),$$

where Γ_{ab} is a positive constant, $\mathbf{g} = \mathbf{v} - \mathbf{v}'$, and $\mathbf{I}$ is the unit dyadic. We denote by F_j' the function $F_j(\mathbf{v}', t)$.

It is not difficult to show[2] that $\partial F_a/\partial t = 0$ if and only if F_a and F_b are the equilibrium, or Maxwell, distributions F_{a0}, F_{b0}. In the vicinity of equilibrium we may write $F_j = F_{j0}[1 + f_j(\mathbf{v}, t)]$. Neglecting terms quadratic in f_j we then obtain from (1) the linear equation

$$F_{a0} \frac{\partial f_a}{\partial t} = \sum_b \frac{\partial}{\partial \mathbf{v}} \cdot \int d^3v' F_{b0}' F_{a0} \left(\frac{\partial f_a}{\partial \mathbf{v}} - \frac{m_a}{m_b} \frac{\partial f_b'}{\partial \mathbf{v}'} \right) \cdot \mathbf{Q}_{ab}. \tag{2}$$

We refer to the quantity $F_{j0} f_j$ as the perturbation from equilibrium.

Equation (2) as it stands is in fact a pair of coupled equations for f_a and f_b. Due to the quite small value of the electron–ion mass ratio, the equations however are only very weakly coupled. Thus, for example, the effect of the ion perturbation on the electron perturbation is small when compared with the effect of the ions and electrons in the unperturbed equilibrium distributions.

In the following we consider the equation for the ions. The treatment of the electron equation is quite similar, and the modifications necessary for this case are indicated later (Sec. V). In the approximation $m_i \gg m_e$ it can be shown[3] that the ions act like a single component gas. In this case Eq. (2) becomes, for the ions,

$$F_{i0} \frac{\partial f_i}{\partial t} = \frac{\partial}{\partial \mathbf{v}} \cdot \int d^3v' F_{i0}' F_{i0} \left(\frac{\partial f_i}{\partial \mathbf{v}} - \frac{\partial f_i'}{\partial \mathbf{v}'} \right) \cdot \mathbf{Q}_{ii}. \tag{3}$$

Our purpose here is to develop an expansion theorem based on (3). We follow the standard method of assuming solutions to (3) of the form $f(\mathbf{v}, t) = g_\lambda(\mathbf{v}) \exp(-\lambda t)$. This reduces (3) to the form $L(\mathbf{v}) g_\lambda(\mathbf{v}) = -\lambda g_\lambda(\mathbf{v})$, where L is a three-dimensional integro–differential operator. In Sec. I we show that (3) requires $\mathrm{Re}\,\lambda > 0$ and $\mathrm{Im}\,\lambda = 0$ as we would expect physically. In Sec. II, we introduce a spherical harmonic expansion which replaces the three-dimensional equation by an infinite set of uncoupled equations, $L_{lm} g_{lm} = -\lambda_{lm} g_{lm}$, where $L_{lm}(v)$ is a singular integro–differential operator. These are cast into a self-adjoint form in Sec. III by introducing a suitable algebraic transformation on the functions $g_{lm}(v, \lambda_{lm})$. With boundary conditions obtained by combining (3) with the conservation laws, we proceed to find the eigenvalue spectrum which is continuous and for $l = 0, 1$ consists of all $\lambda \geq 0$ and for $l \geq 2$ consists of all $\lambda > 0$.

* Present address: The Battelle Development Corporation, Columbus, Ohio.

[1] M. Rosenbluth, W. M. MacDonald and D. L. Judd, Phys. Rev. **107**, 1 (1957).

[2] D. C. Montgomery and D. A. Tidman, *Plasma Kinetic Theory* (McGraw-Hill Book Company, Inc., New York, 1964), Chaps. 2 and 3.

[3] B. B. Robinson and I. B. Bernstein, Ann. Phys. (N.Y.) **18**, 110 (1962).

The expansion theorem is developed in Sec. IV. (Although the spectral resolution theorem implies the existence of an expansion theorem for self-adjoint operators, there remains the nontrivial task of constructing the expansion explicitly.) Since L_{lm} is singular at $v = 0$ and $v \to \infty$, we temporarily replace the interval $0 \leq v < \infty$ by the interval $0 < v_1 \leq v \leq v_2 < \infty$ and show that L_{lm} generates a complete orthonormal set on this interval. To return to the original interval and thus obtain the desired expansion theorem, we use the above completeness property together with an extension of the theory of singular differential equations. This finally yields a set of functions $\{\Psi_{lm}(\mathbf{v}, \lambda_{lm})\}$ that is complete with respect to functions square integrable in velocity space. Since the spectrum is continuous, the expansion has the form of a generalized Fourier integral. For $v \gg (3kT/m_i)^{\frac{1}{2}}$ we have found asymptotic forms of the expansion functions.

I. PROPERTIES OF THE KINETIC EQUATION

For convenience we drop the subscript "i" from F_{0i} and f_i. If f satisfies the conditions

$$\lim_{v\to 0} vf = 0, \quad \lim_{v\to\infty} F_0 \frac{\partial f}{\partial v} = 0, \tag{4}$$

it is possible to show that the linear kinetic equation (3) conserves number, momentum, and kinetic energy densities.

Introducing $f(\mathbf{v}, t) = g_\lambda(\mathbf{v}) \exp(-\lambda t)$ in (3), we find

$$-\lambda F_0 g_\lambda = \frac{\partial}{\partial \mathbf{v}} \cdot \int d^3v' F_0 F_0' \left(\frac{\partial g_\lambda}{\partial \mathbf{v}} - \frac{\partial g_\lambda'}{\partial \mathbf{v}'}\right) \cdot \mathbf{Q}. \tag{5}$$

Multiplying (5) by g_λ^*, where * denotes complex conjugate, and then integrating over $\mathbf{v}$ we have, after integrating by parts,

$$\lambda \int d^3v F_0 |g_\lambda|^2 = \int d^3v \frac{\partial g_\lambda^*}{\partial \mathbf{v}} \cdot \int d^3v' F_0 F_0' \left(\frac{\partial g_\lambda}{\partial \mathbf{v}} - \frac{\partial g_\lambda'}{\partial \mathbf{v}'}\right) \cdot \mathbf{Q}$$
$$- \int d^3v \frac{\partial}{\partial \mathbf{v}} \cdot g_\lambda^* \int d^3v' F_0 F_0' \left(\frac{\partial g_\lambda}{\partial \mathbf{v}} - \frac{\partial g_\lambda'}{\partial \mathbf{v}'}\right) \cdot \mathbf{Q}. \tag{6}$$

The second term on the right in (6) vanishes provided g_λ satisfies the second of conditions (4) and

$$\lim_{v\to 0} v^{\frac{1}{2}} g_\lambda(\mathbf{v}) = 0. \tag{7}$$

Assuming these conditions hold, we exchange v and v' in (6), noting that $\mathbf{Q}(\mathbf{v}, \mathbf{v}') = \mathbf{Q}(\mathbf{v}', \mathbf{v})$. We add the result to (6), obtaining

$$2\lambda \int d^3v F_0 |g_\lambda|^2 = \iint d^3v\, d^3v' F_0 F_0' \left(\frac{\partial g_\lambda^*}{\partial \mathbf{v}} - \frac{\partial g_\lambda'^*}{\partial \mathbf{v}'}\right) \cdot \mathbf{Q} \cdot \left(\frac{\partial g_\lambda}{\partial \mathbf{v}} - \frac{\partial g_\lambda'}{\partial \mathbf{v}'}\right). \tag{8}$$

Since $\mathbf{Q}$ is a positive quadratic form, it follows that the right side of (8) is positive or zero. Hence $\mathrm{Re}\,\lambda \geq 0$, and clearly $\mathrm{Im}\,\lambda = 0$.

Employing standard methods[4] we can find from (8) the most general form of g_λ when $\lambda = 0$;

$$g_0(\mathbf{v}) = a_1 v^2 + \mathbf{a}_2 \cdot \mathbf{v} + a_3, \tag{9}$$

with a_1, $\mathbf{a}_2$, and a_3 arbitrary constants.

II. EXPANSION IN SPHERICAL HARMONICS

From (5) we have

$$-\lambda F_0 g_\lambda = \frac{\partial}{\partial \mathbf{v}} \cdot \left(F_0 \frac{\partial g_\lambda}{\partial \mathbf{v}} \cdot \int d^3v' F_0' \mathbf{Q}\right) - \frac{\partial}{\partial \mathbf{v}} \cdot \left(F_0 \int d^3v' F_0' \frac{\partial g_\lambda'}{\partial \mathbf{v}'} \cdot \mathbf{Q}\right). \tag{5}$$

To perform a parts integration on the second term, consider the quantity

$$\frac{\partial}{\partial \mathbf{v}'} \cdot (F_0' g_\lambda' \mathbf{Q}) = F_0' g_\lambda' \frac{\partial}{\partial \mathbf{v}'} \cdot \mathbf{Q} + F_0' \frac{\partial g_\lambda'}{\partial \mathbf{v}'} \cdot \mathbf{Q} + g_\lambda' \frac{\partial F_0'}{\partial \mathbf{v}'} \cdot \mathbf{Q}. \tag{10}$$

With $F_0 = N_0(\alpha_i/\pi)^{\frac{3}{2}} \exp(-\alpha_i v^2)$, $\alpha_i = m_i/2kT$, we have $\partial F_0'/\partial \mathbf{v}' = -2\alpha_i \mathbf{v}' F_0'$. Also $\mathbf{v}' \cdot \mathbf{Q} = \mathbf{v} \cdot \mathbf{Q}$, so (10) gives

$$\int d^3v' F_0' \frac{\partial g_\lambda'}{\partial \mathbf{v}'} \cdot \mathbf{Q} = \int d^3v' \frac{\partial}{\partial \mathbf{v}'} \cdot (F_0' g_\lambda' \mathbf{Q}) - \int d^3v' F_0' g_\lambda' \frac{\partial}{\partial \mathbf{v}'} \cdot \mathbf{Q} + 2 \frac{\partial}{\partial \mathbf{v}} \cdot \int d^3v' F_0' g_\lambda' \mathbf{Q}. \tag{11}$$

The first term on the right in (11) vanishes if g_λ satisfies (4). Using the relations

$$\mathbf{Q} = \Gamma_{ii} \frac{\partial^2 g}{\partial \mathbf{v}\, \partial \mathbf{v}}, \quad \frac{\partial}{\partial \mathbf{v}'} \cdot \mathbf{Q} = -2\Gamma_{ii} \frac{\partial}{\partial \mathbf{v}} \left(\frac{1}{g}\right),$$

and introducing a dimensionless time τ and dimensionless velocity $\mathbf{c}$ in (3),

$$\tau (= \tau_i) = 4\pi N_0 \Gamma_{ii} (\alpha_i/\pi)^{\frac{3}{2}} t, \quad \mathbf{c} = \alpha_i^{\frac{1}{2}} \mathbf{v},$$

the kinetic equation takes the form

$$-4\pi\lambda e^{-c^2} g_\lambda(\mathbf{c}) = \frac{\partial}{\partial \mathbf{c}} \cdot \left(e^{-c^2} \frac{\partial g_\lambda}{\partial \mathbf{c}} \cdot \frac{\partial^2}{\partial \mathbf{c}\partial \mathbf{c}} \int d^3c' e^{-c'^2} |\mathbf{c} - \mathbf{c}'|\right) - 2 \frac{\partial}{\partial \mathbf{c}} \cdot \left(e^{-c^2} \frac{\partial}{\partial \mathbf{c}} \int d^3c' e^{-c'^2} g_\lambda' |\mathbf{c} - \mathbf{c}'|^{-1}\right) - 2 \frac{\partial}{\partial \mathbf{c}} \cdot \left(\mathbf{c} e^{-c^2} \cdot \frac{\partial^2}{\partial \mathbf{c}\partial \mathbf{c}} \int d^3c' e^{-c'^2} g_\lambda' |\mathbf{c} - \mathbf{c}'|\right). \tag{12}$$

[4] Reference 2, p. 85.

The time τ is measured in units of the "Spitzer self-collision time,"[5] and $c = |\mathbf{c}|$ is in units of the rms thermal speed.

Equation (12), in three dimensions, may be replaced by a set of uncoupled equations in one dimension by introducing the spherical harmonic expansion

$$g_\lambda(\mathbf{c}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} g_{lm}(c, \lambda_{lm}) Y_l^m(\theta, \phi).$$

We find (see Appendix A)

$$\int d^3c' e^{-c'^2} g'_\lambda |\mathbf{c} - \mathbf{c}'|^{-1} = \sum_{l,m} \frac{4\pi}{2l+1} Y_l^m R_{lm}, \quad (13)$$

$$\int d^3c' e^{-c'^2} g'_\lambda |\mathbf{c} - \mathbf{c}'| = \sum_{l,m} \frac{4\pi}{4l^2-1} Y_l^m S_{lm}, \quad (14)$$

$$\int d^3c' e^{-c'^2} |\mathbf{c} - \mathbf{c}'| = 4\pi T, \quad (15)$$

where

$$R_{lm}(c) = \int_0^c dc' c' \left(\frac{c'}{c}\right)^{l+1} e^{-c'^2} g'_{lm} + \int_c^\infty dc' c' \left(\frac{c}{c'}\right)^l e^{-c'^2} g'_{lm},$$

$$S_{lm}(c) = \int_0^c dc' c'^3 \left(\frac{c'}{c}\right)^{l-1} \left[\frac{2l-1}{2l+3}\left(\frac{c'}{c}\right)^2 - 1\right] e^{-c'^2} g'_{lm} + \int_c^\infty dc' c'^3 \left(\frac{c}{c'}\right)^l \left[\frac{2l-1}{2l+3}\left(\frac{c}{c'}\right)^2 - 1\right] e^{-c'^2} g'_{lm},$$

$$T(c) = \tfrac{1}{4} e^{-c^2} + \left(c + \frac{1}{2c}\right) \frac{\pi^{\frac{1}{2}}}{4} \operatorname{erf}(c),$$

with

$$\operatorname{erf}(c) \equiv 2\pi^{-\frac{1}{2}} \int_0^c dx \exp(-x^2).$$

The expressions for R_{lm} and S_{lm} were found by Rosenbluth *et al.*,[1] using a different method, for the axially symmetric case $m = 0$.

Combining (12)–(15) we find

$$-\lambda e^{-c^2} \sum_{l,m} Y_l^m g_{lm} = \sum_{l,m} \frac{\partial}{\partial \mathbf{c}} \cdot \left(e^{-c^2} \frac{\partial}{\partial \mathbf{c}} (Y_l^m g_{lm}) \cdot \frac{\partial^2 T}{\partial \mathbf{c} \partial \mathbf{c}}\right) - \sum_{l,m} \frac{2}{2l+1} \frac{\partial}{\partial \mathbf{c}} \cdot \left(e^{-c^2} \frac{\partial}{\partial \mathbf{c}} (Y_l^m R_{lm})\right) - \sum_{l,m} \frac{2}{4l^2-1} \frac{\partial}{\partial \mathbf{c}} \cdot \left(\mathbf{c} e^{-c^2} \cdot \frac{\partial^2}{\partial \mathbf{c} \partial \mathbf{c}} (Y_l^m S_{lm})\right). \quad (16)$$

Since $T(c)$ is isotropic we can write

$$\partial^2 T / \partial \mathbf{c} \partial \mathbf{c} = \hat{\mathbf{e}}_c \hat{\mathbf{e}}_c T'' + (\mathbf{I} - \hat{\mathbf{e}}_c \hat{\mathbf{e}}_c) c^{-1} T',$$

where primes now denote total differentiation with respect to c. Performing the indicated angular differentiations in (16) and then employing the orthogonality property of the spherical harmonics, we obtain the "radial" equation

$$-\lambda_{lm} g_{lm} = T'' \frac{d^2 g_{lm}}{dc^2} + \left[T''' + \left(\frac{2}{c} - 2c\right) T''\right] \times \frac{dg_{lm}}{dc} - \frac{l(l+1)}{c^3} T' g_{lm} - \frac{2}{2l+1}\left[R''_{lm} + \left(\frac{2}{c} - 2c\right) R'_{lm} - \frac{l(l+1)}{c^2} R_{lm}\right] - \frac{2}{4l^2-1}\left(cS'''_{lm} + (3 - 2c^2) S''_{lm} - \frac{l(l+1)}{c^2}(cS'_{lm} - S_{lm})\right). \quad (17)$$

The index m is clearly superfluous and is deleted in the following. Performing the primed differentiations in (17) we find

$$\begin{aligned} -\lambda_l g_l = {} & \left(\frac{\pi^{\frac{1}{2}}}{4c^3} \operatorname{erf}(c) - \frac{1}{2c^2} e^{-c^2}\right) \frac{d^2 g_l}{dc^2} \\ & + \left[\left(\frac{1}{2c^3} + \frac{2}{c}\right) e^{-c^2} - \left(\frac{1}{c^4} + \frac{2}{c^2}\right) \frac{\pi^{\frac{1}{2}}}{4} \operatorname{erf}(c)\right] \frac{dg_l}{dc} \\ & + \left[2e^{-c^2} - \frac{l(l+1)}{c^3} \times \left\langle \frac{1}{4c} e^{-c^2} + \left(1 - \frac{1}{2c^2}\right) \frac{\pi^{\frac{1}{2}}}{4} \operatorname{erf}(c) \right\rangle\right] g_l \\ & + \frac{4c^3 (l+1)(l+2)}{(2l+1)(2l+3)} \int_0^c dc' e^{-c'^2} \left(\frac{c'}{c}\right)^{l+4} g_l(c') \\ & - \frac{4c}{2l+1}\left(1 + \frac{c^2 l(l-1)}{2l-1}\right) \times \int_0^c dc' e^{-c'^2} \left(\frac{c'}{c}\right)^{l+2} g_l(c') \\ & - \frac{4c^3 l(l-1)}{4l^2-1} \int_c^\infty dc' e^{-c'^2} \left(\frac{c}{c'}\right)^{l-3} g_l(c') \\ & - \frac{4c}{2l+1}\left(1 - \frac{c^2 (l+1)(l+2)}{2l+3}\right) \times \int_c^\infty dc' e^{-c'^2} \left(\frac{c}{c'}\right)^{l-1} g_l(c'). \end{aligned} \quad (18)$$

For boundary conditions we use conditions (4), which were obtained from the conservation laws. Although (7) is stronger than the first of conditions (4), we see below that the solutions of (18) which satisfy (4) also satisfy (7).

III. THE SPECTRUM

If we introduce the transformation

$$g_l(c, \lambda_l) = c^{-1} e^{c^2/2} \psi_l(c, \lambda_l), \quad (19)$$

[5] L. Spitzer, Jr., *Physics of Fully Ionized Gases* (Interscience Publishers, Inc., New York, 1962), 2nd ed., pp. 132–136.

we obtain from (18) the formally self-adjoint equation

$$\frac{d}{dc}\left(P\frac{d\psi_l}{dc}\right) + (Q_l + \lambda_l)\psi_l + \int_0^\infty dc' K_l(c, c')\psi_l(c') = 0, \tag{20}$$

with

$$P(c) = \frac{\pi^{\frac{1}{2}}}{4c^3}\operatorname{erf}(c) - \frac{1}{2c^2}e^{-c^2},$$

$$Q_l(c) = \left(\frac{3}{c^5} - \frac{1}{c}\right)\frac{\pi^{\frac{1}{2}}}{4}\operatorname{erf}(c) + \left(\frac{7}{2} - \frac{3}{2c^4} - \frac{1}{c^2}\right)e^{-c^2} - \frac{l(l+1)}{c^3}\left[\frac{1}{4c}e^{-c^2} + \left(1 - \frac{1}{2c^2}\right)\frac{\pi^{\frac{1}{2}}}{4}\operatorname{erf}(c)\right],$$

$$K_l(c, c') = \frac{4}{2l+1}e^{-\frac{1}{2}(c^2+c'^2)}\left[\frac{(l+1)(l+2)}{2l+3}\binom{c^3}{c'^3} - \binom{c}{c'} - \frac{l(l+1)}{2l-1}\binom{cc'^2}{c'c^2}\right]\left[\frac{(c/c')^l}{(c'/c)^l}\right](c \le c').$$

Transforming (4) via (19), we have

$$\lim_{c\to\infty} e^{-c^2/2}\frac{d\psi_l}{dc} = 0, \tag{21}$$

$$\lim_{c\to 0}\psi_l = 0. \tag{22}$$

We later show (Appendix B) that the problem (20)–(22) is self-adjoint.

The spectrum of (20) is that set of numbers $\{\lambda_l\}$ such that (20) has nontrivial solutions which satisfy (21) and (22). In view of the self-adjoint property the λ_l are real. We have already seen that λ_l must be positive for $l \geq 2$, and positive or zero for $l = 0, 1$.

We can find the spectrum of (20) by first considering the related problem

$$(d/dc)[P(dy_l/dc)] + (Q_l + \lambda_l)y_l = 0 \tag{23}$$

with conditions on the functions $y_l(c)$ identical to (21) and (22). Clearly P, dP/dc, and Q_l are bounded and continuous for all finite c except possibly near $c = 0$. For $c \ll 1$ we have

$$P(c) = \tfrac{1}{3} - \tfrac{1}{5}c^2 + O(c^4),$$

$$Q_l(c) = \tfrac{17}{5} + O(c^2) - l(l+1)\left(\frac{1}{3c^2} - \frac{1}{15} + O(c^2)\right).$$

In general we can write

$$P(c) = \frac{1}{2c^3}\int_0^c dx(e^{-x^2} - e^{-c^2})$$

and thus $P(c) > 0$ for all $c < \infty$. It follows that for $l \neq 0$, (23) has a regular singular point at $c = 0$.

For c small, (23) has the asymptotic solutions

$$y_l(c) \simeq c^{l+1},\ c^{-l} \quad (c \ll 1). \tag{24}$$

The first of these satisfies (22) for all l. It also satisfies the stronger condition obtained from (7),

$$\lim (c \to 0)c^{-\frac{1}{2}}y_l(c) = 0.$$

The second solution satisfies neither condition. This is clear for $l \neq 0$. For $l = 0$, the second solution is a constant which cannot be zero since the solutions (24) are linearly independent.

For c sufficiently large and for $\lambda_l \neq 0$, (23) takes the asymptotic form

$$\frac{d^2y_l}{dc^2} - \frac{3}{c}\frac{dy_l}{dc} + \frac{4\lambda_l c^3}{\pi^{\frac{1}{2}}}y_l = 0 \quad (c \text{ large}, \lambda_l \neq 0).$$

We find ($\lambda_l > 0$)

$$y_l(c) \simeq A_l c^{\frac{3}{4}}\cos(\nu_l c^{\frac{5}{2}} - \gamma_l), \tag{25}$$

with $\nu_l \equiv 4\lambda_l^{\frac{1}{2}}/5\pi^{\frac{1}{4}}$.

Given l and λ_l, Eq. (23) has only one solution which satisfies the condition at $c = 0$. This solution thus contains only one arbitrary constant and it follows that A_l and γ_l in (25) are not independent. Whatever the relation between A_l and λ_l is, (25) satisfies (21) for all positive γ_l. Hence the spectrum of (20) contains all positive λ_l.

To determine the spectrum of the problem in Eqs. (20)–(22), we first note that the solutions of the problem in Eqs. (21)–(23) do not exist in Hilbert space; hence we must introduce a transformation in order to apply a Hilbert space theorem. We recognize the symmetric differential and integral operators in (20) as, respectively, closed and completely continuous operators. It is possible to introduce an algebraic transformation on (20) such that the respective closed and completely continuous properties are retained, while the solutions of the transformed problem in Eqs (21)–(23) now do exist in Hilbert space. Since the transformation has destroyed the symmetry of the problem, we employ a recent extension[6] of Weyl's perturbation theorem which applies to asymmetric linear operators. This extension states that[6] the addition of a completely continuous operator to a closed operator does not change the limit points of the spectrum of the closed operator. With our knowledge of the spectral properties of Eqs. (21)–(23) and in view of the symmetry of Eq. (20), it now follows that the spectrum of each l- component of the linearized Fokker–Planck equation contains all positive real λ_l; we showed earlier that the spectrum is empty for $\lambda_l < 0$.

For $l = 0$ and $l = 1$ we found, respectively, two-fold and three-fold degenerate zero eigenvalues, corresponding to a shift to an equilibrium different from

[6] I. M. Glazman, *Direct Methods of Qualitative Spectral Analysis of Singular Differential Operators* (Daniel Davey & Co., New York, 1966).

that originally postulated. For $l \geq 2$ the eigenvalues are positive and real and are $(2l+1)$-fold degenerate, corresponding to the $2l+1$ different spherical harmonics of order l.

IV. THE EXPANSION THEOREM

With the spectral properties of the kinetic equation in hand, we can now proceed to the development of an expansion theorem based on (20). Our method is essentially an extension of the theory of singular ordinary differential equations, based on a theory due to Weyl and Levinson,[7] to include singular integro–differential equations with Hilbert–Schmidt kernels.

Since the eigenfunctions of the kinetic equation are bounded and continuous on every finite interval, it is natural to pursue an expansion formula for functions $u(c)$ square integrable on the interval Δ: $0 \leq c < \infty$. As with the earlier theory we first establish an expansion formula on a finite subinterval δ of Δ, δ: $a \leq c \leq b$, $0 < a$, $b < \infty$, so that the singularities of the linear operator are external to δ. The expansion formula we seek is then obtained by taking $\delta \to \Delta$ in a suitable manner. In the following we mean by L_Δ the integro–differential operator in (20) and by L_δ the operator obtained when the lower and upper limits of the integral in (20) are replaced by a and b, respectively. In the following, the index l is retained only where it is necessary to avoid confusion.

We have already seen that P, P', and Q_l are continuous on δ, and that $K_l(c, c')$ is bounded and integrable on the square $a \leq c \leq b$, $a \leq c' \leq b$. Tamarkin[8] has shown that, subject to these conditions, the solutions of $L_\delta\phi = -\lambda\phi$ which satisfy homogeneous boundary conditions at $c = a$ and $c = b$ form a complete orthogonal and normalizable set of eigenfunctions $(h_{\delta n})$ on δ, with an associated denumerable sequence of real eigenvalues $(\lambda_{\delta n})$. Assuming the $h_{\delta n}$ to be normalized, the expansion formula on δ is thus

$$u(c) = \sum_{n=0}^{\infty} h_{\delta n} \int_a^b dc\, u(c) h^*_{\delta n}(c), \tag{26}$$

where $u(c)$ is any function square integrable on δ.

We now use the Weyl–Levinson theory to take $\delta \to \Delta$. Since the subsequent development of the expansion theorem is in every respect a duplication of the earlier theory, we display only the salient features.

Given λ, the most general solution of $L_\delta\phi = -\lambda\phi$ is a linear combination of the two linearly independent solutions, say ϕ_1, ϕ_2. Thus, we can write

$$h_{\delta n}(c, \lambda_{\delta n}) = r_{\delta n1}\phi_1(c, \lambda_{\delta n}) + r_{\delta n2}\phi_2(c, \lambda_{\delta n}), \tag{27}$$

where $r_{\delta n1}$ and $r_{\delta n2}$ are complex constants. With (27), (26) becomes

$$u(c) = \sum_{n=0}^{\infty} \sum_{j,k=1}^{2} r_{\delta nj} r^*_{\delta nk} \phi_j \int_a^b dc\, u(c) \phi^*_k. \tag{28}$$

Following Levinson[7] we define an Hermitian, positive semidefinite matrix ρ_δ, called the spectral matrix, with elements $\rho_{\delta jk}$ which consist of step functions with jumps at the eigenvalues $\lambda_{\delta n}$ given by

$$\rho_{\delta jk}(\lambda_{\delta n} + 0) - \rho_{\delta jk}(\lambda_{\delta n} - 0) = r_{\delta nj} r^*_{\delta nk}.$$

Let $\rho_\delta(\lambda + 0) = \rho_\delta(\lambda)$, and let $\rho_\delta(0)$ be the zero matrix. We use the spectral matrix to replace the infinite series in (28) by a Lebesgue–Stieltjes integral

$$u(c) = \int_{-\infty}^{\infty} \sum_{j,k=1}^{2} \phi_j(c, \lambda) \tilde{u}_k(\lambda)\, d\rho_{\delta jk}(\lambda), \tag{29}$$

where

$$\tilde{u}_k = \int_a^b dc\, u(c) \phi^*_k(c, \lambda).$$

As $\delta \to \Delta$ (that is, $a \to 0$, $b \to \infty$), ρ_δ approaches a limit matrix ρ_Δ. To find ρ_Δ, let $\lambda = \mu + i\omega$, $\omega > 0$, and let $\chi_a = \phi_1 + m_a(\lambda)\phi_2$ be a solution of $L_\delta\phi = -\lambda\phi$ satisfying the homogeneous boundary condition

$$\cos \alpha\chi(a) + \sin \alpha P(a)\chi'(a) = 0,$$

and similarly let $\chi_b = \phi_1 + m_b(\lambda)\phi_2$ be a solution of the same equation satisfying

$$\cos \beta\chi(b) + \sin \beta P(b)\chi'(b) = 0.$$

Clearly $m_a = r_{\delta n2}(a)/r_{\delta n1}(a)$ and similarly for m_b. As $a \to 0$ and $b \to \infty$, m_a and m_b approach limiting values in the complex m plane denoted, respectively, by $m_0(\lambda)$ and $m_\infty(\lambda)$. These limiting values are clearly determined by the behavior of ϕ_1 and ϕ_2 for small and large c, for λ complex.

For ϕ_1 and ϕ_2 to be linearly independent, it is necessary and sufficient that their Wronskian equal a nonzero constant, say one;

$$P(\phi_1\phi_2' - \phi_2\phi_1') = 1. \tag{30}$$

This last is satisfied if ϕ_1 and ϕ_2 satisfy the conditions

$$\phi_1(s, \lambda) = \sin \sigma, \qquad \phi_2(s, \lambda) = \cos \sigma,$$
$$P(s)\phi_1'(s, \lambda) = -\cos \sigma, \quad P(s)\phi_2'(s, \lambda) = \sin \sigma,$$

where s is an interior point of δ and $0 \leq \sigma < \pi$. These conditions are also sufficient to ensure that ϕ_1, ϕ_2 are entire functions of λ for each fixed c on δ (this follows from Tamarkin[8]).

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

[8] J. D. Tamarkin, Trans. Am. Math. Soc. **29**, 755 (1927).

With these properties secured, we can find[7] the limit values m_0 and m_∞ and hence the limit matrix ρ_Δ, whose elements are given by

$$\rho_{\Delta jk}(\lambda) - \rho_{\Delta jk}(\eta) = \lim_{\omega \to +0} \frac{1}{\pi} \int_\eta^\lambda \operatorname{Im} M_{jk}(\mu + i\omega)\, d\mu, \tag{31}$$

where

$$M_{11}(\lambda) = \frac{1}{m_0 - m_\infty},$$

$$M_{12}(\lambda) = M_{21}(\lambda) = \frac{1}{2}\frac{m_0 + m_\infty}{m_0 - m_\infty},$$

$$M_{22}(\lambda) = \frac{m_0 m_\infty}{m_0 - m_\infty}.$$

To find the M_{jk} we need asymptotic forms of ϕ_1 and ϕ_2 for large and small c. These are given by (24) and (25), as may be verified by direct substitution. Taking ϕ_1, ϕ_2 to be asymptotic, respectively, to c^{-l}, c^{l+1} for c small, we apply the homogeneous boundary condition to χ_a and then take $a \to 0$ to find

$$\begin{aligned} m_0 &= \infty \qquad (l \neq 0), \\ m_0 &= -\cot \alpha \quad (l = 0). \end{aligned} \tag{32}$$

Thus for $l \neq 0$ only M_{22} can have a nonzero imaginary part and, consequently, only ϕ_2 contributes to the expansion formula (29). When $l = 0$ both solutions are regular at $c = 0$ and the limit matrix is not determined until we specify α. The boundary condition (22) dictates the choice $\alpha = 0$.

For c large we take

$$\phi_2 \simeq A_l c^{\frac{3}{4}} \cos(\nu_l c^{\frac{5}{2}} - \gamma_l) \quad (\nu_l = 4\lambda_l^{\frac{1}{2}}/5\pi^{\frac{1}{2}},\ \lambda_l > 0)$$

and find ϕ_1 by integrating (30);

$$\phi_1 \simeq -\frac{8c^{\frac{3}{4}} \sin[\nu_l(c^{\frac{5}{2}} - c_0^{\frac{5}{2}})]}{5\pi^{\frac{1}{2}} A_l \nu_l \cos(\nu_l c_0^{\frac{5}{2}} - \gamma_l)},$$

where c_0 is a constant of integration. Applying the homogeneous boundary condition to $\chi_b \equiv \phi_1 + m_b(\lambda)\phi_2$ and then taking $b \to \infty$ with $\operatorname{Im} \lambda > 0$ we find

$$m_\infty = \frac{8i \exp[i(\nu_l c_0^{\frac{5}{2}} - \gamma_l)]}{5\pi^{\frac{1}{2}} A_l^2 \nu_l \cos(\nu_l c_0^{\frac{5}{2}} - \gamma_l)}. \tag{33}$$

Combining (31)–(33) we have finally

$$d\rho(\lambda_l) \equiv d\rho_{\Delta 22}(\lambda_l) = \frac{2\, d\lambda}{\pi^{\frac{5}{2}} A_l^2(\lambda_l)\lambda_l^{\frac{1}{2}}} \quad (\lambda_l > 0). \tag{34}$$

Since the spectrum is empty for $\lambda_l < 0$, $\rho(\lambda_l)$ is constant on this range. The expansion formula (29) becomes

$$u(c) = \int_0^\infty \phi_2(c, \lambda_l)\tilde{u}(\lambda_l)\, d\rho(\lambda_l) \tag{35}$$

with

$$\tilde{u} = \int_0^\infty u(c)\phi_2^*(c, \lambda_l)\, dc.$$

The expansion converges in the mean for all functions $u(c)$ square integrable on $(0, \infty)$. If the spectral function ρ is not continuous at $\lambda_l = 0$, this point will contribute to the integral in (35).

We return to the description of perturbations from equilibrium. If $\exp(-c^2)f(\mathbf{c}, 0)$ is square integrable in velocity space, then from (19) and (35) we have

$$e^{-c^2} f(\mathbf{c}, \tau) = \sum_{l,m} \int_0^\infty d\rho(\lambda_l) \Psi_{lm} \tilde{F}_{lm} e^{-\lambda_l \tau} \tag{36}$$

with

$$\tilde{F}_{lm}(\lambda_l) = \tfrac{1}{2}(2l + 1) \int d^3c \Psi_{lm}^* f(\mathbf{c}, 0)$$

and

$$\Psi_{lm}(\mathbf{c}, \lambda_l) = c^{-1} e^{-c^2/2} \psi_l(c, \lambda_l) Y_l^m(\theta, \phi).$$

The functions ψ_l correspond to ϕ_2 in (35) and are the solutions of (20) satisfying (22).

We have defined the density, mean velocity, and kinetic temperature of the ion gas as being proportional respectively to the first three moments of the equilibrium distribution F_0. If ρ is continuous at $\lambda_l = 0$ this point does not contribute to the expansion formula (36) and the eigenfunctions (9) for $\lambda_l = 0$ are not contained in the expansion. By virtue of the conservation laws, the functions (9) are then orthogonal to (36). Thus (36) is complete only if ρ has a jump at $\lambda_l = 0$ for $l = 0, 1$. It follows that the exclusion of (9) from (36) yields an expansion which is complete with respect to all square integrable perturbations conserving N_0, $\langle \mathbf{v} \rangle$, and kT.

V. THE ELECTRON KINETIC EQUATION

We have developed an expansion theorem based on the uncoupled kinetic equation (3) for the ions. The extension to the electron kinetic equation is straightforward and requires only a little algebra.

As indicated earlier, we decouple the electron kinetic equation from the ion equation by dropping the term $(m_e/m_i)\partial f_i'/\partial \mathbf{v}'$ in (2). This amounts to neglecting the effect of the ion perturbation on the electron perturbation, but retains the effects of encounters with ions in the thermal distribution.

The uncoupled equation conserves electron number density provided (4) holds, but does not conserve momentum or kinetic energy in the electron gas. This is as it should be, since a substantial portion of the electron momentum, and a small amount of the energy, is lost to the ions.

Applying the methods of Sec. I we find as before $\mathrm{Re}\,\lambda \geq 0$, $\mathrm{Im}\,\lambda = 0$, and for $\lambda = 0$ we find $g_0 = \text{const}$, corresponding to (9). The remainder of the development proceeds as before. A spherical harmonic expansion yields a set of singular integro-differential equations, and the transformation (19) brings these into self-adjoint form. As before, the expansion formula has the form of a Riemann-Stieltjes, or generalized Fourier, integral.

If, for example, the ions are protons, then we can take $\Gamma_{ee} = \Gamma_{ei}$. If we use α_e in place of α_i in the definition of τ and c, then the electron equations may be obtained from the ion equations by replacing $T(c)$ in (15) by $T(c) + \alpha^{\frac{1}{2}} T(\alpha^{-\frac{1}{2}} c)$, where $\alpha \equiv \alpha_e/\alpha_i = m_e/m_i$.

VI. DISCUSSION

We have used boundary conditions obtained by requiring the solutions of the kinetic equation to be consistent with the conservation laws. The Hilbert space then emerged as a natural function space for the framework of the mathematical development. The question persists (see, e.g., the discussion in Uhlenbeck and Ford[9]) as to whether square integrability should be a requirement on the distribution functions from the beginning. In the light of the present work, this condition does not appear to be necessary, and for our purposes, it would not have been sufficient. To see this we note the condition $\int d^3v\, |F_0 f|^2 < \infty$ leads to

$$\lim_{v \to 0} v^{\frac{3}{2}} f = 0, \tag{37}$$

which is weaker than the corresponding condition (4). Since both solutions of (20) satisfy (37) for $l = 0$, it would be possible to have an expansion theorem for solutions of the kinetic equation which are square integrable but do not satisfy the conservation laws.

ACKNOWLEDGMENTS

The author is indebted to Professor R. K. Osborn, Professor C. L. Dolph, and Professor F. C. Shure for many helpful discussions during the course of this work, which represents part of a Ph.D. dissertation submitted to the university of Michigan (1966).

This research was supported in part by the United States Atomic Energy Commission and by the Advanced Research Projects Agency, Project Defender, ARPA Order No. 675, was monitored by the United States Army Research Office (Durham) under Contract DA-31-124-ARO-D-403.

[9] G. E. Uhlenbeck and G. W. Ford, *Lectures in Statistical Mechanics* (American Mathematical Society, Providence, Rhode Island, 1963), p. 88.

APPENDIX A. INTEGRALS

Let Θ be the angle between $\mathbf{c}$ and $\mathbf{c}'$. From the generating function relation for Legendre polynomials, we have

$$|\mathbf{c} - \mathbf{c}'|^{-1} = (c^2 + c'^2 - 2cc' \cos\Theta)^{-\frac{1}{2}} = c^{-1} \sum_{l=0}^{\infty} \left(\frac{c'}{c}\right)^l P_l(\cos\Theta) \quad (c' \leq c), \tag{A1}$$

and similarly for $c \leq c'$. Writing

$$g_\lambda(\mathbf{c}) = \sum g_{lm} Y_l^m$$

and employing the addition theorem for Legendre polynomials, we obtain Eq. (13).

To find Eq. (14) we use the relation

$$(1 + x^2 - 2xy)^{\frac{1}{2}} = \int dx x (1 + x^2 - 2xy)^{-\frac{1}{2}} - y \int dx (1 + x^2 - 2xy)^{-\frac{1}{2}}. \tag{A2}$$

Combining (A1) and (A2) we have

$$|\mathbf{c} - \mathbf{c}'| = (c^2 + c'^2 - 2cc' \cos\Theta)^{\frac{1}{2}} = c^{-1} \sum_{l=0}^{\infty} \left(\frac{(c'/c)^{l+2}}{l+2} - \frac{(c'/c)^{l+1}}{l+1} \cos\Theta \right) P_l(\cos\Theta) \quad (c' \leq c), \tag{A3}$$

and similarly for $c \leq c'$. From the pure recurrence relation for Legendre polynomials, we have

$$\cos\Theta P_l(\cos\Theta) = \frac{l}{2l+1} P_{l-1}(\cos\Theta) + \frac{l+1}{2l+1} P_{l+1}(\cos\Theta). \tag{A4}$$

Combining (A3) and (A4) and then using the addition theorem as before we find Eq. (14).

Finally, Eq. (15) is obtained from Eq. (14) with $g_{lm} = \delta_{l0}\delta_{m0}$ with δ_{jk} the Kronecker delta.

APPENDIX B. THE SELF-ADJOINT PROPERTY

Let L_δ be either (i) the differential operator in Eq. (20) or (ii) the integro-differential operator in Eq. (20) defined as in Sec. IV on the closed interval δ: $a \leq c \leq b$, $0 < a, b < \infty$. Let L_Δ be similarly defined on $(0, \infty)$. Then in either case (i) or case (ii) there exists a complete orthonormal set of functions $\{h_{\delta n}\}$ on δ, generated by $L_\delta \phi = -\lambda\phi$ with homogeneous boundary conditions at a and b.[7,8]

Let f and g be any nonzero functions square integrable on $(0, \infty)$, and consider the inhomogeneous

problems

$$L_\delta u + \Lambda u = f, \quad L_\delta v + \Lambda^* v = g \tag{B1}$$

with the same homogeneous boundary conditions at a and b as in the above homogeneous problem. Let $\operatorname{Im}\Lambda \neq 0$ so that Λ will not belong to the spectrum of the set $\{h_{\delta n}\}$. Then the problems (B1) have nontrivial solutions[8]

$$u(c) = \int_{-\infty}^{\infty} \sum_{j,k=1}^{2} \frac{\phi_j(c,\lambda)\tilde{u}_k(\lambda)\, d\rho_{\delta jk}(\lambda)}{\Lambda - \lambda},$$

$$v(c) = \int_{-\infty}^{\infty} \sum_{j,k=1}^{2} \frac{\phi_j(c,\lambda)\tilde{v}_k(\lambda)\, d\rho_{\delta jk}(\lambda)}{\Lambda^* - \lambda},$$

with

$$\tilde{u}_k = \int_a^b f(c)\phi_k^*(c,\lambda)\, dc, \quad \tilde{v}_k = \int_a^b g(c)\phi_k^*(c,\lambda)\, dc. \tag{B2}$$

Now, let

$$u_\mu = \int_{-\mu}^{\mu} \sum_{j,k=1}^{2} \frac{\phi_j \tilde{u}_k\, d\rho_{\delta jk}}{\Lambda - \lambda},$$

$$v_\mu = \int_{-\mu}^{\mu} \sum_{j,k=1}^{2} \frac{\phi_j \tilde{v}_k\, d\rho_{\delta jk}}{\Lambda^* - \lambda},$$

with $\tilde{u}_k$ and $\tilde{v}_k$ as in (B2). Multiplying u_μ by g^* and v_μ^* by f and then integrating over δ we have

$$\int_a^b g^*(c)u_\mu(c)\, dc = \int_{-\mu}^{\mu} \sum_{j,k=1}^{2} \frac{\tilde{u}_k(\lambda)\, d\rho_{\delta jk}(\lambda)}{\Lambda - \lambda} \int_a^b g^*(c)\phi_j(c,\lambda)\, dc$$

$$= \int_{-\mu}^{\mu} \sum_{j,k=1}^{2} \frac{\tilde{u}_k(\lambda)\tilde{v}_j^*(\lambda)\, d\rho_{\delta jk}(\lambda)}{\Lambda - \lambda}, \tag{B3}$$

and similarly

$$\int_a^b f(c)v_\mu^*(c)\, dc = \int_{-\mu}^{\mu} \sum_{j,k=1}^{2} \frac{\tilde{v}_k^*(\lambda)\tilde{u}_j(\lambda)\, d\rho_{\delta jk}(\lambda)}{\Lambda - \lambda}. \tag{B4}$$

Since (24) and (25) are asymptotic solutions for both the differential and the integro-differential equations, the limit matrix ρ_Δ is the same in either case. Taking $\delta \to \Delta$ in (B3) and (B4) we have $\rho_{\delta jk} \to \rho_{\Delta 22}$ and thus

$$\int_0^\infty g^*(c)u_\mu(c)\, dc = \int_{-\mu}^{\mu} \frac{\tilde{u}_2\tilde{v}_2^*\, d\rho_{\Delta 22}}{\Lambda - \lambda} = \int_0^\infty f(c)v_\mu^*(c)\, dc. \tag{B5}$$

After taking $\mu \to \infty$ and employing (B1), (B5) becomes

$$\int_0^\infty (L_\Delta v)^* u\, dc = \int_0^\infty (L_\Delta u)v^*\, dc,$$

which is the desired result.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Real Spinor Fields

DAVID HESTENES*
Palmer Laboratory, Princeton University, Princeton, New Jersey

(Received 22 March 1966)

The Dirac equation is expressed entirely in terms of geometrical quantities by providing a geometrical interpretation for the $(-1)^{\frac{1}{2}}$ which appears explicitly in the Dirac equation. In the modification of the Dirac electron theory which ensues, the $(-1)^{\frac{1}{2}}$ appears as the generator of rotations in the spacelike plane orthogonal to the plane containing the electron current and spin vectors. This amounts to a further "relativistic" constraint on the spinor theory and so may be expected to have physical consequences. It does not, however, conflict with well-substantiated features of the Dirac theory.

INTRODUCTION

IN 1928, Dirac proposed a relativistically invariant first-order wave equation for the electron.[1] Dirac's theory has led to a complex of physical explanations and predictions at once so surprising and convincing that it has gained general acceptance among physicists today. The previously perplexing phenomena of electron spin was not only accounted for, but fine details of the hydrogen spectra and an accurate value of the electron magnetic moment were calculated without arbitrary assumptions. Moreover, after some theoretical trauma, it was realized that Dirac's equation entails the existence of a positively charged electron—at just about the same time that such a particle was discovered experimentally. We do well to understand precisely what features of the Dirac equation entail these remarkable results.

* Present address: Physics Department, Arizona State University, Tempe, Arizona.

[1] P. A. M. Dirac, Proc. Roy. Soc. (London) 117, 610 (1928).

problems

$$L_\delta u + \Lambda u = f, \quad L_\delta v + \Lambda^* v = g \tag{B1}$$

with the same homogeneous boundary conditions at a and b as in the above homogeneous problem. Let $\operatorname{Im}\Lambda \neq 0$ so that Λ will not belong to the spectrum of the set $\{h_{\delta n}\}$. Then the problems (B1) have nontrivial solutions[8]

$$u(c) = \int_{-\infty}^{\infty} \sum_{j,k=1}^{2} \frac{\phi_j(c,\lambda)\tilde{u}_k(\lambda)\, d\rho_{\delta jk}(\lambda)}{\Lambda - \lambda},$$

$$v(c) = \int_{-\infty}^{\infty} \sum_{j,k=1}^{2} \frac{\phi_j(c,\lambda)\tilde{v}_k(\lambda)\, d\rho_{\delta jk}(\lambda)}{\Lambda^* - \lambda},$$

with

$$\tilde{u}_k = \int_a^b f(c)\phi_k^*(c,\lambda)\, dc, \quad \tilde{v}_k = \int_a^b g(c)\phi_k^*(c,\lambda)\, dc. \tag{B2}$$

Now, let

$$u_\mu = \int_{-\mu}^{\mu} \sum_{j,k=1}^{2} \frac{\phi_j \tilde{u}_k\, d\rho_{\delta jk}}{\Lambda - \lambda},$$

$$v_\mu = \int_{-\mu}^{\mu} \sum_{j,k=1}^{2} \frac{\phi_j \tilde{v}_k\, d\rho_{\delta jk}}{\Lambda^* - \lambda},$$

with $\tilde{u}_k$ and $\tilde{v}_k$ as in (B2). Multiplying u_μ by g^* and v_μ^* by f and then integrating over δ we have

$$\int_a^b g^*(c)u_\mu(c)\, dc = \int_{-\mu}^{\mu} \sum_{j,k=1}^{2} \frac{\tilde{u}_k(\lambda)\, d\rho_{\delta jk}(\lambda)}{\Lambda - \lambda} \int_a^b g^*(c)\phi_j(c,\lambda)\, dc$$

$$= \int_{-\mu}^{\mu} \sum_{j,k=1}^{2} \frac{\tilde{u}_k(\lambda)\tilde{v}_j^*(\lambda)\, d\rho_{\delta jk}(\lambda)}{\Lambda - \lambda}, \tag{B3}$$

and similarly

$$\int_a^b f(c)v_\mu^*(c)\, dc = \int_{-\mu}^{\mu} \sum_{j,k=1}^{2} \frac{\tilde{v}_k^*(\lambda)\tilde{u}_j(\lambda)\, d\rho_{\delta jk}(\lambda)}{\Lambda - \lambda}. \tag{B4}$$

Since (24) and (25) are asymptotic solutions for both the differential and the integro-differential equations, the limit matrix ρ_Δ is the same in either case. Taking $\delta \to \Delta$ in (B3) and (B4) we have $\rho_{\delta jk} \to \rho_{\Delta 22}$ and thus

$$\int_0^\infty g^*(c)u_\mu(c)\, dc = \int_{-\mu}^{\mu} \frac{\tilde{u}_2\tilde{v}_2^*\, d\rho_{\Delta 22}}{\Lambda - \lambda} = \int_0^\infty f(c)v_\mu^*(c)\, dc. \tag{B5}$$

After taking $\mu \to \infty$ and employing (B1), (B5) becomes

$$\int_0^\infty (L_\Delta v)^* u\, dc = \int_0^\infty (L_\Delta u)v^*\, dc,$$

which is the desired result.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Real Spinor Fields

DAVID HESTENES*
Palmer Laboratory, Princeton University, Princeton, New Jersey

(Received 22 March 1966)

The Dirac equation is expressed entirely in terms of geometrical quantities by providing a geometrical interpretation for the $(-1)^{\frac{1}{2}}$ which appears explicitly in the Dirac equation. In the modification of the Dirac electron theory which ensues, the $(-1)^{\frac{1}{2}}$ appears as the generator of rotations in the spacelike plane orthogonal to the plane containing the electron current and spin vectors. This amounts to a further "relativistic" constraint on the spinor theory and so may be expected to have physical consequences. It does not, however, conflict with well-substantiated features of the Dirac theory.

INTRODUCTION

IN 1928, Dirac proposed a relativistically invariant first-order wave equation for the electron.[1] Dirac's theory has led to a complex of physical explanations and predictions at once so surprising and convincing that it has gained general acceptance among physicists today. The previously perplexing phenomena of electron spin was not only accounted for, but fine details of the hydrogen spectra and an accurate value of the electron magnetic moment were calculated without arbitrary assumptions. Moreover, after some theoretical trauma, it was realized that Dirac's equation entails the existence of a positively charged electron—at just about the same time that such a particle was discovered experimentally. We do well to understand precisely what features of the Dirac equation entail these remarkable results.

* Present address: Physics Department, Arizona State University, Tempe, Arizona.

[1] P. A. M. Dirac, Proc. Roy. Soc. (London) 117, 610 (1928).

To produce a wave equation which is both first order in time and relativistically invariant, Dirac constructed the matrix algebra which bears his name. As Dirac averred, this leads to an explanation of electron spin "without arbitrary assumptions." But one mysterious feature of the electron wavefunction seems to be left unexplained. Why is it a complex function? In Dirac's equation, which is largely determined by the requirement of relativistic invariance, why does an uninterpreted $(-1)^{\frac{1}{2}}$ appear explicitly? As the increasing theoretical importance of antiparticle conjugation tends to show, the appearance of this $(-1)^{\frac{1}{2}}$ is no triviality.

We submit that the $(-1)^{\frac{1}{2}}$ in Dirac's equation can be interpreted geometrically and that the reason for its appearance in physics is inseparable from that of spin. To appreciate this viewpoint it is necessary to understand the full geometrical significance of the Dirac algebra.

The metrical properties of space–time can be represented by introducing appropriate rules for the multiplication of space–time vectors. The result is a Clifford algebra which can be thought of as an algebra of directions in space–time. It is conveniently called the *real* Dirac algebra, because it is isomorphic to the algebra of Dirac matrices over the *real* numbers. The requirements of relativity can be satisfied by writing all physical equations in terms of the real Dirac algebra. Relativity provides no justification for the use of additional "complex" numbers such as those in the "complex Dirac algebra."

In contrast to the "complex matrix algebra of Dirac," the real Dirac algebra has a thoroughly geometrical significance. By re-expressing Dirac's theory in terms of the real Dirac algebra, we find that the $(-1)^{\frac{1}{2}}$ which appears explicitly in Dirac's equation acquires a geometrical interpretation as the generator of rotations in a spacelike plane. Moreover, the orientation of this plane is described by the electron "spin." In this way spin and "complex" numbers are combined in a single geometric entity.

The theory presented in this paper is algebraically isomorphic to Dirac's; it can be provided with an equivalent physical interpretation as well. It differs from Dirac's in that all its algebraic ingredients have a geometrical significance determined by properties we attribute to space–time. This produces a clarification of the Dirac theory, especially as regards the role of complex numbers. It also introduces new possibilities for modifying the theory. But to avoid prejudging the issues, a fairly conservative approach is adopted in this paper. Only a relatively minor amendment to the physical interpretation of the Dirac theory is suggested. In subsequent papers, upon the firm base of the Dirac theory, more general theories will be constructed which incorporate additional physical facts and exploit the geometrical and physical interpretations of $(-1)^{\frac{1}{2}}$ given here.

This paper is divided into three sections. In Sec. A the *real* Dirac algebra is briefly described. In Sec. B the algebraic expression for a *real* spinor field is given. For a Dirac particle, the wavefunction and its symmetries are provided with both physical and geometrical interpretations. The only departure from Dirac theory is to be found in a new and simple distinction between positive and negative energy states. This innovation requires a new expression for the charge–current density. However, since in the one-particle approximation this current is the same as Dirac's current, any consequences which this alteration may entail is manifested only in quantum electrodynamical calculations—somewhat beyond the scope of this paper. Finally, connection with the Dirac theory is made in Sec. C. Readers who are convinced that Sec. B is correct will not find it necessary to read Sec. C.

A. SPACE–TIME ALGEBRA

We use the geometric algebra developed by Hestenes.[2] However, without assuming prior familiarity with space–time algebra (STA), this section tries to give an explanation of the algebra which will be sufficient for the purposes of this paper. To do it succinctly, it is assumed that the reader is well acquainted with the matrix form of the Dirac algebra. He is then asked to re-interpret this algebra geometrically in a prescribed way. If the reader finds the treatment here excessively difficult to follow or too sparse to be satisfying, he is referred to the more detailed discussion given in Ref. 2.

1. Vectors

Physicists are accustomed to thinking of the γ_μ ($\mu = 0, 1, 2, 2$) as four by four matrices which are the four components of single world vector in space–time. Instead, the reader is asked to think of the γ_μ as a frame of four orthonormal vectors in space–time. Think of γ_0 as a unit vector in the forward light cone and the γ_i ($i = 1, 2, 3$) as a right-handed set of spacelike vectors. Thus, γ_0 is the unit normal to the (3-dimensional) spacelike hyperplane spanned by the γ_i. Moreover, any world vector $\mathbf{A}$ can be written as a linear combination of the γ_μ,

$$\mathbf{A} = A^\mu \gamma_\mu . \tag{1.1}$$

[2] D. Hestenes, *Space-Time Algebra* (Gordon and Breach Science Publishers, Inc., New York, 1966).

The A^μ are the components of the vector $\mathbf{A}$ with respect to the set of base vectors $\{\gamma_\mu\}$.

Although the reader is asked to think of the γ_μ as vectors instead of matrices, they still retain the multiplication rules possessed by the Dirac matrices. This makes new demands on our understanding of geometry. Presumably, the matrix multiplication is clear enough, but, in view of the geometrical significance of vectors, it would seem that, in order to justify multiplication of vectors, we must supply it with a geometric interpretation. This can, in fact, be done. Assuming that two vectors $\mathbf{A}$ and $\mathbf{B}$ can be multiplied "like matrices," the product $\mathbf{AB}$ can be understood geometrically by separating it into symmetric and antisymmetric parts:

$$\mathbf{AB} = \mathbf{A}\cdot\mathbf{B} + \mathbf{A}\wedge\mathbf{B}, \tag{1.2a}$$

$$\mathbf{A}\cdot\mathbf{B} = \tfrac{1}{2}(\mathbf{AB} + \mathbf{BA}), \tag{1.2b}$$

$$\mathbf{A}\wedge\mathbf{B} = \tfrac{1}{2}(\mathbf{AB} - \mathbf{BA}). \tag{1.2c}$$

A special notation has been introduced for the symmetric and antisymmetric products because each has a geometrical interpretation independent of the other. The quantity $\mathbf{A}\cdot\mathbf{B}$ is a scalar; the dot signifies the familiar *inner product* of vectors. The quantity $\mathbf{A}\wedge\mathbf{B}$ is commonly called a *bivector* or *2-vector*; the wedge signifies the *outer product* of vectors. Although the outer product is not so familiar to physicists, it was invented and interpreted geometrically many years before the invention of matrix algebra. For the purposes of this paper, it is not so important that the reader appreciate the geometrical interpretation of the outer product[3]; he need only believe that such an interpretation can be supplied, in order to appreciate that by (1.1) the different geometrical products $\mathbf{A}\cdot\mathbf{B}$ and $\mathbf{A}\wedge\mathbf{B}$ are united in a single product $\mathbf{AB}$. Thus, while the product $\mathbf{AB}$ of two vectors is not itself a vector, it is nevertheless composed of quantities with geometrical significance every bit as definite as that of vectors. Let the reader be assured that, in a similar way, the geometric character of the product of any number of vectors can be divined.

It may help the reader to see (1.2) written out for the base vectors γ_μ. Then Eq. (1.2a) becomes

$$\gamma_\mu\gamma_\nu = \gamma_\mu\cdot\gamma_\nu + \gamma_\mu\wedge\gamma_\nu. \tag{1.3a}$$

Because of the orthogonality of the γ_μ,

$$\gamma_\mu\gamma_\nu = \gamma_\mu\cdot\gamma_\nu, \quad \text{if} \quad \mu = \nu, \tag{1.3b}$$

$$\gamma_\mu\gamma_\nu = \gamma_\mu\wedge\gamma_\nu, \quad \text{if} \quad \mu \neq \nu. \tag{1.3c}$$

The set of inner products of the base vectors $\{g_{\mu\nu} \equiv \gamma_\mu\cdot\gamma_\nu\}$ is the so-called "metric tensor" expressed in the frame $\{\gamma_\mu\}$. Using the metric tensor, (1.2b) allows us to write down an equation which is familiar to everyone acquainted with the Dirac algebra,

$$g_{\mu\nu} = \tfrac{1}{2}(\gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu). \tag{1.4}$$

To display explicitly the signature we attribute to the metric of space–time, we write out the nonvanishing components of (1.4),

$$\gamma_0^2 = 1, \quad \gamma_i^2 = -1 \quad (i = 1, 2, 3). \tag{1.5}$$

A word of warning is in order to help the reader guard against misunderstanding. It is evident that the word "vector" is used in the algebraic sense (as an element of a linear space) rather than in the tensor sense (as a set of elements with certain transformation properties). The word "scalar" is also used in the algebraic sense (as an element of the number field over which a linear space is defined) rather than in the tensor sense (as an invariant of some group of transformations). For instance, we would call a component of the vector in (1.1) a scalar; it is the inner product of two vectors.

$$A_\mu = \mathbf{A}\cdot\gamma_\mu = A^\mu g_{\mu\nu}. \tag{1.6}$$

(Note that we use the usual convention of tensor analysis to raise and lower indices.) It is also worth saying that A_μ must be a real number (or function). We do not allow the field of complex numbers, for this would introduce a $(-1)^{\frac{1}{2}}$ without geometrical significance. Soon it should be clear that we already have many algebraic objects with a well-defined geometrical significance which can play the role of $(-1)^{\frac{1}{2}}$.

By taking all linear combinations (over the real numbers) of all linearly independent polynomials of the γ_μ, we obtain the elements of a geometric algebra with 16 linearly independent elements. We can call it the "*real* Dirac algebra" to emphasize its similarity to Dirac's matrix algebra except as regards the number field. Or, we can call it a "vector algebra" to emphasize that all the elements can be constructed out of vectors. Perhaps it is best to call it a "space–time algebra" to emphasize its use as an algebraic representation of the primitive geometrical properties of space–time.

Once we have the space–time algebra, we need only define differentiation to have a complete geometric calculus for space–time. Let $\{x^\mu;\ \mu = 0, 1, 2, 3\}$ be a set of inertial coordinates[4] in space–time. The fundamental vector differential operator, written as $\Box$, can

[3] The geometrical significance of the outer product is discussed in Pt. I of Ref. 2.

[4] Among other names, these coordinates are sometimes called Minkowski, pseudo-Euclidean, Euclidean, or Cartesian coordinates. We prefer the physical name, since in these coordinates the graph of the world line of a point particle with no forces acting on it is a straight line.

be introduced by the formula

$$\Box = \gamma^\mu \partial_\mu, \quad \text{where} \quad \partial_\mu \equiv \partial/\partial x^\mu. \tag{1.7}$$

This operator was first introduced into physics by Dirac.[1] In his effort to construct a "relativistically invariant" first-order differential equation for the electron wavefunction, he was led, in effect, to take the square root of the d'Alembertian $\Box^2$. The appropriate square root is $\Box$, as our notation indicates. But, by now the reader needs hardly be reminded of the difference between our interpretation of the γ^μ and Dirac's.

It is preferable to call $\Box$ the *gradient* operator, because the name agrees with common parlance when $\Box$ operates on scalars. For example, by (1.7), the gradient of the coordinate function x^μ is

$$\Box x^\mu = \gamma^\mu. \tag{1.8}$$

This shows that γ^μ is a vector pointing in the direction of maximum increase in x^μ; the magnitude of γ^μ indicates the rate of change of x^μ. It may be noted that the meaning of γ^μ is surreptitiously altered to the point where γ^μ is to be thought of as a vector field, assigning a tangent vector $\gamma^\mu(\mathbf{x})$ to each point $\mathbf{x}$ in space-time. However, to be consistent with the definition of inertial coordinates, γ^μ must be a constant vector field, i.e.,

$$\partial_\nu \gamma^\mu = 0. \tag{1.9}$$

So, in a sense, the γ^μ at different points are equivalent.

We are accustomed to think of the gradient of a scalar field as the normal to an "equipotential" surface. The gradient of a vector field $\mathbf{A}$ is interpreted differently. To get at this interpretation, we follow (1.2) and decompose $\Box\mathbf{A}$ into symmetric and antisymmetric parts.

$$\Box\mathbf{A} = \Box \cdot \mathbf{A} + \Box \wedge \mathbf{A}. \tag{1.10}$$

The scalar parts $\Box \cdot \mathbf{A}$ is the usual *divergence* of a vector, and $\Box \wedge \mathbf{A}$ is the *curl*. If $\mathbf{A}$ is the electromagnetic vector potential, then $F = \Box \wedge \mathbf{A}$ is the electromagnetic field strength.[5]

2. Conjugation

We are familiar with the operations of transpose, complex conjugation, and Hermitian conjugation in matrix algebra. Analogous "conjugation" operations can be defined for the real Dirac algebra. But first we must get more insight into the structure of the algebra.

Consider the product $\mathbf{A}_1\mathbf{A}_2 \cdots \mathbf{A}_r$ of r-vectors $\mathbf{A}_1, \cdots, \mathbf{A}_r$. The part of $\mathbf{A}_1\mathbf{A}_2 \cdots \mathbf{A}_r$ which is antisymmetric under the interchange of any two vectors is called an r-vector; it is a generalization of (1.2c) and is written $\mathbf{A}_1 \wedge \cdots \wedge \mathbf{A}_r$. This quantity vanishes if the vectors $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_r$ are linearly dependent. Since in space-time we can find at most four linearly independent vectors, we must have $r \leq 4$.

Let us call an element of the real Dirac algebra a d-number. It can be proved that any d-number is expressible as a linear combination of r-vectors simply by using linearly independent r-vectors constructed from the γ_μ and the convention that scalars be called 0-vectors. Thus, any d-number ψ can be written

$$\psi = \psi_S + \psi_V + \psi_B + \psi_T + \psi_P, \tag{2.1}$$

where the subscripts S, V, B, T, P mean, respectively, scalar (0-vector) part, vector (1-vector) part, bivector (2-vector) part, trivector or pseudovector (3-vector) part, pseudoscalar (4-vector) part. The decomposition (2.1) is analogous to the separation of a complex number into real and imaginary parts, which partly explains why ψ is sometimes called a "hypercomplex" number.

By reversing the products of all vectors in the Dirac algebra, we obtain from ψ a new d-number $\tilde{\psi}$.

$$\tilde{\psi} = \psi_S + \psi_V - \psi_B - \psi_T + \psi_P. \tag{2.2}$$

This is an invariant kind of "conjugation" in the sense that it is independent of any basis in the algebra.

The unit pseudoscalar is so important that we represent it by the special symbol i. We can express i in terms of the γ_μ,

$$i = \gamma_0 \wedge \gamma_1 \wedge \gamma_2 \wedge \gamma_3 = \gamma_0\gamma_1\gamma_2\gamma_3. \tag{2.3}$$

Reasons for using the symbol i for this quantity are apparent in the discussion that follows.

With the help of i we can define an operation which reverses the direction of all vectors in space-time. We call space-time conjugation[6] the operation which maps ψ into $\bar{\psi}$, where

$$\bar{\psi} = -i\psi i = \psi_S - \psi_V + \psi_B - \psi_T + \psi_P. \tag{2.4}$$

A d-number ψ is called *even* if $\bar{\psi} = \psi$ and *odd* if $\bar{\psi} = -\psi$.

The set of all even d-numbers forms a subalgebra which can be identified as the Pauli algebra. An inertial frame determines a right-handed set of base vectors σ_i $(i = 1, 2, 3)$ in the Pauli algebra by the definitions

$$\sigma_i \equiv \gamma_i\gamma_0. \tag{2.5}$$

[5] The equations of electrodynamics are formulated in terms of the real Dirac algebra in Pt. II of Ref. 2.

[6] The word "reflection" is reserved for an operation which inverts *points* of space-time. The more general term "conjugate" is used for this operation which reverses the direction of tangent vectors but leaves space-time points unchanged.

From (2.3), it is clear that

$$\sigma_1\sigma_2\sigma_3 = i. \tag{2.6}$$

It is useful to introduce special symbols for two other kinds of conjugation,

$$\psi^* \equiv \gamma_0\psi\gamma_0, \tag{2.7}$$

$$\psi^\dagger \equiv \gamma_0\tilde{\psi}\gamma_0. \tag{2.8}$$

Clearly ψ^* and $\psi^\dagger$ depend on the choice of γ_0. The dagger symbol in (2.8) is particularly appropriate because $\psi^\dagger$ corresponds to the Hermitian conjugate of ψ in the usual matrix representations of the Dirac algebra.

Notice that the operation (2.7) changes $\{\sigma_i\}$ and also $\{\gamma_i\}$ into left-handed frames without affecting γ_0. For this reason, the operation which takes ψ into ψ^* is called space conjugation.[6]

3. Frames and Lorentz Rotations

It can be proved that every field of orthonormal frames $\{\mathbf{e}_\mu(x)\}$ with the same orientation as $\{\gamma\}$ can be obtained from $\{\gamma_\mu\}$ by a field of *Lorentz rotations*[7] expressed as follows:

$$\mathbf{e}_\mu(\mathbf{x}) = R(\mathbf{x})\gamma_\mu\tilde{R}(\mathbf{x}), \tag{3.1}$$

where R is a d-number field with the properties[8]

$$R\tilde{R} = 1, \quad \text{and} \quad R = \bar{R}. \tag{3.2}$$

It may help the reader to observe that (3.1) has the same form as a similarity transformation, which in Dirac's matrix algebra is used to transform from one matrix representation of the γ_μ to another. Here, of course, the interpretation is different, and the operator R is allowed to be a function of the space–time point $\mathbf{x}$. Except for a sign, R is uniquely determined by the $\mathbf{e}_\mu$ and the γ_μ.[9] Because, by (3.1), R determines the transformation of the γ_μ into the $\mathbf{e}_\mu$, *it is sometimes convenient to refer to R itself as a Lorentz rotation.*

The conditions (3.2) imply that R can be written in the form

$$R = \pm e^{-\frac{1}{2}\phi}, \tag{3.3}$$

where ϕ is a bivector field. Conversely, every d-number of the form (3.3) represents a unique Lorentz rotation.

We call (3.1) a *spatial rotation* if $R = R^*$ and a *special timelike rotation* if $R = R^\dagger$. Any Lorentz rotation can be expressed as a spatial rotation followed by a special timelike rotation. This means that R can always be written in the form

$$R = e^{-\frac{1}{2}\chi}e^{-\frac{1}{2}i\theta}, \tag{3.4}$$

where χ and θ are bivectors satisfying $\chi^* = -\chi$, $\theta^* = -\theta$.

Any Lorentz transformation can be expressed as some combination of space–time conjugation (2.4), space conjugation (2.7), and a Lorentz rotation (3.1).

B. SPINOR FIELDS

At this point, it is hoped that the reader is sufficiently prepared to grapple with a reformulation of Dirac's electron theory in terms of our "space–time calculus." This theory can, in a certain sense, be derived from Dirac's theory. But, the derivation is tedious. Moreover, the end result is relatively simple and can be set forth and understood as a physical theory in its own right. Therefore, we present and discuss the reformulated theory directly. The derivation has been relegated to the end of this paper where it can be studied separately.

The electron wavefunction has certain suggestive features which may have a physical significance transcending the Dirac equation, so we begin by discussing the wavefunction and its interpretation without reference to a "wave equation." Then we write down the reformulated Dirac equation and discuss some of the implications it has for the interpretation of the wavefunction. The theory given here is algebraically isomorphic to Dirac's theory. However, it differs in that all the algebraic quantities involved have a definite geometrical significance determined by the properties already ascribed to space–time in our construction of the space–time algebra. This, in turn, suggests certain modifications of the physical theory, although, except for a relatively minor exception, we resist the temptation to pursue them here and are content with showing how the physical features of the Dirac theory are to be formulated in our new language.

4. The Wavefunction

We begin simply by writing down the most succinct expression for a spinor field,

$$\psi(\mathbf{x}) = e^{\frac{1}{2}\mu(\mathbf{x})}, \tag{4.1}$$

where $\mu(\mathbf{x})$ is an *even d-number* field. Let us call $\psi(\mathbf{x})$ a *real* spinor field to emphasize that it is expressed in terms of the real Dirac algebra and so is susceptible to a geometrical interpretation. "Real spinors" are equivalent to "Dirac spinors" in the sense that the two can be put into one-to-one correspondence. This is proved in Sec. C, but a simple check on the assertion can be made immediately by noting that since μ is an even d-number it has eight linearly independent components. This agrees with the fact that a Dirac

[7] Lorentz rotations are discussed in detail in Part IV of Ref. 2.
[8] We usually suppress the argument x.
[9] This is shown explicitly by Eq. (17.24) of Ref. 2.

spinor is a column matrix with four complex components.

While (4.1) may well be the most appropriate expression of a spinor field for certain kinds of mathematical analysis, to explain its geometrical and physical significance we must decompose a spinor into its "geometrically distinct" parts. Therefore, write μ in the form

$$\mu(\mathbf{x}) = \alpha(\mathbf{x}) + i\beta(\mathbf{x}) + \phi(\mathbf{x}), \tag{4.2}$$

where α and β are scalars and ϕ is a bivector in the Dirac algebra. Also define

$$\rho(\mathbf{x}) = e^{\alpha(\mathbf{x})} > 0 \tag{4.3}$$

and use (3.3) to write ψ in the form[8]

$$\psi = (\rho e^{i\beta})^{\frac{1}{2}} R. \tag{4.4}$$

Now observe that ψ determines four "current vectors" $\mathbf{J}_\mu$ as follows:

$$\mathbf{J}_\mu \equiv \psi \gamma_\mu \tilde{\psi} = \rho \mathbf{e}_\mu . \tag{4.5}$$

The $\mathbf{e}_\mu$ are the orthonormal vectors defined by (3.1); $\mathbf{J}_0$ is a "positive" timelike vector,[10] equivalent to the *probability current* density of Dirac theory. Following Dirac, we must interpret $\rho(\mathbf{x})$ as the probability density in the instantaneous rest frame at the point $\mathbf{x}$. More briefly, $\rho(\mathbf{x})$ is the *proper probability density*. To conserve probability, the divergence of $\mathbf{J}_0$ must vanish, i.e.,

$$\Box \cdot \mathbf{J}_0 = 0. \tag{4.6}$$

This provides a restriction on the wave equation for ψ.

The vector $\mathbf{J}_3$ describes the orientation of the electron "spin." In commonly used language, it is "the expectation value of the electron spin." The essential property of the "spin vector" is that it is orthogonal to $\mathbf{J}_0$. The labeling of the "spin vector" with the subscript three, rather than with one or two is merely a convention.

The vectors $\mathbf{J}_1$ and $\mathbf{J}_2$ do not inherit independent physical interpretations from the Dirac theory. However, the plane containing $\mathbf{J}_1$ and $\mathbf{J}_2$, which can be described by the bivector $\mathbf{J}_1 \wedge \mathbf{J}_2$, does have physical significance. The abstract complex plane in which the phase factor of the Dirac wavefunction is defined corresponds to this plane. The $(-1)^{\frac{1}{2}}$ in the electron wavefunction can now be interpreted geometrically as the generator of rotations in the plane of $\mathbf{J}_1 \wedge \mathbf{J}_2$. Therefore, we propose that this plane be called the *phase plane* of the wavefunction. It can also fairly be called the *spin plane*, because it is completely determined by the spin and current vectors. This can be seen easily by noting that, by virtue of (2.3), the generator $\mathbf{e}_2\mathbf{e}_1$ of rotations in the phase plane can be written

$$\mathbf{e}_2\mathbf{e}_1 = i\mathbf{e}_0\mathbf{e}_3 . \tag{4.7}$$

The notions of spin and phase are now inextricably united. The "phase" describes the magnitude of a rotation in the phase plane and the "spin" describes the orientation of the phase plane.

So far we understand the factors ρ and R in the wavefunction (4.4). The scalar ρ describes the probability density, and the Lorentz rotation R describes in a single unit the orientation of the probability current vector and the spin and the phase of the wavefunction. But what can be the meaning of the factor $e^{i\beta}$? Evidently, it does not play a role in determining the observables just mentioned because it does not appear in the expression (4.5) for the current vectors. Inversely, given the "observables" $\mathbf{J}_\mu$, we can find ρ and R, but to determine ψ uniquely we need one more bilinear function of ψ, namely,

$$\psi\tilde{\psi} = \rho e^{i\beta}. \tag{4.8}$$

Now there is one "observable" appearing in the Dirac theory of which we have given no account, namely, the sign of the energy. This "observable" is described by the quantity $e^{i\beta}$. We suppose that $e^{i\beta}$ describes the relative admixture of positive and negative energy components of ψ. More precisely, we take $e^{i\beta(\mathbf{x})}$ to be the relative probability of observing a particle at the point $\mathbf{x}$; so that if $e^{i\beta} = 1$ everywhere, ψ describes a pure *one-particle* state, and if $e^{i\beta} = -1$ everywhere ψ describes a pure *antiparticle* state. In view of the interpretation we have given to ρ the scalar part of $\psi\tilde{\psi}$,

$$(\psi\tilde{\psi})_S = \rho \cos\beta, \tag{4.9}$$

can be interpreted as the *proper particle density* of the spinor field ψ. Thus, a negative value of this quantity indicates the likelihood of observing an antiparticle. Of course, in a quantized version of our theory all the physical observables we have discussed are vacuum expectation values of quantum field operators.

This is a good place to summarize once again the relation of our theory to Dirac's. The theories are equivalent algebraically, but our theory admits a geometrical interpretation of physical quantities which is absent in Dirac's theory. We can, if we wish, give a physical interpretation of our theory which is precisely equivalent to Dirac's. However, the geometric significance which our theory endows to physical quantities suggests other possibilities. Just the same, in this paper we proffer only a relatively

[10] "Positive" means that $\mathbf{J}_0$ is in the forward light cone. This then can be conveniently expressed by $\mathbf{J}_0 \cdot \gamma_0 > 0$.

minor amendment to Dirac's interpretation by identifying a particular factor in the wavefunction as responsible for the admixture of positive and negative energy states. In all honesty, it must be admitted that this idea seems to produce some difficulty with the "superposition principle" which is not fully understood. On the other hand, we encounter more to be said in its favor when we study the Dirac equation in the next section.

With the notions we already have at our disposal, we can discuss the nonrelativistic limit of the electron wavefunction without reference to a specific wave equation. Earlier we identified the γ_μ with an inertial frame. To "fix the phase" of ψ, we can identify the γ_μ in (4.5) with the "laboratory frame." Now suppose the electron is in a pure positive energy state, so $e^{i\beta} = 1$, and use (3.3) to write the wavefunction in the form

$$\psi = \rho^{\frac{1}{2}} e^{-\frac{1}{2}\phi}. \tag{4.10}$$

The factor $\rho^{\frac{1}{2}}$ is the amplitude of the wavefunction and the bivector $\frac{1}{2}\phi$ may be thought of as a "relativistic phase." The "relativistic phase factor" can be decomposed by using (3.4) to express it as a spatial rotation $e^{-\frac{1}{2}i\theta(\mathbf{x})}$ in the laboratory frame followed by a Lorentz transformation $e^{-\frac{1}{2}\chi(\mathbf{x})}$ which, without rotation, takes the laboratory frame into the instantaneous rest frame of the electron:

$$\psi = \rho^{\frac{1}{2}} e^{-\frac{1}{2}\chi} e^{-\frac{1}{2}i\theta}. \tag{4.11}$$

In the nonrelativistic approximation, the factor $e^{-\frac{1}{2}\chi}$ is negligible, so the state of the electron can be described by the effective wavefunction

$$\psi' = \rho^{\frac{1}{2}} e^{-\frac{1}{2}i\theta}. \tag{4.12}$$

This is equivalent to the Pauli wavefunction. The bivector $\frac{1}{2}i\theta(\mathbf{x})$ has an orientation which describes the orientation of the electron "spin" and magnitude which is just the "scalar phase" of the wavefunction. If, as a further approximation, it is supposed that the precession of the spin plane is negligible, then we can write

$$\tfrac{1}{2}i\theta(\mathbf{x}) = i\sigma_3\delta(\mathbf{x}) = \gamma_2\gamma_1\delta(\mathbf{x}), \tag{4.13}$$

where $\delta(x)$ is a "scalar phase." The effective wavefunction can be written as

$$\psi''(\mathbf{x}) = \rho^{\frac{1}{2}} \exp\left[-\gamma_2\gamma_1\delta(\mathbf{x})\right]. \tag{4.14}$$

This is equivalent to the Schrödinger wavefunction. However, in (4.14) the phase factor is to be thought of as a rotation in a spacelike plane (the spin plane) rather than as a function in some abstract complex space. In the "Schrödinger approximation" the orientation of the spin plane is a constant.

5. Symmetries of the Wavefunction

For ψ the equivalent of the *Dirac equation* for an electron is

$$\Box\psi = (m\psi\gamma_0 + e\mathbf{A}\psi)i\sigma_3. \tag{5.1}$$

Here, m is the mass, e is the charge of the electron, and $\mathbf{A}$ is the electromagnetic vector potential. In this section, we discuss certain symmetries of ψ associated with the Dirac equation. A *symmetry* of a spinor field $\psi(\mathbf{x})$ is a mapping of the field ψ onto itself which preserves the wave equation for ψ or changes it in a definite and physically meaningful way.

We have already seen that ψ determines a frame of tangent vectors $\mathbf{J}_\mu(\mathbf{x}) = \psi\gamma_\mu\tilde{\psi}$ at each point $\mathbf{x}$ of space–time, and inversely, except for a factor $e^{i\beta}$, the $\mathbf{J}_\mu$ determine ψ. A transformation on ψ induces a corresponding transformation on the $\mathbf{J}_\mu$. Because we are familiar with the representation of directions by vectors, a transformation of the $\mathbf{J}_\mu$ can be interpreted geometrically. This enables us to give a geometric interpretation to transformations of ψ.

There are two distinct kinds of geometrical transformations. First, there is a *transformation of the tangent vectors* $\mathbf{J}_\mu(\mathbf{x})$ at a point $\mathbf{x}$ of space–time into a new set of vectors $\mathbf{J}'_\mu(\mathbf{x})$ at the same point $\mathbf{x}$. As we see, charge conjugation is a transformation of this kind. Second, there is a *transformation of points* in space–time, wherein the tangent vectors at a point $\mathbf{x}$ of space–time are mapped into "equivalent" vectors at a *different* point $\mathbf{x}'$. Displacements are the simplest transformations of this kind. Symmetries of a spinor field can be interpreted geometrically as some combination of these two kinds of transformations.

First note that (5.1) is invariant under

$$\psi \to \psi\gamma_0, \tag{5.2}$$

which induces

$$\mathbf{J}_0 \to \mathbf{J}_0, \quad \mathbf{J}_i \to -\mathbf{J}_i. \tag{5.3}$$

This transformation tells us that the Dirac equation does not distinguish (or couple) even and odd spinor fields—a fact which is not discovered and so is not interpreted in the usual form of the Dirac theory. Because of this equivalence of even and odd fields, we may, without further comment, confine the rest of our discussion to transformations which leave ψ even.

An operation which changes the sign of the electromagnetic coupling while leaving the rest of the Dirac equation invariant is called *charge conjugation*. By inspection, this can be accomplished only by multiplying (5.1) on the right by some linear combination of σ_1 and σ_2. Therefore, charge conjugation has the general form

$$\begin{aligned} C:\ \psi \to \psi_C &= \psi\sigma_2 \exp(i\sigma_3\phi_C) \\ &= \psi \exp(-i\sigma_2\tfrac{1}{2}\pi) \exp(i\sigma_3\phi_C i), \end{aligned} \tag{5.4}$$

where ϕ_C is a constant scalar. This induces a rotation of π of the $\mathbf{J}_i$ about some axis orthogonal to $\mathbf{J}_3$. For instance, if $\phi_C = 0, \pi$,

$$\mathbf{J}_0 \to \mathbf{J}_0, \quad \mathbf{J}_1 \to -\mathbf{J}_1, \quad \mathbf{J}_2 \to \mathbf{J}_2, \quad \mathbf{J}_3 \to -\mathbf{J}_3. \tag{5.5}$$

We can interpret (5.4) itself as a rotation of the $\mathbf{J}_\mu$, but this is not the whole truth, because (5.4) does not leave $\psi\tilde{\psi}$ invariant,

$$\psi_C\tilde{\psi}_C = -\psi\tilde{\psi}. \tag{5.6}$$

This is an important point, for it shows that charge conjugation changes the sign of the particle density (4.7). To be more specific, note that the factor

$$\exp(-i\sigma_2 \tfrac{1}{2}\pi) = -i\sigma_2$$

in (5.4) reverses the direction of the spin vector $\mathbf{J}_3$ in (5.5) but does not affect the expression $\psi\tilde{\psi}$, while the factor i changes the sign of the particle density but does not affect $\mathbf{J}_3$.

In Minkowski space-time, a point $\mathbf{x}$ with coordinates x^μ can be represented by a "position vector" $\mathbf{x} = x^\mu\gamma_\mu$. "Space reflection" of points in the hyperplane orthogonal to γ_0 can be represented by the mapping

$$P\colon \mathbf{x} \to \mathbf{x}^* = \gamma_0\mathbf{x}\gamma_0. \tag{5.7}$$

The form of the Dirac equation will be preserved if the point transformation (5.7) induces the following transformation of fields:

$$P\colon \mathbf{A}(\mathbf{x}) \to \mathbf{A}^*(\mathbf{x}^*), \tag{5.8}$$

$$P\colon \psi(\mathbf{x}) \to \psi_P(\mathbf{x}) = \psi^*(\mathbf{x}^*)e^{i\sigma\phi_P}. \tag{5.9}$$

We must have $\phi_P = 0$ or π, if we require that successive applications of (5.9) map $\psi(\mathbf{x})$ back to itself. The transformation P is commonly called *parity conjugation*. It induces the transformations,

$$P\colon \mathbf{J}_0(\mathbf{x}) \to \mathbf{J}_0^*(\mathbf{x}^*), \tag{5.10}$$

$$P\colon \mathbf{J}_3(\mathbf{x}) \to -\mathbf{J}_3^*(\mathbf{x}^*). \tag{5.11}$$

Because the components of the tangent vector $\mathbf{J}_0$ in (5.10) transform in the same way as the components of the position vector $\mathbf{x}$ in (5.7), $\mathbf{J}_0$ is usually said to "transform as a vector" under parity conjugation. Because $\mathbf{J}_3$ transforms with opposite sign it is said to "transform as an axial vector." It is not necessary for us to subscribe to this distinction between "vector" and "axial vector," first, because it is not needed to distinguish $\mathbf{J}_0$ from $\mathbf{J}_3$, and second, because it is purely gratuitous, for by using (5.2) we can keep $\mathbf{J}_0$ fixed and change the sign of $\mathbf{J}_3$ at will without affecting the Dirac equation.

A time-reversal transformation T can be constructed in analogy to parity conjugation. The transformation of the position vector is

$$T\colon \mathbf{x} \to \bar{\mathbf{x}}^* = -\mathbf{x}^*. \tag{5.12}$$

The corresponding transformation of the wavefunction is

$$T\colon \psi(\mathbf{x}) \to \psi_T(\mathbf{x}) = i\psi^*(-\mathbf{x}^*)\exp(i\sigma_3\phi_T). \tag{5.13}$$

As long as $\mathbf{A}$ is unique, the Dirac equation determines ψ to within a factor of a constant Lorentz rotation which can be fixed by a choice of the γ_μ, so that (4.5) yields definite $\mathbf{J}_\mu$. But Maxwell's equation determines A only up to a *gauge transformation*,

$$\mathbf{A} \to \mathbf{A}' = \mathbf{A} + \Box\chi. \tag{5.14}$$

If the Dirac equation is to be left invariant, then (5.14) must be accompanied by the transformation

$$\psi \to \psi' = \psi e^{i\sigma_3\chi}. \tag{5.15}$$

This induces the transformation

$$\begin{aligned} &\mathbf{J}_0 \to \mathbf{J}_0, \quad \mathbf{J}_3 \to \mathbf{J}_3, \\ &\mathbf{J}_1 \to \mathbf{J}_1\cos 2\chi - \mathbf{J}_2\sin 2\chi, \\ &\mathbf{J}_2 \to \mathbf{J}_2\cos 2\chi + \mathbf{J}_1\sin 2\chi, \end{aligned} \tag{5.16}$$

which shows that electromagnetic interactions do not distinguish $\mathbf{J}_1$ from $\mathbf{J}_2$.

Closely related to the electromagnetic gauge transformations are the conservation laws found by calculating the divergence of the $\mathbf{J}_\mu$ using (5.1) to evaluate

$$[(\Box\psi)\gamma_\mu\tilde{\psi} + \psi\gamma_\mu(\Box\psi)\tilde{}\,]_S = (\Box\psi\gamma_\mu\tilde{\psi})_S = \Box\cdot\mathbf{J}_\mu,$$

we find that

$$\Box\cdot\mathbf{J}_\mu = ei\mathbf{A}\wedge\mathbf{e}_3\wedge\mathbf{e}_0\wedge\mathbf{J}_\mu. \tag{5.17}$$

The expression on the right vanishes identically if $\mu = 0$ or 3. Therefore, the "probability current" $\mathbf{J}_0$ and the "spin current" $\mathbf{J}_3$ are conserved.

If, as has been suggested, the expression (4.9) is interpreted as the proper particle density, then the charge–current density $\mathbf{J}(\mathbf{x})$ must be

$$\mathbf{J}(\mathbf{x}) = e\mathbf{J}_0(\mathbf{x})\cos\beta(\mathbf{x}). \tag{5.18}$$

If the total charge is to be conserved, then the divergence of $\mathbf{J}$ must vanish, i.e.,

$$\Box\cdot\mathbf{J} = 0. \tag{5.19}$$

Since, as (5.17) shows, the Dirac equation implies that the divergence of $\mathbf{J}_0$ also vanishes, (5.19) implies that

$$\mathbf{J}_0\cdot\Box\beta = 0. \tag{5.20}$$

Let us see what this condition means.

Imagine that we begin with a charge distribution $e\rho(\mathbf{x})$ in the rest frame at the point $\mathbf{x}$. As the electromagnetic interaction is turned on, there is a redistribution of charge due to the creation of particle–antiparticle pairs. The new charge density is $e\rho(\mathbf{x})\cos\beta(\mathbf{x})$. Thus, insofar as the problem of renormalization is to calculate the redistribution of charge due to interaction, it is the problem of calculating $\beta(\mathbf{x})$. Equation (5.20) says that the gradient of β is orthogonal to $\mathbf{J}_0$, which is just the reasonable condition that the redistribution of charge is orthogonal to the direction of the current.

Heretofore $\mathbf{J}_0$ has been the only invariant of the Dirac field used in the definition of the charge–current density. Naturally, difficulties were encountered because $\mathbf{J}_0$ is a positive timelike vector and does not change sign under charge conjugation. In the past these difficulties have been surmounted by reinterpreting the wavefunction as a quantized field operator. By contrast, the current vector $\mathbf{J}$ defined in (5.18) achieves the proper behavior under charge conjugation without resorting to the strenuous expedient of second quantization. Whether or not the added condition on the wavefunction which this approach entails will conspire to give a correct account of renormalization is a matter ultimately to be determined by calculation. Unfortunately, such a calculation is a many-particle problem which is beyond the scope of this paper.

6. Expansion in Plane Waves

The Dirac equation for a free field is

$$\Box\psi = m\psi\gamma_0 i\sigma_3. \tag{6.1}$$

The two positive energy solutions to this equation can be written

$$\psi_i^{(+)} = u_i e^{i\sigma_3\mathbf{p}\cdot\mathbf{x}}, \tag{6.2}$$

where the u_i ($i = 1, 2$) are constant Lorentz rotations. Substituting (6.2) into (6.1) we find

$$\mathbf{p}u_i = mu_i\gamma_0$$

or, since $u_i\tilde{u}_i = 1$,

$$\mathbf{p} = mu_i\gamma_0\tilde{u}_i. \tag{6.3}$$

The solutions u_1 and u_2, described as "spin up" and "spin down," respectively, can be written in the explicit form

$$u_1 = e^{-\frac{1}{2}\chi}, \tag{6.4}$$

$$u_2 = e^{-\frac{1}{2}\chi}i\sigma_2. \tag{6.5}$$

Corresponding to these solutions, the spin vector is

$$\mathbf{e}_3 = u_1\gamma_3\tilde{u}_1 = -u_2\gamma_3\tilde{u}_2. \tag{6.6}$$

The negative energy solutions to (6.1) can be written

$$\psi_i^{(-)} = v_i e^{i\sigma_3\mathbf{p}\cdot\mathbf{x}}, \tag{6.7}$$

where $i = 1, 2$, and

$$v_1 = iu_1^* = ie^{\frac{1}{2}\chi}, \tag{6.8}$$

$$v_2 = iu_2^* = ie^{\frac{1}{2}\chi}i\sigma_2. \tag{6.9}$$

Substituting (6.7) into (6.1), we get

$$\mathbf{p}v_i = mv_i\gamma_0, \quad \mathbf{p}u_i^* = -mu_i\gamma_0,$$

or

$$\mathbf{p} = -m(u_i\gamma_0\tilde{u}_i)^*. \tag{6.10}$$

Since $u_i\gamma_0\tilde{u}_i$ is a positive timelike vector, the energy $p_0 = \mathbf{p}\cdot\gamma_0$ is negative in (6.10).

A general "free-particle" solution $\psi(x)$ of the Dirac equation can be expanded in terms of the plane-wave solutions,

$$\psi(\mathbf{x}) = \int\frac{d^3p}{(2\pi)^{\frac{3}{2}}}\left(\frac{m}{p_0}\right)^{\frac{1}{2}}\sum_{i=1}^{2}(u_i a_i e^{i\sigma_3\mathbf{p}\cdot\mathbf{x}} + v_i b_i e^{-i\sigma_3\mathbf{p}\cdot\mathbf{x}}). \tag{6.11}$$

In this expression $p_0 = \mathbf{p}\cdot\gamma_0 > 0$, and the a_i (and b_i) are "complex numbers" of the form

$$a_i = \alpha_i + i\sigma_3\beta_i,$$

where α_i and β_i are scalars. It is clear that here the role of $(-1)^{\frac{1}{2}}$ in the usual form of the Dirac theory is taken over by $i\sigma_3 = \gamma_1\gamma_2$, the generator of rotations in the 12-plane. The plane-wave "phase factor" $\exp(i\sigma_3\mathbf{p}\cdot\mathbf{x})$ is literally a rotation in the plane containing γ_1 and γ_2. It induces a rotation in the plane of $\mathbf{J}_1$ and $\mathbf{J}_2$. The expansion (6.11) is basically an expansion of $\psi(\mathbf{x})$ in terms of rotations in the 12-plane.

C. CONNECTION WITH DIRAC THEORY

7. Bilinear Forms

In STA the Dirac equation for an electron interacting with an electromagnetic field was written

$$\Box\Psi = (m + e\mathbf{A})\Psi i, \tag{7.1}$$

where

$$\Psi = \Psi\tfrac{1}{2}(1 + \sigma_3). \tag{7.2}$$

Equation (7.1) is algebraically equivalent to the Dirac equation in its matrix form.[11] But there is a significant

[11] A spinor can be defined as a solution to the equations $\Psi\tilde{\Psi} = \Psi i\tilde{\Psi} = 0$. For every such solution, there is a bivector σ_3 (with $\sigma_3^2 = 1$) so that (7.2) holds. Equation (7.1) implies that σ_3 must be constant. The definition of spinor given here is equivalent to the definition (given in Ref. 2) of a spinor as an element of a minimal ideal in the Dirac algebra. Actually, there are an infinite number of minimal ideals. The electron wavefunction is an element of a particular minimal ideal characterized by σ_3, as (7.2) shows. As our analysis in this paper shows, the significance of σ_3 is to be found in its relation to the spin of the electron.

difference in that the uninterpreted $(-1)^{\frac{1}{2}}$ which appears in the Dirac theory has been replaced by a geometrically significant "root of minus one," namely, the pseudoscalar i. This substitution in no way abrogates the well-known and verified features of the Dirac equation. However, it does give all elements of the Dirac equation a well-defined geometrical meaning, and it is just this feature which enables us to replace Ψ by the simple wavefunction ψ which we have already discussed in detail. Actually, as we have seen, it is $i\sigma_3$ rather than i which is properly identified as the $(-1)^{\frac{1}{2}}$ which appears in the Dirac and Schrödinger equations. In the form (7.1), $i\sigma_3$ and i seem indistinguishable for, according to (7.2), $\Psi i = \Psi i\sigma_3$. But as we take Ψ apart, the difference between these two roots of minus one will become important.

We begin our study of Ψ by analyzing the structure of certain bilinear forms of Ψ which are given physical interpretation in the Dirac theory. The vector part of $\Psi\gamma_0\tilde{\Psi}$ is just that vector which was identified by Dirac as the probability current $\mathbf{J}_0$ of the electron field. Note that $(\Psi\gamma_0\tilde{\Psi})^{\sim} = (\Psi\gamma_0\tilde{\Psi})$, so by virtue of (2.2) $\Psi\gamma_0\tilde{\Psi}$ has only scalar, vector, and pseudoscalar parts. Hence, we can write

$$\Psi\gamma_0\tilde{\Psi} = \Psi\Psi^{\dagger}\gamma_0 = \mathbf{J}_0 + \rho e^{i\beta}. \tag{7.3}$$

Here the scalar part is $\rho\cos\beta$ and the pseudoscalar part is $i\rho\sin\beta$.

Now note that $(\tilde{\Psi}\Psi)^{\sim} = \tilde{\Psi}\Psi$ so that $\tilde{\Psi}\Psi$ can have only scalar, vector, and pseudoscalar parts. But, because of (7.2),

$$\tilde{\Psi}\Psi = \tfrac{1}{2}(1 - \sigma_3)\tilde{\Psi}\Psi\tfrac{1}{2}(1 + \sigma_3),$$

which implies that the scalar and pseudoscalar parts of $\tilde{\Psi}\Psi$ must vanish. By taking the inner product of $\tilde{\Psi}\Psi$ with γ_0 and comparing with (7.3), we find

$$\gamma_0 \cdot (\tilde{\Psi}\Psi) = (\gamma_0\tilde{\Psi}\Psi)_S = (\Psi\gamma_0\tilde{\Psi})_S = \rho\cos\beta.$$

It follows that

$$\tilde{\Psi}\Psi = \rho\cos\beta\gamma_0(1 + \sigma_3) = \rho(\gamma_0 - \gamma_3)\cos\beta. \tag{7.4}$$

Obviously, $\tilde{\Psi}\Psi$ is a null vector.

Now, using (7.3) and (7.4), we can evaluate $\Psi\gamma_0\tilde{\Psi}\Psi$ in two different ways,

$$\Psi\gamma_0\tilde{\Psi}\Psi = (\mathbf{J}_0 + \rho e^{i\beta})\Psi = \Psi 2\rho\cos\beta.$$

Hence

$$(\mathbf{J}_0 - \rho e^{-i\beta})\Psi = 0. \tag{7.5}$$

Multiplying (7.5) on the left by $(\mathbf{J}_0 + \rho e^{i\beta})$, we find

$$\mathbf{J}_0^2 = \rho^2. \tag{7.6}$$

If $\mathbf{J}_0^2 \neq 0$ then $\rho^2 \neq 0$ and we can write

$$\mathbf{J}_0 = \rho\mathbf{v}, \tag{7.7}$$

where $\mathbf{v}$ is a unit timelike vector. Let us follow out the implications of (8.7) and afterwards, at the end of Sec. 8, discuss the case when $\mathbf{J}_0^2 = 0$ but $\mathbf{J}_0 \neq 0$.

If ρ is identified with the proper probability of the field, then $\rho > 0$. It follows $\mathbf{J}_0 > 0$ (i.e., when $\mathbf{J}_0$ does not vanish, it is a positive timelike vector).[12] We now write (7.3) in the form

$$\Psi\gamma_0\tilde{\Psi} = \rho(e^{i\beta} + \mathbf{v}), \tag{7.8a}$$

where

$$\rho > 0, \quad \mathbf{v} > 0, \quad \mathbf{v}^2 = 1. \tag{7.8b}$$

We can also construct bivector and trivector quantities which are bilinear in Ψ. Observe that, by virtue of (2.2), $(\Psi i\gamma_0\tilde{\Psi})^{\sim} = -\Psi\gamma_0\tilde{\Psi}$ shows that $\Psi i\gamma_0\tilde{\Psi}$ can have only bivector and trivector parts. By using (7.4) and (7.8) to reduce $\Psi i\gamma_0\tilde{\Psi}\Psi\gamma_0\tilde{\Psi}$ in two different ways, we find that $\Psi i\gamma_0\tilde{\Psi}$ can be written in the following form:

$$\Psi i\gamma_0\tilde{\Psi} = \rho i\mathbf{s}(\mathbf{v}e^{i\beta} + 1), \tag{7.9a}$$

where the only new quantity $\mathbf{s}$ is a spacelike vector orthogonal to $\mathbf{v}$, i.e.,

$$\mathbf{s}^2 = -1, \quad \mathbf{s}\cdot\mathbf{v} = 0. \tag{7.9b}$$

The vector $\mathbf{s}$ describes the orientation of the electron spin.

8. Structure of Dirac Spinors

We can take advantage of the fact that, because $\mathbf{v}$ and $\mathbf{s}$ are orthogonal unit vectors, there exists a field of Lorentz rotations which enables us to identify $\mathbf{v}$ with the $\mathbf{e}_0$ and $\mathbf{s}$ with the $\mathbf{e}_3$ in Eq. (3.1). We write

$$\mathbf{v} = \mathbf{e}_0 = R\gamma_0\tilde{R}, \tag{8.1}$$

$$\mathbf{s} = \mathbf{e}_3 = R\gamma_0\tilde{R}. \tag{8.2}$$

So far, since we have not specified $\mathbf{e}_1$ and $\mathbf{e}_2$, R is not uniquely determined; Eqs. (8.1) and (8.2) are unchanged if R is replaced by $R' = Re^{i\sigma_2\chi}$, where χ is any scalar function.

Now define ψ by

$$\psi = (\rho e^{i\beta})^{\frac{1}{2}}R \tag{8.3}$$

and write

$$\Psi = \psi U. \tag{8.4}$$

Using (7.2) and (7.5), we can write

$$\tfrac{1}{2}(1 + \mathbf{v})e^{-\frac{1}{2}i\beta}\Psi\tfrac{1}{2}(1 + \sigma_3) \tag{8.5}$$

which, when translated into a condition on U, becomes

$$U = \tfrac{1}{2}(1 + \gamma_0)U\tfrac{1}{2}(1 + \sigma_3). \tag{8.6}$$

By separating U into an even part $U_E = \frac{1}{2}(U + \bar{U})$ and odd part $U_O = \frac{1}{2}(U - \bar{U})$, we see that (8.6) can

[12] See Theorem 6 of Appendix B in Ref. 2.

be written

$$U = \tfrac{1}{2}(1 + \gamma_0)M\tfrac{1}{2}(1 + \sigma_3), \tag{8.7}$$

where $M = U_E + \gamma_0 U_O$ is even. By separating M into a part $M_1 = \frac{1}{2}(M + M^*)$ which commutes with γ_0, and a part $M_2 = \frac{1}{2}(M - M^*)$ which anticommutes with γ_0, we see that

$$\begin{aligned} U &= \tfrac{1}{2}(1 + \gamma_0)(M_1 + M_2\sigma_3)\tfrac{1}{2}(1 + \sigma_3) \\ &= \tfrac{1}{4}(M_1 + M_2\sigma_3)(1 + \gamma_0)(1 + \sigma_3). \end{aligned} \tag{8.8}$$

By writing $N = (M_1 + M_2\sigma_3)$, substituting (8.8) into (8.4), and using the condition that (7.3) must be satisfied, we find that

$$N\tilde{N} = NN^\dagger = 1 \tag{8.9}$$

which means that N is a spatial rotation. Therefore, we can absorb N in the Lorentz rotation R of (8.3) and write Ψ in the canonical form

$$\Psi = \psi\tfrac{1}{4}(1 + \gamma_0)(1 + \sigma_3). \tag{8.10}$$

It should be emphasized that (8.10) has been derived from the definition of Ψ as a spinor and independent of any particular form for the Dirac equation.

To replace the Dirac equation for Ψ by an equation for the simpler wavefunction ψ, substitute (8.10) into (7.1), separate the equation into even and odd parts, multiply the odd part on the right by γ_0, and add it again to the even part. The result is Eq. (5.1).

We have one piece of unfinished business. To analyze the case where $\mathbf{J}_0$ is a null vector we return to Eq. (7.6) and argue as in the non-null case. We find that if

$$\mathbf{J}_0^2 = 0, \quad \mathbf{J}_0 \neq 0 \tag{8.11}$$

then Ψ can be written in the canonical form

$$\Psi = \rho^{\frac{1}{2}}R\gamma_0 2^{-\frac{1}{2}}(1 + \sigma_3) = \rho^{\frac{1}{2}}R2^{-\frac{1}{2}}(\gamma_0 - \gamma_3), \tag{8.12}$$

where R is a Lorentz rotation and ρ is a positive scalar. It is important to study the difference between (8.10) and (8.12). By the algebraic operations described in the last paragraph, (8.10) can be reduced to (8.3). But an analogous reduction of (8.12) is not possible. We can do no more than rewrite it in the equivalent even form

$$\psi = \rho^{\frac{1}{2}}R\,2^{-\frac{1}{2}}(1 - \sigma_3). \tag{8.13}$$

This quantity has perhaps as much right to be called a spinor as the quantity (8.3). In fact it can be written as the sum of two spinors of the type (8.3)

$$\rho^{\frac{1}{2}}R2^{-\frac{1}{2}}(1 - \sigma_3) = (\tfrac{1}{2}\rho)^{\frac{1}{2}}R + (\tfrac{1}{2}\rho e^{i\pi})^{\frac{1}{2}}Re^{-i\sigma_3\frac{1}{2}\pi}.$$

Therefore, if the sum of any two spinors is again a spinor, then both (8.3) and (8.13) must be spinors. Still, we have not discussed spinors of the type (8.13) because they cannot be solutions of the Dirac equation for an electron. The Dirac equation does not allow (8.11) to be satisfied. Spinors of the type (8.13) are equivalent to the "two-component spinors" commonly used to describe neutrinos.

D. PROSPECTS

To bring the program initiated here to its logical conclusion, it is necessary to carry out a general reassessment of the role of complex numbers in elementary particle theory. The fact that the $(-1)^{\frac{1}{2}}$ has a geometrical interpretation in electron theory strongly suggests that a similar interpretation can be given to every appearance of $(-1)^{\frac{1}{2}}$ in the basic equations of physics, though it is by no means evident that the $(-1)^{\frac{1}{2}}$ will have the same meaning in each instance. We have learned that the $(-1)^{\frac{1}{2}}$ in the electron wavefunction is inextricably tied up with the spin. Yet the $(-1)^{\frac{1}{2}}$ which seems to play an essential role in the quantization of fermion fields has no evident connection with spin. It will take a careful analysis of quantum electrodynamics to resolve this apparent conflict. Again, the tieup of $(-1)^{\frac{1}{2}}$ and spin suggests that analytic continuation of scattering amplitudes cannot be properly understood without taking spin into account.[13] On the other hand, there may be geometrically different kinds of analytic continuation, for there are several different geometrical roots of minus one in the Dirac algebra.

Anyone who plays the game of theoretical physics with the rules suggested here will not allow a $(-1)^{\frac{1}{2}}$ in his theory unless it has a geometrical significance grounded in space–time. This may be regarded as another constraint imposed by space–time on permissible physical equations which had already been restricted by the requirement of relativistic covariance or invariance.

We have exhibited several bilinear "observables" of a spinor field ψ. The Dirac theory supplies a physical interpretation of $\mathbf{J}_0 = \psi\gamma_0\tilde{\psi}$ and of the general orientation of the frame $\{\mathbf{J}_i = \psi\gamma_i\tilde{\psi}\}$. We have suggested an interpretation of $\psi\tilde{\psi}$ which awaits final justification. To supply a more detailed interpretation of the $\mathbf{J}_\mu$, we must go well beyond the Dirac theory. This possibility will be explored in another paper.

[13] A formulation of the "complex" Lorentz group in terms of the real Dirac algebra is given in Sec. 19 of Ref. 2.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Spin and Isospin

DAVID HESTENES*
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey

(Received 17 November 1965)

Spinor fields can be classified according to the invariants of their derivatives. It is suggested that different invariants describe different interactions of elementary particles. Then the classification of spinor fields becomes a classification of elementary particles and their interactions. A geometric interpretation of isospin is suggested and is used in a model of nucleon interactions. This theory entails an intimate connection between spin and isospin which appears to have some experimental support. Among other things, it connects the pseudoscalar and isospin properties of the pion and accounts for the polarization of the third axis in isospace by the electromagnetic interactions.

1. INTRODUCTION

IN another paper (hereafter called I),[1] the Dirac electron theory was formulated within the framework of the *real* Dirac algebra. Here we examine the possibility of describing nonelectromagnetic interactions within the same framework. Familiarity with the contents and notation of I is a prerequisite for this discussion. We recapitulate some parts of I to establish the motivation and point of departure for the particular generalization of Dirac theory to be considered here.

With spinor field equations we can give a good account of the electromagnetic interactions of protons and electrons. Is it possible to describe all interactions, hence account for all quantum numbers, in terms of spinor fields alone? Since we know the most general expression for a spinor field and discovered in I a good many of its geometrical properties, we should be able to answer this question with a definite yes or no. Surely, this possibility should be carefully investigated, because it is a minimal generalization of the well-established Dirac theory. In this paper we write down the most general equation for a spinor field. By comparing it with the Dirac equation one sees precisely which degrees of freedom are used to describe electromagnetic interactions. The problem is then to see if the remaining degrees of freedom can be used to describe other interactions. The first vaguely encouraging sign is that the number of degrees appears to be sufficient for the description of strong and/or weak interactions. This paper describes an attempt to use some of the unidentified freedoms to describe isospin and pionic interactions of nucleons. The freedom of choice in this matter is greatly limited by the requirement that distinct physical quantities have a distinct geometrical interpretation and vice-versa. So much so that we see no alternative to the identifications made here, where the object is to relate isospin to prominent geometrical properties of spinor fields described in I. Whatever the validity of the specific attempt made here, it is presented in the hope that it encourages others to make a more satisfactory account of the properties of spinor fields.

In I, we learned that a spinor field ψ determines a map of an inertial frame $\{\gamma_\mu;\ \mu = 0, 1, 2, 3\}$ into a set of four orthogonal vector fields $\mathbf{J}_\mu$:

$$\gamma_\mu \to \mathbf{J}_\mu \equiv \psi\gamma_\mu\tilde{\psi}. \tag{1.1}$$

By interpreting the $\mathbf{J}_\mu$ physically we achieve a physical interpretation of ψ. In the usual formulation of the Dirac theory $\mathbf{J}_0$ is identified as the "probability current," $\mathbf{J}_3$ describes the orientation of spin, but $\mathbf{J}_1$ and $\mathbf{J}_2$ are suppressed. In I, $\mathbf{J}_1$ and $\mathbf{J}_2$ were exhumed. Here we enquire into their physical significance. A clue is given by the discussion of charge conjugation in Sec. 6 of I. It can be observed that under charge conjugation the $\mathbf{J}_i$ $(i = 1, 2, 3)$ behave exactly as a frame of base vectors in isospace. This suggests that we interpret the $\mathbf{J}_i$ as "isospin currents." The third axis of isospace is then singled out because $\mathbf{J}_3$ is the "spin vector." We already know from the Dirac equation that $\mathbf{J}_3$ is distinguished from $\mathbf{J}_1$ and $\mathbf{J}_2$ by virtue of the particular form of the electromagnetic interaction.

Having identified the $\mathbf{J}_i$ as isospin currents, the next step is to construct a model of pion–nucleon interactions. We take our clue from an earlier work (hereafter called STA)[2] wherein it was shown that the usual coupling of a pseudoscalar pion to the nucleon can be represented solely within the framework of the real Dirac algebra. This model leads directly to the interpretation of the $\mathbf{J}_\mu$ given above for more general

* Present address: Arizona State University, Tempe, Arizona.
[1] D. Hestenes, J. Math. Phys. 8, 798 (1967).
[2] D. Hestenes, *Space-Time Algebra* (Gordon and Breach Science Publishers, Inc., New York, 1966)

reasons. Moreover, it implies that $\mathbf{J}_0$ is conserved, i.e.,

$$\square \cdot \mathbf{J}_0 = 0 \tag{1.2}$$

and that, neglecting electromagnetic effects, the divergence of $\mathbf{J}_i$ is proportional to the pion field π_i, i.e.,

$$\square \cdot \mathbf{J}_i \sim \pi_i. \tag{1.3}$$

Thus, the π_i can be interpreted as the components of a spacelike vector. In this way, the so-called "pseudoscalar" property of the pion is united with the "isovector" property.

The expression (1.3) also tells us how to distinguish a spinor field representing an electron from a spinor field representing a nucleon. For the electron, but not for the nucleon, the divergence of the spin vector vanishes. This observation suggests a general theory of interactions. The Dirac equation tells us that the variation of $\psi(\mathbf{x})$ as a function of $\mathbf{x}$ is of a very special kind when only electromagnetic interactions are considered, and Eq. (1.3) tells us that the variations of $\psi(\mathbf{x})$ produced by pion interactions are quite different though nevertheless specific. If this description of interactions is correct, then it must be possible to classify the possible variations of a spinor field in a way which is related to a classification of interactions among elementary particles.

Now, to give concreteness to these general remarks, let us examine a specific model. We do not here strive for a complete theory of elementary interactions; we can only hope to unlock the door.

2. NUCLEON FIELD

In STA we wrote the following equation for a nucleon field N with electromagnetic and pionic interactions:

$$\square N = Nim + e\mathbf{A}N\tfrac{1}{2}(1 + \boldsymbol{\sigma}_3)i - giNi\pi\boldsymbol{\gamma}_0, \tag{2.1}$$

where m is the nucleon mass, g is the pion–nucleon coupling constant, and, in the expression $\pi\boldsymbol{\gamma}_0 = \pi^i\boldsymbol{\gamma}_i\boldsymbol{\gamma}_0 = \pi^i\boldsymbol{\sigma}_i$, the π^i are the components of the pion field. The nucleon field N may be thought of as a 4×2 matrix rather than the usual 8×1 matrix. "Isospin operators" multiply N on the right; the eigenstates of $\boldsymbol{\sigma}_3$ are $p = N\frac{1}{2}(1 + \boldsymbol{\sigma}_3)$ and $n = N\frac{1}{2}(1 - \boldsymbol{\sigma}_3)$, respectively the proton and neutron components of the nucleon field.

Equation (2.1) is algebraically equivalent to the nucleon equation usually used in the field theory of direct pseudoscalar pion interactions.[3] But (2.1) has additional properties which accrue because the pion–nucleon coupling is expressed entirely in terms of the real Dirac algebra. In contrast to the usual theory where pionic interactions are composed of completely independent space–time and isospace parts, Eq. (2.1) can be given a thoroughgoing space–time interpretation. To elucidate this, we follow the method of Sec. C in I in order to find an expression for the nucleon field which is more perspicuous than N.

We write N as a generalization of the expression (9.9) in I for an electron field:

$$N = \psi\tfrac{1}{4}(1 + \boldsymbol{\gamma}_0)(1 + \boldsymbol{\sigma}_3)u. \tag{2.2}$$

Here, as in I, ψ describes the energy and spin of the field. The factor $u = u(\mathbf{x})$ is a spatial rotation (i.e., $u\tilde{u} = uu^\dagger = 1$), but in (2.2) it can be regarded as a local rotation in isospace which determines the relative magnitude of the proton and neutron components of the nucleon field at each point of space–time. It was suggested in I that a factor $e^{i\beta}$ describes the admixture of positive and negative energy states in the total wavefunction. Therefore, a continuous variation in $e^{i\beta}$ describes a continuous transition of particles into antiparticles. In complete analogy, the factor u now describes continuous transitions between proton and neutron states.

Continuing the procedure used in I, we substitute (2.2) into (2.1) and eliminate the projection operators $\frac{1}{2}(1 + \boldsymbol{\gamma}_0)$ and $\frac{1}{2}(1 + \boldsymbol{\sigma}_3)$ to get the equation

$$\square(\psi u) = im\psi\boldsymbol{\sigma}_3 u\boldsymbol{\gamma}_0 + e\mathbf{A}\psi\tfrac{1}{2}(\boldsymbol{\sigma}_3 u + u\boldsymbol{\sigma}_3)i + g\psi u\pi. \tag{2.3}$$

Let us define

$$\mathcal{N} = \psi u = (\rho e^{i\beta})^{\frac{1}{2}}R \tag{2.4}$$

and

$$\boldsymbol{\sigma}_3' = \boldsymbol{\sigma}_3'(\mathbf{x}) = \boldsymbol{\gamma}_3'\boldsymbol{\gamma}_0 \equiv \tilde{u}\boldsymbol{\sigma}_3 u, \tag{2.5}$$

so (2.3) can be written as

$$\square\mathcal{N} = m\mathcal{N}\boldsymbol{\gamma}_0 i\boldsymbol{\sigma}_3' + e\mathbf{A}\mathcal{N}\tfrac{1}{2}(\boldsymbol{\sigma}_3 + \boldsymbol{\sigma}_3')i + g\mathcal{N}\pi. \tag{2.6}$$

The spinor field $\mathcal{N}$ describes the nucleon. As before, we can form four vector fields from $\mathcal{N}$:

$$\mathbf{J}_\mu \equiv \mathcal{N}\boldsymbol{\gamma}_\mu\tilde{\mathcal{N}} = \rho\mathbf{e}_\mu. \tag{2.7}$$

As before, $\mathbf{J}_3$ is the nucleon "probability current." But now, $\mathbf{J}_3$ cannot be identified with the "spin current," which is

$$\mathbf{J}_3' = \mathcal{N}\boldsymbol{\gamma}_3'\tilde{\mathcal{N}} = \psi\boldsymbol{\gamma}_3\tilde{\psi} \equiv \rho\mathbf{s}. \tag{2.8}$$

However, there is a connection between $\mathbf{J}_3$ and $\mathbf{J}_3'$; when $\mathbf{J}_3 = \mathbf{J}_3'$ the nucleon is in a proton eigenstate, and when $\mathbf{J}_3 = -\mathbf{J}_3'$ the nucleon is in a neutron eigenstate. The $\mathbf{J}_i$ form a basis for a "local isospace" at each space–time point, and the projection of the "spin vector" on the third axis of this isospace

[3] A recent account of meson field theory is given by K. Nishijima, *Fundamental Particles* (W. A. Benjamin, Inc., New York, 1963).

describes the admixture of proton and neutron components in the nucleon field.

Calculating the divergence of the $\mathbf{J}_\mu$, we find

$$\Box \cdot \mathbf{J}_0 = 0, \tag{2.9a}$$

$$\Box \cdot \mathbf{J}_i = ei\tfrac{1}{2}(\mathbf{J}_3 + \mathbf{J}_3') \wedge \mathbf{e}_\mu \wedge \mathbf{e}_0 \wedge \mathbf{A} + g\rho \cos\beta\pi_i . \tag{2.9b}$$

Equation (2.9) shows that if electromagnetic effects are neglected, the divergence of $\mathbf{J}_i$ is proportional to $\boldsymbol{\pi}_i$, as stated in the introduction to this paper. It is worth emphasizing that this is a consequence of expressing the nucleon equation entirely in terms of the real Dirac algebra. The existence of the last term in (2.9b) can be put to experimental test. We come back to this point later.

The nucleon equation can be thought of as a definition of invariants of the nucleon field. To see this more clearly, multiply (2.6) on the right by $\tilde{\mathcal{N}}$ to get

$$(\Box\mathcal{N})\tilde{\mathcal{N}} = \rho[im\mathbf{s} + e\mathbf{A}e^{i\beta}\tfrac{1}{2}(\mathbf{e}_3 + \mathbf{s})\mathbf{e}_0 i + g\boldsymbol{\pi}'], \tag{2.10}$$

where

$$\boldsymbol{\pi}' \equiv R\boldsymbol{\pi}\tilde{R} = \pi^i \mathbf{e}_i .$$

Let us define vectors $\mathbf{a}$ and $\mathbf{b}$, respectively, called the *divergence* and *torsion* vectors of $\mathcal{N}$, by the equation

$$(\Box\mathcal{N})\tilde{\mathcal{N}} = \mathbf{a} + i\mathbf{b}. \tag{2.11}$$

By comparison with (2.10),

$$\mathbf{a} = \rho[g\boldsymbol{\pi}' + e\tfrac{1}{2}\{-(\mathbf{e}_0\mathbf{e}_3 + \mathbf{e}_0\mathbf{s}) \cdot \mathbf{A} \sin\beta + i(\mathbf{e}_0\mathbf{e}_3 + \mathbf{e}_0\mathbf{s}) \wedge \mathbf{A} \cos\beta\}], \tag{2.12}$$

$$\mathbf{b} = \rho[m\mathbf{s} + e\tfrac{1}{2}\{(\mathbf{e}_0\mathbf{e}_3 + \mathbf{e}_0\mathbf{s}) \cdot \mathbf{A} \cos\beta - i(\mathbf{e}_0\mathbf{e}_3 + \mathbf{e}_0\mathbf{s}) \wedge \mathbf{A} \sin\beta\}]. \tag{2.13}$$

Now let us attempt to interpret these expressions. The divergence and torsion vectors describe geometrically distinguishable properties of a spinor field. Pionic interactions occur *only* through the divergence vector. They are distinguishable from electromagnetic interactions which proceed through *both* vectors. The spin vector is associated *only* with the torsion vector.

Instead of representing the nucleon field by the wavefunction $\mathcal{N}$, we can represent it by the wavefunction ψ. From (2.3) we get the equation

$$\Box\psi = m\psi\boldsymbol{\gamma}_0 i\boldsymbol{\sigma}_3 + e\mathbf{A}\psi\tfrac{1}{2}(\boldsymbol{\sigma}_3 + \boldsymbol{\sigma}_3'')i + g\psi\boldsymbol{\pi}'' + \boldsymbol{\gamma}^\mu\psi\boldsymbol{\rho}_\mu i, \tag{2.14}$$

where

$$\boldsymbol{\sigma}_3'' \equiv u\boldsymbol{\sigma}_3\tilde{u}, \tag{2.15}$$

$$\boldsymbol{\pi}'' \equiv u\boldsymbol{\pi}\tilde{u}, \tag{2.16}$$

$$\boldsymbol{\rho}_\mu \equiv -i(\partial_\mu u)\tilde{u}. \tag{2.17}$$

The last term in (2.14) has the form of "vector mesons" $\boldsymbol{\rho}_\mu$ coupled to the nucleon field. Like $\boldsymbol{\pi}'$, $\boldsymbol{\rho}_\mu$ can be identified as a vector in isospace. The experimental resonance with the proper quantum numbers to be identified with the $\boldsymbol{\rho}_\mu$ is the rho meson. In (2.17) it is shown explicitly that the $\boldsymbol{\rho}_\mu$ are effectively the directional derivatives of the "operator" which mixes proton and neutron components of the nucleon field.

Equation (2.14) has implications for the electromagnetic structure of nucleons, for it says that the electromagnetic potential $\mathbf{A}$ is coupled to the nucleon field by the "charge operator"

$$Q \equiv ie\tfrac{1}{2}(\boldsymbol{\sigma}_3 + \boldsymbol{\sigma}_3''). \tag{2.18}$$

Since according to (2.15), $\boldsymbol{\sigma}_3''$ is not constant, the "effective charge" is "smeared out." The nature of this effect is determined by the rho meson, because

$$\partial_\mu Q = \tfrac{1}{2}e[\boldsymbol{\sigma}_3'', \boldsymbol{\rho}_\mu]. \tag{2.19}$$

3. DISCUSSION

The work in this paper is directed toward the ultimate goal of providing a physical interpretation for all the geometric invariants of a spinor field. We began with the fact that a spinor field ψ determines a map of some inertial frame $\{\boldsymbol{\gamma}_\mu\}$ into the "current vectors" $\mathbf{J}_\mu = \psi\boldsymbol{\gamma}_\mu\tilde{\psi}$. By interpreting the spacelike vectors $\mathbf{J}_i$ as "isospin currents" we have associated isospace with some inertial frame. We may consider the $\boldsymbol{\gamma}_i$ (or, if we wish, the $\mathbf{J}_i$) as a basis in isospace. Having done this much, it is natural to generalize isospace to include the entire four-dimensional space spanned by the $\boldsymbol{\gamma}_\mu$. Also, if we wish, we can identify the $\boldsymbol{\gamma}_\mu$ with some convenient inertial frame, such as the laboratory frame.[4] We can go from one inertial frame to another by the constant Lorentz transformation

$$\boldsymbol{\gamma}_\mu \rightarrow \boldsymbol{\gamma}_\mu' = \tilde{R}\boldsymbol{\gamma}_\mu R \tag{3.1a}$$

($R\tilde{R} = 1$, $\partial_\mu R = 0$). This change of basis leaves the physical vectors $\mathbf{J}_\mu$ unchanged if it is accompanied by the transformation

$$\psi \rightarrow \psi R. \tag{3.1b}$$

If R is a spatial rotation (i.e., if $R = R^*$), then (3.1) is just what is commonly called a rotation in isospace. It may be worth remarking that the transformation (3.1) has nothing to do with the usual "coordinate transformation" unless a correspondence between them is specifically set up.

Having given a physical interpretation to the bilinear invariants of a spinor field, we turn to the

[4] In matrix language this amounts to choosing a particular representation for the Dirac matrices.

problem of providing a physical interpretation for its derivatives. The directional derivatives of a spinor field ψ can always be written in the form

$$\partial_\mu \psi = \tfrac{1}{2}\omega_\mu \psi. \tag{3.2}$$

Each of the four ω_μ can be a general even d-number and so has eight components. Therefore, the most general variation of a spinor field is described by $4 \times 8 = 32$ independent degrees of freedom. We have not attempted to interpret the ω_μ physically. Instead, we have dealt with the simpler equation

$$\Box \psi = \tfrac{1}{2}\omega \psi, \tag{3.3}$$

where

$$\omega = \gamma^\mu \omega_\mu. \tag{3.4}$$

The gradient of ψ is completely determined by ω. A physical interpretation of ω is provided by expressing ω in terms of physical fields. This was done explicitly in the discussion following Eq. (2.11), a discussion which indicates that (3.3) is sufficient to describe electromagnetic and pionic interactions but has room for little else. However, since (3.3) uses only 8 of the possible 32 degrees of freedom, we can get a much more general theory of interactions by using Eq. (3.2). But before attempting such a large generalization, we should study (3.3) thoroughly, especially to see in what ways our interpretation of it can be brought to experimental test.

We have considered a specific model to illustrate the general idea that interactions be described by derivatives of a spinor field and that different kinds of interaction correspond to different invariants of these derivatives. We don't expect this model to be a final theory if only because the coupling constants have not been accounted for and other interactions must be included. Still, as far as we know, it gives a correct account of electromagnetic interactions. We also know that some features of pionic interactions are correctly described by our model. If any special features of the model can be verified experimentally, we have convincing evidence for the general idea that all quantum numbers of elementary particles can be described by symmetries of a real spinor field and its derivatives. Only then can we have confidence that an analysis of spinor fields from first principles leads to a more satisfactory wave equation.

By assuming that the divergence of the renormalized axial vector nucleon current operator in beta decay is proportional to the renormalized pion field operator, Gell-Mann and Lévy[5] derived the formula of Goldberger and Treiman[6] which relates the decay rate of a charged pion to the axial vector coupling constant in beta decay. Adler[7] shows that the same assumption implies consistency conditions involving the strong interactions alone; he finds that they are satisfied experimentally to within 10%. Our model may well provide a theoretical justification for the above assumption, and so be supported by the experimental evidence we have just mentioned. For Eq. (2.9) shows that the divergence of the spacelike (axial vector) nucleon currents are proportional to the components of the pion field. It is necessary only that this relation be preserved under renormalization of the quantized theory. It is even reasonable to *assume* that it be preserved as a condition on renormalization. This condition is similar to, if not identical with, the condition that gauge invariance be preserved by renormalization in quantum electrodynamics. One can even hope that it is a sufficient condition to make it possible to extract finite results from renormalization.

Encouraging as this argument is, we need not and cannot depend on it exclusively. Our model has other special properties which we have already mentioned. For instance, Eq. (2.9) says that the divergence of the spin vector of the nucleon field is proportional to the neutral pion field. Also, recall that the electromagnetic coupling of the Dirac theory is somewhat modified in the nucleon equation, producing a "nonlocal electromagnetic structure" of the nucleon. It should be possible to subject these special features of our model to experimental test. But this requires a more elaborate study than we have attempted here.

We have discussed some general ideas connecting isospin to space–time properties of spinor fields in terms of a specific model. If these ideas are correct, no doubt they must eventually be incorporated into a more ambitious physical model, such as that of Nambu and Jona-Lasinio.[8] By an analogy with superconductivity, these authors suppose that different particles correspond to different solutions of the equation for a single underlying spinor field. To this we would add the idea that isospin is determined by the spacelike currents. $\mathbf{J}_i$ of the underlying spinor field, and perhaps the idea that, neglecting electromagnetic interactions, solutions corresponding to leptons and baryons are distinguished respectively by the vanishing or nonvanishing of the divergence of the $\mathbf{J}_i$.

[5] M. Gell-Mann and M. Lévy, Nuovo Cimento **16**, 705 (1960).

[6] M. Goldberger and S. Treiman, Phys. Rev. **110**, 1178 (1958); *ibid.* **111**, 354 (1958).

[7] S. Adler, Phys. Rev. **137**, B1022 (1965).

[8] Y. Nambu and G. Jona-Lasinio, Phys. Rev. **122**, 345 (1961); *ibid.* **124**, 246 (1961).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

One-Dimensional Three-Body Problem

JAMES HURLEY
University of California, Davis, California

(Received 6 October 1966)

The three-body problem in one dimension is examined to determine for what class of interactions the Schrödinger equation may be solved by separation of variables.

I. INTRODUCTION

THE three-body problem in one dimension is examined to see if there exist interaction potentials for which the Schrödinger equation may be solved by separation of variables. We assume that one of the particles is infinitely massive and defines the origin of the coordinate system. Thus our problem is to discover for what potentials the equation

$$-\frac{\hbar^2}{2m_1}\frac{\partial^2\psi}{\partial x_1^2}-\frac{\hbar^2}{2m_2}\frac{\partial^2\psi}{\partial x_2^2}+V_1(x_1)\psi + V_2(x_2)\psi+V_3(x_1-x_2)\psi=E\psi$$

separates under an appropriate change of variables. We assume that the two light particles are of equal mass. For convenience we relabel the coordinates $x_1 = x$ and $x_2 = y$ so that, after renormalizing the energies, we have

$$-(\partial^2\psi/\partial x^2)-(\partial^2\psi/\partial y^2) + [V_1(x)+V_2(y)+V_3(x-y)]\psi=E\psi.$$

The system is better described by a single point in a two-dimensional space than by two points in one dimension.

We consider the class of all analytic transformation

$$z'=z'(z),$$

where

$$z=x+iy,\quad z'=x'+iy'$$

and seek to find those potentials for which the transformed equation may be solved by separation of variables. For analytic transformations[1]

$$\frac{\partial^2\psi}{\partial x^2}+\frac{\partial^2\psi}{\partial y^2}=\left|\frac{dz'}{dz}\right|^2\left[\frac{\partial^2\psi}{\partial x'^2}+\frac{\partial^2\psi}{\partial y'^2}\right].$$

We see then that we may separate variables in the transformed equation provided

$$|dz/dz'|^2[V_1(x)+V_2(y)+V_3(x-y)-E] = f_1(x')+f_2(y'), \quad (1)$$

where f_1 and f_2 are functions of the indicated variables alone. Our point of view is: given the transformation, find the class of potentials that satisfy Eq. (1). A necessary and sufficient condition is that

$$(\partial^2/\partial x'\partial y')\{|dz/dz'|^2[V_1(x)+V_2(y) + V_3(x-y)-E]\}=0. \quad (2)$$

We assume that the transformation does not depend on the energy (we do not want to search for a new transformation for each energy). We see then that in order for Eq. (2) to be satisfied for all E we require in the first place that

$$(\partial^2/\partial x'\partial y')\,|dz/dz'|^2=0. \quad (3)$$

This is the same condition for which the wave equation separates. It is shown in Ref. 1 that Eq. (3) is satisfied by those transformations satisfying the equation

$$d^3z/dz'^3=\lambda(dz/dz'), \quad (4)$$

where λ is an arbitrary real constant. This is an ordinary differential equation and may easily be integrated to find $z = z(z')$.

II. ROTATION OF COORDINATE AXES

Returning now to Eq. (2) we seek a solution of

$$(\partial^2/\partial x'\partial y')\{|dz/dz'|^2[V_1(x)+V_2(y)+V_3(x-y)]\}=0$$

for a given transformation satisfying Eq. (4).

Carrying out the differentiations, we have

$$g\frac{\partial x}{\partial x'}\frac{\partial x}{\partial y'}\dot{V}_1(x)+\left(\frac{\partial g}{\partial x'}\frac{\partial x}{\partial y'}+\frac{\partial g}{\partial y'}\frac{\partial x}{\partial x'}+g\frac{\partial^2 x}{\partial x'\partial y'}\right)\dot{V}_1(x)+g\frac{\partial y}{\partial x'}\frac{\partial y}{\partial y'}\dot{V}_2(y) + \left(\frac{\partial g}{\partial x'}\frac{\partial y}{\partial y'}+\frac{\partial g}{\partial y'}\frac{\partial y}{\partial x'}+g\frac{\partial^2 y}{\partial x'\partial y'}\right)\dot{V}_2(y)-g\left(\frac{\partial x}{\partial x'}\frac{\partial y}{\partial y'}+\frac{\partial y}{\partial x'}\frac{\partial x}{\partial y'}\right)\dot{V}_3(x-y) + \left[\frac{\partial g}{\partial x'}\left(\frac{\partial x}{\partial y'}-\frac{\partial y}{\partial y'}\right)+\frac{\partial g}{\partial y'}\left(\frac{\partial x}{\partial x'}-\frac{\partial y}{\partial x'}\right)+g\left(\frac{\partial^2 x}{\partial x'\partial y'}-\frac{\partial^2 y}{\partial x'\partial y'}\right)\right]\dot{V}_3(x-y)=0, \quad (5)$$

[1] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill Book Company, Inc., New York, 1953), Vol. I, p. 499.

where $g = |dz/dz'|^2$ and the dot above the V's indicates differentiation with respect to the argument.

We now substitute into Eq. (5) explicit solutions of Eq. (4). Consider first the case $\lambda = 0$. Then

$$z = \tfrac{1}{2}az'^2 + bz' + c. \tag{6}$$

If we further restrict the transformation by setting $a = c = 0$, we have a rotation of coordinates and a change of scale if $|b| \neq 1$. If we let $z = bz'$ in Eq. (5) we have

$$-b_1b_2\dot{V}_1(x) + b_1b_2\dot{V}_2(y) + (b_2^2 - b_1^2)\dot{V}_3(x - y) = 0, \tag{7}$$

where $b = b_1 + ib_2$.

Differentiating with respect to x and y we have

$$(b_2^2 - b_1^2)\dddot{V}_3(x - y) = 0.$$

Either $b_1 = \pm b_2$ (a rotation of the coordinate axes by 45°) or

$$\dddot{V}_3 = 0.$$

Consider first $b_1 \neq \pm b_2$, then

$$V_3 = c_1(x - y)^3 + c_2(x - y)^2 + c^3(x - y) + c_4.$$

The only potentials of interest are symmetric in x and y, so we set $c_1 = c_3 = 0$. Substituting this result back into Eq. (7) we have

$$-b_1b_2[\dot{V}_1[(x) - \dot{V}_2(y)] = \text{const.}$$

Since x and y are independent, $\dot{V}_1$ and $\dot{V}_2$ must separately be constant, so that all three potentials are harmonic oscillator interactions. It is well known that particles interacting by harmonic potentials may be separated by diagonalization (rotation of coordinates).

However, if $b_1 = \pm b_2$ then $V_3(x - y)$ may be chosen arbitrarily. V_1 and V_2 remain harmonic potentials. Thus by rotating the coordinates 45° the potential

$$\tfrac{1}{2}k(x^2 + y^2) + V_3(x - y)$$

becomes

$$\tfrac{1}{2}k(x'^2 + y'^2) + V_3(\sqrt{2}x'),$$

and the problem is reduced to solving two one-dimensional problems, one particle interacting harmonically and the second interacting through the potential $\frac{1}{2}kx'^2 + V_3(\sqrt{2}x')$.

Strangely enough it is possible to extend this result to three dimensions. Consider two particles bound harmonically to an infinitely massive particle at the origin and interacting with each other through a potential which depends only on the distance between the particles. The Schrödinger equation is

$$-\nabla_1^2\psi - \nabla_2^2\psi + \tfrac{1}{2}k(r_1^2 + r_2^2)\psi + V_3(\mathbf{r}_1 - \mathbf{r}_2)\psi = E\psi.$$

Introducing the center of mass and relative coordinates

$$\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2, \quad \mathbf{R} = \tfrac{1}{2}(\mathbf{r}_1 + \mathbf{r}_2),$$

we have

$$-\tfrac{1}{2}\nabla_{\mathbf{R}}^2\psi - 2\nabla_{\mathbf{r}}^2\psi + \tfrac{1}{4}k(r^2 + 4R^2)\psi + V_3(\mathbf{r})\psi = E\psi.$$

Thus the shell model with two particles may be separated and there is no need to use perturbation theory.

III. PARABOLIC COORDINATES

We return now to the class of transformation for which $\lambda = 0$ [see Eq. (6)]. In the previous section we considered the case $z = bz'$. We now take up parabolic coordinates defined by

$$z = \tfrac{1}{2}z'^2.$$

We have set $a = 1$ in Eq. (6). Where $a \neq 1$ we would have a rotation, and change of scale in addition. With this transformation Eq. (5) becomes after some reduction

$$-y\dddot{V}_1 + y\dddot{V}_2 + 3\ddot{V}_2 - 2x\dddot{V}_3 - 3\ddot{V}_3 = 0.$$

To solve for V_3 we divide by y and differentiate with respect to y, obtaining

$$\dddot{V}_2 + 3\frac{\partial}{\partial y}\left(\frac{\ddot{V}_2}{y}\right) - 2x\frac{\partial}{\partial y}\left(\frac{\dddot{V}_3}{y}\right) - 3\frac{\partial}{\partial y}\left(\frac{\ddot{V}_3}{y}\right) = 0.$$

We may eliminate $V_2(y)$ by differentiating with respect to x. Thus

$$2\frac{\partial}{\partial x}\left[x\frac{\partial}{\partial y}\left(\frac{\dddot{V}_3}{y}\right)\right] + 3\frac{\partial^2}{\partial x\partial y}\left(\frac{\ddot{V}_3}{y}\right) = 0.$$

Introducing a change of variables

$$u = x - y, \quad v = x + y,$$

the above equation becomes

$$v(u - v)\frac{d^4V_3(u)}{du^4} - 2u\frac{d^3V_3(u)}{du^3} + 2(v - u)\frac{d^2V_3(u)}{du^2} = 0.$$

Since u and v are independent variables, the equation can be satisfied only if

$$d^2V_3(u)/du^2 = 0.$$

Therefore $V_3(x - y) = c_1(x - y) + c_2$. If we also require that V_3 be symmetric we see that the three-body problem separates only when the two free particles do not interact.

IV. CYLINDRICAL COORDINATES

Returning again to solutions of Eq. (4) we obtain the transformation to cylindrical coordinates by choosing $\lambda = 1$ so that a solution is

$$z = e^{z'}.$$

We then have from Eq. (5)

$$-xy\dot{V}_1 - 3y\dot{V}_1 + xy\dot{V}_2 + 3x\dot{V}_2 + (y^2 - x^2)\dot{V}_3 - 3(x + y)\dot{V}_3 = 0. \quad (8)$$

We may eliminate $V_1(x)$ and $V_2(y)$ by operating with

$$\frac{\partial}{\partial x}\frac{1}{x}\frac{\partial}{\partial y}\frac{1}{y}$$

to obtain

$$\frac{\partial}{\partial x}\left\{\frac{1}{x}\frac{\partial}{\partial y}\left(\frac{1}{y}\left[(y^2 - x^2)^2\dot{V}^3 - 3(x + y)\dot{V}_3\right]\right)\right\} = 0$$

or

$$\frac{\partial}{\partial x}\left\{\frac{1}{x}\frac{\partial}{\partial y}\left(\frac{1}{y}\left[(x + y)f(x - y)\right]\right)\right\} = 0,$$

where

$$f(x - y) = (y - x)\dot{V}_3(x - y) - 3\dot{V}_3(x - y). \quad (9)$$

Carrying out the differentiation we have

$$f' + [(x - y)/xy]f = 0.$$

Since f is a function of $(x - y)$ alone and since x and y are independent variables, the above equation can be satisfied for all x and y only if $f = 0$. Setting $f =$ const in Eq. (9) and solving the resulting ordinary differential equation for V_3, we find

$$V_3(x - y) = c_1 + [c_2/(x - y)^2] + c_3(x - y).$$

Here, c_1 and c_2 are constants of integration and c_3 is proportional to the constant chosen for f.

Again, if we require that V_3 be symmetric ($c_3 = 0$) and ignore the additive constant ($c_1 = 0$), Eq. (8) becomes

$$-xy\dot{V}_1 - 3y\dot{V}_1 + xy\dot{V}_2 + 3x\dot{V}_2 = 0,$$

so that

$$\frac{\partial}{\partial x}\left[\frac{y}{x}(x\dot{V}_1 + 3\dot{V}_1)\right] = 0 \quad \text{or} \quad V_1 = c_4x^2 + \frac{c_5}{x^2} + c_6.$$

Similarly we obtain for V_2

$$V_2 = c_4y^2 + (c_7/y^2) + c_8.$$

Thus the most general potential representing a three-body interaction for which the Schrödinger equation separates in cylindrical coordinates is

$$c_4(x^2 + y^2) + \frac{c_5}{x^2} + \frac{c_7}{y^2} + \frac{c_2}{(x - y)^2}.$$

It is a simple matter to verify that the Schrödinger equation separates, for the above potential is of the form

$$f(r) + [g(\theta)/r^2].$$

Unfortunately this result is not very useful since the two-body potential $1/x^2$ has rather unusual physical characteristics.[2]

V. ELLIPTIC COORDINATES

A calculation similar to that carried out above for rotated, parabolic, and cylindrical coordinates is very tedious for elliptic coordinates. However, the general potential for which the Schrödinger separates in elliptic coordinates is of the form[3]

$$[f(r_1) + g(r_2)]/r_1r_2,$$

where r_1 and r_2 are distances measured from the foci of the elliptic system. Even if functions f and g could be found which would make the potential of the form necessary to describe a three-body interaction [i.e., $V_1(x) + V_2(y) + V_3(x - y)$] the potentials would have unphysical singularities at r_1 and r_2 equal to zero.

VI. CONCLUSION

We have exhibited the most general class of potentials for which the three-body problem in one dimension may be separated by changing variables. [We have confined ourselves to rotated, parabolic, cylindrical, and elliptic coordinates. There may be other potentials possible using a combination of the above. For example, if $\lambda \neq 1$ we would have rotated cylindrical or elliptic coordinates. In these mixed transformations Eq. (5) becomes very difficult to solve.]

Of the interactions obtained the most interesting is the shell model for two particles. This result may be carried over to three dimensions. The only other example (except for the well-known result of three bodies bound by harmonic potentials) is that where all three particles interact through a potential proportional to $1/s^2$, where s is the distance between particles. This potential is of no physical interest however.

ACKNOWLEDGMENT

This work was supported under NASA grant NGR 05-004-006.

[2] See Ref. 1, p. 1665.
[3] See Ref. 1, p. 657.

Quasi-Spin Methods and Pairs of One-Particle CFP in the Seniority Scheme

FREDERICK R. INNES
Air Force Cambridge Research Laboratories, Laurence G. Hanscom Field, Bedford, Massachusetts

(Received 28 March 1966)

The use of creation and destruction operators to form irreducible tensors in quasi-spin (triple tensors) is advanced to give the important points of the seniority scheme for linear and bilinear forms (CFP, matrix elements, and products). A number of such forms as studied by Racah and others are taken up with emphasis on cases diagonal in the seniority number, v. Values of CFP for seniority $v + 1$ can then be determined from those of seniority $v - 1$. The possibilities in quasi-spin are explored to give a number of additional formulas for products of two CFP. Also, Redmond's formula and some specializations of it are found in the seniority scheme and their properties worked out. The calculation of CFP in the seniority scheme by these formulas is described and comparison made with other methods.

I. INTRODUCTION

THE properties of vector-coupled pairs of creation and destruction operators and the connections of such bilinear forms with seniority can be further studied, beyond recent work,[1,2] to develop a number of relations useful in studying CFP (fractional parentage coefficients), more precisely, products of pairs of CFP, in the seniority scheme. The triple tensors of Lawson, Macfarlane, and Judd (which operate in quasi-spin as well as in spin and orbital spaces) are suited to such a study. Explanation by similar means of many of the results of Racah's third paper[3] has been made,[1] but additional results are given below. The possible use of Redmond's[4] and of certain other formulas in the seniority scheme has been realized for some time, but a concise demonstration can now be given through quasi-spin methods. Such formulas are derived from relations in the matrix elements of bilinear forms by a transform process which sums out the tensor rank of the form and inserts a particular and definitive term in this procedure. CFP are basically calculated from these formulas. Those in v and $v + 1$ appear in terms of those in v and $v - 1$ and new terms [Ref. 3(b), Sec. 6(4)] are so taken up as they occur, with increasing values of v.

The work below consists first in outlining some known properties of triple tensors which are linear or bilinear in the creation and destruction operators for a spin-orbital. More complicated forms are not considered nor any which involve more than one group of equivalent particles. The bilinear forms of triple tensors of odd rank lead to known results for the most part (correspondence relations) so that those of even rank remain to be investigated. Connection is made with related formulas already available. The results obtained throw new light on the seniority scheme and make the demonstration and checking of CFP values in this scheme a much simpler matter than has been the case heretofore.

The matrix elements and products of bilinear forms are held diagonal in v since these are the most important cases. They must contain just two intermediate seniority values, as in the orthogonality relations [Ref. 3(b), Eq. (59)]. It may be best to give all such products for a particular intermediate, $v \pm 1$. An effort has been made to cover the cases of interest. The treatment of the subject is sharper after this is done, e.g., in the derivation of correspondence in matrix elements [Ref. 3(b), Eq. (69)].

The general methods used in finding these results involve the Wigner–Eckart theorem, the one-particle submatrix (or reduced matrix) element, the inner product formula, also correspondence, [Ref. 3(b), Eq. (58)], reciprocity, [Ref. 3(b), Eq. (61)]; and the fundamental properties of 3-j and 6-j coefficients. Many of the results could well be considered as Racah transform[5] formulas involving products of pairs of CFP. Conjugation [Ref. 3(b), Eq. (19)] is not explicitly needed for the present purpose, but the orthogonality properties of the CFP are relevant.

Many conventions from earlier work are used here for a more concise presentation.[6] The concepts and definitions are those of Ref. 3(b) in general. All the CFP are of one-particle type, except some two-particle CFP which are clearly indicated, and all are in standard form. For example, the interchanged

[1] R. D. Lawson and M. H. Macfarlane, Nucl. Phys. **66**, 80 (1965).
[2] B. R. Judd, *Second Quantization and Atomic Spectroscopy* (The Johns Hopkins Press, Baltimore, 1967).
[3] G. Racah, Phys. Rev. **62**, 438 (1942); **63**, 367 (1943); **76**, 1352 (1949), denoted a, b, and c, respectively.
[4] P. J. Redmond, Proc. Roy. Soc. (London) **A222**, 84 (1954); A. Hassitt, *ibid.* **A229**, 110 (1955).
[5] J. B. French, Nucl. Phys. **15**, 393 (1960).
[6] F. R. Innes and C. W. Ufford, Phys. Rev. **111**, 194 (1958).

CFP of Ref. 3(b), Eq. (29), has not been used here. An abbreviated form has accordingly become possible and is found for the one-particle CFP throughout. The phases are also those of Ref. 3(b), and since conjugation is excluded, the wavefunctions are in the $\mathfrak{L}$-scheme. They are kept in αSL coupling, but simply expressed, $SL \equiv I$, to save space. Phase determinations are partly left free in the bilinear relations since the wavefunctions enter in pairs, but a standard set of phases is available from Ref. 3(b), Sec. 6.

II. SUBMATRIX ELEMENTS IN QUASI-SPIN

(1) Properties of Triple Tensors

The elementary triple tensors are of rank one-half in quasi-spin and in spin and rank l in orbital space, written $\mathbf{a}^{(qsl)}$. [It is convenient to set $\gamma \equiv qsl$, $\Gamma \equiv QSL$; cf. Eq. (1), below.] The components of such a triple tensor consist of creation or destruction operators for a spin-orbital, plus or minus one-half, respectively, the quasi-spin components of the tensor. The submatrix elements of these are CFP (with proper weights and phases). If they are reduced in quasi-spin, one finds, by the Wigner–Eckart theorem, in case $Q' > Q$,

$$(vI \,\|a^{(\gamma)}\|\, v-1I') = (-)^{2I}(2v(Q+1)[I])^{\frac{1}{2}} \times (l^{v-1}v-1I'|\}l^v vI). \quad (1a)$$

This may be called a fully reduced submatrix element. Other expressions for it may quickly be found by the use of correspondence, reciprocity (or conjugation). The seniority number v has the same content as Q, just as the number of particles n has the same content as M for the present purpose. These quantum numbers $Q \equiv \frac{1}{4}[l] - \frac{1}{2}v$ and its component $M \equiv \frac{1}{2}n - \frac{1}{4}[l]$ are discussed in the references.[1,2] Apart from their appearance in the Clebsch–Gordan and Racah coefficients of the theory, their use also simplifies expression in relations between matrix products or CFP. The other important value for the submatrix element, for $Q' < Q$, may be quickly found by the same theorem

$$(vI \,\|a^{(\gamma)}\|\, v+1I') = (-)^{I+I'+l}\{(v+1)[Q][I']\}^{\frac{1}{2}} \times (l^v vI|\}l^{v+1}v+1I'). \quad (1b)$$

As these are fully reduced matrix elements, they do not contain a number n or M. In this sense the submatrix elements which are found in Ref. 3(b), for example, are partly reduced.

Relevant bilinear operators for which the submatrix elements are n-dependent (partly reduced) are the unit operators $\mathbf{U}^{(K)}$ and $\mathbf{W}^{(K)}$ [see Eq. (4) below; $K \equiv \kappa k$, a pair of ranks, spin and orbital], $\mathbf{W}_0^{(K)} = [K]^{\frac{1}{2}}\mathbf{U}^{(K)}$.[2] For the fully reduced case, the appropriate operators are $\mathbf{X}^{(T)} = (\mathbf{a}^{(\gamma)}\mathbf{a}^{(\gamma)})^{(T)}$, $T \equiv \chi K$; χ, the quasi-spin rank, either zero or one.[2] The matrix elements of these operators display the inner product formula. In sufficiently general form,

$$(\Gamma \,\|X^{(T)}\|\, \Gamma') = (-)^{\Gamma+\Gamma'+T}[T]^{\frac{1}{2}} \sum_{\Gamma''} \begin{Bmatrix} \gamma & T & \gamma \\ \Gamma & \Gamma'' & \Gamma' \end{Bmatrix} \times (\Gamma \,\|a^{(\gamma)}\|\, \Gamma'')(\Gamma'' \,\|a^{(\gamma)}\|\, \Gamma'). \quad (2)$$

The properties of $\mathbf{X}^{(T)}$ may be explored through known properties of $\mathbf{a}^{(\gamma)}$ and $\mathbf{U}^{(K)}$. For this it is useful now to particularize the component, ω, of χ. The results are given by Judd.[2] If $\omega = 0$,

$$X_0^{(T)} = -2^{-\frac{1}{2}}(1-(-)^T)\mathbf{W}_0^{(K)} - (-)^{\chi}[I]^{\frac{1}{2}}\delta(K,0), \quad (3a)$$

and, for any ω, T even

$$X^{(T)} = -[I]^{\frac{1}{2}}\delta(T,0). \quad (3b)$$

But we also introduce $\mathbf{W}_{\pm}^{(K)}$, which involve two creation or two destruction operators. The three bilinear operators of this type are then

$$\mathbf{W}_0^{(K)} = -(a^+a)^{(K)}, \quad \mathbf{W}_+^{(K)} = +2^{-\frac{1}{2}}(a^+a^+)^{(K)}, \quad \mathbf{W}_-^{(K)} = -2^{-\frac{1}{2}}(aa)^{(K)}. \quad (4)$$

Comparable forms may be found in the references.[1,2]

It now becomes easy to rewrite Eq. (3) more explicitly. For odd T,

$$X_0^{(1K)} = -\sqrt{2}\,\mathbf{W}_0^{(K)} + [I]^{\frac{1}{2}}\delta(K,0), \quad (5a)$$

$$X_{\pm 1}^{(1K)} = \pm\sqrt{2}\,\mathbf{W}_{\pm}^{(K)}, \quad (5b)$$

$$X_0^{(0K)} = -\sqrt{2}\,\mathbf{W}_0^{(K)}; \quad (5c)$$

the correspondence relations for bilinear forms. But for even T no operators of type $\mathbf{W}$ occur. These relations are needed to find formulas useful in studying pairs of one-particle CFP. For even T,

$$X_0^{(0K)} = -[I]^{\frac{1}{2}}\delta(K,0), \quad (6a)$$

$$X^{(1K)} = 0, \quad \text{for any } \omega. \quad (6b)$$

Clearly Eq. (5b) relates to two-particle CFP as in Eq. (32), Ref. 3(b), and the fact that K is even in Eq. (5b) and odd in Eq. (6b) shows the allowed term values in l^2. Further, the even value of K in Eq. (5a) relates to Eqs. (67) and (69b) of Ref. 3(b), and the odd value of K in Eq. (5c) to Eqs. (69a) and (70) of Ref. 3(b). Equation (6b) with $\omega = \pm 1$ pertains to the original method for finding CFP, as in Ref. 3(b), Sec. 3, in particular Ref. 3(b), Eq. (11). It is with this and Eq. (6) generally (even T) that much of the following work is concerned in subsection (3). $\mathbf{X}^{(T)}$, T odd, need only be discussed again with emphasis on single values of intermediate v, in matrix products diagonal in v. This is suggested as an extension to nonzero K of the accurate orthogonality relations, Ref. 3(b), Eqs. (59) and (60).

The use of Eqs. (5) and (6) in concert to derive various relations in the seniority scheme will reveal

several connections among pairs of products of CFP. However, these results can be found through the work of Racah. It is in its simplicity that the quasi-spin method is attractive and useful.

The Rajnak–Wybourne formula derives from a trilinear form in $\mathbf{a}^{(\gamma)}$.[7] The analysis of all these forms, and larger ones (the energy among others), must be left for another occasion. The restriction to equivalent particles might still be maintained in such a study and the possibilities for various intermediate v explored.

(2) Matrix Elements of Bilinear Forms

The requirement of the present work is that it yield formulas which are useful in finding CFP in the seniority scheme, in particular, those for the "new terms" of Ref. 3(b), Sec. 6(4). The other parts of the theory are directly found once the CFP are known. The dependence on n is given by correspondence.

Two formulas, or methods, for finding CFP are that of Racah [Ref. 3(b), Eq. (11)] and that of Redmond.[4] (Though neither of these is found in the seniority scheme, it is formally simple to sum over the unspecified quantum numbers, as has often been done before, no doubt, in the past.) These formulas are expressed in terms of double tensors [Ref. 3(a), Sec. 5] as noted above. It is convenient to do so here also. The Clebsch–Gordan and Racah coefficients in quasi-spin are very simple indeed for linear and bilinear forms. They need not be made explicit any longer. However, they are shown in the first example, just below.

Those matrix elements and products which are diagonal in v are of great interest since two intermediate states, $v \pm 1$, may occur. The treatment here is restricted to such cases, but it is a very simple matter to find any nondiagonal case which may be needed.

For the promised example, we note that Eq. (5b) relates two-particle CFP in l^n and l^v. The usual demonstration stems from Ref. 3(b), Secs. 6(2) and 6(4). The emphasis placed below is also rather different from that in other recent work,[1] as seen in the appearance of Racah coefficients in quasi-spin. The fully reduced matrix element diagonal in v is, by Eq. (2),

$$(vI \,\|X^{(1K)}\|\, vI^2) = -(-)^{K+2Q+I+I^2}\sqrt{3}\,[K]^{\frac{1}{2}} \times \sum_{Q^1 I^1} \begin{Bmatrix} q & 1 & q \\ Q & Q^1 & Q \end{Bmatrix} \begin{Bmatrix} l & K & l \\ I & I^1 & I^2 \end{Bmatrix} \times (QI \,\|a^{(\gamma)}\|\, Q^1 I^1)(Q^1 I^1 \,\|a^{(\gamma)}\|\, QI^2). \quad (7)$$

This involves for even K a two-particle CFP in l^v, just as $\pm\sqrt{2}\mathbf{W}_{\pm}^{(K)}$ involves such a CFP in l^n, all according to the definition [Ref. 3(b), Eq. (32)]. The ratio of these is the Wigner–Eckart coefficient in quasi-spin. This equation results finally in a known correspondence relation,

$$(l^{n-2}vI^2l^2K|\}l^nvI) = \left(\frac{(v+1)(v+2)(Q-M+1)(Q+M)}{2n(n-1)Q}\right)^{\frac{1}{2}} \times (l^vvI^2l^2K|\}l^{v+2}vI). \quad (8)$$

If K is odd the CFP are zero [Ref. 3(b), Eq. (11)]. This case, part of Eq. (6b), is conveniently written for $\omega = -1$, and not necessarily diagonal in v,

$$\sum_{v^1 I^1} [I^1]^{\frac{1}{2}} \begin{Bmatrix} l & K & l \\ I^2 & I^1 & I \end{Bmatrix} (l^nv^2I^2|\}l^{n+1}v^1I^1) \times (l^{n+1}v^1I^1|\}l^{n+2}vI) = 0, \quad K \text{ odd}. \quad (9)$$

It can be quickly determined, by correspondence, that there is no formal difference here between a double tensor and a triple tensor formulation, so l^v has been replaced by l^n to yield Eq. (11) of Ref. 3(b) exactly. A convenient transformation of this is made in the next subsection.

At this stage, the matrix elements of Eqs. (5) and (6) should be further considered for particular values of v^1. We introduce the signs $\oplus$ and $\ominus$ to make this distinction for bilinear forms. Now from Ref. 3(b), Sec. 6(6), for any K,

$$(l^nvI \,\|U^{(K)}\ominus\|\, l^nvI') = +\frac{Q-M+1}{[Q]}(l^vvI \,\|U^{(K)}\|\, l^vvI'), \quad (10)$$

and the reduced submatrix element (cleared of n) appears on the right. The $\oplus$-sums depend upon the signature of K. For K odd,

$$(l^nvI \,\|U^{(K)}\oplus\|\, l^nvI') = +\frac{Q+M}{[Q]}(l^vvI \,\|U^{(K)}\|\, l^vvI'). \quad (11a)$$

For K even, not zero,

$$(l^nvI \,\|U^{(K)}\oplus\|\, l^nvI') = -\frac{Q+M}{[Q]}\frac{Q+1}{Q}(l^vvI \,\|U^{(K)}\|\, l^vvI'). \quad (11b)$$

If K is equal to zero, the result is that of Eq. (59b) [Ref. 3(b)], namely,

$$(l^nvI \,\|U^{(0)}\oplus\|\, l^nvI) = +\frac{Q+M}{[Q]}\frac{(v+2[Q])}{v}(l^vvI \,\|U^{(0)}\|\, l^vvI). \quad (12)$$

These forms are convenient and instructive. The two-particle CFP can be similarly expressed. For odd K,

[7] K. Rajnak and B. G. Wybourne, Phys. Rev. 132, 280 (1963).

the $\ominus$-sum is the negative of the $\oplus$-sum, which is

$$(l^{n-2}vI(ll)K\oplus|\}l^nvI') = (-)^{I-I'+1}\left(\frac{[K]}{[I']}\right)^{\frac{1}{2}}[Q]^{-1}\left(\frac{(Q+M)(Q-M+1)}{n(n-1)}\right)^{\frac{1}{2}} \times (l^vvI\,\|U^{(K)}\|\,l^vvI'), \quad (13a)$$

a portion of Eq. (11) in Ref. 3(b). For even K, we only note the usual CFP. If K is even, not zero,

$$(l^{n-2}vIl^2K|\}l^nvI') = (-)^{I-I'+1}\left(\frac{[K]}{[I']}\right)^{\frac{1}{2}}Q^{-1}\left(\frac{(Q+M)(Q-M+1)}{n(n-1)}\right)^{\frac{1}{2}} \times (l^vvI\,\|U^{(K)}\|\,l^vvI'). \quad (13b)$$

This is a rewriting of Eq. (8), useful for checking purposes. For K equal to zero, Eq. (49) of Ref. 3(b) applies. The separate sums, $\oplus$ and $\ominus$, are readily found for these relations.

The use of Eq. (6) has been anticipated in the work above which is essentially given by Racah [Ref. 3(b), Sec. 6(6)] as has been noted. In fact, the appearance of the above equations is readily changed by the use of correspondence. For later reference, we note the matrix elements of Eq. (6) if $\omega = 0$. For even K,

$$\tfrac{1}{2}([l][I])^{\frac{1}{2}}(Q+M)\delta(K,0) = +Q(l^nvI\,\|U^{(K)}\oplus\|\,l^nvI') + \frac{Q+M}{Q-M+1}(Q+1)(l^nvI\,\|U^{(K)}\ominus\|\,l^nvI'); \quad (14a)$$

and for odd K,

$$+(l^nvI\,\|U^{(K)}\oplus\|\,l^nvI') = +\frac{Q+M}{Q-M+1}(l^nvI\,\|U^{(K)}\ominus\|\,l^nvI'). \quad (14b)$$

It may be noted that to place n equal to v in these relations is to find a null result in each case.

With this material in hand, a second step can be undertaken. This is to carry out sums over K, thus obviating the conditions on the signature of K, which prevail in this subsection. The matrix products so obtained are characteristic of the unitary and associative properties of the Racah coefficients.

(3) Transformed Matrix Elements

Under certain restrictions then, these matrix elements can be summed over their ranks, K, odd or even. If they are first multiplied by a Racah coefficient (before summing), this restriction is that the two parameters in one of the triads on K be equal. This already obtains in the elements at hand. The sum rule of interest is very simple. It may be called a partial or incomplete orthogonality relation.[8] The double sign is for even and odd e, respectively,

$$\sum_e [e](1 \pm (-)^e)\begin{Bmatrix} c & a & f \\ a & d & e \end{Bmatrix}\begin{Bmatrix} c & a & g \\ a & d & e \end{Bmatrix} = [g]^{-1}\delta(f,g) \pm (-)^{f+g}\begin{Bmatrix} a & c & f \\ a & d & g \end{Bmatrix}. \quad (15)$$

It yields a matrix product quite different from the matrix element of Eq. (2) above. Its use may be said to display a third type of Racah transform,[5] if the use of the unitary and associative properties of the coefficients separately (it can be derived immediately from these) is considered a first and second type, respectively. The new parameter which appears in the multiplying coefficient for transforms of this sort is seen to be definitive of the CFP in question. The "principal"[9] or "arbitrary"[4] parent in fractional parentage formulas as found in these references does not emerge in the treatment given here.

This transform procedure can be applied to Eq. (9) [Ref. 3(b), Eq. (11)] to yield the very simple formula

$$+\sum_{v^*}(l^{n-2}v^2I^2l|\}^{n-1}v^*I^*)(l^{n-1}v^*I^*|\}l^nvI) = +\sum_{v^1I^1}(-)^{I^1+I^*}([I^1][I^*])^{\frac{1}{2}}\begin{Bmatrix} l & I^* & I^2 \\ l & I^1 & I \end{Bmatrix} \times (l^{n-2}v^2I^2|\}l^{n-1}v^1I^1)(l^{n-1}v^1I^1|\}l^nvI). \quad (16)$$

The definitive term here (and in the following formulas) is I^*. This formula is the companion to the well-known one below. An obvious resemblance to a well-known identity in Clebsch–Gordan coefficients[10] may be remarked in passing.

The two remaining parts of Eq. (6) are given in Eq. (14). Upon taking transforms according to Eq. (15) and combining the results, one can find, after some changes, another very simple formula, that of Redmond.[4,2] It is not necessarily diagonal in v, and so reads, displayed after the fashion of the last formula,

$$+(n+1)\sum_{v^*}(l^nvI|\}l^{n+1}v^*I^*)(l^{n+1}v^*I^*\{|l^nv'I') = +\delta(vI,v'I') + n(-)^{I+I'}([I][I'])^{\frac{1}{2}}\sum_{v^1I^1}\begin{Bmatrix} l & I^* & I \\ l & I^1 & I' \end{Bmatrix} \times (l^nvI\{|l^{n-1}v^1I^1)(l^{n-1}v^1I^1|\}l^nv'I'). \quad (17)$$

The two formulas, Eqs. (16) and (17), may be discussed in terms of χ and ω. If Q and Q' are not equal,

[8] M. Rotenberg, R. Bivins, N. Metropolis, and J. K. Wooten, *The 3-j and 6-j Symbols* (Technology Press, Cambridge, Massachusetts, 1959), p. 14. Note differences from Eq. (15) above.

[9] A. de-Shalit and I. Talmi, *Nuclear Shell Theory* (Academic Press Inc., New York, 1963), p. 274.

[10] L. C. Biedenharn, J. M. Blatt, and M. E. Rose, Rev. Mod. Phys. **24**, 249 (1952), Eq. (18).

then χ must be one. The two formulas are the same, granted reciprocity and correspondence. A change in ω has only that significance. For the diagonal case, $Q = Q'$, Eq. (17) contains both values of χ, but Eq. (16), of course, pertains to $\chi = 1$ only.

Interest in n is much less than in v. If we take Eq. (17) diagonal in v, then the setting of n equal to v removes one of the sums. After minor changes, to bring all CFP to lowest terms, a most useful formula is found, which can, however, also be found directly from Eq. (11), if desired.

$$(v+1)(l^v vI|\}l^{v+1}v + 1I^*)(l^{v+1}v + 1I^*\{|l^v vI') = +\delta(I,I') - v(-)^{I-I'}\frac{([I][I'])^{\frac{1}{2}}}{[Q][I^*]}(l^v vI\{|l^{v-1}v - 1I^*) \times (l^{v-1}v - 1I^*|\}l^v vI') - (l^v vI\{|I^*|\}l^v vI'). \quad (18)$$

The shorthand on the far right-hand side is, generally,

$$(l^n vI\{|I^*|\}l^n v'I') \equiv -n(-)^{I+I'}([I][I'])^{\frac{1}{2}} \times \sum_{v^1 I^1}\begin{Bmatrix} l & I^* & I \\ l & I^1 & I' \end{Bmatrix}(l^n vI\{|l^{n-1}v^1I^1)(l^{n-1}v^1I^1|\}l^n v'I'). \quad (19)$$

This definition is made to fit Eq. (17). The reduced form of it, given in Eq. (18), has the qualities of $(l^v vI \| U^{(K)} \| l^v v'I')$ shown in Eq. (10). Since only v and $v-1$ appear on the right-hand side, and only v and $v+1$ on the other, Eq. (18) is very well suited for explicit study of CFP in the seniority scheme.

(4) Applications

One-particle CFP have been calculated through all terms which occur as far as f^7 in the seniority scheme.[11] Many are known also in the j-j coupling scheme.[12] Nearly all recent calculations of any complexity have been made through the factoring of CFP, the method of Ref. 3(c). Formulas such as those above were not used since it was not known that they were compatible with useful classification schemes. They may, in fact, be used to calculate factors of CFP, as well as entire CFP, in a suitably chosen scheme and without any limiting dependence upon a principal parent, provided the relevant algebra is known. It is also expected that trilinear and other higher forms will be more and more important both in matrix elements and in CFP. Also new sum rules for matrix elements will be developed. Thus, it appears that application of the present methods will be exploratory in nature, in several directions.

The seniority scheme at present allows a free phase to every new term. This freedom is not altered by any quasi-spin considerations, and every new term, *QSL*, enters twice into the expressions given here. However matrix elements or two-particle CFP of even rank K and off-diagonal in v are phase indicators for new terms. The relative phase within each row [of a table of CFP as given, e.g., in Ref. 3(b)] are fixed, of course, as is also the over-all phase of any row but that for a new term.

III. CONCLUSIONS

The simplicity of the seniority scheme as now treated must prove useful in the future development of new classification schemes. The current work in quasi-spin shows the extent of the ideas on seniority given first in Refs. 3(a)–(c). In particular, formulas of the type first noted by Redmond have now found their natural place in the seniority scheme. The connection between these formulas and the method of calculating CFP offered in Ref. 3(b), Sec. 3, has also been established.

The introduction of fully reduced matrix elements (independent of the number of particles) emphasizes once again that important relations need only contain sets of angular momentum quantum numbers for the states involved and not their components. Thus, the idea of correspondence is but another expression of the Wigner–Eckart theorem.[1]

The quasi-spin technique applied to bilinear forms has uncovered several new connections which were implicit, however, in the work of Racah. The result is to simplify the calculation of CFP. Further results can be anticipated in the form of new sum rules for matrix elements and matrix products and new phase determinations in cases of configuration mixing.

ACKNOWLEDGMENTS

The author is indebted to B. R. Judd for timely correspondence on unpublished work and to R. D. Lawson for a conversation on quasi-spin methods. The treatment by G. Racah of Redmond's formula[13] is also of interest for the present standpoint.

[11] C. W. Nielson and G. F. Koster, *Spectroscopic Coefficients for the p^n, d^n, and f^n Configurations* (Technology Press, Cambridge, Massachusetts, 1963).

[12] In Ref. 9, for example.

[13] G. Racah, in *Nuclear Spectroscopy* (Academic Press Inc., New York, 1962), pp. 6, 7.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Theory of Diffraction and Matched Asymptotic Expansions

R. P. KANWAL
Pennsylvania State University, University Park, Pennsylvania

(Received 16 June 1966)

We discuss the theory of diffraction as a singular perturbation problem. Due to the presence of two different lengths—the wavelength and the characteristic geometric length of the scatterer—there exists no uniformly valid series expansion of the scattered field in the entire region. As such, two different expansions are needed; one of them is valid in the near region and the other in the far region. These two series are then matched in an appropriate fashion. It emerges from this analysis that first two terms of the required asymptotic expansion can be obtained by solving two simple problems of the potential theory.

1. INTRODUCTION

THE method of matched asymptotic expansions has grown in recent years. It has been applied to a variety of problems in fluid mechanics.[1] Recently it has also found its way to other branches of mechanics. We have recently studied the dynamic displacements in an elastic space with the help of this technique.[2] In physical problems a general warning of singular behavior can be based upon dimensional reasoning. It turns out that a perturbation solution is uniformly valid in space and time coordinates unless the perturbation parameter m is the ratio of two lengths.

The asymptotic nature of the theory of diffraction is well known.[3] In the study of this theory there arise two lengths—the characteristic geometric length and the wavelength. Their ratio gives the perturbation parameter m. The problem is of physical interest for small as well as large values of m. Here we study the case when m is small. We develop simultaneously two expansions—an inner and an outer expansion. The inner expansion is valid close to the diffracting obstacle and satisfies the boundary conditions on the surface of that obstacle. The outer expansion is valid in the far region and satisfies the radiation condition. The simple principle of constraining the coordinates then helps us in matching these two series in an appropriate region.

Although we have only considered the low-frequency expansions, this method is also applicable for high frequencies if we expand in the inverse powers of m and match the different series appropriately. The latter problem has recently been studied by Ursel,[4] who has rigorously established the asymptotic nature of the ray optics as the short wave limit of the wave optics. His analysis combined with the existence of two characteristic lengths in the theory prove the applicability of the present technique to the short wave asymptotics as well.

2. THE MATHEMATICAL FORMULATION

The classic spatial diffraction problem consists of finding a function ϕ^s exterior to a finite obstacle with a smooth boundary B. This function satisfies the Helmholtz equation

$$(\nabla^2 + k^2)\phi^s = 0, \tag{1}$$

and obeys the boundary condition

$$\phi^s = -\phi^i \quad \text{or} \quad \partial\phi^s/\partial n = -\partial\phi^i/\partial n, \tag{2}$$

at the boundary B; and obeys the radiation condition

$$\lim_{r\to\infty} r\left(\frac{\partial}{\partial r}\phi^s - ik\phi^s\right) = 0. \tag{3}$$

The function ϕ^i gives the incident field which is known everywhere including the boundary B; while ϕ^s gives the scattered field. The quantity k is the wavenumber

$$k = 2\pi/\lambda, \tag{4}$$

and λ is the wavelength.

Let us now nondimensionalize Eq. (1) by introducing a characteristic geometric length a such that the position vector $\mathbf{r}$ is related to the nondimensional vector $\mathbf{r}'$ as

$$\mathbf{r}' = \mathbf{r}/a. \tag{5}$$

Equation (1) then becomes (after dropping the primes)

$$\Delta^2\phi^s + m^2\phi^s = 0, \tag{6}$$

where, of course, ϕ^s has also been nondimensionalized appropriately depending upon the physical nature of ϕ^s. The number m is the same number as we recently encountered in elastodynamics.[2] It is the ratio of the geometric length to the wavelength. This analysis is based on the assumption that $m \ll 1$, i.e., the geometric length is much smaller than the wavelength.

[1] M. Van Dyke, *Perturbation Methods in Fluid Mechanics* (Academic Press Inc., New York, 1964).
[2] R. P. Kanwal, J. Math. Phys. **44**, 275 (1965).
[3] K. O. Friedrichs, Bull. Am. Math. Soc. **61**, 485 (1955).
[4] F. Ursel, Proc. Cambridge Phil. Soc. **62**, 227 (1966).

Let the inner expansion of ϕ^s be given as

$$\phi^s = \sum_{n=0}^{\infty} \phi^s_{(n)}(m)^n. \tag{7}$$

Since ϕ^s_n's are supposed to satisfy only the inner boundary condition (2), we expand ϕ^i also accordingly:

$$\phi^i = \sum_{n=0} \phi^i_{(n)}(m)^n. \tag{8}$$

When we substitute the series (7) and (8) into Eqs. (6) and (2) (in its nondimensional form) we obtain the following system of equations for different orders of m:

$O(1)$ $$\nabla^2\phi^s_{(0)} = 0; \tag{9}$$

$$\phi^s_{(0)} = -\phi^i_{(0)} \quad \text{or} \quad \partial\phi^s_{(0)}/\partial n = -\partial\phi^i_{(0)}/\partial n, \quad \text{at } B; \tag{10}$$

$O(m)$ $$\nabla^2\phi^s_{(1)} = 0; \tag{11}$$

$$\phi^s_{(1)} = -\phi^i_{(1)} \quad \text{or} \quad \partial\phi^s_{(1)}/\partial n = -\partial\phi^i_{(1)}/\partial n, \quad \text{at } B; \tag{12}$$

$O(m^2)$ $$\nabla^2\phi^s_{(2)} - \phi^s_{(0)} = 0; \tag{13}$$

$$\phi^s_{(2)} = -\phi^i_{(2)} \quad \text{or} \quad \partial\phi^s_{(2)}/\partial n = -\partial\phi^i_{(2)}/\partial n, \quad \text{at } B. \tag{14}$$

Thus, to the order $O(1)$ and $O(m)$, we get the classic problems in potential theory.

To study the far region we constrain the coordinates and set

$$\tilde{x} = mx, \quad \tilde{y} = my, \quad \tilde{z} = mz. \tag{15}$$

Equation (6) then becomes

$$\tilde{\nabla}^2\phi^s + \phi^s = 0. \tag{16}$$

If we define the outer expansion for ϕ^s as

$$\phi^s = \sum_{n=0}^{\infty} \psi_{(n)}(m)^n, \tag{17}$$

and substitute in (16) then all the ψ's satisfy the equation

$$\tilde{\nabla}^2\psi_{(n)} + \psi_{(n)} = 0. \tag{18}$$

We take

$$\psi_{(0)} = 0 \tag{19}$$

and

$$\psi_{(n)} = e^{i\tilde{r}}/\tilde{r}. \tag{20}$$

The radiation condition (3) (in its nondimensional form) is thus satisfied for ϕ^s.

There remains the matching problem. It fortunately reduces to certain regularity conditions on $\phi^s_{(n)}$ at infinity. To appreciate this point let us substitute (20) into (17), rewrite the terms in the variables x, y, z; and finally expand ϕ^s in increasing powers of m:

$$\phi^s_{(\text{outer})} = 0 + m\frac{e^{imr}}{mr} + m^2\frac{e^{imr}}{mr} + m^3\frac{e^{imr}}{mr} + \cdots \tag{21}$$

or

$$\phi^s_{(\text{outer})} = \frac{1}{r} + \left(i + \frac{1}{r}\right)m + \left(-\tfrac{1}{2}r + i + \frac{1}{r}\right)m^2 + O(m^3). \tag{22}$$

Comparing (7) and (22) we see that $\phi^s_{(n)}$'s have to approach certain simple functions in the far region. For $\phi^s_{(0)}$ and $\phi^s_{(1)}$ we have very simple regularity conditions. To sum up: The first two terms of the inner expansion, i.e., $\phi^s_{(0)}$ and $\phi^s_{(1)}$ are the solutions of the classic potential equations satisfying simple regularity conditions at infinity. Such solutions are rather well known. For example, similar equations and boundary conditions arise in the study of polarization potential and virtual mass[5]; and there the solutions have been given for many interesting shapes of the obstacles.

ACKNOWLEDGMENTS

This work was sponsored by a grant from the National Science Foundation.

[5] M. Schiffer and G. Szegö, Trans. Am. Math. Soc. 67, 130 (1949).

Closure Relations for the Eigenfunctions of the One-Speed Transport Equation

I. KUŠČER* AND F. SHURE
Department of Nuclear Engineering, The University of Michigan, Ann Arbor, Michigan

(Received 13 May 1966)

Orthogonality relations for the eigenfunctions of the one-speed transport equation are used to derive the corresponding closure relations. These express in a concise form the completeness properties previously proved by Case.

I. INTRODUCTION

THE plane-symmetric normal mode solutions to the equation of steady-state one-speed neutron transport with isotropic scattering possess a variety of completeness properties. These were all proved by Case[1] via construction of a solution to the singular integral equation for the expansion coefficients implied by a hypothetical representation of an arbitrary function in terms of some set of the eigenfunctions. In such an approach the form of the expansion coefficients appears as a by-product of the completeness proof. In a later paper[2] (henceforth referred to as I), a set of orthogonality relations was presented, which may be used to determine expansion coefficients directly. On the other hand, as we show, these relations also provide a completeness proof in a more conventional form. That is, by the use of our knowledge of the normalization coefficients contained in the orthogonality relations, we exhibit the closure property for Case's eigenfunctions by direct calculation. Acquaintance with paper I is assumed, and definitions and notation are borrowed from there.

First, in Sec. II we show that the set of continuum eigenfunctions $\{\phi_\nu(\mu)\}$, $\alpha \leq \nu \leq \beta$, obey a closure relation with respect to the interval $[\alpha, \beta]$ whenever α and β both lie in the open interval $(-1, 1)$. Then, in Sec. III we discuss the lack of uniqueness or "overcompleteness" that arises when α and β are both positive or both negative, and indicate the modifications necessary when α or β take on the special values ± 1.

II. PARTIAL RANGE COMPLETENESS

The expansion coefficient $A(\nu)$ in

$$\psi(\mu) = \int_\alpha^\beta A(\nu)\phi_\nu(\mu)\, d\nu \tag{1}$$

is conveniently calculated by application of the orthogonality relation [Eq. (I.22)]

$$\int_\alpha^\beta \phi_\nu(\mu)\phi_{\nu'}(\mu)\gamma_0(\mu)\, d\mu = \gamma_0(\nu)\Lambda^+(\nu)\Lambda^-(\nu)\delta(\nu - \nu'). \tag{2}$$

(Here and henceforth we choose for the principal-value integral with two singularities in the integrand the same interpretation as used by Case.[1,3]) From the right-hand side of Eq. (1) we obtain

$$\int_\alpha^\beta \phi_{\nu'}(\mu)\gamma_0(\mu)\, d\mu \int_\alpha^\beta A(\nu)\phi_\nu(\mu)\, d\nu = A(\nu')\gamma_0(\nu')\Lambda^+(\nu')\Lambda^-(\nu'). \tag{3}$$

The procedure is legitimate whenever the functions involved are such that the order of integration on the left-hand side of the last equation can be reversed by application of the Poincaré–Bertrand formula.[4]

If the above method is formally applied to the function $\psi(\mu) = \delta(\mu - \mu')$, a closure relation immediately follows:

$$\int_\alpha^\beta \phi_\nu(\mu)\phi_\nu(\mu')\frac{\gamma_0(\mu')}{\gamma_0(\nu)}\frac{d\nu}{\Lambda^+(\nu)\Lambda^-(\nu)} = \delta(\mu - \mu'), \quad \mu, \mu' \in [\alpha, \beta], \quad -1 < \alpha < \beta < 1. \tag{4}$$

Our main goal is to prove this identity in a direct way. We may remark that the convergence of the integral for $\mu \neq \mu'$ is ensured since the factor $\nu[\gamma_0(\nu)]^{-1}$ possesses at most a weak (i.e., integrable) singularity (cf., the discussion at the beginning of Sec. III), whereas the factor $[\Lambda^+\Lambda^-]^{-1}$ is well behaved.

Substituting the expression (I.2c) [i.e., Eq. (2c) of Ref. I. This notation is applied several times throughout this paper] for $\phi_\nu(\mu)$, we obtain two δ-function contributions from the left-hand side of Eq. (4): a term

$$[\lambda^2(\mu)/\Lambda^+(\mu)\Lambda^-(\mu)]\delta(\mu - \mu')$$

due to the product of the δ functions appearing explicitly in ϕ_ν, and a term

$$[\pi c^2\mu^2/4\Lambda^+(\mu)\Lambda^-(\mu)]\delta(\mu - \mu')$$

* Permanent address: Institute of Physics, University of Ljubljana, Ljubljana, Yugoslavia.

[1] K. M. Case, Ann. Phys. (N.Y.) **9**, 1 (1960).

[2] I. Kuščer, N. J. McCormick, and G. C. Summerfield, Ann. Phys. (N.Y.) **30**, 411 (1964).

[3] I. Kuščer and N. J. McCormick, Nucl. Sci. Eng. **23**, 404 (1965).

[4] N. I. Muskhelishvili, *Singular Integral Equations* (P. Noordhoff, Ltd., Groningen, The Netherlands, 1953).

arising from the application of Eq. (I.9). According to Eq. (I.4b), both contributions together give just the right-hand side of (4).

We are left with the task of showing that the remaining part of the integral in (4) vanishes. Since this part has the form

$$\gamma_0(\mu')[I(\mu) - I(\mu')]/(\mu - \mu'), \tag{5}$$

where

$$I(\mu) = \frac{c\mu}{2}\frac{\lambda(\mu)}{\gamma_0(\mu)\Lambda^+(\mu)\Lambda^-(\mu)} + P\int_\alpha^\beta \left(\frac{c\nu}{2}\right)^2 \frac{1}{\gamma_0(\nu)\Lambda^+(\nu)\Lambda^-(\nu)}\frac{d\nu}{\nu - \mu}, \tag{6}$$

we only need to demonstrate that $I(\mu)$ is a constant.

In view of the identities

$$\frac{1}{X_0^+(\nu)} \pm \frac{1}{X_0^-(\nu)} = \frac{c\nu}{2}\frac{\Lambda^-(\nu) \pm \Lambda^+(\nu)}{\gamma_0(\nu)\Lambda^+(\nu)\Lambda^-(\nu)} = \begin{cases} \dfrac{c\nu\lambda(\nu)}{\gamma_0(\nu)\Lambda^+(\nu)\Lambda^-(\nu)}, & (7+) \\ -2\pi i\left(\dfrac{c\nu}{2}\right)^2 \dfrac{1}{\gamma_0(\nu)\Lambda^+(\nu)\Lambda^-(\nu)}, & (7-) \end{cases}$$

which are obtained from Eqs. (I.4b) and (I.17), the integral term on the right-hand side of (6) is equal to

$$-\tfrac{1}{2}[N_0^+(\mu) + N_0^-(\mu)], \tag{8}$$

where

$$N_0(z) \equiv \frac{1}{2\pi i}\int_\alpha^\beta \left[\frac{1}{X_0^+(\nu)} - \frac{1}{X_0^-(\nu)}\right]\frac{d\nu}{\nu - z}. \tag{9}$$

This may be written also as

$$N_0(z) = \frac{1}{2\pi i}\int_C \frac{1}{X_0(z')}\frac{dz'}{z' - z}, \tag{10}$$

where the integration is carried out in the clockwise direction over a closed loop which surrounds the cut (α, β) but leaves the point z outside. We deform this contour into a large circle to obtain

$$N_0(z) = [1/X_0(z)] - 1, \tag{11}$$

where we have used the fact that $X_0(z)$ possesses no zeros in the cut plane and tends to 1 as $z \to \infty$. Therefore,

$$\tfrac{1}{2}[N_0^+(\mu) + N_0^-(\mu)] = \frac{1}{2}\left[\frac{1}{X_0^+(\mu)} + \frac{1}{X_0^-(\mu)}\right] - 1, \tag{12}$$

and by the use of identity (7+) we see that the variable term here just cancels the other term in (6). Thus

$$I(\mu) = 1 \tag{13}$$

remains, so that the expression (5) vanishes, and the closure relation (4) is verified.

III. OVERCOMPLETENESS AND UNDERCOMPLETENESS

We still have to see under what conditions the closure relation (4) is unique, and what happens when $\alpha = -1$ or $\beta = 1$, or both. To this end we reexamine the proof of the preceding section to ascertain when it may be, or may need to be, modified.

Since closure implies the mutual orthogonality of all the eigenfunctions involved, the question of uniqueness is answered simply by trying out all weight functions which produce orthogonality. These functions are listed in paper I; and if they are derived by the constructive method of Case and Zweifel,[5] it can be seen that the list is complete.

The analysis hinges on the behavior of the function $\gamma_0(\nu)$ on the interval $[\alpha, \beta]$ over which it is defined. It can be seen[1] that $\gamma_0(\nu)$ is real finite, and nonvanishing in the interior of the interval, except for the trivial zero at $\nu = 0$. At the end points this function has, at most, weak zeros or weak infinities. That is to say,

$$\gamma_0(\nu) \sim (\beta - \nu)^{\tau(\beta)}, \quad \nu \to \beta, \qquad \sim (\nu - \alpha)^{-\tau(\alpha)}, \quad \nu \to \alpha. \tag{14}$$

As long as $\alpha > -1$ and $\beta < 1$, the exponents here are < 1 in magnitude. In fact $\tau(x)$ is an odd monotone function with $\tau(1) = 1$.

In view of the relation

$$\phi_\nu(-\mu) = \phi_{-\nu}(\mu), \tag{15}$$

no generality is lost by restricting ourselves to the case $\beta > 0$. Let us try, still for $-1 < \alpha < \beta < 1$, to repeat the derivation of Eq. (4), with $X_0(z)$ substituted by an a arbitrary X function obeying the conditions mentioned in I:

$$X(z) = \frac{(a - bz)(c - dz)}{(z - \alpha)(\beta - z)}X_0(z), \quad \text{if} \quad \alpha < 0, \tag{16a}$$

$$X(z) = \frac{(a - bz)}{(\beta - z)}X_0(z), \quad \text{if} \quad \alpha \geq 0, \tag{16b}$$

with arbitrary a, b, c, d. The corresponding functions $\gamma(\nu)$ and $N(z)$ are defined by Eq. (I.15a) and by an equation of the form (9), respectively. In the analog of Eq. (11), an additional term containing z crops up, due to the residues of the integrand at $z = a/b$ and $z = c/d$ (at infinity if $b = 0$ or $d = 0$). Consequently the term corresponding to (5) now does not vanish.

We have tacitly avoided the choice $a/b = c/d = \alpha$ or β in Eq. (16a) or $a/b = \alpha$ in (16b). In view of Eq. (14), whenever $\alpha \leq 0$, such a choice causes the

[5] K. M. Case and P. F. Zweifel, *Linear Transport Theory* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1967).

function $\gamma(\nu)$ to acquire a strong zero at the respective end point, which prevents the integral in Eq. (4) from converging. Thus, although orthogonality is produced by three linearly independent weight functions if $\alpha < 0$, or by two if $\alpha = 0$, in these two cases $\gamma_0(\nu)$ is the only such function allowed in the closure relation (4). This agrees with Case's statement[1] that for any such interval the expansion (1) is unique.

An apparent paradox arises when orthogonality relations with different weight functions are applied to Eq. (1). In contradiction to the established uniqueness of this expansion, different expansion coefficients $A(\nu)$ seem to follow. However, a closer inspection shows that, although orthogonality relations are formally valid for any of those weight functions $\gamma(\mu)$, only one of them, $\gamma_0(\mu)$, can be used for the determination of expansion coefficients. The trouble stems from the weak infinities of $A(\nu)$ and of the "useless" weight functions at one of the end points, where, e.g., a behavior like $(\beta - \nu)^{-\tau(\beta)}$ and $(\beta - \mu)^{-1+\tau(\beta)}$ is encountered. The two exponents add up to -1, whereby the Poincaré–Bertrand formula breaks down,[4] and consequently the equation corresponding to (3) becomes invalid. (The integral over μ on the left-hand side diverges.)

The situation is different if $\alpha > 0$, because then the function

$$\gamma_1(\nu) \equiv [(\nu - \alpha)/(\beta - \nu)]\gamma_0(\nu) \tag{17}$$

is a valid substitute for $\gamma_0(\nu)$, and we indeed get two closure relations. Similarly, for an arbitrary $\psi(\mu)$ we obtain two different expansions (1), one from the orthogonality relation with the weight $\gamma_0(\mu)$, and the other with $\gamma_1(\mu)$. Their difference gives an expansion with the sum zero,

$$\int_\alpha^\beta A(\nu)\phi_\nu(\mu)\,d\nu = 0, \tag{18}$$

$$A(\nu) \propto \frac{c\nu}{2}\,\frac{1}{(\nu - \alpha)\gamma_0(\nu)\Lambda^+(\nu)\Lambda^-(\nu)}, \tag{19}$$

as can be verified directly. We may say that the set $\{\phi_\nu(\mu)\}$, $\alpha \leq \nu \leq \beta$, now is overcomplete (not linearly independent) in the interval $\alpha \leq \mu \leq \beta$.[1]

It should be emphasized that for $0 < \alpha < \beta < 1$ only the two weight functions $\gamma_0(\mu)$ and $\gamma_1(\mu)$ are useful for determining the expansions of $\psi(\mu)$. Linear combinations of γ_0 and γ_1, although formally permitted in the orthogonality relation (2), fail in Eq. (3), because again the Poincaré–Bertrand formula breaks down at one of the end points. For the same reason, none of the orthogonality relations can be applied to linear combinations of the two expansions. In particular, this warning holds for Eqs. (18), (19), which therefore do not contradict orthogonality.

Whenever $\beta = 1$, the proof given in Sec. II breaks down in its first step, since $\gamma_0(\nu) \sim (1 - \nu)$ in the neighborhood of $\nu = 1$ and the integral in (4) diverges. However, in the particular case $0 < \alpha < \beta = 1$ the closure relation (4) is valid with $\gamma_0(\nu)$ replaced by $\gamma_1(\nu)$, Eq. (17). Of course, the proof is the same as before. Hence, here again we have unique expansions in terms of the continuum modes alone, in agreement with Case.

We consider finally the case $\beta = 1$, $\alpha \leq 0$, and try to use again a general X function and a corresponding γ function. It turns out that no such combination leads to a closure relation of the form (4), because either the integral there diverges, or a residue of the reciprocal X function leads to an additional term. We are thus forced to the conclusion that the continuum eigenfunctions alone no longer form a complete set.[1]

However, the set $\{\phi_\nu\}$ may be made complete by the addition of only one (two, in the case $\alpha = -1$, $\beta = 1$) new function, linearly independent of the continuum $\{\phi_\nu\}$. One such function is the discrete eigenfunction

$$\phi_+(\mu) = \tfrac{1}{2}c\gamma_0[1/(\nu_0 - \mu)], \tag{20}$$

which is automatically introduced by repeating the derivation of Sec. II with

$$X_2(z) \equiv [(\nu_0 - z)/(1 - z)]X_0(z), \tag{21}$$

and with the corresponding functions $\gamma_2(\mu)$ and $N_2(z)$. (For simplicity we exclude the case $c = 1$, so that $\nu_0 \neq \infty$.) When proceeding from Eq. (10) to Eq. (11) we get an extra term, namely the residue of the integrand at $z' = \nu_0$. Let us quote the final result for the half-range ($\alpha = 0$, $\beta = 1$):

$$\gamma_2(\mu')\left[\int_0^1 \phi_\nu(\mu)\phi_\nu(\mu')\,\frac{d\nu}{\gamma_2(\nu)\Lambda^+(\nu)\Lambda^-(\nu)} + \phi_+(\mu)\phi_+(\mu')\left(\frac{2}{c\nu_0}\right)^2 \frac{\nu_0 - 1}{X_0(\nu_0)}\right] = \delta(\mu - \mu'). \tag{22}$$

This agrees with what follows from the orthogonality relations (I.A1), (I.A2), (I.A4), if the different notation used there is taken into account:

$$\gamma(z) = \gamma_0(z)/(1 - z) = \gamma_2(z)/(\nu_0 - z).$$

In the full-range case ($\alpha = -1$, $\beta = 1$), two discrete terms are needed for completion, for instance the two discrete eigenfunctions $\phi_+(\mu)$ and $\phi_-(\mu)$. These are introduced by taking

$$X_3(z) \equiv \frac{\nu_0^2 - z^2}{1 - z^2}\,X_0(z) = \frac{\Lambda(z)}{1 - c}. \tag{23}$$

Pursuing the same procedure as before we obtain

$$\mu'\left\{\int_{-1}^{1}\phi_\nu(\mu)\phi_\nu(\mu')\frac{d\nu}{\nu\Lambda^+(\nu)\Lambda^-(\nu)} + [\phi_+(\mu)\phi_+(\mu') - \phi_-(\mu)\phi_-(\mu')]\frac{2}{c\nu_0^2\Lambda'(\nu_0)}\right\} = \delta(\mu-\mu'), \quad (24)$$

where

$$\Lambda'(\nu_0) \equiv \left[\frac{d\Lambda(z)}{dz}\right]_{z=\nu_0} = \frac{c\nu_0}{\nu_0^2-1} - \frac{1}{\nu_0}. \quad (25)$$

The result is seen to correspond to the known orthogonality relations for the full range case.[1]

Recollecting what has been said about completeness of the eigenfunctions we recover the classification given by Case. In short, we can state that the set of continuum modes $\phi_\nu(\mu)$, $\alpha \le \nu \le \beta$, $-1 \le \alpha < \beta \le 1$, with respect to expansions in the interval $\alpha \le \mu \le \beta$,

is overcomplete, whenever this interval does not contain any of the three special points -1, 0, 1;

is complete if the interval contains one of those points;

is undercomplete if it contains two (or all three) of those points, and can be completed by addition of one (or two) further function, e.g., $\phi_+(\mu)$ or $\phi_-(\mu)$ (or both).

IV. COMMENTS

The discussed closure relations are not entirely new, certainly not for the full range and the half range [Eqs. (24) and (22)]. For these two cases they are encountered in connection with the Green's function for the infinite medium and with the solution of the albedo problem for the semi-infinite medium, respectively.[1,5] Hence, the above considerations represent nothing more than a demonstration that the δ function is truly reproduced by the expansions. So is any other function which we would expand.

Instead of deriving the closure relation from the orthogonality relations, one could alternatively prove the closure relation first, and derive the orthogonality relations therefrom. The factor needed in the integrand of Eq. (4) can be constructed in a way very close to that used by Case and Zweifel[5] for the orthogonality relations. Such an approach may have some pedagogic value as a short way of introducing Case's formalism from the start. A brief sketch is given in the following.

The eigenfunctions $\phi_\nu(\mu)$ satisfy the equation

$$(\nu-\mu)\phi_\nu(\mu) = \tfrac{1}{2}c\nu, \quad (26)$$

which we write down twice, for the values μ and μ'. We multiply both sides of the first equation by

$$(2/c\nu)\phi_\nu(\mu')G(\nu)d\nu,$$

and of the second by

$$(2/c\nu)\phi_\nu(\mu)G(\nu)d\nu,$$

where $G(\nu)$ is a function to be determined later. After integration and subtraction we obtain

$$(\mu-\mu')\int_\alpha^\beta \frac{2}{c\nu}\phi_\nu(\mu)\phi_\nu(\mu')G(\nu)\,d\nu = \int_\alpha^\beta \phi_\nu(\mu)G(\nu)\,d\nu - \int_\alpha^\beta \phi_\nu(\mu')G(\nu)\,d\nu. \quad (27)$$

If a closure relation with the "weight" $(2/c\nu)G(\nu)$ exists, the integral on the left-hand side is zero for $\mu \neq \mu'$. Hence we must require that

$$\int_\alpha^\beta \phi_\nu(\mu)G(\nu)\,d\nu = \text{const.} \quad (28)$$

This is a singular integral equation, like those considered by Case. By applying the usual technique[1,5] we find that the solutions $G(\nu)$ indeed are the same as implied by the results of Secs. II and III.

The difference between Eq. (28) and the equation involved in the constructive determination of the orthogonality relations[5] lies in the interchanged role of the variables ν and μ, which means that the equations are the adjoints of each other, in Muskhelishvili's sense.[4] This explains why the number of linearly independent orthogonality relations in general differs from the number of linearly independent closure relations.

Let us conclude with remarks about possible generalizations. A closure relation for the two-media case[2] can immediately be written down. Also the more general scheme with a variable $c(\nu)$, useful with a simple model of energy-dependent neutron transport or of nongrey radiative transfer,[6,7] can be worked out without difficulty. In addition, a generalization to anisotropic scattering[8,9] is also possible.

ACKNOWLEDGMENTS

We owe sincere thanks to Professor P. F. Zweifel and Professor G. C. Summerfield for enlightening discussion and encouragement, and for critical comment which enabled us to avoid some inconsistency.

One of us (I. K.) wishes to express appreciation of the department of Nuclear Engineering, University to Michigan, for the hospitality extended to him during his stay

[6] R. J. Bednarz and J. R. Mika, J. Math. Phys. 4, 1285 (1963); J. R. Mika, Nucl. Sci. Eng. 22, 235 (1965).
[7] J. C. Stewart, I. Kuščer, and N. J. McCormick (to be published).
[8] J. R. Mika, Nucl. Sci. Eng. 11, 415 (1961).
[9] N. J. McCormick and I. Kuščer, J. Math. Phys. 7, 2036 (1966).

Application of Houston's Method to the Sum of Plane Waves over the Brillouin Zone. II. Body-Centered Cubic Lattice

MARIA MIĄSEK
Institute of Physics, Polish Academy of Science, Warsaw, Poland

HELENA DODZIUK
Institute of Organic Chemistry, Polish Academy of Science, Warsaw, Poland

(Received 14 September 1966)

In a previous paper we considered the function $\chi(\mathbf{r}) \equiv (1/N) \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}}$, where the sum runs over the first Brillouin zone of a crystal, and its expansion into series of Cubic Harmonics K_j: $\chi(\mathbf{r}) = \sum_{j=0}^{\infty} g_j(r) \times K_j(\theta, \varphi)$. Houston's method was used in order to find the radial functions $g_j(r)$ for several values of j, for $\chi(\mathbf{r})$ given for the simple cubic and the face-centered cubic lattices. In this paper, the same considerations are applied to $\chi(\mathbf{r})$ given for the body-centered lattice. $g_j(r)$, with $j = 0, 2, 3$, are calculated in the region of small r which is assumed as $0 \leq r \leq 2a$, where a is the lattice constant. In most of the problems of solid-state physics, where the function $\chi(\mathbf{r})$ occurs, it is satisfactory to know its values only for small r, usually not larger than $2a$. The function $g_0(r)$ is calculated using 3-, 6-, and 9-term expansion formulas, $g_2(r)$ and $g_3(r)$ using only 3- and 6-term formulas. Comparing $g_j(r)$ obtained from the formulas with different number of terms it is established that, for r in the region $\langle 0, 2a \rangle$, the 6-term approximation is very good.

1. INTRODUCTION

IN a previous paper[1] we discussed a simple way of integration of the following integrals

$$W = \frac{1}{N} \sum_{\mathbf{k}} \int e^{i\mathbf{k}\cdot\mathbf{r}} \Phi^*(\mathbf{r})\, d\mathbf{r}, \tag{1.1}$$

which occur in certain problems of solid-state physics. The sum over $\mathbf{k}$ runs over the first Brillouin zone of a crystal and we are interested only in cubic structures. In (1.1) we used the name $\chi(\mathbf{r})$ for a function

$$\chi(\mathbf{r}) = \frac{1}{N} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}}. \tag{1.2}$$

This function transforms according to the irreducible representation Γ_1 of the cubic group O_h. The value of W is different from zero only for functions $\Phi(\mathbf{r})$ that also possess Γ_1 symmetry. We have been interested only in such functions.

It was shown that W might be calculated easily when both functions in integrand $\chi(\mathbf{r})$ and $\Phi(\mathbf{r})$ were expanded into series of Cubic Harmonics K_j for Γ_1 representation. $\Phi(\mathbf{r})$ is usually expressed as a combination of atomic orbitals and its expansion into series of Cubic Harmonics is easily established. Our main purpose was to find the radial functions $g_j(r)$ in the expansion of $\chi(\mathbf{r})$

$$\chi(\mathbf{r}) = \sum_j g_j(r) K_j(\theta, \phi). \tag{1.3}$$

It was said before that in most of the solid-state problems the functions $\Phi(\mathbf{r})$ are determined only in the region of small $\mathbf{r}$ so that the integration in (1.1) goes only over small region in the space. Therefore we were interested in the values of $g_j(r)$ only for small r. The region of calculation was assumed as $0 \leq r \leq 2a$, where a is the lattice constant. Furthermore it was established that in such a region only $g_j(r)$ with lowest j's are of importance. Houston's method appeared very useful in order to find $g_j(r)$ with lowest j's and for small r.[2]

In I Houston's method was used in order to find $g_j(r)$ with $j = 0, 2, 3$ in the case of simple cubic and face-centered cubic lattices. This paper deals with the function $\chi(\mathbf{r})$ for body-centered cubic lattice. We calculate the function g_0, g_2, and g_3. In order to check the proper behavior of these $g_j(r)$ for small r we calculate them using expansion formulas with different numbers of terms [Paper I; (2.2), (2.2'), (2.2''), (2.3), (2.3'), (2.4), (2.4')].

2. THE FUNCTIONS $g_j(r)$ FOR A SUM OF PLANE WAVES IN A BODY-CENTERED CUBIC LATTICE

The analytical expression for the function $\chi(\mathbf{r})$ is found from [1; (3.7)] by exact integration over the volume of the Brillouin zone for the body-centered cubic lattice which is given as a rhombic dodecahedron.

[1] M. Miąsek, J. Math. Phys. 7, 139 (1966), hereafter referred to as I.

[2] W. V. Houston, Rev. Mod. Phys. 20, 161 (1948).

TABLE I. The function $g_0(\alpha)/(4\pi)^{\frac{1}{2}}$ for the body-centered cubic structure.

α	$g_0^{(3)}(\alpha)/(4\pi)^{\frac{1}{2}}$	$g_0^{(6)}(\alpha)/(4\pi)^{\frac{1}{2}}$	$g_0^{(9)}(\alpha)/(4\pi)^{\frac{1}{2}}$
0.00	1.0000	1.0000	1.0000
0.05	0.9938	0.9938	0.9938
0.10	0.9755	0.9755	0.9755
0.15	0.9456	0.9456	0.9456
0.20	0.9048	0.9048	0.9048
0.25	0.8542	0.8542	0.8542
0.30	0.7952	0.7952	0.7952
0.35	0.7292	0.7292	0.7292
0.40	0.6579	0.6579	0.6579
0.45	0.5831	0.5831	0.5831
0.50	0.5065	0.5065	0.5065
0.55	0.4299	0.4299	0.4299
0.60	0.3549	0.3549	0.3549
0.65	0.2831	0.2831	0.2831
0.70	0.2159	0.2158	0.2158
0.75	0.1544	0.1542	0.1542
0.80	0.0994	0.0992	0.0992
0.85	0.0517	0.0513	0.0513
0.90	0.0116	0.0111	0.0111
0.95	−0.0206	−0.0214	−0.0214
1.00	−0.0452	−0.0462	−0.0462
1.05	−0.0625	−0.0638	−0.0638
1.10	−0.0729	−0.0747	−0.0747
1.15	−0.0772	−0.0795	−0.0795
1.20	−0.0761	−0.0790	−0.0790
1.25	−0.0707	−0.0742	−0.0743
1.30	−0.0618	−0.0661	−0.0661
1.35	−0.0504	−0.0555	−0.0555
1.40	−0.0374	−0.0433	−0.0434
1.45	−0.0237	−0.0305	−0.0306
1.50	−0.0102	−0.0178	−0.0179
1.55	0.0026	0.0059	−0.0060
1.60	0.0140	0.0048	0.0046
1.65	0.0237	0.0138	0.0136
1.70	0.0313	0.0210	0.0207
1.75	0.0366	0.0260	0.0257
1.80	0.0397	0.0290	0.0286
1.85	0.0406	0.0301	0.0295
1.90	0.0395	0.0294	0.0287
1.95	0.0366	0.0271	0.0264
2.00	0.0323	0.0237	0.0228

TABLE II. The functions $g_2(\alpha)/(4\pi)^{\frac{1}{2}}$ and $g_3(\alpha)/(4\pi)^{\frac{1}{2}}$ for the body-centered cubic structure.

α	$g_2^{(3)}(\alpha)/(4\pi)^{\frac{1}{2}}$	$g_2^{(6)}(\alpha)/(4\pi)^{\frac{1}{2}}$	$g_3^{(3)}(\alpha)/(4\pi)^{\frac{1}{2}}$	$g_3^{(6)}(\alpha)/(4\pi)^{\frac{1}{2}}$
0.00	0	0	0	0
0.05	0	0	0	0
0.10	0	0	0	0
0.15	0	0	0	0
0.20	0.0001	0.0001	0	0
0.25	0.0003	0.0003	0	0
0.30	0.0007	0.0007	0	0
0.35	0.0012	0.0012	0	0
0.40	0.0020	0.0020	−0.0001	−0.0001
0.45	0.0030	0.0030	−0.0001	−0.0001
0.50	0.0043	0.0043	−0.0002	−0.0002
0.55	0.0059	0.0059	−0.0004	−0.0004
0.60	0.0077	0.0077	−0.0006	−0.0006
0.65	0.0097	0.0097	−0.0009	−0.0009
0.70	0.0119	0.0119	−0.0013	−0.0013
0.75	0.0142	0.0142	−0.0019	−0.0019
0.80	0.0165	0.0165	−0.0026	−0.0026
0.85	0.0187	0.0186	−0.0034	−0.0034
0.90	0.0207	0.0205	−0.0044	−0.0044
0.95	0.0223	0.0222	−0.0056	−0.0056
1.00	0.0236	0.0234	−0.0070	−0.0069
1.05	0.0244	0.0241	−0.0085	−0.0084
1.10	0.0247	0.0242	−0.0101	−0.0100
1.15	0.0243	0.0238	−0.0118	−0.0117
1.20	0.0234	0.0227	−0.0135	−0.0134
1.25	0.0219	0.0211	−0.0152	−0.0151
1.30	0.0199	0.0189	−0.0168	−0.0168
1.35	0.0174	0.0163	−0.0183	−0.0183
1.40	0.0145	0.0132	−0.0196	−0.0196
1.45	0.0113	0.0100	−0.0207	−0.0206
1.50	0.0080	0.0066	−0.0214	−0.0213
1.55	0.0047	0.0032	−0.0217	−0.0217
1.60	0.0014	0.0000	−0.0216	−0.0217
1.65	−0.0016	−0.0030	−0.0211	−0.0212
1.70	−0.0043	−0.0055	−0.0202	−0.0203
1.75	−0.0066	−0.0076	−0.0188	−0.0190
1.80	−0.0084	−0.0091	−0.0170	−0.0172
1.85	−0.0097	−0.0100	−0.0149	−0.0151
1.90	−0.0104	−0.0103	−0.0124	−0.0127
1.95	−0.0106	−0.0100	−0.0097	−0.0101
2.00	−0.0103	−0.0091	−0.0068	−0.0073

We have the following expression:

$$\chi^{\mathrm{bcc}}(\mathbf{r}) = \left(\frac{a}{\pi}\right)^3 2(x^4+y^4+z^4-2x^2y^2-2y^2z^2-2z^2x^2)^{-1} \times \left[x\sin\frac{\pi x}{a}\left(\cos\frac{\pi y}{a}\cos\frac{\pi z}{a}-\cos\frac{\pi x}{a}\right) + y\sin\frac{\pi y}{a}\left(\cos\frac{\pi z}{a}\cos\frac{\pi x}{a}-\cos\frac{\pi y}{a}\right) + z\sin\frac{\pi z}{a}\left(\cos\frac{\pi x}{a}\cos\frac{\pi y}{a}-\cos\frac{\pi z}{a}\right)\right]. \quad (2.1)$$

In order to calculate $g_j(r)$ we have to find the expressions for $\chi(\mathbf{r})$ for nine directions $A, B, \cdots, I$. These directions are specified in Paper I and the quantities $\varphi_A, \cdots, \varphi_I$ occurring in the formulas below are given in (I; 4.2).

$$\chi_A^{\mathrm{bcc}} = \frac{2}{\varphi_A^3}\sin\varphi_A(1-\cos\varphi_A),$$

$$\chi_B^{\mathrm{bcc}} = \frac{1}{2\varphi_B^2}\sin\varphi_B(\varphi_B\cos\varphi_B+\sin\varphi_B),$$

$$\chi_C^{\mathrm{bcc}} = \frac{2}{\varphi_C^3}\sin\varphi_C\cos\varphi_C(1-\cos\varphi_C),$$

$$\chi_D^{\mathrm{bcc}} = \frac{2}{9\varphi_D^3}\sin\varphi_D(5-6\sin^2\varphi_D-5\cos\varphi_D + 8\cos\varphi_D\sin^2\varphi_D),$$

$$\chi_E^{\mathrm{bcc}} = \frac{1}{4\varphi_E^3}\sin^2\varphi_E \times[\varphi_E(3-4\sin^2\varphi_E)+\sin\varphi_E\cos\varphi_E],$$

$$\chi_F^{bcc} = \frac{2}{15\varphi_F^3}\sin\varphi_F(-9 + 28\sin^2\varphi_F - 20\sin^4\varphi_F + 9\cos\varphi_F - 16\sin^2\varphi_F\cos\varphi_F),$$

$$\chi_G^{bcc} = \frac{1}{6\varphi_G^3}\sin^3\varphi_G\cos\varphi_G \times (6 + 21\sin^2\varphi_G + 16\sin^4\varphi_G),$$

$$\chi_H^{bcc} = \frac{1}{12\varphi_H^3}\sin^2\varphi_H \times [\varphi_H(9 - 62\sin^2\varphi_H + 116\sin^4\varphi_H - 64\sin^6\varphi_H) + \sin\varphi_H\cos\varphi_H \times (3 - 10\sin^2\varphi_H + 8\sin^4\varphi_H)],$$

$$\chi_I^{bcc} = \frac{1}{10\varphi_I^3}\sin^3\varphi_I\cos\varphi_I \times (10 - 75\sin^2\varphi_I + 144\sin^4\varphi_I - 80\sin^6\varphi_I). \quad (2.2)$$

From the considerations in I, it follows that the g_j are in fact functions of $\alpha = r/a$ so that they depend only on the type of the structure. $g_0(\alpha)$ is calculated using 3-, 6-, and 9-term formulas [I; (2.2), (2.2′), (2.2″)]. The functions are tabulated in Table I. We can draw the same conclusions as in the cases of simple cubic and face-centered cubic lattices. For $\alpha < 1$ the agreement between $g_0^{(3)}$, $g_0^{(6)}$, and $g_0^{(9)}$ is excellent; for $1 < \alpha \leq 2$ the 3-term function $g_0^{(3)}$ differs distinctly from $g_0^{(6)}$ and $g_0^{(9)}$. Let us consider the differences $(g_0^{(6)} - g_0^{(3)})$ and $(g_0^{(9)} - g_0^{(6)})$: at $\alpha = 1.15$ (that is, about the first minimum of g_0)

$$g_0^{(6)} - g_0^{(3)} = -0.0023, \quad g_0^{(9)} - g_0^{(6)} = 0,$$

at $\alpha = 1.85$ (that is, about the second maximum of g_0)

$$g_0^{(6)} - g_0^{(3)} = -0.0105, \quad g_0^{(9)} - g_0^{(6)} = -0.0006.$$

$g_0^{(9)}$ is of course a better approximation of exact g_0 than $g_0^{(6)}$ but we see from the table that in the considered region $g_0^{(9)}$ and $g_0^{(6)}$ differ only very slightly for $\alpha < 2$ so that $g_0^{(6)}$ is a good approximation there.

$g_2(\alpha)$ and $g_3(\alpha)$ are calculated using only 3- and 6-term formulas [I; (2.3), (2.3′), (2.4), (2.4′)] and they are tabulated in Table II. In this case the agreement between $g_j^{(3)}$ and $g_j^{(6)}$ is very good for $\alpha < 1$.

These functions $g_j(r)$ for $0 \leq r \leq 2a$ are very useful in calculations for alkali metals.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Einstein Equations and Electromagnetism

C. LANCZOS
Dublin Institute for Advanced Studies, Dublin, Ireland
and
Department of Astronomy, Yale University, New Haven, Connecticut

(Received 10 October 1966)

Einstein's equations for infinitesimal gravitational fields are investigated from the standpoint of replacing Einstein's flat Minkowskian background space by a strongly agitated Riemannian lattice space, with the added demand that the perturbation field be of very low frequency compared with the lattice frequency. It is found that the strongly nonlinear lattice transmits low-frequency perturbations only if the perturbation field satisfies a set of field equations which become practically identical with Maxwell's equations. While this result demonstrates the close relation between Einstein equations and electromagnetism, it also opens a new door toward the deeper understanding of the quantum phenomena.

1. INTRODUCTION

IN 1916, Einstein established his celebrated field equations

$$R_{ik} = 0 \quad (1.1)$$

as a description of a pure gravitational field.[1] Since Riemannian geometry is apparently void of antisymmetric elements, he came to the conclusion that the realm of electromagnetic phenomena must be beyond the scope of Riemannian geometry. In his last efforts he settled on a theory which abandoned the symmetry of the metrical tensor g_{ik} and replaced it by a general nonsymmetric field.[2]

On the surface it appears that Einstein actually exhausted all the possibilities of Riemannian geometry in his great paper of 1916. A closer examination reveals, however, that certain apparently natural assumptions may have oversimplified the picture. Quite apart from the fact that Einstein's gravitational equations put the matter tensor equal to zero and thus

[1] A. Einstein, Ann. Physik **49**, 769 (1916).

[2] A. Einstein, *The Meaning of Relativity* (Princeton University Press, Princeton, New Jersey, 1955), 5th ed., p. 133.

$$\chi_F^{bcc} = \frac{2}{15\varphi_F^3}\sin\varphi_F(-9 + 28\sin^2\varphi_F - 20\sin^4\varphi_F + 9\cos\varphi_F - 16\sin^2\varphi_F\cos\varphi_F),$$

$$\chi_G^{bcc} = \frac{1}{6\varphi_G^3}\sin^3\varphi_G\cos\varphi_G \times (6 + 21\sin^2\varphi_G + 16\sin^4\varphi_G),$$

$$\chi_H^{bcc} = \frac{1}{12\varphi_H^3}\sin^2\varphi_H \times [\varphi_H(9 - 62\sin^2\varphi_H + 116\sin^4\varphi_H - 64\sin^6\varphi_H) + \sin\varphi_H\cos\varphi_H \times (3 - 10\sin^2\varphi_H + 8\sin^4\varphi_H)],$$

$$\chi_I^{bcc} = \frac{1}{10\varphi_I^3}\sin^3\varphi_I\cos\varphi_I \times (10 - 75\sin^2\varphi_I + 144\sin^4\varphi_I - 80\sin^6\varphi_I). \quad (2.2)$$

From the considerations in I, it follows that the g_j are in fact functions of $\alpha = r/a$ so that they depend only on the type of the structure. $g_0(\alpha)$ is calculated using 3-, 6-, and 9-term formulas [I; (2.2), (2.2′), (2.2″)]. The functions are tabulated in Table I. We can draw the same conclusions as in the cases of simple cubic and face-centered cubic lattices. For $\alpha < 1$ the agreement between $g_0^{(3)}$, $g_0^{(6)}$, and $g_0^{(9)}$ is excellent; for $1 < \alpha \leq 2$ the 3-term function $g_0^{(3)}$ differs distinctly from $g_0^{(6)}$ and $g_0^{(9)}$. Let us consider the differences $(g_0^{(6)} - g_0^{(3)})$ and $(g_0^{(9)} - g_0^{(6)})$: at $\alpha = 1.15$ (that is, about the first minimum of g_0)

$$g_0^{(6)} - g_0^{(3)} = -0.0023, \quad g_0^{(9)} - g_0^{(6)} = 0,$$

at $\alpha = 1.85$ (that is, about the second maximum of g_0)

$$g_0^{(6)} - g_0^{(3)} = -0.0105, \quad g_0^{(9)} - g_0^{(6)} = -0.0006.$$

$g_0^{(9)}$ is of course a better approximation of exact g_0 than $g_0^{(6)}$ but we see from the table that in the considered region $g_0^{(9)}$ and $g_0^{(6)}$ differ only very slightly for $\alpha < 2$ so that $g_0^{(6)}$ is a good approximation there.

$g_2(\alpha)$ and $g_3(\alpha)$ are calculated using only 3- and 6-term formulas [I; (2.3), (2.3′), (2.4), (2.4′)] and they are tabulated in Table II. In this case the agreement between $g_j^{(3)}$ and $g_j^{(6)}$ is very good for $\alpha < 1$.

These functions $g_j(r)$ for $0 \leq r \leq 2a$ are very useful in calculations for alkali metals.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Einstein Equations and Electromagnetism

C. LANCZOS
Dublin Institute for Advanced Studies, Dublin, Ireland
and
Department of Astronomy, Yale University, New Haven, Connecticut

(Received 10 October 1966)

Einstein's equations for infinitesimal gravitational fields are investigated from the standpoint of replacing Einstein's flat Minkowskian background space by a strongly agitated Riemannian lattice space, with the added demand that the perturbation field be of very low frequency compared with the lattice frequency. It is found that the strongly nonlinear lattice transmits low-frequency perturbations only if the perturbation field satisfies a set of field equations which become practically identical with Maxwell's equations. While this result demonstrates the close relation between Einstein equations and electromagnetism, it also opens a new door toward the deeper understanding of the quantum phenomena.

1. INTRODUCTION

IN 1916, Einstein established his celebrated field equations

$$R_{ik} = 0 \quad (1.1)$$

as a description of a pure gravitational field.[1] Since Riemannian geometry is apparently void of antisymmetric elements, he came to the conclusion that the realm of electromagnetic phenomena must be beyond the scope of Riemannian geometry. In his last efforts he settled on a theory which abandoned the symmetry of the metrical tensor g_{ik} and replaced it by a general nonsymmetric field.[2]

On the surface it appears that Einstein actually exhausted all the possibilities of Riemannian geometry in his great paper of 1916. A closer examination reveals, however, that certain apparently natural assumptions may have oversimplified the picture. Quite apart from the fact that Einstein's gravitational equations put the matter tensor equal to zero and thus

[1] A. Einstein, Ann. Physik **49**, 769 (1916).

[2] A. Einstein, *The Meaning of Relativity* (Princeton University Press, Princeton, New Jersey, 1955), 5th ed., p. 133.

give no clue concerning the structure of elementary particles, he makes a further hypothesis, based on phenomenological evidence, which may not hold on a deeper plane. Arguing on the basis of the apparently universal validity of Lorentz transformations, Einstein assumes that in vacuum, far away from matter, the curved nature of the space-time world is only weakly represented. Accordingly he puts the line element in the form

$$g_{ik} = \delta_{ik} + \gamma_{ik} \tag{1.2}$$

(using Minkowskian coordinates), considering γ_{ik} as a small correction term.[3]

The possibility of a strongly curved submicroscopic world, which appears only macroscopically flat, cannot be discounted. In the last few years the author elucidated various aspects of a metric, which was obtained as a weak perturbation of a strongly agitated background field of fourfold periodicity.[4] The fact that under ordinary circumstances this background field does not come in explicit evidence is explainable by the extraordinary smallness of the fundamental lattice constant (obtained by putting the three fundamental universal constants $\hbar$, c, and $8\pi\gamma$ equal to 1).

The mathematical analysis of such a universe is greatly facilitated by the introduction of the principal axes of the matter tensor as basic frame of reference. This is in harmony with the tetrad formulation of general relativity, advocated by Einstein himself, in his theory of distant parallelism.[5] The four unit vectors of the orthogonal tetrad h_{ik}, together with the four eigenvalues σ_a, yield 20 quantities, which are necessary and sufficient for an algebraic representation of metrical tensor *and* matter tensor, at every point of the manifold, according to the equations[6]

$$g_{ik} = h_{ia}h_{ka}, \tag{1.3}$$

$$T_{ik} = \sigma_a h_{ia}h_{ka}. \tag{1.4}$$

[The replacement of R_{ik} by T_{ik} merely changes the eigenvalues σ_a by a constant; the puzzling feature that in (1.4) the sum index a appears in three factors is removed by the general convention that in the counting of sum indices the eigenvalues σ_a are never included.]

It is clear that the validity of Eqs. (1.1), or even that of the "cosmological equations"

$$R_{ik} = \lambda g_{ik} \tag{1.5}$$

would interfere with the construction of the elementary tetrads by removing the uniqueness of the principal axes. We assume that these equations do *not* hold on the submicroscopic level. However, we assume that the four principal axes, although microscopically well determined, are *macroscopically isotropic* (in analogy to the behavior of crystals of cubic symmetry). By this assumption the otherwise very large number of constants, which characterize the microlattice in macroscopic relations, can be greatly reduced. For example, the cosmological equations (1.5), although locally inapplicable, become the correct *macroscopic* description of the lattice field (in view of the extraordinary smallness of the lattice constant, even atomic and subatomic regions have to be considered as macroscopic. We see later that a slight deviation from complete macroscopic isotropy has to be assumed in order to cope with the world of physical facts).

The basic lattice field as it stands is of no direct physical significance, due to its too great uniformity. The world of physical observations is superimposed on the lattice and our first attention is devoted to the "vacuum", i.e., regions of the space-time world, which are free of physical matter and sufficiently far from the central core of material particles to permit a linear approximation. We then are in the domain of Einstein's equation (1.2), which, however, must now be written in the form

$$\hat{g}_{ik} = g_{ik} + \delta g_{ik} \tag{1.6}$$

to which we have to add (since the matter tensor is not zero in our case)

$$\hat{T}_{ik} = T_{ik} + \delta T_{ik}. \tag{1.7}$$

The variation of Eqs. (1.3) and (1.4) yields

$$\delta g_{ik} = h_{ia}\delta h_{ka} + h_{ka}\delta h_{ia}, \tag{1.8}$$

$$\delta T_{ik} = \sigma_a h_{ia}\delta h_{ka} + \sigma_a h_{ka}\delta h_{ia} + h_{ia}h_{ka}\delta\sigma_a. \tag{1.9}$$

The four infinitesimal vectors δh_{ia} are reducible to an infinitesimal tensor of second order of mixed components φ^i_k by putting [cf., Ref. 6, Eq. (4.2)]

$$\delta h_{ia} = \varphi^\mu_i h_{\mu a}. \tag{1.10}$$

In terms of this tensor Eqs. (1.8) and (1.9) become

$$\delta g_{ik} = \varphi^\mu_i g_{\mu k} + \varphi^\mu_k g_{\mu i}, \tag{1.11}$$

$$\delta T_{ik} = \varphi^\mu_i T_{\mu k} + \varphi^\mu_k T_{\mu i} + h_{ia}h_{ka}\delta\sigma_a. \tag{1.12}$$

Einstein, in his theory of distant parallelism, attempted to enrich Riemannian geometry by new elements. In his theory the 16 h_{ia} are the fundamental quantities which form the basis of geometry. In our considerations the elementary tetrads appear as the result of a principal axis transformation, without modifying the basic tenets of Riemannian geometry. The 20 quantities h_{ia} and σ_a are not independent of

[3] A. Einstein, Sitzber. Preuss. Akad. Wiss. **1916**, 688 (1916).
[4] C. Lanczos, J. Math. Phys. **4**, 951 (1963); Phys. Rev. **134**, B476 (1964).
[5] A. Einstein, Math. Ann. **102**, 685 (1930).
[6] C. Lanczos, J. Math. Phys. **7**, 316 (1966), Eqs. (2.16) and (2.17).

each other, because the T_{ik} are expressible in terms of the g_{ik} and their first and second derivatives. This yields 10 partial differential equations of second order between the 20 quantities h_{ia} and σ_a. The remaining 10 equations are provided by the 10 field equations by which we single out a certain class of geometries from the general group of arbitrary Riemannian geometries. In the end we have 20 equations for 20 quantities.

The same conditions return in relation to the perturbation problem. In consequence of the definition of the matter tensor we can express δT_{ik} in terms of δg_{ik} and its first and second derivatives. But then Eq. (1.12) results in 10 partial differential equations of second order between the 20 quantities φ_k^i and $\delta\sigma_a$. These relations are purely caused by the nature of the principal-axis system, and not by the nature of a specific Riemannian geometry. The specification of a definite geometry by field equations adds 10 more equations, which makes our problem uniquely determined. *In the present investigation we do not raise the question of the field equations. Our aim is reduced to the investigation of those equations, which follow from the nature of the principal-axis system, irrespective of the field equations.*

The only exception we make is that we want to assume the universal validity of the scalar equation

$$R = \text{const}, \tag{1.13}$$

which is an exact consequence of the quadratic action principle.[7] In this case

$$\sigma_1 + \cdots \sigma_4 = \text{const} \tag{1.14}$$

and thus the variation of the σ_a can be submitted to the condition

$$\delta\sigma_1 + \cdots \delta\sigma_4 = 0. \tag{1.15}$$

Now, it so happens that for macroscopic effects only the *sum* of the four $\delta\sigma_a$ can play a role, if the basic lattice satisfies the condition of macroscopic isotropy. Since this sum vanishes, we can from the beginning omit the perturbation of the eigenvalues and write (1.12) in the form

$$\delta T_{ik} = \varphi_i^\mu T_{\mu k} + \varphi_k^\mu T_{\mu i}. \tag{1.16}$$

2. THE PERTURBATION PROBLEM

In Einstein's work the perturbation problem appears in his theory of infinitesimal gravitational fields. By putting the metrical tensor in the form (1.2), he obtained the field equations for γ_{ik} in the form of 10 partial differential equations. These equations are derivable from an action principle of the form

$$\delta Q = \delta \int L\, d^4x = 0 \tag{2.1}$$

with a Lagrangian L which can be written in the following form[8]:

$$L = \begin{bmatrix} i\alpha \\ \beta \end{bmatrix}\begin{bmatrix} i\beta \\ \alpha \end{bmatrix} - \begin{bmatrix} \alpha\beta \\ \alpha \end{bmatrix}\begin{bmatrix} ii \\ \beta \end{bmatrix}. \tag{2.2}$$

This quantity has no invariant significance. We can make it, however, to a genuine invariant by writing it in a somewhat different manner.

Since we have to use the process of variation, the use of the symbol δ for the perturbation field would hardly be appropriate. We agree that we denote a small change of a quantity by a bar, thus putting

$$\hat{g}_{ik} = g_{ik} + \bar{g}_{ik}, \tag{2.3}$$

$$\hat{T}_{ik} = T_{ik} + \bar{T}_{ik}, \tag{2.4}$$

$$\hat{\Gamma}_{ik}^m = \Gamma_{ik}^m + \bar{\Gamma}_{ik}^m. \tag{2.5}$$

Whereas the Γ_{ik}^m do not form a tensor, this is different with regard to $\bar{\Gamma}_{ik}^m$, which is a genuine tensor, transforming like a tensor of third order for arbitrary finite coordinate transformations of the g_{ik}. In Einstein's case

$$g_{ik} = \delta_{ik}, \quad T_{ik} = 0, \quad \Gamma_{ik}^m = 0 \tag{2.6}$$

and we see without difficulty that the invariant formulation of the action principle,

$$\delta Q = \delta \int L g^{\frac{1}{2}}\, d^4x \tag{2.7}$$

with

$$L = (\Gamma_{i\alpha}^\beta \Gamma_{k\beta}^\alpha - \Gamma_{ik}^\beta \Gamma_{\alpha\beta}^\alpha) g^{ik}, \tag{2.8}$$

goes over into (2.1), (2.2), if the special conditions (2.6) are satisfied. In our problem, of course, the g_{ik} are very far from being constants, nor is $T_{ik} = 0$. But in the general form (2.7), (2.8) the action principle remains valid even under these generalized conditions, if we add some terms quadratic in $\bar{g}_{iki}$ (cf., note at the end of this paper). It provides us with the matter tensor of the perturbation field, in the form

$$\delta Q = \int \bar{T}^{ik} \delta \hat{g}_{ik} g^{\frac{1}{2}}\, d^4x. \tag{2.9}$$

In Einstein's theory the variation (2.9) is put equal to zero, which yields the field equations $\bar{T}_{ik} = 0$. Our interest in the same action integral stems from the fact that we want to study the relation

$$\bar{T}_{ik} = \varphi_i^\mu T_{\mu k} + \varphi_k^\mu T_{\mu k}, \tag{2.10}$$

and we have good reasons to put this equation in integrated form. There exists a circumstance in our problem, which has no analogy in Einstein's work. The perturbation field of Einstein is erected on a constant metrical field $g_{ik} = \delta_{ik}$; here a distinction between "slow" and "fast" loses all significance. In our present problem it is of vital importance that the

[7] C. Lanczos, Rev. Mod. Phys. **29**, 337 (1957), Eq. (5.22).

[8] H. Weyl, *Space-Time-Matter* (Methuen and Company Ltd., London, 1922), p. 240.

world of physical observations encounters only frequencies which are exceedingly low compared with the very high frequencies of the lattice field. The basic lattice cell is submicroscopic, even compared with atomic dimensions. Hence, our perturbation problem is of a special type, very similar to a situation encountered in broadcasting, where we have a carrier wave of very high frequency, and an amplitude modulation of acoustical frequencies. The equation

$$\bar{g}_{ik} = \varphi_i^\mu g_{\mu k} + \varphi_k^\mu g_{\mu i}, \tag{2.11}$$

which expresses the perturbation of the highly agitated metrical field g_{ik}, is exactly of this type. The lattice field g_{ik} of very high periodic vibrations represents the carrier wave, whereas the slowly changing (and no longer periodic) φ_k^i field represents the amplitude modulation of the basic field. A perturbation of this type is sometimes called "adiabatic"[9] and is characterized by the condition that the perturbation frequencies are so small in comparison to the frequencies of the base field that the perturbation can be considered as practically unchanged over a domain which extends over a lattice cell.

Under such circumstances we are not interested in the field equations in their local manifestations. We perform an integration over the fundamental lattice cell, thus obtaining differential equations with *constant coefficients*, although these constants are in fact only quasi-constants, expressing macroscopic average values, since the microscopically rough fields of the lattice cannot be detected under the prevailing physical conditions. The situation is fully analogous to the problem encountered by Lorentz[10] in his electron theory. Lorentz realized that our macroscopic measurements do not allow the observation of the highly uneven fields of individual particles, but only the collective action of a huge number of such particles. Hence it is permissible to integrate over volumes which are small in macroscopic relations, but large if compared with the size of individual particles. In this manner he could explain the dielectric constant, magnetic permeability and conductivity of materials in Maxwell's phenomenological equations as statistical averages, caused by the concerted action of a large ensemble of particles. We have in our problem an entirely similar situation, with the only difference that, in view of the periodic nature of the basic lattice, the domain of integration is automatically established as that of the fundamental lattice cell.

Lorentz started with a linear set of equations and encountered no difficulty in his averaging process. But in our problem, considering the highly nonlinear nature of the basic field, we cannot take the possibility of a low-frequency perturbation field as self-evident. Perhaps the circumstances are such that *no* such perturbations exist or, if they exist, they can only occur under very special circumstances? In contradistinction to Einstein we start from a basic field which is highly agitated and thus must be the carrier of very large (although fortunately not convertible) energy. In such a case the differential equations describing the perturbation are such that they give rise to a very large "mass term" in the equation. Einstein's equations are homogeneous of second order and no term can occur in them which is not differentiated twice. But in our case the Lagrangian contains, in addition to terms of the form φ_k^i, $_m$ (the comma refers to partial differentiation with respect to x_m) also terms which are proportional to the φ_k^i themselves. Under the assumption of macroscopic isotropy this part of the Lagrangian is reducible to the following three terms:

$$a_1\varphi_k^i\varphi_i^k + a_2\varphi_k^i\varphi_k^i + a_3\varphi_i^i\varphi_k^k. \tag{2.12}$$

The coefficients a_1, a_2, a_3 are defined as quadratic expressions formed out of the $g_{ik,m}$, whereas the terms of the Einsteinian type have coefficients which are formed with the help of the g_{ik} themselves. The order of magnitude of the ratio of the two types of terms can be estimated to be of the order of unity, if the lattice constant is accepted as unit of length. But then the imposed frequencies are of the order of magnitude of the lattice frequencies and the possibility of low-frequency oscillations is precluded from the beginning.

There exists, however, a singular case, defined by the condition

$$a_1 = a_2. \tag{2.13}$$

If this condition is satisfied, then the first two terms combine into

$$a_1(\varphi_k^i + \varphi_i^k)\varphi_i^k = \tfrac{1}{2}a_1(\varphi_k^i + \varphi_i^k)^2. \tag{2.14}$$

Now, the mass term appears only in connection with the *symmetric* combination $\varphi_k^i + \varphi_i^k$, whereas the antisymmetric combination $\varphi_k^i - \varphi_i^k$ drops out. Under these circumstances the low-frequency waves containing $\varphi_k^i + \varphi_i^k$ are blocked, whereas the low-frequency waves containing $\varphi_k^i - \varphi_i^k$ are admitted by the lattice field. Our aim in the following section is to study the nature of these waves.

3. THE ANTISYMMETRIC DIFFERENTIAL OPERATOR

We return to our Lagrangian (2.8) which we wish to study more closely. A simplification occurs now,

[9] S. Tomonaga, *Quantum Mechanics* (North-Holland Publishing Company, Amsterdam, 1962), Vol. I, p. 294.

[10] H. A. Lorentz, Proc. Acad. Amsterdam **5**, 254 (1902); *Collected Papers* (M. Nijhoff, The Hague, 1936), Vol. III, 117.

because in the second term we encounter the vector

$$\Gamma^{\alpha}_{\alpha\beta} = \tfrac{1}{2}(\bar{g}_{ik}g^{ik})_{;\alpha} = \tfrac{1}{2}[(\varphi^{\mu}_{i}g_{\mu k} + \varphi^{\mu}_{k}g_{\mu i})g^{ik}]_{;\alpha} = (\varphi^{\mu}_{\mu})_{;\alpha} \quad (3.1)$$

(the semicolon refers to covariant differentiation). If we can assume in advance that φ^{i}_{k} is antisymmetric, this term drops out. The remaining expression can be written as the sum of two contributions,

$$L_1 = g^{ik}(\varphi^{\mu}_{\nu;i}\varphi^{\nu}_{k;\mu} + \tfrac{1}{2}\varphi^{\mu}_{i;\nu}\varphi^{\nu}_{k;\mu} - \tfrac{1}{2}\varphi^{\mu}_{\nu;i}\varphi^{\alpha}_{\mu;k}) \quad (3.2)$$

and

$$L_2 = \tfrac{1}{2}g^{\alpha\sigma}g^{\beta\rho}g_{\mu\nu}\varphi^{\mu}_{\alpha;\beta}(\varphi^{\nu}_{\rho;\sigma} - \varphi^{\nu}_{\sigma;\rho}). \quad (3.3)$$

In view of the fact that we keep only the antisymmetric part of φ^{i}_{k}, we put

$$\varphi_{ik} = \tfrac{1}{2}(\varphi^{i}_{k} - \varphi^{k}_{i}) = -\varphi_{ki}. \quad (3.4)$$

Moreover, our first task is to obtain the principal part of the differential operator, pertaining to the second partial derivatives; this means that we replace covariant differentiation by ordinary differentiation. According to our general program, we consider the φ_{ik} as slowly changing quantities in view of the smallness of the lattice cells; hence we integrate the rapidly changing factors over the fundamental cell, before carrying out the minimization process. We indicate by a brace $\{\ \}$ the operation of taking the average value of a quantity (after multiplying by $g^{\frac{1}{2}}$) over the lattice cell. We write, for example (denoting the domain of the cell by C),

$$\{g^{ik}\} = \int_C g^{ik}g^{\frac{1}{2}}\,d^4x \Big/ \int_C d^4x. \quad (3.5)$$

In view of the postulated macroscopic isotropy of all four axes, this quantity must be reducible to a single scalar α_0:

$$\{g^{ik}\} = \alpha_0\delta_{ik}. \quad (3.6)$$

The contribution due to L_1 now becomes

$$\int \{L_1\}\,d^4x = \tfrac{1}{2}\alpha_0\int [2\varphi_{\mu\nu,i}\varphi_{\nu i,\mu} + \varphi_{\mu i,\nu}\varphi_{\nu i,\mu} - \varphi_{\mu\nu,i}\varphi_{\nu\mu,i}]\,d^4x. \quad (3.7)$$

Since $\varphi_{a,b}\varphi_{c,d}$ is variationally equivalent to $\varphi_{a,d}\varphi_{c,b}$, we can write the integrand of (3.7) in the form

$$\varphi_{\mu\nu,i}\varphi_{\mu\nu,i} + \varphi_{\mu\nu,\mu}\varphi_{\nu i,i}, \quad (3.8)$$

the variation of which yields

$$-\tfrac{1}{2}\alpha_0\int (2\varphi_{ik,\alpha\alpha} - \varphi_{i\alpha,\alpha k} + \varphi_{k\alpha,\alpha i})\delta\varphi_{ik}\,d^4x. \quad (3.9)$$

We now come to the variation due to L_2. For the evaluation of this quantity we need the following set of constants:

$$\alpha_{ijkmnp} = \{g_{ij}g^{km}g^{np}\}. \quad (3.10)$$

There are altogether 550 coefficients which, however, become reducible to only 8 coefficients by the condition of macroscopic isotropy. They are listed below, in which i, j, k denote 3 arbitrary but different indices and the sum convention is put out of action, each index referring to an individual value between 1 and 4:

$$\begin{aligned} \alpha_1 &= \{g_{ii}g^{ii}g^{ii}\}, & \beta_1 &= \{g_{ij}g^{ij}g^{ii}\}, \\ \alpha_2 &= \{g_{ii}g^{ii}g^{jj}\}, & \beta_2 &= \{g_{ij}g^{ij}g^{kk}\}, \\ \alpha_3 &= \{g_{ii}g^{jj}g^{jj}\}, & \beta_3 &= \{g_{ii}g^{ij}g^{ij}\}, \\ \alpha_4 &= \{g_{ii}g^{jj}g^{kk}\}, & \beta_4 &= \{g_{kk}g^{ij}g^{ij}\}. \end{aligned} \quad (3.11)$$

All the other coefficients not included above must vanish.

The previously introduced constant α_0, defined by (3.6), is expressible in terms of these constants, due to the algebraic identity (summed over j):

$$g_{ij}g^{kj} = \delta_{ik}. \quad (3.12)$$

Hence,

$$\alpha_0 = \alpha_1 + 3\beta_1. \quad (3.13)$$

But we obtain also

$$\alpha_0 = \alpha_2 + \beta_1 + 2\beta_2, \quad (3.14)$$

which yields the relation

$$\alpha_1 + 2\beta_1 = \alpha_2 + 2\beta_2, \quad (3.15)$$

thus reducing the number of available constants to 7.

The calculation shows the emergence of the following fundamental operator, to be denoted by $A_{ik}(\varphi)$:

$$A_{ik}(\varphi) = \varphi_{ik,\alpha\alpha} - \varphi_{i\alpha,k\alpha} + \varphi_{k\alpha,i\alpha}. \quad (3.16)$$

This operator, antisymmetric in i, k, has a number of interesting properties:

(1) Its divergence is identically zero,

$$A_{ik,k}(\varphi) = 0. \quad (3.17)$$

(2) It remains invariant with respect to the following substitution,

$$\varphi_{ik} = \varphi'_{ik} + \Phi_{i,k} - \Phi_{k,i}, \quad (3.18)$$

where Φ_i is an arbitrary vector.

On account of these properties, the operator $A_{ik}(\varphi)$ is a *perfect antisymmetric counterpart of the symmetric operator $B_{ik}(\gamma)$ which appears in Einstein's perturbation problem*[11] (erected on the flat field $g_{ik} = \delta_{ik}$):

$$B_{ik}(\gamma) = \gamma_{\alpha\alpha,ik} + \gamma_{ik,\alpha\alpha} - \gamma_{i\alpha,k\alpha} - \gamma_{k\alpha,i\alpha} - (\gamma_{\alpha\alpha,\beta\beta} - \gamma_{\alpha\beta,\alpha\beta})\delta_{ik}. \quad (3.19)$$

This operator has likewise vanishing divergence. Moreover, it remains invariant with respect to the substitution

$$\gamma_{ik} = \gamma'_{ik} + \Phi_{i,k} + \Phi_{k,i} \quad (3.20)$$

(which here means the freedom of performing arbitrary infinitesimal coordinate transformations, due to the principle of covariance).

[11] W. Pauli, *Theory of Relativity* (Pergamon Press, Inc., New York, 1958), p. 172.

(3) The sum index α in (3.16) assumes the four values of 1 to 4, but we notice that the values $\alpha = i$ and $\alpha = k$ give vanishing contributions. Hence it suffices that α assumes only *two* values, omitting the values i and k.

We can now evaluate the contributions of the 8 coefficients enumerated in (3.11), concerning the integral over $\{L_2\}$. Since all indices are on an equal footing, we can choose without loss of generality some special values for i and k, e.g., $i = 1$, $k = 2$ (together with $i = 2$, $k = 1$), which simplifies the calculations. Our first observation is that the contributions of α_1, α_3, and β_1 cancel out, due to conditions of symmetry. Hence only 5 coefficients contribute to the final result. Furthermore, the coefficients α_2 and β_3, and likewise α_4 and β_4, contribute terms of the same kind and can be combined. This cuts down the number of contributions to 3. The final result appears in the following form:

$$\delta\int\{L_2\}\,d^4x = -\tfrac{1}{2}\int\delta\varphi_{12}\{(\beta_3-\alpha_2)(\varphi_{12,11}+\varphi_{12,22}) + (\beta_4-\alpha_4)(A_{12}+\varphi_{12,33}+\varphi_{12,44}) + \beta_2[4A_{12}-2(\varphi_{12,33}+\varphi_{12,44})]\}\,d^4x. \tag{3.21}$$

The factor of the first term can be changed (on account of the four-dimensionality of the physical world) as follows:

$$\varphi_{12,11}+\varphi_{12,22} = \varphi_{12,\alpha\alpha} - (\varphi_{12,33}+\varphi_{12,44}) \tag{3.22}$$

and now the demand of macroscopic isotropy necessitates that the factor of $\varphi_{12,33}+\varphi_{12,44}$ vanishes. This yields the following added relation between the coefficients α_i and β_i:

$$\alpha_2+\beta_4 = \alpha_4+2\beta_2+\beta_3. \tag{3.23}$$

The resulting δQ becomes under these conditions

$$\delta Q = -\tfrac{1}{2}\int\delta\varphi_{ik}[(\beta_1+8\beta_2+\beta_3)A_{ik}(\varphi) + (\beta_1+2\beta_2+\beta_3)\varphi_{ik,\alpha\alpha}]\,d^4x. \tag{3.24}$$

If we want the entire operator reducible to $A_{ik}(\varphi)$, the factor of the second term has to vanish. This yields one more condition to be satisfied by the β_i:

$$\beta_1+2\beta_2+\beta_3 = 0. \tag{3.25}$$

The 8 coefficients of (3.11) are thus reduced to 5.

Up to now, we have only considered that part of the variational integral which contained the second derivatives of φ_{ik}. In principle, the first derivatives could also be present. They would have to be derived from a term in the Lagrangian which has the form

$$\alpha_{ijkmn}\varphi_{ij}\varphi_{km,n}.$$

Since, however, there is no tensor of fifth order, which could satisfy the principle of macroscopic isotropy, we must have $\alpha_{ijkmn} = 0$. Hence the first derivatives do not appear in the resulting field equations.

Thus far we have obtained the following result concerning the perturbation integral (2.9):

$$\delta Q = 2\int\{\bar{T}_{\mu i}g^{\mu k}\}\delta\varphi_{ik}\,d^4x = -3\beta_2\int A_{ik}(\varphi)\delta\varphi_{ik}\,d^4x. \tag{3.26}$$

On the other hand, relation (2.10) demands

$$\{\bar{T}_{\mu i}g^{\mu k}\}\delta\varphi_{ik} = [\{T_{i\nu}g^{\mu k}\}\varphi_{\nu\mu}+\{T_{\nu\mu}g^{\mu k}\}\varphi_{\nu i}]\delta\varphi_{ik}. \tag{3.27}$$

If on the right side we integrate over the lattice cell, making use of the principle of macroscopic isotropy, we observe that the only contribution could be a "mass term" in the equation, but we agreed in advance that the antisymmetric equation be free of such a term [cf., Eq. (2.13)].

Under these circumstances we would arrive at the homogeneous equation $A_{ik}(\varphi) = 0$ as the expression of the desired field equation, were it not for the fact that, for reasons beyond our understanding at the present time, we must assume that the principle of macroscopic isotropy cannot be exactly satisfied. A slight remaining anisotropy causes the emergence of a small right side of our equation, in which only *three* axes play an equivalent role, while the fourth axis behaves differently. Let us assume that we have for any two individual indices i and j (equal or unequal):

$$\{T_{ii}g^{jj}\} = \lambda\delta_i^j+\alpha\rho_i^j \tag{3.28}$$

(no summation over i or j), where α is a very small constant, whereas

$$\rho_i^j = 0 \quad (i\neq j), \qquad \rho_1^1 = \rho_2^2 = \rho_3^3 = 1, \quad \rho_4^4 = -3. \tag{3.29}$$

The last term in (3.28) represents the deviation of the metrical lattice from macroscopic isotropy. It is equivalent to a correction term in the Lagrangian, which may be written in the form

$$L_3 = \tfrac{1}{2}\alpha\rho_\mu^\mu(\varphi_{\mu i})^2 = \alpha(\varphi_{12}^2+\varphi_{13}^2+\varphi_{23}^2-\varphi_{14}^2-\varphi_{24}^2-\varphi_{34}^2), \tag{3.30}$$

for which we can also put

$$L_3 = \tfrac{1}{2}\alpha\eta_i\eta_j(\varphi_{ij})^2, \tag{3.31}$$

where we define

$$\eta_i = (1,1,1,-1) \tag{3.32}$$

and agree that, in counting sum indices, the η_i are not to be included.

4. THE EQUATIONS FOR THE PERTURBATION FIELD

The final field equations of our perturbation problem appear in the following form

$$\varphi_{ik,\alpha\alpha} - \varphi_{i\alpha,k\alpha} + \varphi_{k\alpha,i\alpha} = \epsilon\eta_i\eta_k\varphi_{ik}, \qquad (4.1)$$

where ϵ is a small (at present unspecified) constant.

From these equations we draw the following conclusions. First of all, the identity (3.17), established earlier for the left side, yields for the antisymmetric tensor φ_{ik} the condition

$$\eta_k\varphi_{ik,k} = 0. \qquad (4.2)$$

This corresponds to the first set of Maxwell's equations, if φ_{ik} is interpreted as the electromagnetic field strength. Furthermore, taking advantage of the second property (3.18) of the operator $A_{ik}(\varphi)$, we make use of the following transformation:

$$\varphi_{ik} = \Phi_{i,k} - \Phi_{k,i} + \epsilon u_{ik} \qquad (4.3)$$

determining the vector Φ_i in such manner that we have

$$u_{ik,k} = 0. \qquad (4.4)$$

For this purpose the equation

$$\Phi_{i,\alpha\alpha} - \Phi_{\alpha,\alpha i} = \varphi_{i\alpha,\alpha} \qquad (4.5)$$

has to be solved. This is always possible, since the right side satisfies the identity

$$\varphi_{i\alpha,\alpha i} = 0, \qquad (4.6)$$

and thus the solution of the equation

$$\Phi_{i,\alpha\alpha} = \varphi_{i\alpha,\alpha} \qquad (4.7)$$

automatically entails $\Phi_{\alpha,\alpha} = 0$, which makes the second term on the left side of (4.5) vanish. (This move is again in full analogy to the method of solving the field equations of infinitesimal fields in Einstein's theory,[12] where, however, this transformation has merely the purpose of normalizing the reference system, but the vector Φ_i drops out from the final result.) In our case the substitution of (4.3) in the field equations (4.1) yields

$$u_{ik,\alpha\alpha} = \eta_i\eta_k(\Phi_{i,k} - \Phi_{k,i} + \epsilon u_{ik}). \qquad (4.8)$$

In view of the smallness of ϵ, the last term on the right side is negligible and we obtain for u_{ik} the inhomogeneous potential equation

$$u_{ik,\alpha\alpha} = \eta_i\eta_k(\Phi_{i,k} - \Phi_{k,i}), \qquad (4.9)$$

where the right side has to satisfy the condition

$$\eta_k(\Phi_{i,kk} - \Phi_{k,ik}) = 0. \qquad (4.10)$$

If we submit Φ_k to the condition

$$\eta_k\Phi_{k,k} = 0 \qquad (4.11)$$

we obtain for Φ_k the wave equation

$$\eta_\alpha\Phi_{k,\alpha\alpha} = 0. \qquad (4.12)$$

This is the customary wave equation for the vector potential Φ_k, whereas (4.11) represents the customary "Lorentz condition".[13] Furthermore, the representation of φ_{ik} in terms of the vector potential according to (4.3)—if the last correction term is omitted—is equivalent to the second set of Maxwellian equations.

It is of interest to observe that, on the basis of the purely elliptic (++++) differential operator (4.1), it is nevertheless possible to obtain the (hyperbolic) wave equation for Φ_i.

The field equations (4.1) possess a Lagrangian, from which they are derivable by variation:

$$L = -\tfrac{1}{2}(\varphi_{ik,\alpha} - \varphi_{i\alpha,k} + \varphi_{k\alpha,i})\varphi_{ik,\alpha} - \tfrac{1}{2}\epsilon\eta_i\eta_k(\varphi_{ik})^2. \qquad (4.13)$$

If we now make the substitution (4.3), considering Φ_i and u_{ik} [submitted to the divergence condition (4.4)] as our new field variables, we obtain L in the following form:

$$L = -\tfrac{1}{2}\epsilon\eta_i\eta_k(\Phi_{i,k} - \Phi_{k,i})^2 - \tfrac{1}{2}\epsilon^2[(u_{ik,\alpha})^2 + 2\eta_i\eta_k(\Phi_{i,k} - \Phi_{k,i})u_{ik}] - \tfrac{1}{2}\epsilon^3\eta_i\eta_k(u_{ik})^2. \qquad (4.14)$$

Considering the smallness of ϵ, which permits us to omit the second and third rows, we recognize in this L the customary Lagrangian of the electromagnetic field.

We thus arrive at the result that, under proper circumstances, the highly agitated metrical lattice acts as a *low pass filter for electromagnetic waves*. The astonishing feature of the derivation is that this result is obtained entirely on the basis of the properties which characterize the principal axes of the matter tensor, independently of any special metrical field equations.

5. CONCLUSIONS

We have investigated the perturbation problem of a Riemannian lattice field of fourfold periodicity, under the condition that the perturbation field is of very low frequency compared with the frequencies of the lattice field. We came to the conclusion that the strongly agitated lattice transmits a low-frequency perturbation only, if that perturbation satisfies a set of partial differential equations. The (macroscopically constant) coefficients of these differential equations are determined by the structure of the basic lattice field. This structure is to a large extent unknown at

[12] J. L. Synge, *Relativity, the General Theory* (North-Holland Publishing Company, Amsterdam, 1960), p. 202.

[13] J. A. Stratton, *Electromagnetic Theory* (McGraw-Hill Book Company, Inc., New York, 1941), p. 24.

the present time and we must be satisfied with conjectures concerning the properties of the basic field, by utilizing the macroscopically observed properties of the perturbation field.

In particular, we obtain the Maxwellian equations as determining equations of the perturbation field, if the basic lattice satisfies—in addition to the principle of macroscopic isotropy—the following three added conditions, discussed more in detail in the text:

$$a_1 = a_2, \tag{2.13}$$

$$\beta_1 + 2\beta_2 + \beta_3 = 0, \tag{3.25}$$

$$\{T_{ii}g^{jj}\} = \lambda\delta_i^j + \alpha\rho_i^j. \tag{3.28}$$

Under these conditions (the inner significance of which escapes us at present), the low-frequency perturbation field admitted by the lattice can be interpreted as the customary electromagnetic field.

The most decisive feature of the present investigation is the circumstance that our result is not based on any specific action principle. Although the perturbation equations possess a (macroscopic) Lagrangian and can be obtained by minimizing an action integral, their true significance lies in a different direction. They express a mathematical property of the fundamental tetrads, which comes in physical appearance in consequence of the slowness of the superposition field. Originally we expected that the coordinate conditions yield 10 equations for 20 quantities, entailing no restriction of the perturbation field. The situation is radically altered, however, by the demand that the lattice transmits a *low-frequency* field. The Maxwell equations express the necessary and sufficient conditions for the realizability of this demand.

This result opens a new perspective toward a deeper understanding of the quantum phenomena, which from the time of Bohr's atomic model seemed to represent a form of over-determination mysteriously superimposed on the classical equations. Apart from the *empirical* necessity of the quantum conditions, no rational principle was ever forwarded to explain in a natural way the inner necessity of such conditions. The present theory is entirely based on rational principles and tries to establish the physical world on the exclusive basis of a genuinely Riemannian world. That a construction, analogous to that of Einstein, is able to arrive at the Maxwellian equations (although based on a positive definite line element) is in itself interesting and demonstrates the inherent power of a Riemannian world. But beyond that the derivation of these equations shows that their validity is *not* a consequence of an action principle (which provides the field equations for the determination of the g_{ik}), but merely a consequence of the nature of the principal axes of the matter tensor.

But then the eigenfunctions of the perturbation field are already determined and the only freedoms left are the *amplitudes* of these eigenfunctions. Hence the selection principle contained in the action principle must of necessity lead to a *quantization of the amplitudes*. The field quantization is thus deeply imbedded in the basic tenets of the theory and an inevitable consequence of the strongly nonlinear nature of the basic metrical lattice.

Note added in proof: For the conclusions of the present paper the Lagrangian (2.8) is of decisive importance. But this Lagrangian yields the perturbation of the matter tensor only if the basic field is free of matter (or at least for $T_{ik} + \lambda g_{ik} = 0$). This condition is not satisfied in our problem. For the general case the Einsteinian invariant (2.8), to be denoted by L_1, has to be augmented by the following L_2 (with the notation $\bar{g}_{ik} = \gamma_{ik}$ and putting $\gamma_{ik}g^{ik} = \gamma$):

$$L_2 = -(R^{ik} - \tfrac{1}{4}Rg^{ik})(\gamma_{i\mu}\gamma_k^\mu - \tfrac{1}{2}\gamma\gamma_{ik}). \tag{5.1}$$

The variation of $L_1 + L_2$ is related to $\bar{T}^{ik}$ as follows:

$$\delta\int(L_1 + L_2)\sqrt{g}\,d^4x = \int(\bar{T}^{ik} + \tfrac{1}{2}T^{ik}\gamma)\delta g_{ik}\sqrt{g}\,d^4x. \tag{5.2}$$

Equation (2.10), which is equivalent to

$$\bar{T}^{ik} + \varphi_\mu^i T^{\mu k} + \varphi_\mu^k T^{\mu i} = 0, \tag{5.3}$$

possesses a Lagrangian in the macroscopic sense, if the macroscopic isotropy condition

$$\{R_i^k\} = \lambda\delta_i^k \tag{5.4}$$

is satisfied. Equation (5.3) is now derivable from the action principle

$$\delta\int L\sqrt{g}\,d^4x = 0, \tag{5.5}$$

where

$$L = (\Gamma_{i\alpha}^\beta\Gamma_{k\beta}^\alpha - \Gamma_{ik}^\alpha\Gamma_{\alpha\beta}^\beta)g^{ik} - \lambda(\varphi_\beta^\alpha\varphi_\alpha^\beta - \varphi_\alpha^\alpha\varphi_\beta^\beta) - R_{ik}g^{\alpha\beta}\varphi_\alpha^i\varphi_\beta^k. \tag{5.6}$$

ACKNOWLEDGMENT

This paper was partially sponsored by the Air Force Office of Scientific Research, Office of Aerospace Research, U.S. Air Force, under AF OSR Grant No. 397–67.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Reduction of Reducible Representations of the Infinitesimal Generators of the Proper, Orthochronous, Inhomogeneous Lorentz Group

J. S. LOMONT
University of Arizona, Tucson, Arizona

H. E. MOSES*
Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts

(Received 2 September 1966)

In the present paper we show explicitly how to reduce reducible representations of the infinitesimal generators of the proper, orthochronous, inhomogeneous Lorentz group. We first construct a basis in which all the reducible representations are expressed as integrals of representations over masses and in which the infinitesimal generators act as in the Foldy–Shirokov realization for nonzero masses and in the Lomont–Moses realization for zero masses. However, the spin operators which appear in the Foldy–Shirokov realizations and the generators of the Euclidean group which appear in the Lomont–Moses realizations are reducible in general. Thus reducible representations are only partially reduced in this basis. On carrying out the reduction of the spin operators and the generators of the Euclidean group, we introduce a second basis such that the reducible representations are completely reduced. By changing the emphasis slightly, the methods of the present paper can be used to obtain the irreducible representations of the generators. One of us (H. E. M.) has already used the methods of the present paper to reduce the electromagnetic vector potential and, in papers which follow the present one, will show how to reduce wavefunctions in general and will also derive the Clebsch–Gordan expansion for the direct product of two massless representations of finite spin and the same sign of energy. From the work of Mautner and Mackey it is known that every reducible unitary ray representation of the proper, orthochronous, inhomogeneous Lorentz group can be reduced to a direct integral of the irreducible representations. The techniques of the present paper thus enable us to carry out the reduction explicitly.

1. INTRODUCTION AND SUMMARY

THE irreducible unitary ray representations of the proper, orthochronous, inhomogeneous Lorentz group were first found and classified by Wigner.[1] One of the immediate problems which suggests itself is whether the reduction of unitary ray representations of the group can be carried out explicitly. It is shown that indeed such a reduction can be made (see the discussion in Ref. 2 based on the work in Refs. 3 and 4).

In the present paper we give a recipe which enables us, in principle at least, to reduce any reducible representation. Actually we solve a more general problem; namely, we show how to reduce the reducible representations of the infinitesimal generators of the group. The representation of the infinitesimal generators of reducible unitary ray representations are, of course, reducible representations of the generators. But not all reducible representations of the generators can be integrated to reducible ray representations of the group. Hence the problem which we solve is more general than the reduction of unitary ray representations.

One of us (H. E. M.) has already used this technique for reducing the electromagnetic vector potential[5] and in the following papers will use it to reduce relativistic wavefunctions generally and also the direct product of two massless representations of the same sign of energy and finite but arbitrary spin as illustrations of the power of the suggested procedure.

A slight shift of emphasis leads to a derivation of the irreducible representations of the infinitesimal generators of the group. For the nonzero-mass cases we are led to the Foldy–Shirokov realizations,[6,7] while for the zero-mass cases we obtain the Lomont–Moses realizations.[8]

We now summarize the procedure for reducing the irreducible representations. The 10 infinitesimal generators of the proper, orthochronous, inhomogeneous Lorentz group are the energy H, the three

* Operated with support from the U.S. Advanced Research Projects Agency.

[1] E. P. Wigner, Ann. Math. **40**, 149 (1939).
[2] J. S. Lomont, J. Math. Phys. **1**, 237 (1960).
[3] F. I. Mautner, Ann. Math. **51**, 1 (1950); **52**, 8 (1950).
[4] G. W. Mackey, "Theory of Group Representations," mimeographed notes, Department of Mathematics, University of Chicago (1955).
[5] H. E. Moses, Nuovo Cimento **42**, 757 (1966).
[6] L. L. Foldy, Phys. Rev. **102**, 568 (1956).
[7] Yu. M. Shirokov, Zh. Eksperim. i Theor. Fiz. **33**, 1196 (1957) [English transl.: Soviet Phys.—JETP **13**, 240 (1961)].
[8] J. S. Lomont and H. E. Moses, J. Math. Phys. **3**, 405 (1962).

components of the angular momentum P_i $(i = 1, 2, 3)$, the three components of the angular momentum J_i, and the three generators corresponding to the space–time rotations $\mathcal{J}_i$. They satisfy the following well-known commutation relations:

$$[H, P_i] = 0, \quad [H, J_i] = 0, \quad [P_i, P_j] = 0,$$

$$[J_i, J_j] = i\sum_k \epsilon_{ijk} J_k, \quad [J_i, P_j] = i\sum_k \epsilon_{ijk} P_k,$$

$$[J_i, \mathcal{J}_k] = i\sum_k \epsilon_{ijk}\mathcal{J}_k, \quad [\mathcal{J}_i, \mathcal{J}_j] = -i\sum_k \epsilon_{ijk} J_k,$$

$$[\mathcal{J}_i, H] = iP_i, \quad [\mathcal{J}_i, P_j] = i\delta_{ij} H. \tag{1.1}$$

In (1.1) ϵ_{ijk} is the usual antisymmetrical three-index symbol.

When the generators are given, they are given in some basis, this basis being characterized by the space of functions in Hilbert space upon which the generators operate. Accordingly, each abstract vector $|\psi\rangle$ is represented by a function $f(\zeta)$ where the variable ζ collectively denotes all the variables, continuous or discrete, upon which the functions in the given representation depend. Then if A is one of the generators, $A\,|\psi\rangle$ is represented by $(Af)(\zeta)$.

We show that there exists a basis such that each vector $|\psi\rangle$ is also represented by a function $F(\mu, \epsilon, \mathbf{p}, \lambda)$, where the vector $\mathbf{p}$ has components p_i $(i = 1, 2, 3)$ each of which takes on every value from $-\infty$ to $+\infty$. The variable μ takes on all the values of the eigenvalues of the mass operator M, where

$$M = [H^2 - P^2]^{\frac{1}{2}} \tag{1.2}$$

with

$$p^2 = \sum_i p_i^2. \tag{1.3}$$

We assume that we are dealing with physically interesting reducible representations so that μ is in the range $0 \leq \mu < \infty$. We also assume that there is no subspace in which all the operators P_i are zero. Since the operator M is known in the ζ basis, the spectrum of M can be found, at least in principle, and has a continuous part and a discrete part in general. Only when $\mu = 0$ is a point eigenvalue need the zero-mass case be considered, for otherwise, when $\mu = 0$ is in the continuous spectrum, the contribution of such components is of zero measure and can thus be ignored.

The variable ϵ takes on the values which occur in the spectrum of the operator sgn H. Since this operator is known in the ζ representation, the eigenvalues which occur can be found. They are, of course, restricted to the values ± 1.

The variable λ is a degeneracy variable which is characterized shortly. Its character and range may depend on μ and ϵ but do not depend on $\mathbf{p}$.

The inner product of two vectors $|\psi\rangle$ and $|\psi^1\rangle$ can be chosen to be

$$\langle \psi^1 \mid \psi\rangle = \sum_\epsilon \int d\mu \int \frac{d\mathbf{p}}{\omega(\mu, p)} \sum_\lambda F^{1*}(\mu, \epsilon, \mathbf{p}, \lambda) F(\mu, \epsilon, \mathbf{p}, \lambda), \tag{1.4}$$

where

$$\omega(\mu, p) = [\mu^2 + p^2]^{\frac{1}{2}}, \tag{1.5}$$

$$p = |\mathbf{p}|. \tag{1.6}$$

In (1.4) the integral over μ is to be interpreted as integration over the continuous spectrum of M and summation over the discrete spectrum. In (1.4) we have also assumed in our notation that λ takes on only discrete values. However, when λ takes on continuous values or represents collectively several variables, the summation over λ is to be replaced by integrals or sums in an appropriate manner.

In terms of the functions $F(\mu, \epsilon, \mathbf{p}, \lambda)$ the state $A\,|\psi\rangle$ is represented by $(AF)(\mu, \epsilon, \mathbf{p}, \lambda)$. In particular, the infinitesimal generators of the inhomogeneous Lorentz group takes the form for $\mu > 0$:

$$(P_i F)(\mu, \epsilon, \mathbf{p}, \lambda) = p_i F(\mu, \epsilon, \mathbf{p}, \lambda),$$

$$(HF)(\mu, \epsilon, \mathbf{p}, \lambda) = \epsilon\omega(\mu, p) F(\mu, \epsilon, \mathbf{p}, \lambda),$$

$$(J_i F)(\mu, \epsilon, \mathbf{p}, \lambda) = -i(\mathbf{p} \times \nabla)_i F(\mu, \epsilon, \mathbf{p}) + (S_i F)(\mu, \epsilon, \mathbf{p}, \lambda),$$

$$(\mathcal{J}_i F)(\mu, \epsilon, p, \lambda) = \epsilon\left[i\omega(\mu, p)\frac{\partial}{\partial p_i} F(\mu, \epsilon, \mathbf{p}, \lambda) + \sum_{j,k} \frac{\epsilon_{ijk} p_j}{\omega(\mu, p) + \mu} (S_k F)(\mu, \epsilon, \mathbf{p}, \lambda)\right]. \tag{1.7}$$

In (1.7) we use $\nabla_i = (\partial/\partial p_i)$. The operators S_i $(i = 1, 2, 3)$ are spin operators which operate on the λ variable only. It is shown that matrices or kernels $S_i(\lambda \mid \lambda')$ exist such that

$$(S_i F)(\mu, \epsilon, \mathbf{p}, \lambda) = \sum_{\lambda'} S_i(\lambda \mid \lambda') F(\mu, \epsilon, \mathbf{p}, \lambda'). \tag{1.8}$$

When λ is a continuous variable the kernels $S_i(\lambda \mid \lambda')$ may be symbolic functions. The kernels or matrices, which in general depend on μ and ϵ but *never* on $\mathbf{p}$ satisfy the commutation rules for spin matrices

$$\sum_{\lambda''} S_i(\lambda \mid \lambda'') S_j(\lambda'' \mid \lambda') = i\sum_k \epsilon_{ijk} S_k(\lambda \mid \lambda'). \tag{1.8a}$$

We show later how the matrices $S_i(\lambda \mid \lambda')$ are found.

For $\mu = 0$, we have the following results:

$$
\begin{aligned}
(P_iF)(0, \epsilon, \mathbf{p}, \lambda) &= p_iF(0, \epsilon, \mathbf{p}, \lambda),\\
(HF)(0, \epsilon, \mathbf{p}, \lambda) &= \epsilon pF(0, \epsilon, \mathbf{p}, \lambda),\\
(J_1F)(0, \epsilon, \mathbf{p}, \lambda) &= \left[-i(\mathbf{p} \times \nabla)_1F(0, \epsilon, \mathbf{p}, \lambda) + \frac{p_1}{p + p_3}(MF)(0, \epsilon, \mathbf{p}, \lambda)\right],\\
(J_2F)(0, \epsilon, \mathbf{p}, \lambda) &= \left[-i(\mathbf{p} \times \nabla)_2F(0, \epsilon, \mathbf{p}, \lambda) + \frac{p_2}{p + p_3}(MF)(0, \epsilon, \mathbf{p}, \lambda)\right],\\
(J_3F)(0, \epsilon, \mathbf{p}, \lambda) &= [-i(\mathbf{p} \times \nabla)_3F(0, \epsilon, \mathbf{p}, \lambda) + (MF)(0, \epsilon, \mathbf{p}, \lambda)],\\
(\mathfrak{K}_1F)(0, \epsilon, \mathbf{p}, \lambda) &= \epsilon\left\{ip\frac{\partial}{\partial p_1}F(0, \epsilon, \mathbf{p}, \lambda) + \frac{p_2}{p + p_3}(MF)(0, \epsilon, \mathbf{p}, \lambda)\right.\\
&\quad \left. + \left[\frac{p_1^2}{p^2(p + p_3)} - \frac{1}{p}\right](T_1F)(0, \epsilon, \mathbf{p}, \lambda) + \frac{p_1p_2}{p^2(p + p_3)}(T_2F)(0, \epsilon, \mathbf{p}, \lambda)\right\},\\
(\mathfrak{K}_2F)(0, \epsilon, \mathbf{p}, \lambda) &= \epsilon\left\{ip\frac{\partial}{\partial p_2}F(0, \epsilon, \mathbf{p}, \lambda) - \frac{p_1}{p + p_3}(MF)(0, \epsilon, \mathbf{p}, \lambda)\right.\\
&\quad \left. + \frac{p_1p_2}{p^2(p + p_3)}(T_1F)(0, \epsilon, \mathbf{p}, \lambda) + \left[\frac{p_2^2}{p^2(p + p_3)} - \frac{1}{p}\right](T_2F)(0, \epsilon, \mathbf{p}, \lambda)\right\},\\
(\mathfrak{K}_3F)(0, \epsilon, \mathbf{p}, \lambda) &= \epsilon\left\{ip\frac{\partial}{\partial p_3}F(0, \epsilon, \mathbf{p}, \lambda) + \frac{1}{p^2}[p_1(T_1F)(0, \epsilon, \mathbf{p}, \lambda) + p_2(T_2F)(0, \epsilon, \mathbf{p}, \lambda)]\right\}. \qquad (1.9)
\end{aligned}
$$

In (1.9) the operators M, T_1, and T_2 can be expressed in terms of matrices or kernels which act on λ.

$$
\begin{aligned}
(MF)(0, \epsilon, \mathbf{p}, \lambda) &= \sum_{\lambda'} M(\lambda \mid \lambda')F(0, \epsilon, \mathbf{p}, \lambda'),\\
(T_1F)(0, \epsilon, \mathbf{p}, \lambda) &= \sum_{\lambda'} T_1(\lambda \mid \lambda')F(0, \epsilon, \mathbf{p}, \lambda'),\\
(T_2F)(0, \epsilon, \mathbf{p}, \lambda) &= \sum_{\lambda'} T_2(\lambda \mid \lambda')F(0, \epsilon, \mathbf{p}, \lambda'). \qquad (1.10)
\end{aligned}
$$

[The operator M in (1.10) is not to be confused with the mass operator M.] The kernels may depend on ϵ but not on $\mathbf{p}$. The kernels satisfy the commutation relations for the generators of the Euclidean group in the plane

$$
\begin{aligned}
&\sum_{\lambda''}[T_1(\lambda \mid \lambda'')T_2(\lambda'' \mid \lambda') - T_2(\lambda \mid \lambda'')T_1(\lambda'' \mid \lambda')] = 0,\\
&\sum_{\lambda''}[T_1(\lambda \mid \lambda'')M(\lambda'' \mid \lambda') - M(\lambda \mid \lambda'')T_1(\lambda'' \mid \lambda')]\\
&\qquad = -iT_2(\lambda \mid \lambda'),\\
&\sum_{\lambda''}[T_2(\lambda \mid \lambda'')M(\lambda'' \mid \lambda') - M(\lambda \mid \lambda'')T_2(\lambda'' \mid \lambda')]\\
&\qquad = iT_1(\lambda \mid \lambda'). \qquad (1.11)
\end{aligned}
$$

In order that the operator $\mathbf{J}^2$ be Hermitian, it is necessary that the eigenvalues of M be integer or half-odd integer.

The way that the kernels are found is discussed shortly.

It is to be noted that the operators are in the Foldy–Shirokov form for the nonzero-mass case and in the Lomont–Moses form in the zero-mass case. However, the operators S_i, M, and T_i are reducible in general as they appear "naturally" in the construction which follows. The reduction of the rotation and Euclidean groups is then carried out in the λ space, this reduction being a well-known one. Thus in the new λ space, the representation of the Lorentz group is completely reduced.

The problem of finding the matrices or kernels of (1.8) and (1.10) are intimately related to the problem of finding the transformation between the basis in which the vector $|\psi\rangle$ is represented by $f(\zeta)$ and that in which it is represented by $F(\mu, \epsilon, \mathbf{p}, \lambda)$.

Let us introduce the set of kets $|\zeta\rangle$ and the set $|\mu, \epsilon, \mathbf{p}, \lambda\rangle$ corresponding to the basis above. As usual we write

$$
\begin{aligned}
f(\zeta) &= \langle \zeta \mid \psi\rangle,\\
F(\mu, \epsilon, \mathbf{p}, \lambda) &= \langle \mu, \epsilon, \mathbf{p}, \lambda \mid \psi\rangle. \qquad (1.12)
\end{aligned}
$$

Since the kets $|\mu, \epsilon, \mathbf{p}, \lambda\rangle$ form a complete set, the following resolution of the identity is possible:

$$
I = \sum_{\epsilon}\int d\mu \int \frac{d\mathbf{p}}{\omega(\mu, p)}\sum_{\lambda}|\mu, \epsilon, \mathbf{p}, \lambda\rangle\langle\mu, \epsilon, \mathbf{p}, \lambda|, \qquad (1.13)
$$

which corresponds to the following orthonormality relations:

$$
\begin{aligned}
&\langle \mu, \epsilon, \mathbf{p}, \lambda \mid \mu', \epsilon', \mathbf{p}', \lambda'\rangle\\
&\qquad = \delta_{\epsilon,\epsilon'}\delta_{\lambda,\lambda'}\delta(\mu - \mu')\delta(\mathbf{p} - \mathbf{p}')\omega(\mu, p). \qquad (1.13a)
\end{aligned}
$$

Equation (1.13) corresponds to the inner product (1.4). From (1.12) and (1.13) we have the relation between $f(\zeta)$ and $F(\mu, \epsilon, \mathbf{p}, \lambda)$,

$$
\begin{aligned}
f(\zeta) = \sum_{\epsilon}\int d\mu \sum_{\lambda}\int \frac{d\mathbf{p}}{\omega(\mu, p)}\\
\times \langle \zeta \mid \mu, \epsilon, \mathbf{p}, \lambda\rangle F(\mu, \epsilon, \mathbf{p}, \lambda). \qquad (1.14)
\end{aligned}
$$

We now indicate how one obtains the transformation functions $(\zeta \mid \mu, \epsilon, \mathbf{p}, \lambda\rangle$ and the matrices which appear in (1.8) and (1.10).

It is convenient to introduce a notation which shows how an operator acts on the variables which we have denoted by ζ. Thus if A is an operator, we define A^ζ by

$$(AF)(\zeta) = A^\zeta f(\zeta). \tag{1.15}$$

We need the operators A^ζ because we want to work on the ζ variable in the transformation functions and the notation $(Af)(\zeta)$ becomes cumbersome.

Let us first obtain all the linearly independent solutions $g(\zeta; \mu, \epsilon, \lambda)$ of the equations

$$\begin{aligned} P_i^\zeta g(\zeta; \mu, \epsilon, \lambda) &= 0, \\ H^\zeta g(\zeta; \mu, \epsilon, \lambda) &= \epsilon\mu g(\zeta; \mu, \epsilon, \lambda), \end{aligned} \tag{1.16}$$

for $\mu > 0$, the three values of i, and both values of ϵ. The variable λ is used to label the linearly independent solutions and for some representations may be a continuous variable. It is shown that solutions of (1.16) always exist. It is also clear that the choice of variable λ is not unique. For example, if λ takes on only two values, suitable linear combinations of the solutions of (1.16) are also linearly independent solutions.

For *any* choice of variable λ we show that

$$J_i^\zeta g(\zeta; \mu, \epsilon, \lambda) = \sum_{\lambda'} S_i(\lambda' \mid \lambda) g(\zeta; \mu, \epsilon, \lambda'), \tag{1.17}$$

where the matrices or kernels S_i satisfy the commutation rules (1.8a) for the angular momentum operators. These kernels depend on μ, ϵ, and the choice of variables λ.

Let us define $\langle \zeta \mid \mu, \epsilon, \mathbf{p}, \lambda)$ by

$$\langle \zeta \mid \mu, \epsilon, \mathbf{p}, \lambda) = [\exp(-i\boldsymbol{\nu} \cdot \boldsymbol{\mathfrak{J}})]^\zeta g(\zeta; \mu, \epsilon, \lambda), \tag{1.18}$$

where

$$\boldsymbol{\nu} \cdot \boldsymbol{\mathfrak{J}} = \sum_i \nu_i \mathfrak{J}_i, \tag{1.19}$$

and the vectors $\mathbf{p}$ and $\boldsymbol{\nu}$ are in a one-to-one correspondence given by

$$\mathbf{p} = -\epsilon\mu\boldsymbol{\nu}(\sinh \nu/\nu), \quad \nu = |\boldsymbol{\nu}|, \tag{1.20}$$

from which

$$p = |\mathbf{p}| = \mu \sinh \nu, \quad \omega(\mu, p) = \mu \cosh \nu. \tag{1.20a}$$

We may consider the function $\langle \zeta \mid \mu, \epsilon, \mathbf{p}, \lambda)$ to be the inner product of the bra $\langle \zeta|$ and a ket $|\mu, \epsilon, \mathbf{p}, \lambda)$. Since the inner product in terms of the ζ representation is known, we can form the inner product $(\mu, \epsilon, \mathbf{p}, \lambda \mid \mu', \varepsilon', \mathbf{p}', \lambda')$. In general

$$\begin{aligned} &(\mu, \epsilon, \mathbf{p}, \lambda \mid \mu', \epsilon', \mathbf{p}', \lambda') \\ &\quad = \omega(\mu, p)\delta_{\epsilon,\epsilon'}\delta(\mu - \mu')\delta(\mathbf{p} - \mathbf{p}')k(\lambda \ \lambda'), \end{aligned} \tag{1.21}$$

where $k(\lambda \mid \lambda')$ is a positive definite matrix or kernel which, in general, is a function of μ and ϵ. The function $\delta(\mu - \mu')$ is a suitably generalized δ function to take into account possible point eigenvalues.

If $k(\lambda \mid \lambda') = \delta_{\lambda,\lambda'}$, then the ket we seek is given by $|\mu, \epsilon, \mathbf{p}, \lambda) = |\mu, \epsilon, \mathbf{p}, \lambda)$. The kernels $S_i(\lambda \mid \lambda')$ of (1.8) are given by (1.17). Furthermore these kernels are a *Hermitian* representation of the infinitesimal generators of rotation group.

If $k(\lambda \mid \lambda') \neq \delta_{\lambda,\lambda'}$, we show in the body of the paper that a knowledge of this matrix enables us to make a new choice of degeneracy variables such that $k(\lambda \mid \lambda') = \delta_{\lambda,\lambda'}$ in the new variables.

It is easy to see that the variables λ which lead to the transformation function are not unique. In practice we try to be sufficiently clever to choose λ so that a redefinition is not necessary and so that the kernels S_i are sufficiently simple in form that they can be reduced to the irreducible representations of the rotation group.

Having indicated how one obtains the transformation functions $\langle \zeta \mid \mu, \epsilon, \mathbf{p}, \lambda\rangle$ for $\mu \neq 0$ and the kernels $S_i(\lambda \mid \lambda')$ we now proceed to obtain the transformation functions for $\mu = 0$ and the kernels $M(\lambda \mid \lambda')$ and $T_i(\lambda \mid \lambda')$. The technique is quite close to that for $\mu \neq 0$.

We first find all the linearly independent solutions $f(\zeta; \epsilon, \lambda)$ of the equations

$$\begin{aligned} &P_1^\zeta f(\zeta; \epsilon, \lambda) = 0, P_2^\zeta f(\zeta; \epsilon, \lambda) = 0, \\ &P_3^\zeta f(\zeta; \epsilon, \lambda) = f(\zeta; \epsilon, \lambda), H^\zeta f(\zeta; \epsilon, \lambda) = \epsilon f(\zeta; \epsilon, \lambda). \end{aligned} \tag{1.22}$$

It is shown that kernels $M(\lambda \mid \lambda')$, $T_1(\lambda \mid \lambda')$, and $T_2(\lambda \mid \lambda')$ always exist such that

$$\begin{aligned} J_3^\zeta f(\zeta; \epsilon, \lambda) &= \sum_{\lambda'} M(\lambda' \mid \lambda) f(\zeta; \epsilon, \lambda'), \\ [-\epsilon\mathfrak{J}_1 - J_2]^\zeta f(\zeta; \epsilon, \lambda) &= \sum_{\lambda'} T_1(\lambda' \mid \lambda) f(\zeta; \epsilon, \lambda'), \\ [-\epsilon\mathfrak{J}_2 + J_1]^\zeta f(\zeta; \epsilon, \lambda) &= \sum_{\lambda'} T_2(\lambda' \mid \lambda) f(\zeta; \epsilon, \lambda'), \end{aligned} \tag{1.23}$$

where the kernels obey the commutation rules (1.11) for the infinitesimal generators of the Euclidean group in the plane. In general the kernels depend on ϵ.

Let us define the function $\langle \zeta \mid 0, \epsilon, \mathbf{p}, \lambda)$ by

$$\begin{aligned} \langle \zeta \mid 0, \epsilon, \mathbf{p}, \lambda) = \{&\exp[i(\boldsymbol{\omega} \cdot \mathbf{J})] \\ &\cdot \exp[iv\mathfrak{J}_3]\}^\zeta f(\zeta; \epsilon, \lambda), \end{aligned} \tag{1.24}$$

where the vector $\mathbf{p}$ is related to the vector $\boldsymbol{\omega}$ and the scalar v by

$$p = |\mathbf{p}| = e^{\epsilon v}, \tag{1.24a}$$

$$p_1 = -p(\sin \omega/\omega)\omega_2, p_2 = p(\sin \omega/\omega)\omega_1, p_3 = p \cos \omega,$$

$$\omega_3 = 0, \quad \omega = |\boldsymbol{\omega}|. \tag{1.24b}$$

In (1.24a) and (1.24b) v is not to be confused with ν of Eq. (1.20) and ω is not to be confused with $\omega(\mu, p)$ of Eq. (1.5).

In a manner similar to that for $\mu = 0$ we can introduce the inner product $(0, \epsilon, \mathbf{p}, \lambda \mid 0, \epsilon', \mathbf{p}', \lambda')$ and one can show

$$(0, \epsilon, \mathbf{p}, \lambda \mid 0, \epsilon', \mathbf{p}', \lambda') = p\delta_{\epsilon,\epsilon'}\delta(\mathbf{p} - \mathbf{p}')k(\lambda \mid \lambda'), \quad (1.25)$$

where $k(\lambda \mid \lambda')$ is a positive definite matrix which may depend on ϵ, but not on $\mathbf{p}$. If $k(\lambda \mid \lambda') = \delta_{\lambda,\lambda'}$, then $|0, \epsilon, \mathbf{p}, \lambda\rangle = |0, \epsilon, \mathbf{p}, \lambda)$, and the kernels of Eq. (1.23) are Hermitian and are those of (1.10). If $k(\lambda \mid \lambda') \neq \delta_{\lambda,\lambda'}$, one can introduce new variables λ for which $k(\lambda \mid \lambda') = \delta_{\lambda,\lambda'}$ in terms of the new variables.

2. INTEGRATED FORM OF THE COMMUTATION RULES

In the next two sections of the present paper it is our objective to verify that the kets $|\mu, \epsilon, \mathbf{p}, \lambda\rangle$ are obtained by the method summarized in the previous section and that the operators corresponding to the infinitesimal generators of the inhomogeneous Lorentz group act as in Eqs. (1.7) and (1.9) in terms of the basis which these kets provide. Toward this end it is useful to give a "dictionary" showing how the infinitesimal generators are transformed by the application of certain unitary operators which, in a sense, are obtained by integrating the infinitesimal generators.

We find the following familiar theorem very useful: Let A and B be any two operators. Then

$$e^{-A}Be^{A} = \sum_{n=0}^{\infty} \frac{\{B, A\}^{(n)}}{n!}, \quad (2.1)$$

where $\{B, A\}^{(n)}$ is defined by induction by means of commutators:

$$\{B, A\}^{(n)} = [\{B, A\}^{(n-1)}, A],$$
$$\{B, A\}^{(0)} = B. \quad (2.2)$$

Then on using (2.1) and the commutation rules for the infinitesimal generators (1.1) we obtain the following results on using an obvious notation for the operators, where $\boldsymbol{\xi}$ $(= \xi_1, \xi_2, \xi_3)$ is any vector with $\xi = |\boldsymbol{\xi}|$:

$$\exp(i\boldsymbol{\xi}\cdot\boldsymbol{\mathcal{J}})H\exp(-i\boldsymbol{\xi}\cdot\boldsymbol{\mathcal{J}}) = H\cosh\xi - (\boldsymbol{\xi}\cdot\mathbf{P})(\sinh\xi/\xi), \quad (2.3)$$

$$\exp(i\boldsymbol{\xi}\cdot\boldsymbol{\mathcal{J}})\mathbf{P}\exp(-i\boldsymbol{\xi}\cdot\boldsymbol{\mathcal{J}}) = \mathbf{P} - H(\sinh\xi/\xi)\boldsymbol{\xi} + (\boldsymbol{\xi}\cdot\mathbf{P})[(\cosh\xi - 1)/\xi^2]\boldsymbol{\xi}, \quad (2.4)$$

$$\exp(i\boldsymbol{\xi}\cdot\boldsymbol{\mathcal{J}})\boldsymbol{\mathcal{J}}\exp(-i\boldsymbol{\xi}\cdot\boldsymbol{\mathcal{J}}) = \boldsymbol{\mathcal{J}}\cosh\xi - (\boldsymbol{\xi}\times\mathbf{J}) \times (\sinh\xi/\xi) - (\boldsymbol{\xi}\cdot\boldsymbol{\mathcal{J}})[(\cosh\xi - 1)/\xi^2]\boldsymbol{\xi}, \quad (2.5)$$

$$\exp(i\boldsymbol{\xi}\cdot\boldsymbol{\mathcal{J}})\mathbf{J}\exp(-i\boldsymbol{\xi}\cdot\boldsymbol{\mathcal{J}}) = \mathbf{J}\cosh\xi + (\boldsymbol{\xi}\times\boldsymbol{\mathcal{J}}) \times (\sinh\xi/\xi) - (\boldsymbol{\xi}\cdot\mathbf{J})[(\cosh\xi - 1)/\xi^2]\boldsymbol{\xi}. \quad (2.6)$$

Finally, if $\mathbf{W}$ is any of the operators $\mathbf{P}$, $\mathbf{J}$, or $\boldsymbol{\mathcal{J}}$, then

$$\exp(-i\boldsymbol{\xi}\cdot\mathbf{J})\mathbf{W}\exp(i\boldsymbol{\xi}\cdot\mathbf{J}) = \mathbf{W}\cos\xi + (\boldsymbol{\xi}\cdot\mathbf{W}) \times [(1 - \cos\xi)/\xi^2]\boldsymbol{\xi} - (\boldsymbol{\xi}\times\mathbf{W})(\sin\xi/\xi). \quad (2.7)$$

3. BASIS VECTORS FOR NONZERO MASS

We now construct the kets $|\mu, \epsilon, \mathbf{p}, \lambda\rangle$ for $\mu \neq 0$. We also show that in the space spanned by these kets, infinitesimal generators behave as in (1.7), i.e., this basis gives the Foldy–Shirokov realization of the infinitesimal generators. Our construction is obtained from a sequence of necessary conditions for the existence of a Hermitian realization for the infinitesimal generators. Since these conditions come from the commutation rules and are abstract in nature, it is not necessary to indicate explicitly that we are working in the ζ representation.

First of all, if the mass operator M of Eq. (1.2) is to be Hermitian, it must have a spectrum, which, in principle at least, can be obtained by working in the ζ representation. As explained in Sec. 1, we assume that the eigenvalues μ are nonnegative. Similarly, the eigenvalues of $H/|H|$, which occur, can be found. The eigenvalues of this operator, which we have denoted by ϵ, can have only the eigenvalues ± 1.

Since the operators M, $H/|H|$, and P_i $(i = 1, 2, 3)$ all commute, we look for a basis in which these operators are spectrally represented simultaneously. The kets which span this basis are tentatively denoted by $|\mu, \epsilon, \mathbf{p}, \lambda)$. We assume that the Hilbert space which is spanned is separable.

We now consider only the subspace of the Hilbert space for fixed values of ϵ and μ where $\mu \neq 0$.

Our first theorem is the following:

Theorem 1: The simultaneous spectrum of the operators P_i includes the point $\mathbf{p} = 0$.

Proof: Since we assume that the operators P_i are Hermitian, they must have a simultaneous eigenvalue, say $\mathbf{p}$. Let us call a corresponding eigenket of M, $H/|H|$, P_i, $|\mu, \epsilon, \mathbf{p}, \lambda)$ as above. But let us also construct the vector $[\exp(i\boldsymbol{\nu}\cdot\boldsymbol{\mathcal{J}})]\,|\mu, \epsilon, \mathbf{p}, \lambda) = |\Phi)$, where $\boldsymbol{\nu}$ and $\mathbf{p}$ are related by (1.20). Then $P_i\,|\Phi) = 0$ for all i. For from (2.4)

$$P_i[\Phi) = \exp(i\boldsymbol{\nu}\cdot\boldsymbol{\mathcal{J}})\exp(-i\boldsymbol{\nu}\cdot\boldsymbol{\mathcal{J}})P_i \times \exp(i\boldsymbol{\nu}\cdot\boldsymbol{\mathcal{J}})\,|\mu, \epsilon, \mathbf{p}, \lambda)$$
$$= \exp(i\boldsymbol{\nu}\cdot\boldsymbol{\mathcal{J}})\{P_i + H(\sinh\nu/\nu)\nu_i + (\boldsymbol{\nu}\cdot\mathbf{P})[(\cosh\nu - 1)/\nu^2]\nu_i\}\,|\mu, \epsilon, \mathbf{p}, \lambda). \quad (3.1)$$

But

$$H = (H/|H|)[P^2 + M^2]^{\frac{1}{2}} \quad (3.2)$$

and thus

$$H\,|\mu, \epsilon, \mathbf{p}, \lambda) = \epsilon\omega(\mu, p)\,|\mu, \epsilon, \mathbf{p}, \lambda). \quad (3.3)$$

Also by definition $P_i\,|\mu, \epsilon, \mathbf{p}, \lambda) = p_i\,|\mu, \epsilon, \mathbf{p}, \lambda)$. On expressing p_i in terms of ν_i in (3.1) one obtains $P_i\,|\Phi) = 0$. Thus we have shown that each of the operators P_i has zero in its spectrum, since $|\Phi)$ is a simultaneous eigenket belonging to these eigenvalues.

Let us denote the kets belonging to the eigenvalue μ of M, ϵ of $H/|H|$, and $\mathbf{p} = 0$ of $\mathbf{P}$ by $|\mu, \epsilon, 0, \lambda)$. The variable λ is a degeneracy variable and is to be chosen in such a manner that as λ runs through its values, we obtain all the linearly independent kets belonging to these eigenvalues of the operators. We then have Theorem 2.

Theorem 2: The vector $\exp(-i\mathbf{v}\cdot\mathbf{\mathfrak{J}})\,|\mu, \epsilon, 0, \lambda)$ is a simultaneous eigenket of the operators M, $H/|H|$, $\mathbf{P}$ belonging to the eigenvalues μ, ϵ, $\mathbf{p}$, respectively, where $\mathbf{p}$ is given by (1.20). Furthermore, the vectors belonging to different values of λ are linearly independent. Finally, any simultaneous eigenket of $M, H/|H|$, and $\mathbf{P}$ having the eigenvalues μ, ϵ, and $\mathbf{p}$ respectively is a linear combination in λ of the vectors

$$\exp(-i\mathbf{v}\cdot\mathbf{\mathfrak{J}})\,|\mu, \epsilon, 0, \lambda).$$

Proof: Since M and $H/|H|$ commute with all of the infinitesimal generators of the group, these operators commute also with the operator $\exp(-i\mathbf{v}\cdot\mathbf{\mathfrak{J}})$. Hence the vector $\exp(-i\mathbf{v}\cdot\mathbf{\mathfrak{J}})\,|\mu, \epsilon, 0, \lambda)$ is a simultaneous eigenket of M and $H/|H|$ with eigenvalues μ and ϵ. The proof of $P_i \exp(-i\mathbf{v}\cdot\mathbf{\mathfrak{J}})\,|\mu, \epsilon, 0, \lambda) = p_i \exp(-i\mathbf{v}\cdot\mathbf{\mathfrak{J}})\,|\mu, \epsilon, 0, \lambda)$ makes use of Eq. (2.4) in much the same way as in the proof of Theorem 1. The proof of the remainder of Theorem 2 is almost obvious, and we do not give it.

We are now in a position to make the following definition:

Definition:

$$|\mu, \epsilon, \mathbf{p}, \lambda) = \exp(-i\mathbf{v}\cdot\mathbf{\mathfrak{J}})\,|\mu, \epsilon, 0, \lambda). \quad (3.4)$$

From the previous theorems these kets span the subspace of the Hilbert space for which $\mu > 0$.

From our construction we see that in general the variables λ depend on μ and ϵ but are independent of $\mathbf{p}$. We also see that our kets depend upon the way in which we chose $|\mu, \epsilon, 0, \lambda)$. But because of the last part of Theorem 2 any other choice is equivalent to merely using linear combinations of our original choice of kets.

The function $g(\zeta, \mu, \epsilon, \lambda)$ of Eqs. (1.16) are simply $g(\zeta; \mu, \epsilon, \lambda) = \langle \zeta \mid \mu, \epsilon, 0, \lambda)$. Equation (3.4) is the same as (1.18).

Having now introduced our kets, we now wish to show how the infinitesimal generators act on these kets. First of all, we have by Theorem 2:

$$P_i\,|\mu, \epsilon, \mathbf{p}, \lambda) = p_i\,|\mu, \epsilon, \mathbf{p}, \lambda),$$
$$H\,|\mu, \epsilon, \mathbf{p}, \lambda) = \epsilon\omega(\mu, p)\,|\mu, \epsilon, \mathbf{p}, \lambda). \quad (3.5)$$

We now find $J_i\,|\mu, \epsilon, p, \lambda)$.

First we need some preliminary definitions and lemmas.

Lemma 1: $J_i\,|\mu, \epsilon, 0, \lambda)$ is a simultaneous eigenstate of the operators M, $H/|H|$, and $\mathbf{P}$ with the eigenvalues μ, ϵ, and 0, respectively.

Proof: That $J_i\,|\mu, \epsilon, 0, \lambda)$ is an eigenstate of M and $H/|H|$ with the stated eigenvalues follows from the fact that these latter operators commute with J_i. That $P_iJ_k\,|\mu, \epsilon, 0, \lambda) = 0$ for all i and k follows from the commutation rules (1.1) for $\mathbf{P}$ and $\mathbf{J}$.

This lemma allows us to define the kernel $S_i(\lambda \mid \lambda')$ in the following way:

Definition:

$$J_i\,|\mu, \epsilon, 0, \lambda) = \sum_{\lambda'} S_i(\lambda' \mid \lambda)\,|\mu, \epsilon, 0, \lambda'). \quad (3.6)$$

The kernels S_i *must* exist since $J_i\,|\mu, \epsilon, 0, \lambda)$ must be a linear combination of the kets $|\mu, \epsilon, 0, \lambda)$ because of the way in which we chose these kets to be the entire linearly independent set corresponding to the eigenvalues μ, ϵ, and $\mathbf{p} = 0$.

Definition:

$$S_i\,|\mu, \varepsilon, \mathbf{p}, \lambda) = \sum_{\lambda'} S_i(\lambda' \mid \lambda)\,|\mu, \varepsilon, \mathbf{p}, \lambda'). \quad (3.7)$$

The following theorem is important:

Theorem 3: The operators S_i satisfy the commutation rules for the angular momentum operators

$$[S_i, S_j] = i\sum_k \epsilon_{ijk}S_k. \quad (3.8)$$

Proof: This theorem is proved by applying the commutation relation $[J_i, J_j] = i\sum_k \epsilon_{ijk}J_k$ to the ket $|\mu, \epsilon, 0, \lambda)$ and using (3.6) and (3.7).

The operators S_i prove to be the spin operators. Their introduction is, of course, motivated by the original work of Wigner using the global "little group" approach.

We now state and prove one of the principal theorems of this section.

Theorem 4: The operators J_i are given in this basis by

$$J_i\,|\mu, \epsilon, \mathbf{p}, \lambda) = [i(\mathbf{p}\times\mathbf{\nabla})_i + S_i]\,|\mu, \epsilon, \mathbf{p}, \lambda), \quad (3.9)$$

where $\nabla_i = \partial/\partial p_i$.

Before proving the theorem we make one more definition suggested by (3.9).

Definition: The components of the orbital angular momentum L_i are defined by

$$L_i\,|\mu, \epsilon, \mathbf{p}, \lambda) = i(\mathbf{p}\times\mathbf{\nabla})_i\,|\mu, \epsilon, \mathbf{p}, \lambda). \quad (3.10)$$

Hence we may write (3.9) as

$$\mathbf{J} = \mathbf{L} + \mathbf{S}. \tag{3.10a}$$

Proof of Theorem 4: We prove (3.9) for $i = 1$. The proof of (3.9) for $i = 2, 3$ is similar.

We start with

$$J_1 |\mu, \epsilon, \mathbf{p}, \lambda) = i(\partial/\partial\alpha) \exp(-i\alpha J_1) \times \exp(-i\boldsymbol{\nu}\cdot\boldsymbol{\mathfrak{J}}) |\mu, \epsilon, 0, \lambda)\}_{\alpha=0}. \tag{3.11}$$

But on replacing $\mathbf{W}$ by $\boldsymbol{\mathfrak{J}}$ in (2.7)

$$\exp(-i\alpha J_1)\exp(-i\boldsymbol{\nu}\cdot\boldsymbol{\mathfrak{J}}) = \exp(-i\alpha J_1)\exp(-i\boldsymbol{\nu}\cdot\boldsymbol{\mathfrak{J}})\exp(i\alpha J_1)\exp(-i\alpha J_1) = \exp(-i\boldsymbol{\nu}'\cdot\boldsymbol{\mathfrak{J}})\exp(-i\alpha J_1), \tag{3.12}$$

where the vector $\boldsymbol{\nu}'$ has the components

$$\boldsymbol{\nu}' = (\nu_1, \nu_2\cos\alpha - \nu_3\sin\alpha, \nu_3\cos\alpha + \nu_2\sin\alpha). \tag{3.12a}$$

On substituting (3.12) into (3.11) and on using (3.6) and the definition (3.7), we obtain

$$J_1 |\mu, \epsilon, \mathbf{p}, \lambda) = i(\partial/\partial\alpha)\exp(-i\alpha S_1) |\mu, \epsilon, \mathbf{p}', \lambda)\}_{\alpha=0}, \tag{3.13}$$

where $\mathbf{p}'$ is related to $\boldsymbol{\nu}'$ as $\mathbf{p}$ is related to $\boldsymbol{\nu}$ through (1.20). Thus from (3.12a)

$$\begin{aligned} p_1' &= p_1, \\ p_2' &= p_2\cos\alpha - p_3\sin\alpha, \\ p_3' &= p_3\cos\alpha + p_2\sin\alpha. \end{aligned} \tag{3.14}$$

Then on carrying out the differentiation in (3.13) one obtains Theorem 4 for $i = 1$.

Theorem 5: The operators $\mathfrak{J}_i$ are given in this basis by

$$\mathfrak{J}_i |\mu, \epsilon, \mathbf{p}, \lambda) = \epsilon\left[-i\omega(\mu, p)\frac{\partial}{\partial p_i} + \frac{(\mathbf{p}\times\mathbf{S})_i}{\omega(\mu, p) + \mu}\right] |\mu, \epsilon, \mathbf{p}, \lambda). \tag{3.15}$$

Proof: On using (2.5), (3.4), (3.6), and (3.7) we have

$$\begin{aligned} \mathfrak{J}_i |\mu, \epsilon, \mathbf{p}, \lambda) &= \mathfrak{J}_i \exp(-i\boldsymbol{\nu}\cdot\boldsymbol{\mathfrak{J}}) |\mu, \epsilon, 0, \lambda) \\ &= \exp(-i\boldsymbol{\nu}\cdot\boldsymbol{\mathfrak{J}})\exp(i\boldsymbol{\nu}\cdot\boldsymbol{\mathfrak{J}})\mathfrak{J}_i \times \exp(-i\boldsymbol{\nu}\cdot\boldsymbol{\mathfrak{J}}) |\mu, \epsilon, 0, \lambda) \\ &= \cosh\nu\exp(-i\boldsymbol{\nu}\cdot\boldsymbol{\mathfrak{J}})\mathfrak{J}_i |\mu, \epsilon, 0, \lambda) - \frac{\sinh\nu}{\nu}\exp(-i\boldsymbol{\nu}\cdot\boldsymbol{\mathfrak{J}})(\boldsymbol{\nu}\times\mathbf{J})_i |\mu, \epsilon, 0, \lambda) - \nu_i\frac{\cosh\nu - 1}{\nu^2}\exp(-i\boldsymbol{\nu}\cdot\boldsymbol{\mathfrak{J}})(\boldsymbol{\nu}\cdot\boldsymbol{\mathfrak{J}}) |\mu, \epsilon, 0, \lambda) \\ &= \cosh\nu\exp(-i\boldsymbol{\nu}\cdot\boldsymbol{\mathfrak{J}})\mathfrak{J}_i\exp(i\boldsymbol{\nu}\cdot\boldsymbol{\mathfrak{J}}) |\mu, \epsilon, \mathbf{p}, \lambda) - \frac{\sinh\nu}{\nu}(\boldsymbol{\nu}\times\mathbf{S})_i |\mu, \epsilon, \mathbf{p}, \lambda) - \nu_i\frac{\cosh\nu - 1}{\nu^2}(\boldsymbol{\nu}\cdot\boldsymbol{\mathfrak{J}}) |\mu, \epsilon, \mathbf{p}, \lambda). \end{aligned} \tag{3.16}$$

Again applying (2.5) to the expression

$$\exp(-i\boldsymbol{\nu}\cdot\boldsymbol{\mathfrak{J}})\mathfrak{J}_i\exp(i\boldsymbol{\nu}\cdot\boldsymbol{\mathfrak{J}}),$$

which occurs in the last of Eqs. (3.16) we obtain

$$\begin{aligned} \mathfrak{J}_i |\mu, \epsilon, \mathbf{p}, \lambda) &= \cosh^2\nu\,\mathfrak{J}_i |\mu, \epsilon, \mathbf{p}, \lambda) + (\sinh\nu\cosh\nu/\nu)(\boldsymbol{\nu}\times\mathbf{J})_i |\mu, \epsilon, \mathbf{p}, \lambda) - \nu_i(\boldsymbol{\nu}\cdot\boldsymbol{\mathfrak{J}}) \times [\cosh\nu(\cosh\nu - 1)/\nu^2] |\mu, \epsilon, \mathbf{p}, \lambda) - (\sinh\nu/\nu) \times (\boldsymbol{\nu}\times\mathbf{S})_i |\mu, \epsilon, \mathbf{p}, \lambda) - \nu_i[(\cosh\nu - 1)/\nu^2] \times (\boldsymbol{\nu}\cdot\boldsymbol{\mathfrak{J}}) |\mu, \epsilon, \mathbf{p}, \lambda). \end{aligned} \tag{3.17}$$

We may regard (3.17) as being an equation for $\mathfrak{J}_i |\mu, \epsilon, \mathbf{p}, \lambda)$ which we can solve by bringing over to the left-hand side the first of the terms on the right.

By using (3.10a) and combining terms we obtain

$$\begin{aligned} -\sinh^2\nu\,\mathfrak{J}_i |\mu, \epsilon, \mathbf{p}, \lambda) &= (\cosh\nu\sinh\nu/\nu)(\boldsymbol{\nu}\times\mathbf{L})_i |\mu, \epsilon, \mathbf{p}, \lambda) + (\sinh\nu/\nu)(\cosh\nu - 1)(\boldsymbol{\nu}\times\mathbf{S})_i |\mu, \epsilon, \mathbf{p}, \lambda) - \nu_i(\sinh^2\nu/\nu^2)(\boldsymbol{\nu}\cdot\boldsymbol{\mathfrak{J}}) |\mu, \epsilon, \mathbf{p}, \lambda). \end{aligned} \tag{3.18}$$

Now

$$(\boldsymbol{\nu}\cdot\boldsymbol{\mathfrak{J}}) |\mu, \epsilon, p, \lambda) = i(\partial/\partial\alpha)\exp(-i\alpha\boldsymbol{\nu}\cdot\boldsymbol{\mathfrak{J}}) \times |\mu, \epsilon, 0, \lambda)\}_{\alpha=1} = i(\partial/\partial\alpha) |\mu, \epsilon, \mathbf{p}', \lambda)\}_{\alpha=1}, \tag{3.19}$$

where $\mathbf{p}'$ is a function of α through

$$\mathbf{p}' = -\epsilon\mu(\boldsymbol{\nu}/\nu)\sinh\alpha\nu. \tag{3.20}$$

It follows that

$$(\nu_i/\nu^2)(\boldsymbol{\nu}\cdot\boldsymbol{\mathfrak{J}}) |\mu, \epsilon, \mathbf{p}, \lambda) = -i\epsilon\omega(\mu, p) \times (p_i/p^2)\mathbf{p}\cdot\boldsymbol{\nabla} |\mu, \epsilon, \mathbf{p}, \lambda). \tag{3.21}$$

On using (3.10) and (3.21) and in (3.18) one finally obtains Eq. (3.15).

Theorem 6: The kets $|\mu, \epsilon, \mathbf{p}, \lambda)$ satisfy the orthonormality relations

$$(\mu, \epsilon, \mathbf{p}, \lambda \,|\, \mu', \epsilon', \mathbf{p}', \lambda') = \omega(\mu, p)\delta_{\epsilon,\epsilon'}\delta(\mu - \mu') \times \delta(\mathbf{p} - \mathbf{p}')k(\lambda \,|\, \lambda'), \tag{3.22}$$

where $k(\lambda \,|\, \lambda')$ is a positive definite matrix (or kernel of a positive definite integral operator) which may depend on μ and ϵ but is independent of $\mathbf{p}$ or $\mathbf{p}'$. We prove Theorem 6 by means of a series of lemmas.

Lemma I: Let us define the kernel

$$k(\mu, \epsilon, \mathbf{p}, \lambda \,|\, \mu', \epsilon', \mathbf{p}', \lambda') = (\mu, \epsilon, \mathbf{p}, \lambda \,|\, \mu', \epsilon', \mathbf{p}', \lambda').$$

Then the kernel k is Hermitian and positive definite in terms of the arguments μ, ϵ, $\mathbf{p}$, λ considered collectively, i.e.,

$$k(\mu, \epsilon, \mathbf{p}, \lambda \,|\, \mu', \epsilon', \mathbf{p}', \lambda') = k^*(\mu', \epsilon', \mathbf{p}', \lambda' \,|\, \mu, \epsilon, \mathbf{p}, \lambda), \tag{3.23}$$

$$\sum_{\epsilon}\sum_{\epsilon'}\int d\mu\int d\mu'\sum_{\lambda}\sum_{\lambda'}\int d\mathbf{p}\int d\mathbf{p}' f^*(\mu, \epsilon, \mathbf{p}, \lambda) \times f(\mu', \epsilon', \mathbf{p}', \lambda')k(\mu, \epsilon, \mathbf{p}, \lambda \,|\, \mu', \epsilon', \mathbf{p}', \lambda') \geq 0 \tag{3.24}$$

for any function $f(\mu, \epsilon, \mathbf{p}, \lambda)$. The equality holds only if $f(\mu, \epsilon, \mathbf{p}, \lambda) = 0$.

Proof: Equation (3.23) follows from the definition of $k(\mu, \epsilon, \mathbf{p}, \lambda \mid \mu', \epsilon', \mathbf{p}', \lambda')$ and the properties of the inner products of the kets $|\mu, \epsilon, \mathbf{p}, \lambda)$. To prove (3.24) we construct the state

$$|\Phi) = \sum_{\epsilon} \int d\mu \sum_{\lambda} \int dp\, f(\mu, \epsilon, \mathbf{p}, \lambda)\, |\mu, \epsilon, \mathbf{p}, \lambda).$$

The requirement that $(\Phi \mid \Phi) \geq 0$ in Hilbert space where the equality holds only if $|\Phi) = 0$ leads to (3.24). Actually one should put some conditions of quadratic integrability on the functions f, but with these conditions the lemma holds.

Lemma II:

$$k(\mu, \epsilon, \mathbf{p}, \lambda \mid \mu', \epsilon', \mathbf{p}', \lambda') = \delta_{\epsilon,\epsilon'}\delta(\mu - \mu')\delta(\mathbf{p} - \mathbf{p}') \times k(\mu, \epsilon, \mathbf{p}{:}\lambda \mid \lambda'), \quad (3.25)$$

where $k(\mu, \epsilon, \mathbf{p}{:}\, \lambda \mid \lambda')$ is a positive definite kernel in terms of the variables λ, λ' which in general depends on μ, ϵ, and $\mathbf{p}$.

Proof: The presence of the δ functions on the right-hand side of (3.25) comes about in the usual fashion because ϵ, μ, and $\mathbf{p}$ are eigenvalues of operators. For example, following the usual proof,

$$(\mu, \epsilon, \mathbf{p}, \lambda|\, M\, |\mu', \epsilon', \mathbf{p}', \lambda') = \mu(\mu, \epsilon, \mathbf{p}, \lambda \mid \mu', \epsilon', \mathbf{p}', \lambda') = \mu'(\mu, \epsilon, \mathbf{p}, \lambda \mid \mu', \epsilon', \mathbf{p}', \lambda').$$

From this equation we obtain

$$(\mu - \mu')k(\mu, \epsilon, \mathbf{p}, \lambda \mid \mu', \epsilon', \mathbf{p}', \lambda') = 0.$$

The solution of this equation is

$$k(\mu, \epsilon, \mathbf{p}, \lambda \mid \mu', \epsilon', \mathbf{p}', \lambda') = \delta(\mu - \mu')A(\mu{:}\, \epsilon, \mathbf{p}, \lambda \mid \epsilon', \mathbf{p}', \lambda'),$$

where A is in general a function of μ and is a positive definite kernel in the remaining variables ϵ, $\mathbf{p}$, λ as follows from the positive definiteness of

$$k(\mu, \epsilon, \mathbf{p}, \lambda \mid \mu', \epsilon', \mathbf{p}', \lambda').$$

One continues the proof in an obvious manner. In (3.25) the δ functions involving $\mathbf{p}$ are Dirac deltas because the variables $\mathbf{p}$ belong to the continuous spectrum, since we are assuming that the Hilbert space is separable and hence a nondenumerable set of eigenkets must belong to the continuous spectrum.

The next few lemmas are proved using the requirement that the infinitesimal generators are Hermitian. Then if A is any of the infinitesimal generators of any Hermitian operator which is constructed from them, we must have

$$(\mu, \epsilon, \mathbf{p}, \lambda|\, A\, |\mu', \epsilon', \mathbf{p}', \lambda') = [(\mu', \epsilon', \mathbf{p}', \lambda'|\, A\, |\mu, \epsilon, \mathbf{p}, \lambda)]^*. \quad (3.26)$$

Lemma III: The spin operators S_i of Eq. (3.7) are Hermitian.

Proof: Before proceeding to the proof we wish to note that this lemma does *not* state that the kernels $S_i(\lambda \mid \lambda')$ are Hermitian in the variables λ. The lemma deals with the notion of Hermiticity in the complete Hilbert space as given by (3.26). To prove Lemma III it is useful to introduce the four Hermitian operators which play such an important role, especially in the global treatment. They are

$$w_0 = (\mathbf{P} \cdot \mathbf{J}),$$
$$w_i = [HJ_i + (\mathbf{P} \times \mathbf{J})_i] \quad (i = 1, 2, 3). \quad (3.27)$$

In terms of our basis they act in the following way:

$$w_0\, |\mu, \epsilon, \mathbf{p}, \lambda) = (\mathbf{p} \cdot \mathbf{S})\, |\mu, \epsilon, \mathbf{p}, \lambda),$$
$$w_i\, |\mu, \epsilon, \mathbf{p}, \lambda) = -\epsilon\left[\frac{\mathbf{p}(\mathbf{p} \cdot \mathbf{S})}{\omega(\mu, p) + \mu} - \mu\mathbf{S}\right] |\mu, \epsilon, \mathbf{p}, \lambda). \quad (3.28)$$

The lemma follows immediately upon using w_0 and w_i in (3.26), using the fact that the operators P_i, M, $(H/|H|)$ commute with the operators S.

Lemma IV: The operators L_i are Hermitian and the operator $\exp(i\boldsymbol{\xi} \cdot \mathbf{L})$ is thus unitary.

Proof: This lemma follows from the Hermiticity of J_i, (3.10a) and the previous lemma.

Definition: Let $\boldsymbol{\xi} = (\xi_1, \xi_2, \xi_3)$ define a set of three real numbers. We define the rotation matrix $R(\boldsymbol{\xi})$ to be the rotation matrix associated with a rotation of magnitude $\xi = |\boldsymbol{\xi}|$ about an axis in the direction given by $\boldsymbol{\xi}$. Specifically the matrix elements of $R(\boldsymbol{\xi})$ are given by

$$R_{ij}(\boldsymbol{\xi}) = \delta_{ij} \cos\xi - \frac{\xi_i\xi_j}{\xi^2}(\cos\xi - 1) + \sum_k \epsilon_{ijk}\xi_k \frac{\sin\xi}{\xi}. \quad (3.29)$$

Lemma V:

$$\exp(i\boldsymbol{\xi} \cdot \mathbf{L})\, |\mu, \epsilon, \mathbf{p}, \lambda) = |\mu, \epsilon, \mathbf{p}', \lambda), \quad (3.30)$$

where

$$\mathbf{p}' = R(\boldsymbol{\xi})\mathbf{p}. \quad (3.30a)$$

In Eq. (3.30a) we use the obvious notation of a matrix acting on a vector to form another vector.

Proof: From the commutation relations

$$[L_i, P_j] = i\sum_k \epsilon_{ijk}P_k, \quad (3.31)$$

which follows from the fifth of Eqs. (1.1) and Eq. (3.10a) and from Eq. (2.1) we obtain

$$\exp(-i\boldsymbol{\xi} \cdot \mathbf{L})P_i \exp(i\boldsymbol{\xi} \cdot \mathbf{L}) = \sum_j R_{ij}(\boldsymbol{\xi})P_j. \quad (3.32)$$

It follows that $\exp[i\boldsymbol{\xi}\cdot\mathbf{L}]\,|\mu,\epsilon,\mathbf{p},\lambda)$ is a simultaneous eigenstate of M, $(H/|H|)$, and P_i with eigenvalues μ, ϵ, and $\mathbf{p}_i'$, respectively, where $\mathbf{p}'$ is given in terms of $\mathbf{p}$ by (3.30a). From Theorem 2 and the definition (3.4) it follows that $\exp[i\boldsymbol{\xi}\cdot\mathbf{L}]\,|\mu,\epsilon,\mathbf{p},\lambda)$ is a linear combination in λ of the kets $|\mu,\epsilon,\mathbf{p}',\lambda)$, i.e.,

$$\exp(i\boldsymbol{\xi}\cdot\mathbf{L})\,|\mu,\epsilon,\mathbf{p},\lambda) = \sum_{\lambda'} C(\xi,\boldsymbol{\xi}/\xi,\mu,\epsilon,\mathbf{p}{:}\lambda,\lambda') \times |\mu,\epsilon,\mathbf{p}',\lambda'). \quad (3.32a)$$

We obtain C in the following way. Let us differentiate Eq. (3.32a) with respect to ξ. One obtains the differential equation

$$\frac{\boldsymbol{\xi}\cdot\mathbf{L}}{\xi}\sum_{\lambda'} C\left(\xi,\frac{\boldsymbol{\xi}}{\xi},\mu,\epsilon,\mathbf{p}{:}\lambda,\lambda'\right)|\mu,\epsilon,\mathbf{p}',\lambda') = \sum_{\lambda'}\left[\frac{\partial}{\partial\xi}C\left(\xi,\frac{\boldsymbol{\xi}}{\xi},\mu,\epsilon,\mathbf{p}{:}\lambda,\lambda'\right)\right]|\mu,\epsilon,\mathbf{p}',\lambda') + \sum_{\lambda'} C\left(\frac{\boldsymbol{\xi}}{\xi},\mu,\epsilon,p{:}\lambda'\right)\frac{\partial}{\partial\xi}|\mu,\epsilon,\mathbf{p}',\lambda'). \quad (3.33)$$

We now use (3.10), (3.29), (3.30a), and the linear independence of the kets $|\mu,\epsilon,\mathbf{p},\lambda)$ to show that

$$(\partial/\partial\xi)C(\xi,\boldsymbol{\xi}/\xi,\mu,\epsilon,\mathbf{p}{:}\lambda,\lambda') = 0,$$

from which it follows that C is independent of ξ. But when $\xi = 0$ we must have $\exp[i\boldsymbol{\xi}\cdot\mathbf{L}] = I$, where I is the identity operator. Thus $C(\xi,\boldsymbol{\xi}/\xi,\mu,\epsilon,\mathbf{p}{:}\,\lambda,\lambda') = \delta_{\lambda,\lambda'}$.

Lemma VI: The kernel $k(\mu,\epsilon,\mathbf{p}{:}\,\lambda\,|\,\lambda')$ of Eq. (3.25) is independent of the direction of $\mathbf{p}$.

This lemma leads to the definition of the kernel $m(\mu,\epsilon,p{:}\,\lambda\,|\,\lambda')$ to indicate the independence of the direction of $\mathbf{p}$:

Definition:

$$m(\mu,\epsilon,p{:}\,\lambda\,|\,\lambda) = k(\mu,\epsilon,\mathbf{p}{:}\,\lambda\,|\,\lambda'). \quad (3.34)$$

Proof of Lemma VI: From Lemma V

$$(\mu,\epsilon,\mathbf{p},\lambda|\exp(i\boldsymbol{\xi}\cdot\mathbf{L})|\,\mu',\epsilon',\mathbf{p}',\lambda') = (\mu,\epsilon,\mathbf{p},\lambda\,|\,\mu',\epsilon',R(\boldsymbol{\xi})\mathbf{p}',\lambda') = (\mu,\epsilon,R(-\boldsymbol{\xi})\mathbf{p},\lambda\,|\,\mu',\epsilon',\mathbf{p}',\lambda'). \quad (3.35)$$

From (3.35) using the form for the inner product of the ket vectors we obtain

$$\delta_{\epsilon,\epsilon'}\delta(\mu-\mu')\delta[\mathbf{p}-R(\boldsymbol{\xi})\mathbf{p}']k(\mu,\epsilon,\mathbf{p}{:}\lambda\,|\,\lambda') = \delta_{\epsilon,\epsilon'}\delta(\mu-\mu')\delta[R(-\boldsymbol{\xi})\mathbf{p}-\mathbf{p}'] \times k(\mu,\epsilon,R(-\boldsymbol{\xi})\mathbf{p}{:}\lambda\,|\,\lambda'). \quad (3.36)$$

But a well-known property of the three-dimensional Dirac δ function is that $\delta(R\mathbf{x}) = \delta(\mathbf{x})$, where $\mathbf{x}$ is any vector and R is any rotation matrix. Thus

$$\delta[\mathbf{p}-R(\boldsymbol{\xi})\mathbf{p}'] = \delta[R(-\boldsymbol{\xi})\mathbf{p}-\mathbf{p}'].$$

Hence, from (3.36) we must have for *any* $\boldsymbol{\xi}$

$$k(\mu,\epsilon,\mathbf{p}{:}\,\lambda\,|\,\lambda') = k(\mu,\epsilon,R(-\boldsymbol{\xi})\mathbf{p}:\lambda\,|\,\lambda'). \quad (3.37)$$

Let us pick $\boldsymbol{\xi}$ in such a way that $R(-\boldsymbol{\xi})\mathbf{p} = \hat{\mathbf{p}}$, where

$$\hat{\mathbf{p}} = (0,0,p).$$

Then $k(\mu,\epsilon,\mathbf{p}{:}\lambda\,|\,\lambda') = k(\mu,\epsilon,\hat{\mathbf{p}}{:}\lambda\,|\,\lambda')$. The lemma follows from the fact that $\hat{\mathbf{p}}$ is independent of the direction of $\mathbf{p}$.

Definition: Let us define the operators N_i by

$$N_i\,|\mu,\epsilon,\mathbf{p},\lambda) = -i\epsilon\omega(\mu,p)(\partial/\partial p_i)\,|\mu,\epsilon,\mathbf{p},\lambda). \quad (3.38)$$

Lemma VII: The operators N_i are Hermitian

Proof: The lemma follows from Lemma III, Eq. (3.15) and the definition (3.38).

We can now finish the proof of Theorem 6.

Proof of Theorem 6: From (3.38), (3.25), and (3.34)

$$(\mu,\epsilon,\mathbf{p},\lambda|\,N_3\,|\mu',\epsilon',\mathbf{p}',\lambda') = -i_{\epsilon'}\delta_{\epsilon,\epsilon'}\delta(\mu-\mu')m(\mu,\epsilon,p{:}\lambda')\,\delta(p_1-p_1') \times \delta(p_2-p_2')\omega(\mu,p')\frac{\partial}{\partial p_3'}\delta(p_3-p_3'). \quad (3.39)$$

On using (3.40) and (3.6) with A replaced by N_i we obtain

$$m(\mu,\epsilon,p{:}\lambda\,|\,\lambda')\omega(\mu,p')\delta'(p_3-p_3') = m(\mu,\epsilon,p'{:}\lambda\,|\,\lambda')\omega(\mu,p)\delta'(p_3-p_3'), \quad (3.40)$$

where $\delta'(x)$ is the derivative of the δ function and $p' = [p_1^2+p_2^2+p_3'^2]^{\frac{1}{2}}$. Let $f(p)$ be any function of p. Then from the properties of the derivative of the δ function we have

$$f(p')\delta'(p_3-p_3') = f(p)\delta'(p_3-p_3') \times [df(p)/dp](p_3/p)\delta(p_3-p_3'). \quad (3.41)$$

On using (3.41) in (3.40) we obtain a differential equation for $m(\mu,\epsilon,p:\lambda\,|\,\lambda')$, namely

$$(d/dp)m(\mu,\epsilon,p:\lambda\,|\,\lambda') = \{p/[\omega(\mu,p)]^2\} \times m(\mu,\epsilon,p:\lambda\,|\,\lambda') \quad (3.42)$$

for which the general solution is

$$m(\mu,\epsilon,p:\lambda\,|\,\lambda') = \omega(\mu,p)k(\lambda\,|\,\lambda'), \quad (3.43)$$

in which $k(\lambda\,|\,\lambda')$ is a constant of integration which, in general, depends on μ and ϵ. From Lemma I $k(\lambda\,|\,\lambda')$ must be a positive-definite kernel in λ.

Definition: If $k(\lambda\,|\,\lambda') = \delta_{\lambda,\lambda'}$ we write

$$|\mu,\epsilon,\mathbf{p},\lambda) = |\mu,\epsilon,\mathbf{p},\lambda\rangle.$$

Theorem 7: The kernels $S_i(\lambda\,|\,\lambda')$ of Eqs. (3.6) and (3.7) are Hermitian in λ when we use the basis $|\mu,\epsilon,\mathbf{p},\lambda\rangle$.

The proof of this theorem is so simple that we omit it. The next part of the present section is used to prove the following:

Theorem 8: The variable λ can always be chosen so that $k(\lambda \mid \lambda') = \delta_{\lambda,\lambda'}$.

Proof: We prove Theorem 8 by showing that if $k(\lambda \mid \lambda') \neq \delta_{\lambda,\lambda'}$ another degeneracy variable can be found so that in this new variable the condition of Theorem 8 is satisfied.

Since $k(\lambda \mid \lambda')$ is positive definite, it has a positive spectrum. That is, there exists a complete set of eigenvectors with components $u_\lambda(\beta)$ such that

$$\sum_{\lambda'} k(\lambda \mid \lambda') u_{\lambda'}(\beta) = \beta u_\lambda(\beta), \tag{3.44}$$

where $\beta > 0$. For simplicity, we assume that the spectrum of the operator, whose kernel is $k(\lambda \mid \lambda')$, is discrete and nondegenerate. More general cases are treated in a similar manner.

The eigenvectors satisfy the following completeness and orthogonality relations:

$$\sum_{\lambda} u^*_\lambda(\beta) u_\lambda(\beta') = \delta_{\beta,\beta'},$$
$$\sum_{\beta} u_\lambda(\beta) u^*_{\lambda'}(\beta) = \delta_{\lambda,\lambda'}. \tag{3.45}$$

Let us define

$$|\mu, \epsilon, 0, \beta\rangle = (\beta)^{-\frac{1}{2}} \sum_{\lambda} |\mu, \epsilon, 0, \lambda) u_\lambda(\beta). \tag{3.46}$$

Then for the kets $|\mu, \epsilon, \mathbf{p}, \beta\rangle$ constructed from $|\mu, \epsilon, 0, \beta\rangle$ through the use of (3.4), the kernel $k(\beta \mid \beta') = \delta_{\beta,\beta'}$ as can be proved easily.

The way that the infinitesimal generators act on the representer $F(\mu, \epsilon, \mathbf{p}, \lambda)$ can be derived from the way that these generators act on the kets in the usual manner.

The irreducible representations of the inhomogeneous Lorentz group corresponding to finite mass can be derived in identical manner by adding more necessary conditions. It is clear that we have used necessary conditions independent of the ζ representation to obtain the basis above. If the representation is to be irreducible, the mass operator M and the operator $(H/|H|)$ must be scalars by Schur's Lemma, since they commute with all of the infinitesimal generators. Thus μ and ϵ can have only one value. Furthermore, it is also seen that a necessary condition for irreducibility is that the kernels of the spin operators $S_i(\lambda \mid \lambda')$ form an irreducible set of generators of the rotation group. It is also easily shown that the representation which is obtained is indeed an irreducible representation of the inhomogeneous Lorentz group completely characterized by μ, ϵ and the irreducible set of generators of the rotation group, i.e., by S^2.

4. BASIS VECTORS FOR ZERO MASS

We now obtain the basis vectors for zero mass and show how the infinitesimal generators act in this basis. As mentioned in the Introduction it is only when the eigenvalue $\mu = 0$ is a point eigenvalue that we must consider the zero-mass representations. For if $\mu = 0$ is in the continuous spectrum of M, the contribution of the zero-mass representations to the inner product (1.4) is of zero measure.

We call "unphysical" those representations in which the operators P_i and H are identically zero, since no particles correspond to such representations. Classically this statement is equivalent to stating that, in contrast to the situation for particles of nonzero mass, massless particles cannot be brought to rest by a suitable change of frame of reference. It might be also mentioned that while these unphysical representations have no direct particle analog, they are still of interest mathematically, since they are representations of the homogeneous Lorentz group. We assume, however, that the reducible representations of interest to us do not contain unphysical representations.

Theorem 1: If the simultaneous spectrum of the operators P_i contains the point $\mathbf{p} = 0$, then the representation is an unphysical representation.

Proof: Let $|0)$ be a ket corresponding to the simultaneous eigenvalue zero of the operators P_i. Then

$$P_i |0) = 0 \quad (i = 1, 2, 3) \tag{4.1}$$

and since the mass is zero

$$H |0) = 0. \tag{4.2}$$

Furthermore, from the commutation rules (1.1), if A is any of the infinitesimal generators, then

$$P_i A |0) = 0, \quad HA |0) = 0.$$

Thus all of the vectors $|0)$ span an invariant space characterized by the requirement that $P_i = 0$ in this space. By our definition, this space is unphysical.

Theorem 1 is also stated in the following way.

Theorem 1a: In a physical representation the point $\mathbf{p} = 0$ is not in the simultaneous spectrum of the operators P_i.

We now come to Theorem 2.

Theorem 2: Let $\hat{\mathbf{p}}$ be the point

$$\hat{\mathbf{p}} = (0, 0, 1). \tag{4.3}$$

Then the point $\hat{\mathbf{p}}$ is in the simultaneous spectrum of the operators P_i.

Proof: Let $\mathbf{p}$ be a point in the simultaneous spectrum of the operators P_i. From Theorem 1a $\mathbf{p} \neq 0$. Let us define $|\Phi)$ as being the corresponding eigenket. This eigenket is also an eigenket of the operators

M and $(H/|H|)$ corresponding to the eigenvalues 0 and ϵ, respectively. Let us also define v and $\boldsymbol{\omega}$ by (1.24a) and (1.24b) where the $\mathbf{p}$ of these equations is replaced by the $\mathbf{p}$ which is in the spectrum of the operators P_i. Then $\exp[-iv\mathfrak{J}_3]\exp[-i\boldsymbol{\omega}\cdot\mathbf{J}\,|\,\Phi)$ is an eigenstate of P_i with simultaneous eigenvalues given by $\hat{\mathbf{p}}$. The proof makes use of the integrated form of the commutation rules of Sec. 2 in much the same way as in Theorem 1 of Sec. 3.

Let us denote the eigenkets belonging to the eigenvalue 0 of M, ϵ of $(H/|H|)$, and $\mathbf{p}=\hat{\mathbf{p}}$ of $\mathbf{P}$ by $|0, \epsilon, \hat{\mathbf{p}}, \lambda)$. The variable λ is a degeneracy variable and is to be chosen in such a manner that as λ runs through its values, we obtain all the linearly independent kets belonging to these eigenvalues of the operators. The function $f(\zeta;\epsilon,\lambda)$ is just $\langle\zeta\,|\,0,\epsilon,p,\lambda)$.

Definition: Let us define $|\mu,\epsilon,p,\lambda)$ by

$$|0,\epsilon,\mathbf{p},\lambda) = \exp[i\boldsymbol{\omega}\cdot\mathbf{J}]\exp[iv\mathfrak{J}_3]\,|0,\epsilon,\hat{\mathbf{p}},\lambda). \quad (4.4)$$

Theorem 3: The ket $|0,\epsilon,\mathbf{p},\lambda)$ is a simultaneous eigenket of the operators M, $(H/|H|)$, $\mathbf{P}$ with the eigenvalues 0, ϵ, $\mathbf{p}$ respectively. Any other simultaneous eigenket of these operators with the same eigenvalues is a linear combination in λ of the eigenkets $|0,\epsilon,\mathbf{p},\lambda)$.

Proof: The proof of this theorem is analogous to the proofs of similar theorems of part 3.

We now show how the infinitesimal generators of the group act in this basis. From the previous theorem

$$\begin{aligned} P_i\,|0,\epsilon,\mathbf{p},\lambda) &= p_i\,|0,\epsilon,\mathbf{p},\lambda),\\ H\,|0,\epsilon,\mathbf{p},\lambda) &= \epsilon p\,|0,\epsilon,\mathbf{p},\lambda), \end{aligned} \quad (4.5)$$

and also

$$\begin{aligned} M\,|0,\epsilon,\mathbf{p},\lambda) &= 0,\\ (H/|H|)\,|0,\epsilon,\mathbf{p},\lambda) &= \epsilon\,|0,\epsilon,\mathbf{p},\lambda). \end{aligned} \quad (4.6)$$

To obtain the remaining generators of the inhomogeneous Lorentz group in this basis we introduce the operators B_1, B_2, C.

Definition:

$$B_1 = -(H/|H|)\mathfrak{J}_1 - J_2, \quad B_2 = -(H/|H|)\mathfrak{J}_2 + J_1, \quad C = J_3. \quad (4.7)$$

Theorem 4: B_1, B_2, and C satisfy the commutation rules for the infinitesimal generators of the Euclidean group in the plane; i.e.,

$$[B_1,B_2]=0,\quad [B_1,C]=-iB_2,\quad [B_2,C]=iB_1. \quad (4.8)$$

Proof: The proof is immediately obtained from the definition (4.7) and the commutation rules (1.1).

Lemma: The kets $B_1\,|0,\epsilon,\hat{\mathbf{p}},\lambda)$, $B_2\,|0,\epsilon,\hat{\mathbf{p}},\lambda)$, $C\,|0,\epsilon,\hat{\mathbf{p}},\lambda)$ are all eigenkets of the operators M, $(H/|H|)$, $\mathbf{P}$ with eigenvalues 0, ϵ, $\hat{\mathbf{p}}$, respectively.

Proof: The proof of this lemma is similar to the proof of the corresponding lemma of the preceding section. This lemma allows us to make the following definitions:

Definition: Let us define the kernels $T_1(\lambda\,|\,\lambda')$, $T_2(\lambda\,|\,\lambda')$, $M(\lambda\,|\,\lambda')$ by

$$\begin{aligned} B_2\,|0,\epsilon,\hat{\mathbf{p}},\lambda) &= \sum_{\lambda'} T_1(\lambda'\,|\,\lambda)\,|0,\epsilon,\hat{\mathbf{p}},\lambda'),\\ B_2\,|0,\epsilon,\hat{\mathbf{p}},\lambda) &= \sum_{\lambda'} T_2(\lambda'\,|\,\lambda)\,|0,\epsilon,\hat{\mathbf{p}},\lambda'),\\ C\,|0,\epsilon,\hat{\mathbf{p}},\lambda) &= \sum_{\lambda'} M(\lambda'\,|\,\lambda)\,|0,\epsilon,\hat{\mathbf{p}},\lambda'). \end{aligned} \quad (4.9)$$

Definition: Let us define the operators T_1, T_2, and M by

$$\begin{aligned} T_1\,|0,\epsilon,\mathbf{p},\lambda) &= \sum_{\lambda'} T_1(\lambda'\,|\,\lambda)\,|0,\epsilon,\mathbf{p},\lambda'),\\ T_2\,|0,\epsilon,\mathbf{p},\lambda) &= \sum_{\lambda'} T_2(\lambda'\,|\,\lambda)\,|0,\epsilon,\mathbf{p},\lambda'),\\ M\,|0,\epsilon,\mathbf{p},\lambda) &= \sum_{\lambda'} M(\lambda'\,|\,\lambda)\,|0,\epsilon,\mathbf{p},\lambda'). \end{aligned} \quad (4.10)$$

[The operator M of Eq. (4.10) is not to be confused with the mass operator M.]

Theorem 5: The operators T_1, T_2, and M satisfy the commutation rules for the infinitesimal generators of the Euclidean group in the plane, i.e., Eqs. (4.8).

Proof: The proof follows from the definition given by Eqs. (4.9) and (4.10) and the commutation relations for B_1, B_2, and C given by Eqs. (4.8).

Theorem 6: The infinitesimal generators J_i are given by the following equations:

$$\begin{aligned} J_1\,|0,\epsilon,\mathbf{p},\lambda) &= L_1\,|0,\epsilon,\mathbf{p},\lambda) + \frac{p_1}{p+p_3}M\,|0,\epsilon,\mathbf{p},\lambda),\\ J_2\,|0,\epsilon,\mathbf{p},\lambda) &= L_2\,|0,\epsilon,\mathbf{p},\lambda) + \frac{p_2}{p+p_3}M\,|0,\epsilon,\mathbf{p},\lambda),\\ J_3\,|0,\epsilon,\mathbf{p},\lambda) &= L_3\,|0,\epsilon,\mathbf{p},\lambda) + M\,|0,\epsilon,\mathbf{p},\lambda), \end{aligned} \quad (4.11)$$

where the operators L_i are given by Eq. (3.10).

Proof: We first prove the last of Eqs. (4.11)

$$\begin{aligned} J_3\,|0,\epsilon,\mathbf{p},\lambda) &= \left\{-i\frac{\partial}{\partial\alpha}\exp[i\alpha J_3]\,|0,\epsilon,\mathbf{p},\lambda)\right\}_{\alpha=0}\\ &= \left\{-i\frac{\partial}{\partial\alpha}\exp[i\alpha J_3]\exp[i\boldsymbol{\omega}\cdot\mathbf{J}]\right.\\ &\qquad \left.\times\exp[iv\mathfrak{J}_3]\,|0,\epsilon,\hat{\mathbf{p}},\lambda)\right\}_{\alpha=0}. \end{aligned} \quad (4.12)$$

Now from Sec. 2 and the fact that J_3 commutes with $\mathfrak{J}_3$

$$\begin{aligned} \exp[i\alpha J_3]\exp[i\boldsymbol{\omega}\cdot\mathbf{J}]\exp[iv\mathfrak{J}_3] &= \exp[i\alpha J_3]\\ &\quad\times\exp[i\boldsymbol{\omega}\cdot\mathbf{J}]\exp[-i\alpha J_3]\exp[i\alpha J_3]\exp[iv\mathfrak{J}_3]\\ &= \exp[i\boldsymbol{\omega}'\cdot\mathbf{J}]\exp[iv\mathfrak{J}_3]\exp[i\alpha J_3], \end{aligned} \quad (4.13)$$

where the vector $\boldsymbol{\omega}'$ is given by

$$\omega_1' = \omega_1 \cos\alpha + \omega_2 \sin\alpha,$$
$$\omega_2' = \omega_2 \cos\alpha - \omega_1 \sin\alpha,$$
$$\omega_3' = \omega_3. \tag{4.13a}$$

Then from (4.12), (4.13), (4.9), (4.10), and (4.4) we have

$$J_3 |0, \epsilon, \mathbf{p}, \lambda) = \{-i(\partial/\partial\alpha) |0, \epsilon, \mathbf{p}', \lambda)\}_{\alpha=0} + M |0, \epsilon, \mathbf{p}, \lambda), \tag{4.14}$$

where $\mathbf{p}'$ is related to $\boldsymbol{\omega}'$ precisely as $\mathbf{p}$ is related to ω, i.e.,

$$p_1' = -p \sin\omega(\omega_2'/\omega),$$
$$p_2' = p \sin\omega(\omega_1/\omega),$$
$$p_3' = p \cos\omega = p_3. \tag{4.14a}$$

On taking the derivative with respect to α which is indicated in (4.14), the third of Eqs. (4.11) is obtained.

We now derive the first two of Eqs. (4.11). We note that

$$(\boldsymbol{\omega} \cdot \mathbf{J}) |0, \epsilon, \mathbf{p}, \lambda) = \{-i(\partial/\partial\alpha) \exp[i\alpha(\boldsymbol{\omega} \cdot \mathbf{J})] \times \exp[iv\mathfrak{J}_3] |0, \epsilon, \mathbf{p}, \lambda)\}_{\alpha=1}$$
$$= \{-i(\partial/\partial\alpha) |0, \epsilon, \mathbf{p}', \lambda)\}_{\alpha=1}, \tag{4.15}$$

where

$$p_1' = -p \sin\alpha\omega(\omega_2/\omega),$$
$$p_2' = p \sin\alpha\omega(\omega_1/\omega),$$
$$p_3' = p \cos\alpha\omega. \tag{4.15a}$$

On expressing $\boldsymbol{\omega}$ in terms of $\mathbf{p}$ we thus obtain from (4.15)

$$(\mathbf{p} \times \mathbf{J})_3 |0, \epsilon, \mathbf{p}, \lambda) = i[-p^2(\partial/\partial p_3) + p_3(\mathbf{p} \cdot \nabla)] |0, \epsilon, \mathbf{p}, \lambda). \tag{4.16}$$

Now, also,

$$(\boldsymbol{\omega} \times \mathbf{J})_3 |0, \epsilon, \mathbf{p}, \lambda) = \exp[i\boldsymbol{\omega} \cdot \mathbf{J}] \exp[-i\boldsymbol{\omega} \cdot \mathbf{J}] \times (\boldsymbol{\omega} \times \mathbf{J})_3 \exp[i\boldsymbol{\omega} \cdot \mathbf{J}] \exp[iv\mathfrak{J}_3] |0, \epsilon, \hat{\mathbf{p}}, \lambda). \tag{4.17}$$

But from the formulas of Sec. 2

$$\exp[i\boldsymbol{\omega} \cdot \mathbf{J}](\boldsymbol{\omega} \times \mathbf{J})_3 \exp[-i\boldsymbol{\omega} \cdot \mathbf{J}] = (\boldsymbol{\omega} \times \mathbf{J})_3 \cos\omega - \omega J_3 \sin\omega,$$
$$\exp[-i\boldsymbol{\omega} \cdot \mathbf{J}](\boldsymbol{\omega} \times \mathbf{J})_3 \exp[i\boldsymbol{\omega} \cdot \mathbf{J}] = (\boldsymbol{\omega} \times \mathbf{J})_3 \cos\omega + \omega J_3 \sin\omega. \tag{4.18}$$

Thus

$$(\boldsymbol{\omega} \times \mathbf{J})_3 |0, \epsilon, \mathbf{p}, \lambda) = \cos\omega \exp[i\boldsymbol{\omega} \cdot \mathbf{J}](\boldsymbol{\omega} \times \mathbf{J})_3 \times \exp[iv\mathfrak{J}_3] |0, \epsilon, \hat{\mathbf{p}}, \lambda) + \omega \sin\omega M |0, \epsilon, \mathbf{p}, \lambda)$$
$$= \cos^2\omega(\boldsymbol{\omega} \times \mathbf{J})_3 |0, \epsilon, \mathbf{p}, \lambda) - \omega \cos\omega \sin\omega J_3 \times |0, \epsilon, \mathbf{p}, \lambda) + \omega \sin\omega M |0, \epsilon, \mathbf{p}, \lambda). \tag{4.19}$$

Equation (4.19) is an equation for $(\boldsymbol{\omega} \times \mathbf{J})_3 |0, \epsilon, \mathbf{p}, \lambda)$ which can be easily solved. On replacing $\boldsymbol{\omega}$ by its expression in terms of $\mathbf{p}$ we obtain the following equation, on using our known expression for $J_3 |0, \epsilon, \mathbf{p}, \lambda)$,

$$(p_1 J_1 + p_2 J_2) |0, \epsilon, \mathbf{p}, \lambda) = [-ip_3(\mathbf{p} \times \nabla)_3 + (p - p_3)M |0, \epsilon, \mathbf{p}, \lambda). \tag{4.20}$$

We now consider (4.16) and (4.20) as a pair of simultaneous equations for the unknowns $J_1 |0, \epsilon, \mathbf{p}, \lambda)$ and $J_2 |0, \epsilon, \mathbf{p}, \lambda)$. On solving for these unknowns, we obtain the first two of Eqs. (4.11). Thus we have completed the derivation of the way that the operators corresponding to the components of the angular momentum act in the basis.

However, we now have to make some restrictions on the operator M which is obtained from the representations of the generators of the Euclidean group in two dimensions.

It is easy to show from (4.11) that

$$(\mathbf{P} \cdot \mathbf{J}) |0, \epsilon, \mathbf{p}, \lambda) = M |0, \epsilon, \mathbf{p}, \lambda). \tag{4.21}$$

It then follows that the operator M must be Hermitian, since the operators P_i and J_i are Hermitian.

We use the following definition:

Definition: If the operator M has any eigenvalue which is not an integer or half-odd integer, the operator M is called improper. If M is not *improper*, it is called *proper*.

Definition: If an operator A has a set of eigenvectors which span the Hilbert space, then the operator is said to have a *spectral decomposition.*

This latter definition is a somewhat unrigorous form of the definition of spectral decomposition used by mathematicians. In most physical applications it is assumed that in general the Hermiticity of an operator assures one that the operator has a spectral decomposition. While this assumption is valid if the space is finite-dimensional or if the operator is bounded, it is not valid in general. We assume that M has a spectral decomposition.

In Ref. 9 the case has been studied in which M has only one eigenvalue. (This case corresponds to the irreducible representations of the inhomogeneous Lorentz group with finite spin.) A simple extension of the results of Ref. 9 leads to the following theorems:

Theorem 7: For any operator M that has a spectral decomposition, the operators J_i have a spectral decomposition.

Theorem 8: A necessary and sufficient condition that $\mathbf{J}^2$ has a spectral decomposition is that M be a proper operator.

[9] H. E. Moses and R. T. Prosser (to be published).

Theorem 9: A necessary and sufficient condition that all of the operators J_i have the same spectrum is that M be a proper operator.

Theorem 10: A necessary condition that the operators $\exp[i\mathbf{a}\cdot\mathbf{J}]$ form a ray representation of the rotation group, where $a = |\mathbf{a}|$ is the angle of rotation and the direction of $\mathbf{a}$ gives the direction of the axis of rotation, is that M be a proper operator.

The above theorems indicate that the only representations of physical interest are those for which M is a proper operator. Theorem 8 is particularly edifying, for it provides an example of a Hermitian operator which does not have a spectral decomposition when M is an improper operator.

We now show how the operators $\mathcal{J}_i$ act on the kets $|0, \epsilon, \mathbf{p}, \lambda)$. We note that

$$\exp[i\boldsymbol{\omega}\cdot\mathbf{J}]\,\mathcal{J}_3\exp[iv\mathcal{J}_3]\,|0,\epsilon,\hat{\mathbf{p}},\lambda) = -i(\partial/\partial v)\,|0,\epsilon,\mathbf{p},\lambda), \quad (4.22)$$

where $\mathbf{p}$ is related to v through (1.24). It follows immediately that

$$\exp[i\boldsymbol{\omega}\cdot\mathbf{J}]\,\mathcal{J}_3\exp[iv\mathcal{J}_3]\,|0,\epsilon,\hat{\mathbf{p}},\lambda) = -i\epsilon(\mathbf{p}\cdot\nabla)\,|0,\epsilon,\mathbf{p},\lambda). \quad (4.23)$$

But also from (2.7) and (1.24)

$$\begin{aligned}&\exp[i\boldsymbol{\omega}\cdot\mathbf{J}]\,\mathcal{J}_3\exp[iv\mathcal{J}_3]\,|0,\epsilon,\hat{\mathbf{p}},\lambda)\\ &\quad=\exp(i\boldsymbol{\omega}\cdot\mathbf{J})\,\mathcal{J}_3\exp(-i\boldsymbol{\omega}\cdot\mathbf{J})\,|0,\epsilon,\mathbf{p},\lambda)\\ &\quad=[\mathcal{J}_3\cos\omega+(\sin\omega/\omega)(\omega_1\mathcal{J}_2-\omega_2\mathcal{J}_1)]\,|0,\epsilon,\mathbf{p},\lambda)\\ &\quad=[(\mathbf{p}\cdot\mathcal{J})/p]\,|0,\epsilon,\mathbf{p},\lambda). \quad (4.24)\end{aligned}$$

Thus from (4.23) and (4.24)

$$(\mathbf{p}\cdot\mathcal{J})\,|0,\epsilon,\mathbf{p},\lambda) = -i\epsilon p(\mathbf{p}\cdot\nabla)\,|0,\epsilon,\mathbf{p},\lambda). \quad (4.25)$$

Now, also using the results of Sec. 2 (noting that $[(\boldsymbol{\omega}\cdot\mathbf{J}),(\boldsymbol{\omega}\cdot\mathcal{J})]=0$),

$$\begin{aligned}&(\boldsymbol{\omega}\cdot\mathcal{J})\,|0,\epsilon,\mathbf{p},\lambda)\\ &\quad=\exp[i\boldsymbol{\omega}\cdot\mathbf{J}](\boldsymbol{\omega}\cdot\mathcal{J})\exp[iv\mathcal{J}_3]\,|0,\epsilon,\hat{\mathbf{p}},\lambda)\\ &\quad=\cosh v\exp[i\boldsymbol{\omega}\cdot\mathbf{J}]\exp[iv\mathcal{J}_3](\boldsymbol{\omega}\cdot\mathcal{J})\,|0,\epsilon,\hat{\mathbf{p}},\lambda)\\ &\qquad-\sinh v\exp[i\boldsymbol{\omega}\cdot\mathbf{J}]\exp[iv\mathcal{J}_3](\boldsymbol{\omega}\times\mathbf{J})_3\,|0,\epsilon,\hat{\mathbf{p}},\lambda)\\ &\quad=\exp[i\boldsymbol{\omega}\cdot\mathbf{J}]\exp[iv\mathcal{J}_3][\omega_1(\mathcal{J}_1\cosh v-J_2\sinh v)\\ &\qquad+\omega_2(\mathcal{J}_2\cosh v+J_1\sinh v)]\,|0,\epsilon,\hat{\mathbf{p}},\lambda)\\ &\quad=\exp[i\boldsymbol{\omega}\cdot\mathbf{J}]\exp[iv\mathcal{J}_3][-\tfrac{1}{2}\epsilon\omega_1e^{-\epsilon v}B_1\,|0,\epsilon,\hat{\mathbf{p}},\lambda)\\ &\qquad-\tfrac{1}{2}\epsilon\omega_2e^{-\epsilon v}B_2\,|0,\epsilon,\hat{\mathbf{p}},\lambda)\\ &\qquad+\tfrac{1}{2}\omega_1e^{\epsilon v}(\mathcal{J}_1-\epsilon J_2)\,|0,\epsilon,\hat{\mathbf{p}},\lambda)\\ &\qquad+\tfrac{1}{2}\omega_2e^{\epsilon v}(\mathcal{J}_2+\epsilon J_1)\,|0,\epsilon,\hat{\mathbf{p}},\lambda)]\\ &\quad=-\tfrac{1}{2}\epsilon\omega_1e^{-\epsilon v}T_1\,|0,\epsilon,\mathbf{p},\lambda)-\tfrac{1}{2}\epsilon\omega_2e^{-\epsilon v}T_2\,|0,\epsilon,\mathbf{p},\lambda)\\ &\qquad+\tfrac{1}{2}\omega_1e^{\epsilon v}\exp[i\boldsymbol{\omega}\cdot\mathbf{J}][\mathcal{J}_1\cosh v+J_2\sinh v\\ &\qquad-\epsilon J_2\cosh v-\epsilon\mathcal{J}_1\sinh v]\exp[iv\mathcal{J}_3]\,|0,\epsilon,\hat{\mathbf{p}},\lambda)\\ &\qquad+\tfrac{1}{2}\omega_2e^{\epsilon v}\exp[i\boldsymbol{\omega}\cdot\mathbf{J}][\mathcal{J}_2\cosh v-J_1\sinh v\\ &\qquad+\epsilon J_1\cosh v-\epsilon\mathcal{J}_2\sinh v]\exp[iv\mathcal{J}_3]\,|0,\epsilon,\hat{\mathbf{p}},\lambda). \quad (4.26)\end{aligned}$$

Then

$$\begin{aligned}&(\boldsymbol{\omega}\cdot\mathcal{J})\,|0,\epsilon,\mathbf{p},\lambda)\\ &\quad=-\tfrac{1}{2}\epsilon\omega_1e^{-\epsilon v}T_1\,|0,\epsilon,\mathbf{p},\lambda)-\tfrac{1}{2}\epsilon\omega_2e^{-\epsilon v}T_2\,|0,\epsilon,\mathbf{p},\lambda)\\ &\qquad+\tfrac{1}{2}\exp[i\boldsymbol{\omega}\cdot\mathbf{J}](\boldsymbol{\omega}\cdot\mathcal{J})\exp[iv\mathcal{J}_3]\,|0,\epsilon,\hat{\mathbf{p}},\lambda)\\ &\qquad-\tfrac{1}{2}\epsilon\exp[i\boldsymbol{\omega}\cdot\mathbf{J}](\boldsymbol{\omega}\times\mathbf{J})_3\exp[iv\mathcal{J}_3]\,|0,\epsilon,\hat{\mathbf{p}},\lambda). \quad (4.27)\end{aligned}$$

Again using Eq. (2.7),

$$\begin{aligned}(\boldsymbol{\omega}\cdot\mathcal{J})\,|0,\epsilon,\mathbf{p},\lambda) &= \epsilon[-(\boldsymbol{\omega}\times\mathbf{J})_3\cos\omega+\omega J_3\sin\omega\\ &\quad-\omega_1e^{-\epsilon v}T_1-\omega_2e^{-\epsilon v}T_2]\,|0,\epsilon,\mathbf{p},\lambda). \quad (4.28)\end{aligned}$$

We rewrite (4.28) by expressing $\boldsymbol{\omega}$ and v in terms of $\mathbf{p}$. We also use our expressions for J_i. We have

$$\begin{aligned}(\mathbf{p}\times\mathcal{J})_3\,|0,\epsilon,\mathbf{p},\lambda) &= -\epsilon[ip(\mathbf{p}\times\nabla)_3+(p-p_3)M\\ &\quad+(p_1/p)T_2-(p_2/p)T_1]\,|0,\epsilon,\mathbf{p},\lambda). \quad (4.29)\end{aligned}$$

We want now to obtain $(\boldsymbol{\omega}\times\mathcal{J})_3\,|0,\epsilon,\mathbf{p},\lambda)$. From Sec. 2 we have

$$\exp(-i\boldsymbol{\omega}\cdot\mathbf{J})(\boldsymbol{\omega}\times\mathcal{J})_3\exp(i\boldsymbol{\omega}\cdot\mathbf{J}) = (\boldsymbol{\omega}\times\mathcal{J})_3\cos\omega+\mathcal{J}_3\omega\sin\omega, \quad (4.30)$$

$$\exp[-iv\mathcal{J}_3](\boldsymbol{\omega}\times\mathcal{J})_3\exp[iv\mathcal{J}_3] = (\boldsymbol{\omega}\times\mathcal{J})_3\cosh v+(\boldsymbol{\omega}\cdot\mathbf{J})\sinh v. \quad (4.31)$$

Then

$$\begin{aligned}&(\boldsymbol{\omega}\times\mathcal{J})_3\,|0,\epsilon,\mathbf{p},\lambda)\\ &\quad=\omega\sin\omega\exp[i\boldsymbol{\omega}\cdot\mathbf{J}]\mathcal{J}_3\exp[iv\mathcal{J}_3]\,|0,\epsilon,\hat{\mathbf{p}},\lambda)\\ &\qquad+\cos\omega\cosh v\exp[i\boldsymbol{\omega}\cdot\mathbf{J}]\\ &\qquad\times\exp[iv\mathcal{J}_3](\boldsymbol{\omega}\times\mathcal{J})_3\,|0,\epsilon,\hat{\mathbf{p}},\lambda)\\ &\qquad+\cos\omega\sinh v\exp[i\boldsymbol{\omega}\cdot\mathbf{J}]\\ &\qquad\times\exp[iv\mathcal{J}_3](\boldsymbol{\omega}\cdot\mathbf{J})\,|0,\epsilon,\hat{\mathbf{p}},\lambda)\\ &\quad=\omega\sin\omega\exp[i\boldsymbol{\omega}\cdot\mathbf{J}]\mathcal{J}_3\exp[iv\mathcal{J}_3]\,|0,\epsilon,\hat{\mathbf{p}},\lambda)\\ &\qquad-\epsilon\cos\omega(\tfrac{1}{2}e^{-\epsilon v})\exp[i\boldsymbol{\omega}\cdot\mathbf{J}]\\ &\qquad\times\exp[iv\mathcal{J}_3](\omega_1T_2-\omega_2T_1)\,|0,\epsilon,\hat{\mathbf{p}},\lambda)\\ &\qquad+\cos\omega(\tfrac{1}{2}e^{\epsilon v})\exp[i\boldsymbol{\omega}\cdot\mathbf{J}]\exp[iv\mathcal{J}_3]\\ &\qquad\times(\omega_1\mathcal{J}_2+\epsilon\omega_1J_1-\omega_2\mathcal{J}_1+\epsilon\omega_2J_2)\,|0,\epsilon,\hat{\mathbf{p}},\lambda)\\ &\quad=\omega\sin\omega\exp[i\boldsymbol{\omega}\cdot\mathbf{J}]\mathcal{J}_3\exp[iv\mathcal{J}_3]\,|0,\epsilon,\hat{\mathbf{p}},\lambda)\\ &\qquad-\epsilon\cos\omega(\tfrac{1}{2}e^{-\epsilon v})(\omega_1T_2-\omega_2T_1)\,|0,\epsilon,\mathbf{p},\lambda)\\ &\qquad+\cos\omega(\tfrac{1}{2}e^{\epsilon v})\exp[i\boldsymbol{\omega}\cdot\mathbf{J}]\\ &\qquad\times\exp[iv\mathcal{J}_3][(\boldsymbol{\omega}\times\mathcal{J})_3+\epsilon(\boldsymbol{\omega}\cdot\mathbf{J})]\,|0,\epsilon,\hat{\mathbf{p}},\lambda)\\ &\quad=-\epsilon\cos\omega(\tfrac{1}{2}e^{-\epsilon v})(\omega_1T_2-\omega_2T_1)\,|0,\epsilon,\mathbf{p},\lambda)\\ &\qquad+\omega\sin\omega\exp[i\boldsymbol{\omega}\cdot\mathbf{J}]\times\mathcal{J}_3\exp[iv\mathcal{J}_3]\,|0,\epsilon,\hat{\mathbf{p}},\lambda)\\ &\qquad+\cos\omega(\tfrac{1}{2}e^{\epsilon v})\exp[i\boldsymbol{\omega}\cdot\mathbf{J}][(\boldsymbol{\omega}\times\mathcal{J})_3\cosh v\\ &\qquad-(\boldsymbol{\omega}\cdot\mathbf{J})\sinh v+\epsilon(\boldsymbol{\omega}\cdot\mathbf{J})\cosh v\\ &\qquad-\epsilon(\boldsymbol{\omega}\times\mathcal{J})_3\sinh v]\exp[iv\mathcal{J}_3]\,|0,\epsilon,\hat{\mathbf{p}},\lambda). \quad (4.32)\end{aligned}$$

Upon some slight rearrangement Eq. (4.32) becomes

$$\begin{aligned}&(\boldsymbol{\omega}\times\mathcal{J})_3\,|0,\epsilon,\mathbf{p},\lambda)\\ &\quad=-\epsilon\cos\omega(\tfrac{1}{2}e^{-\epsilon v})(\omega_1T_2-\omega_2T_1)\,|0,\epsilon,\mathbf{p},\lambda)\\ &\qquad+\omega(\tfrac{1}{2}\sin\omega)\exp[i\boldsymbol{\omega}\cdot\mathbf{J}]\mathcal{J}_3\exp[iv\mathcal{J}_3]\,|0,\epsilon,\hat{\mathbf{p}},\lambda)\\ &\qquad+\tfrac{1}{2}\exp[i\boldsymbol{\omega}\cdot\mathbf{J}]\{\cos\omega[(\boldsymbol{\omega}\times\mathcal{J})_3+\epsilon(\boldsymbol{\omega}\cdot\mathbf{J})]\\ &\qquad+\omega\sin\omega\mathcal{J}_3\}\exp[iv\mathcal{J}_3]\,|0,\epsilon,\hat{\mathbf{p}},\lambda). \quad (4.33)\end{aligned}$$

On using (4.23) for the second term on the right and (4.30) for the third term on the right we obtain

$$(\boldsymbol{\omega} \times \mathfrak{J})_3 \,|0, \epsilon, \mathbf{p}, \lambda)$$
$$= \tfrac{1}{2}\epsilon \cos \omega[(\boldsymbol{\omega} \cdot \mathbf{J}) - e^{-\epsilon\nu}(\omega_1 T_2 - \omega_2 T_1)] \,|0, \epsilon, \mathbf{p}, \lambda)$$
$$- \tfrac{1}{2}\epsilon\omega \sin \omega i(\mathbf{p} \cdot \boldsymbol{\nabla}) \,|0, \epsilon, \mathbf{p}, \lambda). \quad (4.34)$$

We then express $\boldsymbol{\omega}$ and ν in terms of $\mathbf{p}$ to obtain

$$(p_1\mathfrak{J}_1 + p_2\mathfrak{J}_2) \,|0, \epsilon, \mathbf{p}, \lambda) = \epsilon[-ip(\mathbf{p} \cdot \boldsymbol{\nabla})$$
$$+ ip_3 p(\partial/\partial p_3) - (p_3/p^2)(p_1 T_1 + p_2 T_2) \,|0, \epsilon, \mathbf{p}, \lambda). \quad (4.35)$$

From Eqs. (4.25), (4.29), and (4.35) we may solve for $\mathfrak{J}_i \,|0, \epsilon, \mathbf{p}, \lambda)$. The results are given by Theorem 11.

Theorem 11: The operators $\mathfrak{J}_i$ are given by

$$\mathfrak{J}_1 \,|0, \epsilon, \mathbf{p}, \lambda)$$
$$= \epsilon\left\{-ip\frac{\partial}{\partial p_1} + \frac{p_2}{p + p_3} M + \left[\frac{p_1^2}{p^2(p + p_3)} - \frac{1}{p}\right] T_1 + \frac{p_1 p_2}{p^2(p + p_3)} T_2\right\} |0, \epsilon, \mathbf{p}, \lambda),$$

$$\mathfrak{J}_2 \,|0, \epsilon, \mathbf{p}, \lambda)$$
$$= \epsilon\left\{-ip\frac{\partial}{\partial p_2} - \frac{p_1}{p + p_3} M + \frac{p_1 p_2}{p^2(p + p_3)} T_1 + \left[\frac{p_2^2}{p^2(p + p_3)} - \frac{1}{p}\right] T_2\right\} |0, \epsilon, \mathbf{p}, \lambda),$$

$$\mathfrak{J}_3 \,|0, \epsilon, \mathbf{p}, \lambda)$$
$$= \epsilon\left\{-ip\frac{\partial}{\partial p_3} + \frac{1}{p^2}[p_1 T_1 + p_2 T_2]\right\} |0, \epsilon, \mathbf{p}, \lambda). \quad (4.36)$$

Thus we have finally shown how all the infinitesimal generators act on the kets $|0, \epsilon, \mathbf{p}, \lambda)$. The remaining theorems of this section are proved in a manner very similar to that for the corresponding theorems of the previous section.

Theorem 12:

$$(0, \epsilon, \mathbf{p}, \lambda \,|\, 0, \epsilon', \mathbf{p}', \lambda') = p\delta_{\epsilon,\epsilon'}\delta(\mathbf{p} - \mathbf{p}')k(\lambda \,|\, \lambda'), \quad (4.37)$$

where $k(\lambda \,|\, \lambda')$ is a positive definite kernel which may depend on ϵ but does not depend on $\mathbf{p}$.

Definition: If $k(\lambda \,|\, \lambda') = \delta_{\lambda,\lambda'}$, then we define $|0, \epsilon, \mathbf{p}, \lambda\rangle = |0, \epsilon, \mathbf{p}, \lambda)$.

Theorem 13: It is always possible to choose λ such that $k(\lambda \,|\, \lambda') = \delta_{\lambda,\lambda'}$. Furthermore, in the basis $|0, \epsilon, \mathbf{p}, \lambda\rangle$ the kernels $T_i(\lambda \,|\, \lambda')$ and $M(\lambda \,|\, \lambda')$ are Hermitian in the variable λ.

It is now easy to transcribe our results to the functions $F(0, \epsilon, \mathbf{p}, \lambda)$ of Sec. 1.

The irreducible representations for zero mass can obviously be obtained by adding the requirement that ϵ take on only one value and that the generators of the Euclidean group T_i, M form an irreducible set. In the special case that $T_1 = T_2 = 0$ and M is a scalar (an integer or a half-odd integer) we obtain the representations corresponding to finite spin. Indeed this is the procedure which was used to obtain the results of Ref. 8.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Equations with Connected Kernels for N-particle T Operators*

ROGER G. NEWTON
Indiana University, Bloomington, Indiana

(Received 2 September 1966)

The Lippmann–Schwinger procedure for single particle scattering is generalized for the calculation of n-particle T matrices describing collisions of a particle with a bound state of $(n-1)$ other particles. The relevant T operators are expressed in terms of resolvents of compact operators, the equations for which are uncoupled. The two-body interactions appear in the form of their T operators only.

I. INTRODUCTION

THE Hamiltonian of a scattering system is never a compact[1] operator. Consequently the complete Green's function, or resolvent, $\mathcal{G}(E) = (E - H)^{-1}$ cannot be constructed in a straight forward manner by the Fredholm method or by other procedures that are simple generalizations of matrix inversion. In single-particle problems, the resolvent is therefore obtained by means of the Lippmann–Schwinger equation

$$\mathcal{G}(E) = G(E) + K(E)\mathcal{G}(E) \tag{1}$$

with the kernel

$$K(E) = G(E)V \tag{2}$$

if V is the interaction of the particle with an external force center, and $G(E) = (E - H_0)^{-1}$. The solution of the "integral equation" (1) amounts to writing

$$\mathcal{G} = (1 - K)^{-1}G.$$

Since for a large class of interesting interactions K is compact, the resolvent $(1 - K)^{-1}$ can be constructed by canonical methods, both numerically and for purposes of the study of general properties of $\mathcal{G}(E)$.

When the problem involves more than one particle in interaction, the kernel K of the Lippmann–Schwinger equation is no longer compact. In the instance of two particles without external forces, this difficulty can be circumvented trivially by separating out the center of mass and reducing the problem to an effective single-particle equation. If there are more than two particles, this method no longer works. Apart from the great complications due to the large number of variables involved, this is the essential mathematical difficulty in the quantum mechanical description of three or more particles in mutual interaction. One solution of the difficulty is given by the Faddeev equations[2,3]; another, by Weinberg's approach and its Sugar–Blankenbecler variant.[4] We present here a third method that may be described as a simple generalization of the Lippmann–Schwinger procedure properly understood.

The virtues of the new approach are to be stressed with some caution. Since the proof of the pudding will be in the eating, the relative merits of the various ways of attacking the three (or more) body problem will manifest themselves in their use. However, the possible advantage of this new approach over Weinberg's equations is that the present ones do not contain both the two-body interactions and their T operators, but only the latter.[5] On the other hand, unlike the Faddeev equations, with which they share this feature, they are not *coupled* operator equations. This may be an advantage, particularly for numerical calculations. Of course, once reduced to a specific representation, they cannot avoid such couplings as exist between various angular momentum components, for example, but the number of simultaneous integral equations is much reduced. What is perhaps more important, the generalization to n particles is, in principle, straightforward. That too, is in contrast to the Faddeev method.[6]

* This work was supported in part by the National Science Foundation and by the U.S. Army Research Office (Durham).

[1] Compactness of an operator means the same thing as complete continuity.

[2] L. D. Faddeev, Zh. Eksperim. i Teor. Fiz. **39**, 1459 (1960) [English transl.: Soviet Phys.—JETP **12**, 1014 (1961)]; Dokl. Akad. Nauk SSSR **138**, 565(1961) [English transl.: Soviet Phys.—Doklady **6**, 384 (1961)].

[3] L. D. Faddeev, *Mathematical Aspects of the Three-Body Problem in Quantum Scattering Theory* (Daniel Davey & Company, Inc., New York, 1965).

[4] S. Weinberg, Phys. Rev. **133**, B232 (1964); R. Sugar and R. Blankenbecler, *ibid.* **136**, B472 (1964). See also C. van Winter, Kgl. Danske Videnskab. Selskab, Mat.-fys. Skrifter **2**, no. 8 (1964) and no. 10 (1965).

[5] Another advantage is that the resolvents to be constructed contain no singularities that are not present in the full Green's function. The problem of spurious "homogeneous solutions" does not arise. [See P. G. Federbush, Phys. Rev. **148**, 1551 (1966); J. V. Noble, Phys. Rev. **148**, 1553 (1966); R. G. Newton, *ibid.* **153**, 1502 (1967).]

[6] The Faddeev method has been generalized to more than three particles by L. Rosenberg, Phys. Rev. **140**, B217 (1965); V. A. Alessandrini, J. Math. Phys. **7**, 215 (1966); J. Weyers, Phys. Rev. **145**, 1236 (1966); **151**, 1159 (1966).

In Sec. II we briefly discuss in various ways the troubles with the Lippmann–Schwinger equation for more than one particle. In Sec. III we give the formal solution of the difficulties. In Sec. IV we prove that the equations obtained contain only "connected" kernels, and we examine their compactness. In Sec. V we discuss the simple example of three particles, one of which is infinitely massive. Section VI contains a discussion of the general three-body problem, and Sec. VII, of the n-body problem.

II. DISCONNECTED DIAGRAMS

Let us first briefly look at the two-particle problem, with $V(|\mathbf{r}_1 - \mathbf{r}_2|)$ as the inter-particle potential. Because V depends only on the distance between the particles, it commutes with the total momentum operator $\mathbf{P}$. The two-body Green's function

$$G(E) = [E - (p_1^2/2m_1) - (p_2^2/2m_2)]^{-1}$$
$$= [E - (p^2/2\mu) - (P^2/2M)]^{-1},$$

in which

$$\mathbf{P} = \mathbf{p}_1 + \mathbf{p}_2, \quad \mathbf{p} = \tfrac{1}{2}(\mathbf{p}_1 - \mathbf{p}_2),$$
$$M = m_1 + m_2, \quad \mu = m_1 m_2/(m_1 + m_2)$$

also commutes with $\mathbf{P}$. Hence so does the product $K = GV$. As a consequence K cannot be a compact operator in the two-particle Hilbert space. This can be seen either from the fact that, in the momentum representation, K must contain a δ function, expressing the conservation of total momentum, or from a lemma[7] whose relevant corollary states that if K is compact and commutes with a Hermitian operator whose entire spectrum forms a continuum, then $K = 0$.

Now, because both the kernel and the inhomogeneity G in the Lippmann–Schwinger equation (1) contain the same momentum conserving δ function, it must be present in $\mathcal{G}$ and can be factored out. The remaining factor then obeys a Lippmann–Schwinger equation which looks just like that for a one-particle problem, and the difficulty has been removed.

Mathematically speaking we are looking at the subspace of the two-particle Hilbert space which is an eigenspace of the total momentum.[8] On this subspace the kernel of the Lippmann–Schwinger equation is compact.

Now, when there are more than two particles, the difficulties are of a similar nature. But they cannot be removed so simply. Because of the presence of binary

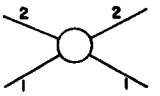


FIG. 1. The diagram for the interactions of two particles.

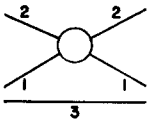


FIG. 2. A disconnected three-particle diagram.

potentials,[9] the kernel of the Lippmann–Schwinger equation is a sum of terms, each of which commutes with a different momentum. Therefore it is not compact. But, except for the total momentum, there is no δ function that can be factored out to remove the trouble. If we draw a diagram such as in Fig. 1 for the interaction of two particles, then the lack of compactness of a three-particle kernel can be attributed to the presence of "disconnected diagrams" such as in Fig. 2, and their repetition (in a familiar perturbation theory language). The fact that particle 3 does not interact is responsible for the δ function that expresses the conservation of its momentum.

Abstractly speaking the general situation is as follows. Let us suppose we have two operators A and B that satisfy these conditions:

(a) There exists a set of mutually commuting Hermitian operators with continuous spectra only, that falls into three disjoint classes. The members p_a of the first commute with B, but not with A:

$$[A, p_a] \neq 0, \quad [B, p_a] = 0;$$

those of the second, p_b, commute with A, but not with B:

$$[A, p_b] = 0, \quad [B, p_b] \neq 0;$$

those of third, p_c, commute with both A and B:

$$[A, p_c] = [B, p_c] = 0.$$

(b) On the eigenspaces of p_b and p_c (i.e., where p_b and p_c are fixed numbers) A has the Hilbert–Schmidt property.[10] Similarly, B has the Hilbert–Schmidt property on the eigenspaces of p_a and p_c. Furthermore, the products AB and BA are in the Hilbert–Schmidt class on the eigenspaces of p_c.

Under these conditions A and B are not compact, and there generally does not exist an invariant

[7] R. G. Newton, *Scattering Theory of Waves and Particles* (McGraw-Hill Book Company Inc., New York, 1966), p. 203.

[8] Strictly speaking there is, of course, no such eigenspace. We do not know how to formulate this statement rigorously in Hilbert space language. It can be seen below what exactly is meant in the momentum representation.

[9] Note that this has nothing to do with assuming that there are two-body forces *only*. The disconnected diagram troubles appear as soon as binary forces are present, regardless of the existence of three-body (or more) potentials.

[10] The operator A has the Hilbert–Schmidt property if $\mathrm{tr}\, AA^\dagger < \infty$. We then also say that A is in L^2. This property implies compactness.

subspace on which the operator $C = A + B$ is compact either. That is to say, C cannot be made compact by "factoring out a δ function." If we try to construct $(1 - \gamma C)^{-1}$ by power series expansion in γ, the trouble becomes visible in the form of "disconnected diagrams."

Using a representation in which all the p's are diagonal, we have

$$(p_a', p_b', p_c', \alpha' | A | p_a'', p_b'', p_c'', \alpha'') = \delta(p_b' - p_b'')\delta(p_c' - p_c'')A(p_b', p_c'; p_a', \alpha'; p_a'', \alpha''),$$

$$(p_a', p_b', p_c', \alpha' | B | p_a'', p_b'', p_c'', \alpha'') = \delta(p_a' - p_a'')\delta(p_c' - p_c'')B(p_a', p_c'; p_b', \alpha'; p_b'', \alpha'').$$

Condition (b) says that $A(p_b', p_c'; p_a', \alpha'; p_a'', \alpha'')$, considered as an integral operator on the variables p_a and α, is in L^2,

$$\int dp_a'\, dp_a''\, d\alpha'\, d\alpha''\, |A(p_b', p_c'; p_a', \alpha'; p_a'', \alpha'')|^2 < \infty.$$

Analogously for B,

$$\int dp_b'\, dp_b''\, d\alpha'\, d\alpha''\, |B(p_a', p_c'; p_b', \alpha'; p_b'', \alpha'')|^2 < \infty.$$

Furthermore, if we write the operator product:

$$(p_a', p_b', p_c', \alpha' | AB | p_a'', p_b'', p_c'', \alpha'') = \delta(p_c' - p_c'')(AB)(p_c'; p_a', p_b', \alpha'; p_a'', p_b'', \alpha''),$$

where

$$(AB)(p_c'; p_a', p_b', \alpha'; p_a'', p_b'', \alpha'') \equiv \int d\alpha'''\, A(p_b', p_c'; p_a', \alpha'; p_a'', \alpha''') \times B(p_a'', p_c'; p_b', \alpha'''; p_b'', \alpha''),$$

then $(AB)(p_c'; p_a', p_b', \alpha'; p_a'', p_b'', \alpha'')$, considered as an integral operator on the variables p_a, p_b, and α, is in L^2,

$$\int dp_a'\, dp_a''\, dp_b'\, dp_b''\, d\alpha'\, d\alpha'' \times |(AB)(p_c'; p_a', p_b', \alpha'; p_a'', p_b'', \alpha'')|^2 < \infty.$$

The problem now is to construct the resolvent $(1 - C)^{-1}$ in terms of resolvents of compact operators only.

III. THE CONSTRUCTION OF THE RESOLVENT

The construction of $(1 - C)^{-1}$ proceeds in analogy with the Lippmann–Schwinger procedure. The latter amounts to using the algebraic identity

$$(E - H)^{-1} = (E - H_0 - V)^{-1} = [1 - (E - H_0)^{-1}V]^{-1}(E - H_0)^{-1};$$

that is,

$$\mathfrak{G} = (1 - K)^{-1}G$$

from which we get the Lippmann–Schwinger equation by writing

$$(1 - K)^{-1} = 1 + K(1 - K)^{-1}. \tag{3}$$

We proceed similarly here:

$$\begin{aligned}(1 - A - B)^{-1} &= [1 - (1 - A)^{-1}B]^{-1}(1 - A)^{-1} \\ &= [1 - B - A(1 - A)^{-1}B]^{-1}(1 - A)^{-1} \\ &= (1 - B)^{-1}[1 - A(1 - A)^{-1}B(1 - B)^{-1}]^{-1} \times (1 - A)^{-1}. \end{aligned} \tag{4}$$

Now because of our assumptions on the operators A and B, each of the three terms into which $(1-A-B)^{-1}$ has been factored is the resolvent of an L^2 operator on an invariant subspace (i.e., after appropriate δ functions have been factored out). Each of these can therefore be constructed by the analog of a Lippmann–Schwinger equation with a Hilbert–Schmidt kernel, such as (3).

Let us introduce the notation

$$T(A) \equiv T_A \equiv A(1 - A)^{-1},$$
$$R(A) \equiv R_A \equiv 1 + T_A = (1 - A)^{-1},$$

as convenient and suggestive abbreviations. Then our result (4) can be written

$$\begin{aligned} R_{A+B} &= R_A R(T_B T_A) R_B \\ &= R_B R(T_A T_B) R_A. \end{aligned} \tag{4'}$$

The corresponding T operator is given by

$$\begin{aligned} T_{A+B} &= T_A R(T_B T_A) R_B + T_B R(T_A T_B) R_A \\ &= R_A R(T_B T_A) T_B + R_B R(T_A T_B) T_A, \end{aligned} \tag{5}$$

which can also be written

$$T_{A+B} = T_A + T_B + T^{(c)}_{A+B}, \tag{6}$$

where

$$T^{(c)}_{A+B} \equiv T_A T_B + R_A T(T_B T_A) R_B.$$

Because $T^{(c)}_{A+B}$ contains no terms that commute both with p_a and p_b, it is expected to be "connected" on the eigenspaces of p_c. We say that $T^{(c)}_{A+B}$ is the *connected part* of T_{A+B}.

The construction of T_{A+B} or of R_{A+B} is thus accomplished by the successive construction of R_A, R_B (or T_A, T_B), and then, in terms of these, of $R(T_A T_B)$ or $T(T_A T_B)$. These are the essentially new

resolvents or T operators. Note that they do not involve A or B themselves, but only T_A and T_B.

The procedure can be repeated and applied to more than two terms. We find after some straight forward algebraic manipulations that

$$R_{A+B+C} = R_A R(T_B T_A) R_B R(T_C T_A) R_D R(T_C T_B) R_C \tag{7}$$

with

$$D \equiv T(T_C T_B) T(T_C T_A) - R(T_C T_B) T_C T^{(c)}_{A+B} R(T_C T_A),$$

and similarly

$$T_{A+B+C} = T_{A+B} + R_A T(T_C T_A) + R_B T(T_C T_B) + T^{(c)}_{A+B+C}, \tag{8}$$

where

$$T^{(c)}_{A+B+C} = R_{A+B} T^{(c)}(T_C T_A + T_C T_B) + T_B T(T_C T_A) + T_A T(T_C T_B) + R_{A+B} T_E R(T_C T_A + T_C T_B),$$
$$E \equiv R[T_C(T_A + T_B)] T_C T^{(c)}_{A+B}.$$

In general, we can always reduce the resolvent of a sum of $n+1$ operators to that of a sum of n

$$R\left(\sum_{i=1}^{n+1} A_i\right) = R\left(\sum_i^n A_i\right) R\left[T(A_{n+1}) T\left(\sum_1^n A_i\right)\right] R(A_{n+1}). \tag{9}$$

In the next step $T(\sum_1^n A_i)$ must be written as a sum of disconnected pieces and a connected term

$$T\left(\sum_1^n A_i\right) = T^{(d)} + T^{(c)} \tag{10}$$

so that

$$R\left(\sum_1^{n+1} A_i\right) = R\left(\sum_1^n A_i\right) R\left[T(A_{n+1}) T^{(d)}\left(\sum_1^n A_i\right)\right] R\left[T(A_{n+1}) T^{(c)}\left(\sum_1^n A_i\right) T(A_{n+1}) T^{(d)}\left(\sum_1^n A_i\right)\right] \times R\left[T(A_{n+1}) T^{(c)}\left(\sum_1^n A_i\right)\right] R(A_{n+1}). \tag{11}$$

At this point the calculation is reduced to that of

$$R\left(\sum_1^n A_i\right)$$

and of

$$R\left[T(A_{n+1}) T^{(d)}\left(\sum_1^n A_i\right)\right].$$

Everything else in (11) is connected.

IV. PROOF OF CONNECTEDNESS

In our applications of these general results, the operators A, B, C, etc., are all of the form of Lippmann–Schwinger kernels. Schematically, V_1 may be the potential for particle 1, V_2 the potential for particle 2, and G the two-particle Green's function. Then in the momentum representation[11]

$$A(E; \mathbf{k}_1, \mathbf{k}_2; \mathbf{k}_1', \mathbf{k}_2') = \delta(\mathbf{k}_2 - \mathbf{k}_2') A(E - k_2^2; \mathbf{k}_1, \mathbf{k}_1'),$$
$$B(E; \mathbf{k}_1, \mathbf{k}_2; \mathbf{k}_1', \mathbf{k}_2') = \delta(\mathbf{k}_1 - \mathbf{k}_1') B(E - k_1^2; \mathbf{k}_2, \mathbf{k}_2'). \tag{12}$$

with

$$A(E; \mathbf{k}, \mathbf{k}') = \tilde{V}_1(\mathbf{k} - \mathbf{k}')/(E - k'^2),$$
$$B(E; \mathbf{k}, \mathbf{k}') = \tilde{V}_2(\mathbf{k} - \mathbf{k}')/(E - k'^2).$$

(The tildes indicate Fourier transforms.) Similarly for the operators $T(A)$ and $T(B)$. If $T_1(E; \mathbf{k}, \mathbf{k}')$ and $T_2(E; \mathbf{k}, \mathbf{k}')$ are off-the-energy-shell two-body T matrices, then

$$T_A(E; \mathbf{k}_1, \mathbf{k}_2; \mathbf{k}_1', \mathbf{k}_2') = \delta(\mathbf{k}_2 - \mathbf{k}_2') \frac{T_1(E - k_2^2; \mathbf{k}_1, \mathbf{k}_1')}{E - k_1^2 - k_2^2},$$

$$T_B(E; \mathbf{k}_1, \mathbf{k}_2; \mathbf{k}_1', \mathbf{k}_2') = \delta(\mathbf{k}_1 - \mathbf{k}_1') \frac{T_2(E - k_1^2; \mathbf{k}_2, \mathbf{k}_2')}{E - k_1^2 - k_2^2}. \tag{13}$$

Consequently, we have

$$(AB)(\mathbf{k}_1, \mathbf{k}_2; \mathbf{k}_1', \mathbf{k}_2') = A(E - k_2^2; \mathbf{k}_1, \mathbf{k}_1') \times B(E - k_1'^2; \mathbf{k}_2, \mathbf{k}_2')$$

and similarly for $T_A T_B$. Therefore, if the kernels $A(E; \mathbf{k}, \mathbf{k}')$ and $B(E; \mathbf{k}, \mathbf{k}')$ are in L^2 *uniformly* in E, then it follows immediately that the kernel AB is in L^2.

Now, for complex E it is quite simple to see that A and B are in L^2. For real E it has been shown by Jost and Pais[12] that A^2 and B^2 are in the Hilbert–Schmidt class under very general and weak conditions on the potentials. Since the energy appears in their proof (in the coordinate representation) only as a phase factor, which in turn is eliminated right from the start, their demonstration shows the necessary *uniformity* in E at the same time. So we have the result that AB contains no disconnected diagrams and furthermore, $(AB)^2$ is L^2 for real E.

The situation for $T_A T_B$ is more complicated. Let us first take E complex. Then, as an integral operator on $\mathbf{k}_1$, $(1 - A)^{-1}$ is bounded uniformly with respect to

[11] In order not to complicate the formulas in this section unnecessarily we use such units for the $\mathbf{k}_i$ that the kinetic energy of the ith particle is simply k_i^2.

[12] R. Jost and A. Pais, Phys. Rev. **82**, 840 (1951).

the real part of E. But since

$$\text{tr}\,(ab)(ab)^\dagger = \text{tr}\,(a^\dagger a)(bb^\dagger) \leq \|b\|^2\,\text{tr}\,a^\dagger a,$$

it follows that $T_A = A(1-A)^{-1}$ has the Hilbert–Schmidt property uniformly with respect to the real part of E. Consequently $T_A T_B$ is in L^2 when E is complex. This proves that $(1 - T_A T_B)^{-1}$ is "connected." Disconnected diagrams cannot appear when E is made real.

But disconnectedness is not all one has to worry about here. The foregoing argument is not sufficient to show that for real E the operator $T_A T_B$ is in L^2 or even compact. The main difficulty comes from the fact that if V_1 leads to bound states then, as a function of E, $T_1(E; \mathbf{k}, \mathbf{k}')$ has poles at negative values of $E = E_n$, $n = 1, 2, \cdots$. Hence for every value of $E \geq E_1$ the integral operator $T_1(E - k_2^2; \mathbf{k}_1, \mathbf{k}_1')$ fails to exist for certain values of $\mathbf{k}_2$. The situation with such bound states present is the one of particular interest. If there are no bound states in any of the two-particle systems, then no scattering target can be formed.

Rather than attempting an independent proof that for real E the operator $T_A T_B$ is compact, we simply refer the reader to the paper by Faddeev.[3] The same problem arises, of course, in the context of the Faddeev equations, and his proof is applicable in the present context as well (see particularly Sec. 6 of Ref. 3).

It is clear that all the arguments of this section are easily generalized to particles in mutual interaction. That involves a relabeling of the variables only. Furthermore, they are readily generalized to more than two sets of $\mathbf{k}$'s and more than two operators. For example, we may have

$$A(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3; \mathbf{k}_1', \mathbf{k}_2', \mathbf{k}_3') = \delta(\mathbf{k}_2 - \mathbf{k}_2')\delta(\mathbf{k}_3 - \mathbf{k}_3')A(E - k_2^2 - k_3^2; \mathbf{k}_1, \mathbf{k}_1'),$$

$$B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3; \mathbf{k}_1', \mathbf{k}_2', \mathbf{k}_3') = \delta(\mathbf{k}_1 - \mathbf{k}_1')\delta(\mathbf{k}_3 - \mathbf{k}_3')B(E - k_1^2 - k_3^2; \mathbf{k}_2, \mathbf{k}_2'),$$

$$C(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3; \mathbf{k}_1', \mathbf{k}_2', \mathbf{k}_3') = \delta(\mathbf{k}_1 - \mathbf{k}_1')\delta(\mathbf{k}_2 - \mathbf{k}_2')C(E - k_1^2 - k_2^2; \mathbf{k}_3, \mathbf{k}_3'),$$

and then prove the connectedness of the product ABC.

V. THREE-BODY PROBLEM WITH ONE PARTICLE FIXED

As an example, let us consider a three-body problem with particle 3 infinitely massive and fixed at the origin. Then

$$H = H_0 + V_{1+2} + V_{12}, \quad V_{1+2} = V_1 + V_2. \tag{14}$$

The first step is to factor out $G = (E - H_0)^{-1}$:

$$E - H = (E - H_0)(1 - A - B) \tag{15}$$

with

$$A = GV_{1+2}, \quad B = GV_{12}. \tag{16}$$

Equation (4′) tells us that

$$\mathcal{G} = (E - H)^{-1} = R_{A+B}G = R_B R_C R_A G, \tag{17}$$

where

$$C = T_A T_B.$$

Therefore we must construct R_A, T_A, R_B, and T_B.

The first two are the Green's function and T operator of the problem in which particles 1 and 2 do not interact with one another:

$$1 - R_A = 1 - (1 - GV_{1+2})^{-1} = (E - H_1 - H_2)^{-1}V_{1+2} = GT_{1+2}$$

and H_1 and H_2 commute. This Green's function can be constructed from the solutions of the two one-particle problems with H_1 and H_2, and so can the corresponding T operator. For example, for values of E such that the spectra of H_1 and of $(E - H_2)$ are disjoint we may write

$$(E - H_1 - H_2)^{-1} = \frac{1}{2\pi i}\int_{\mathcal{C}} dz(z - H_1)^{-1} \times (E - z - H_2)^{-1}, \tag{18}$$

where the contour $\mathcal{C}$ separates the spectra of the operators H_1 and $(E - H_2)$, leaving that of H_1 on its left.[13]

The other two operators, R_B and T_B, are the Green's function and T operator of the two-body problem in which particles 1 and 2 do not see particle 3. If we write

$$H_0 = H_{0CM} + H_{012},$$

where H_{0CM} is the kinetic energy of the center of mass of particles 1 and 2, and H_{012} is the kinetic energy of the relative motion, then

$$1 - R_B = 1 - (1 - GV_{12})^{-1} = -(E - H_{0CM} - H_{012} - V_{12})^{-1}V_{12} = -GT_{12}.$$

In the invariant subspace in which the center-of-mass momentum of 1 and 2 is fixed, this is a one-particle Green's function. We may therefore assume that it and the corresponding T operator have been constructed.

We are left with

$$1 - R_C = (E - H_0 - T_{1+2}GT_{12})^{-1}T_{1+2}GT_{12}, \tag{19}$$

which involves the Green's function of an effective energy dependent Hamiltonian whose interaction part is given by

$$\mathcal{H}' = T_{1+2}GT_{12}. \tag{20}$$

[13] Equation (18) was given by L. Bianchi and L. Favella [Nuovo Cimento **34**, 1825 (1964)], but their proof is more complicated than necessary. All that is needed is to apply $(E - H_1 - H_2)$ to both sides of (18) and then to carry out the integral by Cauchy's residue theorem.

The construction of this Green's function contains no disconnected diagrams, because for complex values of E (on the physical sheet) the operator $GT_{1+2}GT_{12}$ is in L^2. All the important properties of this three-body problem that differ from those of the three individual two-body problems are contained in the function $\mathfrak{G} = (E - H_0 - \mathcal{H}')^{-1}$ which is the solution of an integral equation with compact kernel:

$$\mathfrak{G} = G + G\mathcal{H}'\mathfrak{G}. \tag{21}$$

The pseudo-Hamiltonian $\mathcal{H}'$ has a simple physical interpretation. It clearly describes the scattering of particles 1 and 2 by one another, followed by free propagation and subsequent scattering of both particles by the center. Because this single term $\mathcal{H}'$ involves both particles and the center, it leads to no disconnected graphs.

Rearrangement collisions in which initially particles 1 and 2 are bound together and finally particle 1 or 2 is bound with the center, are described by matrix elements of the operator[14]

$$T_{\underline{b}a} = V_{12} + V_{12}\mathfrak{G}V_{1+2} \tag{22}$$

between bound-state eigenfunctions of $H_0 + V_{1+2}$ and $H_0 + V_{12}$. We find that

$$GT_{\underline{b}a} = B + T(T_AT_B) \tag{23}$$

or

$$T_{\underline{b}a} = V_{12} + T_{12}\mathfrak{G}T_{1+2} \tag{23'}$$

so that the only problem is to construct the Green's function $\mathfrak{G}$ for the interaction $\mathcal{H}'$.[15]

VI. THE FULL THREE-BODY PROBLEM

The general three-body problem is more complicated than the one outlined above because, when the mass of the third particle is finite, then T_{1+2} cannot be constructed so easily. One then has to fall back on Eqs. (7) and (8), with the operators $A = GV_{12}$, $B = GV_{13}$, and $C = GV_{23}$. These contain resolvents of compact operators only (at least for complex energies) and so there are no disconnected diagram difficulties.

Rearrangement collisions of the type $(12)3 \rightarrow (13)2$ are described by matrix elements of the operator[14]

$$T_{\underline{b}a} = V_{13} + V_{23} + (V_{12} + V_{23})\mathfrak{G}(V_{13} + V_{23}) \tag{24}$$

between eigenstates of $H_0 + V_{12}$ and $H_0 + V_{13}$ in which particles 1 and 2 or 1 and 3, respectively, are bound. We then get

$$GT_{\underline{b}a} = B + R(T_CT_B)R_CR(T_AT_B)R_DR(T_AT_C) - 1 \tag{25}$$

with

$$D = T(T_AT_C)T(T_AT_B) + R(T_AT_C)T_AT^{(c)}_{B+C}R(T_AT_B).$$

Similarly, for a dissociation collision $(12)3 \rightarrow 123$ we need matrix elements of the operator

$$T_{c\underline{a}} = V_{13} + V_{23} + (V_{12} + V_{13} + V_{23})\mathfrak{G}(V_{13} + V_{23}), \tag{26}$$

and we get

$$GT_{c\underline{a}} = R_CR(T_BT_C)R_BR(T_AT_B)R_DR(T_AT_C) - 1. \tag{27}$$

All the essentially new properties of the S matrix for three-body collisions are contained in the resolvent R_D. The latter can be written as a Green's function of a pseudo-Hamiltonian whose interaction term is given by

$$\mathcal{H}' = T_{12}G_{23,12}[T_{23}GT_{13} + T_{23}GT_{12}GT_{13} + T_{13} \times (1 + GT_{23})G_{23,13}T_{23}(1 + GT_{13})](1 + GT_{12}G_{13,12}T_{13}) \tag{28}$$

with the auxiliary Green's functions

$$\begin{aligned} G_{23,12} &= (E - H_0 - T_{23}GT_{12})^{-1}, \\ G_{23,13} &= (E - H_0 - T_{23}GT_{13})^{-1}, \\ G_{13,12} &= (E - H_0 - T_{13}GT_{12})^{-1}. \end{aligned} \tag{29}$$

VII. THE n-BODY PROBLEM

It is clear how we proceed in the n-body problem, using the procedure of Eqs. (9) to (11). Unfortunately it is not possible to reduce the $(n + 1)$-body problem in any simple way to the n-body problem. But if we know not only the T operators for n particles but also their split-up into connected and disconnected parts, then it is straight forward (but complicated) to calculate the T operator, or the Green's function, for $(n + 1)$ particles. The main difficulty in the next step is the separation of the $(n + 1)$-body T operator into connected and disconnected parts. That involves all the lower-order terms and no simple iterative scheme, based on the connected and disconnected parts of the n-body T operator only, exists.

The result is that the calculation becomes rapidly more and more cumbersome as the number of particles increases. Even though in principle the present method allows the calculation of any collision or reaction amplitude for n particles by solving integral equations with compact kernels only, if n is large it is surely not practical. There is no hope that this method has anything to contribute to the many-body problem.

[14] See, for example, p. 485 of Ref. 7.

[15] It is only fair to point out that, in this simple special case in which one particle is infinitely massive, the Faddeev equations can also be uncoupled and then lead to the same equations we get with the present method.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

A Set of Harmonic Functions for the Group $SU(3)$ as Specialized Matrix Elements of a General Finite Transformation

T. J. NELSON
Institute for Atomic Research and Department of Physics, Iowa State University, Ames, Iowa

(Received 26 September 1966)

A straightforward method of converting the commutation rules of the generators of a group into a set of differential equations is developed. The detailed development is applied to $SU(3)$. Differential operators having the correct commutators are constructed and then combined into the quadratic and cubic combinations of the two Casimir operators. The solution of the resulting set of differential equations is the matrix element of a general finite transformation with the right-hand state specialized to the isotopic spin singlet. The set of such matrix elements is equivalent to the set of $SU(3)$ harmonic functions derived by Beg and Ruegg. The isotopic spin-hypercharge content of the irreducible representations of $SU(3)$ is deduced from the form of the matrix element, though prior knowledge of the fact that each irreducible representation of $SU(3)$ possesses an isotopic spin singlet is assumed.

INTRODUCTION

A SET of functions which carry all the irreducible representations of $SU(3)$ and some representations of $U(3)$ has been derived by Beg and Ruegg.[1] These authors considered differential operators in a three-dimensional complex space in analogy with $SU(2)$. It has been found in the present work that the same functions can be derived by a different procedure which starts by considering the matrix element of a general finite $SU(3)$ transformation between suitably restricted states. The specialization of one of the states to an isotopic spin singlet eliminates the dependence of the matrix element on three of the eight variables of the general transformation, and differential operators which represent the eight generators of $SU(3)$ are derived in the five remaining variables. Differential equations which represent the two Casimir operators are formed and solved, and in this way the original matrix element is evaluated. Finally the isotopic spin-hypercharge content of the irreducible representations of $SU(3)$ is deduced from the resulting expression for the matrix element.

1. $SU(3)$ TRANSFORMATIONS

As a model of the Lie algebra of $SU(3)$, a set of eight Hermitian, traceless, 3×3 matrices are used. These matrices are the ones used by Pursey[2] in his treatment of the irreducible representations of $SU(3)$.

$$T_1=\frac{1}{2}\begin{pmatrix}0&1&0\\1&0&0\\0&0&0\end{pmatrix},\quad N=\begin{pmatrix}0&0&1\\0&0&0\\1&0&0\end{pmatrix},$$

$$T_2=\frac{1}{2}\begin{pmatrix}0&-i&0\\i&0&0\\0&0&0\end{pmatrix},\quad N'=\begin{pmatrix}0&0&-i\\0&0&0\\i&0&0\end{pmatrix},$$

$$T_3=\frac{1}{2}\begin{pmatrix}1&0&0\\0&-1&0\\0&0&0\end{pmatrix},\quad M=\begin{pmatrix}0&0&0\\0&0&1\\0&1&0\end{pmatrix},$$

$$U=\frac{1}{3}\begin{pmatrix}1&0&0\\0&1&0\\0&0&-2\end{pmatrix},\quad M'=\begin{pmatrix}0&0&0\\0&0&-i\\0&i&0\end{pmatrix}. \tag{1.1}$$

In the physics of elementary particles it is well known that T_i are interpreted as the generators of isotopic spin and U is the generator for hypercharge so that the electric charge of a particle is given by $Q = e(T_3 + 1/2U)$. The four remaining operators connect the isotopic spin-hypercharge multiplets, thereby enlarging the group from $SU(2) \times U(1)$ to $SU(3)$.

A general finite $SU(3)$ transformation can be expressed in the form

$$\mathcal{C} = e^{-i\beta U}e^{-i\alpha_3 T_3}e^{-i\alpha_2 T_2}e^{-i\gamma T_3}e^{-i(\nu N+\nu' N'+\mu M+\mu' M')}. \tag{1.2}$$

However, with the use of the identity[3]

$$e^{A}Be^{-A}=\sum_{n=0}^{\infty}\frac{1}{n!}[A,\}^n B], \tag{1.3}$$

[1] M. Beg and H. Ruegg, J. Math. Phys. 6, 677 (1965).
[2] D. L. Pursey, Proc. Roy. Soc. (London) A275, 284 (1963).
[3] E. Merzbacher, *Quantum Mechanics* (John Wiley & Sons, Inc., New York, 1961), p. 162.

where

$$\begin{aligned}[A, \}^0 B] &= B,\\ [A, \}^1 B] &= [A, B],\\ [A, \}^2 B] &= [A, [A, B]],\\ &\cdot\\ &\cdot\\ &\cdot\end{aligned} \tag{1.4}$$

it can be shown that

$$e^{-i\alpha_3' T_3} e^{-i\alpha_2' T_2} e^{-i\gamma' T_3} e^{-i\nu N} e^{i\gamma' T_3} e^{i\alpha_2' T_2} e^{i\alpha_3' T_3}$$

$$= \exp\left\{-i\nu\begin{bmatrix}\cos(\frac{1}{2}\alpha_2')\cos[\frac{1}{2}(\gamma'+\alpha_3')]N\\ +\cos(\frac{1}{2}\alpha_2')\sin[\frac{1}{2}(\gamma'+\alpha_3')]N'\\ +\sin(\frac{1}{2}\alpha_2')\cos[\frac{1}{2}(\gamma'-\alpha_3')]M\\ +\sin(\frac{1}{2}\alpha_2')\sin[\frac{1}{2}(\gamma'-\alpha_3')]M'\end{bmatrix}\right\}. \tag{1.5}$$

Therefore, this form can be used to represent the last exponential in Eq. (1.2), so that with some redefinition of the variables and with the observation that the product of two $SU(2)$ transformations is, by the group property, a third $SU(2)$ transformation, Eq. (1.2) becomes

$$\mathcal{C} = e^{-i\beta U} e^{-i\alpha_3 T_3} e^{-i\alpha_2 T_2} e^{-i\gamma T_3} e^{-i\nu N} e^{i\gamma' T_3} e^{i\alpha_2' T_2} e^{i\alpha_3' T_3}. \tag{1.6}$$

The set of matrix elements of $\mathcal{C}$ between states of the same irreducible representation of $SU(3)$ can be regarded as the most general set of harmonic functions which carry that representation. It is in precisely this sense that the rotation matrices $D^j_{mm'}$ are the most general $SU(2)$ harmonic functions. The mathematical complexity of the general problem for $SU(3)$ has not yet been resolved, but if one of the states is restricted to the isotopic spin singlet component, the problem is much easier. Since every irreducible representation possesses such a component,[2] no loss of generality in the representations of $SU(3)$ is incurred. The matrix element of $\mathcal{C}$ with a general state on the left and the isotopic spin singlet component ($t = 0$) on the right is denoted as ψ

$$\psi = \langle c(3)c(2); t\mu, u|\ \mathcal{C}\ |c(3)c(2); 00, u_0\rangle \tag{1.7}$$

in anticipation that ψ is one of the harmonic functions under consideration. Here u is the eigenvalue of U, the hypercharge, and μ is the eigenvalue of T_3. The two numbers $c(3)$ and $c(2)$ label the irreducible representations and are the eigenvalues of the two Casimir operators.[2]

$$\begin{aligned}C^{(2)} &= T_1^2 + T_2^2 + T_3^2 + \tfrac{3}{4}U^2\\ &\quad + \tfrac{1}{4}(N^2 + N'^2 + M^2 + M'^2),\\ C^{(3)} &= \tfrac{1}{2}T_1([N, M]_+ + [N', M']_+)\\ &\quad + \tfrac{1}{2}T_2(-[N, M']_+ + [N', M]_+)\\ &\quad + \tfrac{1}{2}T_3(N^2 + N'^2 - M^2 - M'^2)\\ &\quad + U\begin{Bmatrix}2(T_1^2 + T_2^2 + T_3^2) - \frac{1}{2}U^2\\ -1 - \frac{1}{4}(N^2 + N'^2 + M^2 + M'^2)\end{Bmatrix}.\end{aligned} \tag{1.8}$$

The simplification that results from choosing the isotopic spin singlet component on the right is that the dependence of ψ on γ', α_2', and α_3' drops out.

$$\begin{aligned}\psi &= \langle c(3)c(2); t\mu, u|\ T\ |c(3)c(2); 00, u_0\rangle\\ &= \langle c(3)c(2); t\mu, u|\ e^{-i\beta U} e^{-i\alpha_3 T_3} e^{-i\alpha_2 T_2} e^{-i\gamma T_3} e^{-i\nu N}\\ &\qquad \times |c(3)c(2); 00, u_0\rangle.\end{aligned} \tag{1.9}$$

In the $SU(2)$ analogy, the dependence of $D^j_{mm'}(\phi, \theta, \chi)$ on χ drops out when m' is taken to be zero.

The boundary conditions of ψ in the remaining five variables can be inferred from the fundamental representation Eqs. (1.1). There the eigenvalues of the T_i are $\pm\frac{1}{2}$, 0, those of U are $+\frac{1}{3}$, $-\frac{2}{3}$, and those of N are ± 1, 0. Since in any representation the eigenvalues of a generator must be linear combinations with integer coefficients of its eigenvalues in the fundamental representation, ψ must be periodic over 6π in β, 4π in α_3, α_2, and γ and 2π in ν. Furthermore, ψ must be regular in the five variables because of the unitary nature of T. The next section shows how differential operators that have the commutation rules of $SU(3)$ can be constructed in the five variables.

2. DIFFERENTIAL OPERATORS OF $SU(3)$

If a set of differential operators can be found which, operating on ψ, have the effect of placing each of the eight generators of $SU(3)$ somewhere in the matrix element, then the Casimir operators of Eq. (1.8) can be used to form two partial differential equations for ψ. The boundary conditions of ψ, its regularity, and the unitarity condition explained below then suffice for the complete evaluation of ψ except for an overall constant multiplier which can be determined from the condition that ψ must go to $\delta_{t,0}\delta_{\mu,0}\delta_{\mu,\mu_0}$ when its arguments go to zero.

In the following let G stand for one of the eight generators and let $D(G)$ stand for a differential operator which represents G. Then if

$$\begin{aligned}D(G)\psi &= D(G)\,\langle c(3)c(2); t\mu, u|\ T\ |c(3)c(2); 00, u_0\rangle\\ &= -\langle c(3)c(2); t\mu, u|\ GT\ |c(3)c(2); 00, u_0\rangle\end{aligned} \tag{2.1}$$

it becomes a trivial matter to show that

$$[D(G), D(G')]\psi = -\langle c(3)c(2); t\mu, u|\, [G, G']T\,|c(3)c(2); 00, u_0\rangle. \quad (2.2)$$

Thus the differential operators $D(G)$ have the commutation rules of the generators of $SU(3)$ at least when operating on ψ. Now let

$$D(U) = d^{\beta}(U)\frac{1}{i}\frac{\partial}{\partial\beta} + d^{\alpha_3}(G)\frac{1}{i}\frac{\partial}{\partial\alpha_3} + d^{\alpha_2}(G)\frac{1}{i}\frac{\partial}{\partial\alpha_2} + d^{\gamma}(G)\frac{1}{i}\frac{\partial}{\partial\gamma} + d^{\nu}(G)\frac{1}{i}\frac{\partial}{\partial\nu}, \quad (2.3)$$

where the coefficients $d^{\beta}(G), \cdots, d^{\nu}(G)$ are each functions of the five variables.

Because the three generators T_i acting on the right-hand state of ψ produce zero, it is expedient to commute the results of the differentiations through to the right of T. In this way five equations in $d^{\beta}(G), \cdots, d^{\nu}(G)$ result from the coefficients of the generators N, N', M, M', and U. That is,

$$\begin{aligned} D(G)\psi &= -\langle c(3)c(2); t\mu, u|\, T(T^{-1}GT)\,|c(3)c(2); 00, u_0\rangle \\ &= -\langle c(3)c(2); t\mu, u|\, T \\ &\quad \times (NN'MM'U)\mathcal{A}\begin{pmatrix} d^{\beta}(G) \\ d^{\alpha_3}(G) \\ d^{\alpha_2}(G) \\ d^{\gamma}(G) \\ d^{\nu}(G) \end{pmatrix} \\ &\quad \times |c(3)c(2); 00, u_0\rangle, \end{aligned} \quad (2.4)$$

where $\mathcal{A}$ is a 5×5 matrix; the elements of which are functions of the five variables. The finite transformation of any generator by any other generator is tabulated in the Appendix. Using this table it can be shown that

$$\mathcal{A} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ \frac{1}{2}\sin 2\nu & \frac{1}{4}\cos\alpha_2 \sin 2\nu & 0 & \frac{1}{4}\sin 2\nu & 0 \\ 0 & \frac{1}{2}\sin\alpha_2 \sin\gamma \sin\nu & \frac{1}{2}\cos\gamma \sin\nu & 0 & 0 \\ 0 & -\frac{1}{2}\sin\alpha_2 \cos\gamma \sin\nu & \frac{1}{2}\sin\gamma \sin\nu & 0 & 0 \\ \frac{1}{4}(1 + 3\cos 2\nu) & -\frac{3}{8}\cos\alpha_2(1 - \cos 2\nu) & 0 & -\frac{3}{8}(1 - \cos 2\nu) & 0 \end{pmatrix} \quad (2.5)$$

$\det\{\mathcal{A}\} = -\frac{1}{16}\sin^2\nu \sin 2\nu \sin\alpha_2$. It is efficient at this point to invert $\mathcal{A}$. Then as $T^{-1}GT$ is evaluated for each generator G with the help of the Appendix, the problem of finding the coefficients $d^{\beta}(G), \cdots, d^{\nu}(G)$ becomes a simple matrix multiplication

$$\mathcal{A}^{-1} = \begin{pmatrix} 0 & \frac{3}{2}\csc 2\nu(1 - \cos 2\nu) & 0 & 0 & 1 \\ 0 & 0 & 2\csc\nu \csc\alpha_2 \sin\gamma & -2\csc\nu \csc\alpha_2 \cos\gamma & 0 \\ 0 & 0 & 2\csc\nu \cos\gamma & 2\csc\nu \sin\gamma & 0 \\ 0 & \csc 2\nu(1 + 3\cos 2\nu) & -2\csc\nu \cot\alpha_2 \sin\gamma & 2\csc\nu \cot\alpha_2 \cos\gamma & -2 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}. \quad (2.6)$$

It should be obvious that the operators $D(T_i)$ will take the usual form with γ, α_2, and α_3 interpreted as Euler angles. Similarly, it is evident that $D(U) = 1/i(\partial/\partial\beta)$. The complete calculation shows

$$D(T_1) = -\cot\alpha_2 \cos\alpha_3\frac{1}{i}\frac{\partial}{\partial\alpha_3} - \sin\alpha_3\frac{1}{i}\frac{\partial}{\partial\alpha_2} + \csc\alpha_2 \cos\alpha_3\frac{1}{i}\frac{\partial}{\partial\gamma},$$

$$D(T_2) = -\cot\alpha_2 \sin\alpha_3\frac{1}{i}\frac{\partial}{\partial\alpha_3} + \cos\alpha_3\frac{1}{i}\frac{\partial}{\partial\alpha_2} + \csc\alpha_2 \sin\alpha_3\frac{1}{i}\frac{\partial}{\partial\gamma},$$

$$D(T_3) = \frac{1}{i}\frac{\partial}{\partial\alpha_3},$$

$$D(U) = \frac{1}{i}\frac{\partial}{\partial\beta},$$

$$\begin{aligned} D(N) &= \tfrac{3}{2}\sin\left(\beta + \frac{\alpha_3}{2} + \frac{\gamma}{2}\right)\cos\frac{\alpha_2}{2}\tan\nu\frac{1}{i}\frac{\partial}{\partial\beta} \\ &\quad - \sin\left(\beta + \frac{\alpha_3}{2} + \frac{\gamma}{2}\right)\sec\frac{\alpha_2}{2}\cot\nu\frac{1}{i}\frac{\partial}{\partial\alpha_3} \\ &\quad - 2\cos\left(\beta + \frac{\alpha_3}{2} + \frac{\gamma}{2}\right)\sin\frac{\alpha_2}{2}\cot\nu\frac{1}{i}\frac{\partial}{\partial\alpha_2} \\ &\quad - \sin\left(\beta + \frac{\alpha_3}{2} + \frac{\gamma}{2}\right)\left(\cos\frac{\alpha_2}{2}\tan\nu + \sec\frac{\alpha_2}{2}\cot\nu\right)\frac{1}{i}\frac{\partial}{\partial\gamma} \\ &\quad + \cos\left(\beta + \frac{\alpha_3}{2} + \frac{\gamma}{2}\right)\cos\frac{\alpha_2}{2}\frac{1}{i}\frac{\partial}{\partial\nu}, \end{aligned}$$

$$
\begin{aligned}
D(N') = &-\tfrac{3}{2}\cos\left(\beta+\frac{\alpha_3}{2}+\frac{\gamma}{2}\right)\cos\frac{\alpha_2}{2}\tan\nu\,\frac{1}{i}\frac{\partial}{\partial\beta}\\
&+\cos\left(\beta+\frac{\alpha_3}{2}+\frac{\gamma}{2}\right)\sec\frac{\alpha_2}{2}\cot\nu\,\frac{1}{i}\frac{\partial}{\partial\alpha_3}\\
&-2\sin\left(\beta+\frac{\alpha_3}{2}+\frac{\gamma}{2}\right)\sin\frac{\alpha_2}{2}\cot\nu\,\frac{1}{i}\frac{\partial}{\partial\alpha_2}\\
&+\cos\left(\beta+\frac{\alpha_3}{2}+\frac{\gamma}{2}\right)\left(\cos\frac{\alpha_2}{2}\tan\nu\right.\\
&\qquad\left.+\sec\frac{\alpha_2}{2}\cot\nu\right)\frac{1}{i}\frac{\partial}{\partial\gamma}\\
&+\sin\left(\beta+\frac{\alpha_3}{2}+\frac{\gamma}{2}\right)\cos\frac{\alpha_2}{2}\frac{1}{i}\frac{\partial}{\partial\nu},
\end{aligned}
$$

$$
\begin{aligned}
D(M) = &\tfrac{3}{2}\sin\left(\beta-\frac{\alpha_3}{2}+\frac{\gamma}{2}\right)\sin\frac{\alpha_2}{2}\tan\nu\,\frac{1}{i}\frac{\partial}{\partial\beta}\\
&+\sin\left(\beta-\frac{\alpha_3}{2}+\frac{\gamma}{2}\right)\csc\frac{\alpha_2}{2}\cot\nu\,\frac{1}{i}\frac{\partial}{\partial\alpha_3}\\
&+2\cos\left(\beta-\frac{\alpha_3}{2}+\frac{\gamma}{2}\right)\cos\frac{\alpha_2}{2}\cot\nu\,\frac{1}{i}\frac{\partial}{\partial\alpha_2}\\
&-\sin\left(\beta-\frac{\alpha_3}{2}+\frac{\gamma}{2}\right)\left(\sin\frac{\alpha_2}{2}\tan\nu\right.\\
&\qquad\left.+\csc\frac{\alpha_2}{2}\cot\nu\right)\frac{1}{i}\frac{\partial}{\partial\gamma}\\
&+\cos\left(\beta-\frac{\alpha_3}{2}+\frac{\gamma}{2}\right)\sin\frac{\alpha_2}{2}\frac{1}{i}\frac{\partial}{\partial\nu},
\end{aligned}
$$

$$
\begin{aligned}
D(M') = &-\tfrac{3}{2}\cos\left(\beta-\frac{\alpha_3}{2}+\frac{\gamma}{2}\right)\sin\frac{\alpha_2}{2}\tan\nu\,\frac{1}{i}\frac{\partial}{\partial\beta}\\
&-\cos\left(\beta-\frac{\alpha_3}{2}+\frac{\gamma}{2}\right)\csc\frac{\alpha_2}{2}\cot\nu\,\frac{1}{i}\frac{\partial}{\partial\alpha_3}\\
&+2\sin\left(\beta-\frac{\alpha_3}{2}+\frac{\gamma}{2}\right)\cos\frac{\alpha_2}{2}\cot\nu\,\frac{1}{i}\frac{\partial}{\partial\alpha_2}\\
&+\cos\left(\beta-\frac{\alpha_3}{2}+\frac{\gamma}{2}\right)\left(\sin\frac{\alpha_2}{2}\tan\nu\right.\\
&\qquad\left.+\csc\frac{\alpha_2}{2}\cot\nu\right)\frac{1}{i}\frac{\partial}{\partial\gamma}\\
&+\sin\left(\beta-\frac{\alpha_3}{2}+\frac{\gamma}{2}\right)\sin\frac{\alpha_2}{2}\frac{1}{i}\frac{\partial}{\partial\nu}. \qquad (2.7)
\end{aligned}
$$

In fact, these operators have the commutation rules of the generators of $SU(3)$ as they stand, and not merely when operating on ψ.

A ninth operator,

$$\frac{2}{i}\frac{\partial}{\partial\gamma}-\frac{1}{i}\frac{\partial}{\partial\beta},$$

commutes with all the operators of Eqs. (2.7). Since $\frac{1}{2}U - T_3$ commutes with N, it may be seen that

$$
\begin{aligned}
D^{(1)}\psi &= \left(\frac{2}{i}\frac{\partial}{\partial\gamma}-\frac{1}{i}\frac{\partial}{\partial\beta}\right)\psi\\
&= u_0\psi. \qquad (2.8)
\end{aligned}
$$

Beg and Ruegg interpret the emergence of this ninth operator as implying that ψ will also carry some representations of $U(3)$, and that invariance under this $U(3)$ will impose an extra and unwelcome conserved quantum number. As shown later, u_0 is unique for each irreducible representation of $SU(3)$ and thereby serves as one of the two numbers needed to specify such a representation.

When two harmonic functions ψ, ψ' which depend on the same variables are combined in an outer product, the eigenvalues of $D^{(1)}$ are certainly additive. Such a product is really the direct product of two representation matrices having the form given by Eq. (1.9). However, when two harmonic functions ψ, ψ' which depend on different variables and correspond to basis vectors in different spaces are combined in an outer product, the eigenvalues of $D^{(1)}$ cannot be conserved in general for the following reason. When the differential operators are combined in the forms of the Casimir operators, it is seen that $D^{(3)}$ is a function of $D^{(2)}$ and $D^{(1)}$, effectively giving a cubic equation for $D^{(1)}$ in terms of $D^{(2)}$ and $D^{(3)}$. If the eigenvalues of $D^{(1)}$ were additive, a nonlinear relation would exist among the $SU(3)$ operators of the spaces being combined and of the resulting product space. But in an outer product, the eigenvalues of the $SU(3)$ operators are additive so that no such nonlinear relation can exist. In the next section the differential equations for ψ are formed from the Casimir operators and solved.

3. $SU(3)$ HARMONIC FUNCTIONS

Some caution must be used in combining the differential operators of Eq. (2.7) into quadratic or cubic combinations. A product $D(G)D(G')$ acting on ψ has the effect of replacing T by $G'GT$, and a product $D(G)D(G')D(G'')$ replaces T by $-G''G'GT$. The result is that the minus sign is the only significant point for the following reasons. In $C^{(2)}$ all generators occur squared so their order is immaterial. In $C^{(3)}$ in the quantities multiplied by the T_i there are anticommutators among the generators N, N', M, and M' so the order is unimportant in their formation. Each of the T_i commutes with the quantity it multiplies so that the order of each T_i and its factor is arbitrary. The quantity multiplied by U is formed by the squares of the generators and it commutes with U. Hence the

only deviation is the over-all minus sign in the formation of $D^{(3)}$.

The results of combining the differential operators according to Eq. (1.8) are

$$D^{(2)} = \csc^2\nu\left(-\frac{\partial^2}{\partial\alpha_2^2} - \cot\alpha_2\frac{\partial}{\partial\alpha_2} - \csc^2\alpha_2\left(\frac{\partial^2}{\partial\alpha_3^2}+\frac{\partial^2}{\partial\gamma^2}\right) + 2\cot\alpha_2\csc\alpha_2\frac{\partial^2}{\partial\alpha_3\partial\gamma}\right) - \frac{1}{4}\frac{\partial^2}{\partial\nu^2} - \frac{1}{4}(3\cot\nu - \tan\nu)\frac{\partial}{\partial\nu} - \tan^2\nu\left(\frac{3}{4}\frac{\partial}{\partial\beta} - \frac{1}{2}\frac{\partial}{\partial\gamma}\right)^2 - \frac{3}{4}\left(\frac{\partial}{\partial\beta} - \frac{\partial}{\partial\gamma}\right)^2 - \frac{1}{4}\frac{\partial^2}{\partial\gamma},$$

$$D^{(3)} = D^{(1)}(1 - \tfrac{1}{4}D^{(1)2} + D^{(2)}), \tag{3.1}$$

so that ψ satisfies the differential equations

$$\begin{aligned} D^{(2)}\psi &= c(2)\psi, \\ D^{(3)}\psi &= -c(3)\psi \\ &= D^{(1)}[1 - \tfrac{1}{4}D^{(1)2} + D^{(2)}]\psi \\ &= u_0[1 - \tfrac{1}{4}u_0^2 + c(2)]\psi. \end{aligned} \tag{3.2}$$

Now,

$$\begin{aligned}\psi &= \langle c(3)c(2);\, t\mu,\, u|\; e^{-i\beta U}e^{-i\alpha_3 T_3}e^{-i\alpha_2 T_2}e^{-i\gamma T_3}e^{-i\nu N} \\ &\qquad\times |c(3)c(2);\, 00,\, u_0\rangle \\ &= \langle c(3)c(2);\, t\mu,\, u|\; e^{-i\alpha_2 T_2}e^{-i\nu N}|c(3)c(2);\, 00,\, u_0\rangle \\ &\qquad\times e^{-i\beta u}e^{-i\alpha_3 u}e^{\frac{1}{2}i\gamma(u_0-u)}\end{aligned} \tag{3.3}$$

and it is evident that

$$\left[-\frac{\partial^2}{\partial\alpha_2^2} - \cot\alpha_2\frac{\partial}{\partial\alpha_2} - \csc^2\alpha_2\left(\frac{\partial^2}{\partial\alpha_3^2}+\frac{\partial^2}{\partial\gamma^2}\right) + 2\cot\alpha_2\csc\alpha_2\frac{\partial^2}{\partial\alpha_3\partial\gamma}\right]\psi = t(t+1)\psi, \tag{3.4}$$

so that if a solution is assumed of the form

$$\psi = A(\alpha_2)\csc\nu N(\nu)e^{-i\beta u}e^{-i\alpha_3\mu}e^{\frac{1}{2}i\gamma(u_0-u)}, \tag{3.5}$$

then $A(\alpha_2)$ and $N(\nu)$ must satisfy the equations

$$\left\{-\frac{d^2}{d\alpha_2^2} - \cot\alpha_2\frac{d}{d\alpha_2} + \csc^2\alpha_2(\mu^2 + \tfrac{1}{4}[u-u_0]^2) - 2\cot\alpha_2\csc\alpha_2\mu\tfrac{1}{2}(u-u_0)\right\}A(\alpha_2) = t(t+1)A(\alpha_2),$$

$$\begin{aligned}&\left\{-\frac{d^2}{d(2\nu)^2} - \cot 2\nu\frac{d}{d(2\nu)}\right. \\ &\quad + \csc^2 2\nu\{\tfrac{1}{2} + \tfrac{1}{2}[u+\tfrac{1}{2}u_0]^2 + 2t(t+1)\} \\ &\quad \left. + \cot 2\nu\csc 2\nu\{\tfrac{1}{2} - \tfrac{1}{2}[u+\tfrac{1}{2}u_0]^2 + 2t(t+1)\}\right\}N(\nu) \\ &\qquad = [\tfrac{3}{4} - \tfrac{3}{16}u_0^2 + c(2)]N(\nu).\end{aligned} \tag{3.6}$$

Now because ψ must be regular in α_2 and ν and periodic over 4π in α_2 and 2ν, the solutions of these equations are the reduced rotation matrices $d^j_{mm'}$.[4] Hence the solution for $A(\alpha_2)$ is the form

$$A(\alpha_2) = d^t_{\mu,\frac{1}{2}(u-u_0)}(\alpha_2). \tag{3.7}$$

As stated earlier, when $t = 0$ only one value of the hypercharge $u = u_0$ is permitted. For the time being, let

$$\tfrac{3}{4} - \tfrac{3}{16}u_0^2 + c(2) = j(j+1), \tag{3.8}$$

then the solution for $N(\nu)$ must be of the form

$$N(\nu) = d^j_{\frac{1}{2}(u+\frac{1}{2}u_0)+t+\frac{1}{2},\frac{1}{2}(u+\frac{1}{2}u_0)-t-\frac{1}{2}}(2\nu). \tag{3.9}$$

From Eq. (3.9) it can be seen, by specializing to the case $t = 0$, that

$$\begin{aligned} \tfrac{3}{4}u_0 &= \tfrac{1}{2}\text{ integer}, \\ u_0 &= \tfrac{2}{3}\text{ integer} \\ &= \tfrac{2}{3}(n-m). \end{aligned} \tag{3.10}$$

It will suffice if n and m range over the nonnegative integers. Again specializing to the case $t = 0$, it can be seen from Eq. (3.9) that

$$j \geq \tfrac{1}{2}|n-m| + \tfrac{1}{2} \tag{3.11}$$

so that for fixed $u = u_0$ and $t = 0$, j must take on all values from $\frac{1}{2}|n-m| + \frac{1}{2}$ by integer steps. Different values of j of course specify different irreducible representations of $SU(3)$. Therefore the choice

$$j = \tfrac{1}{2}(n+m) + \tfrac{1}{2} \tag{3.12}$$

is permitted because equality in Eq. (3.11) results by choosing the appropriate one of n or m to be zero and because for fixed u_0, j increases by steps of unity as n and m are varied. The two nonnegative integers n and m specify an irreducible representation of $SU(3)$ by specifying u_0 and j. In fact, with the considerable help of hindsight, n and m have been chosen to be exactly the same as used by Pursey.[2]

Let $a(n, m)$ be some number which is determined shortly. Then

$$\begin{aligned}\psi^{n,m}_{t\mu,u} &= (2t+1)^{\frac{1}{2}} \\ &\quad\times a(n,m)\csc\nu\, d^{\frac{1}{2}(n+m+1)}_{\frac{1}{6}(n-m+3u+6t+3),\frac{1}{6}(n-m+3u-6t-3)}(2\nu) \\ &\quad\times d^t_{\mu,\frac{1}{2}u-\frac{1}{3}(n-m)}(\alpha_2)e^{(i/3)(n-m)\gamma}e^{-i\alpha\mu_3}e^{-i(\beta+\frac{1}{2}\gamma)u}.\end{aligned} \tag{3.13}$$

The factor $(2t+1)^{\frac{1}{2}}$ in Eq. (3.13) is required because the $D(G)$ are Hermitian in the scalar product $(\psi^{n'm'}_{t'\mu',u'}, \psi^{nm}_{t\mu,u})$ defined by integrating $\psi^{n'm'*}_{t'\mu',u'}\psi^{nm}_{t\mu,u}$ over 6π in β, 4π in α_3, α_2, and γ, and 2π in ν with weight function $\sin\alpha_2\sin^2\nu\sin 2\nu$. The scalar product vanishes unless the two sets of labels are the same. Since all unitary transformations within a given representation can be generated by Hermitian combinations of the differential operators, the scalar product must be independent of t, μ, and u. By comparing these functions with those of Beg and Ruegg, it can be seen that apart

[4] D. Brink and G. Satchler, *Angular Momentum* (Oxford University Press, Fair Lawn, New Jersey, 1962).

from trivial differences, they are the same if the following interpretation is made:

$$\begin{aligned}\beta &= \tfrac{3}{2}\phi_1,\\ \alpha_3 &= \phi_2 - \phi_3,\\ \alpha_2 &= 2\xi,\\ \gamma &= -\phi_1 - \phi_2 - \phi_3,\\ \nu &= \theta. \end{aligned} \tag{3.14}$$

Here ϕ_1, ϕ_2, ϕ_3, ξ, and θ are the variables used by Beg and Ruegg. Therefore, the functions derived by these authors are essentially the matrix elements of the form of Eq. (1.9).

Because T goes to unity when the arguments of ψ go to zero, $a(n, m)$ is found by imposing the condition

$$\psi^{nm}_{t\mu,u} \xrightarrow[\arg\to 0]{} \delta_{t,0}\delta_{\mu,0}\delta_{u,u_0}. \tag{3.15}$$

It is a simple matter to show that

$$\psi^{nm}_{t\mu,u} \xrightarrow[\arg\to 0]{} a(n, m)(-)\delta_{t,0}[(n+1)(m+1)]^{\frac{1}{2}} \times \delta_{\mu,0}\delta_{\frac{1}{2}u-\frac{1}{2}(n-m),0}; \tag{3.16}$$

hence

$$a(n, m) = -[(n+1)(m+1)]^{-\frac{1}{2}}. \tag{3.17}$$

As a bonus of this work, the isotopic spin-hypercharge content of all the irreducible representations of $SU(3)$ is specified by the form of the matrix element Eq. (3.13). The details of this are presented in the next section.

4. IRREDUCIBLE REPRESENTATIONS OF $SU(3)$

To see the content of the representations it is convenient to introduce two auxiliary numbers k and l defined by

$$k = \tfrac{1}{3}n + \tfrac{2}{3}m, \quad l = \tfrac{2}{3}n + \tfrac{1}{3}m. \tag{4.1}$$

From Eqs. (3.10) and (3.12) it can be seen that

$$u_0 = 2(l-k), \quad j = \tfrac{1}{2}(l+k+1). \tag{4.2}$$

Now by inspection of Eqs. (3.7) and (3.9), it is found for fixed t, u increases by steps of two from its minimum to its maximum value, whereas for fixed u, t goes by steps of one between its limits. Equations (3.7) and (3.9) imply the set of inequalities

$$\begin{aligned} -t \le \tfrac{1}{2}u - (l-k) \le +t,\\ -\tfrac{1}{2}(l+k+1) \le \tfrac{1}{2}(u+l-k) + t + \tfrac{1}{2} \le \tfrac{1}{2}(l+k+1),\\ -\tfrac{1}{2}(l+k+1) \le \tfrac{1}{2}(u+l-k) - t - \tfrac{1}{2} \le \tfrac{1}{2}(l+k+1). \end{aligned} \tag{4.3}$$

By using Eqs. (4.1) and (4.3) it can be shown that, for fixed t,

$$\max\{2(l-k-t), 2(t-l)\} \le u \le \min\{2(k-t), 2(l-k+t)\} \tag{4.4}$$

and for fixed u,

$$|l-k-\tfrac{1}{2}u| \le t \le \min\{(k-\tfrac{1}{2}u), (l+\tfrac{1}{2}u)\}. \tag{4.5}$$

Moreover, Eqs. (3.2), (3.8), and (4.2) imply that

$$\begin{aligned} c(2) &= \tfrac{1}{4}(l+k+1)(l+k+3) + \tfrac{3}{4}(l-k)^2 - \tfrac{3}{4}\\ &= \tfrac{1}{4}(k+l)(k+l+4) + \tfrac{3}{4}(k-l)^2 \end{aligned} \tag{4.6}$$

and

$$\begin{aligned} c(3) &= 2(k-l)\{1 - (k-l)^2 + \tfrac{1}{4}(k+l) \times (k+l+4) + \tfrac{3}{4}(k-l)^2\}\\ &= 2(k-l)(k+1)(l+1). \end{aligned} \tag{4.7}$$

All of the results found in this section can be found in the same notation in the paper by Pursey.[2] However, the method used here has rederived them except for the stipulation that the existence of an isotopic spin singlet component in each irreducible representation of $SU(3)$ was assumed, not proved. If the present treatment is to be made wholly independent of other treatments, this information would have to be obtained in some other way.

ACKNOWLEDGMENTS

It is a pleasure to thank Professor R. H. Good, Jr., for suggesting this investigation and to Professor D. L. Pursey for his help and encouragement.

This work was performed in the Ames Laboratory of the United States Atomic Energy Commission. Contribution No. 1965.

APPENDIX

The details which were used repeatedly in this work are presented in Tables I and II. These are the commutation rules of the generators of $SU(3)$ and the result of a finite transformation of any generator by any other generator. It is hoped that these tables may be of use in the solution of the general problem.

TABLE I. The commutation rules of $SU(3)$.

G	$[G, T_1]$	$[G, T_2]$	$[G, T_3]$	$[G, U]$	$[G, N]$	$[G, N']$	$[G, M]$	$[G, M']$
T_1		iT_3	$-iT_2$	0	$\frac{i}{2}M'$	$-\frac{i}{2}M$	$\frac{i}{2}N'$	$-\frac{i}{2}N$
T_2			iT_1	0	$\frac{i}{2}M$	$\frac{i}{2}M'$	$-\frac{i}{2}N$	$-\frac{i}{2}N'$
T_3				0	$\frac{i}{2}N'$	$-\frac{i}{2}N$	$-\frac{i}{2}M'$	$\frac{i}{2}M$
U					iN'	$-iN$	iM'	$-iM$
N						$i(2T_3+3U)$	$i2T_2$	$i2T_1$
N'							$-i2T_1$	$i2T_2$
M								$i(-2T_3+3U)$
M'								

TABLE II. The finite transformation of any generator by any other generator.

G	$e^{i\alpha_1 T_1} G e^{-i\alpha_1 T_1}$	$e^{i\alpha_2 T_2} G e^{-i\alpha_2 T_2}$	$e^{i\alpha_3 T_3} G e^{-i\alpha_3 T_3}$	$e^{i\beta U} G e^{-i\beta U}$	$e^{i\nu N} G e^{-i\nu N}$	$e^{i\nu' N'} G e^{-i\nu' N'}$	$e^{i\mu M} G e^{-i\mu M}$	$e^{i\mu' M'} G e^{-i\mu' M}$
T_1	T_1	$\cos\alpha_2 T_1 + \sin\alpha_2 T_3$	$\cos\alpha_3 T_1 - \sin\alpha_3 T_2$	T_1	$\cos\nu T_1 + \frac{1}{2}\sin\nu M'$	$\cos\nu' T_1 - \frac{1}{2}\sin\nu' M$	$\cos\mu T_1 + \frac{1}{2}\sin\mu N'$	$\cos\mu' T_1 - \frac{1}{2}\sin\mu' N$
T_2	$\cos\alpha_1 T_2 - \sin\alpha_1 T_3$	T_2	$\cos\alpha_3 T_2 + \sin\alpha_3 T_1$	T_2	$\cos\nu T_2 + \frac{1}{2}\sin\nu M$	$\cos\nu' T_2 + \frac{1}{2}\sin\nu' M'$	$\cos\mu T_2 - \frac{1}{2}\sin\mu N$	$\cos\mu' T_2 - \frac{1}{2}\sin\mu' N'$
T_3	$\cos\alpha_1 T_3 - \sin\alpha T_2$	$\cos\alpha_2 T_3 - \sin\alpha_2 T_1$	T_3	T_3	$\frac{1}{4}(3+\cos 2\nu)T_3 - \frac{3}{8}(1-\cos 2\nu)U + \frac{1}{4}\sin 2\nu N'$	$\frac{1}{4}(3+\cos 2\nu')T_3 - \frac{3}{8}(1-\cos 2\nu')U - \frac{1}{4}\sin 2\nu' N$	$\frac{1}{4}(3-\cos 2\mu)T_3 + \frac{3}{8}(1-\cos 2\mu)U - \frac{1}{4}\sin 2\mu M'$	$\frac{1}{4}(3+\cos 2\mu')T_3 + \frac{3}{8}(1-\cos 2\mu')U + \frac{1}{4}\sin 2\mu' M$
U	U	U	U	U	$\frac{1}{4}(1+3\cos 2\nu)U - \frac{1}{2}(1-\cos 2\nu)T_3 + \frac{1}{2}\sin 2\nu N'$	$\frac{1}{4}(1+3\cos 2\nu')U - \frac{1}{2}(1-\cos 2\nu')T_3 - \frac{1}{2}\sin 2\nu' N$	$\frac{1}{4}(1+3\cos 2\mu)U + \frac{1}{2}(1-\cos 2\mu)T_3 + \frac{1}{2}\sin 2\mu M'$	$\frac{1}{4}(1+3\cos 2\mu')U + \frac{1}{2}(1-\cos 2\mu')T_3 - \frac{1}{2}\sin 2\mu' M$
N	$\cos\frac{\alpha_1}{2}N - \sin\frac{\alpha_1}{2}M'$	$\cos\frac{\alpha_2}{2}N - \sin\frac{\alpha_2}{2}M$	$\cos\frac{\alpha_3}{2}N - \sin\frac{\alpha_3}{2}N'$	$\cos\beta N - \sin\beta N'$	N	$\cos 2\nu' N + \sin 2\nu' T_3 + \frac{3}{2}\sin 2\nu' U$	$\cos\mu N + 2\sin\mu T_2$	$\cos\mu' N + 2\sin\mu' T_1$
N'	$\cos\frac{\alpha_1}{2}N' + \sin\frac{\alpha_1}{2}M$	$\cos\frac{\alpha_2}{2}N' - \sin\frac{\alpha_2}{2}M'$	$\cos\frac{\alpha_3}{2}N' + \sin\frac{\alpha_3}{2}N$	$\cos\beta N' + \sin\beta N$	$\cos 2\nu N' - \sin 2\nu T_3 - \frac{3}{2}\sin 2\nu U$	N'	$\cos\mu N' - 2\sin\mu T_1$	$\cos\mu' N' + 2\sin\mu' T_2$
M	$\cos\frac{\alpha_1}{2}M - \sin\frac{\alpha_1}{2}N'$	$\cos\frac{\alpha_2}{2}M + \sin\frac{\alpha_2}{2}N$	$\cos\frac{\alpha_3}{2}M + \sin\frac{\alpha_3}{2}M'$	$\cos\beta M - \sin\beta M'$	$\cos\nu M - 2\sin\nu T_2$	$\cos\nu' M + 2\sin\nu' T_1$	M	$\cos 2\mu' M - \sin 2\mu' T_3 + \frac{3}{2}\sin 2\mu' U$
M'	$\cos\frac{\alpha_1}{2}M' + \sin\frac{\alpha_1}{2}N$	$\cos\frac{\alpha_2}{2}M' + \sin\frac{\alpha_2}{2}N'$	$\cos\frac{\alpha_3}{2}M' - \sin\frac{\alpha_3}{2}M$	$\cos\beta M' + \sin\beta M$	$\cos\nu M' - 2\sin\nu T_1$	$\cos\nu' M' - 2\sin\nu' T_2$	$\cos 2\mu M' - \sin 2\mu T_3 - \frac{3}{2}\sin 2\mu U$	M'

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Necessary Conditions on Radial Distribution Functions

F. Y. Wu
Virginia Polytechnic Institute, Blacksburg, Virginia

H. T. Tan* and E. Feenberg
Arthur Holly Compton Laboratory of Physics, Washington University, Saint Louis, Missouri

(Received 19 September 1966)

Certain necessary conditions must be met by any function introduced to serve as a radial distribution function of a uniform N-particle system. One large class of necessary conditions is based on the statement that the expectation value of a potential energy (for an arbitrary potential function between pairs of particles) cannot fall below the classical minimum potential energy of the system. To convert this statement into a family of useful inequalities, we have evaluated the classical potential energies of close-packed crystals for a linear combination of Yukawa and Coulomb two-particle interactions. The numerical evaluation of lattice sums is performed by two procedures: (1) direct summation over the lattice (suitable for short-range potentials), and (2) an adaptation of the Ewald summation procedure suitable for long-range potentials. Results are given for the Coulomb and Yukawa potentials and also for the potentials $1/r(r+a)$ and $1/r(r+a)^2$. Two simple approximate forms are developed both giving close lower bounds on the classical potential energy of interacting particles forming a regular crystalline lattice.

1. INTRODUCTION

A GENERAL problem in the quantum theory of many-particle systems is the characterization of admissible density matrices and distribution functions.[1,2]. One aspect of this problem is the development of necessary conditions on the radial distribution function of an extended uniform system.[2,3] In this study we determine a family of necessary conditions based on the evaluation of the minimum classical potential energy of a system of particles interacting in pairs through a linear combination of Coulomb and Yukawa potentials. To begin, the physical system is specified as N particles in a cubical box of volume Ω subject to the limiting process $N, \Omega \to \infty$, while $\rho = N/\Omega$ is held constant. The given Hamiltonian operator H and periodic boundary conditions complete the specification. Let $\psi(1, 2, \cdots, N)$ represent a normalized trial function subject to the constraint $\mathbf{P}\psi = 0$ (an eigenstate of the center-of-mass momentum with the eigenvalue $\mathbf{P}' = 0$). The two-particle distribution function $p^{(2)}(1, 2)$ for the pure state represented by ψ is the positive definite form

$$p^{(2)}(1, 2) = N(N-1)\int |\psi|^2 \, dv_3 \cdots_N, \qquad (1)$$

the integration including summation over all discrete (spin and iso-spin) coordinates when these are present. The radial distribution function $g(r)$ is introduced by writing $p^{(2)}(1, 2) = \rho^2 g(r_{12})$ and neglecting the slight dependence of $p^{(2)}$ on the direction of $\mathbf{r}_{12}$. The condition $\mathbf{P}' = 0$ has the consequence that $p^{(2)}(1, 2)$ is a function of $\mathbf{r}_{12}$ only. This can be seen by expressing ψ as a $3N$-dimensional Fourier series (periodic in the fundamental cube).

Let $\mathbf{k}$ represent an allowed wave vector defined by the periodic boundary condition. For $k \neq 0$, the liquid structure function is defined as a positive definite form by

$$\begin{aligned} S(k) &= \frac{1}{N}\int |\psi|^2 \left|\sum_l e^{i\mathbf{k}\cdot\mathbf{r}_l}\right|^2 dv_{12\cdots N} \\ &= 1 + \frac{\rho^2}{N}\int g(r_{12}) e^{i\mathbf{k}\cdot\mathbf{r}_{12}} \, d\mathbf{r}_1 \, d\mathbf{r}_2 \\ &= 1 + \rho\int [g(r) - g(\infty)] e^{i\mathbf{k}\cdot\mathbf{r}} \, d\mathbf{r}. \end{aligned} \qquad (2)$$

The third line of Eq. (2) involves the asymptotic limit of $g(r)$ as $r \to \infty$ and $N \to \infty$. The proof that $g(\infty) = 1$, or more precisely $\lim_{N\to\infty} |N[1 - g(\infty)]| \ll 1$, can be developed by relating $S(k)$ in a qualitative manner to fluctuations in the number of particles found in a suitably defined half space within Ω.[4] A more precise derivation follows from three well-known sum rules[5] which have the consequence, $S(k) \leq \hbar k/2mC$ (here C is the velocity of 1st sound at absolute zero and N is assumed infinite). Thus the sum rules imply

$$\lim_{k\to 0} [\lim_{N\to\infty} S(k)] = 0. \qquad (3)$$

Equation (1) imposes the normalization condition

$$\frac{\rho}{\Omega}\int [g(r_{12}) - 1] \, d\mathbf{r}_1 \, d\mathbf{r}_2 = -1 \qquad (4)$$

or

$$\rho\int [g(r) - g(\infty)] \, d\mathbf{r} + N[g(\infty) - 1] = -1. \qquad (5)$$

* Washington University Fellow.

[1] A. J. Coleman, Rev. Mod. Phys. 35, 690 (1963).

[2] C. Garrod and J. K. Percus, J. Math. Phys. 5, 1756 (1964).

[3] M. Yamada, Progr. Theoret. Phys. (Kyoto) 25, 579 (1961).

[4] E. Feenberg, in *Lectures in Theoretical Physics* (University of Colorado Press, Boulder, Colorado, 1965), Vol. VIIC.

[5] P. J. Price, Phys. Rev. 94, 257 (1954).

Equations (2), (3), and (5) require

$$\lim_{N\to\infty} |N[1 - g(\infty)]| = 0. \tag{6}$$

Through the sum rules, this argument involves the restriction that ψ is the correct ground-state eigenfunction ψ_0. The fluctuation argument establishes the less precise result

$$\lim_{N\to\infty} |N[1 - g(\infty)]| \ll 1,$$

but requires only that ψ exhibit reasonably well the essential physical structure expected of the ground-state eigenfunction when the interactions are strongly repulsive for small separations of the particles.

The relation[6,7]

$$S(k) = \hbar k/2mC, \quad k \ll \rho^{\frac{1}{3}} \tag{7}$$

is expected to hold in liquid He^4. This behavior has implications for the manner in which $g(r)$ approaches its asymptotic value as r increases without limit. Equation (2) with $g(\infty) = 1$ can be expressed in the equivalent form

$$S(k) = 4\pi\rho\int_0^\infty \left(\frac{\sin kr}{kr} - 1\right)[g(r) - 1]r^2\,dr$$
$$= \frac{4\pi\rho}{k^3}\int_0^\infty \left(\frac{\sin x}{x} - 1\right)\left[g\left(\frac{x}{k}\right) - 1\right]x^2\,dx. \tag{8}$$

The correct dependence on k as $k \to 0$ occurs only if

$$\overline{g(r)} \simeq 1 - b/r^4, \quad r \gg \bar{\rho}^{\frac{1}{3}}, \tag{9}$$

the bar denoting an average over a small range of r values ($\delta r \sim \bar{\rho}^{\frac{1}{3}}$). Introducing Eq. (7) into Eq. (8) for $k \ll \rho^{\frac{1}{3}}$, we find

$$\frac{\hbar k}{2mc} = 4\pi\rho bk\int_0^\infty \left(1 - \frac{\sin x}{x}\right)\frac{dx}{x^2} \tag{10}$$

and[4,8]

$$b = \frac{\hbar}{2\pi^2\rho mc}. \tag{11}$$

A broad class of necessary conditions on $g(r)$ are defined by the statement[2]

$$N(N-1)\int |\psi|^2 v(\mathbf{r}_1, \mathbf{r}_2)\,dv_{12\ldots N} \geq \min \sum_{i\neq j} v(\mathbf{R}_i, \mathbf{R}_j), \tag{12}$$

in which $v(\mathbf{R}_i, \mathbf{R}_j)$ is a real integrable function and the points $\mathbf{R}_1, \mathbf{R}_2, \cdots$ are all distinct, all located within Ω (or on the boundary of Ω), and otherwise arbitrary. Equation (12) implies

$$\rho^2\int g(r_{12})\,v(\mathbf{R}_1, \mathbf{R}_2)\,d\mathbf{r}_1\,d\mathbf{r}_2 \geq \min \sum_{i\neq j} v(\mathbf{R}_i, \mathbf{R}_j) \tag{13}$$

and also

$$\rho^2\int [g(r_{12}) - 1]v(\mathbf{r}_1, \mathbf{r}_2)\,d\mathbf{r}_1\,d\mathbf{r}_2$$
$$\geq \min\left[\sum_{i\neq j} v(\mathbf{R}_i, \mathbf{R}_j) - \rho^2\int v(\mathbf{r}, \mathbf{r}')\,d\mathbf{r}\,d\mathbf{r}'\right]. \tag{14}$$

These inequalities can be given an immediate physical interpretation. Suppose $v(\mathbf{r}_1, \mathbf{r}_2) = v(r_{12})$ is a potential between particles at points $\mathbf{r}_1$ and $\mathbf{r}_2$, Eqs. (12) and (13) state that the expectation value of

$$V = \sum_{i<j} v(R_{ij})$$

with respect to the normalized state function ψ exceeds the minimum possible value of V (assumed to exist for a definite configuration $\mathbf{R}_1, \mathbf{R}_2, \cdots$). Suppose further that $\int v(\mathbf{R}_i - \mathbf{r})\,d\mathbf{r}$ is independent of the location of $\mathbf{R}_i$ except for points in a negligible fraction of the total volume near the surface; then the left- and right-hand members of Eq. (14) represent two ways of computing the potential energy (doubled) of a system of particles immersed in a uniform charged background of equal total strength and opposite sign. The energy expended to assemble the background charge against the interaction is included. Here we postulate that $-\rho v(\mathbf{r} - \mathbf{r}')\,d\mathbf{r}'$ is the potential energy of a particle at $\mathbf{r}$ interacting with the uniform background charge in the volume element $d\mathbf{r}'$. The corresponding potential for the mutual interaction of charges in elements $d\mathbf{r}_1$ and $d\mathbf{r}_2$ is $\rho^2 v(r_{12})\,d\mathbf{r}_1\,d\mathbf{r}_2$.

Results for the Coulomb potential $v(r) = 1/r$ are known from calculations of the minimum potential energy of a system of electrically charged particles.[9,10] In this case Eq. (14) reduces to the explicit form[11]

$$4\pi\rho\int_0^\infty [1 - g(r)]r\,dr \leq 1.792\left(\frac{4\pi\rho}{3}\right)^{\frac{1}{3}}, \tag{15}$$

or, in conjunction with the normalization condition on $g(r)$,

$$\int_0^\infty [1 - g(r)]r\,dr\left[\int_0^\infty [1 - g(r)]r^2\,dr\right]^{-\frac{2}{3}} \leq \frac{1.792}{3^{\frac{1}{3}}} = 1.243. \tag{16}$$

The object of the present study is to derive additional relations of the above type employing the Yukawa potential and linear combinations of Yukawa potentials:

$$v(r) = \sum_l C_l \frac{e^{-\mu_l r}}{r} \tag{17}$$

and

$$v(r) = \int_0^\infty C(\mu)\frac{e^{-\mu r}}{r}\,d\mu. \tag{18}$$

2. EVALUATION OF LATTICE SUMS INVOLVING THE YUKAWA POTENTIAL

The analysis begins with the classical potential energy

$$V(\mathbf{R}_1, \mathbf{R}_2, \cdots, \mathbf{R}_N) = \frac{1}{2}\sum_{i\neq j}\frac{e^{-\mu R_{ij}}}{R_{ij}}$$
$$+ \tfrac{1}{2}\rho^2\int\frac{e^{-\mu|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}\,d\mathbf{r}\,d\mathbf{r}' - \sum_i \rho\int\frac{e^{-\mu|\mathbf{R}_i-\mathbf{r}|}}{|\mathbf{R}_i-\mathbf{r}|}\,d\mathbf{r}, \tag{19}$$

[6] R. P. Feynman and M. Cohen, Phys. Rev. **102**, 1189 (1956).
[7] K. Huang and A. Klein, Ann. Phys. (N.Y.) **30**, 203 (1964).
[8] J. E. Enderby, T. Gaskell, and N. H. March, Proc. Phys. Soc. (London) **85**, 217 (1965).
[9] E. P. Wigner, Trans. Faraday Soc. **34**, 678 (1938).
[10] K. Fuchs, Proc. Roy. Soc. (London) **A151**, 585 (1935).
[11] E. Feenberg, J. Math. Phys. **6**, 658 (1965).

TABLE I. Lattice energies for the Yukawa potential.

$\beta = \mu r_s$	Lattice type	$-\frac{2r_s}{N} E_a^*$	$-\frac{2r_s}{N} V_a^*$	$-\frac{2r_s}{N} V_b$	$-\frac{2r_s}{N} V_c$	$-\frac{2r_s}{N} V$
	sc			−0.0535	−0.0470	1.7606
0.0	bcc	1.8000	1.8611	−0.0363	−0.0331	1.7917
	fcc			−0.0367	−0.0328	1.7916
	hcp					
	sc			−0.0500	−0.0455	1.31339
0.5	bcc	1.3578	1.4089	−0.0340	−0.0317	1.34334
	fcc			−0.0344	−0.0313	1.34326
	hcp					1.34320
	sc			−0.0409	−0.0401	0.96817
1.0	bcc	1.0215	1.0492	−0.0281	−0.0281	0.99341
	fcc			−0.0283	−0.0275	0.99340
	hcp					0.99335
	sc			−0.0293	−0.0326	0.71112
1.5	bcc	0.7709	0.7731	−0.0204	−0.0225	0.73028
	fcc			−0.0206	−0.0222	0.73035
	hcp					0.73032
	sc			−0.0186	−0.0244	0.52478
2.0	bcc	0.5860	0.5677	−0.0131	−0.0167	0.53805
	fcc			−0.0132	−0.0164	0.53816
	hcp					0.53814
	sc			−0.0105	−0.0169	0.39174
2.5	bcc		0.4191	−0.0075	−0.0115	0.40024
	fcc			−0.0076	−0.0113	0.40036
	hcp					0.40035

* These are independent of lattice type.

in which $\mathbf{R}_1, \mathbf{R}_2, \cdots, \mathbf{R}_N$ are a set of lattice points occupying the entire box. One point coincides with the origin at the center. Calculations are made for (i) $\mu > 0$ and $\mu\Omega^{\frac{1}{3}} \gg 1$ and (ii) $\mu = 0$. The first condition is permitted by the limiting process $\Omega, N \to \infty$, subject to ρ held constant for any $\mu > 0$. Thus for $\mu > 0$ the formula for V may be simplified to

$$V(\mathbf{R}_1, \mathbf{R}_2, \cdots, \mathbf{R}_N) = \tfrac{1}{2}N\left[\sum_j{}' \frac{e^{-\mu R_j}}{R_j} - \frac{4\pi\rho}{\mu^2}\right] \quad (20)$$

the primed summation excluding $R_j = 0$.

The direct evaluation of the sum in Eq. (20) is feasible when $\mu\bar{\rho}^{\frac{1}{3}}$ is sufficiently large (calculations have been made for $\mu\bar{\rho}^{\frac{1}{3}} \geq 0.8$). Also a close lower bound on $V(R_1, \cdots, R_N; \mu)$ can be derived by a simple argument based on the physical picture of N-point charges immersed in a uniform background charge of equal total strength but with opposite sign. Each point charge can be pictured at the center of its own s sphere of radius $r_s = (3/4\pi\rho)^{\frac{1}{3}}$. Adjacent s spheres overlap slightly, but in the following nonrigorous argument the overlap is disregarded. Since the conclusions are verified by accurate numerical calculations, the lack of rigor does not matter. In terms of the s spheres the energy of the system can be split into two parts: the self energy E_a of the charges within the N s spheres and the interaction energy E_b of these spheres.[9] As shown in Appendix A, E_a is easily computed [Eq. (A3)] and E_b is found to be positive. The inference appears safe that E_a is a lower bound to the energy $V(R_1, \cdots, R_N; \mu)$. This is confirmed by the explicit computation shown in Table I.

Ewald's summation procedure[12] is easily adapted to the Yukawa potential and offers the advantage of rapid convergence for small μ. It also provides a simple approximate formula giving a moderately close rigorous upper limit on the lattice sum.

The derivation of the Ewald form is given in Appendix B. Some essential information on the basis vectors (in both lattice and dual spaces) and the numerical values of certain derived quantities are listed in Table II. This information is needed to connect the mathematical analysis with the numerical

TABLE II. Basis vectors and nearest-neighbor distance.

Lattice type	sc	fcc	bcc
$\mathbf{a}_1$	$\rho^{-\frac{1}{3}}(1, 0, 0)$	$(2\rho)^{-\frac{1}{3}}(1, 1, 0)$	$(4\rho)^{-\frac{1}{3}}(1, 1, -1)$
$\mathbf{a}_2$	$\rho^{-\frac{1}{3}}(0, 1, 0)$	$(2\rho)^{-\frac{1}{3}}(0, 1, 1)$	$(4\rho)^{-\frac{1}{3}}(-1, 1, 1)$
$\mathbf{a}_3$	$\rho^{-\frac{1}{3}}(0, 0, 1)$	$(2\rho)^{-\frac{1}{3}}(1, 0, 1)$	$(4\rho)^{-\frac{1}{3}}(1, -1, 1)$
$\mathbf{b}_1$	$\rho^{\frac{1}{3}}(1, 0, 0)$	$(\rho/4)^{\frac{1}{3}}(1, 1, -1)$	$(\rho/2)^{\frac{1}{3}}(1, 1, 0)$
$\mathbf{b}_2$	$\rho^{\frac{1}{3}}(0, 1, 0)$	$(\rho/4)^{\frac{1}{3}}(-1, 1, 1)$	$(\rho/2)^{\frac{1}{3}}(0, 1, 1)$
$\mathbf{b}_3$	$\rho^{\frac{1}{3}}(0, 0, 1)$	$(\rho/4)^{\frac{1}{3}}(1, -1, 1)$	$(\rho/2)^{\frac{1}{3}}(1, 0, 1)$
R_n^*	$\rho^{-\frac{1}{3}}$	$2^{\frac{1}{2}}\rho^{-\frac{1}{3}} = 1.12\rho^{-\frac{1}{3}}$	$\sqrt{3}(4\rho)^{-\frac{1}{3}} = 1.09\rho^{-\frac{1}{3}}$

* Nearest-neighbor distance.

[12] P. P. Ewald, Ann. Physik 54, 519, 557 (1917).

evaluation for simple cubic, face-centered cubic, and body-centered cubic lattices.

The results of Appendix B can be summarized in the formulas

$$V = V_a + V_b + V_c,$$

$$\frac{2}{N}V_a = -\frac{2x_0}{\sqrt{\pi}}\exp\left[-\left(\frac{\mu}{2x_0}\right)^2\right] + \mu\left[1 - E\left(\frac{\mu}{2x_0}\right)\right] + \frac{4\pi\rho}{\mu^2}\left[\exp\left[-\left(\frac{\mu}{2x_0}\right)^2\right] - 1\right],$$

$$\frac{2}{N}V_b = \pi\rho\sum_{\mathbf{p}\neq 0}\frac{\exp\left[-\frac{1}{x_0^2}(\pi^2 h_{\mathbf{p}}^2 + \tfrac{1}{4}\mu^2)\right]}{\pi^2 h_{\mathbf{p}}^2 + \tfrac{1}{4}\mu^2}, \tag{21}$$

$$\frac{2}{N}V_c = \sum_{l\neq 0}\frac{1}{2R_l}\left\{e^{-\mu R_l}\left[1 - E\left(x_0 R_l - \frac{\mu}{2x_0}\right)\right] + e^{\mu R_l}\left[1 - E\left(x_0 R_l + \frac{\mu}{2x_0}\right)\right]\right\}.$$

Both V_b and V_c are positive valued functions; consequently, $V_a < V$ supplies a lower limit on V independent of the lattice type. We choose x_0 to make V_a as large as possible, thus simultaneously minimizing the contributions from $V_b + V_c$. The maximum occurs at $x_0 = \pi^{\frac{1}{2}}\rho^{\frac{1}{3}}$, independent of μ, and has the value

$$V_{a,\max} = -N\rho^{\frac{1}{3}}[1/2p^2 + (1 - 1/2p^2)e^{-p^2} - \pi^{\frac{1}{2}}pE(p)], \tag{22}$$

with

$$p = \mu/2\pi^{\frac{1}{2}}\rho^{\frac{1}{3}} = \mu r_s/6^{\frac{1}{3}}\pi^{\frac{1}{6}} = 0.4547\mu r_s. \tag{23}$$

Numerical results for V and the several partial V's appear in Table I as functions of the parameter $\beta = \mu r_s$ and the lattice type. Some results for the hexagonal close-packed lattice are included.

To introduce these results into Eq. (14) let

$$U(\beta) \equiv -\frac{2r_s}{N}\min\,[V(R_1, \cdots R_N; \mu)], \tag{24}$$

then

$$(4\pi\rho)^{\frac{2}{3}}\int_0^\infty [1 - g(r)]\frac{e^{-\mu r}}{r}r^2\,dr \leq \frac{1}{3^{\frac{1}{3}}}U(\beta). \tag{25}$$

The normalization condition can be used to convert Eq. (25) into

$$\int_0^\infty [1 - g(r)]\frac{e^{-\mu r}}{r}r^2\,dr\left[\int_0^\infty [1 - g(r)]r^2\,dr\right]^{-\frac{2}{3}} \leq \frac{1}{3^{\frac{1}{3}}}U(\beta). \tag{26}$$

Equations (25) and (26) remain valid for linear combinations of Yukawa potentials as in Equations (17) and (18) subject to the constraint $C_l \geq 0$ or $C(\mu) \geq 0$. In particular, the distribution $C(\mu) = e^{-\mu a}$ produces a potential intermediate between the inverse distance and the inverse square of the distance:

$$v_1(r) = 1/r(r + a). \tag{27}$$

Equation (25) is replaced by

$$(4\pi\rho)^{\frac{1}{3}}\int_0^\infty [1 - g(r)]\frac{1}{r(r+a)}r^2\,dr \leq \frac{1}{3^{\frac{2}{3}}}U_1\left(\frac{a}{r_s}\right), \qquad U_1(x) = \int_0^\infty e^{-\beta x}U(\beta)\,d\beta, \tag{28}$$

and Eq. (26) by

$$\int_0^\infty [1 - g(r)]\frac{1}{r(r+a)}r^2\,dr\left[\int_0^\infty [1 - g(r)]r^2\,dr\right]^{-\frac{1}{3}} \leq \frac{1}{3^{\frac{2}{3}}}U_1\left(\frac{a}{r_s}\right). \tag{29}$$

The process of using one potential to generate another can be continued in many ways. In particular, the relation

$$v_2(r) = -\frac{d}{da}\frac{1}{r(r+a)} = \frac{1}{r(r+a)^2} \tag{30}$$

leads to

$$\int_0^\infty [1 - g(r)]\frac{1}{r(r+a)^2}r^2\,dr \leq \tfrac{1}{3}U_2\left(\frac{a}{r_s}\right), \qquad U_2(x) = -\frac{d}{dx}U_1(x) = \int_0^\infty \beta e^{-\beta x}U(\beta)\,d\beta. \tag{31}$$

Numerical results for $U_1(\beta)$ and $U_2(\beta)$ appear in Table III. These functions are represented quite well by the simple formulas

$$U_1(x) \simeq f(x), \quad U_2(x) \sim -\frac{d}{dx}f(x),$$

$$f(x) = \frac{3.49}{1 + 1.948x} + \frac{x\ln x}{1 + 19.33x - 20.33x^3}. \tag{32}$$

Equation (16) is recovered from Eq. (28) by letting $a \to \infty$; this yields

$$\lim_{x\to\infty} x\,U_1(x) = 1.792 \tag{33}$$

in agreement with the limiting value given by the interpolation formula of Eq. (32).

TABLE III. Lattice energies for the potentials v_1 and v_2 (bcc or fcc).

$x = a/r_s$	$U_1(x)$	$U_2(x)$
0.0	3.490	∞
0.25	2.308	3.06
0.50	1.718	1.55
0.75	1.382	1.00
1.00	1.160	0.78
1.50	0.878	0.46
2.00	0.704	0.29
2.50	0.590	0.19
3.00	0.514	0.13
4.00	0.402	0.08

A qualitative summary of the results in Table 1 is contained in the statement that the bcc structure has the lowest energy for small β and the fcc for large β with the crossover near $\beta = 1$. The fcc and hcp structures are remarkably close in energy.

3. THE ABSOLUTE MINIMUM POTENTIAL ENERGY

Some information is available on the stability of these structures against small deformations. All potentials considered here are positive-valued, monotonic decreasing functions of r with positive second derivatives. These properties occur as a special case under the general stability condition found by Born[13,14] for small, slowly varying disturbances. Powers[15] showed that Born's conditions are sufficient to ensure stability against arbitrary small disturbances subject to the limitation that only nearest neighbors need be included in evaluating lattice sums. For the Yukawa potential this means $\beta = \mu r_s \gg 1$.

For $\beta = 0$ stability is proved by the actual numerical evaluation of the frequency spectrum.[16] In this case the fixed background charge (not present in the analysis of Born and Powers) is essential for stability.[17] In general ($\beta \neq 0$), the background charge contributes to stability by generating a restoring force for arbitrary small displacements of the particles from the lattice sites. Since stability is assured at $\beta = 0$ (more generally $\beta \ll 1$) and at $\beta \gg 1$, the smallness of the energy range covered by the three close-packed structures is, in itself, strong evidence for all values of β that the most stable close-packed structure either (i) realizes the configuration for the absolute energy minimum, or (ii) possesses an energy very close to the absolute minimum.

In any event the minimum computed energy establishes a safe upper limit on the allowed value of the integral occurring in Eq. (25); safe in the sense that one does not hesitate to discard any trial function $g(r)$ which violates the inequality. This is enough to make the inequality useful. Applications have been made in theoretical studies of liquid helium[18] and of an alpha-particle model of low-density nuclear matter.[19]

ACKNOWLEDGMENT

This research was supported in part by the National Science Foundation under Grant GP-6210.

APPENDIX A. LOWER BOUND ON $V(\mathbf{R}_1, \mathbf{R}_2, \cdots, \mathbf{R}_N; \mu)$

The self-energy E_a of an s sphere, a uniformly charged sphere plus a point charge at the center, contains two parts: the self energy E_{a1} of the uniform charge, and the interaction energy E_{a2} between the point charge and the uniform charge density. Direct calculation yields

$$E_{a1} = \frac{1}{2}\left(\frac{3}{4\pi r_s^3}\right)^2 \int_{\substack{r_1<r_s \\ r_2<r_s}} \frac{e^{-\mu r_{12}}}{r_{12}}\, d\mathbf{r}_1\, d\mathbf{r}_2$$

$$= \frac{3}{r_s}\int_0^1 e^{-2\mu r_s x}(2x - 3x^2 + x^4)\, dx, \tag{A1}$$

$$E_{a2} = -\frac{3}{4\pi r_s^3}\int_0^{r_s} e^{-\mu r} 4\pi r\, dr, \tag{A2}$$

and

$$E_a = N(E_{a1} + E_{a2}) = N\frac{3}{r_s}\left[\frac{3}{4\beta^5}(1 - \beta^2 - \tfrac{2}{3}\beta^3) - \frac{3}{4\beta^5}(1+\beta)^2 e^{-2\beta} + \frac{e^{-\beta}}{\beta^2}(1+\beta)\right] \tag{A3}$$

with $\beta = \mu r_s$.

The result for the Coulomb potential[20] is obtained in the limit of $\beta \to 0$:

$$\lim_{\beta\to 0} E_a = -N\frac{9}{10 r_s}. \tag{A4}$$

Next, we demonstrate that the interaction energy E_b of the s spheres is positive. It is well known that a spherically symmetric distribution of electric charge produces the same potential outside the distribution as if all the charge were concentrated at the center. A closely related conclusion actually holds also for the Yukawa potential. In fact, an elementary computation yields the potential produced by an s sphere (for definiteness, let the uniform charge be positive) at distance $r > r_s$ from its center as

$$V(r) = F(\beta)(e^{-\mu r}/r) \tag{A5}$$

with

$$F(\beta) = (3/\beta^2)(\cosh\beta - \beta^{-1}\sinh\beta) - 1 > 0. \tag{A6}$$

At all points outside the sphere, the charge distribution can be replaced by a point charge of strength $F(\beta)$ at the center. If one replaces all the s spheres by point charges of the same strength, clearly then $E_b > 0$.

As a further check, we observe that $F(\beta)$ increases monotonically as a function of β. We then expect a larger discrepancy in approximating the minimum

[13] M. Born and K. Huang, *Dynamical Theory of Crystal Lattices* (Oxford University Press, London, 1962), p. 142.
[14] A. A. Maradudin, E. W. Montroll, and G. W. Weiss, *Theory of Lattice Dynamics in the Harmonic Approximation* (Academic Press Inc., New York, 1963).
[15] S. C. Powers, Proc. Cambridge Phil. Soc. **38**, 62 (1942).
[16] C. B. Clark, Phys. Rev. **109**, 1133 (1958).
[17] E. W. Kellerman, Phil. Trans. Roy. Soc. (London) **238**, 513 (1940).
[18] W. E. Massey, Ph.D. thesis, Washington University (1966); Phys. Rev. **151**, 153 (1966).
[19] J. W. Clark and T. P. Wang, Ann. Phys. (N.Y.) **40**, 127 (1966).
[20] E. P. Wigner and F. Seitz, Phys. Rev. **46**, 509 (1934).

potential energy by E_a as β increases. This is indeed the case as seen from Table I.

APPENDIX B. EWALD PROCEDURE FOR THE YUKAWA POTENTIAL

The Ewald procedure involves two transformations.[12] The first is based on the integral

$$\frac{2}{\pi^{\frac{1}{2}}}\int_0^\infty e^{-(Rx-\mu/2x)^2}\,dx = \left(\frac{2\mu}{\pi R}\right)^{\frac{1}{2}}\int_0^\infty e^{-\frac{1}{2}\mu R(y-1/y)^2}dy. \tag{B1}$$

A change of variable $y = e^\theta$ yields $dy = d\sinh\theta + d\cosh\theta$; since $\sinh\theta$ is an odd function of θ the integral reduces to $1/R$. Consequently,

$$\frac{e^{-\mu R}}{R} = \frac{2}{\pi^{\frac{1}{2}}}\int_0^\infty e^{-(R^2x^2+\mu^2/4x^2)}\,dx. \tag{B2}$$

Equations (20) and (B2) combine to give

$$\frac{2V}{N} + \frac{4\pi\rho}{\mu^2} = \frac{2}{\pi^{\frac{1}{2}}}\int_0^\infty \sum_j{}' e^{-R_j^2x^2-\mu^2/4x^2}\,dx. \tag{B3}$$

Next, the range of integration is split into two, $0 \le x \le x_0$ and $x_0 < x < \infty$ with x_0 to be determined. Equation (23) assumes the form

$$\frac{2V}{N} + \frac{4\pi\rho}{\mu^2} = \frac{2}{\pi^{\frac{1}{2}}}\int_{x_0}^\infty \sum_j{}' e^{-R_j^2x^2-\mu^2/4x^2}\,dx + \frac{2}{\pi^{\frac{1}{2}}}\int_0^{x_0}\sum_j e^{-R_j^2x^2-\mu^2/4x^2}\,dx - \frac{2}{\pi^{\frac{1}{2}}}\int_0^{x_0} e^{-\mu^2/4x^2}\,dx. \tag{B4}$$

Also

$$\frac{2}{\pi^{\frac{1}{2}}}\int_0^{x_0} e^{-\mu^2/4x^2}\,dx = \frac{\mu}{\pi^{\frac{1}{2}}}\int_{\mu/2x_0}^\infty e^{-y^2}\frac{dy}{y^2} = \frac{2}{\pi^{\frac{1}{2}}}\,x_0e^{-(\mu/2x_0)^2} - \mu 1 - E(\mu/2x_0)], \tag{B5}$$

$$\frac{2}{\pi^{\frac{1}{2}}}\int_{x_0}^\infty e^{-R^2x^2-(\mu/2x)^2}\,dx = 2\left(\frac{\mu}{2\pi R}\right)^{\frac{1}{2}}e^{-\mu R}\int_{y_0}^\infty e^{-\frac{1}{2}\mu R(y-1/y)^2}\,dy$$

$$= 2\left(\frac{\mu}{2\pi R}\right)^{\frac{1}{2}}e^{-\mu R}\left[\int_{\theta_0}^\infty e^{-2\mu R(\sinh\theta)^2}\,d\sinh\theta + \int_{\theta_0}^\infty e^{-2\mu R((\cosh\theta)^2-1)}\,d\cosh\theta\right]$$

$$= \frac{1}{2R}\,[e^{-\mu R}\{1 - E(Rx_0 - \mu/2x_0)\} + e^{\mu R}\{1 - E(Rx_0 + \mu/2x_0)\}]. \tag{B6}$$

In Eqs. (B5) and (B6), $E(x)$ denotes the error function [defined for negative x by the relation $E(x) = -E(-x)$].

The remaining integral over the range $0 \le x \le x_0$ is evaluated by a theta-function transformation. Three linearly independent lattice vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ are determined by the condition that all lattice points are generated by the linear combinations

$$\mathbf{R}_l = l_1\mathbf{a}_1 + l_2\mathbf{a}_2 + l_3\mathbf{a}_3, \quad l_i = 0, \pm1, \cdots. \tag{B7}$$

Here the j subscript is replaced by the descriptive vector $\mathbf{l}$. The dual vectors

$$b_1 = \frac{a_2 \times a_3}{A}, \quad b_2 = \frac{a_3 \times a_1}{A}, \quad b_3 = \frac{a_1 \times a_2}{A}, \tag{B8}$$

with $A = a_1 \cdot (a_2 \times a_3) = 1/\rho$, satisfy the conditions

$$a_i \cdot b_j = \delta_{ij}, \quad b_1 \cdot (b_2 \times b_3) = 1/A = \rho. \tag{B9}$$

The dual vector space contains all vectors

$$\mathbf{h}_\mathbf{p} = p_1\mathbf{b}_1 + p_2\mathbf{b}_2 + p_3\mathbf{b}_3, \quad p_i = 0, \pm1, \cdots. \tag{B10}$$

The required transformation,[12]

$$\sum_\mathbf{l} e^{-R_l^2x^2} = \frac{\pi^{\frac{3}{2}}\rho}{x^3}\sum_\mathbf{p} e^{-(\pi/x)^2h_\mathbf{p}^2} \tag{B11}$$

converts a slowly converging sum for x either very small or very large into an equivalent rapidly converging sum. Now

$$\frac{2}{\pi^{\frac{1}{2}}}\int_0^{x_0}\sum_\mathbf{l} e^{-R^2x^2-(\mu/2x)^2}\,dx = 2\pi\rho\int_0^{x_0}\sum_\mathbf{p} e^{-[\pi^2h_\mathbf{p}^2+(\mu/2)^2]/x^2}\frac{dx}{x^3} = \pi\rho\sum_\mathbf{p}\frac{e^{-[\pi^2h_\mathbf{p}^2+(\mu/2)^2]/x_0^2}}{\pi^2h_\mathbf{p}^2 + (\mu/2)^2}. \tag{B12}$$

These results are combined in Eq. (21) of the text and used to determine the best value of x_0.

Two Variational Principles for Magnetostatic Calculations

DAVID M. SCHRADER
Chemistry Department, University of Iowa, Iowa City, Iowa

(Received 13 August 1966)

Two variational principles for the estimation of the self-interaction energy of a solenoidal current density are introduced. The two variational principles are direct magnetic analogs of the principles of Thomson and Dirichlet for estimating interactions of charge densities. One of the magnetostatic variational principles provides an upper bound on the self-interaction energy, and the other, a lower bound. It is shown how one may extend the ideas to the estimation of the interaction of dissimilar current densities. Various other refinements and extensions are discussed, and a sample calculation is exhibited.

I. INTRODUCTION

THE self-interaction energy of a steady current $\mathbf{j}$ in a medium of unit permeability is

$$W = \frac{1}{2}\iint \frac{\mathbf{j}(\mathbf{r})\cdot\mathbf{j}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,d\mathbf{r}\,d\mathbf{r}', \tag{1}$$

which may also be written as

$$W = \frac{1}{8\pi}\int \mathbf{H}^2(\mathbf{r})\,d\mathbf{r} \tag{2}$$

$$= \frac{1}{8\pi}\int (\nabla\times\mathbf{A}(\mathbf{r}))^2\,d\mathbf{r} \tag{3}$$

$$= \frac{1}{2}\int \mathbf{j}(\mathbf{r})\cdot\mathbf{A}(\mathbf{r})\,d\mathbf{r}, \tag{4}$$

where the magnetic field $\mathbf{H}$ and vector potential $\mathbf{A}$ satisfy

$$\nabla\times\mathbf{H} = 4\pi\mathbf{j}, \tag{5}$$

$$\nabla\cdot\mathbf{H} = 0, \tag{6}$$

$$\nabla\times(\nabla\times\mathbf{A}) = 4\pi\mathbf{j}. \tag{7}$$

The practical calculator attempts to choose the most tractable integral from among the right-hand sides of Eqs. (1)–(4) for a numerical evaluation of W. Unfortunately, all four forms are equally unpleasant. Equation (1) involves a six-fold integration; and while Eqs. (2)–(4) contain only three-fold integrals, two first-order differential equations [Eqs. (5) and (6)] must be solved in order to use Eq. (2), and one second-order differential equation [Eq. (7)] must be solved in order to use Eq. (3) or (4).

We give here two variational principles for the evaluation of W in which the variational integrals are much easier to evaluate than W itself. One variational principle is a direct analog of Thomson's principle and provides an upper bound on W. The other variational principle provides a lower bound and is analogous to Dirichlet's principle. Both Thompson's principle and Dirichlet's principle have recently been discussed in the literature.[1–3]

II. UPPER BOUND

We are given a current density $\mathbf{j}$ and wish to calculate its self-interaction. Suppose we have calculated an approximate magnetic field $\mathbf{H}$ which satisfies Eq. (5) but not necessarily Eq. (6); i.e.,

$$\nabla\times\tilde{\mathbf{H}} = 4\pi\mathbf{j}. \tag{5'}$$

Now, call $\mathbf{h}$ the (unknown) difference between the approximate field $\tilde{\mathbf{H}}$ and the (unknown) exact field $\mathbf{H}$;

$$\tilde{\mathbf{H}} = \mathbf{H} + \mathbf{h}. \tag{8}$$

Now consider the functional

$$W_> = \frac{1}{8\pi}\int \tilde{\mathbf{H}}^2\,d\mathbf{r} = \frac{1}{8\pi}\int (\mathbf{H}+\mathbf{h})^2\,d\mathbf{r}.$$

From Eqs. (8), (5'), and (5), we have $\nabla\times\mathbf{h} = 0$, hence

$$\int \mathbf{h}\cdot\mathbf{H}\,d\mathbf{r} = \int \mathbf{h}\cdot(\nabla\times\mathbf{A})\,d\mathbf{r}$$
$$= \int \mathbf{A}\cdot(\nabla\times\mathbf{h})\,d\mathbf{r} + \int \mathbf{n}\cdot(\mathbf{h}\times\mathbf{A})\,d\mathbf{s} = 0,$$

where we have used a familiar vector identity, Gauss' theorem, the relationship $\mathbf{H} = \nabla\times\mathbf{A}$, and the assumption that $\mathbf{h}$ vanishes sufficiently strongly on the infinite sphere. This result gives us

$$W_> = \frac{1}{8\pi}\int (\mathbf{H}^2 + \mathbf{h}^2)\,d\mathbf{r} = W + \frac{1}{8\pi}\int \mathbf{h}^2\,d\mathbf{r} \geq W.$$

Hence we have the magnetostatic analog of Thomson's principle:

$$W \leq \frac{1}{8\pi}\int \tilde{\mathbf{H}}^2\,d\mathbf{r}, \quad \nabla\times\tilde{\mathbf{H}} = 4\pi\mathbf{j}. \tag{9}$$

Maxwell[4] gives this variational principle for the special case $\mathbf{j} = 0$.

If $\tilde{\mathbf{H}}$ also satisfies

$$\nabla\cdot\tilde{\mathbf{H}} = 0, \tag{6'}$$

then $\tilde{\mathbf{H}} \equiv \mathbf{H}$, and the equality sign in Eq. (9) holds. Therefore, Eq. (6') is the Euler–Lagrange equation associated with the variational principle (9).

[1] D. M. Schrader, Ph.D. thesis, University of Minnesota (1962).
[2] D. M. Schrader and S. Prager, J. Chem. Phys. **37**, 1456 (1962).
[3] S. Prager and J. O. Hirschfelder, J. Chem. Phys. **39**, 3289 (1963).
[4] J. C. Maxwell, *A Treatise on Electricity and Magnetism* (Clarendon Press, Oxford, England, 1904), 3rd ed., Vol. 1.

III. LOWER BOUND

Suppose we have an approximate vector potential $\tilde{\mathbf{A}}$ which differs from the exact (unknown) vector potential $\mathbf{A}$ by an (unknown) vector $\mathbf{a}$:

$$\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{a},$$

$\mathbf{A}$ satisfies Eq. (7) but $\tilde{\mathbf{A}}$ need not. Consider

$$W_< = \int \mathbf{j}\cdot\tilde{\mathbf{A}}\,d\mathbf{r} - \frac{1}{8\pi}\int(\nabla\times\tilde{\mathbf{A}})^2\,d\mathbf{r}$$
$$= \int \mathbf{j}\cdot(\mathbf{A}+\mathbf{a})\,d\mathbf{r} - \frac{1}{8\pi}\int[(\nabla\times\mathbf{A})^2 + 2(\nabla\times\mathbf{A})\cdot(\nabla\times\mathbf{a}) + (\nabla\times\mathbf{a})^2]\,d\mathbf{r}. \quad (10)$$

Now from Eq. (7) we have

$$\int \mathbf{j}\cdot\mathbf{a}\,d\mathbf{r} = \frac{1}{4\pi}\int \mathbf{a}\cdot[\nabla\times(\nabla\times\mathbf{A})]\,d\mathbf{r}$$
$$= \frac{1}{4\pi}\int(\nabla\times\mathbf{a})\cdot(\nabla\times\mathbf{A})\,d\mathbf{r} - \int \mathbf{n}\cdot[\mathbf{a}\times(\nabla\times\mathbf{A})]\,d\mathbf{s}.$$

Hence we see that, for $\mathbf{a}$ well behaved, Eq. (10) becomes

$$W_< = \int \mathbf{j}\cdot\mathbf{A}\,d\mathbf{r} - \frac{1}{8\pi}\int[(\nabla\times\mathbf{A})^2 + (\nabla\times\mathbf{a})^2]\,d\mathbf{r}$$
$$= W - \frac{1}{8\pi}\int(\nabla\times\mathbf{a})^2\,d\mathbf{r} \le W.$$

Therefore, we have the magnetostatic analog of Dirichlet's principle:

$$W \ge \int \mathbf{j}\cdot\tilde{\mathbf{A}}\,d\mathbf{r} - \frac{1}{8\pi}\int(\nabla\times\mathbf{A})^2\,d\mathbf{r}. \quad (11)$$

The associated Euler–Lagrange equation is

$$\nabla\times(\nabla\times\tilde{\mathbf{A}}) = 4\pi\mathbf{j}. \quad (7')$$

It is worth noting that the variational principle (11) is gauge invariant. A similar variational principle is given by Brown[5] in terms of the magnetization and a trial field.

IV. DISSIMILAR CURRENT DENSITIES

The identity

$$\mathbf{j}_1(\mathbf{r})\cdot\mathbf{j}_2(\mathbf{r}') = \tfrac{1}{4}\{[\mathbf{j}_1(\mathbf{r}) + \mathbf{j}_2(\mathbf{r})]\cdot[\mathbf{j}_1(\mathbf{r}') + \mathbf{j}_2(\mathbf{r}')] - [\mathbf{j}_1(\mathbf{r}) - \mathbf{j}_2(\mathbf{r})]\cdot[\mathbf{j}_1(\mathbf{r}') - \mathbf{j}_2(\mathbf{r}')]\} \quad (12)$$

enables us to express the more general integral

$$W' = \iint \frac{\mathbf{j}_1(\mathbf{r})\cdot\mathbf{j}_2(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,d\mathbf{r}\,d\mathbf{r}' \quad (13)$$

as the difference of two integrals of the form of Eq. (1). Therefore, a lower bound of W' is obtained if we estimate

$$\iint \frac{[\mathbf{j}_1(\mathbf{r}) + \mathbf{j}_2(\mathbf{r})]\cdot[\mathbf{j}_1(\mathbf{r}') + \mathbf{j}_2(\mathbf{r}')]}{|\mathbf{r}-\mathbf{r}'|}\,d\mathbf{r}\,d\mathbf{r}'$$

by making an appropriate application of the lower-bound variational principle [Eq. (11)] and

$$\iint \frac{[\mathbf{j}_1(\mathbf{r}) - \mathbf{j}_2(\mathbf{r})]\cdot[\mathbf{j}_1(\mathbf{r}') - \mathbf{j}_2(\mathbf{r}')]}{|\mathbf{r}-\mathbf{r}'|}\,d\mathbf{r}\,d\mathbf{r}',$$

using the appropriate form of Eq. (9). Reversing the procedure yields an upper bound on the integral W'.

[5] W. F. Brown, Jr., J. Phys. Soc. Japan, Suppl. B-I, 17, 540 (1962).

V. OTHER EXTENSIONS

A. Optimum Partitioning of the Current

Equation (12) may be generalized to

$$\mathbf{j}_1(\mathbf{r})\cdot\mathbf{j}_2(\mathbf{r}') = \frac{1}{4\gamma}\{[\mathbf{j}_1(\mathbf{r}) + \gamma\mathbf{j}_2(\mathbf{r})]\cdot[\mathbf{j}_1(\mathbf{r}') + \gamma\mathbf{j}_2(\mathbf{r}')] - [\mathbf{j}_1(\mathbf{r}) - \gamma\mathbf{j}_2(\mathbf{r})]\cdot[\mathbf{j}_1(\mathbf{r}') - \gamma\mathbf{j}_2(\mathbf{r}')]\}, \quad (14)$$

where γ is an arbitrary finite parameter. Suppose we have found trial fields $\tilde{\mathbf{H}}_1$ and $\tilde{\mathbf{H}}_2$ that satisfy

$$\nabla\times\tilde{\mathbf{H}}_m = 4\pi\mathbf{j}_m, \quad m = 1, 2,$$

and trial potentials $\tilde{\mathbf{A}}_1$ and $\tilde{\mathbf{A}}_2$ appropriate to $\mathbf{j}_1$ and $\mathbf{j}_2$, respectively. Denote

$$T_{mn} = \frac{1}{8\pi}\int \tilde{\mathbf{H}}_m\cdot\tilde{\mathbf{H}}_n\,d\mathbf{r},$$

$$D_{mn} = \frac{1}{2}\int(\mathbf{j}_1\cdot\tilde{\mathbf{A}}_2 + \mathbf{j}_2\cdot\tilde{\mathbf{A}}_1)\,d\mathbf{r} - \frac{1}{8\pi}\int(\nabla\times\tilde{\mathbf{A}}_1)\cdot(\nabla\times\tilde{\mathbf{A}}_2)\,d\mathbf{r}.$$

Then an upper bound on W' [Eq. (13)] is

$$W' \le \frac{1}{4}\left\{\frac{1}{\gamma}(T_{11} - D_{11}) + 2(T_{12} + D_{12}) + \gamma(T_{22} - D_{22})\right\}.$$

Mimimizing with respect to γ gives

$$W' \le \tfrac{1}{2}\{(T_{12} + D_{12}) + [(T_{11} - D_{11})(T_{22} - D_{22})]^{\frac{1}{2}}\}. \quad (15)$$

A lower bound on W' is

$$W' \ge \tfrac{1}{4}\{(1/\gamma)(D_{11} - T_{11}) + 2(T_{12} + D_{12}) + \gamma(D_{22} - T_{22})\},$$

which is, for the maximizing γ,

$$W' \ge \tfrac{1}{2}\{(T_{12} + D_{12}) - [(T_{11} - D_{11})(T_{22} - D_{22})]^{\frac{1}{2}}\}. \quad (16)$$

Combining Eqs. (15) and (16),

$$W' \gtrless \tfrac{1}{2}(T_{12} + D_{12}) \pm \tfrac{1}{2}[(T_{11} - D_{11})(T_{22} - D_{22})]^{\frac{1}{2}}. \quad (17)$$

The utility of these variational principles depends upon one's adroitness in guessing trial fields and potentials. One can partly compensate for a particularly inept guess by using Eq. (17), which automatically partitions the current densities in an optimal way. Another useful procedure is to scale optimally the trial fields and potentials.

B. Scaling the Potential

Since the bound given by relation (11) is true for arbitrary $\tilde{\mathbf{A}}$, which vanishes satisfactorily for large $\mathbf{r}$, we may replace $\tilde{\mathbf{A}}$ by $\alpha\tilde{\mathbf{A}}$ and maximize with respect

to α. The result is

$$W \geq \frac{\left(\frac{1}{2}\int \mathbf{j}\cdot\tilde{\mathbf{A}}\,d\mathbf{r}\right)^2}{\frac{1}{8\pi}\int(\nabla\times\tilde{\mathbf{A}})^2\,d\mathbf{r}}$$

$$\geq \int \mathbf{j}\cdot\tilde{\mathbf{A}}\,d\mathbf{r} - \frac{1}{8\pi}\int(\nabla\times\tilde{\mathbf{A}})^2\,d\mathbf{r}. \qquad (18)$$

C. Scaling the Field

We may add to $\tilde{\mathbf{H}}$ an arbitrary lamellar vector $\tilde{\mathbf{H}}'$ without disturbing the constraint in (9). That is,

$$\nabla\times\tilde{\mathbf{H}} = 4\pi\mathbf{j}$$

is invariant to

$$\tilde{\mathbf{H}} \to \tilde{\mathbf{H}} + \beta\tilde{\mathbf{H}}',$$

where $$\nabla\times\tilde{\mathbf{H}}' = 0 \qquad (19)$$

minimizing $$\frac{1}{8\pi}\int(\tilde{\mathbf{H}} + \beta\tilde{\mathbf{H}}')^2\,d\mathbf{r}$$

with respect to β gives

$$W \leq \frac{1}{8\pi}\left[\int\tilde{\mathbf{H}}^2\,d\mathbf{r} - \frac{(\int\tilde{\mathbf{H}}\cdot\tilde{\mathbf{H}}'\,d\mathbf{r})^2}{\int\tilde{\mathbf{H}}'^2\,d\mathbf{r}}\right] \leq \frac{1}{8\pi}\int\tilde{\mathbf{H}}^2\,d\mathbf{r}.$$

Since Eq. (19) is trivially easy to satisfy, this procedure might have some utility in calculations.

D. Non-Constant Permeability

All the relations [except Eq. (1)] may be generalized for the case of a variable magnetic permeability in a straightforward way.

VI. SAMPLE CALCULATION

Consider a charge density

$$\rho(\mathbf{r}) = (4 - r)e^{-r}\sin\theta/4\pi,$$

which rotates about the z axis with unit angular velocity. The current density is

$$\mathbf{j} = \hat{\varphi}(4 - r)e^{-r}\sin\theta/4\pi, \qquad (19')$$

where $\hat{r}$, $\hat{\theta}$, and $\hat{\varphi}$ are the unit vectors in spherical coordinates. It is easily shown that

$$\mathbf{A} = \hat{\varphi}re^{-r}\sin\theta, \qquad (20)$$

$$\mathbf{H} = 2\hat{r}e^{-r}\cos\theta + \hat{\theta}(r - 2)e^{-r}\sin\theta \qquad (21)$$

satisfy

$$\nabla\times\mathbf{A} = \mathbf{H}, \quad \nabla\times\mathbf{H} = 4\pi\mathbf{j}.$$

Also, a little calculating shows that

$$\frac{1}{2}\int\mathbf{j}\cdot\mathbf{A}\,d\mathbf{r} = \frac{1}{8\pi}\int\mathbf{H}^2\,d\mathbf{r} = \frac{1}{4}. \qquad (22)$$

To test our lower-bound variational principle (11), we choose

$$\tilde{\mathbf{A}} = \hat{\varphi}r^2e^{-\alpha r}\sin^2\theta, \qquad (23)$$

where α is a variational parameter. It is easy to show that

$$\frac{1}{2}\int\mathbf{j}\cdot\tilde{\mathbf{A}}\,d\mathbf{r} = \frac{9\pi}{4}\frac{4\alpha - 1}{(\alpha + 1)^6},$$

$$\frac{1}{8\pi}\int(\nabla\times\tilde{\mathbf{A}})^2\,d\mathbf{r} = \frac{24}{(2\alpha)^5},$$

and that

$$\left(\frac{1}{2}\int\mathbf{j}\cdot\tilde{\mathbf{A}}\,d\mathbf{r}\right)^2\Big/\left(\frac{1}{8\pi}\int(\nabla\times\tilde{\mathbf{A}})^2\,d\mathbf{r}\right)$$

has a maximum of approximately 0.2133 for $\alpha = 1.596$. Of course, we may have improved this result with a more flexible variational form for $\tilde{\mathbf{A}}$ than (23).

In choosing a trial field $\tilde{\mathbf{H}}$ to insert into (9), we must satisfy the constraint

$$\nabla\times\tilde{\mathbf{H}} = \hat{\varphi}(4 - r)e^{-r}\sin\theta. \qquad (24)$$

It is easy to show that

$$\tilde{\mathbf{H}} = -\hat{r}\cos\theta(r - 4)re^{-r}$$

satisfies Eq. (24), and is at the same time distinct from the correct magnetic field $\mathbf{H}$ [Eq. (21)]. Calculation shows that

$$\frac{1}{8\pi}\int\tilde{\mathbf{H}}^2\,d\mathbf{r} = \frac{7}{16} \simeq 0.4375.$$

To improve this unimpressive result, we add to $\tilde{\mathbf{H}}$ a lamellar vector $\tilde{\mathbf{H}}'$, where

$$\tilde{\mathbf{H}}' = \nabla\phi, \quad \phi = e^{-\beta r}\cos\theta,$$

and β is a variational parameter. The integrals that appear in (20) are

$$\int\tilde{\mathbf{H}}\cdot\tilde{\mathbf{H}}'\,d\mathbf{r} = \frac{-32\pi\beta^2}{(\beta + 1)^5},$$

$$\int\tilde{\mathbf{H}}'\,d\mathbf{r} = \frac{5\pi}{3\beta}.$$

Our result is that

$$\frac{1}{8\pi}\left(\int\tilde{\mathbf{H}}\cdot\tilde{\mathbf{H}}'\,d\mathbf{r}\right)^2\Big/\int\tilde{\mathbf{H}}'\,d\mathbf{r}$$

has a maximum of 0.0749 for $\beta = 1$, giving an upper bound of 0.3625.

VII. APPLICATIONS

Aside from the obvious application of calculating bounds on the interactions of classical current densities and calculating approximate fields and potentials for known currents, one might use these variational principles to estimate the dipolar interaction integrals that occur in quantum mechanical investigations of atoms and molecules. It should be possible to arrange a sum of electrostatic repulsion integrals as a single integral of the form of the right-hand side of Eq. (1), where $\mathbf{j}$ is a solenoidal vector in some multi-dimensional space.

ACKNOWLEDGMENTS

Professor S. Prager provided the author with the original stimulation for this work by introducing him to the electrostatic variational principles. The author is grateful to Professor S. Prager and to C. F. Lebeda for helpful discussions.

This research was supported in part by the Research Corporation and the University of Iowa Foundation.

Analytic Continuation of the Off-Energy-Shell Scattering Amplitude

J. NUTTALL
Department of Physics, Texas A & M University, College Station, Texas

(Received 31 May 1966)

It is shown that the two-body scattering amplitude $\langle \mathbf{k}| T(E) |\mathbf{k}'\rangle$ may be analytically continued in E through the physical cut into a region of meromorphy, provided that the potential satisfies certain requirements. The residue at a pole is a separable operator whose form agrees with previous work. The proof also demonstrates the existence of a region of meromorphy of $\langle \mathbf{k}| T(E) |\mathbf{k}'\rangle$ as a function of seven complex variables.

1. INTRODUCTION

RECENTLY there has been a good deal of interest in the off-energy-shell two-body T matrix $\langle \mathbf{k}| T(E) |\mathbf{k}'\rangle$, since this quantity enters into the Watson[1]–Faddeev[2] equations for the three-particle scattering. In particular, a number of authors[3–5] have considered approximate forms which may be valid for energies near to a resonance. This question involves an understanding of how the T matrix is continued through the physical cut onto the second sheet of the E plane. Previously, most discussions have been restricted to the case of partial wave amplitudes, although Lovelace[6] has stated without proof some results about $\langle \mathbf{k}| T(E) |\mathbf{k}'\rangle$.

In this paper, we show that, with certain restrictions on the potential, $\langle \mathbf{k}| T(E) |\mathbf{k}'\rangle$ may be continued in E through the physical cut into a region where it is analytic save for possible poles which would correspond to resonances. In deriving this result we show that $\langle \mathbf{k}| T(E) |\mathbf{k}'\rangle$ is a meromorphic function of the seven complex variables $k_x, k_y, k_z, k'_x, k'_y, k'_z, E$ in a certain domain which includes the physical region. We hope that our results are useful in studying the analytic properties of the operator $T(E)$ for three particles, which might then lead to a better understanding of three-body resonances.

2. ANALYTICITY OF $\langle \mathbf{k}| V |\mathbf{k}'\rangle$

To derive properties such as we have mentioned, we need to place some restrictions on the form of the potential. We assume that it is local and is a superposition of Yukawa potentials, so that

$$\langle \mathbf{k}| V |\mathbf{k}'\rangle = V(\mathbf{k} - \mathbf{k}'), \tag{1}$$

$$V(\mathbf{k}) = \int_\mu^\infty d\eta \delta(\eta)/(k^2 + \eta^2). \tag{2}$$

As Lovelace[6] does, we insist that

$$\int_\mu^\nu \delta(\eta)\, d\eta$$

is of bounded variation on $(\mu, +\infty)$.

The T matrix satisfies the equation[1]

$$\langle \mathbf{k}| T(E) |\mathbf{k}_0\rangle = \langle \mathbf{k}| V |\mathbf{k}_0\rangle + \int d\mathbf{k}' \langle \mathbf{k}| V |\mathbf{k}'\rangle (E - k'^2)^{-1} \langle \mathbf{k}'| T(E) |\mathbf{k}_0\rangle, \tag{3}$$

where the units are chosen so that $\hbar^2/2m = 1$.

Now the analyticity of $\langle \mathbf{k}| T(E) |\mathbf{k}_0\rangle$ as a function of the complex variable E for fixed, real $\mathbf{k}, \mathbf{k}_0$ has already been demonstrated for E in the entire plane cut from 0 to $+\infty$, the physical cut. We wish to continue in E through this cut. This is accomplished by distorting the $\mathbf{k}'$ integration surface to complex values so that $(E - k'^2)$ no longer vanishes for a range of real E. Such a procedure is possible because of the analyticity of $\langle \mathbf{k}| V |\mathbf{k}'\rangle$ which follows from (2).

Let us suppose that we wish to continue in E downward through the real axis between $E = k_1^2$ and $E = k_2^2$, $k_2 > k_1 > 0$. Initially the contour of integration is a three- (real) dimensional chain in the space C^3 of three complex variables $\mathbf{k}$ consisting of all real $\mathbf{k}$. This contour we denote by S_0. We consider a continuous distortion of the initial contour S_0 to a new contour S_1 by way of a set of contours $S_\lambda (0 \leq \lambda \leq 1)$. Only that part of S_0 with $k_1 < |\mathbf{k}| \leq k_2$ is distorted, the remainder staying the same, and we choose to define S_λ by

$$S_\lambda = \{\mathbf{k}: \mathbf{k} = \mathbf{K}[1 - i\lambda\alpha(K)]\} \tag{4}$$

for all $\mathbf{K} \in S_0$ and $\alpha(K)$ is a continuous real function of $\mathbf{K}$ with the property

$$\alpha(K) = 0 \quad \text{for} \quad 0 \leq |\mathbf{K}| \leq k_1 \quad \text{and} \quad k_2 \leq |\mathbf{K}|. \tag{5}$$

In order to be useful, $\alpha(K)$ must be chosen so that

[1] See, for instance, M. L. Goldberger and K. M. Watson, *Collision Theory* (John Wiley & Sons, Inc., New York, 1964).
[2] L. D. Faddeev, *Mathematical Aspects of the Three-Body Problem* (Daniel Davey & Company, Inc., New York, 1965).
[3] K. L. Kowalski, Phys. Rev. Letters **15**, 798 (1965).
[4] H. P. Noyes, Phys. Rev. Letters **15**, 538 (1965).
[5] J. Y. Guennégués, Orsay preprint (1965).
[6] C. Lovelace, in *Strong Interactions and High Energy Physics*, R. G. Moorhouse, Ed. (Oliver and Boyd, London, 1964).

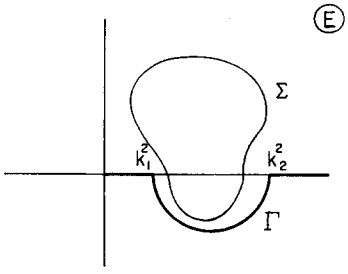


FIG. 1. The complex E plane with Γ being the set of values of k^2 for $\mathbf{k} \in S_1$.

S_λ satisfies the following two requirements.

(i) There exists a connected domain Σ in the complex E plane which includes part of the real axis between k_1^2 and k_2^2 such that $(E - k^2)$ does not vanish for any $\mathbf{k} \in S_1$.

(ii) It is possible to find a connected six- (real) dimensional domain $D \in C^3$ such that $S_\lambda \in D$ for all λ $(0 \leq \lambda \leq 1)$ with the property that $\langle \mathbf{k}| V |\mathbf{k}'\rangle$ is an analytic function of $\mathbf{k}$ and $\mathbf{k}'$ for all $\mathbf{k}, \mathbf{k}' \in D$.

First of all we establish the existence of an $\alpha(K)$ satisfying these properties and then in the next section we show how they are used to continue $\langle \mathbf{k}| T(E) |\mathbf{k}'\rangle$.

The first requirement is satisfied if $\alpha(\mathbf{K})$ is chosen to be positive for $k_1 < |\mathbf{K}| < k_2$ for then on S_1

$$\operatorname{Im} k^2 = -2\alpha(K)\,|\mathbf{K}|^2 < 0. \tag{6}$$

Thus the set of values taken by k^2 as $\mathbf{k}$ varies on S_1 lie on the curve Γ sketched in Fig. 1 and Σ can be chosen as shown.

In connection with the second requirement, we note that if $\alpha(K)$ is small enough, then on all S_λ $(0 \leq \lambda \leq 1)$ $\mathbf{k}$ and $\mathbf{k}'$ are very near to real values which certainly lie inside the domain of analyticity of $\langle \mathbf{k}| V |\mathbf{k}'\rangle$.

The question is how large may we take $\alpha(K)$ and still remain in the domain of analyticity? The larger we take $\alpha(K)$ then the further Σ can project into the lower half E plane.

Let us consider $y = (\mathbf{k} - \mathbf{k}')^2$ with

$$\mathbf{k} = \mathbf{K}[1 - i\lambda\alpha(K)], \quad \mathbf{k}' = \mathbf{K}'[1 - i\lambda'\alpha(K')], \tag{7}$$

where $\mathbf{K}, \mathbf{K}'$ are real and $0 \leq \lambda, \lambda' \leq 1$. We must prove that for no values of $\mathbf{K}$, K' is y real and negative with $y \leq -\mu^2$, for (2) shows that then $\langle \mathbf{k}| V |\mathbf{k}'\rangle$ would not be analytic. We need only consider the case when at least one of $|\mathbf{K}|$, $|\mathbf{K}'|$ lie between k_1 and k_2, say $|\mathbf{K}|$, for otherwise y is real and positive from (5).

To simplify the analysis, take $\alpha(K)$ to be constant between k_1 and k_2, or strictly a continuous function which approximates this very closely. In fact this turns out not to be a restriction. Now consider the case when $|\mathbf{K}'|$ does not lie between k_1 and k_2. In this case $\alpha(K') = 0$ and

$$\begin{aligned} \operatorname{Im} y &= -\lambda\alpha(K)\mathbf{K} \cdot (\mathbf{K} - \mathbf{K}') \\ &= -\lambda\alpha(K)(K^2 - KK'\cos\theta), \end{aligned} \tag{8}$$

where θ is the angle between $\mathbf{K}$ and $\mathbf{K}'$. For $\langle \mathbf{k}| V |\mathbf{k}'\rangle$ to be singular, $\operatorname{Im} y$ must certainly be zero which occurs when

$$\cos\theta = K/K'. \tag{9}$$

To satisfy this condition we need $K' \geq K$ and so $K' \geq k_2$ from our assumption about K'. If $\operatorname{Im} y = 0$ we have

$$\operatorname{Re} y = K'^2 - K^2[1 + \lambda^2\alpha^2(K)]. \tag{10}$$

The minimum value of $\operatorname{Re} y$ is, with $\alpha(K)$ fixed,

$$(\operatorname{Re} y)_{\min} = k_2^2 - k_2^2(1 + \alpha^2) = -\alpha^2 k_2^2.$$

For the required analyticity, we need $(\operatorname{Re} y)_{\min} > -\mu^2$, or

$$\alpha k_2 < \mu. \tag{11}$$

If $|K'|$ lies between k_1 and k_2 we may carry out a similar analysis which results in the same condition as (11). The existence of the domain D follows immediately.

Thus the conclusion is that both the requirements (i) and (ii) are satisfied with $\alpha(K)$ constant between k_1 and k_2 so long as (11) holds. With the optimum choice of α, the curve Γ in the complex E plane is shown in Fig. 2, where the point Q is $E = k_2^2 - y^2 - 2i\mu k_2$. The area swept out by the curve Γ as k_1 and k_2 are varied lies below the real axis and is bounded by a parabola P, the locus of all points A, but does not include the real axis between $-\mu^2$ and $-\frac{1}{4}\mu^2$. We show that the T matrix may be continued analytically into this region.

3. MEROMORPHY OF $\langle \mathbf{k}| T(E) |\mathbf{k}_0\rangle$

To study the analyticity of $\langle \mathbf{k}| T(E) |\mathbf{k}_0\rangle$ we first define the Hilbert space $\mathfrak{H}$ of all measurable functions $f(\mathbf{k})$ defined on S_1 and satisfying

$$\int_{S_1} d\mathbf{k}\,|f(\mathbf{k})|^2 < \infty. \tag{12}$$

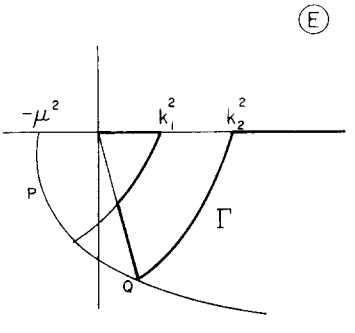


FIG. 2. The form of Γ when $\alpha(K)$ is constant.

The linear operator $a(E)$ acting in $\mathfrak{H}$ is defined by

$$g = a(E)f,$$
$$g(\mathbf{k}) = \int_{S_1} d\mathbf{k}' \langle \mathbf{k}| V |\mathbf{k}'\rangle (E - k'^2)^{-1} f(\mathbf{k}'). \quad (13)$$

Since $V(\mathbf{k}) \sim k^{-2}$ as $|\mathbf{k}| \to \infty$, $g(k)$ also belongs to the Hilbert space. Using arguments identical to those proposed by Lovelace[6] and Weinberg,[7] we may deduce that $a(E)$ is an L_2 operator in $\mathfrak{H}$ and also is operator analytic for $E \in \Sigma$, since the denominator of (13) does not vanish.

Now let us define $t(\mathbf{k}_0, E) \in \mathfrak{H}$ which satisfies the equation

$$t(\mathbf{k}_0, E) = v(\mathbf{k}_0) + a(E)t(\mathbf{k}_0, E). \quad (14)$$

In terms of the square-integrable functions corresponding to $t(\mathbf{k}_0, E)$, $v(\mathbf{k}_0)$, (14) reads

$$t(\mathbf{k}, \mathbf{k}_0, E) = \langle \mathbf{k}| V |\mathbf{k}_0\rangle + \int_{S_1} d\mathbf{k}' \langle \mathbf{k}| V |\mathbf{k}'\rangle (E - k'^2)^{-1} \times t(\mathbf{k}', \mathbf{k}_0, E), \quad (15)$$

which defines $v(\mathbf{k}_0) \in D$.

We now make use of the conclusions of the theorem due to Rellich[8] stated by Lovelace[6] and discussed in the Appendix. Since $a(E)$ is certainly a compact operator analytic for $E \in \Sigma$, $[1 - a(E)]^{-1}$ is a meromorphic operator in that region, and the unique solution of (14) is

$$t(\mathbf{k}_0, E) = [1 - a(E)]^{-1} v(\mathbf{k}_0). \quad (16)$$

If E is not at one of the poles of $[1 - a(E)]^{-1}$, we can use (15) to extend the region of definition of $t(\mathbf{k}, \mathbf{k}_0, E)$ from $\mathbf{k} \in S_1$ to $\mathbf{k} \in D$. In fact $t(\mathbf{k}, \mathbf{k}_0, E)$ is analytic in $\mathbf{k}$ for $\mathbf{k} \in D$. Thus, in particular, (15) holds for real $\mathbf{k}$ and $\mathbf{k}_0$.

Now take a value of $E \in \Sigma$ with $\operatorname{Im} E > 0$. Since $\operatorname{Im} k'^2 < 0$ for $\mathbf{k}' \in S_\lambda$ $(0 \leq \lambda \leq 1)$, $(E - k'^2)^{-1}$ is analytic in $\mathbf{k}'$ for $\mathbf{k}'$ on these contours, as also are $\langle \mathbf{k}| V |\mathbf{k}'\rangle$ and $t(\mathbf{k}', \mathbf{k}_0, E)$ for $\mathbf{k}$, $\mathbf{k}_0$ real. Consequently, the integrand on the right-hand side of (15) is an analytic function of $\mathbf{k}'$ in the neighborhood of S_λ $(0 \leq \lambda \leq 1)$ with $\operatorname{Im} E > 0$. We may therefore distort the contour from S_1 to S_0 without changing the value of the integral (see Fuks[9]) to obtain

$$t(\mathbf{k}, \mathbf{k}_0, E) = \langle \mathbf{k}| V |\mathbf{k}_0\rangle + \int_{S_0} d\mathbf{k}' \langle \mathbf{k}| V |\mathbf{k}'\rangle (E - k'^2)^{-1} \times t(\mathbf{k}', \mathbf{k}_0, E), \quad (17)$$

with $\mathbf{k}$, $\mathbf{k}_0$ real and $\operatorname{Im} E > 0$.

This equation is identical with that satisfied by the usual T matrix $\langle \mathbf{k}| T(E) |\mathbf{k}_0\rangle$ with E just above the physical cut, and thus we may identify $t(\mathbf{k}, \mathbf{k}_0, E)$ with $\langle \mathbf{k}| T(E) |\mathbf{k}_0\rangle$ for $\mathbf{k}$, $\mathbf{k}_0$ real, $\operatorname{Im} E > 0$. It follows that the analytic continuation of $\langle \mathbf{k}| T(E) |\mathbf{k}_0\rangle$ to values of $E \in \Sigma$ with $\operatorname{Im} E < 0$, and also to complex $\mathbf{k} \in D$, is given by $t(\mathbf{k}, \mathbf{k}_0, E)$. Also since $\langle \mathbf{k}| V |\mathbf{k}_0\rangle$ is also analytic in $\mathbf{k}_0$ for $\mathbf{k}_0$, $\mathbf{k} \in D$, $t(\mathbf{k}, \mathbf{k}_0, E)$ is analytic in all seven complex variables $\mathbf{k}$, $\mathbf{k}_0$, E with $\mathbf{k}$, $\mathbf{k}_0 \in D$ and $E \in \Sigma$ except for poles at a discrete set of values of E, which we know[2] cannot lie at points with $\operatorname{Im} E > 0$.

4. FORM OF RESIDUE AT A RESONANCE

We may use the second part of the theorem in the Appendix to study the form of the residue of $t(\mathbf{k}, \mathbf{k}_0, E)$ at a pole $E = E_0$ with $\operatorname{Im} E_0 < 0$. Let us consider the case of a simple pole, and assume that there is only one independent vector $x \in \mathfrak{H}$ such that

$$[a(E_0) - 1]^n x = 0 \quad (18)$$

for any positive integer n. Moreover we assume that

$$a(E_0)x = x. \quad (19)$$

In this case the theorem of the Appendix assures us that the residue of the operator $[1 - a(E)]^{-1}$ at the pole $E = E_0$ is proportional to a projection operator onto the eigenvector x. Thus, using (16), for $\mathbf{k} \in S_1$, $\mathbf{k}_0 \in D$, $t(\mathbf{k}, \mathbf{k}_0, E)$ may be approximated by

$$t(\mathbf{k}, \mathbf{k}_0, E) \approx \frac{x(\mathbf{k})}{E - E_0} \cdot \int_{S_1} d\mathbf{k}' y^*(\mathbf{k}') \langle \mathbf{k}'| V |\mathbf{k}_0\rangle. \quad (20)$$

The function $x(\mathbf{k})$ satisfies (19), which may be written

$$x(\mathbf{k}) = \int_{S_1} d\mathbf{k}' \langle \mathbf{k}| V |\mathbf{k}'\rangle [E_0 - k'^2]^{-1} x(\mathbf{k}') \quad (21)$$

and $y(\mathbf{k})$ is unknown. As before, we may deduce that $x(\mathbf{k})$ is analytic for $\mathbf{k} \in D$. Thus (20) may be continued to all $\mathbf{k} \in D$.

The symmetry of $t(\mathbf{k}, \mathbf{k}_0, E)$ tells us that

$$\int_{S_1} d\mathbf{k}' y^*(\mathbf{k}') \langle \mathbf{k}'| V |\mathbf{k}_0\rangle = A x(\mathbf{k}_0) \quad (22)$$

with A a constant. A detailed argument which proves this may be constructed by starting from the equation

$$\langle \mathbf{k}| T(E) |\mathbf{k}_0\rangle = \langle \mathbf{k}| V |\mathbf{k}_0\rangle + \int d\mathbf{k}' \langle \mathbf{k}| T(E) |\mathbf{k}'\rangle (E - k'^2)^{-1} \langle \mathbf{k}'| V |\mathbf{k}_0\rangle, \quad (23)$$

and repeating the previous steps. We know the solutions of (3) and (23) are the same for real $\mathbf{k}$, $\mathbf{k}_0$ and $\operatorname{Im} E > 0$, so they must be the same everywhere they are defined.

Now we find an equation for the form factor $x(\mathbf{k})$ which involves as far as possible only real $\mathbf{k}$. To do

[7] S. Weinberg, Phys. Rev. **133**, B232 (1964).
[8] F. Rellich, Math. Ann. **113**, 600 (1936); **113**, 677 (1936); **116**, 555 (1939); **117**, 356 (1940); **118**, 462 (1941).
[9] B. A. Fuks, *Functions of Several Complex Variables* (American Mathematical Society, Providence, Rhode Island, 1963).

this we must again move the contour S_1 of (21) back to S_0, but now there is singularity when $k'^2 = E_0$. To study the residue here we change variables from $\mathbf{k}'$ to $z\hat{\mathbf{K}}$. On the contour S_1 we have the relation

$$z = |\mathbf{K}|\,[1 - i\alpha(K)] \quad \text{for all } \hat{\mathbf{K}}. \tag{24}$$

In (21) we may carry out the integration over $\hat{\mathbf{K}}$ to leave an integral over z where z runs along a contour C given by (24), which is such that z^2 is on Γ. We obtain

$$x(\mathbf{k}) = \int_C z^2\, dz(E_0 - z^2)^{-1}F(\mathbf{k}, z). \tag{25}$$

The function $F(\mathbf{k}, z)$ is analytic in z in the region between C and the real axis and is given by

$$F(\mathbf{k}, z) = \int d\hat{\mathbf{K}}\langle \mathbf{k}|\, V\, |z\hat{\mathbf{K}}\rangle x(z\hat{\mathbf{K}}). \tag{26}$$

Now distort the contour C back to the real axis, compensating by taking the residue at $z = E_0^{\frac{1}{2}}$ (which lies in the lower half-plane), to give

$$x(\mathbf{k}) = \int_0^\infty z^2\, dz(E_0 - z^2)^{-1}F(\mathbf{k}, z) - i\pi E_0^{\frac{1}{2}}F(\mathbf{k}, E_0^{\frac{1}{2}}). \tag{27}$$

Transforming back to $\mathbf{k}'$, now real, we obtain

$$x(\mathbf{k}) = -i\pi E_0^{\frac{1}{2}}\int d\hat{\mathbf{K}}\langle \mathbf{k}|\, V\, |E_0^{\frac{1}{2}}\hat{\mathbf{K}}\rangle x(E_0^{\frac{1}{2}}\hat{\mathbf{K}}) + \int_{S_0} d\mathbf{k}'\langle \mathbf{k}|\, V\, |\mathbf{k}'\rangle(E_0 - k'^2)^{-1}x(\mathbf{k}'). \tag{28}$$

To relate this equation to other work, we remark that a solution of (28) may be found of the form

$$x(k\hat{\mathbf{K}}) = f_l(k)Y_l^0(\hat{\mathbf{K}}). \tag{29}$$

Equation (28) then reduces to

$$f_l(k) = -i\pi E_0^{\frac{1}{2}}V_l(k, E_0^{\frac{1}{2}})f_l(E_0^{\frac{1}{2}}) + \int_0^\infty k'^2\, dk'V_l(k, k')(E_0 - k'^2)^{-1}f_l(k'), \tag{30}$$

where

$$V_l(z, z') = \int d(\cos\theta)P_l(\cos\theta) \times \int_\mu^\infty d\eta\sigma(\eta)[z^2 + z'^2 - 2zz'\cos\theta + \eta^2]^{-1}. \tag{31}$$

By comparing (30) with an unnumbered equation of Noyes,[4] the one before his Eq. (4), we deduce that our $f_l(k)$ is proportional to his $f_l(\eta, k)|_{\eta=E_0^{\frac{1}{2}}}$. The residue of the T matrix at its pole given by (20) therefore agrees with that obtained from Noyes' Eq. (13). Thus our work fits in with the partial wave analysis.

5. DISCUSSION

The interesting work of Kowalski[3] and Noyes[4] on approximating the two-body scattering amplitude near a resonance by a separable term needs to be supplemented by a discussion of the analyticity of the remainder. Guennéguès[5] has provided this for the partial-wave amplitude while the present paper does the same for the full amplitude. Our results may easily be seen to imply analyticity of the partial-wave amplitude. Results on the full amplitude have also been stated by Lovelace, but his outline of the method of proof seems different and more complicated.

Our method can undoubtedly be used to extend the region of analyticity beyond that given here. It may also perhaps be useful in simplifying the discussion of questions involving limits as E approaches the real axis. We hope, using the two-body results, to apply the method to study the second-sheet properties of the three-body T-matrix, and thereby understand more about the nature of three-body resonances.

ACKNOWLEDGMENTS

This work was supported by the United States Air Force Office of Scientific Research.

APPENDIX

In this Appendix we discuss a theorem about $[1 - T(\mu)]^{-1}$, where $T(\mu)$ is a compact operator in a Banach space X which is operator analytic in μ for each μ in a connected domain D. The theorem, more general than we need here, includes a result stated by Lovelace,[6] but is based on the work of Rellich.[8] Tiktopoulos[10] has also proved parts of the theorem for Hilbert space. We follow the discussion of Dunford and Schwartz[11] (abbreviated as DS).

In the form in which we require it, the theorem states:

Theorem: If some power of $T(\mu)$ is a compact operator in a Banach space X, operator analytic in μ for each μ in a connected domain D of the complex μ-plane, then either (i) $1 - T(\mu)$ has a bounded inverse for no point in D, or (ii) the inverse exists except at a countable number of isolated points and $[1 - T(\mu)]^{-1}$ is analytic for $\mu \in D$ apart from these points.

In Case ii, in the neighborhood of a point μ_0, where the inverse does not exist, $[1 - T(\mu)]^{-1}$ is given by a Laurent expansion

$$[1 - T(\mu)]^{-1} = \sum_{i=-n}^{\infty} A_i(\mu - \mu_0)^i, \tag{A1}$$

[10] G. Tiktopoulos, Phys. Rev. 133, B1231 (1964).
[11] N. Dunford and J. T. Schwartz, *Linear Operators, Part* I (Interscience Publishers, Inc., New York, 1958).

where A_{-n} acts in the finite-dimensional space X_{μ_0} of vectors $x \in X$ satisfying $[1 - T(\mu_0)]^r x = 0$ for some positive integer r. In fact if ν is the index associated with μ_0, that is the smallest r such that $[1 - T(\mu_0)]^r x = 0$ for all $x \in X_{\mu_0}$, then

$$A_{-n} = A[1 - T(\mu_0)]^{\nu-1} E(\mu_0), \tag{A2}$$

where A is a numerical constant. The projection operator $E(\mu_0)$ projects onto X_{μ_0} and is given by

$$E(\mu_0) = \frac{1}{2\pi i} \int d\lambda [\lambda - T(\mu_0)]^{-1}, \tag{A3}$$

where the integral is taken on a small circle about $\lambda = 1$, not including any other point in the spectrum of $T(\mu_0)$.

All these results are either stated in Chap. VII of DS or else they follow without much difficulty. Only the form of the residue (A2) perhaps needs some discussion, and we sketch below a proof of this result, starting from the facts given by DS.

First of all we note that for μ near enough to μ_0 $E(\mu)$ given by (A3) is a projection operator analytic in μ (DS, VII, 6.9). Also X_{μ_0} may be defined by $X_{\mu_0} = E(\mu_0)X$ (DS, VII, 4.5). If $x_i \in X_{\mu_0}$, $i = 1, \cdots, m$ form a basis for X_{μ_0}, then $E(\mu)x_i$ form a basis for the subspace $E(\mu)X$ for μ near enough to μ_0 (DS, VII, 6.8).

To find the inverse of $\lambda - T(\mu)$ we write

$$\lambda - T(\mu) = \lambda - E(\mu)T(\mu) - (1 - E(\mu))T(\mu) \tag{A4}$$

and consider solving

$$[\lambda - T(\mu)]x = y \tag{A5}$$

for any $y \in X$. We can write any vector in the form $x = x' + x''$, where

$$x' = [1 - E(\mu)]x, \quad x'' = E(\mu)x,$$

so that $x = y$ implies $x' = y'$, $x'' = y''$. Now using (A4) and remembering that $E(\mu)$ and $T(\mu)$ commute, we easily deduce (see Kantorovich and Akilov[12])

$$\begin{aligned} \lambda x' - [1 - E(\mu)]T(\mu)x' &= y', \\ \lambda x'' - E(\mu)T(\mu)x'' &= y''. \end{aligned} \tag{A6}$$

It can be shown that $\{1 - [1 - E(\mu_0)]T(\mu_0)\}$ has an inverse and therefore so does $\{\lambda - [1 - E(\mu)]T(\mu)\}$ for λ, μ sufficiently near to $1, \mu_0$, and the inverse is analytic in λ, μ (DS, VII, 6.4). Consequently we may write

$$x' = \{\lambda - [1 - E(\mu)T(\mu)\}^{-1} y'. \tag{A7}$$

Now the second equation of (A6) is an equation in a space of dimension $m, E(\mu)X$. This equation is discussed in DS, VII, 6.9, where $t_{ij}(\mu)$ is defined by

$$T(\mu)E(\mu)x_i = \sum_{j=1}^{m} t_{ij}(\mu)E(\mu)x_i. \tag{A8}$$

We write

$$x'' = \sum_{i=1}^{m} \alpha_i(\mu)E(\mu)x_i, \quad y'' = \sum_{i=1}^{m} \beta_i(\mu)E(\mu)x_i \tag{A9}$$

and the argument of DS, VII, 6.9 shows that for fixed y, $\beta_i(\mu)$ are analytic functions μ near enough to μ_0.

The equation for x'' may be solved in the form

$$\alpha_i(\mu) = d^{-1}(\lambda, \mu) \sum_{j=1}^{m} [\text{Adj } R(\lambda, \mu)]_{ij} \beta_j(\mu), \tag{A10}$$

where

$$R_{ij} = \lambda \delta_{ij} - t_{ji}(\mu), \tag{A11}$$

$$d(\lambda, \mu) = \det (R_{ij}). \tag{A12}$$

Each element of the adjoint matrix and also the determinant $d(\lambda, \mu)$ are analytic functions of λ, μ near enough to $1, \mu_0$.

The singular terms in the inverse of $[\lambda - T(\mu)]$ come from the expression for x''. We know that $d(1, \mu_0) = 0$, so that $d(1, \mu)$ must have an n-fold zero at $\mu = \mu_0$, but n is not related to the type of zero of $d(\lambda, \mu_0)$ at $\lambda = 1$, which is actually a ν-fold zero (DS, VII, 3.16). However, the residue operator A_{-n} acts in X_{μ_0}, and because of the analyticity of the adjoint matrix and $\beta_i(\mu)$, we have

$$A_{-n}y'' = \text{const} \times \sum_{i,j=1}^{m} [\text{Adj } R(1, \mu_0)]_{ij} \beta_j(\mu_0) E(\mu_0) x_i. \tag{A13}$$

However, this is the same expression that would be obtained for the residue of the inverse of $[\lambda - T(\mu_0)]$ when expanded in a Laurent series in λ. Thus, from DS, VII, 3.16, we have

$$A_{-n} = A[1 - T(\mu_0)]^{\nu-1} E(\mu_0). \tag{A14}$$

In the case in which we apply the theorem, that of an L_2 operator in Hilbert space, most of the conclusions of the theorem follow immediately from Fredholm theory, for the series for numerator and denominator are analytic functions of λ, μ.

[12] L. V. Kantorovich and G. P. Akilov, *Functional Analysis in Normed Spaces* (The Macmillan Company, New York, 1964), p. 519.

Phenomenological Theory of Magnetic Hysteresis

A. C. PIPKIN AND R. S. RIVLIN
Division of Applied Mathematics, Brown University, Providence, Rhode Island

(Received 16 June 1966)

Constitutive equations specifying magnetization as a functional of the magnetic field strength history are discussed. Limitations on the form of the response functional arising from material symmetries are described. It is pointed out that even in initially isotropic materials, an applied magnetic field lying always in one plane may produce magnetization normal to that plane. Various integral and differential approximations to the response functionals are considered.

1. INTRODUCTION

IN the present paper we discuss the formulation of constitutive equations suitable for the description of the gross observable behavior of magnetic materials and the analogous behavior of dielectric and ferroelectric materials. We are concerned here with the expression for the magnetization as a functional of the history of the magnetic field strength. Interactions between electric and magnetic effects, such as those considered in a previous paper[1] for cases in which memory effects are not important, are not considered, nor do we consider the coupling between mechanical and magnetic effects. We have discussed constitutive equations for magnetomechanical effects earlier,[2] but again only for cases in which past histories can be ignored.

The basic constitutive assumption which we use here (Sec. 2) is that the magnetization of a given particle of material at any instant is completely determinate if the magnetic field strength at that particle is known for all times up to the instant considered. The magnetization is thus a functional of the magnetic field history.

Any symmetries which the material may possess in its initial field-free state imply certain restrictions on the form of the response functional in the constitutive relation. These restrictions can be rendered explicit by following the methods devised by Green and Rivlin[3] and Rivlin.[4] The method is presented here (Sec. 9) in the generalized form given by Wineman and Pipkin.[5] Specific results for initially isotropic and transversely isotropic materials are presented as illustrations of the method. The analogous results for the crystal classes can be written down by using the tables of invariants given by Smith and Rivlin[6] and Smith[7] and following the directions given in Sec. 9. Although the procedure leads trivially to the desired canonical forms, the actual forms obtained are in certain cases quite complicated, and in our view, little purpose is served by listing them for each of the crystal classes. Rather, the form for any crystal class can be obtained when required by anyone wishing to investigate its implications.

Experience has shown that if an effect is allowed by material symmetry, then it can eventually be produced experimentally in a suitable material unless thermodynamic restrictions rule out the possibility of such behavior. In Sec. 3 we briefly discuss a transverse magnetic phenomenon which is possible in isotropic materials and which has not previously been reported, so far as we are aware.

With particular reference to ferromagnetic behavior, we show in Sec. 4 how the postulate of rate-independence can be made explicit in the constitutive equation. The results are analogous to those we have given[8] for stress-deformation relations in rate-independent materials with memory.

Assumptions about the smoothness of dependence of magnetization on the field history can lead to approximate expressions for the response functionals in terms of after-effect integrals. Such approximations are discussed in Sec. 5. The related descriptions of small effects superposed on large and a possible specialization to account for rate effects in nearly rate-independent cases are considered in Secs. 6 and 7, respectively. Differential approximations are discussed in Sec. 8.

Modifications necessary when electric rather than magnetic effects are to be considered are outlined in Sec. 10.

[1] A. C. Pipkin and R. S. Rivlin, J. Math. Phys. **1**, 542 (1960).
[2] A. C. Pipkin and R. S. Rivlin, Mem. Acad. Lincei **8**, 1 (1966).
[3] A. E. Green and R. S. Rivlin, Arch. Ratl. Mech. Anal. **1**, 1 (1957).
[4] R. S. Rivlin, Z. Angew. Math. Phys. **12**, 447 (1961).
[5] A. S. Wineman and A. C. Pipkin, Arch. Ratl. Mech. Anal. **17**, 184 (1964).
[6] G. F. Smith and R. S. Rivlin, Arch. Ratl. Mech. Anal. **15**, 169 (1964).
[7] G. F. Smith, Quart. Appl. Math. (to be published).
[8] A. C. Pipkin and R. S. Rivlin, Z. Angew. Math. Phys. **16**, 313 (1965).

Thermodynamic restrictions on the forms of the response functionals are not considered in the present paper, and we have accordingly omitted variables such as temperature or entropy from the constitutive equation, for notational convenience.

2. MEMORY FUNCTIONALS

We consider a particle of material which is initially free from magnetization. We are concerned with the relation between the magnetization $\mathbf{M}(t)$ at this particle at time t and the field strength $\mathbf{H}(t-\tau)$, $\tau \geq 0$, at times up to and including t. Leaving aside all other influences, we suppose that only one value of the magnetization is compatible with any given field history $\mathbf{H}(t-\tau)$, $\tau \geq 0$:

$$\mathbf{M}(t) = \mathcal{M}[\mathbf{H}(t-\tau)\,|_{\tau=0}^{\infty}]. \tag{2.1}$$

With the understanding that the argument function $\mathbf{H}(t-\tau)$ of the functional $\mathcal{M}$ is defined on the interval $0 \leq \tau < \infty$, henceforth we usually omit explicit mention of this in the notation.

In some of the following sections it is convenient to illustrate the discussion in terms of the more restricted form which the functional $\mathcal{M}$ must take if the material is initially isotropic. Material symmetry restrictions are considered in detail in Sec. 9. In particular, we show that if the material is initially isotropic, then $\mathcal{M}$ must be expressible in the form

$$\mathcal{M} = \mathcal{M}^{(1)}[\mathbf{H}(t-\tau); I_1, I_2] + \mathcal{M}^{(2)}[\mathbf{H}(t-\tau_1) \times \mathbf{H}(t-\tau_2); I_1, I_2], \tag{2.2}$$

where the functions I_1 and I_2 are defined by

$$I_1 = \mathbf{H}(t-\tau_1) \cdot \mathbf{H}(t-\tau_2),$$
$$I_2 = \mathbf{H}(t-\tau_1) \cdot \mathbf{H}(t-\tau_2) \times \mathbf{H}(t-\tau_3). \tag{2.3}$$

The functionals $\mathcal{M}^{(1)}$ and $\mathcal{M}^{(2)}$ are linear with respect to their respective first-argument functions. Their dependence on the functions I_1 and I_2 is arbitrary so far as restrictions imposed by material symmetry are concerned. The functional $\mathcal{M}^{(1)}$ has the further property that if $\mathbf{H}$ is expressed in terms of a constant basis $\mathbf{e}^{(i)}$ by $\mathbf{H} = H_i\mathbf{e}^{(i)}$, then

$$\mathcal{M}^{(1)}[\mathbf{e}^{(i)}H_i(t-\tau); I_1, I_2] = \mathbf{e}^{(i)}\mathcal{M}^{(1)}[H_i(t-\tau); I_1, I_2]. \tag{2.4}$$

The functional $\mathcal{M}^{(2)}$ has the same property with respect to its vector argument.

To make the meaning of the latter property more obvious, it is convenient to express the linear functionals $\mathcal{M}^{(\alpha)}$ as integrals. Then (2.2) becomes

$$\mathcal{M} = \int_0^\infty A[I_1, I_2; \tau]\mathbf{H}(t-\tau)\,d\tau + \int_0^\infty \int_0^\infty B[I_1, I_2; \tau_1, \tau_2]\mathbf{H}(t-\tau_1) \times \mathbf{H}(t-\tau_2)\,d\tau_1\,d\tau_2. \tag{2.5}$$

Here A and B are distributions with respect to τ or τ_1 and τ_2, which are functionals of the functions I_1 and I_2. The relation (2.5) is to be regarded only as a more transparent notation for the relation (2.2). There is no loss of generality in taking the kernel B in (2.5) to be antisymmetric with respect to interchange of τ_1 and τ_2.

3. TRANSVERSE EFFECTS

Before discussing possible mathematical representations of the basic constitutive functional $\mathcal{M}$, we consider briefly an interesting consequence of the fundamental assumption (2.1). More particularly, we consider the restricted form (2.2) which $\mathcal{M}$ must take in materials which have the highest possible degree of symmetry, those which are initially isotropic.

We note that if the axis of $\mathbf{H}(t-\tau)$ is independent of τ, as in the usual measurements of hysteresis loops, then $\mathbf{H}(t-\tau_1) \times \mathbf{H}(t-\tau_2)$ is zero, and $\mathbf{M}(t)$ has the same axis as $\mathbf{H}(\tau)$, as one would expect. This fact is more apparent from the less abstract notation (2.5).

However, we note further that if the axis of $\mathbf{H}(t-\tau)$ varies in direction, but lies always in the same plane, then $\mathbf{H}(t-\tau_1) \times \mathbf{H}(t-\tau_2)$ will be perpendicular to this plane for some values of τ_1 and τ_2. In this case $\mathbf{M}(t)$ will, in general, have a component perpendicular to the plane of $\mathbf{H}(\tau)$, and it should be possible to observe magnetization in this direction.

This conclusion is valid for an isotropic material, whether or not it has a center of symmetry. The possibility of this effect is a consequence of the fact that $\mathbf{M}$ and $\mathbf{H}$ are axial rather than polar vectors.

4. RATE-INDEPENDENT BEHAVIOR

The usual hysteresis diagram for a ferromagnetic material implies that the magnetization $\mathbf{M}(t)$ corresponding to a history $\mathbf{H}(t-\tau)$, $\tau \geq 0$, is independent of the rate at which that history is executed. This idealization neglects magnetic after-effect phenomena in which the magnetization continues to change for some time after a sudden change in the field strength. However, for sufficiently slow changes in the field strength, the idealization is a convenient one.

To make the hypothesis of rate-independence explicit in the constitutive equation, we begin by specifying the history $\mathbf{H}(t-\tau)$, $\tau \geq 0$, in terms of its path in $\mathbf{H}$-space and the rate of traversal of this path. The arc length $s(\tau)$ traversed up to time τ is defined by

$$s(\tau) = \int_{-\infty}^{\tau} [\dot{\mathbf{H}}(\xi) \cdot \dot{\mathbf{H}}(\xi)]^{\frac{1}{2}}\,d\xi, \tag{4.1}$$

where the dot denotes differentiation with respect to ξ. Because we restrict our attention to histories for

which $\mathbf{H} = 0$ initially, no question of convergence of the integral with respect to its infinite limit arises. We use $S = s(t)$ to denote the arc length at time t.

The field path in H-space may be described by specifying $\mathbf{H}$ parametrically in terms of the arc length s, as $\mathbf{H}(s)$, $0 \leq s \leq S$. The rate of traversal is specified by the function $s(\tau)$, $-\infty < \tau \leq t$. The constitutive assumption (2.1) is then equivalent to the relation

$$\mathbf{M}(t) = \mathcal{N}[\mathbf{H}(s)\,|_{s=0}^{S};\ s(t-\tau)\,|_{\tau=0}^{\infty}]. \tag{4.2}$$

Here it is convenient to specify the domains of the argument functions explicitly.

The specialization to rate-independent materials is now trivial. A rate-independent material is one for which the functional $\mathcal{N}$ is independent of $s(t-\tau)$, $\tau \geq 0$,

$$\mathbf{M}(S) = \mathcal{N}[\mathbf{H}(s)\,|_{s=0}^{S}]. \tag{4.3}$$

In the case of an initially isotropic material, the functional $\mathcal{N}$ must be of a form analogous to (2.2) or (2.5), with argument functions which are now defined on the range $0 \leq s \leq S$.

Relations in incremental form may be useful. Let $d\mathbf{M}(S)$ and $d\mathbf{H}(S)$ denote the changes in $\mathbf{M}$ and $\mathbf{H}$ from "time" S to time $S + dS$, and notice that $|d\mathbf{H}(S)/dS| = 1$. The field path up to time $S + dS$ is determined by the path up to time S, the direction $d\mathbf{H}(S)/dS$, and the increment dS. Hence, we obtain

$$d\mathbf{M}(S) = \mathcal{N}'[\mathbf{H}(s)\,|_{s=0}^{S};\ d\mathbf{H}(S)/dS]\,dS. \tag{4.4}$$

Here $\mathcal{N}'$ is a functional of the path $\mathbf{H}(s)$, $0 \leq s \leq S$, and a function of the direction $d\mathbf{H}(S)/dS$. An essential feature of the relation (4.4) is that $\mathcal{N}'$ need not be linear in $d\mathbf{H}(S)/dS$; linear dependence is in no way required by rate independence. A relation of the form (4.3) can be recovered by integrating (4.4),

$$\mathbf{M}(S) = \int_0^S \mathcal{N}'[\mathbf{H}(s)\,|_{s=0}^{S^*};\ d\mathbf{H}(S^*)/dS^*]\,dS^*. \tag{4.5}$$

Restrictions imposed by material symmetries on relations of the form (4.4) are easily derived, by following methods of the type to be described in Sec. 9. We note that the arc length s is invariant under orthogonal transformations. In the case of initially isotropic materials, the functional $\mathcal{N}'$ in (4.4) must be expressible in the form

$$\begin{aligned}\mathcal{N}' = {} & \mathcal{N}_1'[\mathbf{H}(s);\ J_1, J_2, J_3, J_4] \\ & + \mathcal{N}_2'[\mathbf{H}(s_1) \times \mathbf{H}(s_2);\ J_1, J_2, J_3, J_4] \\ & + \mathbf{H}'(S)\mathcal{N}_3'[J_1, J_2, J_3, J_4] \\ & + \mathbf{H}'(S) \times \mathcal{N}_4'[\mathbf{H}(s);\ J_1, J_2, J_3, J_4],\end{aligned} \tag{4.6}$$

where the functions J are defined by

$$\begin{aligned}&J_1 = \mathbf{H}(s_1) \cdot \mathbf{H}(s_2), \quad J_2 = \mathbf{H}(s_1) \cdot \mathbf{H}(s_2) \times \mathbf{H}(s_3), \\ &J_3 = \mathbf{H}'(S) \cdot \mathbf{H}(s), \quad J_4 = \mathbf{H}'(S) \cdot \mathbf{H}(s_1) \times \mathbf{H}(s_2),\end{aligned} \tag{4.7}$$

and $\mathbf{H}'(S) = d\mathbf{H}(S)/dS$. Each of the functionals $\mathcal{N}_1'$, $\mathcal{N}_2'$, and $\mathcal{N}_4'$ is linear and has the property (2.4) with respect to its vector argument. Integral notation for these functionals analogous to (2.5) can be introduced.

5. INTEGRAL APPROXIMATIONS

We now return to the general relation (2.1), and its isotropic form (2.5). It is possible to obtain more explicit characterizations of the kernels A and B in (2.5) by making one or another of various assumptions about the smoothness of the dependence of $\mathbf{M}(t)$ on the history $\mathbf{H}(t-\tau)$, $\tau \geq 0$. Such assumptions are directed toward justifying integral representations of the type

$$\begin{aligned}A = {} & A_0(\tau) + \int_0^\infty \int_0^\infty A_1(\tau, \xi_1, \xi_2) \\ & \times I_1(t-\xi_1, t-\xi_2)\,d\xi_1\,d\xi_2 \\ & + \int_0^\infty \int_0^\infty \int_0^\infty A_2(\tau, \xi_1, \xi_2, \xi_3) \\ & \times I_2(t-\xi_1, t-\xi_2, t-\xi_3)\,d\xi_1\,d\xi_2\,d\xi_3 \\ & + \int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty A_{11}(\tau, \xi_1, \xi_2, \xi_3, \xi_4) \\ & \times I_1(t-\xi_1, t-\xi_2)\,I_1(t-\xi_3, t-\xi_4) \\ & \times d\xi_1\,d\xi_2\,d\xi_3\,d\xi_4 + \cdots +,\end{aligned} \tag{5.1}$$

analogous to the Taylor series expansion of a function of a finite number of variables. The precise nature of the smoothness assumptions influences the nature of the kernels A_0, A_1, etc., the type of integration involved, and the sense in which the expansion approximates the functional A. Conversely, the assumption that an expansion of the type (5.1) is possible implies statements about the smoothness of the dependence of $\mathbf{M}(t)$ on $\mathbf{H}(t-\tau)$, $\tau \geq 0$, and we prefer to consider such expansions from this point of view.

It appears unlikely that highly nonlinear behavior such as saturation in ferromagnets can be described conveniently in terms of sums of multiple integrals. On the other hand, integral representations do appear to be useful in the description of small amplitude (i.e., linear or nearly linear) rate effects such as the frequency dependence of magnetic permeability in electromagnetic waves. Thus, for example, by discarding all terms which are nonlinear in the history $\mathbf{H}(t-\tau)$, $\tau \geq 0$, from (2.5) and (5.1) we obtain the usual after-effect integral

$$\mathbf{M}(t) = \int_0^\infty A_0(\tau)\mathbf{H}(t-\tau)\,d\tau. \tag{5.2}$$

If a model of this form is found to be useful at relatively low intensities in a given material, then a quadratic approximation may be adequate at somewhat higher intensities. Thus, with an expansion of the type (5.1) for the functional B in (2.5), we obtain

$$\mathbf{M}(t) = \int_0^\infty A_0(\tau)\mathbf{H}(t-\tau)\,d\tau + \int_0^\infty\int_0^\infty B_0(\tau_1,\tau_2)\mathbf{H}(t-\tau_1)\times\mathbf{H}(t-\tau_2)\,d\tau_1\,d\tau_2. \tag{5.3}$$

The cubic term in such an expansion is

$$\int_0^\infty\int_0^\infty\int_0^\infty A_1(\tau,\xi_1,\xi_2)\mathbf{H}(t-\tau)\times[\mathbf{H}(t-\xi_1)\cdot\mathbf{H}(t-\xi_2)]\,d\xi_1\,d\xi_2\,d\tau, \tag{5.4}$$

and higher-order terms may be written analogously.

6. SMALL EFFECTS SUPERPOSED ON LARGE EFFECTS

When the magnetization $\mathbf{M}_0(t)$ corresponding to some simple base history $\mathbf{H}_0(t-\tau)$, $\tau \geq 0$, is known, it may be convenient to express the magnetization $\mathbf{M}(t)$ corresponding to a slightly different history in terms of $\mathbf{M}_0(t)$ and the perturbation $\mathbf{h} = \mathbf{H} - \mathbf{H}_0$. With certain implicit smoothness assumptions, this can be accomplished by means of a Frechet expansion of the basic response functional $\mathcal{M}$,

$$\begin{aligned}\mathbf{M}(t) &= \mathcal{M}[\mathbf{H}_0(t-\tau)+\mathbf{h}(t-\tau)]\\ &= \mathcal{M}[\mathbf{H}_0(t-\tau)] + \delta\mathcal{M}[\mathbf{H}_0(t-\tau)\,|\,\mathbf{h}(t-\tau)] + O(\|\mathbf{h}\|^2).\end{aligned} \tag{6.1}$$

Here $\delta\mathcal{M}$ is the Frechet differential of $\mathcal{M}$ at the history $\mathbf{H}_0$, with respect to $\mathbf{h}$, and $\|\mathbf{h}\|$ is some appropriate norm of the perturbation history. The first term on the right is $\mathbf{M}_0(t)$. The Frechet differential is a linear functional of $\mathbf{h}$ whose components can conveniently be represented in the form of integrals as

$$\delta\mathcal{M}_i = \int_0^\infty M_{ij}[\mathbf{H}_0(t-\xi)\,|_{\xi=0}^{\infty};\ \tau]h_j(t-\tau)\,d\tau. \tag{6.2}$$

Here the kernels M_{ij} are distributions which are functions of the base history $\mathbf{H}_0$. The forms of these kernals are, as usual, subject to restrictions imposed by material symmetries.

Toupin and Rivlin[9] have used representations of this general type, including electrical effects, in the discussion of electro-magneto-optical effects. In this application, the base history $\mathbf{H}_0$ is independent of time, and the functionals over $\mathbf{H}_0$ reduce to ordinary functions.

7. NEARLY RATE-INDEPENDENT MATERIALS

An approach to the description of rate effects in ferromagnetic materials can be based on the formalism (6.1). For such materials we suppose that if the base history $\mathbf{H}_0$ involves sufficiently slow changes in the field strength, then the dependence on $\mathbf{H}_0$ in (6.1) can be regarded as rate-independent. Thus, with the notation introduced in Sec. 4, we obtain

$$M_i(t) = \mathcal{N}_i[\mathbf{H}(s)|_{s=0}^{S}] + \int_0^\infty N_{ij}[\mathbf{H}_0(s)|_{s=0}^{S};\tau]\times h_j(t-\tau)\,d\tau + O(\|\mathbf{h}^2\|). \tag{7.1}$$

The first term on the right might, for example, represent the magnetization found in the determination of a hysteresis curve with field path $\mathbf{H}_0(s)$, $0 \leq s \leq S$. The second term would then represent rate-dependent behavior such as that observed in small oscillations of the field strength about the basic field path $\mathbf{H}_0(s)$, $0 \leq s \leq S$, or transient effects observed when the basic field path is traversed in short jumps.

8. DIFFERENTIAL APPROXIMATIONS

When the kernels in an integral approximation such as (5.3) can be expressed in terms of a finite number of exponentials with various relaxation times, it may be possible to convert the integral representation into a differential equation relating $\mathbf{M}(t)$ and $\mathbf{H}(t)$. A simpler and more general approach is possible if the field history $\mathbf{H}(t-\tau)$, $\tau \geq 0$, is well represented, over an interval of τ which is large in comparison to the largest relaxation time, by its Taylor series expansion,

$$\mathbf{H}(t-\tau) = \sum \frac{1}{n!}(-\tau)^n\mathbf{H}^{(n)}(t). \tag{8.1}$$

In such cases, by using (8.1) in (2.5), we obtain a relation of the form

$$\mathbf{M}(t) = \sum \alpha_n\mathbf{H}^{(n)}(t) + \sum\sum \alpha_{mn}\mathbf{H}^{(m)}(t)\times\mathbf{H}^{(n)}(t), \tag{8.2}$$

where

$$(-1)^n n!\,\alpha_n = \int_0^\infty A[I_1, I_2;\tau]\tau^n\,d\tau \tag{8.3}$$

and

$$(-1)^{m+n}m!\,n!\,\alpha_{mn} = \int_0^\infty\int_0^\infty B[I_1, I_2;\tau_1,\tau_2]\tau_1^m\tau_2^n\,d\tau_1\,d\tau_2. \tag{8.4}$$

If B is antisymmetric with respect to interchange of τ_1 and τ_2, then $\alpha_{nm} = -\alpha_{mn}$. With integral expansions of the type (5.1) for A and B, the moments α can be expressed in a correspondingly more explicit form.

[9] R. A. Toupin and R. S. Rivlin, Arch. Ratl. Mech. Anal. 7, 434 (1961).

9. MATERIAL SYMMETRY RESTRICTIONS

We now return to the consideration of the restrictions which are imposed on the functional $\mathcal{M}$ in (2.1) by material symmetries. We note first that because $\mathbf{M}$ and $\mathbf{H}$ are axial vectors, their images under an orthogonal transformation $\mathbf{R}$ are $R\mathbf{RM}$ and $R\mathbf{RH}$, respectively, where R is the determinant of $\mathbf{R}$. Now, let $\mathbf{M}(t)$ be the magnetization associated with the field history $\mathbf{H}(t-\tau)$, $\tau \geq 0$, by (2.1), and let $\mathbf{R}$ be a symmetry transformation for the material. Then the magnetization corresponding to the history $R\mathbf{RH}(t-\tau)$, $\tau \geq 0$, is $R\mathbf{RM}(t)$,

$$R\mathbf{RM}(t) = \mathcal{M}[R\mathbf{RH}(t-\tau)]. \tag{9.1}$$

With (2.1), this implies that the functional $\mathcal{M}$ must be of such form as to satisfy the relation

$$\mathcal{M}[R\mathbf{RH}(t-\tau)] = R\mathbf{R}\mathcal{M}[\mathbf{H}(t-\tau)], \tag{9.2}$$

for each history $\mathbf{H}(t-\tau)$, $\tau \geq 0$, and each symmetry transformation $\mathbf{R}$ for the material considered. The relevant group of transformations $\{R\mathbf{R}\}$ is a subgroup of the proper orthogonal group and need not necessarily be a subgroup of the group $\{\mathbf{R}\}$.

The restrictions stated in implicit form by (9.2) can be made explicit in canonical forms such as that shown in (2.2) for isotropic materials. The form of such a representation depends upon the group of transformations $\{R\mathbf{R}\}$. The method of finding such canonical representations, described below, was originated by Green and Rivlin[3] and Rivlin,[4] and extended by Wineman and Pipkin[5] who removed some inessential function-theoretic limitations on the generality of the method.

Let $I_\alpha(\mathbf{a}, \mathbf{b}, \mathbf{c}, \cdots)$, $\alpha = 1, 2, \cdots, A$, be the typical multilinear invariants of the generic vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}, \cdots$, for the group $\{R\mathbf{R}\}$. These invariants can be found in tables given by Smith and Rivlin[6] and Smith.[7] Let $\mathbf{H}(t-\tau_n)$ be denoted, for abbreviation, by $\mathbf{H}_n$. By substituting $\mathbf{H}_1, \mathbf{H}_2, \cdots$, for $\mathbf{a}, \mathbf{b}, \cdots$, respectively, in the typical invariants I_α, we obtain a set of invariants of the form $I_\alpha(\mathbf{H}_1, \mathbf{H}_2, \cdots)$, $\alpha = 1, \cdots, A$. With $\boldsymbol{\varphi}$ an arbitrary auxiliary vector, we also obtain the basic form-invariant vectors $\mathbf{f}^{(\alpha)}$ defined by

$$f_i^{(\alpha)} = \partial I_\alpha(\boldsymbol{\varphi}, \mathbf{H}_1, \mathbf{H}_2, \cdots)/\partial \varphi_i. \tag{9.3}$$

We note that these vectors are independent of $\boldsymbol{\varphi}$. Then, the functional $\mathcal{M}$ satisfies (9.2) if and only if it can be expressed in the form

$$\mathcal{M} = \sum_{\alpha=1}^{A} \mathfrak{L}^{(\alpha)}[\mathbf{f}^{(\alpha)}; I_1, \cdots, I_A], \tag{9.4}$$

where $\mathfrak{L}^{(\alpha)}$ is a functional of the functions $\mathbf{f}^{(\alpha)}$ and $I_1, \cdots, I_A$, linear in $\mathbf{f}^{(\alpha)}$. Furthermore, $\mathfrak{L}^{(\alpha)}$ has the property that if $\mathbf{e}^{(i)}$ are constant vectors, then

$$\mathfrak{L}^{(\alpha)}[\mathbf{e}^{(i)} f_i^{(\alpha)}; I_1, \cdots, I_A] = \mathbf{e}^{(i)} \mathfrak{L}^{(\alpha)}[f_i^{(\alpha)}; I_1, \cdots, I_A]. \tag{9.5}$$

Thus, in component form, $\mathcal{M}$ is given by

$$\mathcal{M}_i = \sum_{\alpha=1}^{A} \mathfrak{L}^{(\alpha)}[f_i^{(\alpha)}; I_1, \cdots, I_A]. \tag{9.6}$$

It should be emphasized that $\mathfrak{L}^{(\alpha)}$ is the same functional for each choice of the subscript i.

As an illustration, we consider the case of isotropic materials. In this case $\{\mathbf{R}\}$ is the full or proper orthogonal group, and $\{R\mathbf{R}\}$ is consequently the proper orthogonal group. There are two typical invariants in this case, $\mathbf{a} \cdot \mathbf{b}$ and $\mathbf{a} \cdot \mathbf{b} \times \mathbf{c}$. Accordingly, the invariants $I_\alpha(\mathbf{H}_1, \mathbf{H}_2, \cdots)$, $\alpha = 1, 2$ are those listed in (2.3). The vectors $\mathbf{f}^{(\alpha)}$ are defined by

$$f_i^{(1)} = \partial(\boldsymbol{\varphi} \cdot \mathbf{H})/\partial \varphi_i = H_i(t-\tau) \tag{9.7}$$

and

$$f_i^{(2)} = \partial(\boldsymbol{\varphi} \cdot \mathbf{H}_1 \times \mathbf{H}_2)/\partial \varphi_i = [\mathbf{H}(t-\tau_1) \times \mathbf{H}(t-\tau_2)]_i. \tag{9.8}$$

The canonical form (9.4) for this case is shown in (2.2).

As a second illustration, we consider transversely isotropic materials. If we let the x_3 coordinate axis coincide with the axis of rotational symmetry, then the typical multilinear invariants are $\mathbf{a} \cdot \mathbf{b}$, $\mathbf{a} \cdot \mathbf{b} \times \mathbf{c}$, and a_3. It follows that the invariant functions $I_\alpha(\mathbf{H}_1, \mathbf{H}_2, \cdots)$, $\alpha = 1, 2, 3$, for this group are those defined in (2.3), together with

$$I_3 = H_3(t-\tau). \tag{9.9}$$

Correspondingly, the basic form-invariant vectors $\mathbf{f}^{(1)}$ and $\mathbf{f}^{(2)}$ are those given in (9.7) and (9.8), respectively, and $\mathbf{f}^{(3)}$ is a unit vector along the x_3 axis,

$$f_i^{(3)} = \delta_{3i}. \tag{9.10}$$

Then, the canonical form (9.4) for this case is

$$\begin{aligned}\mathcal{M} = {} & \mathfrak{L}^{(1)}[\mathbf{H}(t-\tau); I_1, I_2, I_2] \\ & + \mathfrak{L}^{(2)}[\mathbf{H}(t-\tau_1) \times \mathbf{H}(t-\tau_2); I_1, I_2, I_3] \\ & + \mathbf{f}^{(3)} \mathfrak{L}^{(3)}[1; I_1, I_2, I_3].\end{aligned} \tag{9.11}$$

The property (9.5) has been used in factoring out the constant argument $\mathbf{f}^{(3)}$ of $\mathfrak{L}^{(3)}$.

The 32 crystal classes can be divided into eleven families, such that the group $\{R\mathbf{R}\}$ is the same for classes in a given family. One class in each family has no reflectional symmetries, and for this class the transformation groups $\{\mathbf{R}\}$ and $\{R\mathbf{R}\}$ are the same. In order to use the tables of Smith and Rivlin[6] and Smith[7] to find the basic multilinear invariants for a given crystal class, one refers not to the table for

that class, but to the table for the associated class with rotational symmetries $\{\boldsymbol{R}\mathbf{R}\}$ only. The eleven families of crystal classes are listed in Table I, the

TABLE I. Families of crystal classes.

Family 1:	C_1, S_1
2:	C_2, C_{1h}, C_{2h}
3:	V, C_{2v}, V_h
4:	C_4, S_4, C_{4h}
5:	D_4, V_d, C_{4v}, D_{4h}
6:	C_3, C_{3i}
7:	D_3, C_{3v}, D_{3d}
8:	C_6, C_{3h}, C_{6h}
9:	D_6, D_{3h}, C_{6v}, D_{6h}
10:	T, T_h
11:	O, T_d, O_h

class in each family which has no reflectional symmetries being given first. The invariants for each family except number 11 are given in the paper by Smith and Rivlin[6]; those for family 11 are in Smith's paper.[7]

It was shown by Smith and Rivlin[10] that the strain-energy functions for elastic deformations of crystals take eleven different forms depending on the crystal class. The eleven families into which the crystal classes thus divide are the same as those obtained in the present paper.

10. ELECTRICAL EFFECTS

Many of the preceding remarks apply with minor changes to the formulation of constitutive equations for dielectric and ferroelectric behavior. The electric field strength $\mathbf{E}(t)$ at a given particle at time t is assumed to be uniquely determined when the dielectric displacement $\mathbf{D}(t-\tau)$, $\tau \geq 0$, at that particle at times up to and including t is known,

$$\mathbf{E}(t) = \mathfrak{E}[\mathbf{D}(t-\tau)]. \tag{10.1}$$

The restrictions imposed on the form of the functional $\mathfrak{E}$ by symmetry under a group of transformations $\{\mathbf{R}\}$ are stated in implicit form in the equation

$$\mathfrak{E}[\mathbf{R}\mathbf{D}(t-\tau)] = \mathbf{R}\mathfrak{E}[\mathbf{D}(t-\tau)], \tag{10.2}$$

and the solution to this equation can be written down by following the procedure outlined in Sec. 9. The complications arising from the fact that $\mathbf{M}$ and $\mathbf{H}$ are axial vectors do not appear in the present case, in which we deal with polar vectors $\mathbf{E}$ and $\mathbf{D}$. For example, in the case of materials which are initially holohedral isotropic (i.e., $\{\mathbf{R}\}$ is the full orthogonal group), the inner product is the only typical basic multilinear invariant. The constitutive relation (10.1) can accordingly be written in the more explicit form

$$\mathbf{E}(t) = \mathfrak{E}[\mathbf{D}(t-\tau);\ \mathbf{D}(t-\tau_1)\cdot\mathbf{D}(t-\tau_2)], \tag{10.3}$$

where $\mathfrak{E}$ is linear in $\mathbf{D}(t-\tau)$ and has the property (9.5). In (10.3) there is no term of the type which gives rise to the transverse magnetic effect discussed in Sec. 3.

The discussion in Secs. 4–8 applies with only slight modification to the response functional $\mathfrak{E}$, and need not be repeated.

ACKNOWLEDGMENT

The results presented in this paper were obtained in the course of research sponsored by the Office of Ordnance Research, United States Army, under Contract DA-31-124-ARO(D)-416 with Brown University.

[10] G. F. Smith and R. S. Rivlin, Trans. Am. Math. Soc. **88**, 175 1958).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Exponential Stability of Non-uniform Guiding Center Plasma

ABRAHAM KADISH
Courant Institute of Mathematical Sciences, New York University, New York, New York

(Received 14 April 1966)

Sufficient conditions are derived for the exponential stability of non-uniform collisionless guiding center plasma when the magnetic field in equilibrium is unidirectional. These conditions are local analogs of the criteria obtained for uniform plasma equilibria. Growth rates slower than exponential ones are considered.

1. INTRODUCTION

IN this article, we investigate the linear stability of a plasma in a strong magnetic field, employing the guiding center equations derived by Grad.[1]

Our definition of stability is taken to be the boundedness of solutions of the initial-value problem for the linearized system of equations. In the past, investigations of non-uniform plasma have characteristically been restricted to special directions of propagation, or small ratios of fluid to magnetic pressure.[2-4] No such restrictions are made here.

Two of the conditions which we derive for exponential stability were obtained, in the special case of Gaussian distributions, by Furth[5] and by Sagdeev and Vedenov.[6] Chandrasekhar, Kaufman, and Watson[7] could have, but did not, obtain similar inequalities using the marginal stability criteria which they employed for Gaussian distributions in their study of a uniform cylindrical plasma column with circular cross section. In general, however, if the equilibrium distributions are not monotone in the energy, an argument based on marginal stability is known to be incorrect.

The technique we employ for the derivation of exponential stability criteria is that of demonstrating the existence of a unique solution to a partial differential equation for a Laplace transformed variable. Criteria which are sufficient for exponential stability are obtained. Some of these conditions have been shown to be necessary for stability by Grad.[8] His definition of stability, however, is different from the one employed here. Our sufficient conditions are local analogs of those obtained for the stability of uniform plasma equilibria.[9]

It should be pointed out that the absence of exponentially growing modes does not, in general, imply boundedness of solutions to the initial-value problem. In fact, we show that when symmetry considerations dictate solutions which are independent of the direction along the equilibrium magnetic field lines (i.e., propagation orthogonal to this field), the non-uniformity of the equilibrium produces an algebraic instability. This instability is not an exponential one.

We also treat, as a limiting case of the technique employed here, the stability of uniform cylindrical plasma columns of arbitrary cross section.

2. EQUATIONS

The model we employ is given below and was derived by Grad.[1]

$$\left(\frac{\partial}{\partial t} + \nabla \cdot [\mathbf{U} + v\boldsymbol{\beta}]\right) f^{\pm} = -\frac{\partial}{\partial v}\left\{\left[\boldsymbol{\beta} \cdot \alpha^{\pm}\left(\frac{\nabla \cdot \mathbf{P}^{+}}{\rho^{+}} - \frac{\nabla \cdot \mathbf{P}^{-}}{\rho^{-}}\right) + \boldsymbol{\beta} \cdot \nabla(\tfrac{1}{2}U^2 - \mu B) + v(\mathbf{U} \cdot \mathbf{k})\right] f^{\pm}\right\}, \quad (1)$$

$$0 = (\partial \mathbf{B}/\partial t) + \nabla \times (\mathbf{B} \times \mathbf{U}), \quad (2)$$

$$0 = \rho \mathbf{B} \times \frac{d\mathbf{U}}{dt} + \mathbf{B} \times \nabla \cdot (\mathbf{P}^{+} + \mathbf{P}^{-}) - \frac{\mathbf{B}}{\mu_0} \times [(\nabla \times \mathbf{B}) \times \mathbf{B}]. \quad (3)$$

In these equations all quantities are time averaged over a Larmor period. The distributions $f^{\pm}$ are functions of μ and v, the magnetic moment per unit mass and the kinetic velocity parallel to the magnetic field $\mathbf{B}$ ($\boldsymbol{\beta} = \mathbf{B}/|\mathbf{B}|$). In this treatment they are normalized as mass densities ($\int f \, d\mu \, dv = \rho$). The quantities $\mathbf{P}^{\pm}$ and $\rho^{\pm}$ are pressure tensors and mass densities.

[1] H. Grad, in *Proceedings of the Symposium on Electromagnetics and Fluid Dynamics of Gaseous Plasma* (Polytechnic Institute of Brooklyn, Brooklyn, New York, 1961).
[2] A. B. Mikhailovskii, Zh. Eksperim. i Teor. Fiz. **44**, 1552 (1963) [English transl.: Soviet Phys.—JETP **17**, 1043 (1963)].
[3] N. A. Krall and M. N. Rosenbluth, Phys. Fluids **6**, 254 (1963).
[4] A. A. Galeev, Zh. Eksperim. i Teor. Fiz. **44**, 1920 (1963) [English transl.: Soviet Phys.—JETP **17**, 1292 (1963)].
[5] H. P. Furth, Phys. Fluids **6**, 48 (1963).
[6] R. Sagdeev and A. Vedenov, in *Plasma Physics and the Problem of Controlled Thermonuclear Reactions*, M. A. Leontovich, Ed. (Pergamon Press, Inc., New York, 1958), Vol. III, pp. 332–339.
[7] S. Chandrasekhar, A. N. Kaufman, and K. M. Watson, Proc. Roy. Soc. (London) **A245**, 435 (1958).
[8] H. Grad, Phys. Fluids **9**, 225 (1966).
[9] A. Kadish, Phys. Fluids **9**, 514 (1966).

(Superscripts + and − refer to ions and electrons, respectively.) The macroscopic flow velocity orthogonal to **B** is **U**, and **k** is the curvature of the magnetic field. The numbers $\alpha^{\pm}$ are defined by

$$\alpha^{\pm} = \frac{-(m/e)^{\mp}}{(m/e)^{+} - (m/e)^{-}}.$$

Equations (1) are the collisionless Liouville equations for the ion and electron guiding centers. Equations (2) are Maxwell's equations for the time evolution of the magnetic field in which, consistent with the lowest-order guiding center theory, the electric field has been replaced by $\mathbf{B} \times \mathbf{U}$. However, $E_{\|}$ is contained in the terms $(\nabla \cdot \mathsf{P})/\rho$ of the Liouville equations. The last equations are the magnetohydrodynamic equations for the components of velocity orthogonal to **B**. In these, the current has been replaced by $(1/\mu_0)\nabla \times \mathbf{B}$.

We now assume an equilibrium which is independent of the direction along its unidirectional magnetic field vector $\mathbf{B}_0 = (B_0, 0, 0)$ and linearize Eqs. (1)–(3) about this state. If the dependence of the solutions is harmonic in the direction of $\mathbf{B}_0$, then we obtain, after a Laplace transform in time, with $f^{\pm}_{0,v} = (\partial/\partial v) f^{\pm}_0$,

$$\begin{aligned}(v - z)g^{\pm} &+ \{B_{\|}/B_0[\alpha^{\pm}(P^0_{\perp}/\rho_0)_{\Delta} - \mu B_0] + \alpha^{\pm}\langle v^2\rangle\} f^{\pm}_{0,v} \\ &+ \mathbf{B}_{\perp}/B_0 \cdot (\alpha^{\pm}(\nabla^* P^0_{\|}/\rho_0)_{\Delta} \\ &+ \{\alpha^{\pm}[(P^0_{\perp} - P^0_{\|})/\rho_0]_{\Delta} - \mu B_0\}\nabla^* B_0/B_0) f_{0,v} \\ &+ (\mathbf{e} + v\mathbf{B}_{\perp}) \cdot \nabla^*(f^{\pm}_0/B_0) = (1/ik)g^{\pm}(0), \quad (4)\end{aligned}$$

$$-(z\mathbf{B}_{\perp} + \mathbf{e}) = (1/ik)\mathbf{B}_{\perp}(0),$$

$$-zB_{\|} + \nabla^* \cdot \mathbf{e} = (1/ik)B_{\|}(0), \quad (5)$$

$$-z\mathbf{e} - v_A^2 \mathbf{B}_{\perp} + (B_0/\rho_0)\nabla^* p = (1/ik)\mathbf{e}(0). \quad (6)$$

In Eqs. (4)–(6) $\mathbf{B}_{\perp}$ and $B_{\|}$ are the perturbations in **B** orthogonal and parallel to B_0, respectively. The linearization of the total orthogonal pressure

$$P_{\perp} + \tfrac{1}{2}B^2/\mu_0 = P_{\perp} + P_M$$

is given by p. The wavelength (along B_0) of the solutions is k^{-1} $(k > 0)$. The operator ∇^* is defined by

$$\nabla^* = \frac{1}{ik}\left(0, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}\right) = \frac{1}{ik}\nabla;$$

$z = s/ik$, where s is the Laplace transform variable. In addition, the following have been employed.

$$(q)_{\Sigma} = q^{+} + q^{-}, \quad (q)_{\Delta} = q^{+} - q^{-},$$

$$\mathbf{e} = B_0\mathbf{U}, \quad g^{\pm} = f^{\pm} - (B_{\|}/B_0)f^{\pm}_0,$$

$$\langle v^2\rangle = \int v^2\, d\mu\, dv(g^{+}/\rho_0^{+} - g^{-}/\rho_0^{-}),$$

$$v_A^2 = (1/\rho_0)(2P^0_M + P^0_{\perp} - P^0_{\|}).$$

In the above, a 0 superscript or subscript implies an equilibrium parameter. Its absence implies a perturbation. (In the treatment which follows, this subscript is often dropped for convenience when there is no possibility of confusion.) In the following, a bit of shorthand is convenient. We define the expressions

$$I_l = \int_{-\infty}^{+\infty} \frac{dv}{v - z} \int_0^{\infty} \mu^l \left[\left(\frac{f_{0,v}}{\rho_0^{2-l}}\right)^{+} + (-1)^l \left(\frac{f_{0,v}}{\rho_0^{2-l}}\right)^{-}\right] d\mu,$$

$$l = 0, 1, 2.$$

After taking the appropriate moments of Eqs. (4) and performing some elementary manipulations we find

$$\begin{aligned}\langle v^2\rangle &+ \left\{\frac{B_{\|}}{B_0}\left(\frac{P^0_{\perp}}{\rho_0}\right)_{\Delta}\right. \\ &+ \frac{B_{\perp}}{B_0} \cdot \left[\left(\frac{\nabla^* P^0_{\|}}{\rho_0}\right)_{\Delta} + \left(P^0_{\perp} - \frac{P^0_{\|}}{\rho_0}\right)_{\Delta}\frac{\nabla^* B_0}{B_0}\right] \\ &\left. + \frac{B_0}{(|\alpha|\,\rho_0)^{\pm}}\frac{I_1}{I_0}\nabla^* \cdot \left(\frac{B_{\perp}}{B_0}\right)\right\} = j_1, \quad (7)\end{aligned}$$

$$\rho_0(z^2 - v_A^2)(\mathbf{B}_{\perp}/B_0) + \nabla^* p = j_2,$$

$$[B_0^2(I_2 - I_1^2/I_0) + 2(P^0_M + P^0_{\perp})]\nabla^* \cdot (B_{\perp}/B_0) + p = j_3,$$

where j_l $(l = 1, 2, 3, \cdots)$ are functionals of the equilibrium and the initial data. From these equations, we readily obtain

$$\nabla \cdot \left(\frac{\nabla p}{D_{\|}(x; z)}\right) + k^2 \frac{p}{D_{\perp}(x; z)} = j_4, \quad (8)$$

where

$$D_{\|}(x; z) = \rho_0(x)[z^2 - v_A^2(x)],$$

$$D_{\perp}(x; z) = 2(P^0_M + P^0_{\perp}) + B_0^2(x)(I_2 - I_1^2/I_0).$$

The motivation for the notation is to be found in the earlier treatment of the homogeneous plasma.[9] There, stability of uniform equilibria against spatially harmonic perturbations with wave vectors $\mathbf{k} = (k_{\|}, k_{\perp}, 0)$ and $\mathbf{B}_0 = (B_0, 0, 0)$ was considered. The dispersion relation obtained was

$$k_{\|}^2 D_{\|}(z) - k_{\perp}^2 D_{\perp}(z) = 0, \quad \operatorname{Im} z > 0. \quad (9)$$

The existence of such a z was shown to be necessary and sufficient for exponential instability.

We now determine conditions under which Eq. (8) has a unique solution which is analytic in z for $\operatorname{Im} z > 0$. As an inversion of the Laplace transform readily shows, this condition corresponds to an absence of exponential growth.

3. UNIQUENESS

The result we derive in this section is the following.

Theorem: If $\int_0^{\infty} \mu^k f^{\pm}_{0,v}(\mu, v)\, d\mu$ are Hölder continuous in v for $k = 0, 1, 2$ and if

$$f^{\pm}_0(\mu, v) = a^{\pm}(x)g^{\pm}\{b_{\pm}(x)\mu, c^2v^2\},$$

then the following set of conditions is sufficient for the uniqueness of solutions of Eq. (8):

(1) $$v_A^2(x) > 0,$$

(2) $$D_\perp(x; 0) > 0,$$

(3)
$$\left\{\int_0^\infty \mu\, d\mu\left(\frac{f_{0,v}}{\rho_0}\right)_\Delta\right\}^2 < \left\{\int_0^\infty d\mu\left(\frac{f_{0,v}}{\rho_0^2}\right)_\Sigma\right\}\left\{\int_0^\infty \mu^2\, d\mu(f_{0,v})_\Sigma\right\},$$
for $v \neq 0$,

(4) $$I_0 \neq 0, \quad \operatorname{Im} z \geq 0.$$

The restriction on the partial moments of $f_{0,v}^{\pm}$ in the hypothesis permits a definition of I_k which is continuous up to the line $\operatorname{Im} z = 0$ when taken through values in the domain $\operatorname{Im} z > 0$. Actually, for exponential stability it suffices to have the integrals as limits of sequences of Hölder continuous functions. This is easily seen using the maximum principle for analytic functions. The second restriction of $f_0^{\pm}$ to functions of the form prescribed is more serious. It states that the spatial dependence of the equilibrium distributions may appear only as a normalization factor and as a coefficient of the magnetic moment per unit mass. The major effect of this restriction is to keep the ratio of $P_{\|}^0/\rho_0$ independent of space. This restriction is necessary for the success of the treatment which follows. It might be relaxed if a particular distribution pair and the spatial dependence of $P_{\|}^0/\rho_0$ were given. However, the technique we employ here does not yield general results without this restriction.

Condition (1) is easily recognized as the fire-horse inequality,

$$2P_M^0 + P_\perp^0 > P_{\|}^0.$$

Condition (2) is the velocity gradient inequality (a generalization of the mirror machine inequality)

$$P_M^0 + P_\perp^0 > \tfrac{1}{6}(P^{02}/P_{\|}^0)\eta,$$

where η is a dimensionless functional of the scaled distributions. Condition (3), if satisfied, guarantees that under the conditions of the hypothesis the integrand of the integral which appears in Eq. (11) is never real and negative. Condition (4) assures us that if the vector field **B** does not grow then the moment $\langle v^2 \rangle$ is exponentially stable. It also ensures the analyticity in z of the expression

$$D_{\|} - \lambda^2 D_\perp.$$

As shown in our earlier paper, a necessary condition for the validity of condition (4) is that

$$\int_{-\infty}^{+\infty} dv \int_0^\infty d\mu\left[\left(\frac{\partial}{\partial v^2}\frac{f_0}{\rho_0^2}\right)^+ + \left(\frac{\partial}{\partial v^2}\frac{f_0}{\rho_0^2}\right)^-\right] < 0. \quad (10)$$

Taken together with condition (3), this is sufficient. For functions which are nonincreasing in v^2 or which are isotropic in velocity space, this inequality is necessary and sufficient for (4) to obtain the following proof.

Proof: The proof of the theorem is immediate. We assume that for $z = z_0$, there is a function ϕ such that

$$\nabla \cdot \frac{\nabla\phi}{D_{\|}(z_0)} + k^2 \frac{\nabla\phi}{D_\perp(z_0)} = 0,$$

which vanishes at infinity. If this is so, then

$$\int\left[-\frac{|\nabla\phi|^2}{D_{\|}(z_0)} + k^2\frac{|\phi|^2}{D_\perp(z_0)}\right] d\mathbf{x} = 0.$$

Consider the function

$$F(z) = \int \frac{d\mathbf{x}}{I(z_0)}\left[I_0(z)D_{\|}(z)\left|\frac{\nabla\phi}{D_{\|}(z_0)}\right|^2 - k^2 I_0(z)D_\perp(z)\left|\frac{\phi}{D_\perp(z_0)}\right|^2\right]$$
$$= c(z_0)\int \frac{d\mathbf{x}}{\rho_0(x)}\left[I_0(z)D_{\|}(z)\,|\xi_1|^2 - k^2 I_0(z)D_\perp(z)\,|\xi_2|^2\right]. \quad (11)$$

Obviously $F(z_0)/c(z_0) = 0$, and on the arc $z = Re^{i\theta}$, $0 \leq \theta \leq \pi$; $R \gg 1$

$$\frac{F(z)}{c(z_0)} \sim \int d\mathbf{x}\,|\xi_1|^2 > 0.$$

We now proceed exactly as in Ref. 9 and find that if for z real and different from zero

$$\operatorname{Im} F(z)/c(z_0) = 0$$

and

$$\int \frac{d\mathbf{x}}{\rho} I_0' Q = \int \frac{d\mathbf{x}}{\rho}\cdot I_0'\Big\{D_{\|}\,|\xi_1|^2 - k^2[2(P_M + P_\perp) + B_0^2 I_2']\,|\xi_2|^2\Big\} \geq 0, \quad (12)$$

where I_k' is the principal value of I_k, then

$$\operatorname{Re} F(z)/c(z_0) > 0.$$

If the expression in Eq. (12) is negative, then using Schwarz's inequality on $\operatorname{Im} F(z)/c(z_0)$, some algebra yields that $\operatorname{Re} F(z)/c(z_0)$ is larger than the sum of a positive term and

$$\frac{1}{4\int k^2\,|\xi_2|^2\,B^2(\operatorname{Im} I_1)^2(d\mathbf{x}/\rho)} \times\left[\int I_0' Q\frac{d\mathbf{x}}{\rho}\int k^2\,|\xi_2|^2\,B^2(\operatorname{Im} I_1)^2\frac{d\mathbf{x}}{\rho} - \int(\operatorname{Im} I_0)Q\frac{d\mathbf{x}}{\rho}\int k^2\,|\xi_2|^2\,I_0'(B^2\operatorname{Im} I_2)\frac{d\mathbf{x}}{\rho}\right].$$

Employing the form of $f_0^{\pm}$ prescribed in the hypothesis,

we find that the bracketed expression may be written as

$$\frac{B^2}{P_\perp^2}\frac{P_\parallel}{\rho}\left(\int I_0' Q\,\frac{dx}{\rho}\right)\left(\int k^2\,|\xi_2|^2\,\frac{P_\perp^2}{P_\parallel}\,dx\right)$$
$$[(\operatorname{Im} I_1)^2 - (\operatorname{Im} I_0)(\operatorname{Im} I_2)] > 0.$$

Therefore, all intersections with the real line of the mapping of the boundary of the upper half z plane by the function $F(z)/c(z_0)$ lie to the right of the origin. [The intersection due to $z = 0$ is taken into account by conditions (1) and (2).] Therefore, the principle of the argument of analytic function theory implies that $F(z)/c(z_0)$ has no zeros in the upper half plane.

4. EXISTENCE AND ANALYTICITY

Assuming the uniqueness of p, we now prove its existence. We rewrite Eq. (8) in the form

$$\Delta p + k^2\frac{D_\parallel(\infty, z)}{D_\perp(\infty, z)}\,p = j_4^* + a_1(x, z)p + a_2(x, z)\cdot\nabla p, \quad (13)$$

$$a_1(x, z) = k^2\left[\frac{D_\parallel(\infty, z)}{D_\perp(\infty, z)} - \frac{D_\parallel(x, z)}{D_\perp(x, z)}\right],$$

$$\mathbf{a}_2(x, z) = D_\parallel(x, z)\nabla[D_\parallel(x, z)]^{-1},$$

$$j_4^* = D_\parallel(x, z)j_4.$$

We now take Fourier transforms of Eq. (12). After inverting the transforms, we find that

$$p(x; z) = \Psi(x; z) + \int K(|x - y|, y; z)p(y, z)\,d\mathbf{y}. \quad (14)$$

Here, $\Psi(x; z)$ represents the contribution of the inhomogeneity j_4^*.

$$2\pi K(|x - y|, \mathbf{y}, z) = [a_1(y, z) - \nabla\cdot a_2(y, z)] \times J(|x - y|) - a_2(y, z)\cdot\nabla J(|x - y|),$$

$$J(|x - y|) = -\int_0^\infty \frac{r\,dr}{r^2 - k^2[D_\parallel(\infty, z)/D_\perp(\infty, z)]}\,J_0(r\,|x - y|).$$

Here $J_0(t)$ is the zeroth-order Bessel function. It is easy to show that the kernel has a singularity at $|x - y| = 0$ of the form $|x - y|^{-1}$, but (assuming that the equilibrium does not tend to its state at infinity too slowly) is well behaved for $|x - y| \to \infty$. The singularity at $|x - y| = 0$ prevents the immediate application of the Fredholm alternative theorem for L_2 kernels. However, the iterated kernel is square integrable, and hence the theorem applies, provided the initial perturbations are L_2.

It should be noted that the existence of a solution for all z such that $\operatorname{Im} z > 0$ does not imply exponential stability. For this, one needs analyticity. Hence, we now show that the solution is analytic in z for

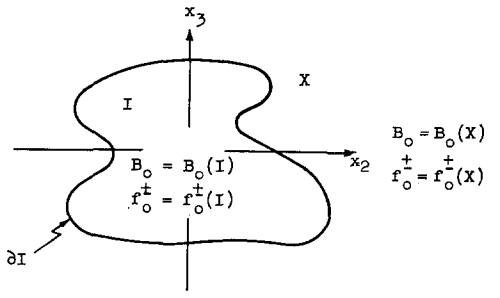


FIG. 1. The arbitrary cross section of the uniform plasma column I.

$\operatorname{Im} z > 0$. It suffices to show that dp/dz exists. Thus, consider the function ξ_δ defined as

$$\xi_\delta(z) = [\xi(z + \delta z) - \xi(z)]/\delta z.$$

We have

$$p_\delta(x, z) = \Psi_\delta^*(x; z) + \int dy K(|x - y|, y; z)p_\delta(y; z)$$

with

$$\Psi_\delta^*(x; z) = \Psi_\delta(x; z) + \int K_\delta(|x + y|, y, z)p(y, z + \delta z)\,dy.$$

Assuming reasonable smoothness of the equilibrium state, we may pass to the limit $\delta z \to 0$, thus obtaining the same integral equation (modulo the inhomogeneity) obtained for p.

5. STABILITY OF THE UNIFORM PLASMA COLUMN

We now consider an equilibrium such as that illustrated in Fig. 1. The equilibrium parameters are taken to be uniform in the domains I and X, but are assumed to suffer a discontinuity across the boundary of I; ∂I. The cross section of I is such that ∂I has a continuously turning tangent (piecewise smoothness of ∂I would suffice). Treating the equilibrium as the limiting case of a smoother configuration and using the familiar "pillbox" argument on the domain D (see Fig. 2) with Eq. (8), we find that across the line ∂I we must have

$$[n\cdot p/D_\parallel] = 0,$$

where $[q]$ is the jump in q across ∂I. $\mathbf{n}$ is the normal to $\partial\mathbf{I}$, drawn outward from $\mathbf{I}$. Therefore, if ϕ is a solution to the homogeneous equation for $z = z_0$, then proceeding as before, we find that $F(z)$ as defined by Eq. (11) must vanish at $z = z_0$. The remainder of the

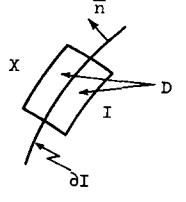


FIG. 2. The domain of the pillbox used to derive conservation conditions across ∂I.

stability argument now parallels that of the general case.

6. ALGEBRAIC STABILITY

We now consider a situation omitted in the previous discussion of stability—perturbations independent of the direction along $\mathbf{B}_0$. In this situation $k = 0$ and the linearized equations are given by

$$0 = (\partial B_{\parallel}/\partial t) + \nabla \cdot \mathbf{e}; \quad 0 = \partial \mathbf{B}_{\perp}/\partial t,$$
$$0 = (\rho_0/B_0)(\partial \mathbf{e}/\partial t) + \nabla p,$$
$$0 = \frac{\partial}{\partial t} g^{\pm} + (\mathbf{e} + v\mathbf{B}_{\perp}) \cdot \nabla g_0^{\pm} + g_0^{\pm} \frac{B^{\pm}}{B_0} \cdot \left\{ \alpha^{\pm} \left(\frac{\nabla P_{\parallel}^0}{\rho_0} \right)_{\Delta} + \left[\alpha^{\pm} \left(\frac{P_{\perp} - P_{\parallel}^0}{\rho_0} \right)_{\Delta} - \mu B_0 \right] \nabla B_0 \right\}. \tag{15}$$

The time invariance of $\mathbf{B}_{\perp}$ implied by Eqs. (15) indicates an algebraic growth in $g^{\pm}$ due to gradients in the equilibrium configuration. The realization of this growth when k is a transform variable is at present undetermined.

ACKNOWLEDGMENTS

The author wishes to thank Professor H. Grad for suggesting this problem, and Dr. E. M. Barston for many helpful discussions.

This work was supported by the United States Atomic Energy Commission under Contract AT(30-1)-1480.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Plane Wave Propagation in Kinetic Theory*

LAWRENCE SIROVICH AND JAMES K. THURBER
Division of Applied Mathematics, Brown University, Providence, Rhode Island
Department of Applied Mathematics, Brookhaven National Laboratories, Upton, New York

(Received 4 August 1966)

A number of results for plane wave propagation in kinetic theory are obtained. Among these are the following. It is shown that the Boltzmann equation for bounded collision operator has a cutoff frequency beyond which plane waves cease to exist. The analytic continuation of the dispersion relation is discussed and asymptotic results beyond the critical frequency are obtained. Using a kinetic model it is shown that the low-frequency formal power series for complex wavenumber in terms of frequency (and vice versa) are divergent. The last is contrary to the recent result obtained for the rigid sphere Boltzmann equation.

1. INTRODUCTION

ONE of the customary devices for examining the nature of a theory is through the study of plane wave propagation. In cases where a solution may be shown to be composed only of plane waves, such a study fully discloses the theory. In rarefied gas dynamics, a continuous spectrum appears in addition to the point spectrum and hence plane waves only partially reveal the theory. However, on physical grounds one expects the continuous spectrum not to play a role except at high frequencies.[1] Cercignani[2] and Weitzner[3] have shown the contrary to be true for the Krook model (this conclusion should hold for any model of the Gross–Jackson[4] type)—but this is a shortcoming of the model rather than a physical effect. In this paper our attention is focused on plane waves only, and more specifically on the low- and high-frequency regimes.

Following the general practice, we define a plane wave as a perturbation solution which has exponential space–time dependence $\sim e^{\sigma t + sx}$. For generality we regard both σ and s as complex; however, only two special cases are of physical interest. When $s = ik$, k real, the situation corresponds to a pure initial-value problem and we refer to this as a free wave. When $\sigma = i\omega$, ω real, the situation corresponds to an oscillating wall (boundary-value problem) and we refer to this as a forced wave. This terminology is

* This work was supported by the U.S. Office of Naval Research under Contract Nonr 562(39) and by the U.S. Atomic Energy Commission.

[1] We use the term frequency generically to signify a modulus of oscillation in either space or time.

[2] C. Cercignani, Ann. Phys. (N.Y.) **30**, 154 (1964).

[3] H. Weitzner, in *Proceedings of Fourth International Conference on Rarefied Gasdynamics* (Academic Press Inc., New York, 1965).

[4] E. P. Gross and E. A. Jackson, Phys. Fluids **5**, 432 (1959).

stability argument now parallels that of the general case.

6. ALGEBRAIC STABILITY

We now consider a situation omitted in the previous discussion of stability—perturbations independent of the direction along $\mathbf{B}_0$. In this situation $k = 0$ and the linearized equations are given by

$$0 = (\partial B_{\parallel}/\partial t) + \nabla \cdot \mathbf{e}; \quad 0 = \partial \mathbf{B}_{\perp}/\partial t,$$
$$0 = (\rho_0/B_0)(\partial \mathbf{e}/\partial t) + \nabla p,$$
$$0 = \frac{\partial}{\partial t} g^{\pm} + (\mathbf{e} + v\mathbf{B}_{\perp}) \cdot \nabla g_0^{\pm} + g_0^{\pm} \frac{B^{\pm}}{B_0} \cdot \left\{ \alpha^{\pm} \left(\frac{\nabla P_{\parallel}^0}{\rho_0} \right)_{\Delta} + \left[\alpha^{\pm} \left(\frac{P_{\perp} - P_{\parallel}^0}{\rho_0} \right)_{\Delta} - \mu B_0 \right] \nabla B_0 \right\}. \tag{15}$$

The time invariance of $\mathbf{B}_{\perp}$ implied by Eqs. (15) indicates an algebraic growth in $g^{\pm}$ due to gradients in the equilibrium configuration. The realization of this growth when k is a transform variable is at present undetermined.

ACKNOWLEDGMENTS

The author wishes to thank Professor H. Grad for suggesting this problem, and Dr. E. M. Barston for many helpful discussions.

This work was supported by the United States Atomic Energy Commission under Contract AT(30-1)-1480.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Plane Wave Propagation in Kinetic Theory*

LAWRENCE SIROVICH AND JAMES K. THURBER
Division of Applied Mathematics, Brown University, Providence, Rhode Island
Department of Applied Mathematics, Brookhaven National Laboratories, Upton, New York

(Received 4 August 1966)

A number of results for plane wave propagation in kinetic theory are obtained. Among these are the following. It is shown that the Boltzmann equation for bounded collision operator has a cutoff frequency beyond which plane waves cease to exist. The analytic continuation of the dispersion relation is discussed and asymptotic results beyond the critical frequency are obtained. Using a kinetic model it is shown that the low-frequency formal power series for complex wavenumber in terms of frequency (and vice versa) are divergent. The last is contrary to the recent result obtained for the rigid sphere Boltzmann equation.

1. INTRODUCTION

ONE of the customary devices for examining the nature of a theory is through the study of plane wave propagation. In cases where a solution may be shown to be composed only of plane waves, such a study fully discloses the theory. In rarefied gas dynamics, a continuous spectrum appears in addition to the point spectrum and hence plane waves only partially reveal the theory. However, on physical grounds one expects the continuous spectrum not to play a role except at high frequencies.[1] Cercignani[2] and Weitzner[3] have shown the contrary to be true for the Krook model (this conclusion should hold for any model of the Gross–Jackson[4] type)—but this is a shortcoming of the model rather than a physical effect. In this paper our attention is focused on plane waves only, and more specifically on the low- and high-frequency regimes.

Following the general practice, we define a plane wave as a perturbation solution which has exponential space–time dependence $\sim e^{\sigma t + sx}$. For generality we regard both σ and s as complex; however, only two special cases are of physical interest. When $s = ik$, k real, the situation corresponds to a pure initial-value problem and we refer to this as a free wave. When $\sigma = i\omega$, ω real, the situation corresponds to an oscillating wall (boundary-value problem) and we refer to this as a forced wave. This terminology is

* This work was supported by the U.S. Office of Naval Research under Contract Nonr 562(39) and by the U.S. Atomic Energy Commission.

[1] We use the term frequency generically to signify a modulus of oscillation in either space or time.

[2] C. Cercignani, Ann. Phys. (N.Y.) **30**, 154 (1964).

[3] H. Weitzner, in *Proceedings of Fourth International Conference on Rarefied Gasdynamics* (Academic Press Inc., New York, 1965).

[4] E. P. Gross and E. A. Jackson, Phys. Fluids **5**, 432 (1959).

the same used by us in our investigations of sound propagation.[5,6]

Formal power series methods can be employed in analyzing low-frequency plane waves, and these are by far the simplest to consider. This approach proved useful in a study which compared the Boltzmann equation, model equations, and various macroscopic approximations.[7] Such formal expansions, which can only be presumed valid for low frequencies, have in addition a variety of other uses, especially in solving problem. It is therefore natural to enquire as to the convergence of such expansions. In two recent studies of the rigid sphere Boltzmann equation,[8,9] it was shown that these expansions are convergent. In Sec. 6, using a kinetic model, we prove the contradictory result; i.e., the expansion is divergent.[10] The reason behind this conflict is that the collision frequency ν which is a constant for the model equations of Gross and Jackson,[4] behaves as $|\xi|$, $|\xi|$ large, for rigid spheres. For reasons given in Sec. 6, our conjecture is that the formal power series of $\sigma(k)$ and $s(\omega)$ diverge for $\nu(\xi) \to |\xi|^{\alpha}$, $0 \leq \alpha < 1$, $|\xi|$ large. And that they converge for $\nu(\xi) \to |\xi|$, $|\xi|$ large (exponents larger than unity are not of physical interest).

Detailed results for $\sigma(k)$, $s(\omega)$ for all values of the arguments do not seem attainable without numerical aid.[5,6] For large values of frequency, however, a variety of results may be obtained. In Sec. 3 we consider the finite moments method of Wang Chang and Uhlenbeck.[11] (This method, which is macroscopic in approach, is the same used in the exhaustive numerical investigations of Pekeris *et al.*[12]) We show in this case that for large ω, Im $s(\omega) \sim \omega$ and Re $s(\omega)$ is bounded. The latter effect accounts for the poor agreement this theory has with experiment.[5,6]

In considering the initial value problem for model equations, we have shown that plane wave propagation ceases past some critical frequency.[13] More recently Cercignani,[2] using the Krook model, has shown the same result for shear waves. In Sec. 4 we consider Boltzmann equations with bounded collision operators (this includes the kinetic models of Gross and Jackson[4]) and show these always have a critical frequency beyond which plane waves do not exist.

In Sec. 5 we consider the formulation of a specific problem and from it we introduce the analytic continuation of the dispersion relation. This results in relations for $\sigma(k)$ and $s(\omega)$ past the critical frequency. This notion has proven valuable in the numerical studies of Refs. 5 and 6, and in fact leads to plots of speed and attenuation rate which are in close agreement with experiment. Using the analytically continued dispersion relation we obtain asymptotic forms for $\sigma(k)$ and $s(\omega)$. In particular, for forced sound waves we obtain

$$s \sim \frac{-|\omega|}{[\ln \omega^2]^{\frac{1}{2}}}\left\{1 - \frac{3\pi i}{2 \ln \omega^2}\right\}.$$

2. FORMULATION

We follow the notation of an earlier paper.[5] In brief, the linearized Boltzmann equation, in one dimension, for the perturbed distribution function g is written as

$$\left(\frac{\partial}{\partial t} + \xi_1 \frac{\partial}{\partial x}\right) g = \mathfrak{L}(g) = \int f_*^0 [g'_* + g' - g_* - g] \times B(\theta, |\xi_* - \xi|)\, d\theta\, d\epsilon\, d\xi_*. \quad (2.1)$$

This equation is made dimensionless with respect to an as yet unspecified frequency $\nu = 1/\tau$, as follows:

$$t' = \nu t, \quad x' = x\nu/(RT_0)^{\frac{1}{2}},$$
$$\xi' = \xi/(RT_0)^{1/2}B' = \rho_0 \nu B/m. \quad (2.2)$$

Substituting and then removing the primes we obtain the dimensionless equation

$$\left(\frac{\partial}{\partial t} + \xi_1 \frac{\partial}{\partial x}\right) g = L(g) = \int \Omega_* [g'_* + g' - g_* - g] B\, d\theta\, d\epsilon\, d\xi_* \quad (2.3)$$

with

$$\Omega = e^{-\xi^2/2}/(2\pi)^{\frac{1}{2}}. \quad (2.4)$$

Introducing the inner product (asterisk denotes complex conjugate)

$$(p, q) = \int \omega p^*(\xi) q(\xi)\, d\xi, \quad (2.5)$$

one can show that for

$$\Psi_i = S^{r(i)}_{l(i)+\frac{1}{2}}(\xi^2/2)\xi^{l(i)} \times P_{l(i)}(\xi_1/\xi) \Big/ \left[\frac{2^{l(i)+1}\Gamma\left(r(i) + l(i) + \frac{3}{2}\right)}{\pi^{\frac{1}{2}} r(i)!\,(2l(i)+1)}\right]^{\frac{1}{2}}, \quad (2.6)$$

$$(\Psi_i, \Psi_j) = \delta_{ij}. \quad (2.7)$$

---

[5] L. Sirovich and J. K. Thurber, J. Acoust, Soc. Am. 37, 329 (1965).

[6] L. Sirovich and J. K. Thurber, in *Proceedings of Fourth International Conference on Rarefied Gasdynamics* (Academic Press Inc., New York, 1965).

[7] L. Sirovich, Phys. Fluids 6, 10 (1963).

[8] J. A. McLennan, Phys. Fluids 8, 1580 (1965).

[9] A. A. Arseniev, Zh. Vitch. Mat. i Mat. Phys. (Moscow) 5, 864 (1965).

[10] See also the papers by L. Sirovich and J. Thurber, in "Notes of a in Summer Conference Statistical Mechanics," Brookhaven National Laboratories (1965).

[11] C. S. Wang Chang and G. Uhlenbeck, "On the Propagation of Sound in Monoatomic Gases," Univ. Mich. Eng. M999 (1952).

[12] C. Pekeris, Z. Alterman, L. Finkelstein, and K. Frankowski, Phys. Fluids 5, 1608 (1962).

[13] L. Sirovich and J. K. Thurber, "Sound Propagation According to Kinetic Models," NYU Inst. Math. Sci. Rept. AFOSR-1380 MF-17 (1961) [see also *Rarefied Gas Dynamics* (Academic Press Inc., New York, 1963)].

S_q^p and P_l are the Laguerre and Legendre polynomials, respectively. The notation $r(i)$, $l(i)$ signifies that the double subscript rl has been reduced to a single subscript i.[14] For $B = B(\theta)$ Wang Chang and Uhlenbeck[11] have shown ψ_i to be eigenfunctions of L, i.e., for $B = B(\theta)$

$$L(\Psi_i) = \lambda_i \Psi_i, \tag{2.8}$$

where

$$\lambda_i = (\!(\Psi_i, L(\Psi_i))\!). \tag{2.9}$$

More generally we define

$$\Lambda_{ij} = (\!(\Psi_i, L(\Psi_j))\!). \tag{2.10}$$

We find it convenient to introduce the projection operator

$$P_N = \sum_{i \le N} \Psi_i(\Psi_i). \tag{2.11}$$

Hence

$$g = P_\infty g = \sum_{i=1}^{\infty} a_i \Psi_i \tag{2.12}$$

with

$$a_i = (\Psi_i, g). \tag{2.13}$$

Plane wave solutions to the Boltzmann equation are of the form

$$g = e^{\sigma t + sx} g(\xi), \tag{2.14}$$

where we allow both σ and s to be complex. In order for such solutions to exist we must find σ and s such that the homogeneous equation

$$(\sigma + s\xi_1 - L)g = 0 \tag{2.15}$$

is satisfied. Various aspects of this problem have been considered in the literature. With the exception of low-frequency phenomena, these results are generally obtainable only by numerical methods. In this paper we analytically show certain general features of both high- and low-frequency wave propagation.

In keeping with common practice we refer to the relation in σ and s which results from

$$(\sigma + s\xi_1 - L) = 0 \tag{2.16}$$

as the dispersion relation. As mentioned in the Introduction we refer to the case when $\mathrm{Re}\,(s) = 0$ as the free-wave problem, and to case when $\mathrm{Re}\,(\sigma) = 0$ as the forced-wave problem.

3. TRUNCATED BOLTZMANN EQUATION

In order to solve (2.16) Wang Chang and Uhlenbeck[11] consider the truncated Boltzmann equation

$$P_N(\sigma + s\xi_1 - L)P_N g = 0. \tag{3.1}$$

In considering the forced wave problem they show for Maxwell molecules, that the coefficients in the formal power series of $s = s(i\omega)$ can be exactly determined by successive truncations. Pekeris and co-workers[12] in a numerical investigation consider the exact dispersion relation which results from (3.1), for relatively large N. In this section we give an analytical description of the truncated dispersion relation (3.1) for high-frequency phenomena.

Equation (3.1) leads to the equivalent $N \times N$ coefficient system

$$(\sigma \mathbf{1} + s\mathbf{A} - \mathbf{\Lambda})\mathbf{a} = 0, \tag{3.2}$$

where

$$A_{ij} = (\Psi_i, P_N \xi_1 \Psi_j), \quad i,j \le N \tag{3.2a}$$

and $\mathbf{\Lambda}$ is defined by (2.10), and $\mathbf{a}$ by (2.13). It is convenient to define the finite inner product equivalent to (2.5)

$$\langle \mathbf{a}, \mathbf{b} \rangle = \sum_{i=1}^{N} a_i^* b_i. \tag{3.3}$$

We now observe that $\mathbf{\Lambda}$ is negative semidefinite,

$$\langle \mathbf{a}, \mathbf{\Lambda a} \rangle = (\!(P_N g, L(P_N g))\!) \le 0 \tag{3.4}$$

the latter relation being a well-known property of the Boltzmann equation. Also as is well known, equality in (3.4) holds if and only if $N = 3$.[15]

From this it follows that there exists an $(N-3)$-order square matrix $\tilde{\mathbf{S}}$ such that

$$\mathbf{S} = \begin{pmatrix} 1_3 & 0 \\ 0 & \tilde{\mathbf{S}} \end{pmatrix}$$

diagonalizes Λ under similarity (1_n denotes the $n \times n$ unit matrix),

$$\mathbf{S\Lambda S^{-1}} = \mathbf{D}.$$

$\mathbf{D}$ is a nonpositive diagonal matrix in which the first three diagonal elements are zero.

Because of its natural importance we focus attention on the forced wave problem. We therefore wish to determine $s = s(\omega)$ such that

$$0 = \det(i\omega \mathbf{1} + s\mathbf{A} - \mathbf{\Lambda}) = \det(\mathbf{1}i\omega + s\mathbf{SAS^{-1}} - \mathbf{D}). \tag{3.5}$$

Setting $s/i\omega = z$, we seek $z(\omega)$ for $\omega \to \infty$ in

$$\det[\mathbf{1} + z\mathbf{H} - (i/\omega)\mathbf{D}] = 0, \tag{3.6}$$

where $\mathbf{H}$ is the real symmetric matrix

$$\mathbf{H} = \mathbf{S^{-1}AS}.$$

Since the degree of (3.6) remains unchanged as

[14] The degree of Ψ_{rl} is $2r + l$. A variety of reductions can be based on the degree of the polynomials.

[15] Since the first three Ψ_i are linear combinations of the collision invariants.

$\omega \to \infty$, z possesses a power series in $1/\omega$.[16] Since $\mathbf{H}$ is real symmetric, and from the form of (3.6), we have

$$z(\infty) \neq 0,$$
$$\operatorname{Im} z(\infty) = 0.$$

Further it is easily shown that (3.6) is even in z. Finally, if we denote the normalized eigenvectors of the matrix of (3.5) by $\mathbf{u}$, we easily have

$$\operatorname{Im}(z) = -\langle \mathbf{u}, \mathbf{\Lambda u}\rangle/\omega\langle \mathbf{u}, \mathbf{Au}\rangle.$$

Since

$$\langle \mathbf{u}, \mathbf{\Lambda u}\rangle \leq 0$$

and

$$\langle \mathbf{u}, \mathbf{Au}\rangle \to -1/z(\infty),$$

we have

$$\operatorname{sgn}(\operatorname{Im} z) = -\operatorname{sgn}\big(z(\infty)\omega\big).$$

Therefore we have, for $\omega \to \infty$,

$$z = \frac{s}{i\omega} \sim \pm\left(\frac{1}{c_0} - \frac{i\alpha}{\omega}\right) \tag{3.7}$$

with c_0 and α nonnegative, or

$$s \sim \pm(i\omega/c_0 + \alpha). \tag{3.8}$$

This has the basic property of forced waves, i.e., that they decay in propagating to either $\pm\infty$. A noteworthy feature of the truncated Boltzmann equation is that it leads to Re s bounded as (3.8) indicates.

Denoting the viscosity of a gas by μ and the pressure by p, we specify the ν, (2.2), by

$$p/\mu.$$

Therefore, if $\bar{\omega}$ denotes the dimensional frequency we have

$$\omega = \bar{\omega}\mu/p = 1/r. \tag{3.9}$$

The quantity r was introduced by Greenspan[17] and is the parameter used in plots of sound results. We now note that as $\omega \to \infty$ the sound speed and attenuation rate are given by

$$\operatorname{Im} z/(\tfrac{5}{3})^{\frac{1}{2}} = a/a_0 \sim \text{const}, \tag{3.10}$$
$$\operatorname{Re} z \sim \tau$$

(a_0 denotes the adiabatic sound speed).

We note in passing that the free wave results follow if one inverts (3.7) and (3.8).

4. BOLTZMANN EQUATION WITH BOUNDED COLLISION OPERATOR

A useful class of equations in kinetic theory are the kinetic models.[4] To obtain these we first introduce

$$l = \operatorname{lub} -\langle \mathbf{x}, \mathbf{\Lambda x}\rangle/\langle \mathbf{x}, \mathbf{x}\rangle,$$

where

$$\Lambda_{ij} = (\Psi_i, \mathfrak{L}\Psi_j), \quad i, j \leq N + 1$$

[$\mathfrak{L}$ is the dimensional collision operator introduced in (2.1)]. Then taking

$$\nu = l$$

in (2.2), the kinetic models are defined to be

$$\left(\frac{\partial}{\partial t} + \boldsymbol{\xi}\cdot\frac{\partial}{\partial \mathbf{x}} + 1\right)g = P_N(L + 1)P_N g = Kg. \tag{4.1}$$

It is clear by construction that the norm $\|K\|$ of $\mathbf{A}$ satisfies

$$0 < \|K\| \leq 1. \tag{4.2}$$

The discussion which now follows also applies to Boltzmann equations for which K satisfies (4.2).

We focus attention on the forced wave problem

$$(i\omega + s\xi_1 + 1)g = Kg, \tag{4.3}$$

and for convenience we take

$$(g, g) = 1.$$

Taking the inner product of (4.3) with g, the real and imaginary parts may be written as

$$s_r = [-1 + (g, Kg)]/(g, \xi_1 g), \tag{4.4}$$
$$(g, \xi_1 g) = -\omega/s_i, \tag{4.5}$$

where we have written $s = s_s + is_i$. Equation (4.5) states that $(g, \xi_1 g) \geq 0$ for waves moving to the right and $(g, \xi_1 g) \leq 0$ for waves moving to the left. With (4.4) we have $s_r < 0$ for waves moving to the right and $s_r > 0$ for waves moving to the left. This is a basic feature of forced waves.

We can in full generality focus attention only on waves moving to the right, and also take $\omega > 0$. Defining

$$z = x + iy = (1 + i\omega)/s,$$

substituting for s in (4.4) and (4.5), and eliminating x, we find

$$y = -\omega(g, \xi_1 g)(g, Kg)/[\omega^2 + (g, Lg)^2]. \tag{4.6}$$

From this it is clear that $y \geq 0$.

We now show that $(g, \xi_1 g)$ is bounded away from zero. For if $(g, \xi_1 g) = 0$ then

$$i\omega = (g, Kg) - 1$$

implies

$$\omega = 0$$

and

$$g = a_1\Psi_1 + a_2\Psi_2 + a_3\Psi_3;$$

i.e., g is composed of only the collisional invariants, and therefore $(1 - K)g = 0$. The Boltzmann equation now becomes

$$s\xi_1 g = 0,$$

[16] If the roots are not distinct a Puiseux series may result.

[17] M. Greenspan, J. Acoust. Soc. Am. **28**, 644 (1956).

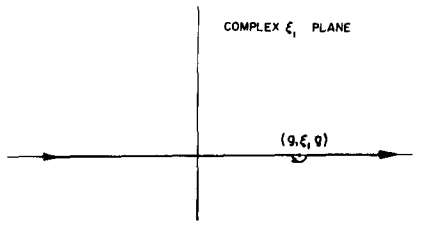


FIG. 1. Path of integration in the complex ξ_1 plane.

which implies a_i, $i = 1, 2, 3$ are constant. These may be eliminated by linearization about an appropriate Maxwellian in the definition of g.

We next show that $(1 + i\omega + s\xi)$, for $\omega \pm 0$, does not vanish for plane waves. For on substituting from (4.4) and (4.5)

$$1 + i\omega + s\xi_1 = 1 + \frac{-1 + (g, Kg)}{(g, \xi_1 g)}\xi_1 + i\left[\omega - \frac{\omega}{(g, \xi_1 g)}\xi_1\right]. \quad (4.7)$$

And hence this expression can be zero only for g, such that $(g, Kg) = 0$. But once K is nonnegative this implies that g belongs to the null space of K and for such a g,

$$(1 + i\omega + s\xi_1)g = 0.$$

This in turn implies that $g = 0$ except at the point $\xi_1 = (g, \xi_1 g)$ and hence $(g, g) = 1$ is violated. This proves our assertion. We may therefore form

$$g = [1/(1 + i\omega + s\xi_1)]Kg$$

and from this

$$1 = (g, [1 + i\omega + s\xi_1]^{-1}Kg) = \int \frac{\Omega g^* Kg\, d\xi}{1 + i\omega + s\xi_1}. \quad (4.8)$$

The argument leading to (4.6) then states that ω and s for plane waves stay to one side of the cut of the integral in (4.8).

Finally, we demonstrate that, in order for (4.7) to be satisfied, ω must be bounded. [Then (4.4) and (4.5) imply that s is also bounded.] Using (4.7) in (4.8), it is easily seen that as $\omega \to \infty$

$$\int \frac{\Omega g^* Kg\, d\xi}{1 + i\omega + s\xi_1} \sim \frac{1}{i\omega}\int \frac{\Omega g^* Kg\, d\xi}{1 - \xi_1/(g, \xi_1 g)},$$

where the path of integration is shown in Fig. 1. Therefore, using the Plemelj formula[18] we see that if $(g, \xi_1 g)$ is bounded as $\omega \to \infty$, the integral is $O(1/\omega)$ and if $(g, \xi_1 g)$ is unbounded the integral also vanishes since $(g, g) = 1$. Hence in either case (4.8) is violated as $\omega \to \infty$ and therefore ω and hence s must remain bounded for plane waves.

[18] In using the Plemelj formula we are assuming that g and Kg are, for example, Holder continuous.

A similar discussion can be achieved by even more elementary means for the free-wave problem.

The appearance of a cutoff in kinetic models for free-wave propagation was demonstrated in an earlier paper[13] and more recently was demonstrated by Cercignani[2] for the Krook equation for shear waves.

5. ANALYTIC CONTINUATION OF THE DISPERSION RELATION AND HIGH-FREQUENCY EXPANSIONS

The solution to both the initial- and boundary-value problems may be regarded as resulting from a normal mode analysis. Briefly, in this method one sums over the discrete modes and integrates over the continuum modes. For Boltzmann equations of the type considered in the previous section the discrete modes vanish at high frequencies and only the continuum modes remain. A more straightforward method of dealing with such problems is by means of transforms. The two approaches are, of course, equivalent.[19] For certain purposes the latter method seems superior. To anticipate the results of this section we first show that the discrete modes have an analytic continuation past the critical frequency. This is important since the cutoff frequency of the equations of Sec. 4 are mathematical rather than physical.[20] For example, the cutoff value of frequency for the Gross–Jackson type models increases with the index N in (4.1). The analytically continued mode furnishes an approximation of propagation speed and attenuation rate when a physical wave appears even though the theory excludes waves at such frequencies. It has been found for sound propagation that the analytic continuation is in excellent agreement with experiment.[5]

Since our only intention is to introduce the analytic continuation, we do so by means of the simplest problem. Consider the pure initial-value problem for

$$\left(\frac{\partial}{\partial t} + ik\xi_1 + 1\right)g = P_N(L + 1)P_N g = \sum_{m,n \le N} a_n \beta_{mn} \Psi_n,$$
$$g(t = 0) = g_0(k). \quad (5.1)$$

This is just (4.1) for "monochromatic" initial data. Introducing the Laplace transform, we obtain

$$(\sigma + ik\xi_1 + 1)g = \sum_{m,n \le N} a_m \beta_{mn} \Psi_n + g_0(k), \quad (5.2)$$

where the same letter has been used for the transformed dependent variable, and σ for the transform variable. Taking σ with a sufficiently large real part, we may

[19] K. M. Case, Ann. Phys. (N.Y.) 7, 349 (1959).

[20] Plane wave cutoff very possibly may occur and a "physically correct" Boltzmann should predict it. The equations of Sec. 4, however, predict the onset of "cutoff" at too low a frequency.

divide by $(\sigma + ik\xi_1 + 1)$ in (5.2) and form the moments

$$a_i = (\Psi_i, [\sigma + ik\xi_1 + 1]^{-1}[P_N(L+1)P_N g + g_0]). \tag{5.3}$$

Then introducing the following definitions,

$$u_i = (\Psi_i, \{\sigma + ik\xi_1 + 1\}^{-1} g_0),$$
$$C_{ij} = (\Psi_i, \{\sigma + ik\xi_1 + 1\}^{-1} \sum_{n \le N} \beta_{jn}\Psi_n, \quad i,j \le N \tag{5.4}$$

the solution for a_i $(i = 1, N)$ is given by

$$\mathbf{a} = \frac{1}{2\pi i}\int_{\uparrow} (\mathbf{1} - \mathbf{C})\mathbf{u}e^{\sigma t}\, d\sigma. \tag{5.5}$$

The arrow denotes the Bromwich path in the right half of the σ plane.[21]

Each of the entries in the matrix $\mathbf{C}$ is of the form

$$(1+\sigma)\int \frac{p(\xi)\Omega\, d\xi}{\sigma + 1 + s\xi_1} = \lambda \int \frac{p(\xi)\Omega\, d\xi}{\lambda - i\xi_1},$$
$$\lambda = (1+\sigma)/is, \tag{5.6}$$

where $p(\xi)$ is a polynomial in ξ. For generality we have introduced the complex wavenumber s. By straightforward manipulation, (5.6) may be reduced to a polynomial in λ plus a polynomial in λ multiplied by

$$M = \lambda \int \Omega\, d\xi/(\lambda - i\xi_1). \tag{5.7}$$

The explicit form of each C_{ij} can be given,[6] but this is of no interest here. The integral in (5.7) actually defines two functions M^+ and M^- for $\operatorname{Re}\lambda > 0$, $\operatorname{Re}\lambda < 0$, respectively. (This lies at the root of the cutoff shown in the last section.) On the other hand, in distorting the Bromwich path of (5.5), one makes use of the Cauchy integral theorem. This requires the analytic continuation of the integrand of (5.5) and hence of M.[22] For example, in the above initial-value problem if $k > 0$, M^- must be used in (5.5). In distorting the integration path to the left of $\operatorname{Re}\sigma = -1$ $(\operatorname{Re}\lambda < 0)$, M^- must still be used; i.e., the analytic continuation of the function M^- defined by (5.7). This yields the analytic continuation of the discrete spectra past the cutoff (i.e., in this case as k passes the critical wavenumber k_c). It should be noted, however, that the last section proves that no plane wave now actually exists in the theory—but for reasons mentioned earlier a plane wave may still physically exist.

In practice there is a somewhat simpler method of solving the above problem. If we form the system

$$(\Psi_i, \{\sigma + ik\xi_1 + 1\}g) = (\Psi_i, P_N(L+1)P_N g + g_0), \quad i = 1, N-2, \tag{5.8}$$

this system of $(N-2)$ equations by a proper arrangement of the Ψ_i, will only contain a_i for $i = 1, N$. Then, augmenting (5.8) by (5.3) for $i = N-1, N$ we have a determined system. The advantage to this mixed formulation is that the M function now occurs only in only two rows of the coefficient matrix of $\mathbf{a}$.

Using the analytic continuation we now obtain the high-frequency expansions for the roots of the dispersion relation. This is now carried out specifically for the Krook model $(N = 3)$.

$$\left(\frac{\partial}{\partial t} + \xi_1\frac{\partial}{\partial x} + 1\right)g = \Psi_1(\Psi_1, g) + \Psi_2(\Psi_2, g) + \Psi_3(\Psi_3, g). \tag{5.9}$$

Forming the matrix system according to the procedure already given we are then led to the consideration of the zeros of

$$\det(\mathbf{1} - \mathbf{C}) = 0.$$

For (5.9) this can be shown to be[13]

$$0 = 1 + \frac{i}{s}\left[\frac{\lambda^3(M-1)}{6} + \frac{11M}{6\lambda} - \frac{2\lambda M}{3} + \frac{5\lambda}{6}\right] - \frac{1}{s^2}\left[\frac{2M^2}{3\lambda^2} - \frac{2M(M-1)}{3} + \frac{\lambda^2(M-1)}{3} - \frac{2M}{3} + 1\right] + \frac{i}{s^3}\left[-\frac{2M(M-1)}{3\lambda} + \frac{\lambda(M-1)}{6} + \frac{M}{6}\right]. \tag{5.10}$$

The function M can be shown to be given by[6]

$$M^{\pm} = \pm(\tfrac{1}{2}\pi)^{\frac{1}{2}}\lambda e^{\lambda^2/2}[1 \mp \phi(\lambda/\sqrt{2})], \quad \operatorname{Re}\lambda \gtrless 0,$$

where

$$\phi(x) = \left(\frac{2}{\pi}\right)^{\frac{1}{2}}\int_0^{x\sqrt{2}} e^{-t^2/2}\, dt \tag{5.11}$$

is the error function.

The asymptotic expansions are easily found to be

$$M^{\pm}(\lambda) = 1 - \frac{1}{\lambda^2} + \frac{1\cdot 3}{\lambda^4} \mp \cdots, \quad |\arg\lambda| < \tfrac{3}{4}\pi,$$
$$M^{\pm}(\lambda) = (-)^{\frac{3}{2}\pm\frac{1}{2}}(2\pi)^{\frac{1}{2}}\lambda e^{\lambda^2/2} + 1 - \frac{1}{\lambda^2} + \frac{1\cdot 3}{\lambda^4} \mp \cdots,$$
$$\tfrac{1}{4}\pi < |\arg\lambda| < \tfrac{5}{4}\pi. \tag{5.12}$$

[21] The use of transforms for equations of the type (4.1) has been given a mathematically rigorous basis in the following: L. Sirovich and J. K. Thurber, Quart. Appl. Math. (to be published).

[22] Since this also requires the analytic continuation of u, this step cannot be accomplished unless g_0 is analytic. It is felt that this difficulty is mathematical rather than physical. One should be able to overcome this by approximating the data by analytic data (polynomials, for example). In Ref. 21 it is shown that the initial value is well posed for kinetic models, and the above remarks therefore follow.

It should be noted that the rays

$$\arg \lambda = \pm \tfrac{3}{4}\pi$$

are Stokes lines for M.

From (5.11) and (5.10) we see that λ bounded implies that s and σ are bounded. To find high-frequency roots we therefore restrict attention to large λ. Due to the nature of the asymptotics this must be carried out in three regions,

I. $|\arg \lambda| < \frac{3}{4}\pi$: Substituting (5.12) into (5.10) and solving the cubic we find

$$s \sim -\frac{i}{\lambda} + \frac{i}{\lambda^3},$$
$$s \sim -\frac{i}{\lambda} \pm \frac{1}{\lambda^2}\left(\frac{5}{3}\right)^{\frac{1}{2}} + \frac{i8}{3\lambda^3}. \tag{5.13}$$

Solving for the free wave roots

$$\sigma \sim -k^2,$$
$$\sigma \sim \pm(\tfrac{5}{3})^{\frac{1}{2}}ik - k^2, \tag{5.14}$$

and for the forced wave roots

$$s \sim \pm(i\omega)^{\frac{1}{2}}(1 + i\omega),$$
$$s \sim \pm(\tfrac{2}{5})^{\frac{1}{2}}i\omega(1 + 2i\omega). \tag{5.15}$$

By inspection we see that these are the low-frequency or Chapman–Enskog expansions[7] and we do not pursue these further.

II. $|\arg \lambda| \sim \frac{3}{4}\pi$. The previous discussion showed that no high-frequency roots lie along rays such that $|\arg \lambda| < \frac{3}{4}\pi$. There are, however, roots in the neighborhood of $|\arg \lambda| = \frac{3}{4}\pi$ as $\lambda \to \infty$. The analysis in this case is tedious and we consider instead the following model,[23]

$$\left(\frac{\partial}{\partial t} + \xi_1 \frac{\partial}{\partial x} + 1\right) g = \Psi_1(\Psi_1, g). \tag{5.16}$$

The dispersion relation for this model is easily found to be

$$s = -iM/\lambda. \tag{5.17}$$

It is sufficient to consider the ray $\arg \lambda = \frac{3}{4}\pi$, and the function M^+. Then along this ray

$$\lambda = \rho e^{i(3\pi/4)},$$
$$\rho = |\lambda|.$$

Substituting into (5.17) and using the appropriate asymptotic form of M^+, we have

$$is/(2\pi)^{\frac{1}{2}} = e^{-i(\rho^2/2)} + O(1/\rho), \tag{5.18}$$

and therefore

$$|is/(2\pi)^{\frac{1}{2}}| = 1 + O(1/\rho).$$

The free-wave roots are

$$k \sim (2\pi)^{\frac{1}{2}},$$
$$-\operatorname{Re} \sigma = \operatorname{Im} \sigma \sim 2\pi n^{\frac{1}{2}}, \tag{5.19}$$

with integers $n \to \infty$. This is relatively low-frequency propagation. For the forced-wave roots, on the other hand, we find

$$s \sim -i(2\pi)^{\frac{1}{2}}e^{-i\omega^2/4\pi}, \tag{5.20}$$

with $\omega \to \infty$.

The Krook model (5.9) has in addition to the above another set of qualitatively similar roots.[13]

III. $\pi > |\arg \lambda| > \frac{3}{4}\pi$. For this region the exponential is the dominant term in M. The solutions to the cubic in s [(5.10)] are (using M^+),

$$s = -i(2\pi)^{\frac{1}{2}}(\tfrac{1}{6}\lambda^4)(1 - 4/\lambda^2)e^{\lambda^2/2} + O(e^{\lambda^2/2}), \tag{5.21}$$

$$s = \frac{4i(2\pi)^{\frac{1}{2}}}{\lambda^2}\left(1 + \frac{3}{\lambda^2}\right)e^{\lambda^2/2} + O\left(\frac{e^{\lambda^2/2}}{\lambda^6}\right), \tag{5.22}$$

$$s = (i/\lambda) + O(1/\lambda^2). \tag{5.23}$$

The last root [(5.23)] is seen to be of low frequency and therefore is of no interest in this discussion. To solve (5.21) and (5.22) we take logarithms, for example, for (5.21) to obtain

$$2 \ln is = 2 \ln [\tfrac{1}{6}(2\pi)^{\frac{1}{2}}] + 8 \ln \lambda + \lambda^2 + 4\pi in + O(1/\lambda^2). \tag{5.24}$$

The branches of the logarithm appear through the term $4\pi in$ (n takes on integer values). The resulting equations may then be solved iteratively without difficulty. For free waves we find to lowest order in the real and imaginary part,

$$\sigma_n \sim -|k|\,(\ln k^2)^{\frac{1}{2}} - \pi in\,|k|/(\ln k^2)^{\frac{1}{2}}. \tag{5.25}$$

It is clear that σ_0 is the continuation of the hydrodynamical diffusion root (since this must be real). Numerical estimates may be made which show that $\sigma_{\pm 1}$ are the continuation of the hydrodynamical sound propagation roots.

For the forced-wave problem we find

$$s_n^{\pm} \sim -\frac{|\omega|}{(\ln \omega^2)^{\frac{1}{2}}} - i\frac{(4n \pm 1)}{2}\pi\frac{|\omega|}{(\ln \omega^2)^{\frac{3}{2}}}. \tag{5.26}$$

Here we consider only the waves moving to the right. The complex wavenumber s_0^+ corresponds to the hydrodynamical diffusion and s_1^- to the hydrodynamical sound propagation. We see therefore that

$$\text{speed of sound} \sim [\ln \omega^2]^{\frac{3}{2}}, \tag{5.27}$$
$$\text{attenuation rate} \sim |\omega|/[\ln \omega^2]^{\frac{1}{2}}. \tag{5.28}$$

[23] An extensive study of this model is to be found in Ref. 13. Since this model preserves (only) the continuity equation, it is referred to as the isosteric model.

This is a marked contrast to the asymptotic behavior found by means of the Wang Chang–Uhlenbeck method, see Eqs. (3.9), (3.10).

Although (5.26) results in particular from the Krook model, it holds with slight modification quite generally for models of the type (4.1). To see this recall the discussion in connection with (5.8). It was pointed out there that the dispersion relation is of degree 2 in the function M. Because of this one finds, as with (5.10), that there are only two high-frequency roots for $\pi \geq |\arg \lambda| > \frac{3}{4}\pi$. With the exception of "cross section" constants these take the same form as (5.21), (5.22) and lead to results similar to (5.25), (5.26).

The results contained in (5.25) and (5.26) can be used in a rough way to obtain a plot of speed and attenuation rate versus frequency. To do this, one plots the roots of the Navier–Stokes dispersion relation and the appropriate branches of (5.25) or (5.26) for all values of frequency. A patch of these two plots turns out to be in good agreement with experiment.

6. LOW-FREQUENCY EXPANSIONS

We now consider the nature of the formal power series expansions of $\sigma(s)$ [or $s(\sigma)$]. Our analysis is constructive and therefore, in order to avoid heavy analysis, we consider the following simplified model[23]:

$$[(\partial/\partial t) + \xi_1(\partial/\partial x) + 1]g = a_1\Psi_1. \tag{6.1}$$

Taking the inner product of (6.1) with $\Psi_1 = 1$ we obtain

$$a_1 = a_1 \int \frac{\Omega\, d\xi}{\sigma + s\xi_1 + 1}. \tag{6.2}$$

Carrying out the integration of ξ_2 and ξ_3 we are left with the dispersion relation [see (5.17)]

$$\sigma = \lambda \int_{-\infty}^{\infty} \frac{e^{-\zeta^2/2}\, d\zeta}{\lambda - i\zeta} - 1 = M(\lambda) - 1, \tag{6.3}$$

where as in the last section

$$\lambda = (1 + \sigma)/is. \tag{6.4}$$

As can be easily verified the function M satisfies

$$dM/d\lambda = (\lambda + 1/\lambda)M - \lambda. \tag{6.5}$$

Setting

$$-\kappa = s^2 \tag{6.6}$$

and using (6.3) and (6.4) we find after some manipulation

$$2\frac{d\sigma}{d\kappa} = \frac{\sigma}{\kappa} + \frac{1}{\kappa} + \frac{1}{\sigma}. \tag{6.7}$$

Further setting

$$\phi = \sigma/\kappa \tag{6.8}$$

we find

$$2\kappa^2\phi(d\phi/d\kappa) + \kappa\phi^2 = \phi + 1. \tag{6.9}$$

Formally writing

$$\phi = \sum_{r=0}^{\infty} a_r \kappa^r \tag{6.10}$$

and substituting we find

$$a_{r+1} = (r + 1)\sum_{j=0}^{r} a_{r-j}a_j. \tag{6.11}$$

For the isosteric model, only one discrete root exists,[13] and in fact $\phi = 0 = \kappa$ satisfies (6.3). From (6.9) we see that $\phi = -1$ for $\kappa = 0$, so that $a_0 = -1$. An induction argument immediately yields that $a_{2r} > 0$ and $a_{2r+1} < 0$. Induction also shows that $|a_r| > r!$, for $r > 1$. We therefore conclude that $\sigma(s)$ [or $s(\sigma)$] is not analytic at the origin.

This last result provides an interesting contrast to the recent work of MacLennan[8] and Arseniev[9] (an account of this is given in the review article of Guiraud[24]). They show for the rigid sphere Boltzmann equation that $\sigma(s)$ [or $s(\sigma)$] are analytic at the origin. Preliminary calculations indicate that analyticity hinges on the relative growths of $|\sigma + s\xi_1|$ and the collision frequency $\nu(\xi)$ for large $|\xi|$. The conjecture is that $\sigma(s)$ is not analytic for

$$\lim \nu(\xi) \to |\xi|^{\alpha}, \quad \alpha < 1,$$

and that analyticity is obtained for $\alpha = 1$. The latter corresponds to the rigid sphere case.

Note added in proof: Since submitting our paper for publication we have obtained a proof of the above conjecture. As a sidelight of this, one may show that, under requirements on the data, the Chapman–Enskog expansion is divergent for $\alpha < 1$ and convergent for $\alpha = 1$. These results will appear in a forthcoming paper.

[24] J. P. Guiraud, "Kinetic Theory and Rarefied Gas Dynamics," ONERA Rept. (1966); also to appear in *Fifth Symposium on Rarefied Gas Dynamics* (Academic Press Inc., New York, to be published).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Construction of the Irreducible Representations of SU_4

R. N. SEN*
Istituto di Fisica Teorica, University of Naples and Istituto Nazionale di Fisica Nucleare, Naples, Italy

(Received 22 March 1966)

The irreducible representations of SU_4 are explicitly constructed from those of SU_3 in the basis which is relevant for elementary particle physics. The technique seems to be applicable for the construction of the representations of SU_n from those of SU_{n-1} for any n.

1. INTRODUCTION

TWO of the crucial steps in the postwar development of strong interaction physics were: (i) the realization that strong interactions are characterized by the conservation (additive) of a new quantum number, and (ii) that the latter could be explained by the invariance of the interaction involved under a group larger than the isotopic rotation group. An invariance group makes more predictions than a conserved quantum number, and the group SU_3 has turned out to be suitable for describing strong interaction symmetries.

If there exists a shorter-range interaction of sufficient strength, or indeed an entire hierarchy of interactions of increasingly smaller range, it may be expected to reveal itself with the increase in available particle energies. In fact, this question will become urgent fairly soon, when the new generation of particle accelerators goes into action. Should such interactions be discovered, newer quantum numbers and larger symmetry schemes may be required to accommodate the data of experiments at higher energies. If, at the first stage, only one new quantum number is required, the choice narrows down to the Lie groups of rank 3.

Among the Lie groups of rank 3 the group SU_4 occupies a special position in view of the physical requirement that in any higher symmetry scheme the SU_3-quantum numbers should remain well defined[1]: The states within any irreducible representation of SU_4 can be characterized by the SU_3 quantum numbers $(\lambda\mu)YIK$ in current use, plus *only one* new quantum number, which would correspond to the conserved quantity in "suprastrong" interactions. The relationship of SU_4 to SU_3 is, in this sense, the same as that of SU_3 to SU_2, the isotopic rotation group. Moreover, the same relationship holds for any SU_n relative to SU_{n-1}. Thus the hierarchy of interactions giving rise to the fewest quantum numbers would correspond to a chain of unitary symmetries.

Physical states of a symmetric system can be classified according to *representations* of the symmetry group, and, within each irreducible representation, by the physically relevant quantum numbers. The construction of the irreducible representations of a symmetry group in the basis formed by the physical quantum numbers is therefore a problem of considerable practical significance.[2] In the present paper we solve this problem for SU_4 for possible application to elementary-particle physics. We take as our basic datum the known representations of SU_3, and endeavor to explain the technique in sufficient detail to allow calculation of the representations of any SU_n from those of SU_{n-1}. It is worth stating that no further information is required for these calculations; in particular, the Wigner–Eckart factorizations which are required in the work may be performed *ab initio*, as the normalizations of the Clebsch–Gordan coefficients involved are not relevant.

2. LIE ALGEBRA OF SU_3 AND ITS IRREDUCIBLE REPRESENTATIONS

In the present section we fix some notations and quote some known formulas for SU_3 for subsequent use. None of the finer points in the derivation of these formulas will be mentioned, as all of them reappear in our calculations for SU_4.

The eight operators which form the Lie algebra of SU_3 may be taken as $\mathbf{Y}$, $\mathbf{I}_\pm$, $\mathbf{I}_0$, $\mathbf{U}_\pm$, and $\mathbf{V}_\pm$, with the

* Present address: Department of Physics, University of Wisconsin–Milwaukee, Milwaukee, Wisconsin.

[1] SU_4 symmetry in elementary particles has been considered by several authors. See, for example, P. Tarjanne and V. L. Teplitz, Phys. Rev. Letters **11**, 447 (1963); P. Tarjanne, Phys. Rev. **136**, B1532 (1964); J. P. Antoine, D. Speiser, and R. J. Oakes, Phys. Rev. **141**, 1542 (1966).

[2] The quantum numbers dictated by physical considerations do not necessarily form a complete set, as for example in the application of SU_3 to nuclear physics. This may give rise to a host of complex problems. See, for example, G. Racah, in *Group-Theoretical Concepts and Methods in Elementary Particle Physics*, F. Gürsey, Ed. (Gordon and Breach Science Publishers, New York, 1964), pp. 31–36; H. Goldberg, "The α-content of Nuclei According to the Shell Model," thesis, Jerusalem (1961); R. N. Sen, "Construction of the Irreducible Representations of SU_3, thesis, Jerusalem (1963); V. Bargmann and M. Moshinsky, Nucl. Phys. **18**, 697 (1960); **23**, 177 (1961).

commutation relations:

$$\left.\begin{aligned}[\mathbf{I}_0, \mathbf{I}_\pm] &= \pm\mathbf{I}_\pm,\\ [\mathbf{I}_+, \mathbf{I}_-] &= 2\mathbf{I}_0.\end{aligned}\right\} \tag{2.1}$$

$$\left.\begin{aligned}[\mathbf{Y}, \mathbf{U}_\pm] &= \mathbf{U}_\pm,\\ [\mathbf{Y}, \mathbf{V}_\pm] &= -\mathbf{V}_\pm.\end{aligned}\right\} \tag{2.2}$$

$$\left.\begin{aligned}[\mathbf{I}_0, \mathbf{U}_\pm] &= \pm\tfrac{1}{2}\mathbf{U}_\pm,\\ [\mathbf{I}_+, \mathbf{U}_-] &= \mathbf{U}_+,\\ [\mathbf{I}_-, \mathbf{U}_+] &= \mathbf{U}_+.\end{aligned}\right\} \tag{2.3}$$

$$\left.\begin{aligned}[\mathbf{I}_0, \mathbf{V}_\pm] &= \pm\tfrac{1}{2}\mathbf{V}_\pm,\\ [\mathbf{I}_+, \mathbf{V}_-] &= \mathbf{V}_+,\\ [\mathbf{I}_-, \mathbf{V}_+] &= \mathbf{V}_-.\end{aligned}\right\} \tag{2.4}$$

$$\left.\begin{aligned}[\mathbf{V}_+, \mathbf{U}_+] &= \mathbf{I}_+,\\ [\mathbf{V}_-, \mathbf{U}_-] &= -\mathbf{I}_-.\end{aligned}\right\} \tag{2.5}$$

$$\left.\begin{aligned}[\mathbf{U}_+, \mathbf{V}_-] &= \mathbf{I}_0 + \tfrac{3}{2}\mathbf{Y},\\ [\mathbf{U}_-, \mathbf{V}_+] &= \mathbf{I}_0 - \tfrac{3}{2}\mathbf{Y}.\end{aligned}\right\} \tag{2.6}$$

It is understood that all commutators which have not been written down are zero. Note that, apart from the operators $\mathbf{I}_0$, $\mathbf{I}_\pm$, $\mathbf{Y}$ which form the Lie algebra of its U_2 subgroup, the group SU_3 contains two sets of infinitesimal operators which transform as tensors of rank $\frac{1}{2}$ under the isospin group $\mathbf{I}_0$, $\mathbf{I}_\pm$.

An irreducible representation of SU_3 is characterized by two positive integers $(\lambda\mu)$; the corresponding Young diagram has row lengths $\lambda + \mu$ and μ. Within an irreducible representation a complete labeling of the states is accomplished by the numbers Y, K, and I, where Y and K are, respectively, the eigenvalues of $\mathbf{Y}$ and $\mathbf{I}_0$, and $I(I+1)$ is the eigenvalue of $\vec{\mathbf{I}}^2$. In this scheme the operators $\mathbf{Y}$ and $\mathbf{I}_0$ are diagonal and the nonvanishing matrix elements of $\mathbf{I}_\pm$ are given by the well-known formulas

$$(YIK \pm 1\,|\mathbf{I}_\pm|\,YIK) = [(I \mp K)(I \pm K + 1)]^{\frac{1}{2}}. \tag{2.7}$$

The nonvanishing matrix elements of $\mathbf{U}_\pm$ are[3]

$$(Y + 1I + \tfrac{1}{2}K \pm \tfrac{1}{2}\,|\mathbf{U}_\pm|\,YIK)$$

$$= \frac{1}{6}\left\{\frac{(I \pm K + 1)(4\lambda + 2\mu - 3Y - 6I)(3Y + 6I + 2\lambda - 2\mu + 6)(3Y + 6I + 2\lambda + 4\mu + 12)}{6(2I+1)(2I+2)}\right\}^{\frac{1}{2}}, \tag{2.8a}$$

$$(Y + 1I - \tfrac{1}{2}K \pm \tfrac{1}{2}\,|\mathbf{U}_\pm|\,YIK)$$

$$= \mp\frac{1}{6}\left\{\frac{(I \mp K)(4\lambda + 2\mu - 3Y + 6I + 6)(2\mu - 2\lambda - 3Y + 6I)(3Y - 6I + 2\lambda + 4\mu + 6)}{6 \cdot 2I(2I+1)}\right\}^{\frac{1}{2}}. \tag{2.8b}$$

Each of these matrix elements is the product of a Wigner coefficient[4] and a reduced matrix element of $\mathbf{U}$. We have chosen the phases of the latter to make all of them positive, and the signs in the right-hand sides of Eqs. (2.8a, b) come only from the Wigner coefficients.

The matrices of $\mathbf{U}$ and $\mathbf{V}$ are related by the Hermiticity conditions $(\mathbf{U}_+)^\dagger = \mathbf{V}_-$, $(\mathbf{U}_-)^\dagger = -\mathbf{V}_+$. The nonvanishing matrix elements of $\mathbf{V}_\pm$ are

$$(Y - 1I + \tfrac{1}{2}K \pm \tfrac{1}{2}\,|\mathbf{V}_\pm|\,YIK)$$

$$= -\frac{1}{6}\left\{\frac{(I \pm K + 1)(4\lambda + 2\mu - 3Y + 6I + 12)(2\mu - 2\lambda - 3Y + 6I + 6)(3Y - 6I + 2\lambda + 4\mu)}{6(2I+1)(2I+2)}\right\}^{\frac{1}{2}}, \tag{2.9a}$$

$$(Y - 1I - \tfrac{1}{2}K \pm \tfrac{1}{2}\,|\mathbf{V}_\pm|\,YIK)$$

$$= \mp\frac{1}{6}\left\{\frac{(I \mp K)(4\lambda + 2\mu - 3Y - 6I + 6)(3Y + 6I + 2\lambda - 2\mu)(3Y + 6I + 2\lambda + 4\mu + 6)}{6 \cdot 2I(2I+1)}\right\}^{\frac{1}{2}}. \tag{2.9b}$$

[3] Y. Lehrer (unpublished, ca. 1950); R. N. Sen, Ref. 2; G. Racah, Ref. 2; G. E. Baird and L. C. Biedenharn, J. Math. Phys. 4, 1449 (1963).

[4] In this paper we use the term "Wigner coefficient" to denote the Clebsch–Gordan coefficient for the three-dimensional rotation group; the term "Clebsch–Gordan coefficient" itself refers only to other groups, e.g., SU_3.

3. TENSOR OPERATORS UNDER SU_3

It is obvious that a set of quantities $\mathbf{T}^{(\lambda\mu)}_{YIK}$ can be said to form a "tensor operator" of the type $(\lambda\mu)$ under SU_3 if they satisfy the commutation relations

$$[\mathbf{Y}, \mathbf{T}^{(\lambda\mu)}_{YIK}] = Y\mathbf{T}^{(\lambda\mu)}_{YIK}, \tag{3.1a}$$

$$[\mathbf{I}_0, \mathbf{T}^{(\lambda\mu)}_{YIK}] = K\mathbf{T}^{(\lambda\mu)}_{YIK}, \tag{3.1b}$$

$$[\mathbf{I}_\pm, \mathbf{T}^{(\lambda\mu)}_{YIK}] = [(I \mp K)(I \pm K + 1)]^{\frac{1}{2}}\mathbf{T}^{(\lambda\mu)}_{YIK\pm1}, \tag{3.1c}$$

$$[\mathbf{U}_\pm, \mathbf{T}^{(\lambda\mu)}_{YIK}] = a_\pm\mathbf{T}^{(\lambda\mu)}_{Y+1,I+\frac{1}{2},K\pm\frac{1}{2}} + b_\pm\mathbf{T}^{(\lambda\mu)}_{Y+1,I-\frac{1}{2},K\pm\frac{1}{2}}, \tag{3.1d}$$

$$[\mathbf{V}_\pm, \mathbf{T}^{(\lambda\mu)}_{YIK}] = c_\pm\mathbf{T}^{(\lambda\mu)}_{Y-1,I+\frac{1}{2},K\pm\frac{1}{2}} + d_\pm\mathbf{T}^{(\lambda\mu)}_{Y-1,I-\frac{1}{2},K\pm\frac{1}{2}}, \tag{3.1e}$$

where

$$a_\pm = (\lambda\mu Y + 1I + \tfrac{1}{2}K \pm \tfrac{1}{2}\,|\mathbf{U}_\pm|\,\lambda\mu YIK), \tag{3.2a}$$

$$b_\pm = (\lambda\mu Y + 1I - \tfrac{1}{2}K \pm \tfrac{1}{2}\,|\mathbf{U}_\pm|\,\lambda\mu YIK), \tag{3.2b}$$

$$c_\pm = (\lambda\mu Y - 1I + \tfrac{1}{2}K \pm \tfrac{1}{2}\,|\mathbf{V}_\pm|\,\lambda\mu YIK), \tag{3.2c}$$

$$d_\pm = (\lambda\mu Y - 1I - \tfrac{1}{2}K \pm \tfrac{1}{2}\,|\mathbf{V}_\pm|\,\lambda\mu YIK). \tag{3.2d}$$

Owing to the fact that in the decomposition of the direct product of two arbitrary irreducible representations of SU_3 many representations occur with multiplicities greater than unity, the systematic development of a tensor algebra for SU_3 is not as straightforward as for the three-dimensional rotation group. However, we are only interested in the representations (10) and (01) of SU_3, the direct products with which decompose as

$$(\lambda\mu) \times (10) = (\lambda + 1\mu) + (\lambda - 1\mu + 1) + (\lambda\mu - 1), \tag{3.3}$$

$$(\lambda\mu) \times (01) = (\lambda - 1\mu) + (\lambda + 1\mu - 1) + (\lambda\mu + 1). \tag{3.4}$$

It is therefore clear that the Wigner–Eckart theorem can be stated for tensors $\mathbf{T}^{(10)}$ and $\mathbf{T}^{(01)}$ under SU_3 exactly as for the three-dimensional rotation group. Therefore, in considering matrix elements, the selection rules on the SU_3 quantum numbers $(\lambda\mu)$ for these tensors are those which are implied by Eqs. (3.3) and (3.4).

4. LIE ALGEBRA OF SU_4

It can be easily verified from the vector diagram of SU_4 that the Lie algebra of SU_4 consists of:

(a) the operators $\mathbf{Y}$, $\mathbf{I}_0$, $\mathbf{I}_\pm$, $\mathbf{U}_\pm$, and $\mathbf{V}_\pm$ of SU_3;

(b) an operator $\mathbf{Q}$ which commutes with all the above;

(c) two sets of operators $(\mathbf{X}_0, \mathbf{X}_\pm)$ and $(\mathbf{Z}_0, \mathbf{Z}_\pm)$ which transform, respectively, as tensors of types (10) and (01) under SU_3; more precisely

$$\mathbf{X}_0 = \mathbf{T}^{(10)}_{-\frac{2}{3},0,0}, \quad \mathbf{X}_\pm = \mathbf{T}^{(10)}_{\frac{1}{3},\frac{1}{2},\pm\frac{1}{2}}, \tag{4.1}$$

$$\mathbf{Z}_0 = \mathbf{T}^{(01)}_{\frac{2}{3},0,0}, \quad \mathbf{Z}_\pm = \mathbf{T}^{(01)}_{-\frac{1}{3},\frac{1}{2},\pm\frac{1}{2}}, \tag{4.2}$$

where the subscripts stand for the Y, I, and K indices, in that order. The commutators of $\mathbf{X}$ and $\mathbf{Z}$ with the SU_3 operators can be read off from the above definitions, in conjunction with Eqs. (2.8), (2.9), (2.10), and (3.2). The remaining nonvanishing commutators are

$$[\mathbf{Q}, \mathbf{X}_0] = \mathbf{X}_0, \quad [\mathbf{Q}, \mathbf{X}_\pm] = \mathbf{X}_\pm, \tag{4.3}$$

$$[\mathbf{Q}, \mathbf{Z}_0] = \mathbf{Z}_0, \quad [\mathbf{Q}, \mathbf{Z}_\pm] = \mathbf{Z}_\pm; \tag{4.4}$$

(i.e., in a scheme in which $\mathbf{Q}$ is diagonal, the operators $\mathbf{X}$ increase the eigenvalue of $\mathbf{Q}$ by 1, and the operators $\mathbf{Z}$ decrease the eigenvalue of $\mathbf{Q}$ by 1).

$$[\mathbf{Z}_0, \mathbf{X}_+] = -\tfrac{1}{8}\mathbf{U}_+, \tag{4.5a}$$

$$[\mathbf{Z}_0, \mathbf{X}_-] = \tfrac{1}{8}\mathbf{U}_-, \tag{4.5b}$$

$$[\mathbf{X}_0, \mathbf{Z}_+] = -\tfrac{1}{8}\mathbf{V}_+, \tag{4.5c}$$

$$[\mathbf{X}_0, \mathbf{Z}_-] = -\tfrac{1}{8}\mathbf{V}_-, \tag{4.5d}$$

$$[\mathbf{X}_+, \mathbf{Z}_+] = \tfrac{1}{8}\mathbf{I}_+, \tag{4.5e}$$

$$[\mathbf{X}_-, \mathbf{Z}_-] = -\tfrac{1}{8}\mathbf{I}_-, \tag{4.5f}$$

and furthermore

$$[\mathbf{X}_0, \mathbf{Z}_0] = \tfrac{1}{8}\mathbf{Y} - \tfrac{1}{6}\mathbf{Q}, \tag{4.6a}$$

$$[\mathbf{Z}_+, \mathbf{X}_-] = \tfrac{1}{8}\mathbf{I}_0 - \tfrac{1}{16}\mathbf{Y} - \tfrac{1}{3}\mathbf{Q}, \tag{4.6b}$$

$$[\mathbf{Z}_-, \mathbf{X}_+] = \tfrac{1}{8}\mathbf{I}_0 + \tfrac{1}{16}\mathbf{Y} + \tfrac{1}{3}\mathbf{Q}. \tag{4.6c}$$

We write down explicitly a few of the commutation relations implied by (4.1) and (4.2), as we use them extensively in the following:

$$[\mathbf{V}_\pm, \mathbf{X}_0] = 0, \tag{4.7a}$$

$$[\mathbf{U}_\pm, \mathbf{X}_0] = \mathbf{X}_\pm, \tag{4.7b}$$

$$[\mathbf{U}_\pm, \mathbf{Z}_0] = 0, \tag{4.7c}$$

$$[\mathbf{V}_\pm, \mathbf{Z}_0] = -\mathbf{Z}_\pm. \tag{4.7d}$$

Finally, we give for reference the vector diagram of SU_4 and the identifications. The diagram is composed of the 12 vectors $\pm\boldsymbol{\alpha}$, $\pm\boldsymbol{\beta}$, $\pm\boldsymbol{\gamma}$, $\pm\boldsymbol{\lambda}$, $\pm\boldsymbol{\mu}$, $\pm\boldsymbol{\nu}$, where

$$\boldsymbol{\alpha} = (1/\sqrt{3}, 0, 0), \quad \boldsymbol{\beta} = (1/2\sqrt{3}, 1/2, 0),$$
$$\boldsymbol{\gamma} = (1/2\sqrt{3}, -1/2, 0), \quad \boldsymbol{\lambda} = (0, 1/3, -\sqrt{2}/3),$$
$$\boldsymbol{\mu} = (1/2\sqrt{3}, -1/6, -\sqrt{2}/3),$$
$$\boldsymbol{\nu} = (1/2\sqrt{3}, 1/6, \sqrt{2}/3).$$

The identifications are

$$\mathbf{I}_0 = \sqrt{3}\mathbf{H}_1, \mathbf{Y} = 2\mathbf{H}_2, \mathbf{Q} = (3/\sqrt{2})\mathbf{H}_3, \mathbf{I}_\pm = 2\sqrt{2}\mathbf{E}_{\pm\boldsymbol{\alpha}},$$
$$\mathbf{U}_+ = 2\sqrt{2}\mathbf{E}_{\boldsymbol{\beta}}, \mathbf{U}_- = 2\sqrt{2}\mathbf{E}_{-\boldsymbol{\gamma}}, \mathbf{V}_+ = 2\sqrt{2}\mathbf{E}_{\boldsymbol{\gamma}},$$
$$\mathbf{V}_- = 2\sqrt{2}\mathbf{E}_{-\boldsymbol{\beta}}, \mathbf{X}_+ = -\mathbf{E}_{-\boldsymbol{\nu}}, \mathbf{X}_- = \mathbf{E}_{-\boldsymbol{\mu}}, \mathbf{X}_0 = \mathbf{E}_{-\boldsymbol{\lambda}},$$
$$\mathbf{Z}_+ = \mathbf{E}_{\boldsymbol{\mu}}, \mathbf{Z}_- = \mathbf{E}_{\boldsymbol{\nu}}, \mathbf{Z}_0 = -\mathbf{E}_{\boldsymbol{\lambda}}.$$

5. BASIS FOR THE REPRESENTATIONS OF SU_4

As SU_4 is a 15-parameter group of rank 3, six quantum numbers are required (by Racah's theorem[5]) to characterize the states within an irreducible representation (fgh) (f, g, h are the projections of the highest weight along the fundamental dominant weights; row lengths of the corresponding Young diagram are $f+g+h$, $g+h$, h). If we first restrict SU_4 to U_3 we obtain the quantum numbers $Q(\lambda\mu)$; subsequent characterization of the states within the $(\lambda\mu)$-representation of SU_3 gives the usual quantum numbers Y, I, and K, thus completing the set. We write an SU_4 state within the (fgh) representation as $|Q\lambda\mu YIK)$, omitting the numbers f, g, and h unless they are required for clarity.

6. SELECTION RULES AND THE NONVANISHING MATRIX ELEMENTS OF SU_4 GENERATORS

In the above basis the matrix elements of the SU_3 operators are as earlier, and the operator $\mathbf{Q}$ is diagonal. Thus only the matrix elements of $\mathbf{X}$ and $\mathbf{Z}$ have to be calculated. From the various selection rules it is easily seen that there are 30 different nonvanishing matrix elements of $\mathbf{X}_0$, $\mathbf{X}_\pm$, $\mathbf{Z}_0$, and $\mathbf{Z}_\pm$. For instance, those of $\mathbf{X}_0$ and $\mathbf{X}_+$ are

$$(Q+1\lambda+1\mu\,Y-\tfrac{2}{3}IK\,|\mathbf{X}_0|\,Q\lambda\mu YIK),$$
$$(Q+1\lambda-1\mu+1\,Y-\tfrac{2}{3}IK\,|\mathbf{X}_0|\,Q\lambda\mu YIK),$$
$$(Q+1\lambda\,\mu-1\,Y-\tfrac{2}{3}IK\,|\mathbf{X}_0|\,Q\lambda\mu YIK),$$
$$(Q+1\lambda+1\mu Y+\tfrac{1}{3}I\pm\tfrac{1}{2}K+\tfrac{1}{2}\,|\mathbf{X}_+|\,Q\lambda\mu YIK),$$
$$(Q+1\lambda-1\mu+1\,Y+\tfrac{1}{3}I\pm\tfrac{1}{2}K+\tfrac{1}{2}|\mathbf{X}_+|\,Q\lambda\mu YIK),$$
$$(Q+1\lambda\mu-1\,Y+\tfrac{1}{3}I\pm\tfrac{1}{2}K+\tfrac{1}{2}\,|\mathbf{X}_+|\,Q\lambda\mu YIK).$$

The Wigner–Eckart theorem factorizes these 30 quantities into six reduced matrix elements, which do not involve Y, I, and K, and the SU_3 Clebsch–Gordan coefficients of the reduction of $(\lambda\mu)\times(10)$ and $(\lambda\mu)\times(01)$. (The latter can again be factorized into a so-called isoscalar factor and a Wigner coefficient.) While the coefficients involved can undoubtedly be picked up from the literature, it is simple and instructive to carry out the Wigner–Eckart factorization *ab initio*. This is our first step.

7. FACTORIZATION OF THE NONVANISHING MATRIX ELEMENTS

Let us start with $\mathbf{X}_0$. We take the matrix element of the commutator $[\mathbf{V}_+,\ \mathbf{X}_0]=0$ between the state $(Q+1\lambda+1\mu Y-\frac{5}{3}I+\frac{1}{2}K+\frac{1}{2}|$ on the left and the state $|Q\lambda\mu YIK)$ on the right. This gives

$$\begin{aligned}(Q+1\lambda+1\mu Y-\tfrac{5}{3}I+\tfrac{1}{2}K+\tfrac{1}{2}&\\ |\mathbf{V}_+|\,Q+1\lambda+1\mu Y-\tfrac{2}{3}IK)&\\ \times(Q+1\lambda+1\mu Y-\tfrac{2}{3}IK\,|\mathbf{X}_0|\,Q\lambda\mu YIK)&\\ =(Q+1\lambda+1\mu Y-\tfrac{5}{3}I+\tfrac{1}{2}K+\tfrac{1}{2}&\\ |\mathbf{X}_0|\,Q\lambda\mu Y-1I+\tfrac{1}{2}K+\tfrac{1}{2})&\\ \times(Q\lambda\mu Y-1I+\tfrac{1}{2}K+\tfrac{1}{2}\,|\mathbf{V}_+|\,Q\lambda\mu YIK).&\end{aligned}$$

Substituting the matrix elements of $\mathbf{V}_+$ and cancelling common factors, we have

$$\begin{aligned}(4\lambda+2\mu-3Y+6I+18)^{\frac{1}{2}}&\\ \times(Q+1\lambda+1\mu Y-\tfrac{2}{3}IK\,|\mathbf{X}_0|\,Q\lambda\mu YIK)&\\ =(4\lambda+2\mu-3Y+6I+12)^{\frac{1}{2}}(Q+1\lambda+1\mu&\\ Y-\tfrac{5}{3}I+\tfrac{1}{2}K+\tfrac{1}{2}\,|\mathbf{X}_0|\,Q\lambda\mu Y-1I+\tfrac{1}{2}K+\tfrac{1}{2})&\end{aligned}$$

or

$$\begin{aligned}&\frac{(Q+1\lambda+1\mu Y-\frac{5}{3}I+\frac{1}{2}K+\frac{1}{2}\,|\mathbf{X}_0|\,Q\lambda\mu}{}\\ &\qquad\qquad Y-1I+\tfrac{1}{2}K+\tfrac{1}{2})\\ &\overline{\qquad\qquad(4\lambda+2\mu-3Y+6I+18)^{\frac{1}{2}}}\\ &=\frac{(Q+1\lambda+1\mu Y-\frac{2}{3}IK\,|\mathbf{X}_0|\,Q\lambda\mu YIK)}{(4\lambda+2\mu-3Y+6I+12)^{\frac{1}{2}}}.\end{aligned}$$

Since the matrix elements of $\mathbf{X}_0$ must be independent of K, this is a relation of the type

$$f'(Q\lambda\mu,\ Y-1I+\tfrac{1}{2})=f'(Q\lambda\mu,\ YI),$$

which means that f' is a function of $3Y+6I$; thus

$$\begin{aligned}(4\lambda+2\mu-3Y+6I+12)^{-\frac{1}{2}}&(Q+1\lambda+1\mu Y-\tfrac{2}{3}\\ &IK\,|\mathbf{X}_0|\,Q\lambda\mu YIK)=f'(Q\lambda\mu,\,3Y+6I).\end{aligned}\tag{7.1}$$

Note that f' cannot be a function of Y or of I alone, except a constant.

Next, consider the matrix element

$$(Q+1\lambda+1\mu Y-\tfrac{5}{3}I-\tfrac{1}{2}K+\tfrac{1}{2}\ |[\mathbf{V}_+,\mathbf{X}_0]|\,Q\lambda\mu YIK)=0.$$

Exactly as above, this gives us the relation

$$\begin{aligned}(4\lambda+2\mu-3Y-6I+6)^{-\frac{1}{2}}&\\ \times(Q+1\lambda+1\mu Y-\tfrac{2}{3}IK\,|\mathbf{X}_0|\,Q\lambda\mu YIK)&\\ =(4\lambda+2\mu-3Y-6I+12)^{-\frac{1}{2}}(Q+1\lambda+1\mu&\\ Y-\tfrac{5}{3}I-\tfrac{1}{2}K+\tfrac{1}{2}\,|\mathbf{X}_0|\,Q\lambda\mu Y-1I-\tfrac{1}{2}K+\tfrac{1}{2})&\end{aligned}$$

or

$$\begin{aligned}(4\lambda+2\mu-3Y-6I+6)^{-\frac{1}{2}}&(Q+1\lambda+1\mu\ Y-\tfrac{2}{3}\\ &IK\,|\mathbf{X}_0|\,Q\lambda\mu YIK)=\varphi'(Q\lambda\mu,\,3Y-6I).\end{aligned}\tag{7.2}$$

Now, multiply (7.1) by $(4\lambda+2\mu-3Y-6I+6)^{-\frac{1}{2}}$, (7.2) by $(4\lambda+2\mu-3Y+6I+12)^{-\frac{1}{2}}$, write the resulting right-hand sides as $f(Q\lambda\mu,\,3Y+6I)$ and

[5] G. Racah, Rend. Lincei **8**, 108 (1950); in *Proceedings of the International Congress of Mathematicians* (American Mathematical Society, Providence, R.I., 1950), Vol. II, p. 467.

$\varphi(Q\lambda\mu, 3Y - 6I)$ respectively, and then compare the resulting equations. This gives

$$[(4\lambda + 2\mu - 3Y - 6I + 6) \times (4\lambda + 2\mu - 3Y + 6I + 12)]^{-\frac{1}{2}} \times (Q + 1\lambda + 1\mu\, Y - \tfrac{2}{3} IK\, |\mathbf{X}_0|\, Q\lambda\mu YIK) = f(Q\lambda\mu, 3Y + 6I) = \varphi(Q\lambda\mu, 3Y - 6I), \quad (7.3)$$

which means that the left-hand side is independent of Y and I. Thus we can write

$$(Q + 1\lambda + 1\mu\, Y - \tfrac{2}{3} IK\, |\mathbf{X}_0|\, Q\lambda\mu YIK) = [(4\lambda + 2\mu - 3Y - 6I + 6) \times (4\lambda + 2\mu - 3Y + 6I + 12)]^{\frac{1}{2}} \times (Q + 1\lambda + 1\mu\, \|\mathbf{X}\|\, Q\lambda\mu), \quad (7.4)$$

where the reduced matrix element is independent of Y and I.

It is now a simple matter to use the commutator $[\mathbf{U}_+, \mathbf{X}_0] = \mathbf{X}_+$ to obtain expressions for the matrix elements $(\cdots \lambda + 1\mu \cdots |\mathbf{X}_+| \cdots \lambda\mu \cdots)$ in terms of the reduced matrix element defined by (7.4) and known coefficients. The commutator $[\mathbf{I}_-, \mathbf{X}_+] = \mathbf{X}_-$ can be used to relate the matrix elements of $\mathbf{X}_-$ to those of $\mathbf{X}_+$, but in this case it is simpler to replace the Wigner coefficients by looking at tables. Similar calculations can be performed for all other matrix elements

$$\begin{pmatrix} \cdots \lambda - 1\mu + 1 \cdots |\mathbf{X}| \cdots \lambda\mu \cdots \\ \lambda\mu - 1 \end{pmatrix}$$

of $\mathbf{X}$, and those of $\mathbf{Z}_0$, $\mathbf{Z}_\pm$. The final results are as follows:

$$(Q + 1\lambda + 1\mu\, Y - \tfrac{2}{3} IK\, |\mathbf{X}_0|\, Q\lambda\mu YIK) = [(4\lambda + 2\mu - 3Y - 6I + 6)(4\lambda + 2\mu - 3Y + 6I + 12)]^{\frac{1}{2}} (Q + 1\lambda + 1\mu\, \|\mathbf{X}\|\, Q\lambda\mu), \quad (7.5a)$$

$$(Q + 1\lambda + 1\mu\, Y + \tfrac{1}{3} I + \tfrac{1}{2} K \pm \tfrac{1}{2} |\mathbf{X}_\pm|\, Q\lambda\mu YIK) = \left\{\frac{(I \pm K + 1)(3Y + 6I + 2\lambda - 2\mu + 6)(3Y + 6I + 2\lambda + 4\mu + 12)(4\lambda + 2\mu - 3Y + 6I + 12)}{6(2I + 1)(2I + 2)}\right\}^{\frac{1}{2}} \times (Q + 1\lambda + 1\mu\, \|\mathbf{X}\|\, Q\lambda\mu), \quad (7.5b)$$

$$(Q + 1\lambda + 1\mu\, Y + \tfrac{1}{3} I - \tfrac{1}{2} K \pm \tfrac{1}{2} |\mathbf{X}_\pm|\, Q\lambda\mu YIK) = \mp \left\{\frac{(I \mp K)(2\mu - 2\lambda - 3Y + 6I + 6)(3Y - 6I + 2\lambda + 4\mu + 6)(4\lambda + 2\mu - 3Y - 6I + 6)}{6 \cdot 2I(2I + 1)}\right\}^{\frac{1}{2}} \times (Q + 1\lambda + 1\mu\, \|\mathbf{X}\|\, Q\lambda\mu). \quad (7.5c)$$

$$(Q + 1\lambda - 1\mu + 1Y - \tfrac{2}{3} IK\, |\mathbf{X}_0|\, Q\lambda\mu YIK) = [(3Y + 6I + 2\lambda - 2\mu)(2\mu - 2\lambda - 3Y + 6I + 6)]^{\frac{1}{2}} (Q + 1\lambda - 1\mu + 1\, \|\mathbf{X}\|\, Q\lambda\mu), \quad (7.6a)$$

$$(Q + 1\lambda - 1\mu + 1Y + \tfrac{1}{3} I + \tfrac{1}{2} K \pm \tfrac{1}{2} |\mathbf{X}_\pm|\, Q\lambda\mu YIK) = -\left\{\frac{(I \pm K + 1)(4\lambda + 2\mu - 3Y - 6I)(3Y + 6I + 2\lambda + 4\mu + 12)(2\mu - 2\lambda - 3Y + 6I + 6)}{6(2I + 1)(2I + 2)}\right\}^{\frac{1}{2}} \times (Q + 1\lambda - 1\mu\, \|\mathbf{X}\|\, Q\lambda\mu), \quad (7.6b)$$

$$(Q + 1\lambda - 1\mu + 1Y + \tfrac{1}{3} I - \tfrac{1}{2} K \pm \tfrac{1}{2} |\mathbf{X}_\pm|\, Q\lambda\mu YIK) = \mp \left\{\frac{(I \mp K)(4\lambda + 2\mu - 3Y + 6I + 6)(3Y - 6I + 2\lambda + 4\mu + 6)(3Y + 6I + 2\lambda - 2\mu)}{6 \cdot 2I(2I + 1)}\right\}^{\frac{1}{2}} \times (Q + 1\lambda - 1\mu + 1\, \|\mathbf{X}\|\, Q\lambda\mu). \quad (7.6c)$$

$$(Q + 1\lambda\mu - 1Y - \tfrac{2}{3} IK\, |\mathbf{X}_0|\, Q\lambda\mu YIK) = [(3Y - 6I + 2\lambda + 4\mu)(3Y + 6I + 2\lambda + 4\mu + 6)]^{\frac{1}{2}} (Q + 1\lambda\mu - 1\, \|\mathbf{X}\|\, Q\lambda\mu), \quad (7.7a)$$

$$(Q + 1\lambda\mu - 1Y + \tfrac{1}{3} I + \tfrac{1}{2} K \pm \tfrac{1}{2} |\mathbf{X}_\pm|\, Q\lambda\mu YIK) = -\left\{\frac{(I \pm K + 1)(4\lambda + 2\mu - 3Y - 6I)(3Y + 6I + 2\lambda - 2\mu + 6)(3Y - 6I + 2\lambda + 4\mu)}{6(2I + 1)(2I + 2)}\right\}^{\frac{1}{2}} \times (Q + 1\lambda\mu - 1\, \|\mathbf{X}\|\, Q\lambda\mu), \quad (7.7b)$$

$$(Q + 1\lambda\mu - 1Y + \tfrac{1}{3} I - \tfrac{1}{2} K \pm \tfrac{1}{2} |\mathbf{X}_\pm|\, Q\lambda\mu YIK) = \pm \left\{\frac{(I \mp K)(2\mu - 2\lambda - 3Y + 6I)(3Y + 6I + 2\lambda + 4\mu + 6)(4\lambda + 2\mu - 3Y + 6I + 6)}{6 \cdot 2I(2I + 1)}\right\}^{\frac{1}{2}} \times (Q + 1\lambda\mu - 1\, \|\mathbf{X}\|\, Q\lambda\mu). \quad (7.7c)$$

$$(Q - 1\lambda - 1\mu\, Y + \tfrac{2}{3} IK\, |\mathbf{Z}_0|\, Q\lambda\mu YIK) = [(4\lambda + 2\mu - 3Y - 6I)(4\lambda + 2\mu - 3Y + 6I + 6)]^{\frac{1}{2}} (Q - 1\lambda - 1\mu\, \|\mathbf{Z}\|\, Q\lambda\mu), \quad (7.8a)$$

$$(Q - 1\lambda - 1\mu\ Y - \tfrac{1}{3}I + \tfrac{1}{2}K \pm \tfrac{1}{2}|\mathbf{Z}_\pm|\ Q\lambda\mu YIK)$$
$$= -\left\{\frac{(I \pm K + 1)(2\mu - 2\lambda - 3Y + 6I + 6)(3Y - 6I + 2\lambda + 4\mu)(4\lambda + 2\mu - 3Y - 6I)}{6(2I + 1)(2I + 2)}\right\}^{\frac{1}{2}}$$
$$\times (Q - 1\lambda - 1\mu\ \|\mathbf{Z}\|\ Q\lambda\mu), \quad (7.8b)$$

$$(Q - 1\lambda - 1\mu\ Y - \tfrac{1}{3}I - \tfrac{1}{2}K \pm \tfrac{1}{2}|\mathbf{Z}_\pm|\ Q\lambda\mu YIK)$$
$$= \mp\left\{\frac{(I \mp K)(3Y + 6I + 2\lambda - 2\mu)(3Y + 6I + 2\lambda + 4\mu + 6)(4\lambda + 2\mu - 3Y + 6I + 6)}{6 \cdot 2I(2I + 1)}\right\}^{\frac{1}{2}}$$
$$\times (Q - 1\lambda - 1\mu\ \|\mathbf{Z}\|\ Q\lambda\mu). \quad (7.8c)$$

$$(Q - 1\lambda + 1\mu - 1Y + \tfrac{2}{3}IK\,|\mathbf{Z}_0|\ Q\lambda\mu YIK)$$
$$= [(3Y + 6I + 2\lambda - 2\mu + 6)(2\mu - 2\lambda - 3Y + 6I)]^{\frac{1}{2}}(Q - 1\lambda + 1\mu - 1\ \|\mathbf{Z}\|\ Q\lambda\mu), \quad (7.9a)$$

$$(Q - 1\lambda + 1\mu - 1Y - \tfrac{1}{3}I + \tfrac{1}{2}K \pm \tfrac{1}{2}|\mathbf{Z}_\pm|\ Q\lambda\mu YIK)$$
$$= -\left\{\frac{(I \pm K + 1)(4\lambda + 2\mu - 3Y + 6I + 12)(3Y - 6I + 2\lambda + 4\mu)(3Y + 6I + 2\lambda - 2\mu + 6)}{6(2I + 1)(2I + 2)}\right\}^{\frac{1}{2}}$$
$$\times (Q - 1\lambda + 1\mu - 1\ \|\mathbf{Z}\|\ Q\lambda\mu), \quad (7.9b)$$

$$(Q - 1\lambda + 1\mu - 1Y - \tfrac{1}{3}I - \tfrac{1}{2}K \pm \tfrac{1}{2}|\mathbf{Z}_\pm|\ Q\lambda\mu YIK)$$
$$= \pm\left\{\frac{(I \mp K)(4\lambda + 2\mu - 3Y - 6I + 6)(3Y + 6I + 2\lambda + 4\mu + 6)(2\mu - 2\lambda - 3Y + 6I)}{6 \cdot 2I(2I + 1)}\right\}^{\frac{1}{2}}$$
$$\times (Q - 1\lambda + 1\mu - 1\ \|\mathbf{Z}\|\ Q\lambda\mu). \quad (7.9c)$$

$$(Q - 1\lambda\mu + 1Y + \tfrac{2}{3}IK\,|\mathbf{Z}_0|\ Q\lambda\mu YIK)$$
$$= [(3Y + 6I + 2\lambda + 4\mu + 12)(3Y - 6I + 2\lambda + 4\mu + 6)]^{\frac{1}{2}}(Q - 1\lambda\mu + 1\ \|\mathbf{Z}\|\ Q\lambda\mu), \quad (7.10a)$$

$$(Q - 1\lambda\mu + 1Y - \tfrac{1}{3}I + \tfrac{1}{2}K \pm \tfrac{1}{2}|\mathbf{Z}_\pm|\ Q\lambda\mu YIK)$$
$$= \left\{\frac{(I \pm K + 1)(4\lambda + 2\mu - 3Y + 6I + 12)(2\mu - 2\lambda - 3Y + 6I + 6)(3Y + 6I + 2\lambda + 4\mu + 12)}{6(2I + 1)(2I + 2)}\right\}^{\frac{1}{2}}$$
$$\times (Q - 1\lambda\mu + 1\ \|\mathbf{Z}\|\ Q\lambda\mu), \quad (7.10b)$$

$$(Q - 1\lambda\mu + 1Y - \tfrac{1}{3}I - \tfrac{1}{2}K \pm \tfrac{1}{2}|\mathbf{Z}_\pm|\ Q\lambda\mu YIK)$$
$$= \pm\left\{\frac{(I \mp K)(4\lambda + 2\mu - 3Y - 6I + 6)(3Y + 6I + 2\lambda - 2\mu)(3Y - 6I + 2\lambda + 4\mu + 6)}{6 \cdot 2I(2I + 1)}\right\}^{\frac{1}{2}}$$
$$\times (Q - 1\lambda\mu + 1\ \|\mathbf{Z}\|\ Q\lambda\mu). \quad (7.10c)$$

8. CALCULATION OF THE REDUCED MATRIX ELEMENTS

The Wigner–Eckart factorization of the matrix elements of $\mathbf{X}$ and $\mathbf{Z}$ effected in the last section means that we have satisfied all the commutation relations of the $\mathbf{X}$ and the $\mathbf{Z}$ with the SU_3 operators. We are then left with six reduced matrix elements to calculate, and in order to do this we must utilize the commutators of the $\mathbf{X}$ and the $\mathbf{Z}$ among each other. This computation is simplified if we make the most of homogeneous equations between the matrix elements, and, among the latter, of those equations which contain the smallest number of summands (i.e., the "stretched" cases).

Let us take the commutator $[\mathbf{X}_0, \mathbf{Z}_0] = 0$ and consider its matrix element

$$(Q\lambda\mu\, Y\,|\mathbf{X}_0\mathbf{Z}_0 - \mathbf{Z}_0\mathbf{X}_0|\ Q\lambda - 1\mu - 1\, Y) = 0.$$

Here we can drop the labels I and K, as $\mathbf{X}_0$ and $\mathbf{Z}_0$ are both I scalars. The above quickly leads to the equation

$$(Q\lambda\mu\ \|\mathbf{X}\|\ Q - 1\lambda - 1\mu)$$
$$\times (Q - 1\lambda - 1\mu\ \|\mathbf{Z}\|\ Q\lambda - 1\mu - 1)$$
$$= (Q\lambda\mu\ \|\mathbf{Z}\|\ Q + 1\lambda\mu - 1)$$
$$\times (Q + 1\lambda\mu - 1\ \|\mathbf{X}\|\ Q\lambda - 1\mu - 1). \quad (8.1)$$

Next, we attempt to obtain an equation with one term on each side which contains

$$(Q - 1\lambda - 1\mu\ \|\mathbf{X}\|\ Q\lambda - 1\mu - 1)$$
$$\times (Q\lambda - 1\mu - 1\ \|\mathbf{X}\|\ Q - 1\lambda - 1\mu)$$

on one side, and

$$(Q\lambda\mu\ \|\mathbf{X}\|\ Q + 1\lambda\mu - 1)(Q + 1\lambda\mu - 1\ \|\mathbf{X}\|\ Q\lambda\mu)$$

on the other. To do this we choose the commutator

$$[\mathbf{X}_0, \mathbf{X}_+] = 0$$

and take its matrix elements between the two states with the largest Q difference which occurs in (8.1), i.e., between $Q+1\,\lambda\,\mu-1$ and $Q-1\,\lambda-1\,\mu$, with appropriate values of Y, I, and K. Each term in the commutator now splits into a sum of two terms. Introducing the expressions for the reduced matrix elements, cancelling common factors and combining terms, we find that Y and I disappear from the coefficients, and we are left with the equation

$$\begin{aligned}(\lambda+\mu+2)&(Q+1\lambda\mu-1\,\|\mathbf{X}\|\,Q\lambda\mu)\\ &\times(Q\lambda\mu\,\|\mathbf{X}\|\,Q-1\lambda-1\mu)\\ =(\lambda+\mu)&(Q+1\lambda\mu-1\,\|\mathbf{X}\|\,Q\lambda-1\mu-1)\\ &\times(Q\lambda-1\mu-1\,\|\mathbf{X}\|\,Q-1\lambda-1\mu).\end{aligned}\tag{8.2}$$

Now we multiply the left-hand side of (8.1) with the right-hand side of (8.2), and conversely, and cancel common factors. Then we are left with the equation

$$\begin{aligned}(\lambda+\mu)&(Q-1\lambda-1\mu\,\|\mathbf{Z}\|\,Q\lambda-1\mu-1)\\ &\times(Q\lambda-1\mu-1\,\|\mathbf{X}\|\,Q-1\lambda-1\mu)\\ =(\lambda+\mu+2)&(Q\lambda\mu\,\|\mathbf{Z}\|\,Q+1\lambda\mu-1)\\ &\times(Q+1\lambda\mu-1\,\|\mathbf{X}\|\,Q\lambda\mu).\end{aligned}\tag{8.3}$$

Writing

$$\begin{aligned}(Q\lambda\mu\,\|\mathbf{Z}\|\,Q+1\lambda\mu-1)&(Q+1\lambda\mu-1\,\|\mathbf{X}\|\,Q\lambda\mu)\\ &=A(Q\lambda\mu)\end{aligned}\tag{8.4}$$

and multiplying both sides of (8.3) by $(\lambda+\mu+1)$, we have

$$\begin{aligned}(\lambda+\mu)(\lambda+\mu+1)&A(Q-1\lambda-1\mu)\\ &=(\lambda+\mu+1)(\lambda+\mu+2)A(Q\lambda\mu),\end{aligned}$$

which gives at once

$$A(Q\lambda\mu)=-\frac{a'(Q-\lambda,\mu)}{(\lambda+\mu+1)(\lambda+\mu+2)},\tag{8.5}$$

where a' is a function of the two variables $Q-\lambda$ and μ, and the minus sign is introduced for convenience. Equation (8.5) reduces the quantity A, which is a function of three variables, to a function of two variables. We can reduce it further to a function of a single variable. In order to do this systematically, it is useful to prepare a list of all independent equations corresponding to "stretched" cases which can be derived from the commutator $[\mathbf{X}_0, \mathbf{Z}_0] = 0$. This list (omitting all irrelevant indices and quantum numbers) can be represented as

$$\left.\begin{aligned}(\lambda\mu\,|\mathbf{XZ}|\,\lambda-1\mu-1)&\to(\lambda\mu\,|\mathbf{X}|\,\lambda-1\mu)\\ &\times(\lambda-1\mu\,|\mathbf{Z}|\,\lambda-1\mu-1)\\ (\lambda\mu\,|\mathbf{ZX}|\,\lambda-1\mu-1)&\to(\lambda\mu\,|\mathbf{Z}|\,\lambda\mu-1)\\ &\times(\lambda\mu-1\,|\mathbf{X}|\,\lambda-1\mu-1)\end{aligned}\right\},\tag{8.6a}$$

$$\left.\begin{aligned}(\lambda\mu\,|\mathbf{XZ}|\,\lambda+1\mu-2)&\\ \to(\lambda\mu\,|\mathbf{X}|\,\lambda+1\mu-1)&\\ \times(\lambda+1\mu-1\,|\mathbf{Z}|\,\lambda+1\mu-2)&\\ (\lambda\mu\,|\mathbf{ZX}|\,\lambda+1\mu-2)&\\ \to(\lambda\mu\,|\mathbf{Z}|\,\lambda\mu-1)&\\ \times(\lambda\mu-1\,|\mathbf{X}|\,\lambda+1\mu-2)&\end{aligned}\right\},\tag{8.6b}$$

$$\left.\begin{aligned}(\lambda\mu\,|\mathbf{XZ}|\,\lambda-2\mu+1)&\to(\lambda\mu\,|\mathbf{X}|\,\lambda-1\mu)\\ &\times(\lambda-1\mu\,|\mathbf{Z}|\,\lambda-2\mu+1)\\ (\lambda\mu\,|\mathbf{ZX}|\,\lambda-2\mu+1)&\\ \to(\lambda\mu\,|\mathbf{Z}|\,\lambda-1\mu+1)&\\ \times(\lambda-1\mu+1\,|\mathbf{X}|\,\lambda-2\mu+1)&\end{aligned}\right\},\tag{8.7a}$$

$$\left.\begin{aligned}(\lambda\mu\,|\mathbf{XZ}|\,\lambda-1\mu+2)&\to(\lambda\mu\,|\mathbf{X}|\,\lambda\mu+1)\\ &\times(\lambda\mu+1\,|\mathbf{Z}|\,\lambda-1\mu+2)\\ (\lambda\mu\,|\mathbf{ZX}|\,\lambda-1\mu+2)&\\ \to(\lambda\mu\,|\mathbf{Z}|\,\lambda-1\mu+1)&\\ \times(\lambda-1\mu+1\,|\mathbf{X}|\,\lambda-1\mu+2)&\end{aligned}\right\},\tag{8.7b}$$

$$\left.\begin{aligned}(\lambda\mu\,|\mathbf{XZ}|\,\lambda+2\mu-1)&\\ \to(\lambda\mu\,|\mathbf{X}|\,\lambda+1\mu-1)&\\ \times(\lambda+1\mu-1\,|\mathbf{Z}|\,\lambda+2\mu-1)&\\ (\lambda\mu\,|\mathbf{ZX}|\,\lambda+2\mu-1)&\to(\lambda\mu\,|\mathbf{Z}|\,\lambda+1\mu)\\ &\times(\lambda+1\mu\,|\mathbf{X}|\,\lambda+2\mu-1)\end{aligned}\right\},\tag{8.8a}$$

$$\left.\begin{aligned}(\lambda\mu\,|\mathbf{XZ}|\,\lambda+1\mu+1)&\to(\lambda\mu\,|\mathbf{X}|\,\lambda\mu+1)\\ &\times(\lambda\mu+1\,|\mathbf{Z}|\,\lambda+1\mu+1)\\ (\lambda\mu\,|\mathbf{ZX}|\,\lambda+1\mu+1)&\to(\lambda\mu\,|\mathbf{Z}|\,\lambda+1\mu)\\ &\times(\lambda+1\mu\,|\mathbf{X}|\,\lambda+1\mu+1)\end{aligned}\right\}.\tag{8.8b}$$

Each of these six sets gives one equation. Furthermore, corresponding sets (a) and (b) are related by the fact that the *same* matrix elements of $\mathbf{Z}$ occur in both. If we apply the process which we used to derive Eq. (8.5) to two corresponding sets, we obtain two different equations for the same function; one of these equations allows us to reduce the number of variables from three to two, and *subsequently* the other from two to one. For example, from the set (8.6b), proceeding exactly as we did in order to derive Eq. (8.5) we obtain the relation

$$(\mu-1)A(Q-1\lambda+1\mu-1)=(\mu+1)A(Q\lambda\mu).\tag{8.9}$$

Introducing (8.5) into the above and multiplying both sides by μ, we obtain

$$\begin{aligned}\mu(\mu-1)&a'(Q-\lambda-2,\mu-1)\\ &=(\mu+1)\mu a'(Q-\lambda,\mu)\\ &=a(Q-\lambda-2\mu),\end{aligned}$$

i.e.,

$$a'(Q-\lambda,\mu)=a(Q-\lambda-2\mu)/\mu(\mu+1)$$

or

$$A(Q\lambda\mu)=-\frac{a(Q-\lambda-2\mu)}{\mu(\mu+1)(\lambda+\mu+1)(\lambda+\mu+2)}. \tag{8.10}$$

In the same way, we obtain from the sets (8.7a, b):

$$(Q\lambda\mu\,\|\mathbf{Z}\|\,Q+1\lambda-1\mu+1)\times(Q+1\lambda-1\mu+1\,\|\mathbf{X}\|\,Q\lambda\mu)=B(Q\lambda\mu)=-\frac{b(Q-\lambda+\mu)}{\lambda(\lambda+1)(\mu+1)(\mu+2)}, \tag{8.11}$$

and from the sets (8.8a, b):

$$(Q\lambda\mu\,\|\mathbf{Z}\|\,Q+1\lambda+1\mu)(Q+1\lambda+1\mu\,\|\mathbf{X}\|\,Q\lambda\mu)=C(Q\lambda\mu)=-\frac{c(Q+2\lambda+\mu)}{(\lambda+1)(\lambda+2)(\lambda+\mu+2)(\lambda+\mu+3)}. \tag{8.12}$$

Thus we have almost reduced the problem to that of determining three functions, each of one variable only.

We have now exhausted all useful information we can obtain from off-diagonal matrix elements, and for the next step we must consider nonvanishing diagonal elements. It suffices to take the simplest of the commutators (4.6), namely (4.6a):

$$[\mathbf{X}_0,\mathbf{Z}_0]=\tfrac{1}{8}\mathbf{Y}-\tfrac{1}{6}\mathbf{Q}.$$

The diagonal matrix element of this commutator provides the equation

$$\begin{aligned}0=\tfrac{1}{6}Q-\tfrac{1}{48}(x+y)&+(p+q-x)(p+q-y+6)\\&\times C(Q-1\lambda-1\mu)+(x+p+6)(-y-p)\\&\times B(Q-1\lambda+1\mu-1)+(y+q+6)\\&\times(x+q+12)A(Q-1\lambda\,\mu+1)\\&-(p+q-x+6)(p+q-y+12)C(Q\lambda\mu)\\&-(x+p)(-y-p+6)B(Q\lambda\mu)-(x+q+6)\\&\times(y+q)A(Q\lambda\mu),\end{aligned} \tag{8.13}$$

where we have written, for convenience,

$$\left.\begin{aligned}2\lambda-2\mu=p,\quad 2\lambda+4\mu=q,\\3Y+6I=x,\quad 3Y-6I=y.\end{aligned}\right\} \tag{8.14}$$

Now A, B, and C are independent of x and y. Therefore the terms containing x and y must vanish for all values of x and y, so that the coefficients of x, y, and xy must vanish identically. Thus (8.13) actually gives the following equations: (a) coeff. of $xy=0$, (b) coeff. of $x=0$, (c) coeff. of $y=0$, and (d) remainder $=0$. Written out in full, these equations are, respectively,

$$\begin{aligned}&C(Q-1\lambda-1\mu)-B(Q-1\lambda+1\mu-1)\\&\quad+A(Q-1\lambda\,\mu+1)-C(Q\lambda\mu)+B(Q\lambda\mu)\\&\quad-A(Q\lambda\mu)=0;\end{aligned} \tag{8.15}$$

$$\begin{aligned}&-(p+q+6)C(Q-1\lambda-1\mu)\\&\quad-pB(Q-1\lambda+1\mu-1)+(q+6)\\&\quad\times A(Q-1\lambda\,\mu+1)+(p+q+12)C(Q\lambda\mu)\\&\quad+(p-6)B(Q\lambda\mu)-qA(Q\lambda\mu)=\tfrac{1}{48};\end{aligned} \tag{8.16}$$

$$\begin{aligned}&-(p+q)C(Q-1\lambda-1\mu)-(p+6)\\&\quad\times B(Q-1\lambda+1\mu-1)+(q+12)\\&\quad\times A(Q-1\lambda\,\mu+1)+(p+q+6)C(Q\lambda\mu)\\&\quad+pB(Q\lambda\mu)-(q+6)A(Q\lambda\mu)=\tfrac{1}{48};\end{aligned} \tag{8.17}$$

$$\begin{aligned}&(p+q)(p+q+6)C(Q-1\lambda-1\mu)-p(p+6)\\&\quad\times B(Q-1\lambda+1\mu-1)+(q+6)(q+12)\\&\quad\times A(Q-1\lambda\,\mu+1)-(p+q+6)(p+q+12)\\&\quad\times C(Q\lambda\mu)+p(p-6)B(Q\lambda\mu)-q(q+6)\\&\quad\times A(Q\lambda\mu)=-\tfrac{1}{6}Q.\end{aligned} \tag{8.18}$$

Note that the equation obtained by subtracting (8.17) from (8.16) is identical with (8.15).

Let us write Eq. (8.15) in terms of λ, μ, and the functions a, b, and c, and eliminate all factors from the denominators. It becomes

$$\begin{aligned}&\lambda(\lambda+1)(\lambda+2)(\mu+2)(\lambda+\mu+3)a(Q-\lambda-2\mu)\\&\quad-\lambda(\lambda+1)(\lambda+2)\mu(\lambda+\mu+1)\\&\quad\times a(Q-\lambda-2\mu-3)-(\lambda+2)\mu(\lambda+\mu+1)\\&\quad\times(\lambda+\mu+2)(\lambda+\mu+3)b(Q-\lambda+\mu)\\&\quad+\lambda(\mu+2)(\lambda+\mu+1)(\lambda+\mu+2)(\lambda+\mu+3)\\&\quad\times b(Q-\lambda+\mu-3)+\lambda\mu(\mu+1)(\mu+2)\\&\quad\times(\lambda+\mu+1)c(Q+2\lambda+\mu)-(\lambda+2)\mu(\mu+1)\\&\quad\times(\mu+2)(\lambda+\mu+3)c(Q+2\lambda+\mu-3)=0.\end{aligned} \tag{8.19}$$

Set $\mu=0$ in the above. This gives

$$2\lambda(\lambda+1)(\lambda+2)(\lambda+3)a(Q-\lambda)+2\lambda(\lambda+1)(\lambda+2)(\lambda+3)b(Q-\lambda-3)=0$$

or

$$a(x)=-b(x-3),$$

i.e.,

$$b(x)=-a(x+3). \tag{8.20}$$

Next, set $\lambda=0$ in (8.19). This gives

$$b(Q+\mu)+c(Q+\mu-3)=0$$

or

$$\begin{aligned}c(x)&=-b(x+3)\\&=a(x+6).\end{aligned} \tag{8.21}$$

Thus the task of determining the three functions a, b, and c is reduced to that of determining a *single* function.

Next, set $\lambda = \mu = 1$ in (8.19) and introduce the relations (8.20) and (8.21). This gives

$$a(Q+9) - 5a(Q+6) + 10a(Q+3) - 10a(Q) + 5a(Q-3) - a(Q-6) = 0. \quad (8.22)$$

Define

$$\Delta a(Q) = a(Q) - a(Q-3). \quad (8.23)$$

Then (8.22) is

$$\Delta^5 a(Q) = 0,$$

a simple linear difference equation of the fifth order, the solution of which involves five arbitrary constants.

Eliminating $C(Q\lambda\mu)$ from Eqs. (8.16) and (8.17), setting $\lambda = \mu = 1$ as before and using the relations (8.10)–(8.12), (8.20), and (8.21), we arrive at the equation

$$\Delta^4 a(Q+6) = -1/12, \quad (8.24)$$

which gives us one constant of integration. However, we can proceed further. Setting $\lambda = \mu = 1$ in (8.18) we obtain the equation

$$-\tfrac{5}{18}Q = 3a(Q+9) - 5a(Q+6) + 5a(Q-3) - 3a(Q-6). \quad (8.25)$$

Now, multiplying (8.22) by -3, adding to (8.25) and cancelling common factors, we obtain the equation

$$\Delta^3 a(Q+6) = -\tfrac{1}{36}Q, \quad (8.26)$$

the solution of which is

$$a(x) = -(1/2^5 3^6)(x-\alpha)(x-\beta)(x-\gamma) \times (x+\alpha+\beta+\gamma-6), \quad (8.27)$$

where α, β, γ are arbitrary constants. These constants depend on the irreducible representation (fgh) of SU_4 under consideration, and cannot therefore be determined from the commutation relations.

9. SEPARATION OF THE REDUCED MATRIX ELEMENTS AND BOUNDARY CONDITIONS

We have obtained, apart from three undetermined constants, the expressions for the three products of reduced matrix elements

$$(Q\lambda\mu \,\|\mathbf{Z}\|\, Q+1\lambda'\mu')(Q+1\lambda'\mu' \,\|\mathbf{X}\|\, Q\lambda\mu),$$

where $(\lambda'\mu')$ can be $(\lambda+1\mu)$, $(\lambda-1\mu+1)$, or $(\lambda\mu-1)$. The reduced matrix elements of $\mathbf{Z}$ and $\mathbf{X}$ cannot be separated by using the commutation relations. In order to separate them it is necessary to impose additional conditions, for example Hermiticity conditions. We therefore require the conditions (cf. end of Sec. 4)

$$(\mathbf{E}_\lambda)^\dagger = \mathbf{E}_{-\lambda},$$
$$(\mathbf{E}_\mu)^\dagger = \mathbf{E}_{-\mu},$$
$$(\mathbf{E}_\nu)^\dagger = \mathbf{E}_{-\nu}$$

to hold in every representation. With $\mathbf{Q}^\dagger = \mathbf{Q}$, these conditions are consistent with each other and with the Hermiticity conditions in SU_3, and translate themselves into the relation

$$(Q+1\lambda'\mu' \,\|\mathbf{X}\|\, Q\lambda\mu) = -[(Q\lambda\mu \,\|\mathbf{Z}\|\, Q+1\lambda'\mu')]^* \quad (9.1)$$

for all allowed $(\lambda'\mu')$. These statements can be verified from Eqs. (7.5) to (7.10) and the identifications given at the end of Sec. 4.

Using (9.1) and the relations (8.20) and (8.21) we can rewrite Eqs. (8.10)–(8.12) as

$$|(Q+1\lambda+1\mu \,\|\mathbf{X}\|\, Q\lambda\mu)|^2 = \frac{a(Q+2\lambda+\mu+6)}{(\lambda+1)(\lambda+2)(\lambda+\mu+2)(\lambda+\mu+3)}, \quad (9.2a)$$

$$|(Q+1\lambda-1\mu+1 \,\|\mathbf{X}\|\, Q\lambda\mu)|^2 = -\frac{a(Q-\lambda+\mu+3)}{\lambda(\lambda+1)(\mu+1)(\mu+2)}, \quad (9.2b)$$

$$|(Q+1\lambda\mu-1 \,\|\mathbf{X}\|\, Q\lambda\mu)|^2 = \frac{a(Q-\lambda-2\mu)}{\mu(\mu+1)(\lambda+\mu+1)(\lambda+\mu+2)}, \quad (9.2c)$$

where $a(x)$ is given by (8.27).

One way of determining the constants α, β, and γ is as follows. The highest weight of the representation (fgh) of SU_4 is

$$W = \left[\frac{1}{2\sqrt{3}}(f+g+h), \frac{1}{6}(f+g-h), \frac{1}{6\sqrt{2}}(f-2g-h)\right].$$

The weights with the largest value of Q in the representation are equivalent to the highest weight. One such weight is

$$M = W - (g+h)\mu = \left[\frac{1}{2\sqrt{3}}f, \frac{1}{6}(f+2g), \frac{1}{6\sqrt{2}}(f+2g+3h)\right],$$

and therefore the highest value of Q in the representation (fgh) is

$$Q_{\max} = \tfrac{1}{4}(f+2g+3h). \quad (9.3)$$

Furthermore, it may be ascertained that weights with Q value equalling $Q_{\max}$ carry the SU_3 quantum numbers

$$\lambda = f, \quad \mu = g. \quad (9.4)$$

We may now observe that the left-hand sides of (9.2a, b, c) must vanish for $Q = Q_{\max}$. This provides three conditions which can be satisfied by choosing appropriate values for the integration constants. We determine α, β, γ by requiring the corresponding factors to vanish for $Q = \tfrac{1}{4}(f+2g+3h)$, $\lambda = f$,

$\mu = g$ in the right-hand sides of (9.2a), (9.2b), and (9.2c), respectively. This gives

$$\alpha = \tfrac{1}{4}(9f + 6g + 3h + 24),$$
$$\beta = \tfrac{1}{4}(-3f + 6g + 3h + 12),$$
$$\gamma = \tfrac{1}{4}(-3f - 6g + 3h),$$

so that the final expression for $a(x)$ becomes

$$a(x) = (1/2^5 3^6)[x - \tfrac{1}{4}(9f + 6g + 3h) - 6] \times [x - \tfrac{1}{4}(-3f + 6g + 3h) - 3] \times [x - \tfrac{1}{4}(-3f - 6g + 3h)] \times [x + \tfrac{1}{4}(3f + 6g + 9h) + 3]. \quad (9.5)$$

There remain three other boundary conditions, namely, those deriving from the requirement that the left-hand side of (9.2a), (9.2b), and (9.2c) should vanish for $Q + 1 = Q_{\min}$. As there are no further disposable constants, these conditions must be satisfied automatically, and thus provide a check on the calculations. The minimum value of Q is $Q_{\min} = -\tfrac{1}{4}(3f + 2g + h)$, and the corresponding values of λ and μ are g and h, respectively.

Apart from arbitrary phase factors, Eqs. (9.1), (9.2a, b, c), and (9.5) solve our problem.

ACKNOWLEDGMENT

The author would like to thank Professor H. Umezawa for many useful discussions.

This work was supported in part by the U.S. Air Force Office of Scientific Research under contract No. AF 61(052)-826 through the European Office of Aerospace Research (OAR), U.S. Air Force.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

General Method for the Inverse Scattering Problem at Fixed Energy

PIERRE C. SABATIER
Laboratoire de Physique Théorique et Hautes Energies, Orsay, France

(Received 13 August 1966)

A method for studying the inverse scattering problem at fixed energy is given; it enables one to get at a fairly large class of potentials $\mathcal{C}$. It is shown that the problem has in $\mathcal{C}$ an infinity of solutions, depending on an infinity of parameters. If the study is restricted to the potentials of $\mathcal{C}$ which can be continued as even analytic functions in a circle centered at the origin, the problem has only one solution. A linear approximation is given and is shown to yield inverse formulas of the Born approximation. For even potentials of $\mathcal{C}$, the relevance of this approximation is studied as follows: Comparison is made between a given static potential and the potential obtained from the scattering amplitude through the inverse formula. The standard deviation between these two potentials is shown to go to zero as the energy goes to ∞. Miscellaneous properties of potentials, wavefunctions, and Jost functions are given in the framework of this method.

INTRODUCTION

IN two previous papers,[1,2] we studied successively the asymptotic behavior and the analytic properties of the potentials derived from a given set of phase shifts in the framework of Newton's method.[3] We showed in paper I that if the phase shifts fulfill a very weak condition,[4] namely $|\delta_l| < Cl^{-3-\epsilon}$, as $l \to \infty$, there is one potential, and only one, which goes to zero faster than $r^{-2+\epsilon}$ as $r \to \infty$. All the potentials equivalent to this one go to zero like $r^{-\frac{3}{2}}$ as $r \to \infty$.

In paper II, we showed that Newton's potentials can be continued in the r complex plane, and we found several limitations of their analytic properties. We also studied their Jost functions, and we came to the general conclusion that the potentials obtained with Newton's method belong to a narrow class of potentials.

In this paper, we give a very general method for constructing potentials, and then restrict the study to a narrower class for which it is easy to solve the inverse problem at fixed energy with a method analogous to Newton's. This class is, however, so general that the inverse problem has an infinity of solutions.

[1] P. C. Sabatier, J. Math Phys. 7, 1515 (1966). Hereafter referred to as I.
[2] P. C. Sabatier, J. Math. Phys. 7, 2079 (1966), Hereafter referred to as II.
[3] R. G. Newton. J. Math. Phys. 3, 75 (1962).
[4] Throughout this paper, we mean by ϵ a positive number, which can be made arbitrarily small, but not equal to zero. We use C as a general constant. Both are not meant to have the same value every time they are used. We mean by bound an upper bound for the absolute value.

$\mu = g$ in the right-hand sides of (9.2a), (9.2b), and (9.2c), respectively. This gives

$$\alpha = \tfrac{1}{4}(9f + 6g + 3h + 24),$$
$$\beta = \tfrac{1}{4}(-3f + 6g + 3h + 12),$$
$$\gamma = \tfrac{1}{4}(-3f - 6g + 3h),$$

so that the final expression for $a(x)$ becomes

$$a(x) = (1/2^5 3^6)[x - \tfrac{1}{4}(9f + 6g + 3h) - 6] \times [x - \tfrac{1}{4}(-3f + 6g + 3h) - 3] \times [x - \tfrac{1}{4}(-3f - 6g + 3h)] \times [x + \tfrac{1}{4}(3f + 6g + 9h) + 3]. \quad (9.5)$$

There remain three other boundary conditions, namely, those deriving from the requirement that the left-hand side of (9.2a), (9.2b), and (9.2c) should vanish for $Q + 1 = Q_{\min}$. As there are no further disposable constants, these conditions must be satisfied automatically, and thus provide a check on the calculations. The minimum value of Q is $Q_{\min} = -\tfrac{1}{4}(3f + 2g + h)$, and the corresponding values of λ and μ are g and h, respectively.

Apart from arbitrary phase factors, Eqs. (9.1), (9.2a, b, c), and (9.5) solve our problem.

ACKNOWLEDGMENT

The author would like to thank Professor H. Umezawa for many useful discussions.

This work was supported in part by the U.S. Air Force Office of Scientific Research under contract No. AF 61(052)-826 through the European Office of Aerospace Research (OAR), U.S. Air Force.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

General Method for the Inverse Scattering Problem at Fixed Energy

PIERRE C. SABATIER
Laboratoire de Physique Théorique et Hautes Energies, Orsay, France

(Received 13 August 1966)

A method for studying the inverse scattering problem at fixed energy is given; it enables one to get at a fairly large class of potentials $\mathcal{C}$. It is shown that the problem has in $\mathcal{C}$ an infinity of solutions, depending on an infinity of parameters. If the study is restricted to the potentials of $\mathcal{C}$ which can be continued as even analytic functions in a circle centered at the origin, the problem has only one solution. A linear approximation is given and is shown to yield inverse formulas of the Born approximation. For even potentials of $\mathcal{C}$, the relevance of this approximation is studied as follows: Comparison is made between a given static potential and the potential obtained from the scattering amplitude through the inverse formula. The standard deviation between these two potentials is shown to go to zero as the energy goes to ∞. Miscellaneous properties of potentials, wavefunctions, and Jost functions are given in the framework of this method.

INTRODUCTION

IN two previous papers,[1,2] we studied successively the asymptotic behavior and the analytic properties of the potentials derived from a given set of phase shifts in the framework of Newton's method.[3] We showed in paper I that if the phase shifts fulfill a very weak condition,[4] namely $|\delta_l| < Cl^{-3-\epsilon}$, as $l \to \infty$, there is one potential, and only one, which goes to zero faster than $r^{-2+\epsilon}$ as $r \to \infty$. All the potentials equivalent to this one go to zero like $r^{-\frac{3}{2}}$ as $r \to \infty$.

In paper II, we showed that Newton's potentials can be continued in the r complex plane, and we found several limitations of their analytic properties. We also studied their Jost functions, and we came to the general conclusion that the potentials obtained with Newton's method belong to a narrow class of potentials.

In this paper, we give a very general method for constructing potentials, and then restrict the study to a narrower class for which it is easy to solve the inverse problem at fixed energy with a method analogous to Newton's. This class is, however, so general that the inverse problem has an infinity of solutions.

[1] P. C. Sabatier, J. Math Phys. 7, 1515 (1966). Hereafter referred to as I.
[2] P. C. Sabatier, J. Math. Phys. 7, 2079 (1966), Hereafter referred to as II.
[3] R. G. Newton. J. Math. Phys. 3, 75 (1962).
[4] Throughout this paper, we mean by ϵ a positive number, which can be made arbitrarily small, but not equal to zero. We use C as a general constant. Both are not meant to have the same value every time they are used. We mean by bound an upper bound for the absolute value.

The set of equivalent potentials depend on an infinity of independent parameters.

In Sec. 1, we show that a generalization of Newton's approach to the inverse scattering problem enables one to build a very large class of potentials—more precisely, of all the potentials analytic in some circle of the r complex plane, centered at the origin and cut along the negative axis.

If $K(z, z)$ is associated to the potential by

$$K(z, z) = -\tfrac{1}{2}z\int_0^z \rho V(\rho)\, d\rho$$

a potential $V(z)$ is characterized by the coefficients c_μ of the expansion

$$K(z, z) = \sum_\mu c_\mu \phi_\mu(z) u_\mu(z),$$

where the functions $\phi_\mu(z)$ and $u_\mu(z)$ are the regular solutions of the Schrödinger equation with and without the potential. The index μ may be either a discrete or a continuous index.

In Sec. 2, we restrict the class of potentials by assuming $|c_\mu|$ to be bounded by $C\mu^{\frac{1}{2}-\epsilon}$. This condition enables us to get at the asymptotic behavior of the wavefunctions which yield an infinite set of linear equations relating the phase shifts to the coefficients c_μ. When the μ are allowed to include all the integral numbers, it is possible to choose arbitrarily all the other c_μ and to solve the problem. There is therefore an infinity of solutions. In Newton's method all the c_μ are equal to zero for nonintegral μ. As we have seen in I, the solutions of the problem depend only on one parameter. For a potential analytic in the neighborhood of the origin, the c_l and the $c_{l+\frac{1}{2}}$ are allowed to be different from zero. It follows that a study symmetric to Newton's one will be that of potentials with $c_{l+\frac{1}{2}} \neq 0$, $c_\mu = 0$ otherwise. These are even analytic potentials in the neighborhood of the origin. We show that the problem has one solution and only one in this case—It should be borne in mind that the bounds of the c_μ hold, so that we can only obtain a subclass of the even potentials.

Miscellaneous problems are studied in Sec. 3. We first give a short study of the linear approximation, obtained by keeping only the first-order terms at any step of our method. This approximation is identical with the usual Born approximation. We then review some properties of the functions involved in scattering theory and which can be derived in the framework of our method. We successively study the analytic properties of the wavefunctions for a very general class of potentials, the asymptotic properties of the potentials and the analytic properties of the Jost functions for the class defined in Sec. 2, the construction of transparent potentials and of solvable examples in the class defined in Sec. 2. We end this section with a study of the Born approximation for even potentials belonging to the class of Sec. 2. These potentials are given by the inverse sine transform of the scattering amplitude continued by 0 outside the physical angular range. Some comments on the approximation of a static potential by means of a potential of this kind are then made.

1. A VERY LARGE CLASS OF POTENTIALS

Let us start with a function $f(r, r')$ defined by the partial differential equation:

$$[D_0(r) - D_0(r')]f(r, r') = 0,$$
$$f(0, r') = f(r, 0) = 0, \tag{1.1}$$

where

$$D_0(r) \equiv r^2(\partial^2/\partial r^2 + 1). \tag{1.2}$$

It is easy to see that this equation is satisfied by any function of the following form:

$$f(r, r') = \sum_\mu c_\mu u_\mu(r) u_\mu(r'), \tag{1.3}$$

where the functions $u_\mu(r)$ are given by

$$u_\mu(r) = (\tfrac{1}{2}\pi r)^{\frac{1}{2}} J_{\mu+\frac{1}{2}}(r). \tag{1.4}$$

In (1.4), μ is either a discrete or a continuous index (in which case the sum is to be replaced by an integral), which can take any real value larger than -1. Let us now assume that the following bound holds for any μ:

Assumption I: There exist

$$k \text{ and } \alpha > 0 \Rightarrow c_\mu/\mu!\,\mu! < C\mu^k\alpha^{-2\mu}. \tag{1.5}$$

We know that the functions $u_\mu(z)$ are equal to the product of $z^{\mu+1}$ by an entire function. Using this property, it is easy to continue $f(z, z')$ for any finite value of z' (or z) as an analytic function of z (or z') inside a circle ($|z\,z'| < \alpha^2$), except for cuts issued from the origin. We put these cuts along the negative real axes of the z and z' complex planes. To find a determination for $f(z, z')$, it is sufficient to choose a determination for $\log z$ and $\log z'$, since the failure of analyticity comes from terms $e^{\mu \log z}$ or $e^{\mu \log z'}$.

We hereafter call $\mathfrak{D}_0$ a domain in the complex plane comprising a circle with the center at the origin and a radius $(\alpha - \epsilon)$, except for a cut along the negative real axis. Obviously, if z and z' lie in the domain $\mathfrak{D}_0$ of the complex plane, the function $f(z, z')$ is analytic both in z and z'.

Let us now define from $f(r, r')$ a function $K(r, r')$ by the integral equation

$$K(r, r') = f(r, r') - \int_0^r dr_1 r_1^{-2} K(r, r_1) f(r_1, r'). \quad (1.6)$$

We intend to get at the analytic structure of $K(r, r')$. In order to avoid nonessential difficulties, we put a first restriction on the functions $f(r, r')$ by assuming $\mu \geq 0$, although it is easy to extend the analysis until $\mu = -\frac{1}{2} + \epsilon$. The method we used in our previous paper (II) can then be applied to (1.6) with only trivial modifications if z and z' are assumed to belong to $\mathfrak{D}_0$.

Let us retrace our argument: Introduce the two functions

$$\phi_r(r') \equiv \phi(r, r') = (rr')^{-1} f(r, r'), \quad (1.7)$$

$$K_r(r') = (rr')^{-1} K(r, r'). \quad (1.8)$$

We apply the Fredholm method to the following equation, equivalent to (1.8)[5]:

$$K_r(r') = \phi_r(r') - \int_0^r dr_1 K_r(r_1) \phi(r_1, r'). \quad (1.9)$$

We obtain, therefore, the value of $K_r(r')$:

$$K_r(r') = \phi_r(r') + [\mathfrak{D}(r)]^{-1} \mathfrak{D}(r, r'), \quad (1.10)$$

where

$$\mathfrak{D}(r) = 1 + \sum_{m=1}^{\infty} (m!)^{-1} \times \int_0^r \cdots \int_0^r \Phi \begin{pmatrix} r_1, r_2 \cdots r_m \\ r_1, r_2 \cdots r_m \end{pmatrix} dr_1 \cdots dr_m, \quad (1.11)$$

$$\mathfrak{D}(r, r') = -\sum_{m=0}^{\infty} (m!)^{-1} \times \int_0^r \cdots \int_0^r \phi(r, \rho) \Phi \begin{pmatrix} \rho, r_1 \cdots r_m \\ r', r_1 \cdots r_m \end{pmatrix} d\rho\, dr_1 \cdots dr_m. \quad (1.12)$$

In these formulas the notation Φ stands for the Fredholm determinant associated with the kernel $\phi(r, r')$.

The formulas (1.10), (1.11), and (1.12) enable us to continue analytically $K_r(r')$ when r and r' get complex values (z and z'). $\phi(z, z')$ is an analytic function of z and z' when z and z' belong to $\mathfrak{D}_0$; it is bounded, say, by $C(|z|, |z'|)$ for any $|z|$ smaller than $|\alpha|^2/|z'|$. From this it follows that any term of (1.11) can be given an analytic continuation for any z in $\mathfrak{D}_0$, and any term of (1.12) can be continued as an analytic function of z and z' for z and z' in $\mathfrak{D}_0$. Besides, if we choose rays issued from the origin as contours of integration, and apply well-known inequalities to the integrals and the determinants involved in $\mathfrak{D}(z)$ and $\mathfrak{D}(z, z')$, we find that the terms of order m in $\mathfrak{D}(z)$ and $\mathfrak{D}(z, z')$ are, respectively, bounded by

$$|z|^m [C(|z|, |z|)]^m m^{\frac{1}{2}m} \quad (z \in \mathfrak{D}_0),$$

$$|z|^{m+1} C(|z|, |z|)^{m+1} C(|z|, |z'|)(m+1)^{\frac{1}{2}(m+1)} \quad (z, z' \in \mathfrak{D}_0).$$

The series involved in (1.11) and (1.12) are therefore uniformly convergent for any couple of numbers z and z' in $\mathfrak{D}_0$. The function $\mathfrak{D}(z)$ is therefore analytic for z in $\mathfrak{D}_0$. $\mathfrak{D}(z, z')$ is an analytic function[6] of z and z' when z and z' are taken in $\mathfrak{D}_0$.

It follows from these results that $zV(z)$, as defined below, is analytic in the domain $\mathfrak{D}_0$ of the z complex plane, except at isolated points z_i. Since the Neumann[7,8] series for Eq. (1.10) converges at least inside a nonvanishing circle centered at the origin of the z complex plane ($|z|\, C(|z|, |z|) < 1$), none of these poles can be at the origin.

It is easy to see that our argument can easily be extended to any connected domain $\mathfrak{D}$ where the function $\phi(z, z')$ is an analytic function of z and z', provided that this domain includes the origin. In its original form, the argument presupposes also that it is possible to reach any point of the domain on a ray issued from the origin, but this condition may be avoided by contour integrations. It is now easy to show that the following partial differential equation holds in the domain of analyticity $\mathfrak{D}^*$ of $K(z, z')$:

$$D(z)K(z, z') - D_0(z')K(z, z') = 0, \quad (1.13)$$

$$K(z, 0) = K(0, z') = 0, \quad (1.14)$$

where $D(z)$ is equal to $D_0(z) - z^2 V(z)$ and $V(z)$ to

$$V(z) = -2z^{-1}(d/dz)z^{-1}K(z, z). \quad (1.15)$$

In order to prove our point, we have only to apply $D(z)$ and $D_0(z')$ to $K(z, z')$ as given by (1.10) and to show through straightforward but tedious differentiations and integrations by part that the right-hand side of (1.13) satisfies the homogeneous version of (1.10). Inside the domain of analyticity $\mathfrak{D}^*$ of $K(z, z')$, the solution of (1.10) in complex form is unique, so that the only solution of the homogeneous equation is zero. Obviously this argument fails at the poles z_i. The remarks we made in (II) can be applied again in the present case. Q.E.D.

[5] For all details, please refer to (II).

[6] It is easy to obtain a little more information: z lying in $\mathfrak{D}_0$, the function $\mathfrak{D}(z, z')$ is an analytic function of z and z' provided that $|zz'| < \alpha^2$ and z' be not on the cut.

[7] F. G. Tricomi, *Integral Equations* (Interscience Publishers, Inc., New York, 1957).

We now define the functions

$$\phi_\nu(z) = u_\nu(z) - \int_0^z dz_1 z_1^{-2} K(z, z_1) u_\nu(z_1) \quad (\mathrm{Re}\,\nu > -1), \tag{1.16}$$

and show easily from (1.13), with the same argument as in (II), that the functions $\phi_\nu(z)$ are the regular solutions of the Schrödinger equation

$$D(z)\phi_\nu(z) = \nu(\nu+1)\phi_\nu(z), \quad \phi_\nu(0) = 0. \tag{1.17}$$

Inserting (1.3) into (1.6), we observe that $K(r, r')$ may be expanded as

$$K(r, r') = \sum_\mu c_\mu \mathcal{K}_\mu(r) u_\mu(r'),$$

where

$$\mathcal{K}_\mu(r) = u_\mu(r) - \int_0^r dr_1 r_1^{-2} K(r, r_1) u_\mu(r_1).$$

Comparison with (1.16) leads to the bilinear expansion

$$K(z, z') = \sum_\mu c_\mu \phi_\mu(z) u_\mu(z'), \tag{1.18}$$

from which it follows that

$$\phi_\nu(z) = u_\nu(z) - \sum_\mu L_\nu^\mu(z) c_\mu \phi_\mu(z) \quad (\mathrm{Re}\,\nu > -1), \tag{1.19}$$

where

$$L_\nu^\mu(z) = \int_0^z dz_1 z_1^{-2} u_\nu(z_1) u_\mu(z_1) \quad [\mathrm{Re}\,(\nu+\mu) > -1] \tag{1.20}$$

or

$$L_\nu^\mu(z) = \frac{u_\nu(z)u'_\mu(z) - u_\mu(z)u'_\nu(z)}{\mu(\mu+1) - \nu(\nu+1)} \quad [\mathrm{Re}\,(\nu+\mu) > -1]. \tag{1.21}$$

At this point, we have completely solved the Schrödinger equation for a potential defined by (1.15), since we know, from (1.6) and (1.16), the regular solutions of all the partial-wave equations. It is interesting to describe the class of potentials which can be studied in this way. We restrict the function $f(r, r')$ to the class defined by (1.3) and (1.5), which is narrower than the class of all the solutions of (1.1). Even with this restriction, the class of potentials is very large, since it involves all the potentials analytic in some circle centered at the origin and cut along the negative real axis. If the radius α of $\mathfrak{D}_0$ is finite, we have obviously to make a further assumption.

Assumption II: The potential and the functions $\phi_l(r)$ can be continued on the whole positive real axis.

For this we have to impose in particular Assumption III.

Assumption III: We discard the sets of c_μ which yield a pole on the positive real axis in $\mathfrak{D}_0$, since the order of the pole of $rV(r)$ would be at least equal to 2, and the scattering problem might lack sense.

Actually, provided such a pole is not at $+\infty$, it is still possible to define mathematically the problem by avoiding the pole on all the contours of integration. As we already remarked in paper (II), this may pave the way for a study of singular potentials.

If the c_μ are bounded in such a way that

Assumption IV:

$$|c_\mu/\mu!\,\mu!|^{\mu^{-1}} \to 0 \quad \text{as} \quad \mu \to \infty,$$

the very weak Assumption III is sufficient to ensure that the Schrödinger equation is defined and solved on the whole positive real axis.

We still have to impose a further condition.

Assumption V:

$$r^{1+\epsilon}V(r) \quad \text{goes to zero as} \quad r \to \infty.$$

Unfortunately, this assumption cannot be related to conditions on the c_μ. In the following sections, we give asymptotic bounds to the c_μ which are *sufficient* to ensure Assumption V, but they are by no way necessary conditions. It is a pity to notice that very simple potentials as $(1+r^2)^{-1}$, for instance, escape even Assumption IV, since they have single order poles in the complex plane.

If two functions $V(z)$ holomorphic in a domain $\mathfrak{D}_0$ are different, they give rise to different potentials $V(r)$. It is of interest to ask whether different sets of c_μ lead to different functions $V(z)$. In order to answer this question, we first assume that all the μ's are rational and that it is possible to find a common denominator m for the two different sets. Let us then define, with suitable determinations, the quantities

$$\zeta = z^{1/m}, \quad \zeta' = z'^{1/m}. \tag{1.22}$$

The functions $u_\mu(z)$ [or, $u_\mu(z')$] are entire functions of ζ (or, ζ'). $f(z, z')$, with Assumption I, is an analytic function of ζ and ζ' in the circle $(0, |\alpha|^{1/m})$. Equation (1.9) gives place, with obvious notations, to

$$\bar{K}_\zeta(z') = \bar{\phi}_\zeta(\zeta') - m\int_0^\zeta \bar{K}_\zeta(\zeta_1)\bar{\phi}(\zeta_1, \zeta')\zeta_1^{m-1}\, d\zeta_1. \tag{1.23}$$

It is clear that $\bar{K}_\zeta(\zeta_1)$ is equal to

$$\bar{K}_\zeta(\zeta_1) = \overline{\mathfrak{D}}(\zeta, \zeta_1)/\overline{\mathfrak{D}}(\zeta), \tag{1.24}$$

where $\overline{\mathfrak{D}}(\zeta)$ and $\overline{\mathfrak{D}}(\zeta, \zeta_1)$ are analytic functions of ζ and ζ_1 when ζ and ζ_1 belong to the circle $(0, |\alpha|^{1/m})$.

We incidentally remark that a glance at Eq. (1.23)

suggests the method which may be used to extend the values of μ towards the negative real values—more exactly until $\mu = -m^{-1}E[\frac{1}{2}(m-1)]$. When such values of μ are involved in the expansion (1.3), we make the following change of functions:

$$\phi(\zeta_1, \zeta') = \zeta_1^{-E[\frac{1}{2}(m-1)]}\zeta^{-E[\frac{1}{2}(m-1)]}\phi^*(\zeta_1, \zeta'),$$
$$K_\zeta(\zeta') = \zeta^{-E[\frac{1}{2}(m-1)]}\zeta^{1-E[\frac{1}{2}(m-1)]}K_\zeta^*(\zeta'). \quad (1.25)$$

The function $K_\zeta^*(\zeta')$ is solution of an equation whose kernel is regular, which allows us to apply to (1.25) the same arguments as above. However, as we wrote above, we discard this case and assume $\mu \geq 0$.

It follows from Eq. (1.24) that $\zeta V(\zeta^m)$ is a function of ζ analytic in a circle centered at the origin. Conversely, let us assume that $zV(z)$ is an holomorphic function of ζ inside a circle ($|\zeta| < |\alpha|^{1/m}$), and let us define $K(z, z)$ as

$$K(z, z) = -\tfrac{1}{2}z\int_0^z \rho V(\rho)\, d\rho. \quad (1.26)$$

We show in Appendix A that $K(z, z)$ can be expanded in a series of products $u_{p/m}\phi_{p/m}$ and $u_{(p+\frac{1}{2})/m}\phi_{(p+\frac{1}{2})/m}$

$$K(z, z) = \sum_{p=0}^{\infty} c_{p/m}u_{p/m}\phi_{p/m} + \sum_{p=0}^{\infty} c_{(p+\frac{1}{2})/m}u_{(p+\frac{1}{2})/m}\phi_{(p+\frac{1}{2})/m} \quad (1.27)$$

and this expansion is unique.

The coefficients c_μ are therefore in this case unambiguously associated with the potential. When the μ's are not rational, either their ratios are rational, in which case we can give an argument analogous to the above, or they are not, in which case the nature of the branch point of $k(z, z)$ is different.

We conclude that, in general, two different sets of c_μ lead to different potentials.

2. INVERSE SCATTERING PROBLEM

We now intend to solve the inverse scattering problem at fixed energy. In order to relate the phase shifts to the coefficients c_μ, we lay a very restrictive condition on the asymptotic behavior of these coefficients.

Assumption VI:

$$|c_\mu| < C\mu^{\frac{1}{3}-\epsilon}. \quad (2.1)$$

With this assumption,[8] the series involved in (1.19) are uniformly convergent as $z \to \infty$, so that it is possible to replace the functions by their asymptotic behavior:

$$\phi_\nu(r) \sim (2i)^{-1}[f_1(\nu)e^{ir} - f_2(\nu)e^{-ir}],$$
$$L_\mu^\nu(r) \sim L_\mu^\nu(\infty) = \frac{\sin((\nu-\mu)(\frac{1}{2}\pi))}{(\nu+\frac{1}{2})^2 - (\mu+\frac{1}{2})^2}, \quad (2.2)$$

[8] The argument of the proof is the same as in paper (II).

where $f_1(\nu)$ and $f_2(\nu)$ are the Jost functions. For real ν, they reduce to

$$f_1(\nu) = A_\nu e^{i\delta_\nu}e^{\frac{1}{2}i\pi\nu}, \quad f_2(\nu) = [f_1(\nu)]^* \quad (\mathrm{Im}\,\nu = 0). \quad (2.3)$$

We obtain the expansion of the Jost functions

$$f_1(\nu) = e^{-\frac{1}{2}i\pi\nu} - \sum_\mu \frac{\sin((\nu-\mu)(\frac{1}{2}\pi))}{\nu(\nu+1) - \mu(\mu+1)}\, c_\mu A_\mu e^{i\delta_\mu}e^{-\frac{1}{2}i\mu\pi},$$
$$f_2(\nu) = e^{\frac{1}{2}i\pi\nu} - \sum_\mu \frac{\sin((\nu-\mu)(\frac{1}{2}\pi))}{\nu(\nu+1) - \mu(\mu+1)}\, c_\mu A_\mu e^{-i\delta_\mu}e^{\frac{1}{2}i\mu\pi}, \quad (2.4)$$

and, for a real μ_0 (larger than -1)

$$A_{\mu_0}e^{i\delta_{\mu_0}}e^{-\frac{1}{2}i\pi\mu_0} = e^{-\frac{1}{2}i\pi\mu_0} - \sum_\mu L_{\mu_0}^\mu(\infty)c_\mu A_\mu e^{i\delta_\mu}e^{-\frac{1}{2}i\mu\pi}. \quad (2.5)$$

2.1 Definition and Resolution of the Inverse Problem

In order to define the inverse problem, we should specify the class of potentials in which we look for the solution. We suppose that the potentials satisfy Assumptions III and VI, and furthermore, that the μ's are rational and *include all the integral numbers*. In these conditions, we prove that it is possible to choose arbitrarily, within some very large conditions of validity, all the c_μ for which μ is not an integral number, and solve the problem. We use hereafter the notation $L_{\mu'}^\mu$ for $L_{\mu'}^\mu(\infty)$.

Proof: Let $\hat{\phi}$ be the vector with components $A_\mu e^{i[\delta_\mu - \frac{1}{2}\mu\pi]}$. We can consider $\hat{\phi}$ as a sum of $\hat{\phi}_1$, with components $A_l e^{i(\delta_l - \frac{1}{2}l\pi)}$, and $\hat{\phi}_0$, with the other components. In the same way, the vector $\hat{u}$, with components $e^{-\frac{1}{2}i\pi\mu_0}$, is considered as the sum of $\hat{u}_1$ and $\hat{u}_0$. Equation (2.5), where we successively give to μ_0 all the values we allow μ to take, can be written in matrix notations as

$$\hat{\phi} = \hat{u} - L\mathbf{C}\hat{\phi}, \quad (2.6)$$

where $\mathbf{C}$ is the diagonal matrix with components c_μ, L the matrix with components $L_{\mu'}^\mu$. We now divide L into the four parts which couple the vectors $\hat{\phi}_0$ and $\hat{\phi}_1$ to the vectors $\hat{u}_0$ and $\hat{u}_1$:

$$\hat{\phi}_1 = \hat{u}_1 - L_{(1)}^{(1)}\mathbf{C}_1\hat{\phi}_1 - L_{(1)}^{(0)}\mathbf{C}_0\hat{\phi}_0, \quad (2.7)$$
$$\hat{\phi}_0 = \hat{u}_0 - L_{(0)}^{(1)}\mathbf{C}_1\hat{\phi}_1 - L_{(0)}^{(0)}\mathbf{C}_0\hat{\phi}_0. \quad (2.8)$$

We now choose *arbitrarily* the matrix $\mathbf{C}_0$, assuming only for convenience that

$$|c_\mu| < C\mu^{-\frac{5}{3}-\epsilon}. \quad (2.9)$$

It is, in general, possible to inverse the matrix

$(1 + L_{(0)}^{(0)}\mathbf{C}_0)$. If the elements of $\mathbf{C}_0$ are small enough, the Neumann series converges. If not, and except for special $\mathbf{C}_0$, the Fredholm method[9] gives the inverse matrix. Use of it in (2.8) yields

$$\hat{\phi}_0 = (1 + L_{(0)}^{(0)}\mathbf{C}_0)^{-1}[\hat{u}_0 - L_{(0)}^{(1)}\mathbf{C}_1\hat{\phi}_1]. \quad (2.10)$$

Inserting (2.10) in (2.7), we obtain

$$\hat{\phi}_1 = \hat{\xi}_1 - [L_{(1)}^{(1)} + S_{(1)}^{(1)}]\mathbf{C}_1\hat{\phi}_1, \quad (2.11)$$

where we put

$$\hat{\xi}_1 = \hat{u}_1 - L_{(1)}^{(0)}\mathbf{C}_0(1 + L_{(0)}^{(0)}\mathbf{C}_0)^{-1}\hat{u}_0, \quad (2.12)$$

$$S_{(1)}^{(1)} = -L_{(1)}^{(0)}\mathbf{C}_0(1 + L_{(0)}^{(0)}\mathbf{C}_0)^{-1}L_{(0)}^{(1)}. \quad (2.13)$$

Let $\xi_l e^{i(\Delta_l - \frac{1}{2}l\pi)}$ be the components of $\hat{\xi}_1$, let $S_l^{l'}$ be the components of $S_{(1)}^{(1)}$. The components of $L_{(1)}^{(1)}$ are $L_l^{l'}$. Equation (2.11) is equivalent to the set of equations

$$A_l = \xi_l e^{i(\Delta_l - \delta_l)} - \sum_{l'=0}^{\infty}(L_l^{l'} + S_l^{l'})c_{l'}A_{l'}e^{i(\delta_{l'} - \delta_l)}i^{l-l'}. \quad (2.14)$$

We now consider each of the matrices $L_{(1)}^{(1)}$ and $S_{(1)}^{(1)}$ as a sum of two matrices which couple respectively components with index of same parity and components with index of different parity:

$$\begin{aligned} L_l^{l'} &= i^{l'-l-1}M_l^{l'} + [\pi/2(2l+1)]\delta_l^{l'}, \\ {}^{+}S_l^{l'} &= -i^{l-l'}S_l^{l'} \quad \text{for } l - l' \text{ even, 0 otherwise.} \\ {}^{-}S_l^{l'} &= i^{l-l'+1}S_l^{l'} \quad \text{for } l - l' \text{ odd, 0 otherwise.} \end{aligned} \quad (2.15)$$

Insertion of these quantities in (2.14) yields

$$\begin{aligned} A_l\left[1 + \frac{\pi}{2(2l+1)}c_l - {}^{+}S_l^l c_l\right] &= \xi_l e^{i(\Delta_l - \delta_l)} + i\sum_{l'}(M_l^{l'} + {}^{-}S_l^{l'})c_{l'}A_{l'}e^{i(\delta_{l'} - \delta_l)} \\ &\quad + \sum_{l'} {}^{+}S_l^{l'}c_{l'}A_{l'}e^{i(\delta_{l'} - \delta_l)}. \end{aligned} \quad (2.16)$$

Taking the imaginary part of both sides, we obtain the formula

$$\begin{aligned} \xi_l \sin(\delta_l - \Delta_l) &= \sum_{l'}(M_l^{l'} + {}^{-}S_l^{l'})c_{l'}A_{l'}\cos(\delta_{l'} - \delta_l) \\ &\quad + \sum_{l'} {}^{+}S_l^{l'}c_{l'}A_{l'}\sin(\delta_{l'} - \delta_l). \end{aligned} \quad (2.17)$$

Introducing now the vectorial notations $\hat{\xi}$ and $\hat{a}$, for the vectors with respective components $\xi_l \sin(\delta_l - \Delta_l)$ and $a_l (= c_l A_l \cos\delta_l)$, and introducing the diagonal matrix $\tan\delta$ with components $\tan\delta_l$, we obtain the equation

$$\hat{\xi} = (P + Q)\hat{a} = P(1 + P^{-1}Q)\hat{a}, \quad (2.18)$$

where

$$P = M(1 + M^{-1}\,{}^{-}S + M^{-1}\,{}^{+}S\tan\delta), \quad (2.19)$$

$$Q = \tan\delta[(M + {}^{-}S)\tan\delta - {}^{+}S]. \quad (2.20)$$

[9] See, for example, R. G. Newton, J. Math. Phys. 2, 188 (1961).

According to (2.13), the matrices ${}^{+}S$ or ${}^{-}S$ depend on the arbitrary coefficient in C_0, so that, *in general*, a matrix like

$$(1 + M^{-1}\,{}^{-}S + M^{-1}\,{}^{+}S\tan\delta)$$

can be inverted, say, by the Fredholm method. As a result we can define, and build, a matrix P^{-1}. The matrix Q depends on the phase shifts, so that, *in general*, $(1 + P^{-1}Q)$ can be inverted. We assume that all these inversions are feasible. The coefficients a_l are then given by the formula

$$\hat{a} = (1 + P^{-1}Q)^{-1}P^{-1}\hat{\xi}. \quad (2.21)$$

The knowledge of the a_l enables us to get A_l from (2.16), since, if we take the real part of both members of this equation, we obtain

$$\begin{aligned} A_l\left[1 + \frac{\pi}{2(2l+1)}c_l - {}^{+}S_l^l c_l\right] &= \xi_l\cos(\Delta_l - \delta_l) \\ &\quad - \cos\delta_l\sum_{l'}(M_l^{l'} + {}^{-}S_l^{l'})a_{l'}(\tan\delta_{l'} - \tan\delta_l) \\ &\quad + \cos\delta_l\sum_{l'} {}^{+}S_l^{l'}a_{l'}(1 + \tan\delta_l\tan\delta_{l'}). \end{aligned} \quad (2.22)$$

Obviously, the usual indetermination occurs when the matrix M^{-1} is introduced. This can be studied as in (I).

It follows from our study that, even with the very strong assumptions we made, the indetermination in the inverse scattering problem at fixed energy has at least the power of the denumerable. Since continuous sequences of c_μ would give rise to other classes of potentials, we may be sure that the indetermination of the problem has the power of the continuum.

Suppose now we impose to the potential to be analytic in a nonvanishing circle centered at the origin, and to satisfy Assumption VI. The problem has still a denumerable infinity of solutions since, according to formula (1.27), we have to find coefficients c_l and $c_{l+\frac{1}{2}}$:

$$K(z, z) = \sum_{l=0}^{\infty} c_l u_l\phi_l + \sum_{l=0}^{\infty} c_{l+\frac{1}{2}}u_{l+\frac{1}{2}}\phi_{l+\frac{1}{2}}. \quad (2.23)$$

It follows from the results of Appendix A that the coefficients c_l are equal to zero for an even potential. The potentials obtained by Newton's method are characterized by

$$c_{l+\frac{1}{2}} = 0 \quad (2.24)$$

and they cannot be even potentials. Since the inverse scattering problem in Newton's method is nearly completely determined, it is interesting to study the problem for even potentials.

2.2. Inverse Problem for Even Potentials

We assume that the coefficients c_μ are equal to zero except for $\mu = l + \frac{1}{2}$, and are restricted by Assumptions III and VI.

For the sake of convenience, we introduce a coefficient $c_{-\frac{1}{2}}$, which ultimately will be made equal to zero. Equation (2.5) takes the form

$$A_l e^{i(\delta_l - \frac{1}{2}\pi l)} = e^{-\frac{1}{2}i\pi l} - \sum_p L_l^{p-\frac{1}{2}} c_{p-\frac{1}{2}} A_{p-\frac{1}{2}} e^{i[\delta_{p-\frac{1}{2}} - (p-\frac{1}{2})(\frac{1}{2}\pi)]} \quad (2.25)$$

or

$$A_l = e^{-i\delta_l} - \sum_p L_l^{p-\frac{1}{2}} c_{p-\frac{1}{2}} A_{p-\frac{1}{2}} e^{i[\delta_{p-\frac{1}{2}} - \delta_l]} e^{i(l+\frac{1}{2}-p)(\frac{1}{2}\pi)}. \quad (2.26)$$

If we take the imaginary part of both members, we obtain the equation

$$\begin{aligned} -\tan\delta_l = &\sum_p L_l^{p-\frac{1}{2}} \sin(l + \tfrac{1}{2} - p)(\tfrac{1}{2}\pi) \\ &\times a_{p-\frac{1}{2}}(1 + \tan\delta_{p-\frac{1}{2}} \tan\delta_l) \\ &+ \sum_p L_l^{p-\frac{1}{2}} \cos(l + \tfrac{1}{2} - p)(\tfrac{1}{2}\pi) \\ &\times a_{p-\frac{1}{2}}(\tan\delta_{p-\frac{1}{2}} - \tan\delta_l). \end{aligned} \quad (2.27)$$

In this equation, a fundamental matrix occurs:

$$m_l^p = L_l^{p-\frac{1}{2}} \sin(l + \tfrac{1}{2} - p)(\tfrac{1}{2}\pi) = \tfrac{1}{2}[(l + \tfrac{1}{2})^2 - p^2]^{-1}, \quad (2.28)$$

whereas

$$L_l^{p-\frac{1}{2}} \cos(l + \tfrac{1}{2} - p)(\tfrac{1}{2}\pi) = (-1)^{l-p} m_l^p. \quad (2.29)$$

Insertion of this matrix in (2.27) yields the equation

$$\begin{aligned} -\tan\delta_l = \sum_p m_l^p a_{p-\frac{1}{2}}[1 &+ \tan\delta_{p-\frac{1}{2}} \tan\delta_l \\ &+ (-1)^{l+p}(\tan\delta_{p-\frac{1}{2}} - \tan\delta_l)], \end{aligned} \quad (2.30)$$

whereas it is easy to obtain from (2.5), in the same way as in (I), a relation between the phase shifts of half-integral order:

$$\tan\delta_{p_0+\frac{1}{2}} = \sum_p {}^{(0)}M_{p_0+1}^{p+1} a_{p+\frac{1}{2}}(1 + \tan\delta_{p_0+\frac{1}{2}} \tan\delta_{p+\frac{1}{2}}), \quad (2.31)$$

where

$${}^{(0)}M_{p_0}^p = \begin{cases} (p^2 - p_0^2)^{-1}, & \text{for } p - p_0 \text{ odd}, \\ 0, & \text{otherwise}. \end{cases} \quad (2.32)$$

Let us now remark that, from a result of (I), we know the inverse ${}^{(0)}\gamma$ of the matrix ${}^{(0)}M$, so that we have numbers ${}^{(0)}\gamma_l^{l'}$ such that

$$\sum_p {}^{(0)}M_{2l+1}^{2p} {}^{(0)}\gamma_{2p}^{2q+1} = \delta_{2l+1}^{2q+1}. \quad (2.33)$$

Since m_l^p is equal to twice ${}^{(0)}M_{2l+1}^{2p}$, the matrix m^{-1} with elements $\eta_p^q = \frac{1}{2}\gamma_{2p}^{2q+1}$ is a right inverse of m. It is also a left inverse of m, because ${}^{(0)}M$ is a double-sided inverse of ${}^{(0)}\gamma$. In the same way, we find that if ${}^{(0)}\hat{v}$ is annihilated by M, the vector $\hat{s}$ annihilated by m has the following components:

$$s_p = {}^{(0)}v_{2p}. \quad (2.34)$$

The results of (I) yield

$$\eta_p^q = \begin{cases} -4\pi^{-2}, & \text{for } p = 0, \\ 8\pi^{-2} \dfrac{(2q+1)^2}{(2p)^2 - (2q+1)^2} & \text{for } p \neq 0. \end{cases} \quad (2.35)$$

$$s_p = \begin{cases} \frac{1}{2} & \text{for } p = 0, \\ 1 & \text{for } p \neq 0. \end{cases} \quad (2.36)$$

Let us now denote by $\hat{a}$ the vector with components $a_{p-\frac{1}{2}}$, $\tan\delta$, and $\tan\delta'$ the diagonal matrix with components $\tan\delta_p$ and $\tan\delta_{p-\frac{1}{2}}$, I the diagonal matrix whose elements are

$$I_l^l = \begin{cases} 1 & \text{for } l \text{ even}, \\ -1 & \text{for } l \text{ odd}. \end{cases}$$

Taking into account the value of $a_{-\frac{1}{2}}$, we can write (2.30) and (2.31) in matrix notations:

$$\tan\delta\hat{e} = m\hat{a} + \tan\delta m \tan\delta'\hat{a} + ImI \tan\delta'\hat{a} - \tan\delta ImI\hat{a}, \quad (2.37)$$

$$\tan\delta'\hat{e} = {}^{(0)}M\hat{a} + \tan\delta'{}^{(0)}M \tan\delta'\hat{a} \quad (2.38)$$

and it is possible to rewrite (2.37) as

$$m^{-1}\tan\delta\hat{e} = \alpha\hat{s} + \hat{a} + m^{-1}\tan\delta m \tan\delta'\hat{a} + m^{-1}ImI \tan\delta'\hat{a} - m^{-1}\tan\delta ImI\hat{a}, \quad (2.39)$$

where the value of α has to be chosen so that $a_{-\frac{1}{2}}$ is equal to zero. In order to study this system of equations, it is possible to use either an iteration or a perturbation process. Let us assume for instance that

$$\tan\delta = \lambda\delta, \quad (2.40)$$

where δ is a fixed diagonal matrix, whose elements go to zero as l tends to infinity. We may write for α, $\hat{a}$, and $\tan\delta'$, formal expansions in powers of λ, and make the coefficients of λ^n in (2.38) and (2.39) equal to zero. The convergence of these expansions can be shown if λ is small enough. Proofs are given in Appendix C. It follows from this convergence that a potential exists and is unique at least if all the phase shifts are small enough. We have not tried to show the existence and uniqueness of the potential in the general case.

3. MISCELLANEOUS RESULTS

3.1. Linear Approximation

Our method of studying the inverse scattering problem at fixed energy can be given a simpler form if we limit to the first order the expansions of all the

functions encountered at the various steps of the theory. This gives for the fundamental equations

$$f(r, r') = \sum_{\mu} c_{\mu} u_{\mu}(r) u_{\mu}(r'), \tag{3.1}$$

$$K(r, r') \sim f(r, r'), \tag{3.2}$$

$$V(r) \sim -2r^{-1}(d/dr)r^{-1}f(r, r). \tag{3.3}$$

On the other hand, it is easy to obtain from (2.5) the following exact equations:

$$\sin \delta_{\mu_0} = -\sum_{\mu} L^{\mu}_{\mu_0} c_{\mu} A_{\mu} \sin [\delta_{\mu} - \delta_{\mu_0} + (\tfrac{1}{2}\pi)(\mu_0 - \mu)], \tag{3.4}$$

$$A_{\mu_0} = 1 - \sum_{\mu} L^{\mu}_{\mu_0} c_{\mu} A_{\mu} \cos [\delta_{\mu} - \delta_{\mu_0} + (\tfrac{1}{2}\pi)(\mu_0 - \mu)]. \tag{3.5}$$

The linear form of these equations can be written as

$$\tan \delta_{\mu_0} \sim -\sum_{\mu} L^{\mu}_{\mu_0} a_{\mu} \sin ((\tfrac{1}{2}\pi)(\mu_0 - \mu)), \tag{3.6}$$

$$A_{\mu_0} \sim 1 - \sum_{\mu} L^{\mu}_{\mu_0} a_{\mu} \cos ((\tfrac{1}{2}\pi)(\mu_0 - \mu)), \tag{3.7}$$

where we put

$$a_{\mu} = c_{\mu} A_{\mu} \cos \delta_{\mu} \sim c_{\mu} A_{\mu}. \tag{3.8}$$

These equations make sense only if all the phase shifts and the a_{μ} are small. If we allow the μ's to contain all the integers, we can choose a_{μ} for nonintegral μ, but this choice is not completely arbitrary. To see this point, let us assume that δ_l is equal to $\lambda \delta_l^{(0)}$, and a_{μ}, for nonintegral μ, is equal to $\lambda a_{\mu}^{(0)}$, where $\delta_l^{(0)}$ and $a_{\mu}^{(0)}$ have a reasonable asymptotic behavior. The range of values of λ for which the linear approximation makes sense depends obviously on the choice of $a_{\mu}^{(0)}$ as well as on the $\delta_l^{(0)}$. In other words, the arbitrary a_{μ} should be small enough. Once this condition is fulfilled, it is easy to see that the a_l computed from the δ_l are of the order of λ. In this approximation, Eq. (3.7) reduces, therefore, to

$$A_{\mu_0} = 1. \tag{3.9}$$

It is possible to use (3.6) in order to relate the physical phase shifts to a set of c_{μ} for which μ is different from l, for instance $l + \frac{1}{2}$. The matrix which connects the δ_l to this set should be inversed in order to solve the problem. It may happen that vectors are annihilated by this matrix. We should then use a convenient coefficient in order to be consistent with the linear approximation. For instance, the coefficient which we used for the "special potential" in Newton's method (see I), is consistent with the linear approximation.

If the linear approximation is consistent for the derivation of the c_{μ}, it is consistent in Eqs. (3.2) and (3.3), because the c_{μ} are of the order of λ, and there is always a range of λ for which the Neumann series converges and yields (3.2), from (1.6).

Born Approximation

It is easy to prove that the linear approximation is the same as the well-known Born approximation. We know that in the Born approximation the phase shift is given by

$$\delta_l \sim \tan \delta_l \sim \tfrac{1}{2}\pi \int_0^{\infty} V(\rho) J_{l+\frac{1}{2}}(\rho) J_{l+\frac{1}{2}}(\rho) \rho \, d\rho. \tag{3.10}$$

If $V(\rho)$ is given the form (3.3) and $f(\rho, \rho)$ the form (3.1), we get

$$\tan \delta_l \sim \tfrac{1}{2}\pi^2 \int_0^{\infty} J_{l+\frac{1}{2}}(\rho) J_{l+\frac{1}{2}}(\rho) \frac{\partial}{\partial \rho} \times \sum_{\mu} c_{\mu} J_{\mu+\frac{1}{2}}(\rho) J_{\mu+\frac{1}{2}}(\rho) \, d\rho. \tag{3.11}$$

These integrals are calculated in Appendix B and shown to be equal to

$$\tan \delta_l \sim -\sum_{\mu} \frac{\sin^2 (\frac{1}{2}\pi)(l - \mu)}{(l + \frac{1}{2})^2 - (\mu + \frac{1}{2})^2} c_{\mu}, \tag{3.12}$$

which is identical with (3.6). Q.E.D.

3.2. Some Properties of the Functions Constructed in Secs. 1 and 2

We have already given the analytic properties of the potentials which can be constructed by the general method of Sec. 1. With Assumption I only, we can give the analytic properties of the wavefunctions. From Appendix A, it is easy to see that they are analytic functions of z in the same domain as $zV(z)$, i.e., in $\mathfrak{D}_0$. Now, from formulas (1.19) and (1.21), it is easy to define an analytic continuation of $\phi_{\nu}(z)$ for any value of ν, as a meromorphic function of ν. In the special case of an even potential, the symmetry properties of the cylindrical Bessel functions enable us to see that $\phi_{\nu}(z)$ is an entire function of ν and satisfies the symmetry property:

$$(-1)^l \left[1 + \frac{\pi}{2l} c_{l-\frac{1}{2}}\right] \phi_{l-\frac{1}{2}}(z) = \phi_{-l-\frac{1}{2}}(z). \tag{3.13}$$

Asymptotic Behavior of Potentials

It would be of interest to derive the asymptotic properties of the potentials from the asymptotic properties of the phase shifts. This is easy in the simple case where the c_{μ} include the c_l, which are derived from the phase shifts, whereas all other c_{μ} are bounded by $C\mu^{-\frac{5}{3}}$.

The argument we have given in (I) (Sec. 3.3) can be applied, and it ensures that the functions f and g which we may construct in the present case behave asymptotically as those which were introduced in (I).

It follows that the leading term in the asymptotic behavior of the potential is still given by

$$V(r) \sim -2\pi^{-\frac{1}{2}}(\alpha - \beta)r^{-\frac{3}{2}} \cos(2r - \tfrac{1}{4}\pi) + O(r^{\epsilon-2}). \tag{3.14}$$

It is interesting to notice that we can deduce from the asymptotic behavior of $K(r, r)$ a sum rule. From the definition of $V(r)$ we get

$$\int_0^\infty rV(r)\,dr = -2 \lim_{r\to\infty} [r^{-1}K(r, r)]. \tag{3.15}$$

Now, since we can write from the analysis of (I) (Sec. 3.2)

$$K(r, r) \sim g(r, r) \sim \tfrac{1}{2}(\alpha + \beta)r[1 + O(r^{-\frac{1}{2}})],$$

we easily obtain the sum rule

$$\int_0^\infty rV(r)\,dr = -(\alpha + \beta), \tag{3.16}$$

where α is the arbitrary parameter and β is given as in (I). In the Born approximation, β is

$$\beta \sim \beta_0 = \sum_p \bar{\mu}_{2p} \tan \delta_{2p}. \tag{3.17}$$

Properties of the Jost Functions

For the class of potentials studied in Sec. 2, the properties of the Jost functions can easily be derived from formulas (2.4). It is obvious that the Jost functions are, in general, meromorphic functions, with poles depending on the values of μ. However, there is a remarkable exception. For an even potential (of that class), the Jost function is an entire function of ν of order 1, and finite type. It would be easy to obtain an interpolation formula with the help of the Lagrange Valiron theorem.[10] The location of the zeros of these Jost functions in the ν complex plane may be studied in the same way as, in (II), those of Newton's potentials, and with analogous results.

Transparent Potentials

The transparent potentials, which are associated to a set of phase shifts all equal to zero, can be derived with the help of the method of Sec. 2.1, in the same way as in (I). Obviously there is an infinity of solutions and it is clear that they are in general energy-dependent. In Born approximation, it is easy to write down the coefficients c_μ for a transparent potential: Let us choose a set of numbers c_μ ($\mu \neq l$) such that $|c_\mu| \leqslant C\mu^{-2}$, it follows from (3.6) that the coefficient c_l of the transparent potentials associated to this set of coefficients is given in the Born approximation, by the solution of the matrix equation:

$$\sum_0^\infty M_l^{l'} c_{l'} = \sum_\mu L_l^\mu c_\mu \sin(\tfrac{1}{2}\pi)(l - \mu), \tag{3.18}$$

which yields

$$c_{l'} = \sum_l \gamma_{l'}^l \sum_\mu L_l^\mu c_\mu \sin(\tfrac{1}{2}\pi)(l - \mu) + \alpha v_l, \tag{3.19}$$

where the matrix γ ($= M^{-1}$) and the vector $\hat{v}$ are defined as in (I), and α is an arbitrary parameter. Besides, with the help of well-known formulas,[11] it is easy to show directly that the contribution of $\hat{v}$ to the potential is transparent in the Born approximation.

Solvable Examples

Examples for which the solution can be written down explicitly can be constructed in the same way as in (II).

3.3. Born Approximation for Even Potentials and the General Approximation Problem

From Eqs. (2.35), (2.36), and (1.18) it is easy to show that for an even potential of the class defined in Sec. 2, we have

$$K(r, r) = \frac{4}{\pi^2} r \sum_{p=1}^\infty J_p(r)J_p(r) \sum_{l=0}^\infty \frac{(2p)^2}{(2p)^2 - (2l + 1)^2} \tan \delta_l. \tag{3.20}$$

This can as well be written

$$K(r, r) = \frac{2}{\pi} r \sum_{p=1}^\infty 2pJ_p(r)J_p(r) \times \int_0^\pi \sin 2px \sum_0^\infty \cos(2l + 1)x \tan \delta_l\,dx \tag{3.21}$$

or, using well-known formulas,[12]

$$K(r, r) = \frac{2r^2}{\pi} \int_0^\pi J_1(2r \sin x) \cos x\,dx \times \sum_0^\infty \cos(2l + 1)x \tan \delta_l. \tag{3.22}$$

Let us recall the well-known formula for the scattering amplitude in the Born approximation:

$$f(\theta) \sim \sum_0^\infty (2l + 1)\delta_l P_l(\cos \theta), \tag{3.23}$$

which is equivalent to

$$\tan \delta_l \sim \delta_l \sim \tfrac{1}{2} \int_0^\pi f(\theta)P_l(\cos \theta) \sin \theta\,d\theta. \tag{3.24}$$

[10] R. P. Boas, *Entire Functions* (Academic Press Inc., New York, 1954), p. 221.

[11] Bateman Manuscript Project, *Higher Transcendental Functions* (McGraw-Hill Book Company, Inc., New York, 1953), Formulas 7.15 (2) and 7.7 (30).

[12] Reference 11, formulas 7.15 (38) and 7.2 (51).

Let us now substitute this result in (3.22) and use a well-known formula[13] in order to evaluate the Fourier series. Performing the integral over x successively on the intervals $(0, \frac{1}{2}\pi)$ and $(\frac{1}{2}\pi, \pi)$, with the help of a known formula,[14] we then get the relation between $K(r, r)$ and $f(\theta)$:

$$K(r, r) = \frac{r}{2\pi}\int_0^{\pi} f(\theta) \sin\theta \, d\theta \frac{1 - \cos 2r \sin \frac{1}{2}\theta}{\sin(\frac{1}{2}\theta)}. \tag{3.25}$$

From this we readily derive $V(r)$

$$rV(r) = -\frac{2}{\pi}\int_0^{\pi} f(\theta) \sin(2r \sin(\tfrac{1}{2}\theta)) \sin\theta \, d\theta. \tag{3.26}$$

This remarkable result could be obtained in a much more elegant way. The functions of r $[rJ_p(r)J_p(r)]$ are odd entire functions of order 1 and type 2. Now, using some evaluations of (I) (Sec. 1.1), it is easy to show that

$$c_{l-\frac{1}{2}} \sim \text{const} + O(l^{-1-\epsilon}) \quad (l \to \infty) \tag{3.27}$$

provided only that the phase shifts are bounded by $C^{l-2-\epsilon}$.

We use this bound and a well-known formula[15] in order to derive the asymptotic behavior of $K(r, r)$ for large r, and deduce from the result that $rV(r)$ is an odd entire function of order 1 and type 2.

Let us now study the exponential Fourier transform of $rV(r)$ or, since $rV(r)$ is odd, its sine Fourier transform

$$-u\bar{f}(u) = \int_0^{\infty} rV(r) \sin ru \, dr. \tag{3.28}$$

Provided that $rV(r)$ belongs to $\mathfrak{L}_1$, the Paley–Wiener theorem ensures that[10] $\bar{f}(u)$ is equal to zero for $|u|$ larger than 2. If we settle

$$u = 2 \sin(\tfrac{1}{2}\theta), \tag{3.29}$$

formula (3.28) is nothing but the well-known formula giving the scattering amplitude $f(\theta)$ $[= \bar{f}(u)]$ in terms of the potential. Our result proves that the even potentials correspond to a continuation of $f(\theta)$ on the real axis equal to zero outside $(-2, 2)$. The inversion formula for the sine Fourier transform then readily yields (3.26).

Physical Meaning of the Result

In order to see the physical meaning of our result concerning the analytical properties of the even potentials in the class defined in Sec. 2, let us now recall some well-known properties. For large values of the energy, the scattering amplitude due to a regular static potential, as for instance the nuclear potential, strongly decreases with the angle. In the Born approximation, this amplitude is given by (3.28), or, by introducing explicitly the energy through the wave-number k:

$$2k \sin(\tfrac{1}{2}\theta) f(k, \theta) = \int_0^{\infty} \rho U(\rho) \sin 2k\rho \sin(\tfrac{1}{2}\theta) \, d\rho, \tag{3.30}$$

where $U(\rho)$ is now energy independent.

We know the sine Fourier transform of $\rho U(\rho)$ only on the interval $(0, 2k)$, and the potential we construct is the inverse sine transform of the function equal to this Fourier transform inside this interval and equal to zero outside the interval. In other words, the scattering amplitude can be considered as a measure of the potential, made through a linear filter whose range increases with the energy. The problem is then: Is the potential $U(E, \rho)$ which we construct near the unknown potential $U(\rho)$ from which we start, and how near. We see that this reduces to an old problem,[16] which is still of current interest. For a detailed study of this problem, we refer to classical treatises,[16] and we only give here some criteria to evaluate the standard deviations. It is reasonable to assume that the unknown function $\rho U(\rho)$ belongs to $\mathfrak{L}_2$, and is bounded, with only isolated singularities. As a result it is possible to find a number M so that the Fourier transform $F(u)$ of $\rho U(\rho)$ be bounded as

$$|F(u)| < M/|u|.$$

As a result, if $r_0(k)$ is the square deviation between $U(E, \rho)$ and $U(\rho)$

$$r_0(k) = \int_0^{\infty} [U(\rho) - U(E, \rho)]^2 \, d\rho, \tag{3.31}$$

well-known properties[17] of the Fourier transform lead us to the bound

$$r_0(k) \leq (4k)^{-1}\pi M^2. \tag{3.32}$$

The standard deviation between $U(E, \rho)$ and $U(\rho)$ goes to zero as $E \to \infty$.

The assumptions we made on $U(\rho)$ are very weak. If we assume that the regularity of the unknown potential is sufficient to ensure that its Fourier transform cannot decrease beyond $\theta = \pi$ more slowly than before $\theta = \pi$, it is possible to get a bound on the mean square deviation much stronger than (3.32). Let us assume for instance that $F(u)$ decrease faster

[13] Reference 11, formula 3.10 (2).

[14] Bateman Manuscript Project, *Integral Transforms* (McGraw-Hill Book Company, Inc., New York, 1953), formula 8.4 (4).

[15] Reference 11, formula 7 (15)(38).

[16] See, for example, J. Arsac, *Transformation de Fourier et théorie des distributions* (Dunod, Paris, 1961), p. 225ff.

[17] See Ref. 16, p. 254ff.

then $F(2)(\frac{1}{2}u)^{-p}$, where p is in general much larger than 1, we obtain:

$$r_0(k) \leq 8[k^2 f(k, \pi)]^2 / 2^{2p+1}, \tag{3.33}$$

which gives an order of the deviation as a function of the cross section for $\theta = \pi$.

In previous publications,[18,19] we gave a study of the semiclassical approximations and emphasized the following points:

(1) The validity of the JWKB method, in order to derive the phase shifts, is related to the existence of bounds for the derivatives of the potential.

(2) The validity of the evaluation of $f(\theta)$ as an integral over the momentum, is related to the behavior of the Fourier transform of the phase shifts, and by this way can be related either to the behavior of the scattering amplitude at large angles or to the analytic properties of $V(r)$. The cutoff which is necessary to make this second step of the semiclassical theory exact, is the same as above, i.e., $f(u)$ equal to zero for $u \geq \pi$. If this cutoff is assumed to be valid, it is easy to derive[19] a unique interpolation formula for the phase shifts, from which a unique potential can be calculated in the framework of JWKB theory.

From the mathematical point of view,[20] we see therefore that the conditions for the potentials of the class studied in Sec. 2.1 are by no way more restrictive than the conditions which we assume if we try to solve the inverse problem in the semiclassical theory.

It is unfortunate to find potentials which depend on the energy. However, since their standard deviation from a static potential can be very small, we may hope that wherever the static potential is a smooth function, the potential we obtain is nearly the same, and that only at those points where the potential behaves too steeply (according to Bernstein theorem) to have a finite spectrum, there are energy-dependent oscillations.

APPENDIX A

Let us assume that $zV(z)$ is analytic in a circle $|z| < \Gamma$, and define $K(z, z)$ from $V(z)$ by the following formula:

$$K(z, z) = -\tfrac{1}{2}z \int_0^z uV(u)\, du. \tag{A1}$$

We prove in this Appendix that $K(z, z)$ can be given, in general, by the following expansion:

$$K(z, z) = \sum_0^\infty s_l \phi_l(z) u_l(z) + \sum_0^\infty s_{l+\frac{1}{2}} \phi_{l+\frac{1}{2}}(z) u_{l+\frac{1}{2}}(z) \tag{A2}$$

convergent in the circle $|z| < \Gamma$, and where the expansion coefficients are unique. In order to prove this statement, we need some upper bounds for the Taylor coefficients of the functions involved. We introduce auxiliary functions $c_\lambda(z)$ equal to $z^{-\lambda-1}\phi_\lambda(z)$, and which are solutions of the equation:

$$L(z)c_\lambda(z) \equiv [z^{-\lambda-1}D(z)z^{\lambda+1} - \lambda(\lambda+1)]c_\lambda(z) = 0. \tag{A3}$$

Application of the Frobenius[21] method to this equation leads us to a solution. We look for a development of the functions $c_\lambda(z)$ of the form

$$c_\lambda(z) = \sum_p c_\lambda^p z^{p+\rho}, \quad c_\lambda^0 = N_\lambda \neq 0, \tag{A4}$$

where N_λ are normalization coefficients. Introducing for convenience the analytic function $w(z)$ equal to $z^2(1 - V(z))$, we define a set of coefficients f_λ^n as

$$\begin{aligned} z^{-p-\rho} L z^{p+\rho} &= \sum_{n=0}^\infty f_\lambda^n (p+\rho) z^n \\ &= (p+\rho)(2\lambda + 1 + p + \rho) + w(z). \end{aligned} \tag{A5}$$

Equation (A3) is equivalent to the system:

$$\left\{\begin{array}{l} c_\mu^0 f_\lambda^0(\rho) = \rho(2\lambda + 1 + \rho) = 0, \\ \cdot \\ \cdot \\ \cdot \\ c_\lambda^p f_\lambda^0(\rho + p) + c_\lambda^{p-1} f_\lambda^1(\rho + p - 1) + \cdots c_\lambda^0 f_\lambda^p(\rho) = 0. \end{array}\right. \tag{A6}$$

For the solution of interest, ρ is zero. Besides, the $f_\lambda^n(\rho + p)$ depend only upon n for n larger than 1. We use hereafter the simplified notation f_n.

Let $M(R)$ be the upper bound of $|w'(z)|$ on the circle $|z| = R = \Gamma - \epsilon$; we can get upper bounds for the f_n by the Cauchy integral theorem:

$$|f_{n+1}| = \left| \frac{1}{2\pi i} \int_R \frac{w'(z)}{z^{n+1}}\, dz \right| \leq \frac{M(R)}{R^n} \tag{A7}$$

The following relations define therefore upper

[18] P. C. Sabatier, Nuovo Cimento 37, 1180 (1965).

[19] P. C. Sabatier, Thèse de Doctorat, Faculté des Sciences d'Orsay, Série A, No. 153 (1966).

[20] It should be borne in mind that our study is valid *only* in Born approximation. Even potentials of the class studied in Sec. 2 may have poles in the r complex plane, so that their Fourier transform certainly does not vanish outside a finite interval. In order to study completely the standard deviation of these potentials, one should introduce also deviations from the linear approximations, which we neglect here.

[21] E. L. Ince, *Ordinary Differential Equations* (Dover Publications, Inc., New York, 1956), p. 396.

bounds C_λ^p for the c_λ^p:

$$\begin{cases} C_\lambda^{p+1} = \dfrac{M(R)}{(p+1)(2\lambda+p+2)} \\ \qquad \times \{C_\lambda^p + C_\lambda^{p-1}R^{-1} + \cdots C_\lambda^0 R^{-p}\}, \\ \cdot \\ \cdot \\ \cdot \\ C_\lambda^1 = [M(R)/(2\lambda+2)]C_\lambda^0, \\ C_\lambda^0 = |N_\lambda|. \end{cases} \tag{A8}$$

(A8) can easily be solved:

$$C_\lambda^{p+1} = |N_\lambda| \prod_{n=0}^{n=p} \left[\frac{M(R)}{(n+1)(2\lambda+n+2)} + \frac{n(2\lambda+n+1)}{(n+1)(2\lambda+n+2)R}\right]. \tag{A9}$$

For large enough λ, $RM(R)$ is smaller than $(2\lambda+1)$. If we replace $M(R)$ by $(2\lambda+n+1)R^{-1}$ in the factors of (A8) except that for which $n=0$, we easily get the inequality

$$|c_\lambda^p| \leq C_\lambda^{p+1} < |N_\lambda|\, M(R)(2\lambda+p+2)^{-1}R^{-p}. \tag{A10}$$

Since 0 is an entire function, we can derive for any value of R, and, in particular, for $R < \Gamma - \epsilon$, an inequality analogous to (A10), in which the c_λ^p are to be replaced by the Taylor coefficients of $z^{-\lambda-1}u_\lambda(z)$. Let us now consider the product

$$\tau_{2\lambda}(z) = \phi_\lambda(z)u_\lambda(z) = z^{2\lambda+2}\sum_q \tau_{2\lambda}^{2\lambda+q} z^q N_\lambda^2. \tag{A11}$$

The coefficient $\tau_{2\lambda}^{2\lambda+q}$ $(q > 1)$ can be bounded by

$$R^{-q}B_1(R)B_2(R)\left\{\sum_{p=1}^{q-1}(2\lambda+2+p)^{-1}(2\lambda+q-p+2)^{-1} + (2\lambda+q+2)^{-1}[(B_1(R))^{-1} + (B_2(R))^{-1}]\right\}N_\lambda^2,$$

where $B_1(r)$ and $B_2(r)$ are associated respectively with $\phi_\lambda(z)$ and $u_\lambda(z)$. Comparison of the first series with an integral and some majorizations show readily that this expression can be bounded by

$$|\tau_{2\lambda}^{2\lambda+q}| < R^{-q}B_3(R)\frac{\log(2\lambda+2+q)}{2\lambda+2+q}N_\lambda^2, \tag{A12}$$

where $B_3(R)$ is finite for $R < \Gamma - \epsilon$.

It is clear, now, that $K(z, z)$ is analytic in the circle $|z| < \Gamma$, so that a majorization analogous to (A7) holds for its Taylor coefficients:

$$K(z,z) = z^2\sum_{n=0}^{\infty} k_n z^n, \tag{A13}$$

$$k_n < K(R_1)R_1^{-n} \quad (R_1 < \Gamma). \tag{A14}$$

It is now easy to expand $K(z, z)$ in terms of the functions $\tau_p(z)$: we only need to compare the coefficients of z^{n+2} on both sides of the following relation:

$$\sum_{p=0}^{\infty}\sigma_p\tau_p(z) = z^2\sum_{n=0}^{\infty}k_n z^n. \tag{A15}$$

This leads us to the system

$$\sigma_n\tau_n^n = k_n - \sum_{p=0}^{n-1}\sigma_p\tau_p^n. \tag{A16}$$

If we replace any term in (A16) by an upper bound of its modulus, and use for that (A12) and (A14), with $R_1 = \alpha^{-1}R(\alpha < 1)$, we obtain a system whose solutions yield upper bounds for the $|\sigma_n|$:

$$|\sigma_n| < (\tau_n^n)^{-1}R^{-n}v_n, \tag{A17}$$

where

$$\frac{n+2}{\log(n+2)}v_n = \frac{(n+2)}{\log(n+2)}\alpha^n K(R_1) + B(R)\sum_{p=0}^{n-1}v_p. \tag{A18}$$

We can replace the first term in (A18) by an upper bound $K_1(R_1)$, or $K_2(R)$. The system (A18) is then easy to solve, and we find that v_n can be bounded independently of n, so that

$$|\sigma_n| < (\tau_n^n)^{-1}R^{-n}S(R) \quad (R < \Gamma - \epsilon). \tag{A19}$$

The results given in the beginning of this appendix follow readily. In particular, the method of construction shows clearly that the coefficients s_l and $s_{l+\frac{1}{2}}$, respectively, equal to σ_{2l} and σ_{2l+1}, are unique. If $w(z)$ is an even function, it is easy to see that the f_λ^{2p+1} are equal to zero, therefore the c_λ^{2p+1} are equal to zero, and a rapid investigation shows that the s_l are equal to zero.

Extension to the "Rational" Cases

Let us now assume that $zV(z)$ is an analytic function of z^α, where α is a rational number inside the circle $|z| < \Gamma$. We can obviously limit our study to the case where α is equal to the inverse m^{-1} of an integral number. Let us then put $z = t^m$ and assume that $zV(z)$ is an analytic function of t. Equation (A3) holds with transformations schematically written as

$$z \to t^m,$$
$$D(z) \to \bar{D}(t),$$
$$L(z) \to \bar{L}(t).$$

The Frobenius method can be used as above; that is to say, we write

$$\bar{L}(t)\bar{c}_\lambda(t) = \bar{L}(t)\sum_q c_\lambda^q t^{q+\sigma} = 0;$$

we calculate

$$t^{-q-\sigma}L(t)t^{q+\sigma} = \sum_{k=0}^{\infty} f_\lambda^l(q+\sigma)t^k$$
$$= \frac{q+\sigma}{m}\left(2\lambda + 1 + \frac{q+\sigma}{m}\right) + \bar{w}(t),$$

where $\bar{w}(t)$ is equal to

$$\bar{w}(t) = t^{2m}[1 - \bar{V}(t)]. \tag{A20}$$

Provided that $\bar{w}(0)$ is zero,

$$[(q+\sigma)/m]\{2\lambda + 1 + [(q+\sigma)/m]\}$$

is equal to $f_\lambda^0(q+\sigma)$, whereas $f_\lambda^k(q+\sigma)$ does not depend on k for $k > 0$. We can obtain for the f_λ bounds similar to those of (A7), and the systems of equations which yield the coefficients c^p and C^p are similar to the systems (A6) and (A8). In particular, the system (A8) gives place to

$$\begin{cases} C_\lambda^{p+1} = \dfrac{m^2\bar{M}(\bar{R})}{(p+1)[m(2\lambda+1)+p+1]} \\ \qquad \times [C_\lambda^p + C_\lambda^{p-1}\bar{R}^{-1} + \cdots C_\lambda^0\bar{R}^{-p}] \\ \cdot \\ \cdot \\ C_\lambda^1 = \dfrac{m^2\bar{M}(\bar{R})}{m(2\lambda+1)+1}, \\ C_\lambda^0 = |N_\lambda|. \end{cases} \tag{A21}$$

The theory is therefore the same as above, with the trivial modifications

$$\lambda \to \Lambda = \lambda m,$$
$$M(R) \to m^2\bar{M}(\bar{R}).$$

All derivations can be done readily and lead us to the formula

$$K(z,z) = \sum_{p=0}^{\infty} c_{p/m}u_{p/m}\phi_{p/m} + \sum_{p=0}^{\infty} c_{(p+\frac{1}{2})/m}u_{(p+\frac{1}{2})/m}\phi_{(p+\frac{1}{2})/m}, \tag{A22}$$

which is valid when $zV(z)$ is an analytic function of t.

Remark: Since it is sufficient that $\bar{w}(t)$ is equal to zero for $t = 0$ and since we only need, for this, $z^{(2-m^{-1})}V(z)$ to be finite at the origin, we can extend the theory to this case. The expansion we find for $K(z,z)$ is similar to (A22) but μ can go until $-\frac{1}{2} + \frac{1}{2m}$:

$$K(z,z) = c_{-\frac{1}{2}+1/2m}u_{-\frac{1}{2}+1/2m}\phi_{-\frac{1}{2}+1/2m} + c_{-\frac{1}{2}+1/m}u_{-\frac{1}{2}+1/m}\phi_{-\frac{1}{2}+1/m} + \cdots. \tag{A23}$$

APPENDIX B

Let ϕ_l and ψ_l be, respectively, the solutions of Eq. (1.17) with potentials V_1 and V_2, let us calculate the integral

$$I(r) = \int_0^r x^{-1}\phi_{l'}\psi_{l'}\frac{d}{dx}(x^{-1}\phi_l\psi_l)\,dx, \tag{B1}$$

where l and l' are numbers larger than $-\frac{1}{2}$, and not necessarily integers. Integration by parts yields

$$2I(r) - [x^{-2}\phi_{l'}\psi_{l'}\phi_l\psi_l]_0^r = \int_0^r \left\{x^{-1}\phi_{l'}\psi_{l'}\frac{d}{dx}(x^{-1}\phi_l\psi_l) - x^{-1}\phi_l\psi_l\frac{d}{dx}(x^{-1}\phi_{l'}\psi_{l'})\right\}dx. \tag{B2}$$

Some algebra and use of well-known Wronskian properties show that the right-hand side of (2) is equal to

$$[l(l+1) - l'(l'+1)] \times \int_0^r dx\,\frac{d}{dx}\left[\int_0^x \phi_{l'}\phi_l y^{-2}\,dy \int_0^x \psi_{l'}\psi_l y^{-2}\,dy\right]$$

or

$$[l(l+1) - l'(l'+1)]\left(\phi_{l'}\frac{d}{dx}\phi_l - \phi_l\frac{d}{dx}\phi_{l'}\right) \times \left(\psi_{l'}\frac{d}{dx}\psi_l - \psi_l\frac{d}{dx}\psi_{l'}\right). \tag{B3}$$

When the potentials V_1 and V_2 are regular enough, the functions ϕ_l and ψ_l behave asymptotically as $A_l \sin(r + \delta_l - \frac{1}{2}l\pi)$ and $B_l \sin(r + \epsilon_l - \frac{1}{2}l\pi)$. It follows that $I(\infty)$ reduces to

$$I(\infty) = \frac{1}{2}\frac{A_lA_{l'}B_lB_{l'}}{l(l+1) - l'(l'+1)} \times \sin[\delta_{l'} - \delta_l + (l-l')(\tfrac{1}{2}\pi)] \times \sin[\epsilon_{l'} - \epsilon_l + (l-l')(\tfrac{1}{2}\pi)]. \tag{B4}$$

For V_1 and V_2 equal to zero, the functions ϕ and ψ reduce to the free-particle wavefunctions which are related to Bessel functions through (1.4). Using (B4), (3.12) can be obtained straightforwardly from (3.11).

APPENDIX C

We want to study the convergence of the expansions in powers of λ which may be introduced in order to solve systems (2.37), (2.39). Let us first define the class C_2 of the vectors whose lth component has the bound $C(l+1)^{-2}$, where C is a constant independent

of l. We call $C(\hat{w})$ the norm of the vector $\hat{w}$. Let us now put

$$a_{p-\frac{1}{2}} = -\alpha s_p + b^{(p)}, \tag{C1}$$

$$-m^{-1}\tan \delta Im\hat{s} = \lambda \hat{v}_0,$$

$$-m^{-1}\tan \delta Im I = \lambda D,$$

$$m^{-1}\tan \delta m + m^{-1} Im I = B. \tag{C2}$$

With the help of bounds given in I (Sec. 1.2.3), it is a matter of simple evaluation to show that $\hat{v}_0$ belongs to C_2 and that the application of any of the matrices ${}^{(0)}M$, m^{-1}, B, D, to a vector of C_2 leads us to a vector of C_2. We can write this schematically as

$$C(N\hat{v}) \leq C(N)C(\hat{v}), \tag{C3}$$

where N is the "norm" of the matrix N. Obviously, we can take the same norm C for the four matrices quoted above. Let us now notice that

$${}^{(0)}M\hat{s} = 0. \tag{C4}$$

Substitution of (C1), (C2), (C4), and (2.40) in systems (2.37), (2.39), together with the condition ($\frac{1}{2}\alpha = b^{(0)}$) yields the following systems:

$$\lambda m^{-1}\delta\hat{e} = \hat{b} + B\tan \delta'\hat{b} - 2b^{(0)}B\tan \delta'\hat{s} + 2\lambda b^{(0)}\hat{v}_0 + \lambda D\hat{b},$$

$$\tan \delta'\hat{e} = M\hat{b} + \tan \delta' M \tan \delta'\hat{b} - 2b^{(0)}\tan \delta' M \tan \delta'\hat{s}, \tag{C5}$$

for which we look for a solution of the form

$$\hat{b} = \lambda\hat{b}_1 + \lambda^2\hat{b}_2 + \cdots,$$

$$\tan \delta' = \lambda\hat{\delta}_1 + \lambda^2\hat{\delta}_2 + \cdots. \tag{C6}$$

Substitution of these formal expansions in (C5) yields

$$m^{-1}\delta\hat{e} = \hat{b}_1,$$

$$\hat{\delta}_1\hat{e} = M\hat{b}_1, \tag{C7}$$

for the first order and

$$\hat{b}_n = -B\sum_1^{n-1}\delta_p\hat{b}_{n-p} + 2\sum_1^{n-1} b_p^{(0)}B\delta_{n-p}\hat{s} - 2b_{n-1}^{(0)}\hat{v}_0 - D\hat{b}_{n-1},$$

$$\delta_n\hat{e} = M\hat{b}_n + \sum_{p=1,q=1}^{(n-1)}\delta_p M\delta_q\hat{b}_{n-p-q} - 2\sum_{p=1,q=1}^{(n-1)} b_{n-p-q}^{(0)}\delta_q M\delta_p\hat{s} \tag{C8}$$

for the nth order. This recurrent system enables us to obtain $\hat{b}_n$ from the values of $\hat{b}_p$ and $\hat{\delta}_p$ for $p < n$, and $\delta_n\hat{e}$ from $\hat{b}_n$ and those values.

It is clear that all the vectors involved in (C8) belong to C_2. Let B_p and Δ_p be respectively the norm of $\hat{b}_p$ and $(\delta_p\hat{e})$ for all the values of p lower than n. The norm of $\hat{b}_n$ and δ_n is smaller than the number given by the solution of the system:

$$B_n = 3C\left[B_{n-1} + \sum_{p=1}^{n-1}\Delta_p B_{n-p}\right],$$

$$\Delta_n = C\left[B_n + 3\sum_{p=1,q=1}^{n-1}\Delta_p\Delta_q B_{n-p-q}\right]. \tag{C9}$$

We can always take C larger than 1, so that Δ_n is larger than CB_n, and B_n is smaller than

$$3\left[\Delta_{n-1} + \sum_{p=1}^{n-1}\Delta_p\Delta_{n-p}\right].$$

We are led therefore to study the recurrent equation

$$\Delta_n = 3C\left[\Delta_{n-1} + \sum_{p=1}^{n-1}\Delta_p\Delta_{n-p} + \sum_{p=1,q=1}^{n-1}\Delta_p\Delta_q\Delta_{n-p-q}\right],$$

$$\Delta_1 = \Delta_1. \tag{C10}$$

This system can be studied in turn by putting x^n in factor of both members and performing the summation from 2 to ∞. Let us put

$$xf(x) = \sum_1^{\infty}\Delta_p x^p. \tag{C11}$$

It is clear that the equivalent equation is

$$f(x) - \Delta_1 = 3Cx\{f(x) + [f(x)]^2 + x[f(x)]^3\}. \tag{C12}$$

With the help of well-known results on the second degree equation, it is easy to show that x is a simple[22] analytical function of $[f(x) - \Delta_1]$ in some circle centered at the origin.

It follows that $f(x)$ is analytic in some circle centered at the origin. As a result, Δ_n and B_n are bounded by the Taylor coefficients of an analytic function in some circle centered at the origin. It follows that the formal expansions in powers of λ converge in some circle centered at the origin.

It would be easy to extend our method to the case in which δ_l is bounded by $c(l+1)^{-\alpha}$, where α is a positive number ≤ 2.

[22] E. C. Titchmarsh, *Theory of Functions* (Oxford University Press, New York, 1932), p. 198.

Herglotz–Noether Theorem in Conformal Space–Time*

HUGO D. WAHLQUIST AND FRANK B. ESTABROOK
Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California

(Received 23 August 1966)

The Herglotz–Noether theorem is proved for rigid motions in conformal space–time.

THE theorem stating that isometries are the only rotating rigid motions permitted in special relativity was originally proved by Herglotz[1] and Noether.[2] Several papers devoted to rigid motions in general relativity have considered the possibility of extending the validity of this theorem to curved space–times.[3–6] Using the dyadic formalism,[6,7] we have now succeeded in proving the theorem for all conformally flat manifolds. The approach is quite similar to that used in Ref. 6 to treat rotating rigid motions in Einstein spaces.

The curvature dyadic, $\mathbf{E}$, of the quotient space of a rigid ($\mathbf{S} = 0$) congruence has vanishing time derivative in body-fixed axes ($\boldsymbol{\omega} = \boldsymbol{\Omega}$), i.e., $\dot{\mathbf{E}} = 0$ (vid., Sec. II of Ref. 6); and from the commutation properties of temporal and transverse spatial derivatives for rigid congruences, it follows that the symmetric, traceless dyadic $\mathbf{F}$ defined by

$$\mathbf{F} \equiv \tfrac{1}{2}[\nabla \times \mathbf{E} - \mathbf{E} \times \nabla] \tag{1}$$

must satisfy $\dot{\mathbf{F}} = 0$. Applying the generalized Gauss equation [Eq. (D.34) of Ref. 7] to the quotient space of a rigid congruence in conformally flat ($\mathbf{A} = \mathbf{B} = 0$) manifolds leads to the following expression for $\mathbf{E}$;

$$\mathbf{E} = \mathbf{T} - 3\boldsymbol{\Omega}\boldsymbol{\Omega} + \tfrac{1}{3}(2\rho - \operatorname{Tr}\mathbf{T})\mathbf{I}, \tag{2}$$

where $\boldsymbol{\Omega}$ is the angular velocity of the congruence, $\mathbf{T}$ is the stress dyadic, and ρ is the energy density. Using the curvature equations and Bianchi identities of Ref. 7, an explicit expression for $\mathbf{F}$ can now be derived;

$$\mathbf{F} = 6\boldsymbol{\Omega}(\mathbf{t} - \boldsymbol{\Omega} \times \mathbf{a}) + 6(\mathbf{t} - \boldsymbol{\Omega} \times \mathbf{a})\boldsymbol{\Omega} - 4\boldsymbol{\Omega} \cdot (\mathbf{t} - \boldsymbol{\Omega} \times \mathbf{a})\mathbf{I}, \tag{3}$$

where $\mathbf{a}$ is the acceleration and $\mathbf{t}$ is the momentum density vector. From the time constancy of the two invariants of $\mathbf{F}$ it then follows that

$$(|\boldsymbol{\Omega}|\,|\mathbf{t} - \boldsymbol{\Omega} \times \mathbf{a}|)^{\cdot} = (\boldsymbol{\Omega} \cdot \mathbf{t})^{\cdot} = 0. \tag{4}$$

At this point the degenerate case $\boldsymbol{\Omega} \cdot \mathbf{t} = 0$ must be treated separately. First, we assume $\boldsymbol{\Omega} \cdot \mathbf{t} \neq 0$, which permits defining the unit vectors

$$\hat{\mathbf{u}} \equiv \frac{\boldsymbol{\Omega}}{|\boldsymbol{\Omega}|}, \quad \hat{\mathbf{v}} \equiv \frac{\mathbf{t} - \boldsymbol{\Omega} \times \mathbf{a}}{|\mathbf{t} - \boldsymbol{\Omega} \times \mathbf{a}|}, \tag{5}$$

and from $\dot{\mathbf{F}} = 0$ we find

$$\dot{\hat{\mathbf{u}}} = \dot{\hat{\mathbf{v}}} = 0. \tag{6}$$

However, using the vector equation (D.28) of Ref. 7 we can derive the equality

$$\hat{\mathbf{u}} \cdot (\nabla \times \hat{\mathbf{u}}) = -2(\boldsymbol{\Omega} \cdot \mathbf{t})/\Omega^2, \tag{7}$$

which with the previous results, Eqs. (4)–(6), demands

$$\dot{\boldsymbol{\Omega}} = 0, \tag{8}$$

so that $\boldsymbol{\Omega}$ is a body-fixed vector. In the case $\boldsymbol{\Omega} \cdot \mathbf{t} = 0$ we can arrive at the same result by a different route, since, again with the dyadic equations of Ref. 7, it can be shown that

$$\nabla(\boldsymbol{\Omega} \cdot \mathbf{t}) = 3(\Omega^2)\dot{\boldsymbol{\Omega}} - 3(\boldsymbol{\Omega} \cdot \mathbf{t})\mathbf{a}. \tag{9}$$

Thus, Eq. (8) is valid for all rigid motions in conformally flat manifolds, and it is easily shown that, when $\Omega^2 \neq 0$, this equation entails the further conditions

$$\nabla \times \mathbf{a} = \dot{\mathbf{a}} = 0, \tag{10}$$

which together with the constraint of rigidity, $\mathbf{S} = 0$, are the necessary and sufficient conditions for a Killing vector, or isometric, congruence.

* This paper presents results of one phase of research carried out at the Jet Propulsion Laboratory, California Institute of Technology, under Contract NAS7-100, sponsored by the National Aeronautics and Space Administration.

[1] G. Herglotz, Ann. Physik **31**, 393 (1910).
[2] F. Noether, Ann. Physik **31**, 919 (1910).
[3] G. Salzman and A. H. Taub, Phys. Rev. **95**, 1659 (1954).
[4] F. A. E. Pirani and G. Williams, Séminaire JANET, 5ième année, No. 8 (1961); *ibid.* No. 9 (1962).
[5] R. H. Boyer, Proc. Roy. Soc. (London) **A283**, 343 (1965).
[6] H. D. Wahlquist and F. B. Estabrook, J. Math. Phys. **7**, 894 (1966).
[7] F. B. Estabrook and H. D. Wahlquist, J. Math. Phys. **5**, 1629 (1964).

Theory of Diffraction by a Curved Inhomogeneous Body

JAMES R. WAIT*
Institute for Telecommunication Sciences and Aeronomy, Environmental Science Services Administration, Boulder, Colorado

(Received 21 September 1966)

The diffraction of electromagnetic waves by a convex spherical surface is considered. The problem is formulated in terms of the mutual impedance between two radially oriented electric dipoles for a smooth inhomogeneous surface. The general integral equation is simplified when the surface is sectionally homogeneous. In this case, the result is given as a twofold integral which involves the product of pattern functions associated with a homogeneous spherical surface. The reduction of these pattern functions to manageable form is one of the major objectives of the paper. The final result shows explicitly how the classical Fresnel diffraction pattern is modified by the radius of curvature and the properties of the diffracting surface. The special case where the diffracting surface is perfectly conducting except for a small inhomogeneity near the crest reduces to a very simple formula if the grazing angle is zero. It is shown that this result is compatible with the perturbation theory for scattering from an impedance strip on a conducting plane.

1. INTRODUCTION

THE influence of the curvature of an obstacle on the diffraction factor is an intriguing subject. A direct application of physical-optics theory leads to the Cornu-spiral formulas for the Fresnel diffraction pattern. These classically derived results (e.g., Born and Wolf[1]) show no dependence on the shape or properties of the diffracting edge. However, in 1947, Pekeris[2] showed that the diffraction pattern of an elevated antenna in the vicinity of the earth's shadow could be decomposed into two parts. For angles near grazing, the first term was the classical physical optics result which can be expressed in terms of Fresnel integrals. The second term, which constituted a correction, was an explicit function of the curvature of the diffracting surface. The same development was rediscovered by Fock[3] in a highly celebrated paper. Following similar ideas, Rice[4] has developed a representation for parabolic cylinders, while Wait and Conda[5] have developed the theory and given extensive results for finitely conducting cylindrical obstacles. At the same time, the latter authors extended Fock's theory for the sphere to permit the results to be used when the antenna heights were not restricted to be small (compared with the radius of curvature of the diffracting surface). Some interesting applications of these results to radio propagation are given in a paper by Dougherty and Maloney,[6] who also give some additional numerical data. Finally, we should mention that a useful general discussion of the problem and a summary of various formulas have been given by Logan and Yee.[7]

In this paper, we wish to develop the theory for a diffracting surface, which, while smooth, is inhomogeneous. The problem is treated as an extension of a previous formulation for radio propagation over a mixed-path spherical earth.[8,9] The final results, while quite complicated, may be simplified in the interesting situation where the grazing angle is near zero and the surface inhomogeneity is localized at the crest.

2. MUTUAL IMPEDANCE FORMULATION FOR HOMOGENEOUS SPHERICAL SURFACE

The mutual impedance Z_{ab} between two vertical electric dipoles located at A and B over a spherical earth is considered for an implied time factor $\exp(i\omega t)$. The situation is illustrated in Fig. 1(a), where a vertical cross section is shown. The great circle distance between A and B (measured along the surface of the earth) is d. When the earth is homogeneous, it is permissible to express the mutual impedance in the following form:

$$Z_{ab} = (l_a l_b i\mu_0\omega/2\pi d)e^{-ikd}W(d, Z), \qquad (2.1)$$

where l_a and l_b are the effective lengths of the dipoles A and B, $k = 2\pi/(\text{free-space wavelength})$, $\mu_0 = 4\pi \times 10^{-7}$, and $W(d, Z)$ is a dimensionless attenuation function which is a function of d and the surface impedance of the earth. A contour integral representation for the attenuation function, for zero antenna

* Visiting Professor of Applied Physics, Harvard University, Cambridge, Massachusetts.

[1] M. Born and E. Wolf, *Principles of Optics* (Pergamon Press, Inc., New York, 1959).

[2] C. L. Pekeris, J. Appl. Phys. **18**, 667, 1025 (1947).

[3] V. A. Fock, Usp. Fiz. Nauk SSSR **43**, 587 (1950).

[4] S. O. Rice, Bell System Tech. J. **33**, 417–504 (1954).

[5] J. R. Wait and A. M. Conda, J. Res. Natl. Bur. Std. **63D**, 181 (1959).

[6] H. Dougherty and L. Maloney, Radio Sci. J. Res. NBS/USNC-URSI **68D**, No. 2, 239 (1964).

[7] N. A. Logan and K. S. Yee, in *Electromagnetic Waves*, R. E. Langer, Ed. (University of Wisconsin Press, Madison, 1962), pp. 139–180.

[8] J. R. Wait, Can. J. Phys. **36**, 9 (1958).

[9] J. R. Wait, in *Advances in Radio Research*, J. A. Saxton, Ed. (Academic Press Inc., New York, 1964), Vol. 1. pp. 157–217.

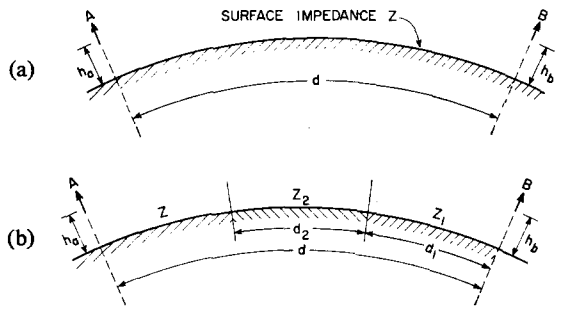


FIG. 1. (a) Cross-sectional view of the curved surface for a homogeneous path between A and B. (b) Cross-sectional view of the curved surface for an inhomogeneous path indicating the meaning of the parameters.

heights, is given by

$$W(d, Z) = W(x, q) = \frac{1}{2}\left(\frac{ix}{\pi}\right)^{\frac{1}{2}} \int_\Gamma \frac{e^{-ixt} w_1(t)}{w_1'(t) - q w_1(t)}\, dt, \tag{2.2}$$

where the contour Γ is in the complex t plane from $\infty \exp[-i(\frac{2}{3}\pi)]$ along a straight line to 0 and then out along the real axis to $+\infty$. Other parameters are $x = (\frac{1}{2}ka)^{\frac{1}{3}}(d/a)$, $q = -i(\frac{1}{2}ka)^{\frac{1}{3}} Z/\eta_0$, $\eta_0 = 120\pi$; a is the earth's radius, while $w_1(t) = \pi^{\frac{1}{2}}[\mathrm{Bi}\,(t) - i\,\mathrm{Ai}\,(t)]$, and $w_1'(t) = \pi^{\frac{1}{2}}[\mathrm{Bi}'\,(t) - i\,\mathrm{Ai}'\,(t)]$ are Airy functions in the standard notation.[10] Various methods are now available for evaluating the contour integral given by $W(x, q)$ above. For purposes of the present paper, it is assumed that $W(x, q)$ is a known quantity.

For arbitrary antenna heights h_a and h_b, the appropriate form for the attenuation function is

$$W(x, q, y_a, y_b) = \tfrac{1}{2} e^{\frac{1}{4}i\pi}\left(\frac{x}{\pi}\right)^{\frac{1}{2}} \int_\Gamma e^{-ixt} F(t, q, y_a, y_b)\, dt, \tag{2.3}$$

where

$$F = w_1(t - y_b) \times \left[v(t - y_a) - \frac{v'(t) - q v(t)}{w_1'(t) - q w_1(t)} w_1(t - y_a)\right] \tag{2.4a}$$

$$= -\tfrac{1}{2} i w_1(t - y_b) \times \left[w_2(t - y_a) - \frac{w_2'(t) - q w_2(t)}{w_1'(t) - q w_1(t)} w_1(t - y_a)\right]. \tag{2.4b}$$

Here, $w_2(t) = \pi^{\frac{1}{2}}[\mathrm{Bi}\,(t) + i\,\mathrm{Ai}\,(t)]$, $v(t) = \pi^{\frac{1}{2}}\,\mathrm{Ai}\,(t)$, $y_a = (2/ka)^{\frac{1}{3}} kh_a$, and $y_b = (2/ka)^{\frac{1}{3}} kh_b$. The identity

$$W(x, q, 0, 0) = W(x, q) \tag{2.5}$$

follows from the Wronskian relation

$$w_1'(t) w_2(t) - w_1(t) w_2'(t) = 2i. \tag{2.6}$$

The preceding is an encapsulated version of the known theory for diffraction by a homogeneous spherical (or cylindrical) boundary. Inherent in the formulation is that $d/a \ll 1$, $(ka) \gg 1$, and both h_a and $h_b \ll d$. However, some of these restrictions may be relaxed as indicated below.

[10] J. C. P. Miller, *The Airy Integral* (Cambridge University Press, Cambridge, England, 1946).

3. INTEGRAL EQUATION FOR THE INHOMOGENEOUS PROBLEM

When the surface impedance Z' is a function of distance α measured from B along the great-circle path towards A, the attenuation function is designated

$$W'(x, q'(\hat{x}), y_a, y_b),$$

where

$$q'(\hat{x}) = -i(\tfrac{1}{2}ka)^{\frac{1}{3}}(Z'/\eta_0). \tag{3.1}$$

An integral equation for W' follows from an application of Lorentz's reciprocal relation. The one-dimensional simplified form[9] is

$$W'(x, q'(\hat{x}), y_a, y_b) = W(x, q, y_a, y_b) + \left(\frac{x}{\pi i}\right)^{\frac{1}{2}} \int_0^x \frac{[q'(\hat{x}) - q]}{[\hat{x}(x - \hat{x})]^{\frac{1}{2}}} \times W(x - \hat{x}, q, y_a, 0) W'(\hat{x}, q'(\hat{x}), 0, y_b)\, d\hat{x}. \tag{3.2}$$

The relevant dimensionless parameters are indicated in Fig. 1(b). As indicated, the unknown attenuation function appears on the right-hand side in the integrand of the integral and on the left-hand side of Eq. (3.2). An explicit solution is readily obtained when the great-circle path is sectionally homogeneous. For example, we choose

$$\begin{aligned} q'(\hat{x}) &= q_1 \quad \text{for} \quad 0 < \hat{x} < x_1 \\ &= q_2 \quad \text{for} \quad x_1 < \hat{x} < x_2 \\ &= q \quad \text{for} \quad x_2 < \hat{x} < x. \end{aligned}$$

An equivalent statement, in terms of surface impedances, is

$$\begin{aligned} Z'(\alpha) &= Z_1 \quad \text{for} \quad 0 < \alpha < d_1 \\ &= Z_2 \quad \text{for} \quad d_1 < \alpha < d_2 \\ &= Z \quad \text{for} \quad d_2 < \alpha < d. \end{aligned}$$

Then, as discussed in detail elsewhere,[9]

$$W'(x, q', y_a, y_b) = W(x, q, y_a, y_b) + \left(\frac{x}{\pi i}\right)^{\frac{1}{2}} (q_1 - q) \times \int_0^{x_1} \frac{W(x - \hat{x}, q, y_a, 0) W(\hat{x}, q_1, 0, y_b)}{[\hat{x}(x - \hat{x})]^{\frac{1}{2}}}\, d\hat{x} + \left(\frac{x}{\pi i}\right)^{\frac{1}{2}} (q_2 - q) \times \int_{x_1}^{x_1 + x_2} \frac{W(x - \hat{x}, q, y_a, 0) W'(\hat{x}, q', 0, y_b)}{[\hat{x}(x - \hat{x})]^{\frac{1}{2}}}\, d\hat{x}, \tag{3.3}$$

where

$$W'(\hat{x}, q', 0, y_b) = W(\hat{x}, q_1, 0, y_b) + \left(\frac{\hat{x}}{\pi i}\right)^{\frac{1}{2}} (q_2 - q_1) \times \int_0^{\hat{x} - x_1} \frac{W(\hat{x} - x', q_1, 0, y_b) W(x', q_2)}{[(\hat{x} - x')x']^{\frac{1}{2}}}\, dx'. \tag{3.4}$$

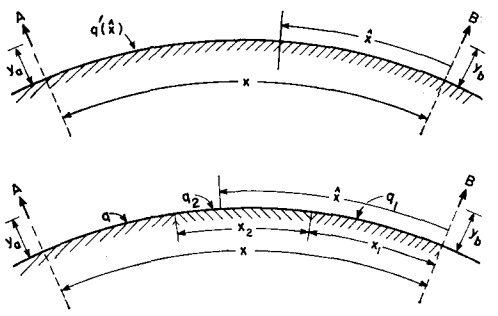


FIG. 2. (a) Cross-sectional view of the curved surface for an inhomogeneous path. (b) Cross-sectional view of the sectionally homogeneous surface indicating the meaning of the various parameters.

The meaning of the various quantities for the sectionally homogeneous case is illustrated in Figs. 2(a) and 2(b) for actual and dimensionless parameters, respectively. The results given by Eqs. (3.3) and (3.4) are a formal statement of the solution. Equivalent expressions have been given previously. Now we consider how to transform these integral formulas into a form which is suitable for numerical evaluation near the "line of sight."

4. SOLUTION FOR LARGE ANTENNA HEIGHTS

We now return to Eq. (2.3) and note that it may be decomposed as follows:

$$W(x, q, y_a, y_b) = W_f + W_g, \tag{4.1}$$

where

$$W_f = \frac{1}{2}\left(\frac{ix}{\pi}\right)^{\frac{1}{2}}\left[-\frac{i}{2}\int_{C_1} e^{-ixt} w_1(t - y_b) w_2(t - y_a)\, dt + \int_{C_2} e^{-ixt} w_1(t - y_b) v(t - y_a)\, dt\right], \tag{4.2}$$

$$W_g = -\frac{1}{2}\left(\frac{ix}{\pi}\right)^{\frac{1}{2}}\left[-\frac{i}{2}\int_{C_1} e^{-ixt}\frac{w_2'(t) - q w_2(t)}{w_1'(t) - q w_1(t)} \times w_1(t - y_a) w_1(t - y_b)\, dt + \int_{C_2} e^{-ixt} \times \frac{v'(t) - q v(t)}{w_1'(t) - q w_1(t)} w_1(t - y_a) w_1(t - y_b)\, dt\right]. \tag{4.3}$$

In these integrals, C_1 is the straight line from $\infty \exp[-i(\frac{2}{3}\pi)]$ to 0 and C_2 is the straight line from 0 to ∞.

By using the asymptotic approximation,

$$w_1(t) \cong e^{-\frac{1}{4}i\pi}(-t)^{-\frac{1}{4}} \exp\{-i[\tfrac{2}{3}(-t)^{\frac{3}{2}}]\} \tag{4.4}$$

valid for $-t \gg 1$, we easily show that

$$e^{-ixt} w_1(t - y_a) w_1(t - y_b) \cong -\frac{i}{(y_a y_b)^{\frac{1}{4}}} e^{-i\Omega_0} e^{-i\chi t} \times \left[1 - \frac{it^2}{4u^2} + O\left(\frac{1}{u^4}\right)\right], \tag{4.5}$$

where

$$\Omega_0 = \tfrac{2}{3} y_a^{\frac{3}{2}} + \tfrac{2}{3} y_b^{\frac{3}{2}},$$
$$\chi = x - y_a^{\frac{1}{2}} - y_b^{\frac{1}{2}},$$
$$u^2 = (y_a y_b)^{\frac{1}{2}}(y_a^{\frac{1}{2}} + y_b^{\frac{1}{2}})^{-1}.$$

This result is valid for both y_a and $y_b \gg 1$ for the important range of integration over t in W_g. Thus, we find that

$$W_g \cong -\frac{1}{2}\frac{x^{\frac{1}{2}}}{(y_a y_b)^{\frac{1}{4}}} e^{-i\Omega_0} \times \left[G(\chi) + \frac{i}{4u^2}\frac{\partial^2 G(\chi)}{\partial \chi^2} + O\left(\frac{1}{u^4}\right)\right], \tag{4.6}$$

where

$$G(\chi) = e^{-i\pi/4}\frac{1}{\pi^{\frac{1}{2}}}\int_0^\infty e^{-i\chi t}\frac{v'(t) - q v(t)}{w_1'(t) - q w_1(t)}\, dt + e^{i\pi/12}\frac{1}{\pi^{\frac{1}{2}}}\int_0^\infty e^{-(\frac{1}{2}\chi t)(3^{\frac{1}{2}} - i)} \times \frac{v'(t) - q e^{-i(\frac{2}{3}\pi)} v(t)}{w_2'(t) - q e^{-i(\frac{2}{3}\pi)} w_2(t)}\, dt. \tag{4.7}$$

In writing $G(\chi)$ in this form, we have made use of certain rotation formulas for the Airy functions [e.g., $w_1(te^{-i(\frac{2}{3}\pi)}) = e^{-\frac{1}{3}i\pi} w_2(t)$ and $w_2(te^{-i(\frac{2}{3}\pi)}) = 2e^{+i\pi/6} v(t)$]. Furthermore, in what follows, we assume that $u^2 \gg 1$ and, thus, only the first term in the square bracket in Eq. (4.6) need be retained.

In expressing W as a sum of two parts as indicated by Eq. (4.1), it is clearly evident that the first term W_f is independent of q which characterizes the properties of the diffracting surface. This suggests that, for the inhomogeneous surface, we write

$$W'(x, q', y_a, y_b) = W_f + W_g', \tag{4.8}$$

where W_f is defined by Eq. (4.2) and W_g' is analogous to Eq. (4.3) but now must involve the properties of the inhomogeneous surface. Taking a cue from Eq. (4.6), we define a function $G'(\chi)$ as follows:

$$W_g' = -\tfrac{1}{2}[x^{\frac{1}{2}}/(y_a y_b)^{\frac{1}{4}}] e^{-i\Omega_0} G'(\chi). \tag{4.9}$$

5. SIMPLIFICATIONS OF THE PATTERN FUNCTION

In order to make use of the integral formulas given by Eqs. (3.3) and (3.4), we note the attenuation function W in the integrands may be approximated asymptotically in a manner analogous to Eq. (4.6). First of all, we note from Eq. (2.3) that

$$W(x, q, 0, y) = W(x, q, y, 0) = \frac{1}{2}\left(\frac{ix}{\pi}\right)^{\frac{1}{2}}\int_\Gamma \frac{e^{-ixt} w_1(t - y)}{w_1'(t) - q w_1(t)}\, dt. \tag{5.1}$$

Now, if y is a sufficiently large positive number, the function $w_1(t - y)$ may be replaced by the first term

of its asymptotic series. This leads to the useful approximation

$$W(x, q, 0, y) \cong \exp\left[-i(\tfrac{2}{3}y^{\frac{3}{2}})\right] V(x - y^{\frac{1}{2}}, q) x^{\frac{1}{4}} y^{-\frac{1}{4}}, \tag{5.2}$$

where

$$V(X, q) = \frac{1}{2\pi^{\frac{1}{2}}} \int_\Gamma \frac{e^{-iXt}}{w_1'(t) - q w_1(t)}\, dt \tag{5.3}$$

may be recognized as the pattern function of "cut-back" factor[11] for a (vertically polarized) source on a curved surface.

A mixed-path pattern function V' may be defined in a manner analogous to Eq. (5.2). For example, if $y_b \gg 1$, we are permitted to replace Eq. (3.4) by the approximate form

$$W'(\hat{x}, q', 0, y_b) \cong \exp\left[-i(\tfrac{2}{3}y_b^{\frac{3}{2}})\right] \times V'(\hat{x} - y_b^{\frac{1}{2}}, q_1, q_2) \hat{x}^{\frac{1}{4}} y_b^{-\frac{1}{4}}, \tag{5.4}$$

where

$$V'(\hat{x} - y_b^{\frac{1}{2}}, q_1, q_2) = V(\hat{x} - y_b^{\frac{1}{2}}, q_1) + \frac{q_2 - q_1}{(\pi i)^{\frac{1}{2}}} \times \int_0^{\hat{x} - x_1} \frac{V(\hat{x} - y_b^{\frac{1}{2}} - x', q_1) W(x', q_2)}{(x')^{\frac{1}{2}}}\, dx'. \tag{5.5}$$

Here we have explicitly indicated the functional dependence of V' on q_1 and q_2 to stress that it is not a function of q.

6. INTEGRAL FORMULA FOR $G'(\chi)$

By using the preceding asymptotic representations for the various attenuation functions and the basic integral formula Eq. (3.3), it is a straightforward matter to show that

$$G'(\chi) \cong G(\chi) - \frac{2(q_1 - q)}{(\pi i)^{\frac{1}{2}}} \int_0^{x_1} V(x - \hat{x} - y_a^{\frac{1}{2}}, q) \times V(\hat{x} - y_b^{\frac{1}{2}}, q_1)\, d\hat{x} - \frac{2(q_2 - q)}{(\pi i)^{\frac{1}{2}}} \times \int_{x_1}^{x_1 + x_2} V(x - \hat{x} - y_a^{\frac{1}{2}}, q) V'(\hat{x} - y_b^{\frac{1}{2}}, q_1, q_2)\, d\hat{x}. \tag{6.1}$$

It should be mentioned that in writing the integral, with limits 0 to x_1, in Eq. (6.1), a certain inaccuracy has been permitted. For example, strictly speaking, the function $W(\hat{x}, q_1, 0, y_b)$ is not well approximated by the form Eq. (5.2) near the lower limit of the integration. However, for the purpose intended, this factor is of minor concern as the contribution from this part of the integrand is small.

7. FRESNEL DIFFRACTION PATTERN

Using Eq. (6.1), the contribution W_g' as defined by Eq. (4.9) may then be computed in terms of the pattern functions for a homogeneous surface. The total field requires adding this to W_f as indicated by Eq. (4.8). As mentioned before, the latter function depends only on the geometry of the diffracting surface. In an illuminating development by Fock,[3] it was shown that

$$W_f \cong \frac{x^{\frac{1}{2}}}{2(y_a y_b)^{\frac{1}{4}}} e^{-i\Omega_0} u f(u\chi) \quad \text{for} \quad \chi > 0$$
$$\cong \tfrac{1}{2} e^{-i\Omega(x)} - \frac{x^{\frac{1}{2}}}{2(y_a y_b)^{\frac{1}{4}}} e^{-i\Omega_0} u f(-u\chi) \quad \text{for} \quad \chi < 0, \tag{7.1}$$

where

$$f(\alpha) = \frac{1}{\pi^{\frac{1}{2}}} \exp\left[+i\left(\alpha^2 + \frac{\pi}{4}\right)\right] \int_\alpha^\infty \exp(-iz^2)\, dz, \tag{7.2}$$

$$\Omega(x) = -\tfrac{1}{12}x^3 + \tfrac{1}{2}x(y_a + y_b) + [(y_a - y_b)^2/4x]. \tag{7.3}$$

This approximate representation for W_f is valid if, as usual, both y_a and y_b are very large. On the other hand, χ may be regarded as finite or small. Consequently, $u\chi$ may be arbitrary with either sign. Actually, if χ is very small (i.e., near grazing) both expressions for W_f above nearly coincide. This is a consequence of the approximate equations

$$u^2 x/(y_a y_b)^{\frac{1}{2}} = 1 + [\chi/(y_a^{\frac{1}{2}} + y_b^{\frac{1}{2}})] \cong 1$$

and

$$\Omega(x) \cong \Omega_0 - u^2\chi^2.$$

8. TOTAL DIFFRACTION PATTERN

Using the approximations valid for small χ, we can combine Eqs. (7.1) and (4.9) and, thus, Eq. (4.8) is written

$$2W' \cong e^{-i\Omega_0}[f(u\chi) - G'(\chi)/u], \tag{8.1}$$

which is our final result. For the conditions stated, the term on the right is the modification of the free-space field of the source A at the receiver's location B. We see immediately that if u is sufficiently large only the first term in the square bracket survives. This limit corresponds with the predictions of physical optics. In this case, we note that the argument $u\chi$ of the Fresnel integral may be expressed as

$$\alpha = u\chi \cong [2ks_a s_b/(s_a + s_b)]^{\frac{1}{2}}(\tfrac{1}{2}\theta), \tag{8.2}$$

where $s_a = (2ah_a)^{\frac{1}{2}}$ and $s_b = (2ah_b)^{\frac{1}{2}}$ are the distance from the terminals A and B to their horizons on the diffracting surface. As illustrated in Fig. 3, θ is the deflection angle which may be positive or negative.

As indicated by Eq. (8.1), $G'(\chi)/u$ is the influence of the curvature and electrical characteristics of the diffracting crest. For sufficiently rounded edges, this term may be neglected as mentioned above. A previous detailed analytical and numerical study of the integral

[11] J. R. Wait and A. M. Conda, IRE Trans. Antennas Propagation, AP-6, No. 4, 348 (1958).

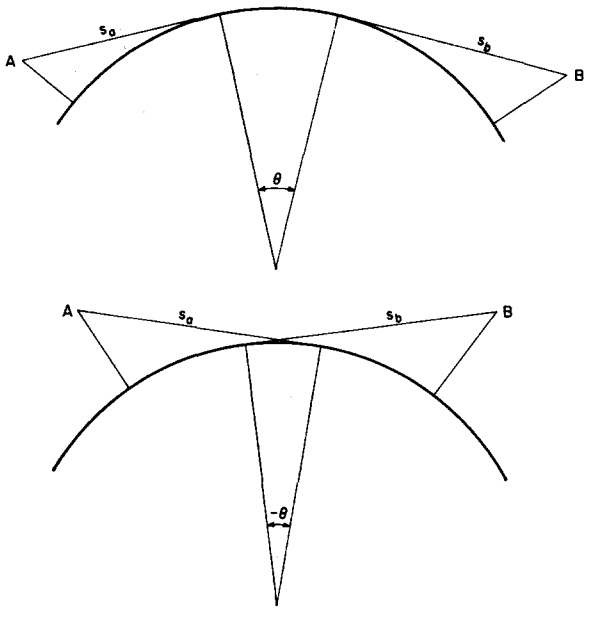


FIG. 3. The important geometrical parameters for near grazing conditions.

$G(\chi)$ for a homogeneous finitely conducting curved surface has already been carried out.[5]

It is important to note that if the parameters u and χ are defined in terms of the length of the tangents s_a and s_b, it is not necessary to require that the antenna heights h_a and h_b are small compared with a which is the radius of curvature of the diffracting edge. Specifically, in this case, we should use the definitions

$$u = \left(\frac{2ks_as_b}{s_a + s_b}\right)^{\frac{1}{2}} \frac{(2/ka)^{\frac{1}{3}}}{2} \quad \text{and} \quad \chi = \left(\frac{ka}{2}\right)^{\frac{1}{3}}\theta. \tag{8.3}$$

Also, as may be verified by comparing the results with a cylindrical model, the results for both the sphere and the cylinder are essentially the same for near grazing conditions.

9. DISCUSSION OF A SPECIAL CASE

In order to give some physical insight into the nature of the present situation, we consider a rather extreme special case. The diffracting surface is imagined to be perfectly conducting except for a small transverse strip at or very near the crest. Also, we choose $\theta = 0$ corresponding to the situation where the line of sight between A and B just grazes the crest. For the conditions as stated, $q = q_1 = 0$ and $\chi = 0$. Then, the square bracket term in Eq. (8.1) has the form

$$[f(0) - G'(0)/u] = \tfrac{1}{2} - G(0)/u + \Delta, \tag{9.1}$$

where

$$\Delta = \frac{2q_2}{u(\pi i)^{\frac{1}{2}}} \int_{x_1}^{x_1+x_2} V(x - \hat{x} - y_a^{\frac{1}{2}}, 0) \times V'(\hat{x} - y_b^{\frac{1}{2}}, 0, q_2)\, d\hat{x}. \tag{9.2}$$

For the case considered, $G(0) = -0.295 + i0.0811$ which is the value adopted from a previous study.[5] In Eq. (9.2), the quantity Δ represents the influence of the inhomogeneity on the diffracting surface. Even in this rather restrictive special case it appears that the integral over $\hat{x}$ would need to be evaluated numerically if the inhomogeneity is strong. However, for a weak perturbation (i.e., sufficiently small values of q_2 or x_2), we may obtain a first-order estimate very simply by noting that the integrand is a slowly varying function if the strip is near the crest. Thus, to within a first order,

$$\Delta \cong [2q_2x_2/u(\pi i)^{\frac{1}{2}}]V(x - \bar{x} - y_a^{\frac{1}{2}}, 0)V(\bar{x} - y_b^{\frac{1}{2}}, 0), \tag{9.3}$$

where $\bar{x} = x_1 + \frac{1}{2}x_2$. If, in fact, the strip is right at the crest $x - \bar{x} \cong y_a^{\frac{1}{2}}$ and $\bar{x} \cong y_b^{\frac{1}{2}}$. Then,

$$\Delta \cong [2q_2x_2/u(\pi i)^{\frac{1}{2}}][V(0, 0)]^2 \cong -\frac{2(i/\pi)^{\frac{1}{2}}(Z_2/\eta_0)kd_2[V(0, 0)]^2}{[2ks_as_b/(s_a + s_b)]^{\frac{1}{2}}}. \tag{9.4}$$

This shows that the influence of the inhomogeneity at the crest does not, to within a first order depend on the curvature of the crest. Not too surprisingly, it has the nature of a cylindrical wave emanating from the crest with an amplitude proportional to the transverse width d_2 of the strip and the impedance contrast Z_2. Its strength is modified by the square of the pattern function V evaluated at the tangent point on the crest. From earlier work,[11] we note that

$$[V(0, 0)]^2 = 0.490.$$

It is particularly interesting to note that the expression for Δ may be deduced from a perturbation analysis for scattering from an impedance strip on a conducting plane with the important exception that the function $[V(0, 0)]^2$ is replaced by unity. The derivation is outlined very briefly in the Appendix.

It is evident from Eq. (9.4) that the contribution from Δ represents a diminutation of the grazing field as the real part of $Z_2 \exp(\frac{1}{4}i\pi)$ has a positive real part in a dissipative dielectric medium.

10. CONCLUDING REMARKS

The results given in this paper should be useful for making estimates of the fields diffracted by smooth inhomogeneous curved surfaces. While the problem has been formulated in terms of radially oriented electric dipoles over a spherical obstacle, it is possible to apply the results to cylindrical bodies if the grazing angles are sufficiently small. Also, while the antenna

heights were assumed to be small compared with the radius of curvature of the diffracting surface, this is not an essential restriction, as indicated in the text.

Finally, it might be mentioned that the present results are readily generalized to horizontal polarization. For example, if the radial electric dipoles are replaced by radial magnetic dipoles, the general formulas are unchanged if the normalized surface impedances Z_i/η_0 are replaced everywhere by normalized surface admittances $Y_i\eta_0$.

APPENDIX. SCATTERING FROM AN IMPEDANCE STRIP

As indicated in Fig. 4, we consider a magnetic line source at M which is parallel to the impedance strip of width d_2 on an otherwise perfectly conducting plane. The magnetic field scattered from the strip has only a z component H_z^s. An appropriate transform representation is

$$H_z^s(x, y) = \int_{-\infty}^{+\infty} \exp\left[-(\lambda^2 - k^2)^{\frac{1}{2}} y\right] A(\lambda) \exp(i\lambda x)\, d\lambda, \tag{A1}$$

where

$$(A\lambda) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} H_z^s(\alpha, 0) \exp(-i\lambda\alpha)\, d\alpha. \tag{A2}$$

The λ integration may be carried out to give[12]

$$H_z^s \cong \frac{\pi^{\frac{1}{2}}}{(2ik\rho)^{\frac{1}{2}}} \frac{iky}{\pi\rho} \int_{-\infty}^{+\infty} e^{ik\rho\cos\phi} H_z^s(\alpha, 0)\, d\alpha e^{-ik\rho}, \tag{A3}$$

which is valid for $|k\rho| \gg 1$. Now, on the surface $y = 0$, for $|\alpha| < \frac{1}{2}d_2$, we may choose

$$H_z^s(\alpha, 0) \cong \frac{e^{-ik\rho_0}}{(k\rho_0)^{\frac{1}{2}}} (R_\parallel - 1) e^{ik\alpha\cos\phi_0}, \tag{A4}$$

where $e^{-jk\rho_0}(k\rho_0)^{-\frac{1}{2}}$ is the normalized strength of the incident wave at the center of the strip, and $R_\parallel$ is the appropriate reflection coefficient which, at the center of the strip, is given by

$$R_\parallel \cong \frac{\sin\phi_0 - Z_2/\eta_0}{\sin\phi_0 + Z_2/\eta_0}, \tag{A5}$$

[12] J. R. Wait, Can. J. Phys. 33, 383 (1955).

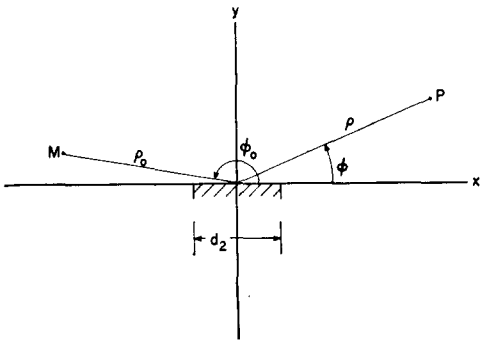


FIG. 4. Cross-sectional view of impedance strip on flat perfectly conducting plane.

where Z_2 is the surface impedance. For present purposes, we may assume that $R_\parallel$ is constant over the width of the strip, although this is not strictly correct.

The integration over α in Eq. (A3) may now be carried out to give

$$\frac{H_z^s}{H_z^p} \cong \frac{-2(i/\pi)^{\frac{1}{2}} kd_2(Z_2/\eta_0)}{[2k\rho\rho_0/(\rho + \rho_0)]^{\frac{1}{2}}} T(\phi), \tag{A6}$$

where

$$T(\phi) = \frac{\sin\phi}{\sin\phi_0 + (Z_2/\eta_0)} \frac{\sin[(\frac{1}{2}kd_2)(\cos\phi + \cos\phi_0)]}{\frac{1}{2}(kd_2)(\cos\phi + \cos\phi_0)}, \tag{A7}$$

$$H_z^p = \exp[-ik(\rho + \rho_0)]/[k(\rho + \rho_0)]^{\frac{1}{2}}. \tag{A8}$$

For the small perturbation case being considered here, Z_2/η_0 may be neglected in the denominator in the expression for $T(\phi)$. Then, as we are interested in the case where both $\pi - \phi_0$ and ϕ are small angles, the quantity $T(\phi)$ is of the order of unity. Then, if we recognize that H_z^p is the strength of the primary field at P, we see that the ratio H_z^s/H_z^p bears a remarkable similarity to Δ as given by Eq. (9.4). [Of course, ρ_0 and ρ are to be identified with s_a and s_b, respectively.] We do not believe that any closer equivalence should be expected in view of the numerous approximations made in treating this idealized special case.

Evaluation of Expressions of the Debye–Waller Form

R. H. TERWIEL
Lorentz Institute of Theoretical Physics, University of Leiden, Leiden, The Netherlands

(Received 19 September 1966)

A very simple evaluation of expressions of the Debye–Waller form is given.

RECENTLY Mermin[1] has given a method of evaluating averages of the Debye–Waller form,

$$\langle e^0\rangle \equiv \operatorname{Tr} e^{-\beta H}e^0/\operatorname{Tr} e^{-\beta H},\quad \beta = 1/k_BT,$$
$$H = \sum \omega_i(a_i^\dagger a_i + \tfrac{1}{2}),\quad 0 = \sum c_i a_i + d_i a_i^\dagger,$$
$$[a_i, a_j^\dagger] = \delta_{ij}, \tag{1}$$

which is considerably simpler than the methods customarily employed. In this note we would like to present an alternative method of evaluation which, in our opinion, is certainly as direct as the one employed by Mermin and which avoids the use of the Baker–Hausdorff theorem[2] and an iteration (induction) procedure followed by resummation of an infinite series. In fact, apart from the cyclic property of the trace one only needs the relations

$$e^{-\beta H}a_i e^{\beta H} = e^{\beta\omega_i}a_i, \tag{2a}$$

$$e^{\lambda 0}a_i e^{-\lambda 0} = a_i - \lambda d_i, \tag{2b}$$

which can both be derived in an elementary way.

Defining

$$F(\lambda) \equiv \langle e^{\lambda 0}\rangle \tag{3}$$

we have

$$dF/d\lambda = \sum f_i \tag{4}$$

with

$$f_i \equiv \langle c_i a_i e^{\lambda 0}\rangle + \langle d_i a_i^\dagger e^{\lambda 0}\rangle \tag{5a}$$
$$= \tfrac{1}{2}\{\langle c_i a_i e^{\lambda 0}\rangle + \langle e^{\lambda 0}c_i a_i\rangle + \langle d_i a_i^\dagger e^{\lambda 0}\rangle + \langle e^{\lambda 0}d_i a_i^\dagger\rangle\}. \tag{5b}$$

Now, since

$$\operatorname{Tr} e^{-\beta H}c_i a_i e^{\lambda 0} = e^{\beta\omega_i}\operatorname{Tr} e^{-\beta H}e^{\lambda 0}c_i a_i,$$

as follows by inserting into the left-hand side the unit operator $e^{\beta H}e^{-\beta H}$ after a_i and applying (2a), we have

$$\langle c_i a_i e^{\lambda 0}\rangle = e^{\beta\omega_i}\langle e^{\lambda 0}c_i a_i\rangle$$

and similarly

$$\langle e^{\lambda 0}d_i a_i^\dagger\rangle = e^{\beta\omega_i}\langle d_i a_i^\dagger e^{\lambda 0}\rangle,$$

so that (5b) can be written in the form

$$f_i = \tfrac{1}{2}(e^{\beta\omega_i} + 1)\{\langle e^{\lambda 0}c_i a_i\rangle + \langle d_i a_i^\dagger e^{\lambda 0}\rangle\}. \tag{6}$$

Since, according to (2b),

$$\langle e^{\lambda 0}c_i a_i\rangle = \langle c_i a_i e^{\lambda 0}\rangle - \lambda c_i d_i\langle e^{\lambda 0}\rangle, \tag{7}$$

we obtain from (6), using the definitions (3) and (5a),

$$f_i = \tfrac{1}{2}(e^{\beta\omega_i} + 1)\{f_i - \lambda c_i d_i F\}, \tag{8}$$

or, according to (4),

$$dF/d\lambda = \lambda(\sum c_i d_i \coth \tfrac{1}{2}\beta\omega_i)F, \tag{9}$$

leading to

$$F(\lambda) = \exp(\tfrac{1}{2}\lambda^2 \sum c_i d_i \coth \tfrac{1}{2}\beta\omega_i). \tag{10}$$

[1] N. D. Mermin, J. Math. Phys. 7, 1038 (1966).
[2] Mermin's formula (2).

Integral Equations in the Theory of Classical Fluids

M. S. WERTHEIM
University of California, Los Alamos Scientific Laboratory, Los Alamos, New Mexico, and Laboratoire de Physique Théorique et Hautes Energies, Orsay (S.-et O.), France

(Received 7 July 1966)

Topological analysis of Mayer diagrams is used to obtain generalizations of the Percus–Yevick and convolution hypernetted chain integral equations to systems of equations involving higher correlation functions. Only the higher direct correlation functions appear in the final integral equations. A distinctive feature of the system obtained is the appearance of derivatives with respect to the density. Numerical comparison with exact virial coefficients is made for the standard test potentials of parallel hard squares and cubes.

I. INTRODUCTION

In recent years considerable progress has been made in the problem of a classical system of particles interacting by a pair potential $V(\mathbf{r}_{ij})$. The most fruitful approach has been the solution of approximate integral equations for the pair distribution function $g_2(\mathbf{r}_1, \mathbf{r}_2)$ These integral equations have been derived in many different ways. Thus the Percus–Yevick (PY) equation was originally derived by Percus and Yevick[1] by the method of collective coordinates. Later a diagrammatic derivation was given by Stell,[2] and a derivation by functional differentiation by Percus.[3] The Yvon–Born–Green (BGY) integral equation[4] was obtained by introducing the superposition approximation into an exact relation between the two particle distribution function $g_2(\mathbf{r}_1, \mathbf{r}_2)$ and the three particle distribution function $g_3(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$. The convolution hypernetted chain (CHNC) integral equation[5] was first obtained by diagrammatic techniques.

A unified method of deriving all the various approximate integral integrations for the pair distribution function $g_2(\mathbf{r}_{12})$ was introduced by Percus.[3] His method of functional Taylor series also provided a natural way of generalizing to systems of integral equations involving the higher n-particle distribution functions g_n. Two slightly different generalizations of the PY and CHNC equations to a system of integral equations involving g_3 have actually been derived by Verlet,[6] who also calculated the fifth and sixth virial coefficients for the standard test cases of parallel hard squares and cubes, for which the first seven virial coefficients are known exactly.[7]

One defect of the g_3 obtained by Verlet is the fact that it is symmetric in only two of the three particles 1, 2, 3. This fact as well as the non-uniqueness of the generalization from the PY and CHNC to the PY 2 and CHNC 2 equations suggested a study of alternative methods of generalizing these integral equations. An alternative to the method of functional differentiation is the graph-theoretical analysis of Mayer diagrams.[8]

The natural way in which the direct correlation function[9] $C_2(\mathbf{r}_1, \mathbf{r}_2)$ enters the PY and CHNC integral equations suggests formulating the generalization in terms of both the higher $g_n(\mathbf{r}_1, \mathbf{r}_2 \cdots \mathbf{r}_n)$ and the higher $C_n(\mathbf{r}_1, \mathbf{r}_2 \cdots \mathbf{r}_n)$.[10] Each C_n is defined in terms of g_n and the $C_{n'}$ or $g_{n'}$ for $n' < n$ by a convolution relation obtained by functional differentiation of the Ornstein–Zernike definition

$$u(\mathbf{r}_1, \mathbf{r}_2) - C_2(\mathbf{r}_1, \mathbf{r}_2) = \rho \int d\mathbf{r}_3 u(\mathbf{r}_1, \mathbf{r}_3) C_2(\mathbf{r}_3, \mathbf{r}_2), \quad (1)$$

where $u(\mathbf{r}_1, \mathbf{r}_2) \equiv g_2(\mathbf{r}_1, \mathbf{r}_2) - 1$, and ρ is the number density.

Several different expansions of C_n in terms of all the $g_{n'}$, $n' = 2, 3, \cdots, \infty$ may be derived.[11] However, in all cases simultaneous truncation of these expansions combined with the defining relations for the C_n leads to an inconsistency which manifests itself by higher orders in ρ of the C_n and g_n not being expressible in terms of Mayer diagrams. This suggested using the exact

[1] J. K. Percus and G. J. Yevick, Phys. Rev. **110**, 1 (1958).
[2] G. Stell, Physica **29**, 517 (1963).
[3] J. K. Percus, Phys. Rev. Letters **8**, 462 (1962).
[4] J. Yvon, *Actualités scientifiques et industrielles* (Hermann & Cie., Paris, 1935), No. 203; M. Born and H. S. Green, Proc. Roy. Soc. (London) **A188**, 10 (1946).
[5] J. M. Van Leeuwen, J. Groeneveld, and J. deBoer, Physica **25**, 792 (1959); M. S. Green, Hughes Aircraft Company Report (1959); E. Meeron, J. Math. Phys. **1**, 192 (1960); T. Morita and K. Hiroike, Progr. Theoret. Phys. (Kyoto) **23**, 1003 (1960); G. S. Rushbrooke, Physica **26**, 259 (1960); L. Verlet, Nuovo Cimento **18**, 77 (1960).
[6] L. Verlet, Physica **30**, 95 (1964); **31**, 965 (1965).
[7] W. G. Hoover and A. G. De Rocco, J. Chem. Phys. **36**, 3141 (1962).
[8] H. D. Ursell, Proc. Cambridge Phil. Soc. **23**, 685 (1927); J. E. Mayer and M. G. Mayer, *Statistical Mechanics* (John Wiley & Sons, Inc., New York, 1940).
[9] L. S. Ornstein and F. Zernike, Proc. Acad. Sci. Amsterdam **17**, 793 (1914).
[10] R. J. Baxter, J. Chem. Phys. **41**, 553 (1964).
[11] M. S. Wertheim, unpublished.

relation between successive C_n[10]

$$\frac{\partial}{\partial \rho} C_n(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_n) = \int d\mathbf{r}_{n+1} C_{n+1}(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_{n+1}) \tag{2}$$

combined with a single closure relation for the highest C_n retained.

As a result the system of equations obtained is complicated by the appearance of derivatives with respect to the density. There are, however, several redeeming features. The most important is the explicit elimination of all the g_n for $n > 2$. Another one is the fact that all the C_n and g_n are automatically symmetric in all n particles.

II. DEFINITIONS AND ELEMENTARY THEOREMS

This section contains some well-known results necessary to our subsequent analysis. Further results, more details, and references to the many original papers may be found in a review article by Stell.[12] In most instances we also follow Stell in terminology.

A diagram consists of a set of points or vertices, representing particles and a set of lines or bonds connecting pairs of particles. The bond represents a function to be specified of the space coordinates of the two particles connected by the bond. Thus an e-bond connecting particles 1 and 2, indicated by a dashed line, means $e(\mathbf{r}_1, \mathbf{r}_2) = \exp[-\beta V(\mathbf{r}_1, \mathbf{r}_2)]$. Here $\beta = [kT]^{-1}$ where T is the Kelvin temperature and k is Boltzmann's constant. In all that follows, particle coordinates $\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_n$ are abbreviated by writing $12 \cdots n$. The f-bond, indicated by a solid line, stands for $f(ij) \equiv e(ij) - 1$. Other bonds are introduced later. A particle is represented by a black circle if there is an integration over the coordinates of the particles. If there is no such integration the particle is represented by a white circle. A diagram is called connected if there is at least one path of bonds from any circle to any other circle. A diagram is called irreducible if it is connected and there is no circle the removal of which causes the diagram to break into two or more fragments. Such a circle is called an articulation circle (AC) if at least one fragment contains no white circle; it is called a cutting circle (CC) if no fragment is free of white circles.

Any quantity which can be written as a set of diagrams of f-bonds can of course be rewritten as a set of diagrams of e-bonds or of a mixture of f-bonds and e-bonds by using $e = f + 1$, where 1 denotes simply the absence of a bond. We refer to a physical quantity as f-irreducible (e-irreducible) if it consists only of irreducible diagrams when written in terms of f-bonds (e-bonds). We refer to a physical quantity as e-f-irreducible if it consists entirely of irreducible diagrams when written using only e-bonds as direct bonds connecting pairs of white circles, f-bonds between black–black and black–white pairs. Diagrams which are not irreducible are called reducible. An irreducible diagram in which there is a direct bond connecting every pair of vertices is called a complete star.

The diagrammatic expression for the excess Helmholtz free energy ΔA of a gas was first derived by Born and Fuchs,[13] who obtained

$$-\beta \Delta A = \sum_{n=2}^{\infty} \frac{\rho^{n-1}}{n!} \sum_i W(S_{in}) S_{in}. \tag{3}$$

S_{in} stands for an irreducible diagram of n black circles. The weight $W(S_{in})$ is equal to the number of distinct ways the same diagram can be drawn using the same n vertices but different f-bonds. The summation on i runs over all irreducible diagrams of n black circles.

The cluster expansions of the ordinary n-particle distribution functions $g_n(12 \cdots n)$ and the n-particle direct correlation functions $C_n(12 \cdots n)$ are obtained from the expansion for $-\beta \Delta A$ in the following way. To obtain the expansion of $\rho^{n-1} g_n(12 \cdots n)$, place n white circles on the vertices of a complete star of n vertices and change all the $\frac{1}{2}n(n-1)$ f-bonds in the complete star to e-bonds. The set of diagrams obtained by placing the n white circles in all possible ways is the expansion of $g_n(12 \cdots n)$. Diagrams in $\beta \Delta A$ that do not contain a complete star of n vertices do not contribute to $g_n(12 \cdots n)$. The expansion of $\rho^{n-1} C_n(12 \cdots n)$ is obtained by simply placing n white circles on any n vertices in all possible ways.

From this procedure it is clear that all the diagrams in C_n are f-irreducible. Similarly all the diagrams in g_n are e-f-irreducible, but not necessarily f-irreducible when all the e-bonds are expanded as $e = f + 1$, 1 denoting simply absence of a bond. An important result, central to all our subsequent topological arguments, is the following: C_n is the f-irreducible part of g_n. Because of its importance a simple proof is given.

Let us formally rewrite the expansion of $-\beta \Delta A$ in terms of e-bonds, using $f = e - 1$. In the order consisting of m-particle terms we obtain e-reducible diagrams and a single e-irreducible diagram, the complete star of m vertices. The latter fact is seen

[12] G. Stell, in *The Equilibrium Theory of Classical Fluids*, H. L. Frisch and J. L. Lebowitz, Eds. (W. A. Benjamin, Inc., New York, 1964).

[13] M. Born and K. Fuchs, Proc. Roy. Soc. (London) **A166**, 391 (1938).

most easily from the grand canonical partition function.

The prescription for obtaining $\rho^{n-1}g_n$ becomes the following: place n white circles on a complete star of n vertices in all possible ways. The prescription for obtaining $\rho^{n-1}C_n$ is unchanged: place n white circles on any n vertices. In terms of f-bonds, all f-irreducible terms originate in the complete star of e-bonds; the other terms merely generate terms which cancel the f-reducible terms in the complete star. However, in the complete star of e-bonds the prescriptions for obtaining g_n and C_n are identical. Hence g_n and C_n differ only by f-reducible terms.

The prescription given for obtaining g_n and C_n may be combined with Eq. (3) into the following formula:

$$\xi_n(12\cdots n) = \sum_{m=0}^{\infty} \frac{\rho^m}{m!} \sum_i W(S^{(\xi)}_{im,n}) S^{(\xi)}_{im,n}, \tag{4}$$

where ξ is C or g, m is the number of black circles in the diagram, the weight $W(S^{(\xi)}_{im,n})$ is the number of ways of drawing the diagram $S^{(\xi)}_{im,n}$ containing m black and n white circles labeled $1, 2, \cdots, n$. The sum over i runs over all $S^{(\xi)}_{im,n}$ fulfilling the following conditions: for $\xi = g$, it is e-f-irreducible and contains a complete star of n white circles connected by direct e-bonds. For $\xi = C$, it is f-irreducible.

Finally we need two simple combinatorial results. Let us suppose that a diagram S containing m black circles can be broken into several pieces S_i containing m_i black circles by simultaneously removing two or more white circles. Suppose that no two of the fragments contain the same set of white circles. Then each fragment can be labeled by the set of white circles it contains, and the weight $W(S)$ becomes simply the product of the weight of partitioning the m black circles into sets of m_i particles (with $\sum m_i = m$) and the weights $W(S_i)$ of the fragments. Thus

$$W(S) = \left[m! / \prod_i (m_i) \right] \prod_i W(S_i)$$

or

$$\frac{W(S)}{m!} = \prod_i \frac{W(S_i)}{m_i!} \tag{5}$$

This means that the coefficient of S is simply the product of the coefficients of the S_i.

Now consider the case in which all white circles are removed simultaneously, and each fragment is connected to every white circle. In this case the fragments cannot be labeled by reference to white circles, and there may be several topologically equivalent fragments. Suppose that there are ν_i fragments of type i, with weight $W(S_i)$ and containing m_i black circles. The result is

$$\frac{W(S)}{m!} = \prod_i \left[\frac{W(S_i)}{m_i!} \right]^{\nu_i} \frac{1}{\nu_i!}. \tag{6}$$

III. THE PY AND CHNC INTEGRAL EQUATIONS

We begin by sketching a diagrammatic derivation of the PY and CHNC equations which serve as a model for generalization to a system involving the higher correlation functions. We want to combine the definition of $C_2(12)$, Eq. (1), with a closure relation.

To accomplish this, we note that diagrams in $g_2(12)$ contain the $e(12)$ bond; when it is destroyed there results a term equal to 1 from the leading term in $g_2(12)$, as well as a set of connected diagrams with two white circles. We denote the subset of all the f-irreducible diagrams in $Q_2(12) = g_2(12)/e(12)$ as $Z_2(12)$, the complementary subset of f-reducible diagrams as $m_2(12)$. Hence

$$g_2(12) = e(12)[1 + m_2(12) + Z_2(12)]. \tag{7}$$

$C_2(12)$ is the f-irreducible part of $g_2(12)$; hence

$$C_2(12) = f(12)[1 + m_2(12)] + e(12)Z_2(12). \tag{8}$$

The PY approximation is obtained by setting $Z_2(12)$ equal to zero.

$$Z_2^{\mathrm{PY}}(12) = 0. \tag{9}$$

The CHNC approximation results by approximating $Z_2(12)$ by the following subset of the diagrams contained in the true $Z_2(12)$: the sum of all diagrams which, by cutting at both white circles simultaneously, are broken into two or more reducible fragments. Each fragment (including the two white circles) is a diagram in $m_2(12)$. Let there be ν_i fragments of type S_{im} containing m_i black circles and having weight $W(S_{im})$ with

$$\sum_i \nu_i m_i = m,$$

the total number of black circles and

$$\sum_i \nu_i = \nu,$$

the total number of fragments. Then the total coefficient of the diagram in $Z_2(12)$ is given by Eq. (6). The set of all such diagrams with a given value of ν is identical with the set of all diagrams in $m_2^{\nu}(12)$. A simple combinatorial argument shows that in $m_2^{\nu}(12)$ the diagram occurs with a total coefficient

$$\prod_i [W(S_i)/m_i!]^{\nu_i}.$$

Hence

$$Z_2(12)^{\mathrm{CHNC}} = \sum_{\nu=2}^{\infty} m_2^{\nu}(12)/\nu!$$

or

$$Z_2(12)^{\mathrm{CHNC}} = e^{m_2(12)} - m_2(12) - 1, \qquad (10)$$

which constitutes the CHNC closure.

IV. THE PY II AND CHNC II SYSTEMS

We now construct an analogous set of relations for $g_3(123)$ and $C_3(123)$. Let us write $g_3(123)$ in terms of f-bonds and then remove the two white circles 1 and 2. This splits each diagram into two fragments, one attached to all three white circles, and one attached only to circles 1 and 2. The latter fragment may be absent—denoted by 1—or it may be any connected diagram with two white circles, hence a diagram in $u(12)$. Since here—and in all subsequent instances in our analysis of the PY II system—the sets of white circles are not the same for the two fragments, we have the simple case of a product. Thus we obtain a factor $g_2(12) = 1 + u(12)$ and a factor containing diagrams with three white circles. Performing the same operation at the pairs of circles 13 and 23, we obtain factors $g_2(13)$ and $g_2(23)$. The three g-bonds suffice to make the diagram e–f-irreducible. Hence

$$g_3(123)/g_2(12)g_2(13)g_2(23)$$

consists of a term with no bonds and a term $Q_3(123)$ containing diagrams attached to all three white circles.

$$g_3(123) = g_2(12)g_2(13)g_2(23)[1 + Q_3(123)]. \qquad (11)$$

$Q_3(123)$ may be characterized as follows: it is the set of connected diagrams with three white circles 1, 2, 3 and containing no network of bonds attached to only a pair of white circles. In other words, from any black circle in $Q_3(123)$ we can reach any white circle by a path of f-bonds without going through another white circle.

We now proceed to classify the diagrams in $Q_3(123)$ according to the cutting circles (CC's) in the diagram. We say that a point A is a CC belonging to the white circle 1, if removing A disconnects 1 from all other white circles. Naturally, there may be more than one CC for 1; however, if there is more than one CC for white circle 1, then there exists a unique maximal CC for 1 which has the property that its deletion leaves more black circles connected to 1 than are left by deletion of any other CC belonging to 1.

We now classify the diagrams in $Q_3(123)$ into four classes, corresponding to three, two, one, or none of the white circles having CC's. Let us examine the case in which CC's exist for all three white circles. We need consider only the three maximal CC's. If the maximal CC's belonging to any two of the white circles are identical, then this black circle is also the maximal CC for the third white circle. Thus there are only two cases: the three maximal CC's are either all identical or all distinct.

In the former case, letting A be the common maximal CC and coloring it white, cutting at A produces three fragments which are diagrams in $u(1A)$, $u(2A)$, and $u(3A)$, respectively. We conclude that the set of all such diagrams in $Q_3(123)$ is given by $\int d(A)u(1A)u(2A)u(3A)$.

In the latter case, letting A, B, C be the maximal CC's for 1, 2, and 3, respectively, coloring them white, and cutting at A, B, and C produces three fragments that are diagrams in $u(1A)$, $u(2B)$, and $u(3C)$, respectively, as well as a fragment which contains the three white circles A, B, and C and is irreducible by virtue of the fact that A, B, and C were chosen to be the maximal CC's for 1, 2, and 3. Since there are no further conditions to be imposed on this fragment, it can be any diagram in $C_3(ABC)$. We conclude that the set of diagrams in $Q_3(123)$ containing three distinct maximal CC's is

$$\int d(A)\, d(B)\, d(C)u(1A)u(2A)u(3A)C_3(ABC).$$

Next consider the case of only two white circles, say 1 and 2, having CC's. Calling the maximal CC's A and B, we note that A and B must be distinct, since otherwise A = B would also be a CC for 3. By an argument analogous to that of the last paragraph these diagrams in $Q_3(123)$ are given by

$$\int d(A)\, d(B)u(1A)u(2B)C_3(AB3).$$

Next there are terms with CC's for only one of the white circles, say circle 1. Coloring the maximal CC white, and cutting the diagram there, we obtain a diagram in $u(1A)$ and a diagram which is characterized as follows: it is f-irreducible, has 3 white circles, and contains no network of f-bonds attached only to white circles 2 and 3. Thus these diagrams are given by $\int d(A)X_3(23 \mid A)u(A1)$, where $X_3(23 \mid A)$ is defined as the subset of $C_3(123)$ consisting of diagrams with no network of bonds attached only to 2 and 3. $X_3(12 \mid 3)$ is symmetric only in the two particles 1 and 2.

Finally, there is the set of f-irreducible diagrams in $Q_3(123)$, which we denote by $Z_3(123)$.

From this point on the complexity of the mathematical expressions makes it preferable to express all equations in diagrammatic form. We use a circular symbol for C_3, for $X_3(12 \mid 3)$ a crescent-shaped symbol with the concave portion of the crescent drawn between circles 1 and 2. $Z_3(123)$ is represented by a concave triangle. A g_2-bond is shown as a dashed line,

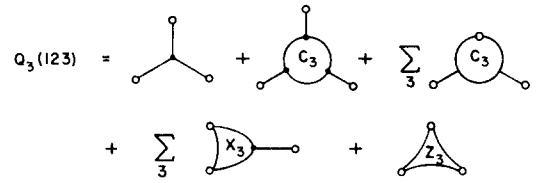


FIG. 1. Topological analysis of $Q_3(123)$.

a u-bond is indicated by a solid line. In our diagrammatic expressions there frequently occur topologically identical terms which differ only by the labeling of the white circles. Where the appropriate sum is obvious we avoid clutter by writing the term only once with white circles unlabeled and preceded by a summation sign, placing the integer denoting the number of terms summed below the summation sign.

The expression just obtained for $Q_3(123)$ is shown in Fig. 1. Equation (11) expresses $g_3(123)$ in terms of $Q_3(123)$. Expanding the three g_2-bonds as $g_2(ij) = 1 + u(ij)$ and retaining all the terms containing enough u-bonds to make the diagram f-irreducible we obtain the expression for $C_3(123)$ shown in Fig. 2. Recalling that $X_3(12 \mid 3)$ is the part of $C_3(123)$ without a network of f-bonds attached only at 1 and 2 leads immediately to the equation for $X_3(12 \mid 3)$ shown in Fig. 3.

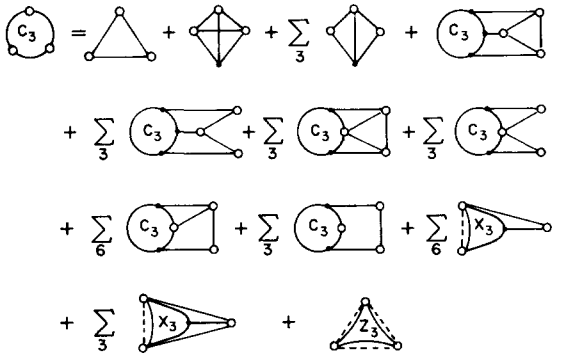


FIG. 2. $C_3(123)$ obtained by calculating the f-irreducible part of $g_3(123)$.

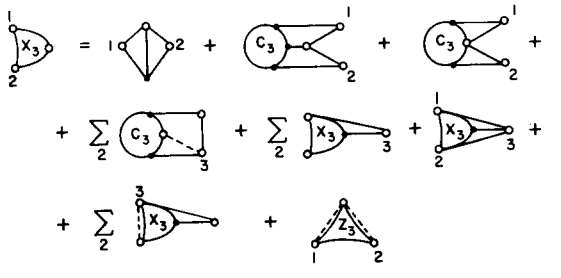


FIG. 3. $X_3(12 \mid 3)$ expressed as the subset of $C_3(123)$ not containing a $u(12)$-bond.

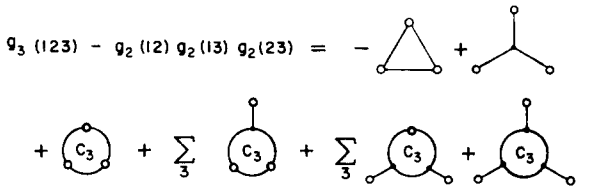


FIG. 4. The relation between $g_3(123)$ and $C_3(123)$ obtained by functional differentiation.

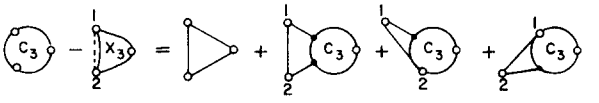


FIG. 5. $X_3(12 \mid 3)$ expressed in terms of $C_3(123)$.

By multiplying the equation of Fig. 1 by

$$g_2(12)g_2(13)g_2(23)$$

and subtracting the equation of Fig. 2 we can obtain a relation for $g_3(123) - C_3(123)$ which does not contain $Z_3(123)$.

A second expression for $g_3(123) - C_3(123)$ is obtained most easily by functional differentiation of (1).[14] It is shown graphically in Fig. 4. After equating the two expressions we find that a symmetric sum of terms, for each of which just one of the white circles is attached by a u-bond to a CC, is equal to zero. We can set each of the groups of terms with particles 1, 2, or 3 having a CC to zero separately. After cancelling the dangling u-bond we find the relation shown in Fig. 5.

Another relation for $C_3(123) - g_2(12)X_3(12 \mid 3)$ can be obtained by multiplying the relation in Fig. 3 by $g_2(12)$ and subtracting from the expression for $C_3(123)$ in Fig. 2. Note that $Z_3(123)$ cancels. When we equate this expression for $C_3(123) - g_2(12)X_3(12 \mid 3)$ to the expression in Fig. 5 we obtain no new results, but just recover the relation of Fig. 5, obtaining a check on the correctness of our calculations.

Actually our way of obtaining the preceding identity has been unnecessarily complicated. It is simpler to construct $C_3(123) - g_2(12)X_2(12 \mid 3)$ by topological analysis. Since $X_2(12 \mid 3)$ is the subset of all diagrams in $C_3(123)$ without a network attached only to white circles 1 and 2, $C_3(123) - g_2(12)X_2(12 \mid 3)$ is the subset of all diagrams in $C_3(123)$ characterized by the following properties: (1) there is a $u(12)$ bond, (2) when the $u(12)$ bond is deleted, the diagram becomes reducible. Further we note that the white circle 3 can be a CC in diagrams contained in

$$[C_3(123) - g_2(12)X_3(12 \mid 3)]/u(12).$$

[14] J. K. Percus, in *The Equilibrium Theory of Classical Fluids*, H. L. Frisch and J. L. Lebowitz, Eds. (W. A. Benjamin, Inc., New York, 1964).

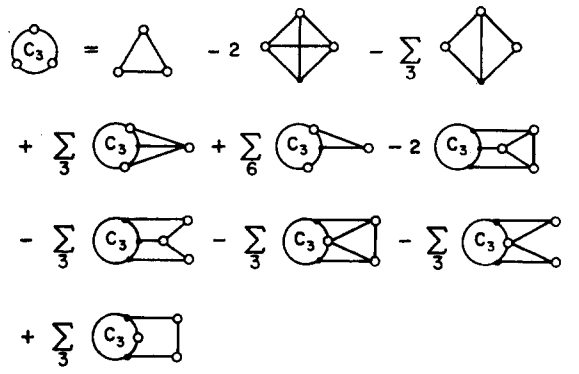


FIG. 6. The PY II closure.

We enumerate all possible topological cases. First, if 3 is a CC, we obtain $u(13)u(23)$. If there are black CC's for both 1 and 2, then the maximal CC's must be distinct. By analogy to previous arguments we obtain a term $\int d(\mathrm{A})\, d(\mathrm{B})u(1\mathrm{A})u(2\mathrm{B})C_3(\mathrm{AB3})$. If only one of the white circles, say 1, has CC's then we obtain a term of the form $\int d(\mathrm{A})u(1\mathrm{A})C_3(\mathrm{A23})$. There is a similar term with 1 and 2 interchanged. Hence the result obtained by direct topological analysis is in agreement with the indirect calculation.

Finally we use the result of Fig. 5 to eliminate the $X_3(\alpha\beta \mid \gamma)$ from Fig. 2. The resulting integral equation contains only C_3, u, and Z_3. The PY II closure consists of setting

$$Z_3^{\mathrm{PY}}(123) = 0, \tag{12}$$

resulting in the integral equation shown in Fig. 6.

To obtain the CHNC II closure, we let $Z_3(123)$ be the sum of all products of connected, reducible diagrams containing three white circles. The sum of all connected reducible diagrams with three white circles and no networks of f-bonds attached only to pairs of white circles is $Q_3(123) - Z_3(123)$. Using the combinatorial theorem used for CHNC,

$$Z_3^{\mathrm{CHNC}}(123) = e^{m_3(123)} - m_3(123) - 1, \tag{13}$$

where $m_3(123) \equiv Q_3(123) - Z_3(123)$. The analogy to the PY and CHNC cases is obvious. More generally, the PY n closure consists of setting $Z_n(12\cdots n) = 0$ in a set of integral equations obtained by an analogous process. The CHNC n systems differ from the PY n system by further including all terms which, when all n white circles are removed simultaneously, do not give rise to a fragment which is both attached to all n white circles, and (including the n white circles) irreducible.

The equation in Fig. 6 is a linear integral equation

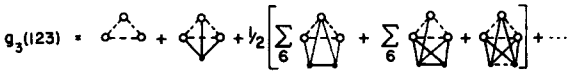


FIG. 7. Expansion of $g_3(123)$ in g_2-bonds and u-bonds.

for $C_3(123)$, which may be solved formally by iteration, leading to an expansion of $C_3(123)$ in irreducible diagrams of u-bonds. Using Fig. 4 we can then obtain the corresponding expansion of $g_3(123)$. The corresponding expansion of the exact $g_2(123)$ has been obtained by Abe[15] and by Verlet.[16] The three lowest orders of their expansion are shown in Fig. 7. The PY II result reproduces the two lowest orders correctly. In the next order it reproduces the first two terms, but omits the last term shown in Fig. 7.

Several comments are in order. Since the closure does not contain the pair potential explicitly, its adequacy can be tested by direct computation from experimental data for g_3 and g_2 at a single density ρ and temperature T subject only to the assumption of negligible non-additive forces. Presently available data are insufficient for such a comparison.

The closure must be combined with the Ornstein–Zernike relation (1), and the relation[10]

$$\frac{\partial}{\partial \rho} C_2(12) = \int d(3) C_3(123) \tag{14}$$

to obtain a closed system. Because of the derivative with respect to ρ in Eq. (14), we cannot solve the system at a single density $\rho \neq 0$. We can only advance in ρ from a known starting solution. This presents no difficulty in the one phase region, and in the gas phase below the critical point. We integrate from $\rho = 0$, using the integrated form of Eq. (14)

$$C_2(12, \rho) = f(12) + \int_0^{\rho} d\rho' \int d(3) C_3(123; \rho'). \tag{15}$$

It is noteworthy that the pair potential makes its only explicit appearance in the $f(12)$ which appears as a constant of integration in Eq. (15).

It seems highly likely that the PY II and CHNC II systems resemble the PY and CHNC equations in not yielding van der Waals loops below the critical point, showing instead an intermediate density region where the equation has no solution.[6] If this proves true, then it is at present not clear how to use the PY II and CHNC II systems in the liquid below the critical point, since we lack a starting solution.

V. VIRIAL COEFFICIENTS

Figure 8 shows the five and six particle diagrams omitted from the exact $C_3(123)$ in PY II approximation. It is convenient to use e-bonds both in constructing the diagram and in reducing the number of diagrams. Also listed are the weight of each diagram, and the value when integrated. The actual value is the

[15] R. Abe, Progr. Theoret. Phys. (Kyoto) **21**, 421 (1959).
[16] L. Verlet, Nuovo Cimento **18**, 77 (1960).

value listed multiplied by $[a^{n-1}/n!]^{\mu}$, where a is the length of a side of the square or cube, n is the number of particles in the diagram, and μ is the dimensionality ($\mu = 2$ or 3). For hard rods all the diagrams shown vanish.

In the following, superscripts indicate the total number of particles in a diagram. The omitted part of $C_3^{(5)}(123)$ is just the leading term in

$$Z_3(123)g_2(12)g_2(23)g_2(13).$$

This term is zero for both hard squares and hard disks, since the e-bonds and f-bonds in the diagram are geometrically incompatible. A certain number of terms in $C_3^{(6)}(123)$ contain the same arrangement of e-bonds and f-bonds and hence are also equal to zero for hard squares and hard disks. The first three terms arise from omission of $Z_3^{(5)}(123)$, the remaining terms are due to omitting $Z_3^{(6)}(123)$ in the closure relation shown in Fig. 7.

Figure 9 shows the additional 5 and 6 particle diagrams included in CHNC II but not in PY II. The inclusion of additional diagrams spoils the exactness of $C_3^{(5)}(123)$ for hard squares and disks. For hard rods the additional diagrams vanish individually. This suggests the conjecture that the CHNC II system is exact for hard rods.

NO.	DIAGRAM	WEIGHT	VALUE INTEGRATED HARD SQUARES	HARD CUBES
1		1	0	34,560
1		9	0	− 2,985,984
2		18	0	8,097,024
3		9	0	8,097,024
4		1	0	0
5		9	0	2,824,704
6		18	0	− 4,764,288
7		9	0	− 5,819,904
8		18	0	11,983,104
9		3	0	10,948,608
10		18	96	5,178,981
11		18	− 43,776	−162,933,888
12		6	− 62,176	− 230,130,048
13		18	74,624	303,309,312
14		18	−192	−13,972,992

FIG. 8. Five and six particle diagrams omitted in PY II approximation. The solid lines are f-bonds; the dashed lines are e-bonds.

NO.	DIAGRAM	WEIGHT	VALUE INTEGRATED HARD SQUARES	HARD CUBES
1		1	960	503,040
1		9	−10,368	− 34,338,816
2		18	25,952	102,826,368
3		9	25,952	102,826,368
4		1	−10,368	− 34,338,816
5		18	−20,896	− 72,115,584
6		18	40,032	155,924,352
7		18	47,552	194,423,424
8		18	−80,576	− 367,036,800

FIG. 9. Five and six particle diagrams included by CHNC II but omitted by PY II. The solid lines are f-bonds; the dashed lines are e-bonds.

Table I shows the fifth virial coefficients for hard rods, squares, cubes, and spheres obtained by the PY[17] and CHNC[17] equations, the PY 2 and CHNC 2 extensions given by Verlet[6] and the PY II and CHNC II systems derived here. Where they differ, the PY theory always yields a better virial coefficient than the corresponding CHNC theory. The more complicated PY II and CHNC II systems derived here give better fifth virial coefficients than the PY 2 and CHNC 2 systems.

TABLE I. The fifth virial coefficient for hard particles calculated from various integral equations.

approximation \ system	rods	squares	cubes	spheres
exact	1	3.7222	3.1597	0.1097
PY	1	4.2361	12.4303	0.121
PY 2	1	3.639	1.701	0.107
PY II	1	3.7222	3.2431	0.1098
CHNC	0.800	1.8042	−6.9394	0.049
CHNC 2	1.033	4.039	5.198	0.122
CHNC II	1	3.6656	2.0301	0.1065

Table II summarizes sixth virial coefficients obtained from the same theories as well as from the PY III and CHNC III systems, which are discussed in the next section. Here the PY 2 equation gives better virial

[17] W. G. Hoover and J. C. Poirier, J. Chem. Phys. **38**, 327 (1963); S. Katsura and Y. Abe, J. Chem. Phys. **39**, 2068 (1963); J. S. Rowlinson, Proc. Roy. Soc. (London) **279**, 147 (1964).

TABLE II. The sixth virial coefficient for hard particles calculated from various integral equations.

approximation \ system	rods	squares	cubes
exact	1	3.02500	−18.8796
PY	1	4.42000	9.2067
PY 2	1	3.020	−21.203
PY II	1	3.11204	−13.5939
PY III	1	3.02500	−18.8777
CHNC	0.70278	0.96500	−2.2133
CHNC II	1	2.94468	−16.9473
CHNC III	1	3.00926	−19.0401

coefficients than the PY II equation. The fact that the situation is the reverse for the fifth virial coefficients suggests that the good sixth virial coefficients yielded by PY 2 may to some extent be accidental. Successively better comparisons could be made by comparing $C_2^{(6)}(12)$ and $C_3^{(6)}(123)$ for the exact system and the two approximations.

The results obtained suggest undertaking a more extensive test by numerical integration of the PY II system. In view of the exact fifth virial coefficient yielded by PY II and the availability of extensive Monte Carlo[18] and molecular dynamics[19] calculations, the hard-disk system seems particularly appropriate for such a test.

In general, the PY n (CHNC n) system yields the first $n+2$ virial coefficients exactly. However, for hard particles, geometrical incompatibility of f-bonds and e-bonds in certain diagrams can make certain omitted terms equal to zero, leading to a number of "free" exact virial coefficients. The possibilities for this are better, the lower the dimensionality of the system. They are also much better for PY n than for CHNC n, although the number of "free" virial coefficients increases sporadically with increasing n in both cases. For hard disks, the CHNC sequence first exhibits a free virial coefficient for the CHNC V system. Here the contribution from $Z^{(8)\mathrm{CHNC}}(123456)$ is zero because a disk cannot overlap six mutually nonoverlapping discs.

VI. THE PY III SYSTEM

Although there is little likelihood that the PY III system can be solved numerically in the near future, we sketch its derivation in order to illustrate how the method of topological analysis can be applied to finding a closure for $C_n(12\cdots n)$. Only a few minor extensions of the techniques used previously are needed.

We begin by factoring $g_4(1234)$. After taking out a complete star of six g_2-bonds, there remain networks of f-bonds attached to three white circles and a network of bonds attached to all four white circles. Thus

$$g_4(1234)=\prod_{\substack{\alpha\neq\beta=\\1,2,3,4}} g_2(\alpha\beta)\left[1+\sum_{\substack{\mu\neq\nu\neq\sigma=\\1,2,3,4}} Q_3(\mu\nu\sigma)+Q_4(1234)\right]. \tag{16}$$

$Q_4(1234)$ is the sum of connected diagrams containing four white circles and having the following properties: there may be CC's but no AC's; there is no network of f-bonds attached to only a pair of white circles.

At this point we could have analyzed $Q_4(1234)$ further into a sum of products of Q_3's and a factor consisting of the sum of connected diagrams with four white circles and having the property that every white circle can be reached from every black circle by a path of f-bonds without going through another white circle. However, this procedure would prove very inconvenient later on.

We now classify the diagrams in $Q_4(1234)$ according to CC's. Let us first consider the case in which there exists at least one 2–2 CC, defined as a CC the removal of which splits the diagram into two connected fragments, each containing two white circles, say 1, 2 and 3, 4. Let A be the minimal CC for the pair 1, 2, i.e., the CC that disconnects more black circles from the pair 1, 2 than any other CC while leaving 1 and 2 connected. Let us color A white, cut at A, and denote by $H(12\mathrm{A})$ the connected fragment containing 1, 2, and A. Since A was chosen as the minimal CC for the pair (1, 2), A has no CC in $H(12\mathrm{A})$. Furthermore, there is no direct $u(12)$-bond. The diagram is then classified according to its CC's. If there are no CC's, it is in $X_3(12\,|\,\mathrm{A})$; if there is one CC, say for particle 1, then we have a diagram in $\int d(\mathrm{B})C_3(2\mathrm{AB})u(1\mathrm{B})$; if there are two distinct CC's for 1 and 2, we have a diagram in

$$\int d\mathrm{B}\; d\mathrm{C}C_3(\mathrm{ABC})u(1\mathrm{B})u(2\mathrm{C}).$$

Finally, if the CC's of 1 and 2 coincide, then they must also coincide with A, since A was assumed to be the minimal CC of the pair (1, 2). In this case we obtain simply $u(1\mathrm{A})u(2\mathrm{A})$.

Letting D be the minimal CC for the pair 3, 4, and coloring D white, we obtain a similar result for the fragment characterized by the white circles 3, 4 and D. Either A and D are identical, or they are connected

[18] W. W. Wood, Los Alamos Scientific Laboratory Report No. LA-2827 (1963).

[19] B. J. Alder and T. E. Wainwright, J. Chem. Phys. **31**, 459 (1959); **33**, 1439 (1960); Phys. Rev. **127**, 359 (1962).

by a u(AD) bond. Thus we can write the total contribution as

$$\sum_3 \int d\mathrm{A}\, d\mathrm{B}\, H(\alpha\beta\mathrm{A})\hat{u}(\mathrm{AB})\mathrm{H}(\gamma\delta\mathrm{B}),$$

where the sum runs over the three ways of partitioning 1, 2, 3, 4 into pairs. The quantity

$$\hat{u}(12) \equiv u(12) + \rho^{-1}\delta(12)$$

is useful for condensing notation; it is indicated in subsequent figures by a zig-zag line.

In the preceding there is one term which is symmetric in all four particles. It is the term

$$\int d\mathrm{A} u(1\mathrm{A})u(2\mathrm{A})u(3\mathrm{A})u(4\mathrm{A}),$$

which actually occurs only once, whereas we have counted it three times. Hence twice this term must be subtracted.

If there is no 2–2 CC, either four, three, two, one, or none of the white circles can have CC's. No two of the CC's can be identical since there is no 2–2 CC. In the case of 4 CC's we obtain

$$\int C_4(\mathrm{ABCD})u(\mathrm{A1})u(\mathrm{B2})u(\mathrm{C3})u(\mathrm{D4})\, d\mathrm{A}\, d\mathrm{B}\, d\mathrm{C}\, d\mathrm{D}.$$

If three of the white circles—say 1, 2, and 3—have CC's, we obtain

$$\int C_4(\mathrm{ABC4})u(\mathrm{A1})u(\mathrm{B2})u(\mathrm{C3})\, d\mathrm{A}\, d\mathrm{B}\, d\mathrm{C}.$$

If there are CC's for white circles 1 and 2, we obtain $\int X_4(34 \mid \mathrm{AB})u(\mathrm{A1})u(\mathrm{B2})\, d\mathrm{A}\, d\mathrm{B}$. Here $X_4(12 \mid 34)$ is the subset of diagrams in $C_4(1234)$ with no network attached only to white circles 1 and 2. Similarly, if 1 is the only white circle with a CC, we obtain

$$\int X_4(234 \mid \mathrm{A})u(\mathrm{A1})\, d\mathrm{A}.$$

Here $X_4(123 \mid 4)$ denotes the subset of diagrams in $C_4(1234)$ containing no network attached only to any of the pairs of white circles 12, 13, or 23. The expression for $Q_4(1234)$ is shown diagrammatically in Fig. 10.

$X_4(12 \mid 34)$ is indicated by a crescent with the concave part between circles 1 and 2. $X_4(123 \mid 4)$ is shown as a double-crescent with 1, 2, and 3 located on the three kinks. The symbol used for $Z_4(1234)$ is diamond-shaped.

We can now write down $C_4(1234)$ by noting that it is the f-irreducible part of $g_4(1234)$, expressing the latter in terms of $Q_4(1234)$ by using (16), expanding the g_2-bonds as $g_2 = u + 1$, and keeping terms with

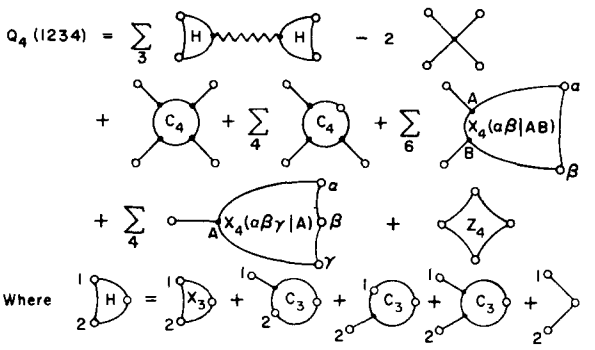


FIG. 10. Topological analysis of $Q_4(1234)$.

enough u-bonds to make the diagram f-irreducible. The calculation is tedious but straightforward and produces a large number of terms. Since this intermediate result still contains the $X_4(\alpha\beta \mid \gamma\delta)$ and $X_4(\alpha\beta\gamma \mid \delta)$, we do not exhibit it explicitly.

We do not use the expressions for $X_4(12 \mid 34)$ and $X_4(123 \mid 4)$ obtained by picking out of our expression for $C_4(1234)$ the subsets with no $u(12)$-bond, and no $u(12)$-, $u(13)$-, or $u(23)$-bonds respectively. It is much simpler to construct $X_4(12 \mid 34)$ and $X_4(123 \mid 4)$ by direct topological analysis. We begin by constructing $C_4(1234) - g_2(12)X_4(12 \mid 34)$.

$C_4(1234) - g_2(12)X_4(12 \mid 34)$ is irreducible, has four white circles, and is characterized by the following properties: (1) there is a $u(12)$-bond, (2) if the $u(12)$-bond is destroyed the diagram becomes reducible. We classify the diagrams obtained by destroying $u(12)$ according to the CC's, remembering that the white circles 3 and 4 can be CC's for 1 and 2. The result obtained for $[C_4(1234) - g_2(12)X_4(12 \mid 34)]/u(12)$ is shown in Fig. 11.

Finally, in order to eliminate the $X_4(\alpha\beta\gamma \mid \delta)$, we must calculate either

$$X_4(12 \mid 34) - g_2(13)g_2(23)X_4(123 \mid 4) \quad \text{or}$$
$$C_4(1234) - g_2(12)g_2(13)g_2(23)X_4(123 \mid 4).$$

The quantity

$$\mathrm{C}_4(1234) - g_2(12)g_2(13)g_2(23)X_4(123 \mid 4)$$

is the set of irreducible diagrams containing four white circles and having the following properties: (1) there are one, two, or three bonds out of the set 12,

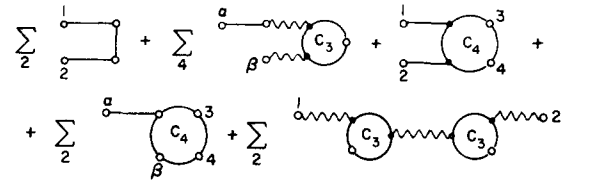


FIG. 11. The result obtained for $[C_4(1234) - g_2(12)X_4(12 \mid 34)]/u(12)$ by topological analysis.

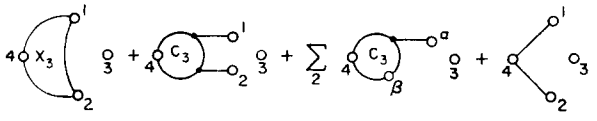


FIG. 12. The quantity $H(12, 3, 4)$.

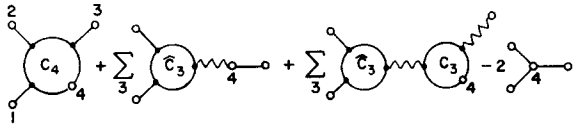


FIG. 13. The quantity $L(123, 4)$.
$\hat{C}_3(123) \equiv C_3(123) + \rho^{-2}\delta(12)\delta(13)$.

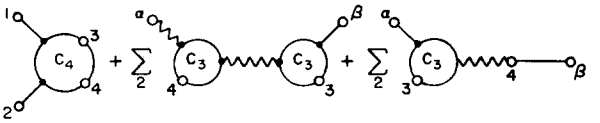


FIG. 14. The quantity $M(12, 3, 4)$.

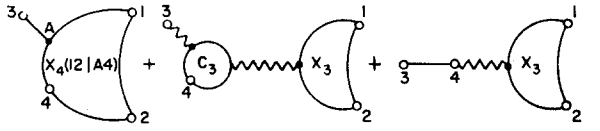


FIG. 15. The quantity $N(12, 3, 4)$.

13, 23; (2) when all the 12, 13, and 23 bonds are destroyed, the diagram becomes reducible. We now consider the reducible diagrams obtained after destroying all 12, 13, 23 bonds and classify them according to the minimum number of bonds in the set 12, 13, 23 which must be added in order to make an irreducible diagram. There are four distinct cases.

The first set consists of all diagrams which require addition of at least a specific pair of u-bonds. Thus addition of $u(13)$ and $u(23)$ is required if and only if 3 is an isolated point. Whether 1 and 2 have CC's clearly does not matter. We denote the latter set of diagrams by $H(12, 3, 4)$. The expression obtained for $H(12, 3, 4)$ by our usual procedure is shown in Fig. 12.

The second set, denoted by $L(123, 4)$, consists of the set of diagrams requiring addition of at least any pair of bonds out of the set 12, 13, 23. $L(123, 4)$ is shown in Fig. 13. This set is characterized by the existence of a CC for each of the three white circles 1, 2, and 3.

The third set consists of all diagrams requiring addition of at least one specific bond. $M(12, 3, 4)$, the set of diagrams requiring addition of at least the $u(12)$-bond, is shown in Fig. 14. This set is characterized by the absence of a CC for white circle 3, and the absence of a 2–2 CC which separates 3 from both 1 and 2.

The fourth set consists of diagrams requiring addition of either of a pair of bonds. $N(12, 3, 4)$, the set of diagrams needing $u(13)$ or $u(23)$ is shown in Fig. 15. The set is characterized by neither 1 nor 2 having a CC, and the presence of a CC that separates 3 from both 1 and 2.

In terms of the quantities just defined we can now write

$$\begin{aligned} &C_4(1234) - g_2(12)g_2(13)g_2(23)X_4(123 \mid 4) \\ &\quad = \sum_3 H(\alpha\beta, \gamma, 4)u(\alpha\gamma)u(\beta\gamma)g_2(\alpha\beta) + L(123, 4)[u(12)u(13)u(23) + \sum_3 u(\alpha\beta)u(\alpha\gamma)] \\ &\qquad + \sum_3 M(\alpha\beta, \gamma, 4)[u(\alpha\beta)g_2(\alpha\gamma)g_2(\beta\gamma) + u(\alpha\gamma)u(\beta\gamma)] + \sum_3 N(\alpha\beta, \gamma, 4)[u(\alpha\gamma)u(\beta\gamma) + u(\alpha\gamma) + u(\beta\gamma)]g_2(\beta\alpha). \end{aligned} \tag{17}$$

The quantity $X_4(12 \mid 34) - g_2(13)g_2(23)X_4(123 \mid 4)$ is the set of all irreducible diagrams containing four white circles and having the following properties: (1) there is no $u(12)$ bond; (2) there are one or two bonds out of the set $u(23)$, $u(13)$; (3) when all the $u(23)$, $u(13)$ bonds are destroyed the diagram becomes reducible.

The diagrams obtained by destroying all $u(12)$ and $u(13)$ bonds can be classified as belonging to H, L, M, or N. Thus we obtain the expression

$$\begin{aligned} &X_4(12 \mid 34) - g_2(13)g_2(23)X_4(123 \mid 4) \\ &\quad = [H(12, 3, 4) + L(12, 3, 4) + M(12, 3, 4)]u(13)u(23) + \sum_2 M(\alpha 3, \beta, 4)u(\alpha 3)g_2(\beta 3) \\ &\qquad + N(12, 3, 4)[u(13)u(23) + u(13) + u(23)] + \sum_2 N(\alpha 3, \beta, 4)u(\beta 3)g_2(\alpha 3). \end{aligned} \tag{18}$$

Multiplying (18) by $g_2(12)$ and subtracting from (17), we obtain an alternative expression for

$$\begin{aligned} C_4(1234) - g_2(12)X_4(12\ 34) &= u(12)\{\sum_2 H(\alpha 3, \beta, 4)u(\beta 3)g_2(\alpha 3) + L(123, 4)[u(13) + u(23)] \\ &\quad + M(12, 3, 4)[u(13) + u(23) + 1] + \sum_2 M(\alpha 3, \beta, 4)u(\beta 3) + \sum_2 N(\alpha 3, \beta, 4)g_2(\alpha 3)\}. \end{aligned} \tag{19}$$

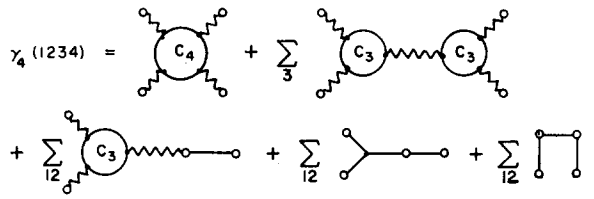


FIG. 16. The relation between $g_4(1234)$ and $C_4(1234)$ obtained by functional differentiation.

We can equate (19) to the alternative expression for $[C_4(1234) - g_2(12)X_4(12 \mid 34)]/u(12)$ shown in Fig. 11, equating to zero separately terms with $u(13)$ and $u(23)$, terms with only $u(13)$, terms with only $u(23)$, and terms with neither. In every case we obtain an identity, i.e., a relation already obtained previously. This confirms the correctness of our calculation.

As a final check, we eliminate the $X_4(\alpha\beta \mid \gamma\delta)$ and $X_4(\alpha\beta\gamma \mid \delta)$ from the expression for $g_4(1234) - C_4(1234)$ obtained from Fig. 10. Of course, $Z_4(1234)$ disappears in the subtraction. The result obtained is just the convolution relation between $g_4(1234)$ and $C_4(1234)$, which is obtained by double functional differentiation of (1) followed by clearing the expression of δ-functions. The result is shown in Fig. 16. The quantity $\gamma_4(1234)$ is defined by

$$\gamma_4(1234) = g_4(1234) - \sum_4 g_3(\alpha\beta\gamma) - \sum_3 g_2(\alpha\beta)g_2(\gamma\delta) + 2\sum_6 g_2(\alpha\beta) - 6.$$

The exact relation in Fig. 4 is used to express g_3 in terms of C_3.

By eliminating the $X_4(\alpha\beta\gamma \mid \delta)$ and $X_4(\alpha\beta \mid \gamma\delta)$ from the relation for $C_4(1234)$ implicit in Fig. 10, followed by setting $Z_4(1234) = 0$, we finally obtain the PY III closure shown in Fig. 17. $\hat{C}_3(123)$ is defined by $\hat{C}_3(123) = C_3(123) + \rho^{-2}\delta(12)\delta(13)$. The quantities $\hat{C}_3$, X_3, and g_2 are retained solely for the sake of more compact notation. Eliminating them casts the equation of Fig. 17 in the form of an integral equation in C_4, C_3, and u. By expanding the g_2-bonds as $g_2 = u + 1$ one verifies easily that all terms are f-irreducible.

To avoid clutter, some diagrams have been drawn as products of two terms; the first being a connected diagram, the second a set of added u-bonds which make it irreducible. Where there are several topologically equivalent diagrams of u-bonds in the second factor we show only one, indicating the others by a summation sign.

VII. CONCLUSION

The method of topological analysis has been applied to the problem of finding a closure to supplement the defining convolution relation between $g_n(12\cdots n)$ and $C_n(12\cdots n)$. In the process we are led to define certain new quantities, namely subsets of the diagrams in $C_n(12\cdots n)$ characterized by the absence of networks of f-bonds between some pairs of white circles. Identities have been derived which eliminate all of the new quantities except $Z_n(12\cdots n)$ which is the subset of diagrams in $C_n(12\cdots n)$ containing no network of f-bonds attached only at any pair of white circles. The PY n closure sets $Z_n(12\cdots n)$ equal to zero. The CHNC closure sums a certain subset of the diagrams in the true $Z_n(12\cdots n)$.

The systems of integral equations derived have several distinctive features. One is the appearance of derivatives with respect to the number density. This makes it necessary to advance in density from a given starting solution. An attractive feature of the theory is the fact that the number of "super-bonds" that appear has been maximized. The intermolecular potential $V(ij)$ appears explicitly only when the system is integrated from $\rho = 0$. Otherwise, instead of $f(ij)$, the value of $C_2(12;\rho)$ at the starting density $\rho = \rho_0$ appears in the PY II and CHNC II equations. This is very much in line with the original idea of Ornstein and Zernike,[9] who introduced $C_2(12)$ as a kind of effective f-function. Thus, one can advance in

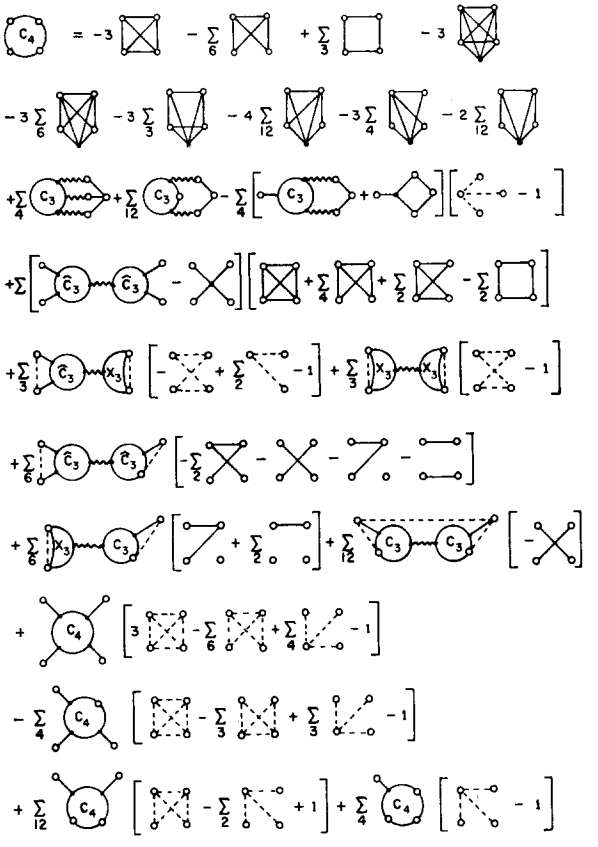


FIG. 17. The PY III closure.

density by numerical integration from any experimentally determined $C_2(12;\rho_0)$ without inquiring into the form of the intermolecular potential.

The procedure given in this paper can be applied for any number of white circles. In general one writes

$$g_n(12\cdots n) = g_n^{(S)}(12\cdots n) + \tilde{g}_n(12\cdots n).$$

Here $g_n^{(S)}$ is the superposition part of $g_n(12\cdots n)$, i.e., the part which can be factored into a product of g_m with $m < n$. The terms in $g_n^{(S)}$ are obtained by enumerating all the ways a complete star of n-vertices can be broken up into complete stars of lower order and associating a factor $\tilde{g}_m(12\cdots m)$ with each complete star of m particles. $Q_n(12\cdots n)$ is defined by

$$Q_n(12\cdots n) = \tilde{g}_n(12\cdots n)/\prod_{\substack{\alpha<\beta= \\ 1,2\cdots n}} g_2(\alpha\beta).$$

ACKNOWLEDGMENT

Part of this work was performed under the auspices of the United States Atomic Energy Commission. This work was begun while the author was a visitor, on PRT leave from Los Alamos Scientific Laboratory, at Universite de Paris, Laboratoire de Physique Théorique, Orsay (S.-et O.), France. The author wishes to thank Professor B. Jancovici, Professor M. M. Lévy, and Professor L. Verlet for their hospitality at Laboratoire de Physique Théorique, Orsay.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Null Fields in Einstein–Maxwell Field Theory. II

J. R. TROLLOPE
University of Alberta, Edmonton, Alberta, Canada

(Received 22 September 1966)

If the propagation vector of an electromagnetic null field is proportional to a gradient, then the metric tensor can be expressed in terms of only four unknown functions. In the present analysis, it is assumed that the divergence of this vector does not vanish. It is shown that the field equations reduce to a single partial differential equation. Particular solutions which represent electromagnetic and purely gravitational waves are given.

1. INTRODUCTION

IN a previous discussion[1] of the null electromagnetic field, it was convenient to introduce a tetrad $(\mathbf{w}_i, \mathbf{a}_i, \mathbf{e}_i, \mathbf{h}_i)$ of real vectors. The vectors $\mathbf{e}_i$ and $\mathbf{h}_i$ are unit and orthogonal to their respective complements while $\mathbf{w}_i$ and $\mathbf{a}_i$ are null with unit inner product. The metric tensor, the electromagnetic tensor, and its dual were shown to have the following forms:

$$\mathsf{g}_{ij} = \mathbf{w}_i\mathbf{a}_j + \mathbf{w}_j\mathbf{a}_i + \mathbf{e}_i\mathbf{e}_j + \mathbf{h}_i\mathbf{h}_j, \tag{1.1}$$

$$\mathsf{F}_{ij} = \mathbf{w}_i\mathbf{e}_j - \mathbf{w}_j\mathbf{e}_i, \tag{1.2}$$

$$^*\mathsf{F}_{ij} = \mathbf{w}_i\mathbf{h}_j - \mathbf{w}_j\mathbf{h}_i. \tag{1.3}$$

The bulk of the analysis was carried out subject to the *ad hoc* condition $\mathbf{w}^i_{;i} = 0$. The reason for this assumption was that it is sufficient (but not necessary) to ensure that $\mathbf{w}_i$ is proportional to a gradient. If the latter is the case, then one can introduce coordinates so that the basis vectors have the following representations:

$$\mathbf{w}_i = e^r(1, 0, 0, 0), \tag{1.4}$$

$$\mathbf{a}_i = e^{-r}(\alpha, 1, \beta, \gamma), \tag{1.5}$$

$$\mathbf{e}_i = e^{-r}(0, 0, 1, 0), \tag{1.6}$$

$$\mathbf{h}_i = e^{-r}(0, 0, 0, 1). \tag{1.7}$$

The symbols α, β, γ, and τ are functions of the coordinates. The form of the metric was determined from (1.1) and the Ricci tensor R_{ij} was computed and subjected to the field equation

$$\mathsf{R}_{ij} = 2\mathbf{w}_i\mathbf{w}_j. \tag{1.8}$$

However, the conservation equation

$$\mathsf{R}^i_{j;i} = 0 \tag{1.9}$$

can be combined with the identity

$$(\mathbf{w}^i\mathbf{w}_i)_{;j} = 0 \tag{1.10}$$

to yield

$$\mathbf{w}^i_{;i}\mathbf{w}_j - \mathbf{w}^i(\mathbf{w}_{i;j} - \mathbf{w}_{j;i}) = 0. \tag{1.11}$$

[1] M. Wyman and R. Trollope, J. Math. Phys. 6, 1995 (1965).

density by numerical integration from any experimentally determined $C_2(12;\rho_0)$ without inquiring into the form of the intermolecular potential.

The procedure given in this paper can be applied for any number of white circles. In general one writes

$$g_n(12\cdots n) = g_n^{(S)}(12\cdots n) + \tilde{g}_n(12\cdots n).$$

Here $g_n^{(S)}$ is the superposition part of $g_n(12\cdots n)$, i.e., the part which can be factored into a product of g_m with $m < n$. The terms in $g_n^{(S)}$ are obtained by enumerating all the ways a complete star of n-vertices can be broken up into complete stars of lower order and associating a factor $\tilde{g}_m(12\cdots m)$ with each complete star of m particles. $Q_n(12\cdots n)$ is defined by

$$Q_n(12\cdots n) = \tilde{g}_n(12\cdots n)/\prod_{\substack{\alpha<\beta= \\ 1,2\cdots n}} g_2(\alpha\beta).$$

ACKNOWLEDGMENT

Part of this work was performed under the auspices of the United States Atomic Energy Commission. This work was begun while the author was a visitor, on PRT leave from Los Alamos Scientific Laboratory, at Universite de Paris, Laboratoire de Physique Théorique, Orsay (S.-et O.), France. The author wishes to thank Professor B. Jancovici, Professor M. M. Lévy, and Professor L. Verlet for their hospitality at Laboratoire de Physique Théorique, Orsay.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Null Fields in Einstein–Maxwell Field Theory. II

J. R. TROLLOPE
University of Alberta, Edmonton, Alberta, Canada

(Received 22 September 1966)

If the propagation vector of an electromagnetic null field is proportional to a gradient, then the metric tensor can be expressed in terms of only four unknown functions. In the present analysis, it is assumed that the divergence of this vector does not vanish. It is shown that the field equations reduce to a single partial differential equation. Particular solutions which represent electromagnetic and purely gravitational waves are given.

1. INTRODUCTION

IN a previous discussion[1] of the null electromagnetic field, it was convenient to introduce a tetrad $(\mathbf{w}_i, \mathbf{a}_i, \mathbf{e}_i, \mathbf{h}_i)$ of real vectors. The vectors $\mathbf{e}_i$ and $\mathbf{h}_i$ are unit and orthogonal to their respective complements while $\mathbf{w}_i$ and $\mathbf{a}_i$ are null with unit inner product. The metric tensor, the electromagnetic tensor, and its dual were shown to have the following forms:

$$\mathsf{g}_{ij} = \mathbf{w}_i\mathbf{a}_j + \mathbf{w}_j\mathbf{a}_i + \mathbf{e}_i\mathbf{e}_j + \mathbf{h}_i\mathbf{h}_j, \tag{1.1}$$

$$\mathsf{F}_{ij} = \mathbf{w}_i\mathbf{e}_j - \mathbf{w}_j\mathbf{e}_i, \tag{1.2}$$

$$^*\mathsf{F}_{ij} = \mathbf{w}_i\mathbf{h}_j - \mathbf{w}_j\mathbf{h}_i. \tag{1.3}$$

The bulk of the analysis was carried out subject to the *ad hoc* condition $\mathbf{w}^i_{;i} = 0$. The reason for this assumption was that it is sufficient (but not necessary) to ensure that $\mathbf{w}_i$ is proportional to a gradient. If the latter is the case, then one can introduce coordinates so that the basis vectors have the following representations:

$$\mathbf{w}_i = e^r(1, 0, 0, 0), \tag{1.4}$$

$$\mathbf{a}_i = e^{-r}(\alpha, 1, \beta, \gamma), \tag{1.5}$$

$$\mathbf{e}_i = e^{-r}(0, 0, 1, 0), \tag{1.6}$$

$$\mathbf{h}_i = e^{-r}(0, 0, 0, 1). \tag{1.7}$$

The symbols α, β, γ, and τ are functions of the coordinates. The form of the metric was determined from (1.1) and the Ricci tensor R_{ij} was computed and subjected to the field equation

$$\mathsf{R}_{ij} = 2\mathbf{w}_i\mathbf{w}_j. \tag{1.8}$$

However, the conservation equation

$$\mathsf{R}^i_{j;i} = 0 \tag{1.9}$$

can be combined with the identity

$$(\mathbf{w}^i\mathbf{w}_i)_{;j} = 0 \tag{1.10}$$

to yield

$$\mathbf{w}^i_{;i}\mathbf{w}_j - \mathbf{w}^i(\mathbf{w}_{i;j} - \mathbf{w}_{j;i}) = 0. \tag{1.11}$$

[1] M. Wyman and R. Trollope, J. Math. Phys. 6, 1995 (1965).

Putting $j = 1$ and making use of (1.4)–(1.7), we conclude that

$$\mathrm{w}^i_{;i} + \mathrm{e}^r{}_{,2} = 0. \tag{1.12}$$

Hence, the *ad hoc* condition $\mathrm{w}^i_{;i} = 0$ requires that τ be independent of x^2.

In the present analysis it is assumed that w_i is proportional to a gradient but $\mathrm{w}^i_{;i} \neq 0$. That is, the metric is of the same form as it was in the previous discussion but τ is no longer independent of x^2.

2. FIELD EQUATIONS

According to (1.1) and (1.4)–(1.7) the metric is given by

$$g_{ij} = \begin{pmatrix} 2\alpha & 1 & \beta & \gamma \\ 1 & 0 & 0 & 0 \\ \beta & 0 & e^{-2r} & 0 \\ \gamma & 0 & 0 & e^{-2r} \end{pmatrix} \tag{2.1}$$

A direct computation provides the following nonzero Christoffel symbols:

$$\left\{\begin{matrix}1\\11\end{matrix}\right\} = -\alpha_{,2}, \quad \left\{\begin{matrix}1\\33\end{matrix}\right\} = \left\{\begin{matrix}1\\44\end{matrix}\right\} = e^{-2r}\tau_{,2},$$

$$\left\{\begin{matrix}1\\13\end{matrix}\right\} = -\tfrac{1}{2}\beta_{,2}, \quad \left\{\begin{matrix}1\\14\end{matrix}\right\} = -\tfrac{1}{2}\gamma_{,2},$$

$$\left\{\begin{matrix}3\\11\end{matrix}\right\} = e^{2r}(\beta\alpha_{,2} + \beta_{,1} - \alpha_{,3}),$$

$$\left\{\begin{matrix}4\\11\end{matrix}\right\} = e^{2r}(\gamma\alpha_{,2} + \gamma_{,1} - \alpha_{,4}),$$

$$\left\{\begin{matrix}3\\12\end{matrix}\right\} = \tfrac{1}{2}e^{2r}\beta_{,2}, \quad \left\{\begin{matrix}4\\12\end{matrix}\right\} = \tfrac{1}{2}e^{2r}\gamma_{,2},$$

$$\left\{\begin{matrix}3\\13\end{matrix}\right\} = \tfrac{1}{2}\beta\beta_{,2}e^{2r} - \tau_{,1}, \quad \left\{\begin{matrix}4\\14\end{matrix}\right\} = \tfrac{1}{2}\gamma\gamma_{,2}e^{2r} - \tau_{,1},$$

$$\left\{\begin{matrix}3\\14\end{matrix}\right\} = \tfrac{1}{2}e^{2r}(\beta\gamma_{,2} + \beta_{,4} - \gamma_{,3}),$$

$$\left\{\begin{matrix}4\\13\end{matrix}\right\} = \tfrac{1}{2}e^{2r}(\gamma\beta_{,2} + \gamma_{,3} - \beta_{,4}),$$

$$\left\{\begin{matrix}3\\23\end{matrix}\right\} = \left\{\begin{matrix}4\\24\end{matrix}\right\} = -\tau_{,2},$$

$$\left\{\begin{matrix}3\\33\end{matrix}\right\} = -\beta\tau_{,2} - \tau_{,3}, \quad \left\{\begin{matrix}4\\44\end{matrix}\right\} = -\gamma\tau_{,2} - \tau_{,4},$$

$$\left\{\begin{matrix}3\\34\end{matrix}\right\} = -\tau_{,4}, \quad \left\{\begin{matrix}4\\43\end{matrix}\right\} = -\tau_{,3},$$

$$\left\{\begin{matrix}3\\44\end{matrix}\right\} = -\beta\tau_{,2} + \tau_{,3}, \quad \left\{\begin{matrix}4\\33\end{matrix}\right\} = -\gamma\tau_{,2} + \tau_{,4},$$

$$\left\{\begin{matrix}2\\11\end{matrix}\right\} = \alpha_{,1} - \sigma\alpha_{,2} + e^{2r}(\beta\alpha_{,3} + \gamma\alpha_{,4} - \beta\beta_{,1} - \gamma\gamma_{,1}),$$

$$\left\{\begin{matrix}2\\12\end{matrix}\right\} = \alpha_{,2} - \tfrac{1}{2}e^{2r}(\beta\beta_{,2} + \gamma\gamma_{,2}),$$

$$\left\{\begin{matrix}2\\13\end{matrix}\right\} = \alpha_{,3} - \tfrac{1}{2}\sigma\beta_{,2} + \beta\tau_{,1} - \tfrac{1}{2}e^{2r}\gamma(\gamma_{,3} - \beta_{,4}),$$

$$\left\{\begin{matrix}2\\14\end{matrix}\right\} = \alpha_{,4} - \tfrac{1}{2}\sigma\gamma_{,2} + \gamma\tau_{,1} - \tfrac{1}{2}e^{2r}\beta(\beta_{,4} - \gamma_{,3}),$$

$$\left\{\begin{matrix}2\\23\end{matrix}\right\} = \tfrac{1}{2}\beta_{,2} + \beta\tau_{,2}, \quad \left\{\begin{matrix}2\\24\end{matrix}\right\} = \tfrac{1}{2}\gamma_{,2} + \gamma\tau_{,2},$$

$$\left\{\begin{matrix}2\\33\end{matrix}\right\} = \beta_{,3} + e^{-2r}(\tau_{,1} + \sigma\tau_{,2}) + \beta\tau_{,3} - \gamma\tau_{,4},$$

$$\left\{\begin{matrix}2\\44\end{matrix}\right\} = \gamma_{,4} + e^{-2r}(\tau_{,1} + \sigma\tau_{,2}) + \gamma\tau_{,4} - \beta\tau_{,3},$$

$$\left\{\begin{matrix}2\\34\end{matrix}\right\} = \tfrac{1}{2}(\gamma_{,3} + \beta_{,4}) + \beta\tau_{,4} + \gamma\tau_{,3},$$

where

$$\sigma = -2\alpha + e^{2r}(\beta^2 + \gamma^2). \tag{2.2}$$

Field equation (1.8) reduces to

$$(-2\tau)_{,ij} - \left\{\begin{matrix}\alpha\\ij\end{matrix}\right\}_{,\alpha} + \left\{\begin{matrix}\alpha\\i\beta\end{matrix}\right\}\left\{\begin{matrix}\beta\\\alpha j\end{matrix}\right\} + 2\left\{\begin{matrix}\alpha\\ij\end{matrix}\right\}\tau_{,\alpha} = 2e^{2r}\delta^1_i\delta^1_j. \tag{2.3}$$

The first equation to be considered is

$$\mathrm{R}_{22} = -2[\tau_{,22} - (\tau_{,2})^2] = 0. \tag{2.4}$$

The first integral of (2.4) gives

$$-\frac{1}{\tau_{,2}} = x^2 + f(x^1, x^3, x^4), \tag{2.5}$$

where f is arbitrary. However, the metric (2.1) is form invariant under the group of transformations of the type

$$\bar{x}^2 = x^2 + \varphi(x^1, x^3, x^4). \tag{2.6}$$

Hence, there is no loss of generality in taking $f \equiv 0$. The second integration leads to

$$e^r = e^\nu/x^2, \tag{2.7}$$

where ν is a function of x^1, x^3, and x^4.

The equation $\mathrm{R}_{23} = 0$ reduces to

$$-\tfrac{1}{2}\beta_{,22} - \beta\tau_{,22} + 2\beta(\tau_{,2})^2 = 0. \tag{2.8}$$

This can be integrated [using (2.7)] to obtain

$$\beta = \xi(x^2)^2 + \xi_0/x^2, \tag{2.9}$$

where the ξ's are independent of x^2. Similarly, $\mathrm{R}_{24} = 0$ implies

$$\gamma = \eta(x^2)^2 + \eta_0/x^2, \tag{2.10}$$

the η's being independent of x^2.

If expressions (2.9) and (2.10) are substituted into the equations $R_{34} = 0$, $\frac{1}{2}(R_{33} - R_{44}) = 0$, one obtains

$$-(\xi_{,4} + \eta_{,3} + 2\xi\nu_{,4} + 2\eta\nu_{,3})x^2 + \tfrac{1}{2}(\xi_{0,4} + \eta_{0,3} + 2\xi_0\nu_{,4} + 2\eta_0\nu_{,4})(x^2)^{-2} + \tfrac{9}{2}\xi_0\eta_0(x^2)^{-4} = 0 \quad (2.11)$$

and

$$-(\xi_{,3} - \eta_{,4} + 2\xi\nu_{,3} - 2\eta\nu_{,4})x^2 + \tfrac{1}{2}(\xi_{0,3} - \eta_{0,4} + 2\xi_0\nu_{,3} - 2\eta_0\nu_{,4})(x^2)^{-2} + \tfrac{9}{4}(\xi_0^2 - \eta_0^2)(x^2)^{-4} = 0. \quad (2.12)$$

Hence, it is necessary that $\xi_0 = \eta_0 = 0$ and

$$u = e^{2\nu}\xi, \quad v = e^{2\nu}\eta \quad (2.13)$$

are harmonic conjugates. One can now consider the equations $R_{33} = 0$, $R_{44} = 0$ separately. The former is

$$R_{33} = -2e^{-2\nu}(\alpha_{,2}x^2 + \alpha) - \nabla^2\nu - (3\xi_{,3} + \eta_{,4} + 2\xi\nu_{,3} - 2\eta\nu_{,4} + 4e^{-2\nu}\nu_{,1})x^2 + 3(\xi^2 + \eta^2)(x^2)^2 = 0, \quad (2.14)$$

where ∇^2 represents the Laplacian with respect to the variables x^3 and x^4. This equation gives the dependence of α on x^2. It implies that

$$\alpha + \alpha_{,2}x^2 = A + Bx^2 + C(x^2)^2, \quad (2.15)$$

where A, B, and C are functions independent of x^2. The integration of (2.15) gives

$$\alpha = (a/x^2) + \alpha_0 + \alpha_1 x^2 + \alpha_2(x^2)^2, \quad (2.16)$$

the a and the α_i being functions of x^1, x^3, and x^4. If this expression for α is now substituted in (2.14), the following relations are obtained:

$$\begin{aligned} &a = \text{arbitrary}, \quad \alpha_0 = -\tfrac{1}{2}e^{2\nu}\nabla^2\nu, \\ &\alpha_1 = -\tfrac{1}{2}e^{2\nu}(\xi_{,3} + \eta_{,4}) - \nu_{,1}, \\ &\alpha_2 = \tfrac{1}{2}e^{2\nu}(\xi^2 + \eta^2). \end{aligned} \quad (2.17)$$

This gives the dependence of α on u, v, and ν and since (2.9)–(2.13) imply that

$$\beta = (x^2)^2 u e^{-2\nu}, \quad \gamma = (x^2)^2 v e^{-2\nu}, \quad (2.18)$$

then the remaining equations must determine the forms of ν and a.

The expressions for R_{44} and R_{12} lead to nothing new. However, the application of (2.17) and (2.18) to R_{13} and R_{14} leads to

$$R_{13} = a_{,3}/(x^2)^2, \quad R_{14} = a_{,4}/(x^2)^2, \quad (2.19)$$

so a depends only on x^1.

Nine of the ten field equations have now been satisfied. The last equation, $R_{11} = 2e^{2\tau}$, takes the following form:

$$R_{11} = R_{11}^{(-2)} \cdot (x^2)^{-2} + R_{11}^{(-1)} \cdot (x^2)^{-1} + R_{11}^{(0)} + R_{11}^{(1)}x^2 + R_{11}^{(2)}(x^2)^2 = 2e^{2\nu}(x^2)^{-2}. \quad (2.20)$$

where the $R_{11}^{(i)}$ are independent of x^2. One can show that all these coefficients are identically zero with the exception of $R_{11}^{(-2)}$. The final equation is

$$-2a_{,1} - 6a\alpha_1 + e^{2\nu}\nabla^2\alpha_0 = 2e^{2\nu}. \quad (2.21)$$

If the expressions in (2.17) are now substituted in (2.21) then we obtain the following partial differential equation which determines ν in terms of u, v, and a:

$$\nabla^2(e^{2\nu}\nabla^2\nu) = -4 + ae^{-2\nu} \times [12(u_{,3} + v_{,4} - 2u\nu_{,3} - 2v\nu_{,4} + \nu_{,1}) - 4(a_{,1}/a)]. \quad (2.22)$$

This equation can be somewhat simplified by putting

$$z = x^3 + ix^4, \quad w = u + iv. \quad (2.23)$$

The result is

$$\frac{\partial^2}{\partial z\partial\bar{z}}\left(e^{2\nu}\frac{\partial^2\nu}{\partial z\partial\bar{z}}\right) + \tfrac{1}{4} = \tfrac{3}{4}a\left[\frac{\partial}{\partial z}we^{-2\nu} + \frac{\partial}{\partial\bar{z}}\bar{w}e^{-2\nu}\right] - \tfrac{1}{4}e^{\nu}(e^{-3\nu}a)_{,1}, \quad (2.24)$$

the bars denote complex conjugates.

3. ELECTROMAGNETIC SOLUTIONS

Since a is an arbitrary function of the single variable x^1, a reasonable *ad hoc* condition is to require that it be zero. Hence, we consider the equation

$$\frac{\partial^2}{\partial z\partial\bar{z}}\left(e^{2\nu}\frac{\partial^2\nu}{\partial z\partial\bar{z}}\right) = -\tfrac{1}{4}. \quad (3.1)$$

The general solution of (3.1) would involve four arbitrary functions; the dependence on two of these functions can be easily obtained. The first integration yields

$$\partial^2\nu/\partial z\partial\bar{z} = e^{-2\nu}[-\tfrac{1}{4}z\bar{z} + f(z) + g(\bar{z})], \quad (3.2)$$

where f and g are arbitrary. If $f(z)$ and $g(\bar{z})$ are chosen to be identically zero, then (3.2) may be written

$$\partial^2\nu/\partial z\partial\bar{z} = -\tfrac{1}{4}\exp(-2\nu + \log z\bar{z}). \quad (3.3)$$

If one now puts

$$-2\nu + \log\tfrac{1}{2}z\bar{z} = \varphi, \quad (3.4)$$

then (3.4) reduces to the Liouville equation[2]

$$\partial^2\varphi/\partial z\partial\bar{z} = e^{\varphi}. \quad (3.5)$$

[2] A. R. Forsyth, *Theory of Differential Equations* (Dover Publications, Inc., New York, 1959), Vol. 6, p. 143.

The general solution to this equation is

$$e^{\varphi} = 2F'G'/(F + G)^2, \tag{3.6}$$

where F and G are arbitrary functions of z and $\bar{z}$, respectively. Hence, a particular solution of (3.1) is

$$e^{2\nu} = z\bar{z}(F + G)^2/4F'G'. \tag{3.7}$$

4. GRAVITATIONAL SOLUTIONS

If we require that R_{ij} be identically zero then the metric will describe a gravitational field. The only modification that the field equations will require is that the term $\frac{1}{4}$ must be removed from (2.24). If it is again assumed that a is identically zero then (2.24) reduces to

$$\frac{\partial^2}{\partial z \partial \bar{z}}\left(e^{2\nu}\frac{\partial^2 \nu}{\partial z \partial \bar{z}}\right) = 0. \tag{4.1}$$

The first integral of this equation is

$$\partial^2\nu/\partial z\partial\bar{z} = e^{-2\nu}[f(z) + g(\bar{z})]. \tag{4.2}$$

This equation can also be transformed into Liouville's equation if either f or g is zero. For example, suppose that g is zero. Put

$$\varphi = -2\nu + \log [2f(z)], \tag{4.3}$$

and obtain

$$\partial^2\varphi/\partial z\partial\bar{z} = -e^{\varphi}. \tag{4.4}$$

The general solution is

$$e^{\varphi} = -2F'G'/(F + G)^2. \tag{4.5}$$

Hence,

$$e^{2\nu} = -4f(F + G)^2/F'G', \tag{4.6}$$

where F and G are arbitrary functions of z and $\bar{z}$, respectively.

Finally, if neither f nor g is zero then one can verify that another particular solution of (4.2),

$$e^{2\nu} = -\tfrac{2}{3}(f + g)^3/f'g'. \tag{4.7}$$

5. SUMMARY OF RESULTS

Let $u(x^1, x^3, x^4)$ and $v(x^1, x^3, x^4)$ be arbitrary solutions of the Cauchy–Riemann equations

$$u_{,3} = v_{,4}, \quad u_{,4} = -v_{,3}, \tag{5.1}$$

and suppose f, F, and G are arbitrary functions of one variable. The metric defined by

$$g_{ij} = \begin{pmatrix} 2\alpha & 1 & \beta & \gamma \\ 1 & 0 & 0 & 0 \\ \beta & 0 & (x^2)^2e^{-2\nu} & 0 \\ \gamma & 0 & 0 & (x^2)^2e^{-2\nu} \end{pmatrix}, \tag{5.2}$$

where

$$2\alpha = -e^{2\nu}\nabla^2\nu - (u_{,3} + v_{,4} - 2uv_{,3} - 2vv_{,4})x^2 + (u^2 + v^2)(x^2)^2, \tag{5.3}$$

$$\beta = (x^2)^2ue^{-2\nu}, \tag{5.4}$$

$$\gamma = (x^2)^2ve^{-2\nu}, \tag{5.5}$$

determines:

(i) an electromagnetic field if

$$e^{2\nu} = \frac{[(x^3)^2 + (x^4)^2][F(x^3 + ix^4) + G(x^3 - ix^4)]^2}{4F'(x^3 + ix^4)G'(x^3 - ix^4)}; \tag{5.6}$$

(ii) a gravitational field if

$$e^{2\nu} = \frac{-4f(x^3 \pm ix^4)[F(x^3 + ix^4) + G(x^3 - ix^4)]^2}{F'(x^3 + ix^4)G'(x^3 - ix^4)}; \tag{5.7}$$

(iii) a gravitational field if

$$e^{2\nu} = \frac{-\frac{2}{3}[F(x^3 + ix^4) + G(x^3 + ix^4)]^3}{F'(x^3 + ix^4)G'(x^3 + ix^4)}. \tag{5.8}$$

In conclusion we would like to express regret of the lack of physical interpretations in this paper. However, it is hoped that these solutions may eventually yield some insight into the physical side of the picture. The solutions seem to have the unsatisfactory property that they are not asymptotically flat but it is possible that a coordinate transformation may remedy this apparent defect.

ACKNOWLEDGMENTS

We would like to thank the Canadian Mathematical Congress for their support of this work. We would also like to express our appreciation to B. E. Smith for assisting with the calculation of the field equations and to D. M. Aitken for suggesting solution (5.8).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Scattering from Coulomb-Like Potentials without Divergences or Cutoffs. I. Formalism: Generalized Lippmann–Schwinger Equations*

GEOFFREY B. WEST†
Institute of Theoretical Physics, Department of Physics, Stanford University, Stanford, California

(Received 6 October 1966)

We present a general method whereby integral equations, which govern the solution to the radial Schrödinger equation, can be derived. The equations are quite flexible and can be arranged so as to define divergenceless iterative procedures. They are therefore particularly useful for Coulombic and singular potentials, neither of which are normally covered by the standard theory. We have carefully examined the standard theory as applied to Coulombic potentials; this is generally characterized by logarithmic divergences. We show that the standard Lippmann–Schwinger equation is no longer applicable and must be replaced by the corresponding *homogeneous* integral equation. Furthermore, we demonstrate that the T matrix, as conventionally defined, vanishes. Upon application of our generalized integral equation, a perturbative result for the scattering amplitude is obtained without the appearance of divergences or cutoffs. In the case of a pure point Coulomb potential, this result agrees very favorably with the exact one. In the modified case, our expression, by virtue of the fact that it does not require knowledge of Coulomb wavefunctions, is much simpler for computational purposes than the standard expressions. One simple example of the application of this method to singular potentials is briefly discussed.

I. INTRODUCTION

IN this paper we present a method for obtaining integral equations, more general than the usual Lippmann–Schwinger (L–S) equation, which govern the solution to the radial Schrödinger equation.[1] The new integral equations are of particular use for those potentials which are not normally covered by the standard theory. These potentials are the Coulombic potentials,[2] i.e., those which behave like $1/r$ as $r \to \infty$, and also potentials which are more singular than $1/r^2$ at the origin. Although one example of the latter class is briefly discussed (see Appendix D), in this paper we are concerned almost entirely with the Coulombic class. Our aim is to obtain integral equations which are amenable to an iterative procedure without introducing divergences. The discussion is limited to scattering solutions only; the extension to bound-state problems appears to be straightforward.

Although scattering of charged particles by a point Coulomb potential can be solved exactly,[3] divergences of a logarithmic nature characteristically arise when the problem is attacked via the Born expansion. These divergences are normally associated with the "infinite extent" of the potential and the distorting effect it has upon the asymptotic form of the scattering solutions,[4,5] (we clarify this in Sec. III). The first Born term, although not strictly convergent, does give the correct magnitude for the scattering amplitude. This observation led Dalitz to the conjecture that the rest of the Born series can be summed to give an unobservable phase factor.[4,5] A proof of this conjecture has recently been indicated by Weinberg,[6] who used Feynman parametrization techniques, similar to those used in quantum electrodynamics, to treat the divergent integrals.

In many problems of a practical nature, the pure Coulomb potential is modified near the origin by a strong short-range potential. This problem may be solved exactly in terms of phase shifts ζ_l (see Appendix C).[3] To obtain this result from the standard L–S approach or from any approach which demands non-Coulombic asymptotic forms, cutoffs, either in the range of the potential or in the number of partial waves, are generally employed.[7,8] These phase shifts ζ_l must be defined in terms of a Coulombic asymptotic form and must therefore be calculated using Coulomb wavefunctions rather than the simpler spherical Bessel functions. This can be a laborious and cumbersome

* The work was supported in part by the U.S. Air Force through the Air Force Office of Scientific Research, Contract AF 49(638)-1389.

† Present address: Laboratory of Nuclear Studies, Cornell University, Ithaca, New York.

[1] For a review of the standard theory see, e.g., T. Wu and T. Ohmura, *Quantum Theory of Scattering* (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1963), pp. 1–56.

[2] The words Coulombic and Coulomb-like are taken to be synonymous here; they refer only to the behavior at infinity. The behavior at zero is always taken to be nonsingular unless otherwise stated.

[3] L. I. Schiff, *Quantum Mechanics* (McGraw-Hill Book Company, Inc., New York, 1955), 2nd ed., pp. 114–121.

[4] R. H. Dalitz, Proc. Roy. Soc. (London) **A206**, 509 (1951).

[5] C. Kacser, Nuovo Cimento **13**, 303 (1959).

[6] S. Weinberg, Phys. Rev. **140**, B516 (1965).

[7] M. Goldberger and K. Watson, *Collision Theory* (John Wiley & Sons, Inc., New York, 1964), pp. 259–269.

[8] J. T. Holdeman and R. M. Thaler, Phys. Rev. **139**, B1186 (1965); W. Ford, *ibid.* **133**, B1616 (1964).

task. One might hope that the problem could be formulated in such a way that one need only evaluate the purely strong interaction phase shifts δ_l defined with respect to asymptotic plane waves, together with relatively simple integrals over the Coulomb potential. This approach has been suggested by Schiff in another context.[9] He was particularly interested in setting up a formalism whereby the electromagnetic structure of the pion could be most easily extracted from $\pi^{\pm} - \alpha$ scattering data, as first suggested by Hofstadter.[9] Schiff found, in a fashion directly analogous to the point Coulomb problem, that if a perturbation expansion be attempted in the Coulomb coupling strength, logarithmic divergences result. To circumvent this difficulty the use of a physical cutoff was proposed. Subsequently a method was suggested by which the result could be made finite and cutoff independent, at least in first order.[10] Antoine[11] has recently shown that it appears impossible to generalize this method consistently to all orders without introducing a model to describe the short-range modification of the Coulomb potential. We review and give a critical discussion of this approach in Sec. II.

In Sec. III we carefully examine the L–S equation for a Coulombic potential. We show how the integrals can be made convergent, in a consistent fashion, by defining either a summability convention or by choosing a definite contour in the cut complex r plane. In this way we are led to a *homogeneous* integral equation which we claim contains both the scattering and bound-state solutions for Coulombic problems. Although this equation may be of interest in other contexts,[12] it is of limited value for computational purposes. An interesting result of this analysis is contained in Appendix A, where we show that the standard T matrix is zero for all Coulombic potentials! Since the standard formalism demands asymptotic plane waves, this result is not quite as surprising as it might first appear.

In Sec. IV, motivated by techniques developed in Sec. III, we introduce a generalized integral equation for scattering. This equation contains two constants and two arbitrary functions. The two constants are related to the boundary conditions while the arbitrary functions may be chosen so as to cancel out any divergence introduced into the integrals by the potential. By a judicious choice of these functions, we are able to write down a Fredholm equation for the Coulombic potential which will give finite results when iterated. We do not discuss the convergence of this Fredholm series. We point out how our method may be used to derive, in a very simple manner, partial-wave Green's functions. An example of this is given in Appendix C, where the Coulomb Green's function is derived.

An approximate expression is obtained for the point Coulomb scattering amplitude which compares favorably with the exact result. This is generalized to situations where the Coulomb potential is strongly modified at the origin. The resulting scattering amplitude is almost identical to that obtained from the standard approach as discussed in Sec. II. It differs in two respects: first, it is completely free from divergences and second, it gives a different magnitude for the cross section in the regions of forward scattering; this latter difference is not of great importance in practical situations. The result is readily applicable to computational problems.

We hope in a later paper to illustrate its usefulness by applying it to analyze recent data[13] on $\pi^{\pm} - \alpha$ scattering in order to obtain the form factor of the pion.

II. DIVERGENCES

1. Example

We motivate our approach by a review of Schiff's method[9] for treating a Coulombic potential V strongly modified at the origin by a (nuclear) potential U. We require the following definitions and equations.

A free wave of wavenumber k, energy $E = k^2/2m$, is denoted by χ and the free Green's function by G_0. The outgoing wave solution for the nuclear part is given by the solution to the L–S equation

$$\phi^+ = \chi + G_0 U \phi^+. \tag{1}$$

Subscripts i and f refer to initial and final states with propagation vectors $\mathbf{k}_i$ and $\mathbf{k}_f$, respectively. In terms of a partial-wave expansion

$$\phi_i^+(\mathbf{r}) = \sum_{l=0}^{\infty} (2l+1) i^l e^{i\delta_l} R_l(r) P_l(\cos\theta_i). \tag{2}$$

Here θ_i is the angle between $\mathbf{k}_i$ and $\mathbf{r}$. Similarly θ_f is the angle between $\mathbf{k}_f$ and $\mathbf{r}$, and θ that between $\mathbf{k}_i$ and $\mathbf{k}_f$.

The scattering amplitude $f_n^+(\theta)$ is defined from the asymptotic form of Eq. (2) (an arrow always denotes an asymptotic form):

$$\phi_i^+(\mathbf{r}) \to e^{ikr_i\cos\theta_i} + (e^{ikr}/r) f_n^+(\theta_i). \tag{3}$$

[9] L. I. Schiff, Progr. Theoret. Phys. (Kyoto), Suppl. Extra Number, 400 (1965); R. Hofstadter and M. M. Sternheim, Nuovo Cimento **38**, 1854 (1965).

[10] This procedure has been suggested by J. S. Bell and independently by F. Gross and D. R. Yennie (private communication).

[11] J. P. Antoine, Nuovo Cimento **44**, 1068 (1966).

[12] For example, in discussing the analytic properties of the solution, see S. Okubo and D. Feldman, Phys. Rev. **117**, 292 (1960); R. A. Mapleton, J. Math. Phys. **2**, 478 (1963); and **3**, 297 (1963).

[13] M. E. Nordberg and K. F. Kinsey, Phys. Letters **20**, 692 (1966).

Demanding

$$R_l(r) \to (1/\rho)\sin(\rho - \tfrac{1}{2}l\pi + \delta_l) \quad (\rho \equiv kr), \quad (4)$$

we find

$$f_n^+(\theta_i) = \frac{1}{2ik}\sum_{l=0}^{\infty}(2l+1)(e^{2i\delta_l} - 1)P_l(\cos\theta_i). \quad (5)$$

For the solution to the complete problem, we similarly have

$$\psi^+ = \chi + G_0(U + V)\psi^+ \quad (6)$$

and

$$\psi_i^+(\mathbf{r}) = \sum_{l=0}^{\infty}(2l+1)i^l e^{i(\eta_l+\zeta_l)}S_l(r)P_l(\cos\theta_i). \quad (7)$$

Knowing

$$S_l(r) \to (1/\rho)\sin(\rho - \tfrac{1}{2}l\pi - n\ln 2\rho + \eta_l + \zeta_l), \quad (8)$$

we have

$$f_{\text{tot}}(\theta_i) = \frac{1}{2ik}\sum_{l=0}^{\infty}(2l+1)e^{2i(\eta_l+\zeta_l)}P_l(\cos\theta_i). \quad (9)$$

Here the Coulomb coupling strength n is Ze^2m/k, where Z is the nuclear charge, and the pure Coulomb phase shifts are defined to be

$$\eta_l = \arg\Gamma(l + 1 + in).$$

Directly from Eqs. (1) and (6), one can obtain the usual expression for scattering from two potentials. The complete T matrix, defined by

$$T_{fi} \equiv \langle\chi_f|\, U + V\,|\psi_i^+\rangle, \quad (10)$$

is[14]

$$T_{fi} = \langle\phi_f^-|\, U\,|\chi_i\rangle + \langle\phi_f^-|\, V\,|\psi_i^+\rangle. \quad (11)$$

Using (3), the Legendre polynomial expansion for a plane wave and the addition theorem for Legendre polynomials, one can show that the T matrix for pure nuclear scattering is

$$\langle\phi_f^-|\, U\,|\chi_i\rangle = -4\pi f_n^+(\theta). \quad (12)$$

Using the formal equation

$$\psi^+ = \phi^+ + (E - H_0 - U - V + i\epsilon)^{-1}V\phi^+, \quad (13)$$

the second term in Eq. (11) is iterated. The first-order term is

$$\langle\phi_f^-|\, V\,|\phi_i^+\rangle = 4\pi\sum_{l=0}^{\infty}(2l+1)P_l(\cos\theta) \times\int_0^{\infty}V(r)e^{2i\delta_l}R_l^2(r)r^2\,dr. \quad (14)$$

Each integral in (14) diverges logarithmically. That this divergence does not arise from an unjustifiable substitution of ϕ_i^+ for ψ_i^+ in Eq. (11) was verified by Schiff. He noted that the exact result is

$$\langle\phi_f^-|\, V\,|\psi_i^+\rangle = 4\pi\sum_{l=0}^{\infty}(2l+1)P_l(\cos\theta) \times\int_0^{\infty}V(r)e^{i(\delta_l+\eta_l+\zeta_l)}R_lS_lr^2\,dr. \quad (15)$$

If we multiply the differential equation[15] for R_l by r^2S_l and that for S_l by r^2R_l, subtract and integrate we obtain

$$\int_a^b V(r)R_lS_lr^2\,dr = [W(R_l, S_l)]_a^b, \quad (16)$$

where a and b are arbitrary points and where we have defined[16]

$$W(R_l, S_l) \equiv r^2[R_lS_l' - S_lR_l']. \quad (17)$$

Using (4) and (8) and the fact that R_l and S_l are regular at the origin we obtain for the integral occurring in (15) the ambiguous result

$$\lim_{r\to\infty}\frac{1}{k}\sin(n\ln 2\rho - \eta_l + \zeta_l + \delta_l). \quad (18)$$

Schiff argues that this cannot oscillate to zero since that would imply no scattering. Expanding (18) in powers of n, we obtain the logarithmic divergence found in Eq. (14). Hence one seems forced to accept a cutoff for the integrals occurring in that equation.

2. Removal of Divergence

The following procedure for removing the divergences in first order has been suggested.[10] Schiff noted that subtracting the first Born Coulomb amplitude in the guise of an infinite sum of $j_l^2(\rho)$ does not remove the divergence. This is equivalent to adding and subtracting a term

$$\tilde{f}_c \equiv \int_0^{\infty}V(r)j_0(qr)r^2\,dr; \quad q \equiv 2k\sin^2(\tfrac{1}{2}\theta). \quad (19)$$

The integrand in (14) now contains the factor $[e^{2i\delta_l}R_l^2(r) - j_l^2(kr)]$. This has the advantage that only terms with $\delta_l \neq 0$ occur and one can therefore expect rapid convergence of the l sum. However, consider the asymptotic form of this factor:

$$e^{2i\delta_l}R_l^2(r) - j_l^2(kr) \to (1/4\rho^2)[e^{2i\rho}(e^{4i\delta_l} - 1)(-1)^{l+1} + 2(e^{2i\delta_l} - 1)]. \quad (20)$$

[14] Reference 7, p. 256.

[15] See Appendix B.

[16] We sometimes refer to W as a Wronskian although our definition differs from the conventional one by the inclusion of the r^2. Note also that a prime denotes differentiation with respect to the variable r ($' \equiv d/dr$), except, of course, when it is used on the symbol r itself; in that case it refers only to a spatial point r' to be distinguished from the point r.

It is the last term $\propto (e^{2i\delta_l} - 1)$ which causes the divergence, so we subtract it. It can be resummed to give a term proportional to $f_n^+(\theta)$. Hence (14) may be rewritten as

$$\frac{1}{4\pi}\langle\phi_f^-|\, V\,|\phi_i^+\rangle = \sum_{l=0}^{\infty}(2l+1)P_l(\cos\theta) \times \int_0^\infty V(r)\left[R_l^2 e^{2i\delta_l} - j_l^2 - \frac{(e^{2i\delta_l}-1)}{2\rho^2}\right] r^2\,dr + \tilde{f}_c + F, \tag{21}$$

where

$$F \equiv 2ikf_n^+(\theta)\frac{1}{2k^2}\int_0^\infty V(r)\,dr. \tag{22}$$

Although F diverges, it has two important properties: (1) it is purely imaginary and (2) it is linear in both V and $f_n^+(\theta)$. Hence when the cross section ($\propto |T_{fi}|^2$) is computed and only terms linear in V are kept, the divergence (22) does not appear. We can therefore ignore the factor F in Eq. (21) and write

$$f_{\text{tot}}(\theta) = f_n^+(\theta) - \tilde{f}_c(\theta) - \sum_{l=0}^{\infty}(2l+1)P_l(\cos\theta) \times \int_0^\infty V(r)\left[R_l^2 e^{2i\delta_l} - j_l^2 - \frac{(e^{2i\delta_l}-1)}{2\rho^2}\right] r^2\,dr. \tag{23}$$

3. Remarks

The following comments and objections are relevant to the above:

(a) Sums and integrals have been freely interchanged although we do not have absolute convergence.

(b) The Born Coulomb term $\tilde{f}_c$, although not divergent, is, as it now stands, not a well-defined quantity.

(c) Similarly, (18) is not well defined but it is *not* divergent. Hence the T matrix, Eq. (11), is not a divergent quantity, although the perturbative result unambiguously is.

(d) It is possible to write an expression for the exact T matrix [Eq. (10)] in terms of partial waves in a fashion analogous to Eq. (16). One can then show directly that Eq. (11) is no longer valid.

(e) Equation (11) is based upon the assumption that initially we have free plane waves χ. Yet, for the final state we have used a wave of very different asymptotic form. Hence to be consistent, either a cutoff should be introduced from the very beginning or a formalism set up with distorted asymptotic waves built in.

(f) Finally, concerning the technique suggested for removal of the divergence, we feel that it cannot be entirely satisfactory until one has shown that it gives unique and consistent results and that there are no observables contained in the neglected terms. The very appearance of divergences sheds doubt upon the validity of the perturbation expansion for small n and large distances. The work of Antoine indicates that it is extremely difficult to substantiate this result in general even with recourse to models for U and V.[11] Note, furthermore, that the introduction of the $1/2\rho^2$ term into the integrand for the case of a pure Coulomb potential introduces a divergence at the origin which is not easily removed.

We complete this section by showing that a treatment of the exact Eq. (15) in a fashion similar to the subtraction procedure outlined above leads to a result differing from the result quoted there, Eq. (23). This demonstration proves to be instructive.

The only difference to first order in V is to replace F by

$$F' = 2ik(\tfrac{1}{2}[f_n K + f_{\text{tot}}\bar{K}]),$$

where

$$K \equiv \frac{1}{2k^2}\int_0^\infty V(r)e^{-in\ln 2\rho}\,dr \tag{24}$$

and $\bar{K}$ is the complex conjugate of K. We now show that not only is K *not* divergent but that it has an *imaginary* contribution of *zeroth* order. Although not necessary, we take for simplicity

$$V(r) = \begin{cases} g/a, & r < a, \\ g/r, & r > a, \end{cases}$$

where $g \equiv 2nk$. Then

$$K = \frac{n}{k}\frac{(2ka)^{-in}}{1-in} + \frac{n}{k}\int_a^\infty e^{-in\ln 2\rho}\frac{dr}{r}. \tag{25}$$

In the second term, we make the replacement $y = \log 2\rho$. The integral then becomes

$$\frac{n}{k}\int_{\ln 2ka}^\infty e^{-iny}\,dy. \tag{26}$$

This is precisely the class of integral to which $\tilde{f}_c$ belongs. In the usual sense of a Riemannian infinite integral, it is "almost" convergent. The strongest statement concerning the convergence of (26) is that it is Césaro-1 summable.[17] This implies that the limit

$$\lim_{\lambda\to\infty}\int_a^\lambda\left(1 - \frac{|y|}{\lambda}\right)(e^{-iny})\,dy \tag{27}$$

exists. The weakest statement and the one usually employed by physicists is that the integral can be made convergent by the use of an integrating factor

[17] E. C. Titchmarsh, *Introduction to the Theory of the Fourier Integral* (Oxford University Press, New York, 1962), Chaps. I and III, Sec. 10.

$e^{-\lambda|y|}$. This corresponds to Césaro-∞ convergence. The consistency of the Césaro technique and several theorems concerning its use are discussed at some length by Titchmarsh. By definition we employ Césaro-∞ summability to define such integrals. We use this rather than the stronger statement (27) purely as a matter of convenience. An equivalent procedure is to give n a small positive imaginary part λ which is taken to zero after integration. Another alternative for such integrals is to define a ray contour in the complex y plane; this contour is rotated from the real axis through the infinitesimal angle λ about the point $y = 0$. (The corresponding contour in the cut complex r plane is a somewhat complicated spiral.)

Under these conditions we find that

$$K = -\frac{i}{k}\frac{(2ka)^{-in}}{1 - in}$$
$$= -\frac{i}{k}[1 + in(1 - \ln 2ka) + O(n^2)].$$

Notice that had we expanded (24) in powers of n, each integral would have contained the logarithmic divergence, *even* under the above mentioned integration conventions.

The result we obtain for f_{tot} in this manner is

$$-in(1 - \ln 2ka)[f_{\text{tot}}(\theta) + f_n(\theta)]$$
$$= \sum (2l + 1)P_l(\cos\theta)$$
$$\times \int_0^\infty V(r)\left[R_l^2 e^{2i\delta_l} - j_l^2 - \frac{1}{2\rho^2}(e^{2i\delta_l} - 1)\right] r^2\, dr. \tag{28}$$

This differs considerably from the previous result. It is, in fact, possible to show that if (28) is to agree with (23) then, to second order in n, $|f_{\text{tot}}|^2 = |f_n|^2$. This equality is clearly not valid and its "derivation" serves only to demonstrate that great caution must be taken when manipulating infinite quantities.

Finally it should be noted that the techniques we have discussed here are intimately related to the definition of the δ function. As an example of this consider the case

$$V(r) = g/r \equiv V_0(r) \tag{29}$$

in Eq. (24). We then have

$$K = \frac{n}{k}\int_{-\infty}^{\infty} e^{-iny}\, dy$$
$$= (n/k)\delta(n), \quad \text{conventionally.}$$

The contour here must be defined above the real axis on the right-hand plane and below the real axis on the left (alternatively allow $n \to n + i\lambda sgny$). This result makes it even more apparent why a simple power series expansion breaks down; K is clearly not an analytic function of n. [Note that although K is now zero, (23) has lost its usefulness in that there is now a divergence in the integrand coming from the $(e^{2i\delta_l} - 1)/2\rho^2$ term.]

III. A NEW INTEGRAL EQUATION

1. Standard Formalism

We have seen that applying the standard formalism to Coulombic problems leads to several ambiguities and paradoxes. It would therefore seem appropriate to examine some of the underlying assumptions of the basic equations.

The simplest argument for the use of an integral equation such as (1) is that the right-hand side is certainly a formal solution to the corresponding differential equation, hence if the integral exists the equation must be correct. The inhomogeneous term χ is included to ensure the correct asymptotic form.

Using the asymptotic form of the free Green's function $G_0(\mathbf{r}, \mathbf{r}')$:

$$G_0(\mathbf{r}, \mathbf{r}') \to \frac{e^{-ikr}}{4\pi r} e^{-ik\hat{\mathbf{r}}\cdot\mathbf{r}'}\{1 + O(r'/r)\}$$

one deduces from the asymptotic form of Eq. (1) the relationship of the T matrix to the scattering amplitude [see Eq. (12)]. An alternative way of seeing this and one that is of particular interest to us is to examine the integral

$$I(\mathbf{r}) \equiv \int G_0(\mathbf{r}, \mathbf{r}')U(\mathbf{r}')\phi^+(\mathbf{r}')\, d\mathbf{r}'. \tag{30}$$

We require the partial-wave expansion of $G_0(\mathbf{r}, \mathbf{r}')$

$$G_0(\mathbf{r}, \mathbf{r}') = \frac{-ik}{4\pi}\sum_{l=0}^{\infty}(2l + 1)P_l(\cos\theta_1)$$
$$\times j_l(kr_<)h_l^{(1)}(kr_>), \tag{31}$$

where θ_1 is the angle between $\mathbf{r}$ and $\mathbf{r}'$; $r_{\substack{>\\(<)}}$ means that the larger (or smaller) of r, r' is to be taken. We then have

$$I(\mathbf{r}) = -ik\sum_{l=0}^{\infty}(2l + 1)P_l(\cos\theta)$$
$$\times i^l e^{i\delta_l}[h_l^{(1)}J_1(r) + j_l J_2(r)], \tag{32}$$

where

$$J_1(r) \equiv \int_0^r j_l(\rho')U(r')R_l(r')r'^2\, dr' \tag{33}$$

and

$$J_2(r) \equiv \int_r^\infty h_l^{(1)}(\rho')U(r')R_l(r')r'^2\, dr'. \tag{34}$$

For scattering solutions, we are interested in those $J_{1,2}(r)$ which, to $O(1/r)$, behave asymptotically like constants. Using the well-known asymptotic forms of R_l, $h_l^{(1)}$, and j_l [see Eq. (4)], it is easily verified that for potentials which fall off faster than $1/r$ this is indeed the case. In fact, to $O(1/r)$, we find $J_2(r) \to 0$ and

$$J_1(r) \to \int_0^\infty j_l(\rho')U(r')R_l' r'^2\, dr' \equiv J_1(\infty), \quad \text{a const.} \tag{35}$$

Whence we obtain

$$I(\mathbf{r}) \to \frac{e^{ikr}}{r}\sum_{l=0}^{\infty}(2l+1)P_l(\cos\theta)e^{i\delta_l}J_1(\infty) = -(e^{ikr}/4\pi r)\langle\chi_f|\, U\, |\phi_i^+\rangle. \tag{36}$$

This verifies the connection between the T matrix and the scattering amplitude [see Eq. (12)].

This discussion gives a very definite meaning to the phrase "infinite extent" as applied to a potential. Potentials of the form $1/r^n$ $(n > 1)$ certainly extend to infinity and cause scattering at asymptotic distances. However, these asymptotic scatterings do not affect the asymptotic form of the scattering solutions to $O(1/r)$ and the result of Eq. (35) is true. The potential $1/r^n$ is therefore to be considered of finite extent. For the Coulomb potential we note that (35) is not true.

2. Coulomb Potentials

For Coulomb potentials we have seen that, even if the L–S equation were to hold, the T matrix is not the scattering amplitude. In fact, we noted that the T matrix is not a well-defined quantity, but suffers an ambiguity of the form of Eq. (18).

Let us examine $I(\mathbf{r})$, as defined by Eq. (30), for the point Coulomb potential V_0, Eq. (29). We limit ourselves for the moment to asymptotic values of $I(\mathbf{r})$.

Using a Wronskian relationship analogous to (16) we find that

$$J_1(r) = W(j_l, F_l) \to (1/k)\sin(n\ln 2\rho - \eta_l), \tag{37}$$

where F_l is the point Coulomb radial wavefunction. Clearly this technique cannot be used to evaluate $J_2(r)$. Instead we replace the integrand in (34) by its asymptotic form. We then require two types of integrals:

$$J_{2A} \equiv \int_r^\infty e^{2i\rho' - in\ln 2\rho'}\frac{dr'}{r'} \tag{38}$$

and

$$J_{2B} \equiv \int_r^\infty e^{in\ln 2\rho'}\frac{dr'}{r'}, \tag{39}$$

where

$$J_2(r) = (n/k)[(-1)^{2l+2}e^{i\eta_l}J_{2A} + e^{-i\eta_l}J_{2B}]. \tag{40}$$

It is easy to see by successive integration by parts that $J_{2A}(r) \to O(1/r)$, and so can be neglected. The integral $J_{2B}(r)$ is precisely of the form discussed in Sec. II [see Eq. (25)]. Using the procedure defined there we obtain

$$J_{2B}(r) \to (i/n)e^{in\ln 2\rho}.$$

We therefore find that

$$I(\mathbf{r}) \to -\frac{1}{\rho}\sum_{l=0}^{\infty}(2l+1)P_l(\cos\theta) \times e^{i\eta_l}i^l\sin(\rho - n\ln 2\rho - \tfrac{1}{2}l\pi + \eta_l). \tag{41}$$

This is exactly the asymptotic form of the point Coulomb wavefunction. This suggests that we do not require an inhomogeneous term χ in the integral equation but rather that the homogeneous equation can, by itself, generate scattering solutions. We now indicate a very simple proof of this conjecture.

We have to show that

$$F_l(r) = -ik[h_l^{(1)}(\rho)J_1(r) + j_l(\rho)J_2(r)] \tag{42}$$

for arbitrary r. We have already noted that

$$J_1(r) = W(j_l, F_l).$$

For $J_2(r)$ we divide the range $[r, \infty]$ into $[r, R] + [R, \infty]$ where we choose R so large that asymptotic values may be used. For the range $[r, R]$ we use the Wronskian relationship; for the range $[R, \infty]$ we replace the integrand by its asymptotic form. Carrying this out explicitly [cf., Eqs. (37)–(41)], one finds that to $O(1/R)$ the ambiguous R dependence cancels out. We can now make the R dependence arbitrarily small by allowing R to become arbitrarily large. We are then left with $J_2(r) = W(h_l^{(1)}, F_l)$ evaluated at r.

Hence the right-hand side of (41) is

$$h_l^{(1)}W(j_l, F_l) - j_l W(h_l^{(1)}, F_l) = F_l W(j_l, h_l^{(1)}) \tag{43}$$

identically.

Furthermore, $W(j_l, h_l^{(1)}) = i/k$. Q.E.D.

3. Remarks

(a) We have shown that the scattering solutions to the point Coulomb potential u_c^+ satisfy a homogeneous integral equation:

$$u_c^+(\mathbf{r}) = \int G_0(\mathbf{r}, \mathbf{r}')V_0(r')u_c^+(\mathbf{r}')\, d\mathbf{r}'. \tag{44}$$

It should be noted that the kernel in Eq. (44) is not

square-integrable and therefore the Fredholm alternative does not apply.[18] This theorem would say that either there is a trivial solution, $u_c = 0$, or there are solutions only for certain eigenvalues. It is, in fact, possible to show that (44) also gives the bound-state solutions.[19] Note that upon letting $n \to 0$ after carrying out the integration, the expected free wave is obtained.

(b) One can show that the way in which the infinite integral has been evaluated is consistent in the sense that the final result is independent of both the number of times, and the order in which, one breaks up the region of integration. We emphasize that the convention we have adopted for the meaning of this integral must be observed at all times. In Appendix A we give an interesting example of what can happen if this convention is broken.

(c) The arguments given to derive (44) could, of course, have been given for the case of a regular potential. The only difference occurs when one evaluates $J_2(r)$. To $O(1/R)$ the usual L–S equation results; as in the Coulomb case the R dependence can be made arbitrarily small by making R arbitrarily large.

(d) An equation of the form (44) also holds for the solution to the potential $(U + V)$. This equation is not very useful for computational purposes. One reason for this is that homogeneous equations must generally be solved by some self-consistent method of successive approximations. Another reason is the difficulty associated with the definition of the scattering amplitude, $f_{\text{tot}}(\theta)$. As we have already remarked above, $J_{1,2}(r)$ are not asymptotically constants. This means that to compute $f_{\text{tot}}(\theta)$ we must have a method which automatically extracts from the integrals the distorted phase factor. If this is not the case, we then have the undesirable task of having to compute indefinite integrals. Finally, we note that (40) essentially implies that $J_1(\infty) = 0$. This is verified in Appendix A. This means that if the usual definition of the T matrix is employed there will be no scattering! This strange result can be understood if $|T_{fi}|^2$ is interpreted as the probability of scattering particles from the plane-wave state i to the plane wave state f by a potential V. For any Coulombic potential this must be zero by virtue of its unique asymptotic form.

With the above discussion in mind we have investigated other possible ways of setting up integral equations. This is discussed in the following section.

IV. GENERALIZED EQUATIONS FOR SCATTERING

1. General Potential[20]

The use of the Wronskian identity (43) in proving that the Coulomb solution obeys an homogeneous integral equation can be generalized to give other types of integral equations for an arbitrary potential. It also gives a very simple method for obtaining partial-wave Green's functions.

For the radial wavefunction R_l we rewrite the Wronskian identity (43) in the form

$$(i/k)R = (h^{(1)}/f_1)f_1 W(j, R) - (j/f_2)f_2 W(h^{(1)}, R), \quad (45)$$

where f_1 and f_2 are arbitrary functions of r.

We show in Appendix B that from the Schrödinger equation the following identity is true:

$$f_1 W(j, R) = [f_1 W(j, R) - f_1' r^2 j R]_a + f_1' r^2 j R - \int_a^r [\{(r^2 f_1')' - U r'^2 f_1\} j + 2 f_1' r'^2 j'] R \, dr. \quad (46)$$

Here, j and R may be interpreted as *any* solution to their respective differential equations.

Using these in Eq. (45) we obtain

$$\begin{aligned} \frac{i}{k} R\left[1 + ikr^2 h^{(1)} j\left(\frac{f_1'}{f_1} - \frac{f_2'}{f_2}\right)\right] &= \frac{h^{(1)}}{f_1}[f_1 W(j, R) - f_1' r^2 j R]_a \\ &\quad - \frac{j}{f_2}[f_2 W(h^{(1)}, R) - f_2' r^2 h^{(1)} R]_b \\ &\quad - \frac{h^{(1)}}{f_1}\int_a^r [\{(r^2 f_1')' - U r'^2 f_1\} j + 2 f_1' r'^2 j'] R \, dr' \\ &\quad - \frac{j}{f_2}\int_b^r [\{(r^2 f_2')' - U r'^2 f_2\} h^{(1)} + 2 f_2' r'^2 h^{(1)\prime}] R \, dr'. \end{aligned} \quad (47)$$

We generalize this equation even further at the end of this section.

The boundary conditions are inserted into this equation by the requirements to be imposed upon R at the points a and b which may be arbitrarily specified. The behavior of R at the points $r = 0$ and $r = \infty$ can usually be obtained from general properties of the corresponding differential equation. Rather than discuss Eq. (47) in general we mention only a few special cases.

(i) $f_1 = f_2 = \text{const}$; $U(r)$ has no singularity at

[18] S. G. Mikhlin, *Integral Equations* (Hindustani Press, Delhi, India, 1960). This reference contains several examples of homogeneous equations which generate a continuum of eigenvalues.

[19] C. C. Grosjean, "Formal Theory of Scattering Phenomena", Institut Interuniversitaires des Sciences Nucleaires Monographie No. 7, Bruxelles (1960), pp. 83–89.

[20] In this section, for ease of reading, we drop the subscript l on occasion.

the origin worse than $1/r^2$ and is of finite extent. We demand

$$R \sim j \quad \text{at} \quad r = 0, \tag{a}$$

$$R \to (1/\rho)\sin(\rho - \tfrac{1}{2}l\pi + \delta_l) \quad \text{at} \quad r = \infty. \tag{b}$$

We then obtain the usual L–S inhomogeneous equation

$$\frac{i}{k}R = je^{-i\delta_l} + h^{(1)}\int_0^r Ur'^2Rj\,dr' - j\int_r^\infty Ur'^2Rh^{(1)}\,dr'. \tag{48}$$

This immediately verifies the partial-wave expansion of $G_0(\mathbf{r}, \mathbf{r}')$, Eq. (31).

For the Coulombic case our previous discussion led us to give the value zero to the ambiguous constant $[W(h^{(1)}, R)]_{r=\infty}$. Procedures which introduce a cutoff C, eventually to be taken to infinity, are easily incorporated into this scheme by setting $b = C$. One then has the *exact* equation

$$\frac{i}{k}R = je^{-i\delta_l - in\ln 2kC} + h^{(1)}\int_0^r Ur'^2Rj\,dr' - j\int_r^C Ur'^2Rh^{(1)}\,dr'. \tag{49}$$

We feel that it is to an equation such as either (49) above or (50) below, rather than to (48) that the work such as that of Kacser[5] and Antoine[11] should be applied.

(ii) $f_1 = f_2 = \text{const}$; $U(r)$ has no singularity at the origin worse than $1/r^2$. Choose $a = b = 0$. Even for Coulombic potentials, a Volterra equation follows:

$$\frac{i}{k}R = \frac{i}{k}j - j\int_0^r Ur'^2Rh^{(1)}\,dr' + h^{(1)}\int_0^r Ur'^2Rj\,dr'. \tag{50}$$

If this is used in a perturbative treatment, we must be careful to renormalize the j_l at each stage. Because logarithmic terms will arise this is not suitable for Coulombic problems.

(iii) $f_1 = f_2$; $U(r)$ is more singular at the origin than $1/r^2$, but is of finite extent. We do not discuss this in any detail. We point out, however, that because the arbitrary functions f occur in the integrand it is often possible to choose them in such a way that they cancel out any divergences arising from the singular nature of the potential in perturbation theory. An example of this is given in Appendix D. Alternatively it is possible to choose f to be a "regularizing function", i.e., a function which effectively parametrizes the potential in such a way that for nonzero values of the parameter all integrals converge while for the zero value, the regularized potential reduces to its original form. This procedure is clearly related to the perati-zation procedure introduced by Khuri and Pais.[21] We do not discuss the connection in this paper.

2. Point Coulomb Potential

We have investigated the above cases for a Coulomb potential and have found that none of them is useful as far as a perturbative scheme is concerned. We find the only useful form to be the following.

(iv) $f_1 \neq f_2$; $U(r) = V_0(r)$; $a = 0$; $b = \infty$. We choose

$$f_1 = (2kr)^{in} \equiv e^{in\ln 2\rho}$$

and

$$f_2 = (2kr)^{-in} \equiv e^{-in\ln 2\rho}.$$

This choice is made for two reasons: (a) we want an inhomogeneous equation and so demand

$$f_2 W(h^{(1)}, F) \to \text{const} \quad \text{at} \quad r = \infty;$$

(b) these f's ensure convergence of the perturbative result.

We obtain the following equation:

$$\frac{i}{k}F[1 - grh^{(1)}j] = \frac{i}{k}j(2\rho)^{in}e^{-i\eta_l} - h^{(1)}(2\rho)^{-in} \times \int_0^r [\{(in - n^2) - gr'\}j + 2inrj'] \times F(2\rho')^{in}\,dr' - j(2\rho)^{in} \times \int_r^\infty [\{(in + n^2) + gr'\}h^{(1)} + 2inrh^{(1)\prime}] \times F(2\rho')^{-in}\,dr'. \tag{51}$$

Because of the smallness of g, the function $(1 - grh^{(1)}j)$ has no zeros on the real axis. Hence we do not introduce spurious poles if we take

$$\chi_c \equiv j_l e^{in\ln 2\rho - i\eta_l}/(1 - grh_l^{(1)}j_l) \tag{52}$$

as the inhomogeneous term in a Fredholm equation for F. Note, further, that

$$(1 - grh_l^{(1)}j_l) \to 1 \quad \text{at} \quad r = \infty.$$

Let us take χ_c as the first-order approximation for F_l. We are only interested in the asymptotic form to $O(1/r)$. In that case the second integral in (51) can be dropped and we need only consider the first integral. Keeping only terms of first order in n (or g) we find that the first integral contributes

$$e^{-i\eta_l}g\int_0^r r'j_l^2e^{2in\ln 2\rho'}\,dr' = e^{-i\eta_l}g\left(\int_0^\infty - \int_r^\infty\right)r'j_l^2e^{2in\ln 2\rho'}\,dr'. \tag{53}$$

[21] N. N. Khuri and A. Pais, Rev. Mod. Phys. 36, 590 (1964).

We evaluate these integrals in accordance with the prescription defined in Sec. II. We then obtain

$$g\int_r^\infty r' j_l^2 e^{2in\ln 2\rho'}\,dr' \to \frac{i}{2k}\,e^{2in\ln 2kr}$$

and[22]

$$g\int_0^\infty r' j_l^2 e^{2in\ln 2\rho'}\,dr'$$

$$= 2^{4in} g\,\frac{\pi}{2k^2}\,\frac{\Gamma(-2in)}{[\Gamma(\tfrac{1}{2}-in)]^2}\,\frac{\Gamma(l+1+in)}{\Gamma(l+1-in)}$$

$$\equiv (i/2k) f_l. \tag{54}$$

Substituting these results back into (51) we obtain

$$e^{i\eta_l} F_l \to \tfrac{1}{2} h_l^{(2)} e^{in\ln 2\rho} + \tfrac{1}{2} h_l^{(1)} f_l e^{-in\ln 2\rho};$$

hence

$$u_c(\mathbf{r}) = \sum_{l=0}^{\infty} (2l+1) P_l(\cos\theta) i^l F_l(r) e^{i\eta_l}$$

$$\to \sum_{l=0}^{\infty} (2l+1) P_l(\cos\theta)$$

$$\times [(-1)^l e^{-i(\rho - n\ln 2\rho)} - f_l e^{i(\rho - n\ln 2\rho)}]. \tag{55}$$

We therefore identify f_l as the S matrix; the scattering amplitude is given by [cf. Eq. (9)]

$$f_c(\theta) = \frac{1}{2ik}\sum_{l=0}^{\infty} (2l+1) P_l(\cos\theta) f_l.$$

The exact result for f_l is the well-known factor[3]

$$e^{2i\eta_l} = \Gamma(l+1+in)/\Gamma(l+1-in).$$

Using standard properties of the gamma function, f_l can be manipulated into the form

$$f_l = [\Gamma(1+2in)\cosh(\pi n) e^{-2i\eta_0}] e^{2i\eta_l}$$

$$= [1 + O(n^2)] e^{2i\eta_l}.$$

We therefore see that our approximation does not differ from the exact result until third order. An alternative but possible easier way to make the comparison is to first sum the partial-wave series for j_l^2 before performing the integral (we can now legitimately interchange sums and integrals). In that case we require the integral

$$\int_0^\infty e^{2in\ln 2\rho} \sin qr\,dr.$$

Using our previously defined contour this is found to be[23]

$$(2k/q)^{2in}[\Gamma(1+2in)/q]\cosh \pi n. \tag{56}$$

[22] A. Erdelyi *et al.*, *Tables of Integral Transforms* (McGraw-Hill Book Company, Inc., New York, 1954), Vol. II, p. 342, Eq. (24).
[23] Reference 22, Vol. I, p. 68.

This gives for the scattering amplitude,

$$[n/2k^2 \sin^2(\tfrac{1}{2}\theta)]\, e^{-in\ln\sin^2(\frac{1}{2}\theta)+i\pi}\,[\Gamma(1+2in)\cosh\pi n].$$

Since the standard result is

$$f_c(\theta) = [n/2k\sin^2(\tfrac{1}{2}\theta)] e^{-in\ln\sin^2(\frac{1}{2}\theta)+i\pi+2i\eta_0}$$

we have verified in this particular case that Σ_l and $\int dr$ may be freely interchanged.

3. Generalization

In order to consider the scattering from the potential $(V+U)$ we must generalize Eq. (47). We first note that there is nothing fundamental about starting from solutions of the Bessel equation; this was only done as a matter of convenience and because we were interested in calculational ease. The Wronskian equation (43) is an identity and is satisfied for any three functions R_l, $R_l^{(1)}$, and S_l say. Furthermore the identities (46) are similarly satisfied by any two functions R_l and S_l, say, *provided* the corresponding Schrödinger equations differ in their potentials by an amount V, say [to be inserted as U in Eq. (46)]. Choose R_l and S_l as defined in Sec. II. Define two functions $R_l^{(1)}$ and $R_l^{(2)}$ with properties similar to $h_l^{(1)}$ and $h_l^{(2)}$. It is then easy to show that S_l satisfies an equation directly analogous to (46) with j_l replaced by R_l, $h_l^{(1)}$ by $R_l^{(1)}$, U by V, and η_l by Δ_l ($\equiv \eta_l + \zeta_l - \delta_l$).

We now carry through the same type of procedure described above for the point Coulomb case and obtain to first order in g:

$$S_l \to R_l e^{in\ln 2\rho - i\Delta_l} - ikR_l^{(1)} e^{-in\ln 2\rho - i\Delta_l}$$

$$\times \int_0^r V r'^2 R_l^2 (2kr')^{2in}\,dr'. \tag{57}$$

Asymptotically, the integral can be rewritten as [cf. (53)–(54)]

$$\int_0^\infty V r^2 R_l^2 (2kr)^{2in}\,dr - \frac{i}{2k}\,e^{2in\ln 2kr}. \tag{58}$$

Hence (57) becomes

$$e^{i(\eta_l+\zeta_l)} S_l \to \tfrac{1}{2} R_l^{(2)} e^{in\ln 2\rho + i\delta_l} - \tfrac{1}{2} R_l^{(1)} e^{-in\ln 2\rho + i\delta_l}(-ik)$$

$$\times \int_0^\infty V r^2 R_l^2 (2kr)^{2in}\,dr. \tag{59}$$

By comparing this to the desired asymptotic form of ψ^+ [see Eqs. (7) and (8)] we deduce that the scattering amplitude is [see Eq. (9)]

$$f_{\text{tot}}(\theta) = -\sum_{l=0}^{\infty} (2l+1) P_l(\cos\theta) e^{2i\delta_l}$$

$$\times \int_0^\infty V r^2 R_l^2 (2kr)^{2in}\,dr. \tag{60}$$

In this form the result is not very useful. We now proceed in a manner very similar to that used in Sec. II when removing the first-order divergence. The integral is rewritten as

$$\int_0^\infty Vr^2(2kr)^{2in}\left[R_l^2e^{2i\delta_l}-j_l^2-\frac{(e^{2i\delta_l}-1)}{2\rho^2}\right]dr + \int_0^\infty j_l^2Vr^2(2kr)^{2in}\,dr+\frac{(e^{2i\delta_l}-1)}{2k^2}\int_0^\infty V(2kr)^{2in}\,dr. \tag{61}$$

The last integral can be rearranged as follows:

$$\int_0^\infty V(2kr)^{2in}\,dr=\int_0^{1/2k}V(2kr)^{2in} + \int_{1/2k}^\infty (V-V_0)(2kr)^{2in}\,dr+\int_{1/2k}^\infty V_0(2kr)^{2in}\,dr. \tag{62}$$

The first two integrals in this equation are clearly finite and well defined in the usual sense; we call their sum (*ink* γ). Expanding the factor $(2kr)^{2in}$ in the integrands of these two integrals, we see that to the lowest order in n, γ is real. The last integral in (62) can be evaluated to give

$$\int_{1/2k}^\infty ge^{2in\ln 2kr}\frac{dr}{r}=ik.$$

Collecting these results together we obtain to the first order in n:

$$f_{\text{tot}}(\theta)=f_n(1+in\gamma)-\int_0^\infty j_0(qr)Vr^2(2kr)^{2in}\,dr - \sum_{l=0}^\infty (2l+1)P_l(\cos\theta) \times\int_0^\infty Vr^2\left[R_le^{2i\delta_l}-j_l^2-\frac{(e^{2i\delta_l}-1)}{2\rho^2}\right]dr \tag{63}$$

$$=f_n(\theta)-f_c(\theta)-\int_0^\infty j_0(qr)(V-V_0)r^2\,dr - \sum_{l=0}^\infty (2l+1)P_l(\cos\theta) \times\int_0^\infty Vr^2\left[R_l^2e^{2i\delta_l}-j_l^2-\frac{(e^{2i\delta l}-1)}{2\rho^2}\right]dr. \tag{64}$$

Notice that in (63) and (64) we have expanded $(2kr)^{2in}$ only when the integral is convergent in the usual sense. The result is clearly very similar to that of Eq. (23) except that now γ is a *finite* number. We have dropped γ in going from (63) to (64) since it does not appear to first order in n when the cross section ($|f_{\text{tot}}|^2$) is computed. Another difference is to be found in the pure Coulomb term which we have implicitly separated out in Eq. (64). This contains the phase factor $e^{-in\ln\sin^2(\frac{1}{2}\theta)}$; this is only significant for forward scattering angles and is in general not important.

Two further points are worth emphasizing:

(a) Had we attempted to expand $(2kr)^{2in}$ in powers of n in Eq. (61) we would have run into the divergence problem and the result would become meaningless.

(b) The procedure outlined above is no longer valid when $V=V_0$. In that case γ *does* diverge and cannot be neglected since there is a corresponding divergence in the terms which we have kept. Equation (60) is still valid and it is only our particular rearrangement of the integral that is failing. It is not difficult to find an alternative procedure which can deal with this particular case. Since this case is not of great interest, we omit it from this discussion.

ACKNOWLEDGMENTS

The author wishes to thank Professor L. I. Schiff for a critical reading of the manuscript and for his continued encouragement throughout the course of this work. He would also like to thank Dr. L. M. Linson and Dr. M. Schick for several illuminating conversations.

APPENDIX A. SOME UNUSUAL RESULTS

We prove here that for Coulomb potentials $J_1(\infty)=0$. This implies [see Eq. (36)] that the T matrix, as usually defined, is exactly zero! We rewrite

$$J_1(\infty)=[J_1(\infty)-J_1(R)]+J_1(R)$$

and choose R asymptotically large. From (37) we have

$$J_1(R)\to(1/k)\sin(n\ln 2kR-\eta_l).$$

Using the definition of $J_1(r)$, Eq. (33), we find, upon inserting asymptotic forms into the integrand, that

$$J_1(\infty)-J_1(R)\to-(1/k)\sin(n\ln 2kR-\eta_l).$$

Making R arbitrarily large we infer that $J_1(\infty)=0$. Q.E.D.

We now show that from an erroneous use of the homogeneous integral equation, the nuclear contribution to the scattering is also zero. We show in Appendix C that the complete scattering amplitude can be written as

$$f_{\text{tot}}=f_c-(1/4\pi)\langle u_{c,f}^-|\,U+V-V_0\,|\psi_0^+\rangle.$$

We substitute into the matrix element the homogeneous equations for both u_c and ψ and obtain

$$\langle u_{c,f}^-|\,U+V-V_0\,|\psi_0^+\rangle=\langle u_{c,f}^-|\,V_0G_0^+(U+V)\,|\psi_0^+\rangle - \langle u_{c,f}^-|\,V_0\,|\psi_0^+\rangle =0. \tag{A1}$$

Since we have "proven" above that $f_c=0$, we deduce that all scattering amplitudes are zero!

This "proof" breaks down, of course, because we have written a finite quantity on the left of (A1) as the difference between two quantities, which, in the normal understanding of the bra-ket notation, would diverge. However, if these matrix elements are given their proper interpretations, namely as asymptotic forms of integrals, one can then show that the standard result, Eq. (9), follows.

APPENDIX B. DERIVATION OF THE WRONSKIAN RELATIONSHIP EQ. (46)

Let R and j be solutions to the following differential equations:

$$(1/r^2)(r^2R')' + \{k^2 - [l(l+1)/r^2]\}R = UR, \quad \text{(B1)}$$

$$(1/r^2)(r^2j')' + \{k^2 - [l(l+1)/r^2]\}j = 0. \quad \text{(B2)}$$

Multiply (B1) by fr^2R, (B2) by fr^2j, subtract and integrate from an arbitrary point a to another arbitrary point b. After performing an integration by parts, we obtain

$$\int_a^b fr^2jUR\,dr = [fW(j,R)]_a^b - \int_a^b f'W(j,R)\,dr. \quad \text{(B3)}$$

Now

$$\int_a^b f'r^2jR'\,dr = [f'r^2jR]_a^b - \int_a^b R[j(r^2f')' + f'r^2j']\,dr.$$

Choosing $b = r$ we therefore deduce the equation quoted in Sec. IV:

$$fW(j,R) = [fW(j,R) - f'r^2jR]_{r=a} + f'r^2jR - \int_a^r [\{(r^2f')' - r^2Uf\}j + 2f'r^2j']R\,dr. \quad (46)$$

Note that Eq. (16) is simply a special case of (46) as can be explicitly seen by setting $f = 1$.

APPENDIX C. COULOMB GREEN'S FUNCTION

As stated in Sec. IV the use of the Wronskian identity is very general and not in any way limited to the functions j and R used in the previous appendix. As an example of this we derive the partial-wave Coulomb Green's function by starting from the two equations,

$$(1/r^2)(r^2F_l')' + \{k^2 - [l(l+1)/r^2]\}F_l = V_0F_l \quad \text{(C1)}$$

and

$$(1/r^2)(r^2S_l')' + \{k^2 - [l(l+1)/r^2]\}S_l = (U+V)S_l. \quad \text{(C2)}$$

Multiply (C1) by r^2S_l, (C2) by r^2F_l, subtract and integrate from a to b, we then obtain

$$\int_a^b F_l(U+V-V_0)S_lr^2\,dr = [W(F_l,S_l)]_a^b. \quad \text{(C3)}$$

Let F_l and G_l be the regular and irregular solutions of (C2) with the following asymptotic forms

$$F_l \to (1/\rho)\sin(\rho - n\ln 2\rho + \tfrac{1}{2}l\pi - \eta_l),$$
$$G_l \to (1/\rho)\cos(\rho - n\ln 2\rho + \tfrac{1}{2}l\pi - \eta_l).$$

One then has $W(F_l, G_l) = i/k$. We now demand that S_l be the regular solution of (C2) with an asymptotic form corresponding to F_l. Using (C3) together with the Wronskian identity (43) we obtain

$$\frac{i}{k}S_l(r) = \frac{i}{k}F_l(r)e^{-i\zeta_l} + G_l(r) \times \int_0^r F_l(r')(U+V-V_0)S_l(r')r'^2\,dr' + F_l(r)\int_r^\infty G_l(r')(U+V-V_0)S_l(r')r'^2\,dr'. \quad \text{(C4)}$$

By summing Eq. (C4) over all partial waves one can obtain the formal equation

$$\psi^+ = u_0^+ + G_c(U+V-V_0)\psi^+, \quad \text{(C5)}$$

where G_c is defined as the Coulomb Green's function. From (C4) one can immediately deduce that

$$G_c(\mathbf{r},\mathbf{r}') = -ik\sum_{l=0}^{\infty}(2l+1)P_l(\cos\theta_1)F_l(kr_<)G_l(kr_>). \quad \text{(C6)}$$

This method for obtaining $G_c(\mathbf{r},\mathbf{r}')$ is very much simpler than the conventional method in which this function is obtained from an eigenfunction expansion.[24]

If one now examines the asymptotic form of (C4) it is straightforward to deduce that the scattering amplitude can be written as

$$f_{\text{tot}} = f_c - (1/4\pi)\langle u_{c,f}^-|\,U+V-V_0\,|\psi_i^+\rangle. \quad \text{(C7)}$$

In terms of partial waves this reduces to Eq. (9). In many of the standard works[3] the point Coulomb term is subtracted out and the result is written as

$$f_{\text{tot}} = f_c + \sum_{l=0}^{\infty}(2l+1)P_l(\cos\theta)e^{2i\eta_l}(e^{2i\zeta_l} - 1). \quad \text{(C8)}$$

APPENDIX D. SIMPLE APPLICATION TO A SINGULAR POTENTIAL

We present here a simple example to illustrate the usefulness of Eq. (47) in deriving integral equations for the singular potential $U = g^2/r^4$. Let $f = f_1 = f_2$ be a solution to the differential equation $(r^2f')' - Ur^2f = 0$. The solution which is finite at the origin is

[24] L. H. Hostler, J. Math. Phys. **5**, 591 (1964).

$f = e^{-g/r}$. Let us take $a = 0$, $b = \infty$ and demand that $R \sim e^{\pm ig/r}$ at $r = 0$ and

$$R \to (1/\rho) \sin(\rho - \tfrac{1}{2}l\pi + \delta_l) \quad \text{at} \quad r = \infty.$$

Using these in (47) we obtain the Fredholm equation:

$$\tilde{R}_l = j_l e^{-i\delta_l} + ikh_l^{(1)} \int_0^r 2gj_l' \tilde{R}_l \, dr' - ikj_l \int_r^\infty 2gh_l^{(1)\prime} \tilde{R}_l \, dr',$$

where $\tilde{R}_l \equiv R_l e^{-g/r}$. This equation may be iterated without introducing divergences. Examining the asymptotic form of this equation we can immediately identify

$$2g \int_0^\infty j_l' R_l e^{-g/r} \, dr$$

as the partial-wave S matrix.

Clearly, other types of equations, such as a Volterra equation, may be deduced by suitable choice of the arbitrary parameters.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Clebsch–Gordan Series in the Exceptional Groups

F. ZACCARIA
Istituto di Fisica Teorica dell'Università, Napoli, Italy

(Received 30 September 1966)

Conditions are given in order that the number of terms and multiplicities in the Clebsch–Gordan series for the exceptional groups are dependent on only one of the irreducible representations whose direct product is decomposed.

A THEOREM on the Clebsch–Gordan series for the classical groups was proved in a previous paper.[1] It states the necessary and sufficient conditions in order that the number of terms and multiplicities in the CG series are dependent on only one of the irreducible representations, whose direct product is decomposed.

The method was an extension of that used by Vitale[2] for all rank-2 groups. It has now been applied also to the remaining groups, namely the exceptional ones of rank >2. The Weyl groups in such cases do not always act in a simple way as in the classical groups, and the detailed computations which have been necessary appeared to be simpler in the Cartan[3] formulation of the root diagrams.

Same notations are kept as in the previous work. In the Appendix details are given about the components of highest weights, their transformations under Weyl reflections, and the inequalities to be maximized with a suitable choice of the elements of the Weyl group. The following are the wanted relations:

$$F(4): \quad p_{1,3} \geq 2p_1' + 2p_2' + 3p_3' + 4p_4',$$
$$p_{2,4} \geq p_1' + 2p_2' + 3p_3' + 3p_4',$$

$$E(6): \quad p_i \geq p_1' + 2p_2' + p_3' + 2p_4' + 2p_5' + 3p_6' \quad (i = 1, \cdots, 6),$$

$$E(7): \quad p_i \geq 2p_1' + p_2' + 2p_3' + 2p_4' + 3p_5' + 3p_6' + 4p_7' \quad (i = 1, \cdots, 7),$$

$$E(8): \quad p_i \geq 2p_1' + 2p_2' + 3p_3' + 3p_4' + 4p_5' + 4p_6' + 5p_7' + 6p_8' \quad (i = 1, \cdots, 8).$$

All such results, including those for the classical groups, have also been obtained by Tits,[4] with a method where the Weyl group is present through its general properties and acts on the root vectors.

ACKNOWLEDGMENTS

The author is grateful to Professor B. Vitale for his continuous guidance and to Professor J. Tits for the interest shown in the problem.

APPENDIX

The weight space is referred to an orthonormal system of vectors e_l.

F(4)

$$M_1 = p_1 + p_2 + \tfrac{3}{2}p_3 + 2p_4,$$
$$M_2 = p_2 + \tfrac{1}{2}p_3 + p_4,$$
$$M_3 = \tfrac{1}{2}p_3 + p_4,$$
$$M_4 = \tfrac{1}{2}p_3,$$

[1] F. Zaccaria, J. Math. Phys. **7**, 1548 (1966).
[2] B. Vitale, Nuovo Cimento **44**, 291 (1966), referred to in the previous paper as No. 3.
[3] E. Cartan, *Oevres complètes* (Paris, 1952), pp. 374ff, where the expressions of the fundamental dominant weights can be found.
[4] J. Tits (private communication).

$f = e^{-g/r}$. Let us take $a = 0$, $b = \infty$ and demand that $R \sim e^{\pm ig/r}$ at $r = 0$ and

$$R \to (1/\rho) \sin(\rho - \tfrac{1}{2}l\pi + \delta_l) \quad \text{at} \quad r = \infty.$$

Using these in (47) we obtain the Fredholm equation:

$$\tilde{R}_l = j_l e^{-i\delta_l} + ikh_l^{(1)} \int_0^r 2gj_l' \tilde{R}_l \, dr' - ikj_l \int_r^\infty 2gh_l^{(1)\prime} \tilde{R}_l \, dr',$$

where $\tilde{R}_l \equiv R_l e^{-g/r}$. This equation may be iterated without introducing divergences. Examining the asymptotic form of this equation we can immediately identify

$$2g \int_0^\infty j_l' R_l e^{-g/r} \, dr$$

as the partial-wave S matrix.

Clearly, other types of equations, such as a Volterra equation, may be deduced by suitable choice of the arbitrary parameters.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Clebsch–Gordan Series in the Exceptional Groups

F. ZACCARIA
Istituto di Fisica Teorica dell'Università, Napoli, Italy

(Received 30 September 1966)

Conditions are given in order that the number of terms and multiplicities in the Clebsch–Gordan series for the exceptional groups are dependent on only one of the irreducible representations whose direct product is decomposed.

A THEOREM on the Clebsch–Gordan series for the classical groups was proved in a previous paper.[1] It states the necessary and sufficient conditions in order that the number of terms and multiplicities in the CG series are dependent on only one of the irreducible representations, whose direct product is decomposed.

The method was an extension of that used by Vitale[2] for all rank-2 groups. It has now been applied also to the remaining groups, namely the exceptional ones of rank >2. The Weyl groups in such cases do not always act in a simple way as in the classical groups, and the detailed computations which have been necessary appeared to be simpler in the Cartan[3] formulation of the root diagrams.

Same notations are kept as in the previous work. In the Appendix details are given about the components of highest weights, their transformations under Weyl reflections, and the inequalities to be maximized with a suitable choice of the elements of the Weyl group. The following are the wanted relations:

$$F(4): \quad p_{1,3} \geq 2p_1' + 2p_2' + 3p_3' + 4p_4',$$
$$p_{2,4} \geq p_1' + 2p_2' + 3p_3' + 3p_4',$$

$$E(6): \quad p_i \geq p_1' + 2p_2' + p_3' + 2p_4' + 2p_5' + 3p_6' \quad (i = 1, \cdots, 6),$$

$$E(7): \quad p_i \geq 2p_1' + p_2' + 2p_3' + 2p_4' + 3p_5' + 3p_6' + 4p_7' \quad (i = 1, \cdots, 7),$$

$$E(8): \quad p_i \geq 2p_1' + 2p_2' + 3p_3' + 3p_4' + 4p_5' + 4p_6' + 5p_7' + 6p_8' \quad (i = 1, \cdots, 8).$$

All such results, including those for the classical groups, have also been obtained by Tits,[4] with a method where the Weyl group is present through its general properties and acts on the root vectors.

ACKNOWLEDGMENTS

The author is grateful to Professor B. Vitale for his continuous guidance and to Professor J. Tits for the interest shown in the problem.

APPENDIX

The weight space is referred to an orthonormal system of vectors e_l.

F(4)

$$M_1 = p_1 + p_2 + \tfrac{3}{2}p_3 + 2p_4,$$
$$M_2 = p_2 + \tfrac{1}{2}p_3 + p_4,$$
$$M_3 = \tfrac{1}{2}p_3 + p_4,$$
$$M_4 = \tfrac{1}{2}p_3,$$

[1] F. Zaccaria, J. Math. Phys. **7**, 1548 (1966).
[2] B. Vitale, Nuovo Cimento **44**, 291 (1966), referred to in the previous paper as No. 3.
[3] E. Cartan, *Oevres complètes* (Paris, 1952), pp. 374ff, where the expressions of the fundamental dominant weights can be found.
[4] J. Tits (private communication).

$$p_1 \geq (S^kM')_2 + (S^kM')_3 + (S^kM')_4 - (S^kM')_1,$$
$$p_2 \geq (S^kM')_3 - (S^kM')_2,$$
$$p_3 \geq -2(S^kM')_4,$$
$$p_4 \geq (S^kM')_4 - (S^kM')_3,$$
$$S^kM = \begin{cases} M - 2M_ie_i, \\ M - (\pm M_i \pm M_j)(\pm e_i \pm e_j), \\ M - \frac{1}{2}(\pm M_1 \pm M_2 \pm M_3 \pm M_4) \\ \qquad \times (\pm e_1 \pm e_2 \pm e_3 \pm e_4) \end{cases} \quad \begin{pmatrix} i,j = 1,\cdots,4 \\ i \neq j \end{pmatrix},$$

with the same sequence of signs in both parentheses of the right-hand side.

E(6)

$$M_1 = \tfrac{2}{3}p_1 + p_2 + \tfrac{4}{3}p_3 + \tfrac{4}{3}p_4 + \tfrac{5}{3}p_5 + 2p_6,$$
$$M_2 = \tfrac{2}{3}p_1 + p_2 + \tfrac{1}{3}p_3 + \tfrac{4}{3}p_4 + \tfrac{5}{3}p_5 + 2p_6,$$
$$M_3 = \tfrac{2}{3}p_1 + p_2 + \tfrac{1}{3}p_3 + \tfrac{4}{3}p_4 + \tfrac{2}{3}p_5 + 2p_6,$$
$$M_4 = \tfrac{2}{3}p_1 + p_2 + \tfrac{1}{3}p_3 + \tfrac{4}{3}p_4 + \tfrac{2}{3}p_5 + p_6,$$
$$M_5 = \tfrac{2}{3}p_1 + p_2 + \tfrac{1}{3}p_3 + \tfrac{1}{3}p_4 + \tfrac{2}{3}p_5 + p_6,$$
$$M_6 = -\tfrac{1}{3}p_1 + p_2 + \tfrac{1}{3}p_3 + \tfrac{1}{3}p_4 + \tfrac{2}{3}p_5 + p_6,$$
$$p_1 \geq (S^kM')_6 - (S^kM')_5,$$
$$p_2 \geq \tfrac{1}{3}[(S^kM')_1 + (S^kM')_2 + (S^kM')_3] - \tfrac{2}{3}[(S^kM')_4 + (S^kM')_5 + (S^kM')_6],$$
$$p_3 \geq (S^kM')_2 - (S^kM')_1,$$
$$p_4 \geq (S^kM')_5 - (S^kM')_4,$$
$$p_5 \geq (S^kM')_3 - (S^kM')_2,$$
$$p_6 \geq (S^kM')_4 - (S^kM')_3,$$
$$S^kM = \begin{cases} M - (M_i - M_j)(e_i - e_j), \\ M + [\frac{1}{3}\sum_l M_l - M_i - M_j - M_k) \\ \qquad \times (e_i + e_j + e_k), \\ M - \frac{1}{3}\sum_l M_l(e_1 + e_2 + \cdots + e_6) \end{cases} \quad \begin{pmatrix} i,j,k = 1,\cdots,6 \\ i \neq j \neq k \end{pmatrix}.$$

E(7)

$$M_1 = p_1 + \tfrac{3}{2}p_2 + \tfrac{3}{2}p_3 + 2p_4 + 2p_5 + \tfrac{5}{2}p_6 + 3p_7,$$
$$M_2 = p_1 + \tfrac{1}{2}p_2 + \tfrac{3}{2}p_3 + 2p_4 + 2p_5 + \tfrac{5}{2}p_6 + 3p_7,$$
$$M_3 = p_1 + \tfrac{1}{2}p_2 + \tfrac{3}{2}p_3 + p_4 + 2p_5 + \tfrac{5}{2}p_6 + 3p_7,$$
$$M_4 = p_1 + \tfrac{1}{2}p_2 + \tfrac{3}{2}p_3 + p_4 + 2p_5 + \tfrac{3}{2}p_6 + 3p_7,$$
$$M_5 = p_1 + \tfrac{1}{2}p_2 + \tfrac{3}{2}p_3 + p_4 + 2p_5 + \tfrac{3}{2}p_6 + 2p_7,$$
$$M_6 = p_1 + \tfrac{1}{2}p_2 + \tfrac{3}{2}p_3 + p_4 + p_5 + \tfrac{3}{2}p_6 + 2p_7,$$
$$M_7 = \tfrac{1}{2}p_2 + \tfrac{3}{2}p_3 + p_4 + p_5 + \tfrac{3}{2}p_6 + 2p_7,$$
$$p_1 \geq (S^kM')_7 - (S^kM')_6,$$
$$p_2 \geq (S^kM')_2 - (S^kM')_1,$$
$$p_3 \geq \tfrac{1}{3}\sum (S^kM')_l - (S^kM')_5 - (S^kM')_6 - (S^kM')_7,$$
$$p_4 \geq (S^kM')_3 - (S^kM')_2,$$
$$p_5 \geq (S^kM')_6 - (S^kM')_5,$$
$$p_6 \geq (S^kM')_4 - (S^kM')_3,$$
$$p_7 \geq (S^kM')_5 - (S^kM')_4,$$
$$S^kM = \begin{cases} M - (M_i - M_j)(e_i - e_j), \\ M + (\frac{1}{3}\sum_l M_l - M_i - M_j - M_k) \\ \qquad \times (e_i + e_j + e_k), \\ M - (\frac{1}{3}\sum_l M_l - M_i)(\sum_{l \neq i} e_l) \end{cases} \quad \begin{pmatrix} i,j,k = 1,\cdots,7 \\ i \neq j \neq k \end{pmatrix}.$$

E(8) $\qquad \left[\sum_{l=1}^{9} M_l = 0\right]$

$$M_1 = p_1 + \tfrac{5}{3}p_2 + 2p_3 + \tfrac{8}{3}p_4 + 3p_5 + \tfrac{10}{3}p_6 + 4p_7 + 5p_8,$$
$$M_2 = \tfrac{2}{3}p_2 - \tfrac{1}{3}p_4 + \tfrac{1}{3}p_6,$$
$$M_3 = -\tfrac{1}{3}p_2 - \tfrac{1}{3}p_4 + \tfrac{1}{3}p_6,$$
$$M_4 = -\tfrac{1}{3}p_2 - \tfrac{1}{3}p_4 - \tfrac{2}{3}p_6,$$
$$M_5 = -\tfrac{1}{3}p_2 - \tfrac{1}{3}p_4 - \tfrac{2}{3}p_6 - p_8,$$
$$M_6 = -\tfrac{1}{3}p_2 - \tfrac{1}{3}p_4 - \tfrac{2}{3}p_6 - p_7 - p_8,$$
$$M_7 = -\tfrac{1}{3}p_2 - \tfrac{1}{3}p_4 - p_5 - \tfrac{2}{3}p_6 - p_7 - p_8,$$
$$M_8 = -\tfrac{1}{3}p_2 - p_3 - \tfrac{1}{3}p_4 - p_5 - \tfrac{2}{3}p_6 - p_7 - p_8,$$
$$M_9 = -p_1 - \tfrac{1}{3}p_2 - p_3 - \tfrac{1}{3}p_4 - p_5 - \tfrac{2}{3}p_6 - p_7 - p_8,$$
$$p_1 \geq (S^kM')_9 - (S^kM')_8,$$
$$p_2 \geq (S^kM')_3 - (S^kM')_2,$$
$$p_3 \geq (S^kM')_8 - (S^kM')_7,$$
$$p_4 \geq (S^kM')_2 + (S^kM')_3 + (S^kM')_4,$$
$$p_5 \geq (S^kM')_7 - (S^kM')_6,$$
$$p_6 \geq (S^kM')_4 - (S^kM')_3,$$
$$p_7 \geq (S^kM')_6 - (S^kM')_5,$$
$$p_8 \geq (S^kM')_5 - (S^kM')_4,$$
$$S^kM = \begin{cases} M - (M_i - M_j)(e_i - e_j), \\ M - (M_i + M_j + M_k)(e_i + e_j + e_k) \\ \qquad + \frac{1}{3}(M_i + M_j + M_k)\sum_l e_l \end{cases} \quad \begin{pmatrix} i,j,k = 1,\cdots,9 \\ i \neq j \neq k \end{pmatrix}.$$

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Permutational Symmetry of Reduced Density Matrices

R. D. POSHUSTA*
Quantum Theory Group, Keele University, Staffordshire, England
(Received 23 September 1966)

The usual restriction on the permutational symmetry of density matrices is removed. An Nth-order density matrix of general permutational symmetry is reduced and the symmetry of the resulting pth-order density matrix is found. This gives rise to necessary conditions for the N-representability of spin-free density matrices. We indicate the importance of these results for symmetry under the groups $GL(n)$, $U(n)$, and $SU(n)$. We generalize the question of N-representability to the tensor algebra, where contraction of tensors plays the role of reduction of density matrices. Symmetry properties of the tensors are studied and density matrices are considered as special cases.

INTRODUCTION

AMONG the many excellent papers on density-matrix theory (Refs. 1 and 2 provide the foundations of the theory as well as a thorough bibliography), not one discusses the most general types of permutational symmetry. All are restricted to the two simplest cases: complete symmetry and/or complete antisymmetry (though some make the restriction only after developing important properties which hold for all symmetry types). It is not necessary to restrict one's attention to these special cases. Neither is it always desirable, e.g., the natural approach to spin-free density matrices[3] is through the Young operators of general permutational symmetry.

This paper lays the foundations of the theory of density matrices with arbitrary permutational symmetry. Special interest is focused on the problem of "percolation" of symmetry from Nth order to pth order under reduction; in other words, we find an answer to the question: What permutational symmetry survives in pth order after reduction from an Nth-order density matrix of known symmetry? The solution of this problem for permutational symmetry also provides a partial solution for symmetry under some continuous groups.

Our results provide conditions for N-representability of density matrices with general permutational symmetry (e.g., spin-free density matrices). Necessary conditions are found but sufficient conditions are more elusive as in the cases of complete symmetry and complete antisymmetry. These sufficient conditions are pursued in a subsequent paper, where we deal with the eigenvalues of various permutation adapted density matrices.

The Dirac bra–ket notation is used here because of its economy; and tensor algebra is employed because of its elegance. Also, it is unnecessarily cumbersome to confine ourselves to density matrices. Instead, we consider a large class of tensor operators, their permutational symmetry, and the effect of contraction on this symmetry. Density matrices are special cases of such tensor operators and reduction of density matrices is a special kind of contraction.

Transition density matrices may have permutational symmetry α on the left and another symmetry β on the right. Such density matrices correspond to transitions between different symmetries, e.g., different spin states if they are spin-free density matrices. In the Appendix we show how α and β may differ without producing zero on reduction.

1. TENSOR ALGEBRA

A complete discussion of tensor algebra[4] cannot be given here. The reader is assumed to have a basic knowledge of the subject. This section serves primarily to present our notation; this is done by reviewing the application of tensor algebra to quantum theory.

Let U be any vector space over the complex field. Elements of U are called kets (written $|\psi\rangle$) and represent the states of a single particle (actually there is some redundancy in this representation, since a ray is sufficient to specify a state). The elements of the N-fold tensor product space $U_0^N = U \otimes U \otimes \cdots \otimes U$ are also called kets (written $|\psi\rangle$) but they represent states of a collection of N particles. The dual space of U, denoted by $U^\dagger$, consists of all linear mappings of U into the complex field. Elements of $U^\dagger$ are called bras (written $\langle\psi|$); they form an alternate representation of the states of a particle. The dual space of U_0^N is

* Present address: Department of Chemistry, Washington State University, Pullman, Washington.

[1] A. J. Coleman, Rev. Mod. Phys. 35, 668 (1963).
[2] R. McWeeny, Rev. Mod. Phys. 32, 335 (1960).
[3] R. D. Poshusta and F. A. Matsen, J. Math. Phys. 7, 711 (1966).
[4] G. D. Mastow, J. H. Sampson, and J. Meyer, *Fundamental Structures of Algebra* (McGraw-Hill Book Company, Inc., New York, 1963).

denoted by U_N^0 and can be shown to be given by $U_N^0 = U^\dagger \otimes U^\dagger \otimes \cdots \otimes U^\dagger$. The space $U_N^N = U_0^N \otimes U_N^0$ consists of tensors of type (N, N) (i.e., tensors which transform cogrediently and contragrediently in N indices relative to transformations of U). Elements of U_N^N are linear mappings of U_0^N into U_0^N and it can be shown that every such mapping is an element of U_N^N. Hence all the operators of quantum mechanics[5] are tensors in U_N^N.

Given a basis, $\{|i\rangle, i = 1, 2, \cdots\}$, of U we choose a basis, $\{\langle i|, i = 1, 2, \cdots\}$, of $U^\dagger$ such that the images, $\langle i \,|\, j\rangle$, are zero if $i \neq j$ and one if $i = j$. Then an arbitrary ket $|\psi\rangle$ is expressible in the basis by means of its image (bra–kets) under the basis bras

$$|\psi\rangle = |i\rangle\langle i \,|\, \psi\rangle$$

(Einstein summation convention). Further we choose $\langle \psi \,|\, \phi\rangle = \langle \phi \,|\, \psi\rangle^*$ so that

$$\langle\psi| = \langle \psi \,|\, i\rangle\langle i|.$$

The tensor product spaces then possess the following bases:

$$U_1^1 = (|i\rangle\langle j|; i = 1, 2, \cdots; j = 1, 2, \cdots),$$
$$U_0^N = (|i(N)\rangle; i_r = 1, 2, \cdots),$$
$$U_N^0 = (\langle i(N)|; i_r = 1, 2, \cdots),$$
$$U_N^N = (|i(N)\rangle\langle j(N)|; i_r = 1, 2, \cdots; j_r = 1, 2, \cdots).$$

Here, we have introduced the abbreviation $i(N)$ for the set of N-ordered indices $i_1, i_2, \cdots, i_N$ and

$$|i(N)\rangle \equiv |i_1\rangle\, |i_2\rangle \cdots |i_N\rangle.$$

In terms of these bases we expand an arbitrary N-particle ket as $|\psi\rangle = |i(N)\rangle\langle i(N) \,|\, \psi\rangle$, and an arbitrary N-particle operator as

$$X = |i(N)\rangle\langle i(N)|\, X \,|\, j(N)\rangle\langle j(N)|.$$

The Nth-order density matrices are examples of tensors of type (N, N). Ordinary density matrices are those tensors of the simple form $|\Psi\rangle\langle\Psi|$. Statistical densities are weighted averages of such simple density matrices:

$$\sum_i w_i |\Psi_i\rangle\langle\Psi_i|.$$

The collection of all statistical densities forms a convex subset[1] of U_N^N. Transition density matrices have the more general form $|\Psi\rangle\langle\Phi|$ but are again tensors of type (N, N).

2. PERMUTATIONS AND SYMMETRY ADAPTED SUBSPACES

In U_0^N we interpret the order of one-particle kets (or orbitals) in terms of particle numbers. That is, in $|i(N)\rangle$ the first particle is represented by (or occupies) $|i_1\rangle$, the second by $|i_2\rangle$ and so forth. A permutation among these particles is denoted by

$$P = \downarrow \begin{pmatrix} 1 & 2 & \cdots & N \\ r_1 & r_2 & \cdots & r_N \end{pmatrix}$$

and we define

$$P\,|i(N)\rangle \equiv |i_{r_1}\rangle\, |i_{r_2}\rangle \cdots |i_{r_N}\rangle,$$

in which the first particle occupies $|i_{r_1}\rangle$, etc., and its dual

$$\langle i(N)|\, P^{-1} \equiv \langle i_{r_N}| \cdots \langle i_{r_2}|\, \langle i_{r_1}|,$$

in which the first particle occupies $\langle i_{r_1}|$. The symmetric group, S_N, of all these permutations spans a subalgebra, A_N, of U_N^N. The full tensor algebra U_N^N is simple and hence contains no proper ideals. But A_N is only semisimple and has a unique decomposition[6] (or reduction) in minimal ideals

$$A_N = \sum_\alpha A^\alpha.$$

This decomposition of A_N induces a decomposition on U_N^N:

$$U_N^N = \sum_\alpha \sum_\beta A^\alpha U_N^N A^\beta,$$

where $A^\alpha U_N^N A^\beta$ is invariant (but not irreducible) under A_N. The subspace $A^\alpha U_N^N A^\beta$ is also an algebra (nilpotent unless $\alpha = \beta$) and is called a symmetry adapted subspace or subalgebra of U_N^N.

Density matrices for systems of indistinguishable particles have their permutational symmetries restricted. For example, fermion density matrices lie in the totally antisymmetric subspace $A^{\{1^N\}} U_N^N A^{\{1^N\}}$, boson density matrices in the symmetric subspace $A^{\{N\}} U_N^N A^{\{N\}}$, and spin-free electron density matrices lie in $A^\alpha U_N^N A^\beta$, where α and β must be of the form $\{2^p, 1^{N-2p}\}$ (see Ref. 3).

3. REDUCTION OF TENSORS

Reduction of tensors of type (N, N) is a kind of contraction. For this purpose the tensor products $|i(N)\rangle$ are partitioned into two parts. The first p members, $|i(p)\rangle \equiv |i_1\rangle\, |i_2\rangle \cdots |i_p\rangle$ are considered separately from the last q members,

$$|i(q)\rangle \equiv |i_{p+1}\rangle \cdots |i_N\rangle \quad (p + q = N),$$

and we write

$$|i(N)\rangle = |i(p)\rangle\, |i(q)\rangle.$$

If X is any tensor of U_N^N, the (q)-trace of X is defined to be

$$\mathrm{Tr}^{(q)} X = \mathrm{Tr}^{(q)}\, (|i(N)\rangle\langle i(N)|\, X \,|\, j(N)\rangle\langle j(N)|)$$
$$= |i(p)\rangle\langle i(p) i(q)|\, X \,|\, j(p) i(q)\rangle\langle j(p)|.$$

[5] Creation and annihilation operators belong to tensor spaces of the form U_N^{N+1}, U_{N+1}^N, etc.

[6] H. Boerner, *Representations of Groups* (North-Holland Publishing Company, Amsterdam, 1963).

Hence $\mathrm{Tr}^{(q)}$ is a linear mapping,

$$\mathrm{Tr}^{(q)}: U_N^N \to U_p^p.$$

The group of permutations on $|i(p)\rangle$ is denoted by $S_{(p)}$ and its algebra by $A_{(p)}$ while the corresponding group and algebra for $|i(q)\rangle$ are $S_{(q)}$ and $A_{(q)}$.

We are concerned in the following with the (q) trace of symmetry-adapted subalgebras $A^\alpha U_N^N A^\beta$. We define

$$U_{\alpha\beta}^{N\to p} = \mathrm{Tr}^{(q)}(A^\alpha U_N^N A^\beta).$$

If $p = N$ we write $U_{\alpha\beta}^N \equiv A^\alpha U_N^N A^\beta$ and if $p = 0$ we write $U_{\alpha\beta}^{N\to 0} = U_0^0$, the complex field.

The central theorems concerning $U_{\alpha\beta}^{N\to p}$ are Theorems 4 and 5 in the following section. They are based on two lemmas which are given below. Theorems 2 and 3 follow easily from the lemmas, but their proofs are not given because they are special cases of Theorems 4 and 5.

Lemma 1: $U_{\alpha\beta}^{N\to p}$ is spanned by $\{\mathrm{Tr}^{(q)}(e^\alpha X_n e^\beta), n = 1, 2, \cdots\}$, where $\{X_n, n = 1, 2, \cdots\}$ is any basis of U_N^N and e^α and e^β are the principal idempotents of A^α and A^β.

Proof: First, note that $A^\alpha = A_N e^\alpha = e^\alpha A_N$ and therefore $U_{\alpha\beta}^N = e^\alpha A_N U_N^N A_N e^\beta$. Second, A_N contains the identity so that $A_N U_N^N = U_N^N$ and hence $U_{\alpha\beta}^N = e^\alpha U_N^N e^\beta$. Finally, let $x \in U_{\alpha\beta}^{N\to p}$ so that there exists an element $X \in U_N^N$ such that $x = \mathrm{Tr}^{(q)}(e^\alpha X e^\beta)$. When X is expressed in terms of the basis elements X_n, the theorem follows.

Lemma 2: For $X \in U_N^N$ and $a \in A_{(p)}$, then

$$a\,\mathrm{Tr}^{(q)} X = \mathrm{Tr}^{(q)}(aX),$$
$$(\mathrm{Tr}^{(q)} X)a = \mathrm{Tr}^{(q)}(Xa).$$

Proof: Let $X = |i(N)\rangle\langle i(N)|\,X\,|j(N)\rangle\langle j(N)|$ and let

$$a = \sum_P a_P P,$$

where $P \in S_{(p)}$. Then

$$\begin{aligned} a\,\mathrm{Tr}^{(q)} X &= a\,|i(p)\rangle\langle i(N)|\,X\,|j(p)i(q)\rangle\langle j(p)| \\ &= \sum_P a_P P\,|i(p)\rangle\langle i(N)|\,X\,|j(p)i(q)\rangle\langle j(p)| \\ &= \mathrm{Tr}^{(q)} \sum_P a_P P\,|i(p)\rangle|i(q)\rangle \\ &\qquad \times \langle i(N)|\,X\,|j(N)\rangle\langle j(N)|) \\ &= \mathrm{Tr}^{(q)}(aX). \end{aligned}$$

Similarly from the right.

Theorem 1: $U_{\alpha\beta}^{N\to p}$ is invariant under $A_{(p)}$ from both sides.

Proof: Let $a \in A_{(p)}$ and let $x \in U_{\alpha\beta}^{N\to p}$. Then there exists an $X \in U_N^N$ such that $x = \mathrm{Tr}^{(q)}(e^\alpha X e^\beta)$. By Lemma 2 we have $ax = \mathrm{Tr}^{(q)}(ae^\alpha X e^\beta)$. But e^α commutes with every element of $A_N \supset A_{(p)}$, so that

$$ax = \mathrm{Tr}^{(q)}(e^\alpha a X e^\beta) \in U_{\alpha\beta}^{N\to p}.$$

Similarly from the right.

Theorem 2:

$$U_{\{1^N\}\{1^N\}}^{N\to p} \subseteq U_{\{1^p\}\{1^p\}}^p.$$

Theorem 3:

$$U_{\{N\}\{N\}}^{N\to p} = U_{\{p\}\{p\}}^p.$$

Theorem 2 applies to fermion density matrices; it shows that the antisymmetry of density matrices is preserved in all orders. Similarly, Theorem 3, which applies to boson density matrices, shows that the symmetry of density matrices is preserved in all orders. But what happens in the general case and especially what sorts of permutational symmetries are permitted for spin-free reduced density matrices? The answer is provided by Theorem 4.

Theorem 3 provides necessary *and* sufficient conditions on reduced tensors while Theorem 2 provides only necessary conditions (as has often been pointed out). Theorem 5 is a generalization of Theorem 3 and provides sufficient conditions for a class of "largely symmetric tensors."

4. DERIVABILITY AND REPRESENTABILITY

A tensor x of type (p, p) is said to be N,α,β-derivable if and only if there exists a tensor X of type (N, N) such that

$$x = \mathrm{Tr}^{(q)}(e^\alpha X e^\beta).$$

Alternatively, x is N,α,β-derivable if and only if $x \in U_{\alpha\beta}^{N\to p}$. Coleman's N-representability of density matrices is a special kind of derivability, namely x is N-representable if and only if

(1) $\mathrm{Tr}\,x = 1$,
(2) $x\,|\psi\rangle = \lambda\,|\psi\rangle \Rightarrow \lambda \geq 0$,
(3) $x^\dagger = x$,
(4) $\begin{cases} x \text{ is } N,\{1^N\},\{1^N\}\text{-derivable (fermions)}, \\ x \text{ is } N,\{N\},\{N\}\text{-derivable (bosons)}. \end{cases}$

Spin-free density matrices satisfy (1), (2), and (3) but have more general requirements on their permutational symmetry.[3]

Theorems 2 and 3 can be restated in terms of derivability:

Theorem 2: $x \in U^p$ is $N,\{1^N\},\{1^N\}$-derivable implies that $x \in U_{\{1^p\}\{1^p\}}^p$. That is, antisymmetry from both

sides is a necessary condition for $N,\{1^N\},\{1^N\}$-derivability.

Theorem 3: $x \in U_p^p$ is $N,\{N\},\{N\}$-derivable if and only if $x \in U_{\{p\}\{p\}}^p$. That is, it is both necessary and sufficient for $N,\{N\},\{N\}$-derivability that x be symmetric from both sides.

Simple proofs of Theorems 2 and 3 are often seen—these proofs rely on the fact that $e^{\{1^N\}}$ and $e^{\{N\}}$ are primitive idempotents of A_N. In general e^α can be resolved into a sum of f^α primitive idempotents in infinitely many equivalent ways (Ref. 6)

$$e^\alpha = \sum_{r=1}^{f^\alpha} e_r^\alpha .$$

Hence $U_{\alpha\beta}^N$ is also resolved into disjoint subspaces

$$U_{\alpha\beta}^N = \sum_r \sum_s e_r^\alpha U_N^N e_s^\beta ,$$

and likewise

$$U_{\alpha\beta}^{N\to p} = \sum_r \sum_s \mathrm{Tr}^{(q)}(e_r^\alpha U_N^N e_s^\beta)$$

(though here the subspaces are not disjoint).

The Young idempotents are a convenient set of primitive idempotents. Let r denote a Young tableau and P_r and N_r its row and column operators,[6,7] then the Young idempotent of r is given by $E_r = N_r P_r$ (or equally well by $P_r N_r$). With r is associated the partition $\alpha = \{\alpha_1, \alpha_2, \cdots\}$ (also called the shape) in which $\alpha_1, \alpha_2, \cdots$ are the lengths of the rows in the tableau. (By convention $\alpha_1 \geq \alpha_2 \geq \cdots$.) A subset of the Young tableaux is sufficient for decomposition of A_N, namely the standard tableaux. A particularly useful set of primitive idempotents for use in contracted tensors is the set of seminormal idempotents.[7] These are defined as follows. Let $E_r^{(p)}$ be the Young idempotent of $A(p)$ constructed from the standard tableau which results when the integers $p+1$, $p+2, \cdots, N$ are struck from the standard tableau r. Then the seminormal idempotents, $e_r^{(p)}$ are defined by:

$$e_r^{(2)} \equiv E_r^{(2)}, \quad e_r^{(3)} \equiv e_r^{(2)} E_r^{(3)} e_r^{(2)}, \cdots,$$
$$e_r^{(p+1)} \equiv e_r^{(p)} E_r^{(p+1)} e_r^{(p)}, \cdots,$$
$$e_r^{(N)} \equiv e_r \equiv e_r^{(N-1)} E_r e_r^{(N-1)}.$$

The partitions of the successive tableaux obtained in this way are said to be embedded.

Theorem 4: Let ξ and ζ be partitions of p and let α and β be partitions of N, then

$$U_{\alpha\beta}^{N\to p} \subseteq \sum_{\xi,\zeta} U_{\xi\zeta}^p ,$$

where the sum extends over all ξ embedded in α and all ζ embedded in β. That is, a necessary condition for a tensor, x, of type (p,p) to be N,α,β-derivable is that $e^\xi x e^\zeta$ vanish if ξ and ζ are not embedded in α and β, respectively.

Proof: (A) In the resolution of e^α and e^β into primitive idempotents choose e_r and e_s to be "adapted" to the chain of subgroups $S_N \supset S_{(N-1)} \supset \cdots \supset S_{(2)}$. (This choice for the primitive idempotents is equivalent to choosing an irreducible representation in which the chain is completely reduced.) That is, choose the idempotents with the property

$$e_\rho^{(p)} e_r^\alpha = e_r^\alpha e_\rho^{(p)} = \begin{cases} e_r^\alpha & \text{if} \quad \rho \text{ is embedded in } r, \\ 0 & \text{otherwise.} \end{cases}$$

Here the labels ρ, r are taken to be standard Young tableaux, and ρ is said to be embedded in r if ρ results from r by striking off the integers $p+1, p+2, \cdots, N$. For example, the seminormal idempotents have this property. The shape, ξ, of ρ embedded in r is embedded in α, the shape of r. Obviously, $\xi = \{\xi_1, \xi_2, \cdots\}$ is embedded in α if and only if $\xi_1 \leq \alpha_1$, $\xi_2 \leq \alpha_2$, etc.

(B) By Lemma 2 we conclude that

$$\begin{aligned} e_\rho^{(p)} \mathrm{Tr}^{(q)}(e_r^\alpha U_N^N e_s^\beta) e_\sigma^{(p)} &= \mathrm{Tr}^{(q)}(e_\rho^{(p)} e_r^\alpha U_N^N e_s^\beta e_\sigma^{(p)}) \\ &= \mathrm{Tr}^{(q)}(e_r^\alpha U_N^N e_s^\beta) \\ &\subseteq e_\rho^{(p)} U_p^p e_\sigma^{(p)} \\ &\subseteq e^\xi U_p^p e^\zeta \\ &\subseteq U_{\xi\zeta}^p , \end{aligned}$$

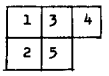
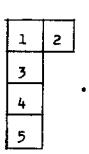
where ξ is embedded in α and ζ in β.

Corollary: Let r and s be two standard Young tableaux of N integers and let ρ and σ be the standard tableaux of p integers obtained by striking $p+1$, $p+2, \cdots, N$ from r and s. Then

$$\mathrm{Tr}^{(q)}(e_r U_N^N e_s) \subseteq e_\rho U_p^p e_\sigma ,$$

where e_r and e_ρ are the seminormal primitive idempotents corresponding to these tableaux. (The idempotents need only be adapted to the subgroup $S_{(p)}$ for the theorem to hold.)

For example, let $|\Psi_r\rangle\langle\Psi_s|$ be symmetry adapted by the seminormal idempotents belonging to

$r =$  and $s =$  .

Then the successive contractions give rise to tensors

[7] D. E. Rutherford, *Substitutional Analysis* (Edinburgh University Press, Edinburgh, 1948).

TABLE 1. Symmetry of $\mathrm{Tr}^{(q)} |\Psi_r\rangle\langle\Psi_s|$.

p	ρ	σ
4	1 3 4 / 2	1 2 / 3 / 4
3	1 3 / 2	1 2 / 3
2	1 / 2	1 2

of type (p, p) whose symmetries ρ, σ are given in Table 1.

In the special case $\alpha = \{1^N\}$ the only partition of p which is embedded in α is $\xi = \{1^p\}$. Hence Theorem 2 is a special case of Theorem 4. Similarly, if $\alpha = \{N\}$, then $\xi = \{p\}$ and Theorem 4 requires that $U^{N\to p}_{\{N\}\{N\}} \subseteq U^{p}_{\{p\}\{p\}}$ which establishes part of Theorem 3. To complete the proof of Theorem 3 we need only show that $U^{p}_{\{p\}\{p\}} \subseteq U^{N\to p}_{\{N\}\{N\}}$. This is a special case of Theorem 5 which we prove next.

For the symmetry $\alpha = \{2^{N/2-S}, 1^{2S}\}$ of electron spin-free density matrices, Theorem 4 establishes that the pth-order reduced density matrices have symmetry components $\xi = \{2^{p/2-S'}, 1^{2S'}\}$, where the possible values of S' are determined by the embedding condition. Here S is the spin quantum number of the Nth-order density matrix and S' that of the pth-order. For example, the quintet, $S = 2$, $\{2, 1^4\}$, state of a six electron system gives rise to a 4th-order density matrix with components: triplet, $S = 1$, $\{2, 1^2\}$, and quintet, $S = 2$, $\{1^4\}$.

Theorem 5: Let $\xi = \{\xi_1, \xi_2, \cdots\}$ be any partition of p and let $\alpha = \{\xi_1 + q, \xi_2, \cdots\}$. Then

$$U^{p}_{\xi\xi} \subseteq U^{N\to p}_{\alpha\alpha},$$

i.e., membership in $U^{p}_{\xi\xi}$ is a sufficient condition for N,α,α-derivability.

Proof: $U^{p}_{\xi\xi}$ is spanned by the basis set

$$\{E^{(p)}_{\rho} |i(p)\rangle\langle j(p)| E^{(p)\dagger}_{\sigma}\},$$

where $i(p)$ and $j(p)$ range over all p-tuples, ρ and σ range over all tableaux of shape ξ, and $E^{(p)}_{\rho}$, $E^{(p)}_{\sigma}$ are the corresponding Young idempotents

$$E^{(p)}_{\rho} = N^{(p)}_{\rho} P^{(p)}_{\rho},$$
$$E^{(p)\dagger}_{\sigma} = P^{(p)}_{\sigma} N^{(p)}_{\sigma}.$$

The theorem is proved if it is shown that every such basis element results from contraction of some element of $U^{N}_{\alpha\alpha}$.

Set $E^{(N)}_{r} = N^{(N)}_{r} P^{(N)}_{r}$ and note that, since ξ is embedded in α according to the rule of the theorem, we have

$$N^{(N)}_{r} = N^{(p)}_{\rho}$$

and

$$P^{(N)}_{r} = P^{(p)}_{\rho} + P^{(p)}_{\rho} \cdot [\text{permutations of } (q) \text{ with } (p)].$$

Also

$$E^{(N)}_{r} = E^{(p)}_{\rho} + E^{(p)}_{\rho} \cdot [\text{permutations of } (q) \text{ with } (p)].$$

A ket such as $|i(p)\rangle$ may contain one or more repeated indices, in which case a permutation of these indices leaves the ket invariant. The set of numbers $\kappa_1, \kappa_2, \cdots, \kappa_k$ which indicates how often $|i_1\rangle$, $|i_2\rangle, \cdots |i_k\rangle$ are repeated forms a partition of p, $\kappa = \{\kappa_1, \kappa_2, \cdots\}$ called the invariance of $|i(p)\rangle$. If $E^{(p)}$ is applied to $|i(p)\rangle$ the result vanishes unless κ equals ξ or precedes ξ in dictionary order[8] [because $|i(p)\rangle$ may not be antisymmetrized in more indices than are distinct]. We consider all invariance types κ in dictionary order from ξ to $\{1^p\}$.

Case 1: Let $\kappa = \xi$. Then define

$$|i(N)\rangle = |i(p)\rangle |i_1\rangle |i_1\rangle \cdots |i_1\rangle$$

so that a permutation of (q) with (p) will either leave $|i(N)\rangle$ unchanged or create a ket $|i'(p)\rangle |i'(q)\rangle$ in which $|i'(p)\rangle$ has invariance greater than ξ. Then

$$E^{(N)}_{r} |i(N)\rangle = E^{(p)} |i(N)\rangle$$

and

$$\mathrm{Tr}^{(q)} (E^{(N)}_{r} |i(N)\rangle\langle j(N)| E^{(N)\dagger}_{s}) = E^{(p)}_{\rho} |i(p)\rangle\langle j(p)| E^{(p)\dagger}_{\sigma},$$

where $i(p)$ and $j(p)$ have invariance κ. The tensors of this type are called type one.

Case 2: Let $|i(p)\rangle$ have invariance κ', the first predecessor of ξ. Then define $|i(N)\rangle = |i(p)\rangle |i_1\rangle \cdots |i_1\rangle$ so that a permutation of (q) with (p) either leaves $|i(N)\rangle$ unchanged, gives a ket of invariance κ as in Case 1, or gives a ket of invariance greater than ξ. Then

$$\mathrm{Tr}^{(q)} (E^{(N)}_{r} |i(N)\rangle\langle j(N)| E^{(N)\dagger}_{s}) = E^{(p)}_{\rho} |i(p)\rangle\langle j(p)| E^{(p)\dagger}_{\sigma} + \text{tensors of type one}$$

for $j(p)$ having invariance κ or κ'. The tensors of type one can be subtracted from both sides of this equation,

[8] Dictionary order is defined in the following way. If $\kappa_1 \neq \xi_1$ then κ precedes ξ if $\kappa_1 < \xi_1$; if $\kappa_1 = \xi_1$ but $\kappa_2 < \xi_2$, then again κ precedes ξ, and so forth. For example, $\{1, 1, 1, 1\} < \{2, 1, 1\} < \{2, 2\} < \{3, 1\} < \{4\}$ and $\{1, 1, 1, 1, 1\} < \{2, 1, 1, 1\} < \{2, 2, 1\} < \{3, 1, 1\} < \{3, 2\} < \{4, 1\} < \{5\}$.

thus showing the existence of a tensor in $U^N_{\alpha\alpha}$ which gives on contraction

$$E^{(p)}_\rho \, |i(p)\rangle\langle j(p)| \, E^{(p)\dagger}_\sigma,$$

where $i(p), j(p)$ have invariance κ or κ'.

In a similar way this result can be extended to κ'', $\cdots$, $\{1, 1, \cdots, 1\}$. Thus every basis element of $U^P_{\xi\xi}$ is an element of $U^{N\to p}_{\alpha\alpha}$.

If $\alpha = \{N\}$, Theorem 5 establishes that $U^P_{\{p\}\{p\}} \subseteq U^{N\to p}_{\{N\}\{N\}}$ which completes the proof of Theorem 3.

5. SYMMETRY UNDER *GL(n)*, *U(n)*, AND *SU(n)*

The preceding results concerning symmetry under S_N bear directly on symmetry under $GL(n)$ and several of its subgroups, e.g., $U(n)$ and $SU(n)$. Let the ket space U have finite dimension and denote this dimension by n. The groups $GL(n)$, $U(n)$, and $SU(n)$ are the matrix groups of all nonsingular, all unitary, and all unitary unimodular $n \times n$ matrices. Clearly these $n \times n$ matrices are irreducible.

Now, consider U_N and the transformations induced on U_N by transformations on U. These induced transformations are N-fold Kronecker products of the original defining matrices and have dimension n^N. They are not irreducible. The complete reduction of U_N is carried out with the symmetric group algebra.

A proof of the reducibility of U_N and a derivation of its irreducible subspaces cannot be given here (but see Ref. 6). Briefly the proof rests on the following: If e is a primitive idempotent of A_N, then eU^N is either zero or an irreducible subspace under $GL(n)$. Conversely if R is an irreducible subspace of U^N then a primitive idempotent of A_N exists such that $R = eU^N$. Moreover, all integral representations of $GL(n)$ are obtained in this way and two irreducible representations are equivalent if and only if the corresponding idempotents e, e' are equivalent. Lastly, an irreducible representation of $GL(n)$ remains irreducible when the matrices are restricted to certain subgroups, e.g., $U(n)$ and $SU(n)$.

Let e^α_τ be a primitive idempotent of A_N. Then $e^\alpha_\tau U^N$ generates an irreducible representation of $GL(n)$. Since equivalent representations correspond to equivalent idempotents, the partitions α characterize irreducible representations up to equivalence. Let Γ^α be an irreducible representation of $GL(n)$ and let $|\psi\rangle$ be a ket of this symmetry. Then $|\psi\rangle \in e^\alpha U^N$. Further, if T is a tensor of type (N, N) and if T transforms under $GL(n)$ according to irreducible symmetries Γ^α from the left and Γ^β from the right, then $T \in U^N_{\alpha\beta}$. Theorem 4 indicates that contraction of T to type (p, p) produces a tensor which transforms under $GL(n)$ according to

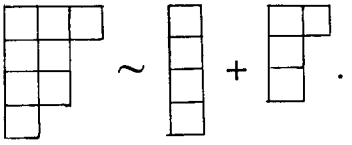
$$\Gamma^l = \sum_{\xi\subset\alpha} \dot{}\, \Gamma^\xi$$

from the left and

$$\Gamma^r = \sum_{\zeta\subset\beta} \dot{}\, \Gamma^\zeta$$

from the right. Here $\xi \subset \alpha$ means ξ embedded in α.

The same remarks apply to $U(n)$ but for $SU(n)$ some of the representations Γ^α, Γ^β are equivalent when $\alpha \neq \beta$. In order that the preceding remarks concerning $GL(n)$ should also apply to $SU(n)$, one need only note that two irreducible representations Γ^α and Γ^β are equivalent if $\alpha = \beta$ or if α differs from β by having one or more columns of length n appended (on the left, of course). For example, let $n = 4$, then $\{2, 1, 1\} \sim \{3, 2, 2, 1\}$ since



One gets all the irreducible representations of $SU(n)$ by considering only such partitions, α, that have less than n rows, i.e., $\alpha = \{\alpha_1, \alpha_2, \cdots, \alpha_{n-1}\}$.

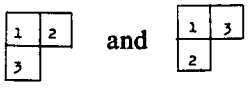
6. CONCLUSIONS

Theorem 4 places important necessary conditions on spin-free reduced density matrices. The restrictions confine the pth-order density matrices to a subspace of the tensor space U^p_p. A relatively small fraction of the partitions of $N - 1$ are embedded in a given partition of N. Hence Theorem 4 places strong conditions on the $(N - 1)$st-order density matrices. Relatively larger fractions of the partitions of $N - 2$, $N - 3$, etc. are embedded in a given partition of N. Hence the conditions imposed by the theorem become weaker as the reduction is continued. When $p = 2$, the two partitions $\{1, 1\}$ and $\{2\}$ are *both* embedded in a given α of N unless $\alpha = \{1^N\}$ or $\alpha = \{N\}$, so that the theorem is weak indeed. Finally, when $p = 1$ the trivial partition $\{1\}$ is embedded in *every* α of N so that the theorem imposes no restriction.

There is, however, an indirect restriction on 1st- and 2nd-order density matrices imposed by this theorem. This arises because of the one–one correspondence between 1st order and $(N - 1)$st order, and between 2nd order and $(N - 2)$nd order. These cases of $p = N - 1$ and $p = N - 2$ are the very ones in which Theorem 4 imposes the strongest conditions. Hence one may yet expect Theorem 4 to give useful conditions on first and second order spin-free density matrices.

This aspect of the problem is to be considered in a later paper.

Theorem 5 provides sufficient conditions for N,α,α-derivability. If these are combined with conditions which guarantee that a tensor of type (N, N) be a density matrix, we have sufficient conditions for N,α,α-representability. The allowed spin-free symmetries of electron systems have no more than two columns. Hence the only application of Theorem 5 to density matrices is $U^{3\to 2}_{\{2,1\}\{2,1\}} \supseteq U^{2}_{\{1,1\}\{1,1\}}$. This suggests that every "triplet" 2nd-order density matrix is 3, "doublet"-representable. Unfortunately this is not true, because one of the conditions that a (3, 3)-type tensor be a spin-free density matrix is that the two structures



have the same weight.[3] Hence the contraction of such a density matrix has both {1, 1} and {2} symmetry. In other words the theorem gives sufficient conditions for N,α,α-derivability of tensors but not for N,α,α-representability of density matrices.

Theorems 4 and 5 completely answer the problem of the permutational symmetries of reduced tensors. The corollary of Theorem 4 shows the way in which one should resolve a tensor into symmetry adapted components so as to be able to specify its symmetry components after reduction. The necessary and sufficient conditions for N,α,β-derivability are not to be found in permutational symmetry. Instead, one must look at the influence of symmetry on other properties.

ACKNOWLEDGMENTS

The author thanks Professor R. McWeeny for providing the facilities and atmosphere for this research, and the National Science Foundation for its generous support.

APPENDIX

In this appendix we are concerned with those spaces $U^{N\to p}_{\alpha\beta}$ which vanish because α and β are "incompatible" symmetries for contraction to order p. As regards density matrices, this corresponds to a type of "forbidden transition." The expectation value of *any* p-particle operator between states of symmetries α and β vanishes if $U^{N\to p}_{\alpha\beta} = 0$.

The kernel of the mapping $\mathrm{Tr}^{(q)}: U^N_N \to U^p_p$ consists of those tensors of type (N, N) whose (q)-trace vanish. The kernel of $\mathrm{Tr}^{(q)}: U^N_{\alpha\beta} \to U^p_p$ may consist of the entire algebra $U^N_{\alpha\beta}$ if the partitions α and β are sufficiently far removed in dictionary order. The special case $U^{N\to 0}_{\alpha\beta}$ is trivial since $U^{N\to 0}_{\alpha\beta} = \delta_{\alpha\beta}U^0_0$. The general case can be studied with the following.

Lemma: Let $\pi \in S_{(q)}$ and let $X \in U^N_N$, then

$$\mathrm{Tr}^{(q)}X = \mathrm{Tr}^{(q)}(\pi X\pi^{-1}).$$

Proof: The proof follows easily from the definitions.

Theorem: $U^{N\to p}_{\alpha\beta} = 0$ if α and β cannot yield the same shape ξ on deletion of p cells.

Proof: Instead of the standard tableaux, use the "anti-standard" tableaux (those in which the integers *descend* from left to right and top to bottom) to form seminormal idempotents which can then be adapted to the subgroup chain on the last integers

$$S_N \supset \cdots \supset S_{(q)} \supset \cdots \{1, (N-1, N)\} \supset 1.$$

Consider the element $x \in e_r U^N_N e_s$ in which e_r and e_s are seminormal units adapted to $S_{(q)}$. Then $e_\rho x = xe_\sigma = x$, where ρ and σ are the anti-standard tableaux which result when the integers $1, 2, \cdots, p$ are struck from r and s. Now

$$\mathrm{Tr}^{(q)}x = \mathrm{Tr}^{(q)}(xe_\sigma),$$

the idempotent e_σ is an element of $A_{(q)}$ and thus can be written as

$$e_\sigma = \sum_\pi c_\pi \pi, \quad \text{where} \quad \pi \in S_{(q)}.$$

By the Lemma we have

$$\begin{aligned}\mathrm{Tr}^{(q)}x &= \sum_\pi c_\pi \mathrm{Tr}^{(q)}(x\pi)\\ &= \sum_\pi c_\pi \mathrm{Tr}^{(q)}(\pi x)\\ &= \mathrm{Tr}^{(q)}(e_\sigma x).\end{aligned}$$

But $e_\sigma e_\rho$, and therefore $e_\sigma x$, vanishes unless $\sigma = \rho$. Hence $U^{N\to p}_{\alpha\beta}$ vanishes if there is no shape ξ of q embedded in both α and β. This proves the Theorem.

The Theorem has important applications to spin-free density matrices. Consider a spin-free transition density matrix of a six nucleon system, between symmetries $\alpha = \{3, 3\}$ (e.g., spin quantum number, $S = \frac{1}{2}$ and isospin quantum number, $T = \frac{1}{2}$) and $\beta = \{2, 1^4\}$ (e.g., spin quantum number, $S = 1$ and isospin quantum number, $T = 2$). Deletion of two cells from each of α and β produces the shapes $\xi = \{2^2\}, \{3, 1\}$, and $\zeta = \{2, 1^2\}, \{1^4\}$. Since there is no pair $\xi = \zeta$, the 2nd-order spin-free density matrix vanishes. Hence no two-particle spin-free force produces a transition between states of symmetries α and β.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 8, NUMBER 4 APRIL 1967

Exponential Operators and Parameter Differentiation in Quantum Physics

R. M. WILCOX
National Bureau of Standards, Boulder, Colorado

(Received 23 August 1966)

Elementary parameter-differentiation techniques are developed to systematically derive a wide variety of operator identities, expansions, and solutions to differential equations of interest to quantum physics. The treatment is largely centered around a general closed formula for the derivative of an exponential operator with respect to a parameter. Derivations are given of the Baker–Campbell–Hausdorff formula and its dual, the Zassenhaus formula. The continuous analogs of these formulas which solve the differential equation $dY(t)/dt = A(t)Y(t)$, the solutions of Magnus and Fer, respectively, are similarly derived in a recursive manner which manifestly displays the general repeated-commutator nature of these expansions and which is quite suitable for computer programming. An expansion recently obtained by Kumar and another new expansion are shown to be derivable from the Fer and Magnus solutions, respectively, in the same way. Useful similarity transformations involving linear combinations of elements of a Lie algebra are obtained. Some cases where the product $e^A e^B$ can be written as a closed-form single exponential are considered which generalize results of Sack and of Weiss and Maradudin. Closed-form single-exponential solutions to the differential equation $dY(t)/dt = A(t)Y(t)$ are obtained for two cases and compared with the corresponding multiple-exponential solutions of Wei and Norman. Normal ordering of operators is also treated and derivations, corollaries, or generalization of a number of known results are efficiently obtained. Higher derivatives of exponential and general operators are discussed by means of a formula due to Poincaré which is the operator analog of the Cauchy integral formula of complex variable theory. It is shown how results obtained by Aizu for matrix elements and traces of derivatives may be readily derived from the Poincaré formula. Some applications of the results of this paper to quantum statistics and to the Weyl prescription for converting a classical function to a quantum operator are given. A corollary to a theorem of Bloch is obtained which permits one to obtain harmonic-oscillator canonical-ensemble averages of general operators defined by the Weyl prescription. Solutions of the density-matrix equation are also discussed. It is shown that an initially canonical ensemble behaves as though its temperature remains constant with a "canonical distribution" determined by a certain fictitious Hamiltonian.

1. INTRODUCTION

OPERATOR identities, expansions, and solutions to differential equations occur widely in quantum physics and have been derived with the aid of a variety of abstract or complicated methods. These include functional analysis, Lie algebra theory, the Feynman ordering-operator calculus,[1] the commutator superoperator, special function theory, and special methods which appear to be of limited use. Although we do not doubt the power and usefulness of some of these methods, it is interesting to see what may be accomplished with a few simple but versatile tools. One of these tools, the parameter differentiation of quantum-mechanical linear operators, has been instructively discussed by Aizu.[2] However, that treatment is confined to obtaining matrix elements and traces involving derivatives. We are here particularly interested in identities involving operators themselves, especially exponential operators. A device which we make extensive use of in establishing identities involving exponential operators is to require that both sides of an equation satisfy the same first-order differential equation and the same initial condition. We refer to this tool as *the differential equation method.*

Another device, used to obtain results for general operator functions from those involving exponential operators, is to assume that the general functions can be expressed as linear combinations of exponential operators. We refer to this tool as *the method of linear superposition.* Special cases of it are Fourier and Laplace series or integrals. This procedure is often easier to apply than the often-used method which constructs general functions from linear combinations of powers of operators,[3] but the set of functions which may be represented by either method appears to be the same. Using mainly these tools, we intend to derive a variety of scattered results in a concise systematic and elementary manner which many physicists may find easier to understand. Nevertheless, we believe that we have occasionally obtained a new application or generalization of a known result.

[1] R. P. Feynman, Phys. Rev. **84**, 108 (1951).
[2] K. Aizu, J. Math. Phys. **4**, 762 (1963).
[3] W. H. Louisell, *Radiation and Noise in Quantum Electronics* (McGraw-Hill Book Company, Inc., New York, 1964), Chap. 3, p. 98.

Although everyone knows how to differentiate an ordinary exponential function, it is not widely realized that a general closed formula exists for the derivative of an exponential operator with respect to a parameter. [See Eq. (2.1) of Sec. 2.] We believe that this formula is sufficiently simple and useful that many physicists and applied mathematicians can profitably commit it to memory. In Sec. 2, this formula is verified by the differential equation method and compared with formulas of Feynman[1] and Kubo.[4] A basic lemma is derived which displays the formal correspondence between parameter differentiation and commutation. The concept of differentiation with respect to an operator is also discussed. In Sec. 3, Eq. (2.1) is shown to be of use for problems in equilibrium quantum statistics. In Sec. 4, Eq. (2.1) and the differential equation method are used to easily derive the Baker–Campbell–Hausdorff (BCH) formula.[5] This important formula determines Z such that $e^A e^B = e^Z$ is identically satisfied. A good review of the history of this formula is given by Weiss and Maradudin.[6] A formula of Zassenhaus,[7] said to be the dual of the BCH formula, is also derived in Sec. 4 in a similar manner. This formula expresses e^{A+B} as an infinite product of exponential operators. A number of examples of the type of problem which exploits the BCH formula occur in connection with the Weyl prescription for converting a classical function to a quantum operator. This is described in Sec. 5. A more direct derivation of a useful formula recently obtained by Daughaday and Nigam[8] is given there, while interesting theorems of McCoy,[9] Moyal, and Wigner[10] are stated without proof.

In Sec. 6, a definition of a Lie algebra is given and examples of various Lie algebras which occur in quantum mechanics are given. Similarity transformations involving operators which are linear combinations of Lie elements (denoted LCLE) are readily obtained by the differential equation method. The results may be used to "change the representation" in quantum-mechanical problems, and are also used in Secs. 7, 8, and 10. In Sec. 7, Eq. (2.1) and the differential equation method are used to obtain closed-form expressions for $Z = \ln(e^A e^B)$ for certain cases where A and B are LCLE's. Generalizations of formulas obtained by Sack[11] and by Weiss and Maradudin[6] are obtained. A corollary to a theorem of Bloch[12] is also derived which is useful for obtaining harmonic-oscillator thermal averages of general operator functions.

In Sec. 8, we obtain solutions to the important operator differential equation, $dY(t)/dt = A(t)Y(t)$. In Sec. 8.1, we obtain the Magnus[7] and Fer[13] solutions in a manner similar to that used to derive the BCH and Zassenhaus formulas, respectively. It should be noted that all four of these formulas are derived by a recursion procedure which manifestly displays the repeated-commutator nature of these expansions and which is quite suitable for computer programming. An expansion recently obtained by Kumar[14] and another new expansion are shown to be derivable from the Fer and Magnus solutions, respectively, in the same way. Both of these expansions may be more suitable than the Zassenhaus expansion for certain purposes. In Sec. 8.2, Lie algebraic solutions in terms of a single exponential are obtained and compared with corresponding multiple-exponential solutions of Wei and Norman.[15] In Sec. 9, solutions of the density matrix equation, Eq. (9.1), are discussed. It is shown that an initially canonical ensemble behaves as though its temperature remains constant with a "canonical distribution" determined by a certain fictitious Hamiltonian.

In Sec. 10, a collection of old and recent results involving normal ordering of operators are efficiently derived with the aid of the results and methods of previous sections. Derivations, corollaries, or generalizations of formulas obtained by Louisell,[3] Heffner and Louisell,[16] Schwinger,[17] McCoy,[18] Kermack and McCrea,[19] and Cohen[20] are given. In Sec. 11, higher derivatives of exponential operators are treated with a formula due to Poincaré.[21] [See Eq. (10.1).] Although

[4] *Lectures in Theoretical Physics* R. Kubo, W. E. Brittin, and L. G. Dunham, Eds. (Interscience Publishers, Inc., New York, 1959), Vol. I, p. 139, Eq. (2.17).

[5] J. E. Campbell, Proc. London Math. Soc. **29**, 14 (1898); H. F. Baker, *ibid.* **34**, 347 (1902); **35**, 333 (1903); **2**, 293 (1904); **3**, 24 (1904); F. Hausdorff, Ber. Verhandl. Saechs. Akad. Wiss. Leipzig, Math.-Naturw. Kl. **58**, 19 (1906); N. Jacobson, *Lie Algebras* (Interscience Publishers, Inc., New York, 1962), Chap. 5, p. 170.

[6] G. H. Weiss and A. A. Maradudin, J. Math. Phys. **3**, 771 (1962).

[7] W. Magnus, Commun. Pure Appl. Math. **7**, 649 (1954).

[8] H. Daughaday and B. P. Nigam, Phys. Rev. **139**, B1436 (1965).

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

[10] J. E. Moyal, Proc. Cambridge Phil. Soc. **45**, 99 (1949); E. P. Wigner, Phys. Rev. **40**, 749 (1932); C. L. Mehta, J. Math. Phys. **5**, 677 (1964).

[11] R. A. Sack, Phil. Mag. **3**, 497 (1958).

[12] F. Bloch, Z. Physik **74**, 295 (1932).

[13] F. Fer, Bull. Classe Sci. Acad. Roy. Belg. **44**, 818 (1958).

[14] K. Kumar, J. Math. Phys. **6**, 1928 (1965).

[15] J. Wei and E. Norman, J. Math. Phys. **4**, 575 (1963).

[16] H. Heffner and W. H. Louisell, J. Math. Phys. **6**, 474 (1965).

[17] *Quantum Theory of Angular Momentum*, J. Schwinger, L. C. Biedenharn, and H. van Dam, Eds. (Academic Press Inc., New York, 1965), pp. 274–276.

[18] N. H. McCoy, Proc. Edinburgh Math. Soc. **3**, 118 (1932).

[19] W. O. Kermack and W. H. McCrea, Proc. Edinburgh Math. Soc. **2**, 220 (1931).

[20] L. Cohen, J. Math. Phys. **7**, 244 (1966).

[21] H. Poincaré, Trans. Cambridge Phil. Soc. **18**, 220 (1899); see also R. Bellman, *Introduction to Matrix Analysis* (McGraw-Hill Book Company, Inc., New York, 1960), p. 103, Ex. 43. This formula is the basis of resolvent theory. See, e.g., Messiah, Ref. 48, Chap. 16, Sec. 3, p. 712.

even greater formal efficiency may be obtained with the Feynman ordering-operator calculus,[1] we have preferred to use the Poincaré formula since it is more firmly rooted in classical analysis and since it readily permits one to differentiate *general* operator functions in a concise manner. It is shown to provide a compact alternative for the derivation of formulas of the type derived by Aizu.[2] We note that such formulas may be used to derive all of the interesting sum rules and hypervirial theorems treated in a recent paper by Morgan and Landsberg.[22]

2. DERIVATIVE OF EXPONENTIAL OPERATOR

If the operator H is a function of a parameter λ, $H \equiv H(\lambda)$, then

$$\frac{\partial}{\partial\lambda} e^{-\beta H} = -\int_0^\beta e^{-(\beta-u)H} \frac{\partial H}{\partial\lambda} e^{-uH}\, du. \quad (2.1)$$

This identity (aside from notation) has been derived and used by Snider in treating a quantum Boltzmann equation.[23] Snider's derivation was based upon the commutator superoperator and an integral representation for the beta function. The present author found Eq. (2.1) independently by employing the parameter-differentiation technique explained by Aizu,[2] together with the well-known expansion[24]

$$e^A B e^{-A} = B + [A, B] + (1/2!)[A, [A, B]] + \cdots. \quad (2.2)$$

The author has applied Eq. (1) to the calculation of the polarizability of a one-dimensional NaCl lattice.[25] Equation (2.1) has also been obtained recently by Kumar,[26] but the full generality of the result is not obvious in that treatment.

Like any operator identity, Eq. (2.1) may be verified by showing that the matrix elements of both sides of the equation are the same in some suitable representation. This may be done, using Aizu's techniques,[2] by choosing a representation in which $H(\lambda)$ is diagonal. However, an easier method is to show that both sides of the equation satisfy the first-order differential equation

$$[\partial F(\beta)/\partial\beta] + HF(\beta) = -(\partial H/\partial\lambda)e^{-\beta H}, \quad (2.3)$$

with the initial condition $F(0) = 0$.

A special case of Eq. (2.1) is an identity of Feynman,[27]

$$\left[\frac{d}{d\epsilon} e^{\alpha+\epsilon\beta}\right]_{\epsilon=0} = \int_0^1 e^{(1-s)\alpha}\beta e^{s\alpha}\, ds, \quad (2.4)$$

where α and β are independent of ϵ. Equation (2.4) has been obtained by Feynman by means of his ordering-operator calculus, and this method may also be used to derive Eq. (2.1).[28] Also, Eq. (2.1) may be obtained from Eq. (4) by considering the Taylor expansion

$$H(\lambda + \epsilon) = H(\lambda) + \epsilon(\partial H/\partial\lambda) + O(\epsilon^2).$$

However, Eq. (2.1) is a more convenient form to use than Eq. (2.4), particularly for differential equations.

An identity of Kubo is a corollary of Eq. (2.1).[4]

$$[A, e^{-\beta\mathcal{H}}] = -\int_0^\beta du e^{-(\beta-u)\mathcal{H}}[A, \mathcal{H}]e^{-u\mathcal{H}}. \quad (2.5)$$

In Eq. (2.5), A and $\mathcal{H}$ are arbitrary operators. This identity has had much use in the theory of irreversible processes. Like Eq. (2.1), it is easily established by the differential equation method, as Montroll has noted.[29] It may also be derived from Eq. (2.1) by making the similarity transformation

$$\begin{aligned} H(\lambda) &= e^{\lambda A}\mathcal{H}e^{-\lambda A}, \\ e^{-\beta H} &= e^{\lambda A}e^{-\beta\mathcal{H}}e^{-\lambda A}. \end{aligned} \quad (2.6)$$

Conversely, Eq. (2.1) follows from Eq. (2.5) by setting $\mathcal{H} = H(\lambda)$ and $A = \partial/\partial\lambda$. Another form of Eq. (2.5), obtained by setting $\beta = -it/\hbar$, occurs as the solution to the Heisenberg equation of motion, $\dot{A}(t) = i\hbar^{-1}[\mathcal{H}, A(t)]$, for $\mathcal{H}$ independent of t. A generalization to the case where $\mathcal{H}$ is time dependent may also be readily established by the differential equation method.[30]

Equations (2.1) and (2.5) may be used to derive a useful basic lemma.

Lemma: If $[A, H(\lambda)] = \partial H/\partial\lambda$, then $[A, f(H)] = \partial f(H)/\partial\lambda$. (Note: in the language of Lie algebra, A is said to be the generator of an infinitesimal transformation due to a change in the parameter λ.)

Proof: Equations (2.1) and (2.5) imply this result for the special case where $f(H) = e^{-\beta H}$, while the general case easily follows by the method of linear superposition.

[22] D. J. Morgan and P. T. Landsberg, Proc. Phys. Soc. (London) **86**, 261 (1965).
[23] R. F. Snider, J. Math. Phys. **5**, 1586 (1964), Appendix B.
[24] See, e.g., Louisell, Ref. 3, p. 101, Eq. (3. 14).
[25] R. M. Wilcox, National Bureau of Standards Report (1964).
[26] Reference 14, Eqs. (46) and (A1).
[27] Reference 1, Eq. (6). A related identity occurs in R. Karplus and J. Schwinger, Phys. Rev. **73**, 1025 (1948), Eq. (I. 8).
[28] The author is indebted to Dr. J. H. Shirley for pointing this out to him.
[29] E. Montroll, in *Lectures in Theoretical Physics*, W. E. Brittin, B. W. Downs, and J. Downs, Eds. (Interscience Publishers, Inc., New York, 1961), Vol. III, p. 259, Eq. (XI.6).
[30] See, e.g., R. L. Peterson, Rev. Mod. Phys. **39**, 69 (1967), Eq. (23).

Another kind of differentiation which frequently occurs in quantum physics is the differentiation of an operator by an operator. This may be defined by means of parameter differentiation as follows.[31] Let $H \equiv H(Q_1, Q_2, \cdots, Q_n)$ be a function of the operators $Q_1, Q_2, \cdots, Q_n$ which need not commute with each other. Then the operator derivative with respect to Q_j is defined by

$$\frac{\partial H}{\partial Q_j} \equiv \lim_{\lambda \to 0} \frac{\partial H}{\partial \lambda}(Q_1, \cdots, Q_j + \lambda, \cdots, Q_n). \quad (2.7)$$

From Eqs. (2.1) and (2.7), it follows that

$$\frac{\partial e^{-\beta H}}{\partial Q_j} = -\int_0^\beta du e^{-(\beta-u)H} \frac{\partial H}{\partial Q_j} e^{-uH}. \quad (2.8)$$

The above basic lemma is also clearly valid with λ replaced by Q_j. To obtain a familiar result which we make use of later, let $A = q$, $H = p$, $\lambda = p/i\hbar$, so that $[q, p] = i\hbar = i\hbar\partial p/\partial p$. Then by the lemma,

$$[q, f(p)] = i\hbar \partial f(p)/\partial p. \quad (2.9)$$

Letting $f(p) = e^{-i\mu p/\hbar}$, Eq. (2.9) becomes

$$[q, e^{-i\mu p/\hbar}] = \mu e^{-i\mu p/\hbar}. \quad (2.10)$$

From Eq. (2.10), it follows by a simple standard argument that if $|q'\rangle$ is an eigenstate of q so that $q|q'\rangle = q'|q'\rangle$, then $e^{-i\mu p/\hbar}$ acts as a displacement operator,[32]

$$e^{-i\mu p/\hbar}|q'\rangle = |q' + \mu\rangle. \quad (2.11)$$

3. APPLICATION OF EQ. (2.1) TO EQUILIBRIUM QUANTUM STATISTICS

Equation (2.1) is well suited for applications to equilibrium quantum statistics, where the thermal average of any operator Q is given by

$$\langle Q \rangle = \mathrm{Tr}\,[e^{-\beta\mathcal{H}}Q]/\mathrm{Tr}\, e^{-\beta\mathcal{H}}. \quad (3.1)$$

In Eq. (1), β denotes $(KT)^{-1}$, where K is Boltzmann's constant and T is the absolute temperature. We assume that $\mathcal{H}$ depends upon n parameters λ_i, $\mathcal{H} = \mathcal{H}(\lambda_1, \lambda_2, \cdots, \lambda_n)$, and that Q is independent of λ_i. Then it easily follows that

$$\frac{\partial\langle Q\rangle}{\partial\lambda_i} = -\int_0^\beta du \left\langle e^{u\mathcal{H}} \frac{\partial\mathcal{H}}{\partial\lambda_i} e^{-u\mathcal{H}} Q \right\rangle + \beta \left\langle \frac{\partial\mathcal{H}}{\partial\lambda_i} \right\rangle \langle Q\rangle$$

$$= -\int_0^\beta du \left\langle e^{u\mathcal{H}} \Delta \frac{\partial\mathcal{H}}{\partial\lambda_i} e^{-u\mathcal{H}} \Delta Q \right\rangle, \quad (3.2)$$

where $\Delta Q \equiv Q - \langle Q\rangle$, etc. If we consider the case where the Hamiltonian has a perturbation λV added to an unperturbed part $\mathcal{H}_0$, so that $\mathcal{H} = \mathcal{H}_0 + \lambda V$, then Eq. (3.2) becomes

$$\frac{\partial\langle V\rangle}{\partial\lambda} = -\int_0^\beta du \langle e^{u\mathcal{H}} \Delta V e^{-u\mathcal{H}} \Delta V\rangle. \quad (3.3)$$

Assuming, for notational convenience, discrete energy levels $\mathcal{H}_i$ or $\mathcal{H}_j$, one easily obtains

$$\frac{\partial\langle V\rangle}{\partial\lambda} = -(\mathrm{Tr}\, e^{-\beta\mathcal{H}})^{-1} \sum_{i,j} |\langle i|\Delta V|j\rangle|^2 \times \int_0^\beta du \exp(-\beta\mathcal{H}_i + u\mathcal{H}_i - u\mathcal{H}_j), \quad (3.4)$$

which is seen to be nonpositive.[33] This result is the quantum-statistical analog of the well-known result that the second-order perturbation energy of the ground state is always negative. The latter result may be obtained from Eq. (3.4) as a special case by evaluating the integral in Eq. (3.4) and letting β become arbitrarily large.

Another application of Eq. (3.2) is to a system with dipole moment $\mathbf{M}$ subjected to an applied field $\mathbf{E}$, so that it is described by the Hamiltonian

$$\mathcal{H} = \mathcal{H}_0 - \mathbf{M}\cdot\mathbf{E}. \quad (3.5)$$

The field-dependent isothermal static susceptibility tensor $\chi_{\alpha\beta}(\mathbf{E})$ for a sample of unit volume is defined by

$$\chi_{\alpha\beta}(\mathbf{E}) \equiv \partial\langle M_\alpha\rangle/\partial E_\beta, \quad (3.6)$$

where $\alpha, \beta = x, y$, or z. It follows from Eqs. (2) and (5) that

$$\chi_{\alpha\beta}(\mathbf{E}) = \int_0^\beta du \langle e^{u\mathcal{H}} \Delta M_\beta e^{-u\mathcal{H}} \Delta M_\alpha\rangle,$$

a result given by Kubo for the zero-field case.[34]

Before leaving this section, we note that thermodynamic perturbation theory may be conveniently based upon Eq. (2.1).[35] For lattice-dynamical systems, a diagrammatic analysis may be developed which is topologically the same as the ones given by Cowley for nonequilibrium quantum systems, and by the author for classical thermostatic systems.[36]

4. THE BAKER–CAMPBELL–HAUSDORFF FORMULA AND RELATED IDENTITIES

To derive the BCH formula and for later applications, we prefer to use the form

$$\partial e^Z/\partial\lambda = \int_0^1 dx e^{xZ} Z'(\lambda) e^{-xZ} e^Z \quad (4.1)$$

[31] See, e.g., Louisell, Ref. 3, p. 108, Eqs. (3.34).
[32] See, e.g., Louisell, Ref. 3, Sec. 1.11.
[33] C. Kittel, *Quantum Theory of Solids* (John Wiley & Sons, Inc., New York, 1963), p. 127, Problem 2.
[34] Kubo, Ref. 4, Eq. (2.48).
[35] L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1958), Chap. 3, Sec. 32, p. 93.
[36] R. A. Cowley, Advan. Phys. **12**, 421 (1963); R. M. Wilcox, Phys. Rev. **139**, A1281 (1965).

obtained from Eq. (2.1) by substituting $Z \equiv Z(\lambda)$ for $-H$, 1 for β, and $1 - x$ for u. We seek to express Z as a power series in λ such that

$$e^{Z} = e^{\lambda A} e^{\lambda B} \tag{4.2}$$

is identically satisfied. Thus,

$$Z = \sum_{n=1}^{\infty} \lambda^n Z_n, \quad Z'(\lambda) = \sum_{n=1}^{\infty} n\lambda^{n-1} Z_n, \tag{4.3}$$

where the Z_n are to be determined. Note that $Z = 0$ when $\lambda = 0$ as required by Eq. (4.2). Differentiating Eq. (4.2) with respect to λ and multiplying both sides from the right by $e^{-Z} = e^{-\lambda B} e^{-\lambda A}$, one obtains

$$\int_0^1 dx e^{xZ} Z'(\lambda) e^{-xZ} = A + e^{\lambda A} B e^{-\lambda A}. \tag{4.4}$$

The quantity $e^{\lambda A} B e^{-\lambda A}$ is easily expanded by Eq. (2.2) in the form

$$e^{\lambda A} B e^{-\lambda A} = \sum_{j=0}^{\infty} \lambda^j (j!)^{-1} \{A^j, B\}, \tag{4.5}$$

where the repeated commutator bracket is defined inductively by[37]

$$\{A^0, B\} = B, \quad \{A^{n+1}, B\} = [A, \{A^n, B\}]. \tag{4.6}$$

Similarly, since $\int_0^1 x^j \, dx = 1/(j+1)$, Eq. (4.5) implies that

$$\int_0^1 dx e^{xZ} Z' e^{-xZ} = \sum_{k=0}^{\infty} \frac{\{Z^k, Z'\}}{(k+1)!}. \tag{4.7}$$

Substituting Eqs. (4.3), (4.5), and (4.7) into Eq. (4.4), we obtain

$$\sum_{k=0}^{\infty} \left\{ \frac{1}{(k+1)!} \left(\sum_{m=1}^{\infty} \lambda^m Z_m \right)^k, \sum_{n=1}^{\infty} n\lambda^{n-1} Z_n \right\} = A + \sum_{j=0}^{\infty} \lambda^j (j!)^{-1} \{A^j, B\}. \tag{4.8}$$

Since Eq. (8) must be satisfied identically in λ, we equate coefficients of λ^j on the two sides of the equation. For $j = 0$, one obtains

$$Z_1 = A + B. \tag{4.9}$$

For $j = 1$, one obtains

$$\{Z^0, 2Z_2\} + \tfrac{1}{2}\{Z_1, Z_1\} = \{A, B\} \tag{4.10}$$

or

$$Z_2 = \tfrac{1}{2}[A, B]. \tag{4.11}$$

For $j = 2$, one obtains

$$\{Z^0, 3Z_3\} + \tfrac{1}{2}\{Z_1, 2Z_2\} + \tfrac{1}{2}\{Z_2, Z_1\} + \tfrac{1}{6}\{Z_1^2, Z_1\} = \tfrac{1}{2}\{A^2, B\}. \tag{4.12}$$

Equation (12) may be simplified by means of Eqs. (4.6), (4.9), and (4.11) to obtain

$$Z_3 = \tfrac{1}{12}[A, [A, B]] + \tfrac{1}{12}[[A, B], B]. \tag{4.13}$$

The BCH formula to third order is obtained by substituting Eqs. (4.3), (4.9), (4.11), and (4.13) into Eq. (4.2) and setting $\lambda = 1$:

$$e^A e^B = \exp\{A + B + \tfrac{1}{2}[A, B] + \tfrac{1}{12}[A, A, B] + \tfrac{1}{12}[[A, B], B] + \cdots\}. \tag{4.14}$$

The recursion scheme based upon Eq. (4.8) may, in principle, be carried out to arbitrarily high order. Weiss and Maradudin have manually calculated Z out to the fifth order,[6] while Richtmyer and Greenspan have calculated Z out to the tenth order by computer.[38] The expansion is not unique due to the existence of the Jacobi identity

$$[[A, B], C] + [[C, A], B] + [[B, C], A] = 0,$$

the identity

$$[[[A, B], C], D] + [[[B, C], D], A] + [[[C, D], A], B] + [[[D, A], B], C] = [[A, C], [B, D]],$$

and others.

It frequently happens that the commutator of A and B commutes with both A and B. In this case, Eq. (4.14) reduces to

$$e^A e^B = e^{A+B+\frac{1}{2}C} = e^{A+B} e^{\frac{1}{2}C} = e^{\frac{1}{2}C} e^{A+B}, \tag{4.15}$$

where $C \equiv [A, B]$. Equation (4.15) has been derived in various ways and is frequently employed in physical problems.[39] Another form of Eq. (4.15), which has been used by Moyal and by Sudarshan, is[10,40]

$$e^B e^{2A} e^B = e^{2A+2B}. \tag{4.16}$$

Equation (4.16) follows from Eq. (4.15) if one interchanges the order of A and B in Eq. (4.15),

$$e^B e^A = e^{A+B} e^{-\frac{1}{2}C}, \tag{4.17}$$

and then multiplies Eq. (17) by Eq. (15) from the right. Another derivation of Eq. (4.16) has been given by Sudarshan.[40]

To derive the Zassenhaus formula, we set

$$e^{\lambda(A+B)} = e^{\lambda A} e^{\lambda B} e^{\lambda^2 C_2} e^{\lambda^3 C_3} \cdots. \tag{4.18}$$

Differentiating both sides of Eq. (4.18) with respect to λ and multiplying it from the right by

$$e^{-\lambda(A+B)} = \cdots e^{-\lambda^3 C_3} e^{-\lambda^2 C_2} e^{-\lambda B} e^{-\lambda A},$$

[37] Note that our convention for the repeated-commutator bracket is opposite to that of Weiss and Maradudin, Ref. 6.

[38] R. D. Richtmyer and S. Greenspan, Commun. Pure Appl. Math. **18**, 107 (1965); The author is indebted to Dr. Greenspan for a private communication.

[39] A partial list of references includes Refs. 3, 8, 9, 10, 18, 19, 48, 50, 52, and 53.

[40] E. C. G. Sudarshan, in *Brandeis Summer Institute Lectures in Theoretical Physics* (W. A. Benjamin, Inc., New York, 1962), p. 181.

one obtains

$$A + B = A + e^{\lambda A}Be^{-\lambda A} + e^{\lambda A} + e^{\lambda B}(2\lambda C_2)e^{-\lambda B}e^{-\lambda A} + e^{\lambda A}e^{\lambda B}e^{\lambda^2 C_2}(3\lambda^2 C_3)e^{-\lambda^2 C_2}e^{-\lambda B}e^{-\lambda A} + \cdots. \tag{4.19}$$

The quantities $e^{\lambda A}Be^{-\lambda A}$, etc., and again expanded by means of Eq. (4.5) so that Eq. (4.19) becomes

$$0 = \sum_{n=1}^{\infty}\frac{\lambda^n}{n!}\{A^n, B\} + 2\lambda\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{\lambda^{m+n}}{m!\,n!}\{A^m, B^n, C_2\} + 3\lambda^2\sum_{k=0}^{\infty}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{\lambda^{k+m+2n}}{k!\,m!\,n!}\{A^k, B^m, C_2^n, C_3\} + \cdots. \tag{4.20}$$

The quantities $\{A^n, B\}$ are defined by Eq. (6), while the quantities $\{A^m, B^n, C_2\}$ are defined inductively as

$$\{A^0, B^n, C_2\} = \{B^n, C_2\}, \tag{4.21a}$$

$$\{A^{m+1}, B^n, C_2\} = [A, \{A^m, B^n, C_2\}]. \tag{4.21b}$$

Higher-order repeated-commutator brackets are similarly defined. Clearly the coefficients of λ^j must vanish in Eq. (20). Setting $j = 1$, one obtains

$$C_2 = -\tfrac{1}{2}[A, B]. \tag{4.22}$$

Setting $j = 2$, one obtains

$$0 = \tfrac{1}{2}\{A^2, B\} + 2\{A, B^0, C_2\} + 2\{A^0, B, C_2\} + 3\{A^0, B^0, C_2^0, C_3\}$$

or, upon using Eqs. (4.6), (4.21), and (4.22),

$$C_3 = \tfrac{1}{3}[B, [A, B]] + \tfrac{1}{6}[A, [A, B]]. \tag{4.23}$$

We note that both the BCH formula and the Zassenhaus formula could have been derived a little more simply by repeatedly differentiating Eqs. (4.2) and (4.18), respectively, and setting $\lambda = 0$ after each differentiation. Although Z_n and C_n may be obtained in a recursive manner by this procedure, this method does not manifestly display the general repeated-commutator nature of the expansion. This method has been used by Huang[41] to obtain the first three factors in the Zassenhaus formula. Huang has used the result to treat quantum deviations from the classical limit of the partition function for both a Bose–Einstein and a Fermi–Dirac gas.

5. WEYL'S PRESCRIPTION FOR QUANTIZING A CLASSICAL FUNCTION

Let p, q denote a conjugate pair of canonical variables of classical mechanics, and let P, Q denote the corresponding quantum operators,[42]

$$[Q, P] = i\hbar. \tag{5.1}$$

Then, on the basis of group-theoretical considerations, Weyl has proposed that the quantum operator $F(P, Q)$ corresponding to a given classical function $f(p, q)$ be represented by the Fourier integral[43]

$$F(P, Q) = \iint_{-\infty}^{\infty} g(\sigma, \tau)e^{i(\sigma P + \tau Q)}\,d\sigma\,d\tau, \tag{5.2}$$

where $g(\sigma, \tau)$ is the Fourier transform of the classical function $f(p, q)$. The generalization of the Weyl prescription to the case of n independent pairs of canonical variables has been considered by Daughaday and Nigam (DN).[8] In effect, this amounts to interpreting the quantities occurring in Eq. (2) as

$$P = (P_1, P_2, \cdots, P_n),$$
$$\sigma = (\sigma_1, \sigma_2, \cdots, \sigma_n),$$
$$d\sigma = d\sigma_1\,d\sigma_2\cdots d\sigma_n,$$
$$\sigma P = \sigma_1 P_1 + \sigma_2 P_2 + \cdots + \sigma_u P_u,$$
etc.

By a rather tortuous process, DN obtain the useful result that the quantum function $F(P, Q)$ corresponding to the classical function

$$f(p, q) = p_1^{\alpha}p_2^{\beta}\cdots p_u^{\gamma}\varphi(q_1, q_2, \cdots, q_n)$$

is given by[8]

$$2^{-(\alpha+\beta+\cdots+\gamma)}[\cdots[[\varphi, P_1]_{+\alpha}, P_2]_{+\beta}\cdots P_n]_{+\gamma}, \tag{5.3}$$

where $[\varphi, P_1]_{+\alpha}$ refers to the anticommutator bracket repeated α times. For example,

$$[\varphi, P_1]_{+2} = [\varphi P_1 + P_1\varphi, P_1]_{+1} = \varphi P_1^2 + 2P_1\varphi P_1 + P_1^2\varphi.$$

In DN, Eq. (3) has been used to obtain the quantum Hamiltonian for the case of two charged spinless particles interacting by retarded fields. We show here that Eq. (3) may be derived much more directly with the aid of Eq. (4.16) written in the form

$$e^{i(\sigma P + \tau Q)} = e^{\frac{1}{2}i\sigma P}e^{i\tau Q}e^{\frac{1}{2}i\sigma P}. \tag{5.4}$$

Equation (5.4) is valid, since σP and τQ both commute with their commutator. The Fourier transform of $f(p, q)$, $g(\sigma, \tau)$, is easily found to be given by

$$i^{\alpha+\beta+\cdots+\gamma}\delta^{(\alpha)}(\sigma_1)\delta^{(\beta)}(\sigma_2)\cdots\delta^{(\gamma)}(\sigma_n)u(\tau), \tag{5.5}$$

where, e.g., $\delta^{(\alpha)}(\sigma)$ denotes the αth derivative of the "Dirac δ function" and $u(\tau)$ is the Fourier transform of $\varphi(q)$. One easily obtains Eq. (5.3) by substituting

[41] K. Huang, *Statistical Mechanics* (John Wiley & Sons, Inc., New York, 1963), p. 217, Eq. (10.60).

[42] Only in Sec. 5 do small p and q denote classical variables.

[43] H. Weyl, *The Theory of Groups and Quantum Mechanics* (E. P. Dutton & Co., Inc., New York, 1931), p. 274.

Eqs. (5.4) and (5.5) into Eq. (5.2), and using the well-known formula of distribution theory,[44]

$$\int_{-\infty}^{\infty} \delta^{(\alpha)}(\sigma') f(\sigma')\, d\sigma' = (-)^{\alpha} f^{(\alpha)}(0).$$

We state without proof some other interesting results involving the Weyl prescription which may be derived with the aid of the BCH formula and the Fourier integral theorem.

Theorem (McCoy[9]): Let $F(P, Q)$ be the quantum operator which corresponds to the classical function $f(p, q)$ according to the Weyl prescription, and let $F_Q(P, Q)$ be the function obtained from $F(P, Q)$ by ordering all Q factors to the left of all P factors with the aid of Eq. (5.1). Then $F_Q(P,Q)$ may be obtained by replacing p and q by P and Q with the Q's to the left of the P's in all terms of the power-series expansion of

$$\left[\exp\left(-\tfrac{1}{2}i\hbar\frac{\partial^2}{\partial p\partial q}\right)\right] f(p, q) = f(p, q) - \frac{i\hbar}{2}\frac{\partial^2 f(p, q)}{\partial p\partial q} - \left(\frac{\hbar}{2}\right)^2 \frac{1}{2!}\frac{\partial^4 f(p, q)}{\partial p^2 \partial q^2} + \cdots. \quad (5.6)$$

In DN, Eq. (5.6) is generalized to the case of more than one degree of freedom, and is used to derive Eq. (5.3).

Theorem (Moyal[10]): Let $F(P, Q)$, $G(P, Q)$, and $C(P, Q)$ be defined in the Weyl manner in terms of the classical functions $f(p, q)$, $g(p, q)$, and $c(p, q)$, respectively, and let

$$[F(P, Q), G(P, Q)] = C(P, Q). \quad (5.7)$$

Then

$$c(p, q) = 2i\left[\sin\frac{\hbar}{2}\left(\frac{\partial}{\partial q_1}\frac{\partial}{\partial p_2} - \frac{\partial}{\partial p_1}\frac{\partial}{\partial q_2}\right)\right] \times f(p_1, q_1) g(p_2, q_2), \quad (5.8)$$

evaluated at $p_1 = p_2 = p$ and $q_1 = q_2 = q$.

In the limit as $\hbar \to 0$, this reduces to the well-known relation between the quantum commutator and the classical Poisson bracket. This result and the corresponding law for operator multiplication have been treated by Mehta[45] with the aid of the BCH formula. Moyal's result originated in a paper which formulates quantum mechanics in terms of phase-space distribution functions, and occurs also in Sudarshan's study of the structure of dynamical theories.[40]

Theorem (Wigner[10]): Let $G(P, Q)$ be any operator defined in the Weyl manner, and let

$$\langle G\rangle \equiv \langle\psi|\, G(P, Q)\, |\psi\rangle \quad (5.9)$$

be the expectation value of G in the state $|\psi\rangle$. Then $\langle Q\rangle$ may be calculated from the phase-space distribution function $f(p, q)$ by means of

$$\langle G\rangle = \iint dp\, dq\, g(p, q) f(p, q), \quad (5.10)$$

where

$$f(p, q) \equiv \frac{1}{2\pi}\int \psi^*(q - \tfrac{1}{2}\hbar\tau) e^{-i\tau p} \psi(q + \tfrac{1}{2}\hbar\tau)\, d\tau, \quad (5.11)$$

with $\psi(q) \equiv \langle q \mid \psi\rangle$.

As shown by Moyal,[10] this result may be easily proved with the aid of Eqs. (2.11), (4.16), and the Fourier integral theorem.

6. LIE ALGEBRAIC SIMILARITY TRANSFORMATIONS

In quantum physics, one frequently encounters sets of operators $X_1, X_2, \cdots, X_n$ such that the commutator of any pair is a linear combination of members of the set according to the rule

$$[X_i, X_j] = c_{ij1}X_1 + c_{ij2}X_1 + \cdots + c_{ijn}X_n. \quad (6.1)$$

The X_i's are said to be elements of a Lie algebra, while the c_{ijk} are called the "structure constants" of the algebra.[46] For our purposes here, we do not require any knowledge of this highly abstract and technical subject. The "closure property" expressed by Eq. (6.1) is sufficient. We seek to perform a similarity transformation on X_j of the form $e^Z X_j e^{-Z}$, where Z is a given linear combination of the X_i's,

$$Z \equiv d_1X_1 + d_2X_2 + \cdots + d_nX_n. \quad (6.2)$$

Now from Eqs. (6.1), (6.2), and (2.2) it follows that $e^Z X_j e^{-Z}$ is also a linear combination of the X_i's,

$$e^Z X_j e^{-Z} = \sum_{i=1}^{n} g_{ji} X_i \equiv \mathbf{g}_j \cdot \mathbf{X}. \quad (6.3)$$

From the basic property of a similarity transformation, it then follows that any function of the X_i's, $F \equiv F(X_1, X_2, \cdots, X_n)$, is transformed to

$$e^Z F e^{-Z} = F(\mathbf{g}_1 \cdot \mathbf{X}, \mathbf{g}_2 \cdot \mathbf{X}, \cdots, \mathbf{g}_n \cdot \mathbf{X}). \quad (6.4)$$

Although the g_{ij}'s may in principle be determined from Eqs. (6.1), (6.2), and (2.2), in practice the infinite summation may become very complicated to perform. In this case, a better approach is to introduce a

[44] M. S. Lighthill, *Introduction to Fourier Analysis and Generalized Functions* (Cambridge University Press, Cambridge, England, 1959), p. 19, Eq. (17).

[45] Mehta, Ref. 10.

[46] See, e.g., H. J. Lipkin, *Lie Groups for Pedestrians*, (John Wiley & Sons, Inc., New York, 1965), or Jacobson, Ref. 5.

parameter x into the exponentials and determine the differential equation which $e^{xZ}X_i e^{-xZ}$ must satisfy. It is seen that this procedure generally leads to a coupled set of linear homogeneous first-order differential equations which, as is well known, may be solved as an eigenvalue problem.

The lowest-dimensional Lie algebra of physical interest is the non-Abelian two-dimensional algebra with elements $\{X, Y\}$ satisfying

$$[X, Y] = Y. \tag{6.5}$$

This algebra has been treated by Sack[11] in connection with his "Taylor's Theorem for Shift Operators" and by Wei and Norman.[15] The latter authors have shown that this algebra occurs in two master-equation problems: (a) a system of simple harmonic oscillators with Landau–Teller transition probabilities; (b) the deuterium exchange reaction with only nearest-neighbor transition probabilities. Other realizations of this algebra are known in quantum mechanics. The set $\{q/\mu, e^{-i\mu p/\hbar}\}$ satisfies Eq. (6.5), as may be seen from Eq. (2.10). Some realizations which involve annihilation and creation operators are $\{-a^\dagger a, a\}$, $\{a^\dagger a, a^\dagger\}$, $\{-\gamma aq, a^2\}$, $\{\gamma a^\dagger q, (a^\dagger)^2\}$, $\{\gamma aq, q^2\}$, and $\{-\gamma a^\dagger q, q^2\}$, where $a = \gamma(q + i\omega^{-1}p)$, $a^\dagger = \gamma(q - i\omega^{-1}p)$, $\gamma = (\omega/2\hbar)^{\frac{1}{2}}$, and $[a, a^\dagger] = 1$. This algebra also occurs as a subalgebra of larger Lie algebras. For example, in the algebra SU_3 there are 12 distinct pairs of elements which satisfy Eq. (6.5).[46]

To illustrate the parameterization method alluded to above, let

$$Z = \alpha X + \beta Y, \tag{6.6}$$

$$G = aX + bY, \tag{6.7}$$

$$G(x) = e^{xZ}Ge^{-xZ} \tag{6.8}$$

$$= a(x)X + b(x)Y. \tag{6.9}$$

Then $G(x)$ satisfies the differential equation

$$G'(x) = [Z, G(x)], \tag{6.10}$$

subject to the initial condition

$$G(0) = G. \tag{6.11}$$

Substituting Eqs. (6.6) and (6.9) into Eq. (6.10) and using Eq. (6.5), one finds that

$$a'(x)X + b'(x)Y = [\alpha b(x) - \beta a(x)]Y. \tag{6.12}$$

Since X and Y are independent operators, as is easily proved from Eq. (6.5), we must have

$$a'(x) = 0, \quad b'(x) = \alpha b(x) - \beta a(x) \tag{6.13}$$

subject to the initial condition

$$a(0) = a, \quad b(0) = b. \tag{6.14}$$

The solution of Eqs. (6.13) and (6.14) is easily found to be given by

$$a(x) = a, \quad b(x) = \beta a/\alpha + (b - \beta a/\alpha)e^{\alpha x}. \tag{6.15}$$

From Eqs. (6.6)–(6.9) and (6.15), we then find that

$$e^{\alpha X+\beta Y}(aX + bY)e^{-\alpha X-\beta Y} = aX + [\beta a\alpha^{-1} + (b - \beta a\alpha^{-1})e^{\alpha}]Y. \tag{6.16}$$

Note that it could have been foreseen from Eqs. (2.2) and (6.5) that $a(x) = a$. In future calculations, we make such possible simplifications at the outset without comment. From Eqs. (6.4) and (6.16), it follows that

$$e^{\alpha X+\beta Y}F(X, Y)e^{-\alpha X-\beta Y} = F[X + \beta\alpha^{-1}(1 - e^{\alpha})Y, e^{\alpha}Y]. \tag{6.17}$$

Special cases of Eq. (6.17) are[47]

$$e^{\alpha X}F(X, Y)e^{-\alpha X} = F(X, e^{\alpha}Y), \tag{6.18}$$

$$e^{\beta Y}F(X, Y)e^{-\beta Y} = F(X - \beta Y, Y). \tag{6.19}$$

A frequently occurring three-dimensional Lie algebra is spanned by the operators $\{P, Q, I\}$ with the commutation relations

$$[P, Q] = cI, \quad [P, I] = [Q, I] = 0. \tag{6.20}$$

It is exemplified by the coordinate-momentum or annihilation-creation operator commutation rules. Since all results for this algebra are well known, we do not explicitly derive them here. However, all such results may be obtained as a special case of the four-dimensional Lie algebra which we consider next.

A Lie algebra of interest for harmonic oscillator problems is spanned by the operators $\{P, Q, W \equiv P^2 + Q^2, cI\}$ with the commutation relations defined by Eq. (6.20) and by[15]

$$[W, P] = -2cQ, \quad [W, Q] = 2cP, \quad [W, I] = 0. \tag{6.21}$$

A physical realization of this algebra is provided by the set $\{p, \omega q, p^2 + \omega^2 q^2, -i\hbar\omega I\}$. An alternative and more convenient set to use is $\{W, X, Y, sI\}$, where

$$X = Q - iP, \quad Y = Q + iP, \quad W = XY - \tfrac{1}{2}sI, \quad s = 2ic. \tag{6.22}$$

Equations (6.21) and (6.22) imply the commutation relations

$$[W, X] = -sX, \quad [W, Y] = sY, \quad [X, Y] = sI. \tag{6.23}$$

[47] Examples of Eqs. (6.18) and (6.19) occur in Louisell, Ref. 3, Theorems 9 and 6, respectively.

A physical realization of this algebra is provided by the set $\{a^\dagger a, a, a^\dagger, I\}$. Let

$$Z = \gamma W + \delta X + \rho Y, \tag{6.24}$$

$$G = gW + dX + rY, \tag{6.25}$$

and let $G(x)$ be defined by Eq. (6.8). Then Eqs. (6.10) and (6.11) are again satisfied. Letting

$$G(x) = gW + d(x)X + r(x)Y + u(x)I,$$

one finds the differential equations

$$d'(x) = -s\gamma\, d(x) + s\delta g,$$
$$r'(x) = s\gamma r(x) - s\rho g,$$
$$u'(x) = s\delta r(x) - s\rho\, d(x),$$

subject to $d(0) = d$, $r(0) = r$, and $u(0) = 0$. The solution is easily found to be given by

$$d(x) = g\delta\gamma^{-1} + (d - g\delta\gamma^{-1})e^{-s\gamma x},$$
$$r(x) = g\rho\gamma^{-1} + (r - g\rho\gamma^{-1})e^{s\gamma x},$$
$$u(x) = \delta\gamma^{-1}(r - g\rho\gamma^{-1})(e^{s\gamma x} - 1) + \rho\gamma^{-1}(d - g\delta\gamma^{-1})(e^{-s\gamma x} - 1). \tag{6.26}$$

From Eqs. (6.4), (6.8), (6.24), (6.25), and (6.26), one finds

$$e^Z F(X, Y)e^{-Z} = F(U, V), \tag{6.27}$$

where

$$U = e^{-s\gamma}X + \rho\gamma^{-1}(e^{-s\gamma} - 1)I, \tag{6.28a}$$

$$V = e^{s\gamma}Y + \delta\gamma^{-1}(e^{s\gamma} - 1)I. \tag{6.28b}$$

Expressed in terms of Q and P by Eqs. (6.22), one finds, with $Z = \alpha P + \beta Q + \gamma W$,

$$e^Z F(P, Q)e^{-Z} = F(R, S), \tag{6.29}$$

where

$$R = P\cos 2c\gamma + Q\sin 2c\gamma - \tfrac{1}{2}I\gamma^{-1}[\alpha(1 - \cos 2c\gamma) - \beta\sin 2c\gamma], \tag{6.30a}$$

$$S = -P\sin 2c\gamma + Q\cos 2c\gamma - \tfrac{1}{2}I\gamma^{-1}[\alpha\sin 2c\gamma + \beta(1 - \cos 2c\gamma)]. \tag{6.30b}$$

Another well-known Lie algebra is that of the angular momentum operators (or rotation generators) $\{J_x, J_y, J_z\}$, where

$$[J_x, J_y] = iJ_z, \quad [J_y, J_z] = iJ_x, \quad [J_z, J_y] = iJ_x. \tag{6.31}$$

(We use units in which $\hbar = 1$.) Let

$$\mathbf{b}(x)\cdot\mathbf{J} \equiv b_x(x)J_x + b_y(x)J_y + b_z(x)J_z = e^{-ix\mathbf{a}\cdot\mathbf{J}}(\mathbf{b}\cdot\mathbf{J})e^{ix\mathbf{a}\cdot\mathbf{J}}, \tag{6.32}$$

where a and b are constant vectors. Then

$$\mathbf{b}'(x)\cdot\mathbf{J} = -i[\mathbf{a}\cdot\mathbf{J}, \mathbf{b}(x)\cdot\mathbf{J}] = [\mathbf{a}\times\mathbf{b}(x)]\cdot\mathbf{J}.$$

Hence $\mathbf{b}'(x) = \mathbf{a}\times\mathbf{b}(x)$ subject to $\mathbf{b}(0) = \mathbf{b}$. The solution is easily found to be given by

$$\mathbf{b}(x) = (\hat{a}\cdot\mathbf{b})\hat{a}(1 - \cos ax) + \mathbf{b}\cos ax + \hat{a}\times\mathbf{b}\sin ax, \tag{6.33}$$

where $a \equiv |\mathbf{a}|$ and $\hat{a} = \mathbf{a}/a$. Hence

$$e^{-i\mathbf{a}\cdot\mathbf{J}}F(J_x, J_y, J_z)e^{-i\mathbf{a}\cdot\mathbf{J}} = F(K_x, K_y, K_z), \tag{6.34}$$

where

$$\mathbf{K} = \hat{a}(\hat{a}\cdot\mathbf{J})(1 - \cos a) + \mathbf{J}\cos a + \mathbf{J}\times\hat{a}\sin a. \tag{6.35}$$

Equations (6.34) and (6.35) specify how a general operator function of J must transform under an arbitrary rotation of the coordinate axes characterized by the vector $\mathbf{a}$. An interesting special case is where $a_x = a$ and $a_y = a_z = 0$. Then Eqs. (6.34) and (6.35) reduce to[48]

$$\exp(-iaJ_x)F(J_x, J_y, J_z)\exp(iaJ_x) = F(J_x, J_y\cos a + J_z\sin a, J_z\cos a - J_y\sin a). \tag{6.36}$$

Another three-dimensional Lie algebra of wide interest is the "split three-dimensional simple algebra" characterized by[15]

$$[E, F] = H, \quad [E, H] = 2E, \quad [F, H] = -2F. \tag{6.37}$$

A physical realization of the set $\{E, F, H\}$ is given by $\{iJ_-, iJ_+, 2J_z\}$, where J_+ and J_- are the angular momentum raising and lowering operators, respectively, $J_+ \equiv J_x + iJ_y$, $J_- \equiv J_x - iJ_y$. These operators occur, not only for ordinary spin, but also for isotopic spin and for quasi-spin in many-fermion systems.[49] Another guise in which these operators occur is where

$$J_- = c^\dagger b, \quad J_+ = b^\dagger c, \quad J_z = \tfrac{1}{2}(b^\dagger b - c^\dagger c), \tag{6.38}$$

and $b^\dagger$, $c^\dagger$, b, and c are the creation and annihilation operators of a two-dimensional harmonic oscillator,

$$[b, b^\dagger] = [c, c^\dagger] = 1, \quad [b, c] = [b, c^\dagger] = 0. \tag{6.39}$$

Another physical realization of the set $\{E, F, H\}$ is given by $\{P^2/2c, Q^2/2c, (QP + PQ)/2c\}$, where $[P, Q] = cI$.[15] Let

$$Z = \alpha E + \beta F + \gamma H, \quad G = aE + bF + gH, \tag{6.40}$$

and let $G(x)$ once more be defined by Eq. (8). Letting

$$G(x) = a(x)E + b(x)F + g(x)H, \tag{6.41}$$

[48] Special cases of Eq. (6.36) are as follows: A. Messiah, *Quantum Mechanics* (John Wiley & Sons, Inc., New York, 1962), Vol. II, p. 578, Ex. 8; C. P. Slichter, *Principles of Magnetic Resonance* (Harper and Row, Inc., New York, 1963), p. 26, Eq. (13).

[49] See, e.g., Schwinger, Ref. 17; Heffner and Louisell, Ref. 16; Messiah, Ref. 48; Lipkin, Ref. 46.

one finds the coupled equations

$$\begin{aligned} a'(x) &= -2\gamma a(x) + 2\alpha g(x), \\ b'(x) &= 2\gamma b(x) - 2\beta g(x), \\ g'(x) &= -\beta a(x) + \alpha b(x). \end{aligned} \tag{6.42}$$

The solution of Eqs. (6.42) subject to $a(0) = a$, $b(0) = b$, $g(0) = g$, is found to be given by

$$\begin{aligned} a(x) &= (a\beta + b\alpha - 2g\gamma)\alpha\rho^{-2}\sinh^2\rho x \\ &\quad + a\cosh 2\rho x + (g\alpha - a\gamma)\rho^{-1}\sinh 2\rho x, \\ b(x) &= (a\beta + b\alpha - 2g\gamma)\beta\rho^{-2}\sinh^2\rho x \\ &\quad + b\cosh 2\rho x + (b\gamma - g\beta)\rho^{-1}\sinh 2\rho x, \\ g(x) &= (a\beta\gamma + b\alpha\gamma - 2g\alpha\beta)\rho^{-2}\sinh^2\rho x \\ &\quad + g + \tfrac{1}{2}(b\alpha - a\beta)\rho^{-1}\sinh 2\rho x, \end{aligned}$$

where

$$\rho \equiv [\gamma^2 - \alpha\beta]^{\frac{1}{2}}.$$

Thus, with Z given by Eq. (40), one finds

$$e^{Z}f(E, F, H)e^{-Z} = f(J, K, L), \tag{6.43}$$

where J, K, and L are defined as

$$\begin{aligned} J &\equiv u(\alpha, \beta, \gamma)E + \beta^2 wF + v(\alpha, \beta, \gamma)H, \\ K &\equiv \alpha^2 wE + u(\alpha, \beta, -\gamma)F - v(\beta, \alpha, -\gamma)H, \\ L &\equiv -2v(\beta, \alpha, \gamma)E + 2v(\alpha, \beta, -\gamma)F \\ &\quad + (1 - 2\alpha\beta w)H, \\ w &\equiv \rho^{-2}\sinh^2\rho, \end{aligned} \tag{6.44}$$

$$\begin{aligned} u(\alpha, \beta, \gamma) &\equiv \cosh 2\rho + \alpha\beta w - \gamma\rho^{-1}\sinh 2\rho, \\ v(\alpha, \beta, \gamma) &\equiv \beta\gamma w - \tfrac{1}{2}\beta\rho^{-1}\sinh 2\rho. \end{aligned} \tag{6.45}$$

In case $\gamma = 0$, w, u, and v simplify to

$$\begin{aligned} w &= (\alpha\beta)^{-1}\sin^2(\alpha\beta)^{\frac{1}{2}}, \\ u(\alpha, \beta, 0) &= \cos^2(\alpha\beta)^{\frac{1}{2}}, \\ v(\alpha, \beta, 0) &= -\tfrac{1}{2}(\beta/\alpha)^{\frac{1}{2}}\sin 2(\alpha\beta)^{\frac{1}{2}}. \end{aligned} \tag{6.46}$$

Some other special cases,

$$\begin{aligned} e^{\alpha E}f(E, F, H)e^{-\alpha E} &= f(E, F + \alpha H + \alpha^2 E, H + 2\alpha E), \\ e^{\beta F}f(E, F, H)e^{-\beta F} &= f(E - \beta H + \beta^2 F, F, H - 2\beta F), \\ e^{\gamma H}f(E, F, H)e^{-\gamma H} &= f(e^{-2\gamma}E, e^{2\gamma}F, H), \end{aligned} \tag{6.47}$$

may also be easily obtained directly from the commutator expansion, Eq. (6.22).

Consider now the six-dimensional Lie algebra whose elements are I, P, Q, P^2, QP, and Q^2, with $[P, Q] = cI$. The most general second-degree polynomial in P and Q is a linear combination of these elements. Let

$$Z = \alpha P^2 + \beta Q^2 + \gamma QP + \delta P + \epsilon Q, \tag{6.48}$$

and let

$$e^{xZ}Pe^{-xZ} = d(x)P + e(x)Q + f(x)I. \tag{6.49}$$

Then $d \equiv d(x)$, $e \equiv e(x)$, and $f \equiv f(x)$ satisfy

$$\begin{aligned} d' &= 2c\alpha e - c\gamma d, \quad e' = c\gamma e - 2c\beta d, \\ f' &= c\delta e - c\epsilon d, \end{aligned}$$

subject to $d(0) = 1$, $e(0) = f(0) = 0$. The solution is given by

$$\begin{aligned} d(x) &= \cosh\lambda x - \gamma c\lambda^{-1}\sinh\lambda x, \\ e(x) &= -2\beta c\lambda^{-1}\sinh\lambda x, \\ f(x) &= -\epsilon c\lambda^{-1}\sinh\lambda x \\ &\quad + (c/\lambda)^2(\epsilon\gamma - 2\beta\delta)(\cosh\lambda x - 1), \end{aligned} \tag{6.50}$$

where

$$\lambda \equiv c[\gamma^2 - 4\alpha\beta]^{\frac{1}{2}}. \tag{6.51}$$

The expression for $e^{xZ}Qe^{-xZ}$ may be obtained from Eqs. (6.49) and (6.50) by making the following substitutions:

$$P \rightleftarrows Q, \quad C \to -c, \quad \alpha \rightleftarrows \beta, \quad \delta \rightleftarrows \epsilon.$$

We conclude this section by considering an infinite-dimensional Lie algebra. This set, which occurs in the work of Kermack and McCrea,[19] consists of P and all functions of Q, where $[P, Q] = cI$. Clearly this set satisfies the closure condition since $[P, \varphi(Q)] = c\varphi'(Q)$ is in the set. Let $\varphi(Q)$ and $f(Q)$ be arbitrary functions, and let

$$\begin{aligned} &\big(\exp\{x[\alpha P + \varphi(Q)]\}\big)[\beta P + f(Q)] \\ &\times \big(\exp\{-x[\alpha P + \varphi(Q)]\}\big) \equiv \beta P + F(Q, x), \end{aligned} \tag{6.52}$$

where α, β, and x are parameters. Then $F(Q, x)$ must satisfy the differential equation

$$\begin{aligned} \partial F(Q, x)/\partial x &= [\alpha P + \varphi(Q), \beta P + F(Q, x)] \\ &= \alpha c\partial F(Q, x)/\partial Q - \beta c\varphi'(Q), \end{aligned} \tag{6.53}$$

subject to the condition that $F(Q, 0) = f(Q)$. The solution is easily found to be given by

$$F(Q, x) = f(Q + \alpha cx) - \beta\alpha^{-1}[\varphi(Q + \alpha cx) - \varphi(Q)]. \tag{6.54}$$

Three special cases of interest are obtained by letting $f(Q)$, β, and α, respectively, go to zero:

$$\begin{aligned} &\big(\exp\{x[P + \varphi(Q)]\}\big)P\big(\exp\{-x[P + \varphi(Q)]\}\big) \\ &\quad = P - \varphi(Q + cx) + \varphi(Q), \end{aligned} \tag{6.55}$$

$$\begin{aligned} &\big(\exp\{x[P + \varphi(Q)]\}\big)f(Q)\big(\exp\{-x[P + \varphi(Q)]\}\big) \\ &\quad = f(Q + cx), \end{aligned} \tag{6.56}$$

$$e^{\varphi(Q)}Pe^{-\varphi(Q)} = P - c\varphi'(Q). \tag{6.57}$$

7. PRODUCT OF LIE EXPONENTIALS

Let A and B be LCLE's and let $e^{A}e^{B} = e^{Z}$. Then from the BCH formula, Eq. (4.14), and the "closure

property," Eq. (6.1), it follows that Z is also a LCLE. In this section, we determine the form of Z, given A and B, for some of the Lie algebras introduced in the previous section. We parametrize the problem by setting

$$e^Z \equiv e^{Z(\lambda)} = e^{\lambda A} e^B, \tag{7.1}$$

subject to $Z(0) = B$. Then differentiating Eq. (7.1) with respect to λ by means of Eq. (4.1) and multiplying from the right by $e^{-Z} = e^{-B} e^{-\lambda A}$, one obtains

$$\int_0^1 dx e^{xZ} Z'(\lambda) e^{-xZ} = A. \tag{7.2}$$

We again consider first the non-Abelian two-dimensional Lie algebra of Eq. (6.5), $[X, Y] = Y$. Let $A = \alpha_1 X + \beta_1 Y$, $B = \alpha_2 X + \beta_2 Y$, and $Z(\lambda) = \alpha X + \beta(\lambda) Y$, where $\alpha = \alpha_2 + \lambda\alpha_1$ and $\beta(0) = \beta_2$. From Eq. (6.16), the quantity $e^{xZ} Z'(\lambda) e^{-xZ}$ is seen to be given by

$$\alpha_1 X + [\beta\alpha_1\alpha^{-1} + (\beta' - \beta\alpha_1\alpha^{-1}) e^{\alpha x}] Y.$$

Then integrating over x in Eq. (7.2) and equating coefficients of Y, one finds that $\beta(\lambda)$ satisfies

$$\alpha_1\beta\alpha^{-1} + (\beta' - \alpha_1\beta\alpha^{-1})\alpha^{-1}(e^\alpha - 1) = \beta_1. \tag{7.3}$$

Defining $u \equiv u(\lambda) = \beta\alpha^{-1}$, Eq. (7.3) becomes

$$(\beta_1 - \alpha_1 u)^{-1} u' = (e^\alpha - 1)^{-1} = e^{-\alpha}(1 - e^{-\alpha})^{-1}, \tag{7.4}$$

which may also be written

$$-\partial[\ln(\alpha_1 u - \beta_1)]/\partial\lambda = \partial[\ln(1 - e^{-\alpha})]/\partial\lambda. \tag{7.5}$$

Integrating Eq. (7.5) and solving for $\beta(\lambda)$, one obtains

$$\beta(\lambda) = \lambda\beta_1\varphi(\lambda\alpha_1, \alpha_2) + \beta_2\varphi(-\alpha_2, -\lambda\alpha_1), \tag{7.6}$$

where

$$\varphi(x, y) \equiv (1 + x^{-1}y)(e^x - 1)(e^{x+y} - 1)^{-1}. \tag{7.7}$$

Hence

$$[\exp(\alpha_1 X + \beta_1 Y)][\exp(\alpha_2 X + \beta_2 Y)] = \exp[(\alpha_1 + \alpha_2) X + \beta(1) Y], \tag{7.8}$$

where $\beta(1)$ is obtained by setting $\lambda = 1$ in Eq. (7.6).

Some special cases of Eq. (7.8) have been derived by Sack by means of his "Taylor's Theorem for Shift Operators" [11]:

$$\exp[\alpha(X + \lambda Y)] = e^{\alpha X} \exp[\lambda(1 - e^{-\alpha}) Y] = \{\exp[\lambda(e^\alpha - 1) Y]\} e^{\alpha X}. \tag{7.9}$$

Sack has applied Eqs. (7.9) to the last four realizations of this algebra listed above Eq. (6.6) in order to obtain a formula for the matrix elements of a Gaussian potential.[50]

[50] Sack, Ref. 11. Matrix elements for generalized Gaussian potentials and other potentials which may be represented as Fourier integrals are obtained in R. M. Wilcox, J. Chem. Phys. **45**, 3312 (1966).

Next, consider the four-dimensional Lie algebra defined by Eqs. (6.23). Let $A = \gamma_1 W + \delta_1 X + \rho_1 Y$, $B = \gamma_2 W + \delta_2 X + \rho_2 Y$, and $Z(\lambda) = \gamma W + \delta(\lambda) X + \rho(\lambda) Y + \sigma(\lambda) I$, where $\gamma = \gamma_2 + \lambda\gamma_1$, $\delta(0) = \delta_2$, $\rho(0) = \rho_2$, and $\sigma(0) = 0$. Then from Eqs. (7.2) and (6.26) one finds, after integrating over x and equating coefficients of X, Y, and I, respectively,

$$\gamma_1\delta\gamma^{-1} + (\delta' - \gamma_1\delta\gamma^{-1})(-\gamma s)^{-1}(e^{-\gamma s} - 1) = \delta_1, \tag{7.10a}$$

$$\gamma_1\rho\gamma^{-1} + (\rho' - \gamma_1\rho\gamma^{-1})(\gamma s)^{-1}(e^{\gamma s} - 1) = \rho_1, \tag{7.10b}$$

$$\sigma' + \delta\gamma^{-1}(\rho' - \gamma_1\rho\gamma^{-1})[(\gamma s)^{-1}(e^{\gamma s} - 1) - 1] + \rho\gamma^{-1}(\delta' - \gamma_1\delta\gamma^{-1})[(-\gamma s)^{-1}(e^{-\gamma s} - 1) - 1] = 0. \tag{7.10c}$$

Equations (7.10a) and (7.10b) are of the same form as Eq. (7.3), so that their solutions may be obtained from Eq. (7.6) by appropriate changes of variables:

$$\delta(\lambda) = \lambda\delta_1\varphi(-\lambda\gamma_1 s, -\gamma_2 s) + \delta_2\varphi(\gamma_2 s, \lambda\gamma_1 s), \tag{7.11}$$

$$\rho(\lambda) = \lambda\rho_1\varphi(\lambda\gamma_1 s, \gamma_2 s) + \rho_2\varphi(-\gamma_2 s, -\lambda\gamma_1 s). \tag{7.12}$$

From Eqs. (7.10a), (7.10b), and (7.10c), one finds that

$$\sigma'(\lambda) = \gamma^{-1}[\partial(\rho\delta)/\partial\lambda - \rho_1\delta - \delta_1\rho]. \tag{7.13}$$

Equation (7.13) is integrated with the aid of Eqs. (7.11) and (7.12) by putting the terms in a form similar to the right side of Eq. (7.5). Setting

$$\tau_1 = \rho_1/\gamma_1, \quad \tau_2 = \rho_2/\gamma_2, \quad \mu_1 = \delta_1/\gamma_1, \quad \mu_2 = \delta_2/\gamma_2, \tag{7.14}$$

one obtains finally

$$\sigma(1) = (\gamma_1 + \gamma_2)^{-1}\rho(1)\delta(1) - \gamma_1\tau_1\mu_1 - \gamma_2\tau_2\mu_2 + \theta, \tag{7.15}$$

where

$$\theta \equiv 2s^{-1}(\tau_1 - \tau_2)(\mu_1 - \mu_2)\sinh(\tfrac{1}{2}s\gamma_1) \times \sinh(\tfrac{1}{2}s\gamma_2)\operatorname{csch}[\tfrac{1}{2}s(\gamma_1 + \gamma_2)]. \tag{7.16}$$

This result may be stated in terms of the X and Y operators as follows: Let $[X, Y] = sI$. Then the equation

$$\exp[\gamma_1(X + \tau_1 I)(Y + \mu_1 I)] \exp[\gamma_2(X + \tau_2 I)(Y + \mu_2 I)] = \exp[(\gamma_1 + \gamma_2)(X + \tau I)(Y + \mu I) + \theta I] \tag{7.17}$$

is identically satisfied provided θ is defined by Eq. (7.16), and provided τ and μ are defined by

$$\tau \equiv \tau_1\psi(\gamma_1, \gamma_2) + \tau_2\psi(-\gamma_2, -\gamma_1), \tag{7.18}$$

$$\mu \equiv \mu_1\psi(-\gamma_1, -\gamma_2) + \mu_2\psi(\gamma_2, \gamma_1), \tag{7.19}$$

where $\psi(x, y)$ is defined by

$$\psi(x, y) \equiv [e^{sx} - 1][e^{s(x+y)} - 1]^{-1}. \tag{7.20}$$

The result may alternatively be stated in terms of the P and Q operators: Let $[P, Q] = cI$. Then the equation

$$\exp\{\gamma_1[(P+\nu_1 I)^2 + (Q+\omega_1 I)^2]\} \times \exp\{\gamma_2[(P+\nu_2 I)^2 + (Q+\omega_2 I)^2]\} = \exp\{(\gamma_1+\gamma_2)[(P+\nu I)^2 + (Q+\omega I)^2] + \theta I\} \tag{7.21}$$

is identically satisfied provided

$$\theta \equiv c^{-1}[(\nu_1+\nu_2)^2 + (\omega_1-\omega_2)^2]\lambda(\gamma_1,\gamma_2), \tag{7.22}$$

$$\nu \equiv \nu_1\chi(\gamma_1,\gamma_2) + \nu_2\chi(\gamma_2,\gamma_1) + (\omega_2-\omega_1)\lambda(\gamma_1,\gamma_2), \tag{7.23}$$

$$\omega \equiv \omega_1\chi(\gamma_1,\gamma_2) + \omega_2\chi(\gamma_2,\gamma_1) + (\nu_1-\nu_2)\lambda(\gamma_1,\gamma_2), \tag{7.24}$$

where

$$\chi(x,y) \equiv \sin(cx)\cos(cy)\csc[c(x+y)], \tag{7.25}$$

$$\lambda(x,y) \equiv \sin(cx)\sin(cy)\csc[c(x+y)]. \tag{7.26}$$

Some special cases of Eqs. (7.17) and (7.21) are of interest:

$$\ln[\exp(\alpha X+\beta Y)\exp(\gamma XY)] = \gamma[X+\beta s(e^{s\gamma}-1)^{-1}][Y+\alpha s(1-e^{-s\gamma})^{-1}] - \tfrac{1}{2}\alpha\beta s\coth(\tfrac{1}{2}s\gamma), \tag{7.27}$$

$$\ln[\exp\gamma(Q^2+P^2)\exp(\alpha P+\beta Q)] = \gamma\{P+\tfrac{1}{2}c[\alpha\cot(\gamma c)+\beta]\}^2 + \gamma\{Q+\tfrac{1}{2}c[\beta\cot(\gamma c)-\alpha]\}^2 - \tfrac{1}{4}(\alpha^2+\beta^2)c\cot(\gamma c). \tag{7.28}$$

Equation (7.28) with $\gamma = 1$ has been obtained by Weiss and Maradudin,[6] who derive the result directly from the BCH formula by a rather intricate summation procedure.

As an easy application of Eq. (7.28), we state a corollary to Bloch's theorem concerning the characteristic function of a harmonic oscillator in thermal equilibrium.[12]

Theorem: Let the thermal average be defined by Eq. (3.1) with $\mathcal{H} = \frac{1}{2}p^2 + \frac{1}{2}\omega^2 q^2$ and $[q, p] = i\hbar$. Then

$$\langle e^{i(\sigma p+\tau q)}\rangle = e^{-\frac{1}{2}\tau^2\langle q^2\rangle}e^{-\frac{1}{2}\sigma^2\langle p^2\rangle}, \tag{7.29}$$

where

$$\langle p^2\rangle = \omega^2\langle q^2\rangle = \tfrac{1}{2}\hbar\omega\coth(\tfrac{1}{2}\beta\hbar\omega).$$

By means of Eq. (5.2), this result may be used to obtain thermal averages for general observables represented in the Weyl manner.[51] Related corollaries to Bloch's theorem have occurred in connection with the scattering of x rays or neutrons by molecules or harmonic lattices,[52] as well as in treatments of the coherent states of the radiation field.[53]

We conclude this section by treating a special case of the infinite-dimensional Lie algebra whose elements are P and all functions of Q, where $[P, Q] = cI$. We seek to find a function $Z \equiv Z(Q, \lambda)$ which satisfies

$$e^Z = e^{\lambda A}e^{-\lambda P}, \tag{7.30}$$

where $A \equiv P + \varphi(Q)$ and $\varphi(Q)$ is an arbitrary given function. Then by differentiating Eq. (7.30) with respect to λ, multiplying both sides from the right by $e^{-Z} = e^{\lambda P}e^{-\lambda A}$, and applying Eq. (6.55), one finds, after some cancellations, that

$$\partial Z/\partial\lambda = \varphi(Q + c\lambda). \tag{7.31}$$

Since Z must vanish when λ does,

$$Z(Q,\lambda) = \int_0^\lambda dx\,\varphi(Q+cx). \tag{7.32}$$

From Eqs. (7.32) and (7.30), one obtains

$$\exp\{\lambda[P+\varphi(Q)]\} = \left[\exp\int_0^\lambda dx\,\varphi(Q+cx)\right]e^{\lambda P}. \tag{7.33}$$

This result is used in Sec. 10 to obtain a normal-ordering expansion of Kermack and McCrea, Eqs. (10.42).

8. SOLUTIONS OF $dY(t)/dt = A(t)Y(t)$

The operator differential-equation system

$$dY(t)/dt = A(t)Y(t), \quad Y(0) = I \tag{8.1}$$

has been extensively studied by mathematicians because of its relevance to the theory of coupled or higher-order ordinary differential equations. Some examples of this equation which occur in quantum physics are the equation of motion for the time-evolution operator, Eq. (9.3), the Bloch equation, $-\partial e^{-\beta\mathcal{H}}/\partial\beta = \mathcal{H}e^{-\beta}$, and master or rate equations.

8.1. Expansions of Magnus and Fer

Instead of dealing with Eq. (8.1) directly, we introduce the iteration parameter λ as

$$dY_\lambda(t)/dt = \lambda A(t)Y_\lambda(t), \quad Y_\lambda(0) = I, \tag{8.1$_\lambda$}$$

and seek to join $Y_0(t) = I$ to $Y_1(t) \equiv Y(t)$. To derive

[51] Equation (29) may also be derived easily from the condition that both sides satisfy the same first-order differential equation with respect to a parameter. Still another derivation of Eq. (29) is obtained by first calculating the matrix elements of $\exp(i\sigma p + i\tau q)$ in the harmonic oscillator representation (as one may do by employing the method which the author used in Ref. 50 for the case where $\sigma = 0$) and then carrying out the trace sum with the aid of the generating function for the Laguerre polynomials.

[52] A. C. Zemach and R. J. Glauber, Phys. Rev. **101**, 118 (1956); Weiss and Maradudin, Ref. 6; A. A. Maradudin, E. W. Montroll, and G. H. Weiss, Solid State Phys. Suppl. 3, 239 (1963); N. D. Mermin, J. Math. Phys. 7, 1038 (1966).

[53] R. J. Glauber, Phys. Rev. **131**, 2766 (1963); Louisell, Ref. 3, p. 244.

the Magnus formula, we assume a solution of the form

$$Y_\lambda(t) = \exp[\Omega(\lambda, t)] \equiv e^{\Omega}, \tag{8.2}$$

where

$$\Omega(\lambda, t) \equiv \sum_{n=1}^{\infty} \lambda^n \Delta_n(t). \tag{8.3}$$

From Eqs. (8.1$_\lambda$), (8.2), and (4.1), it follows that

$$\int_0^1 dx e^{x\Omega} \dot{\Omega} e^{-x\Omega} = \lambda A(t), \tag{8.4}$$

where $\dot{\Omega} \equiv \partial\Omega(\lambda, t)/\partial t$. Using the commutator expansion and integrating over x, as in Eq. (4.7), and substituting in Eq. (8.3), one obtains

$$\left\{\sum_{k=0}^{\infty} \frac{1}{(k+1)!}\left(\sum_{n=1}^{\infty} \lambda^n \Delta_n(t)\right)^k, \sum_{m=1}^{\infty} \lambda^m \dot{\Delta}_m(t)\right\} = \lambda A(t). \tag{8.5}$$

A recursive procedure again results from equating coefficients of λ^j on the two sides of Eq. (8.5). For $j = 1$, one obtains $\dot{\Delta}_1(t) = A(t)$. Hence,

$$\Delta_1(t) = \int_0^t A(\tau)\, d\tau. \tag{8.6}$$

For $j = 2$ one obtains

$$\dot{\Delta}_2(t) + \tfrac{1}{2}[\Delta_1(t), \dot{\Delta}_1(t)] = 0.$$

Hence,

$$\Delta_2(t) = \frac{1}{2}\int_0^t d\sigma \int_0^\sigma d\tau [A(\sigma), A(\tau)]. \tag{8.7}$$

For $j = 3$ one obtains

$$\dot{\Delta}_3 + \tfrac{1}{2}[\Delta_1, \dot{\Delta}_2] + \tfrac{1}{2}[\Delta_2, \dot{\Delta}_1] + \tfrac{1}{6}[\Delta_1, [\Delta_1, \dot{\Delta}_1]] = 0.$$

After carrying out the integration and putting the results in "time-ordered" form, one obtains

$$\Delta_3(t) = \frac{1}{6}\int_0^t dt_1 \int_0^{t_1} dt_2 \int_0^{t_2} dt_3 \times \{[[A_1, A_2], A_3] + [[A_3, A_2], A_1]\}, \tag{8.8}$$

where $A_1 \equiv A(t_1)$, etc. The fourth-order term is similarly calculated. We find

$$\Delta_4(t) = \frac{1}{12}\int_0^t dt_1 \int_0^{t_1} dt_2 \int_0^{t_2} dt_3 \int_0^{t_3} dt_4 \times \{[[[A_3, A_2], A_4], A_1] + [[[A_3, A_4], A_2], A_1] + [[[A_1, A_2], A_3], A_4] + [[[A_4, A_1], A_3], A_2]\}. \tag{8.9}$$

The solution of Eqs. (8.1) to fourth order is given by Eqs. (8.2), (8.3), and (8.6)–(8.9) with $\lambda = 1$.[54] This result is said to be the continuous analog of the BCH formula.

We now derive the Fer formula,[13] which may be said to be the continuous analog of the Zassenhaus formula. Let the solutions of Eqs. (8.1$_\lambda$) be given by

$$Y_\lambda(t) = e^{\lambda S_1} e^{\lambda^2 S_2} e^{\lambda^3 S_3} \cdots, \tag{8.10}$$

where $S_1 \equiv S_1(t)$, etc. Substituting Eq. (8.10) into (8.1$_\lambda$), multiplying from the right by

$$e^{-\lambda^3 S_3} e^{-\lambda^2 S_2} e^{-\lambda S_1} = [Y_\lambda(t)]^{-1},$$

and expanding in terms of repeated commutators as in Eq. (4.20), one obtains

$$\sum_{n=0}^{\infty} \frac{\lambda^{n+1}}{(n+1)!}\{S_1^n, \dot{S}_1\} + \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} \frac{\lambda^{m+2n+2}}{m!\,(n+1)!} \times \{S_1^m, S_2^n, \dot{S}_2\} + \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\sum_{k=0}^{\infty} \frac{\lambda^{m+2n+3k+3}}{m!\,n!\,(k+1)!} \times \{S_1^m, S_2^n, S_3^k, \dot{S}_3\} + \cdots = \lambda A(t), \tag{8.11}$$

where the repeated commutators are again defined by Eqs. (4.21). The recursion scheme based upon Eq. (8.11) leads to

$$S_1(t) = \int_0^t A(\tau)\, d\tau, \tag{8.12}$$

$$S_2(t) = \frac{1}{2}\int_0^t d\sigma \int_0^\sigma d\tau [A(\sigma), A(\tau)], \tag{8.13}$$

$$S_3(t) = \frac{1}{3}\int_0^t dt_1 \int_0^{t_1} dt_2 \int_0^{t_2} dt_3 \times \{[A_2, [A_3, A_1]] + [A_3, [A_2, A_1]]\}. \tag{8.14}$$

Equations (8.10), (8.12), (8.13), and (8.14), with $\lambda = 1$, give the first three factors in the infinite-product solution to Eqs. (8.1). The paper by Fer discusses the convergence of the infinite product and derives recursion relations, but does not obtain explicit expressions for the S_j's.[13]

$$A(t) \equiv e^{-at} b e^{at}, \quad Y_\lambda(t) \equiv e^{-at} U(t), \tag{8.15}$$

where a and b are independent of t. Then $U(t)$ must satisfy

$$dU(t)/dt = (a + \lambda b)U(t), \quad U(0) = I, \tag{8.16}$$

which implies that

$$U(t) = e^{(a+\lambda b)t}. \tag{8.17}$$

Kumar's expansion for $e^{(a+\lambda b)t}$ as an infinite product is given by[14]

$$e^{(a+\lambda b)t} = e^{at} Y_\lambda(t), \tag{8.18}$$

where $Y_\lambda(t)$ is given by Eqs. (8.10), (8.12), (8.13), (8.14), $\cdots$, and (8.15). An alternative expansion for $e^{(a+\lambda b)t}$ as a product of only two exponentials may

[54] Magnus, Ref. 7, and Weiss and Maradudin, Ref. 6, have carried out the calculation to third order. Their third-order terms, though not given in such a symmetrical form, may be shown to be equivalent to ours. A special case of this formula has also been calculated by Kumar, Ref. 14, to third order.

be more suitable for some purposes. It is obtained from the Magnus solution to Eqs. (8.1_λ) and is given by Eqs. (8.18), (8.2), (8.3), (8.6)–(8.9), $\cdots$, and (8.15). These expansions are more complicated than the Zassenhaus expansion, Eq. (4.18), but appear to be more useful when λb is "small" compared with a. They also appear to be more useful for large values of t, provided $U(t)$ is unitary. A fuller heuristic comparison of the two infinite-product forms for the case where $t \to \infty$ has been given by Kumar.[14]

8.2. Lie Algebraic Solutions

We now consider the case where $A(t)$ in Eq. (8.1) may be expressed as a LCLE,

$$A(t) = a_1(t)X_1 + a_2(t)X_2 + \cdots + a_n(t)X_n.$$

Then if a solution of the form

$$Y(t) = e^{\Omega(t)} \tag{8.19}$$

exists, the repeated-commutator form for Ω and the "closure property" of the Lie algebra imply that $\Omega(t)$ is also a linear combination of the X_j's. We will find some closed-form solutions for two Lie algebras of physical interest by a method similar to those employed in previous sections. Our method is also basically the same as the method employed by Wei and Norman to obtain solutions in the product form[15]

$$Y(t) = \exp[g_1(t)X_1] \cdots \exp[g_n(t)X_n] \tag{8.20}$$

if one ignores the technical group-theoretical considerations of that treatment. In practice, however, the single-exponential solutions are more difficult to obtain.

We again consider first the two-dimensional algebra where $[X, Y] = Y$. In Eqs. (8.1), let

$$A(t) \equiv a(t)X + b(t)Y, \tag{8.21}$$

where $a(t)$ and $b(t)$ are arbitrary functions of t, and let

$$\Omega(t) \equiv \alpha(t)X + \beta(t)Y \tag{8.22}$$

in Eq. (8.19). Substituting Eqs. (8.21) and (8.22) into Eq. (8.4) with $\lambda = 1$, using Eq. (6.16), integrating over x, and equating coefficients of X and Y, one obtains

$$\dot{\alpha} = a(t), \quad \dot{\alpha}u + \dot{u}(e^{\alpha} - 1) = b(t), \tag{8.23}$$

where

$$u(t) \equiv \beta(t)/\alpha(t). \tag{8.24}$$

Integrating Eqs. (8.23), one obtains

$$\alpha(t) = \int_0^t a(\tau)\, d\tau, \tag{8.25}$$

$$\beta(t) = [1 - e^{-\alpha(t)}]^{-1}\alpha(t)\gamma(t), \tag{8.26}$$

where

$$\gamma(t) = \int_0^t d\tau b(\tau)e^{-\alpha(\tau)}\, d\tau. \tag{8.27}$$

For the case where $A(t)$ is given by Eq. (8.21), Eqs. (8.19), (8.22), (8.25), (8.26), and (8.27) constitute the solution of Eq. (8.1). By means of Eq. (7.9), this result may be shown to be equivalent to the product form of Wei and Norman,[55]

$$Y(t) = e^{\alpha(t)X}e^{\gamma(t)Y}. \tag{8.28}$$

A solution may also be obtained for the four-dimensional harmonic-oscillator algebra of Eqs. (6.23). In Eqs. (8.1) and (8.19), let

$$A(t) \equiv g(t)W + d(t)X + r(t)Y + u(t)I \tag{8.29a}$$

$$\equiv g(t)[X + w(t)I][Y + v(t)I] + f(t)I, \tag{8.29b}$$

$$\Omega(t) \equiv \gamma(t)W + \delta(t)X + \rho(t)Y + \mu(t)I \tag{8.30a}$$

$$\equiv \gamma(t)[X + \omega(t)I][Y + \nu(t)I] + \varphi(t)I. \tag{8.30b}$$

Then proceeding in the same way as before, using either Eqs. (6.26) or Eq. (6.27), one obtains the differential equations

$$\dot{\gamma} = g,$$
$$g\nu + \dot{\nu}(1 - e^{-s\gamma})s^{-1} = d = gv,$$
$$g\omega + \dot{\omega}(e^{s\gamma} - 1)s^{-1} = r = gw,$$
$$\dot{\varphi} = g(w - \omega)(v - \nu) + f(t),$$

where $\gamma \equiv \gamma(t)$, $g \equiv g(t)$, etc. The solution, subject to the conditions $\gamma(0) = \delta(0) = \rho(0) = \mu(0) = 0$, is given by

$$\gamma(t) = \int_0^t g(\tau)\, d\tau, \tag{8.31}$$

$$\nu(t) = [e^{s\gamma(t)} - 1]^{-1}s\alpha(t), \tag{8.32}$$

$$\omega(t) = [1 - e^{-s\gamma(t)}]^{-1}s\beta(t), \tag{8.33}$$

$$\varphi(t) = \int_0^t d\tau\{g(\tau)[w(\tau) - \omega(\tau)][v(\tau) - \nu(\tau)] + f(\tau)\}, \tag{8.34}$$

where

$$\alpha(t) \equiv \int_0^t d(\tau)e^{s\gamma(\tau)}\, d\tau, \tag{8.35}$$

$$\beta(t) \equiv \int_0^t r(\tau)e^{-s\gamma(\tau)}\, d\tau. \tag{8.36}$$

For the case where $A(t)$ is given by Eqs. (8.29), Eqs. (8.19) and (8.30)–(8.36) constitute the solution of Eq. (8.1). This may be compared with the product form of Wei and Norman,[55]

$$Y(t) = e^{\gamma(t)W}e^{\alpha(t)X}e^{\beta(t)Y}e^{\psi(t)I}, \tag{8.37}$$

[55] Wei and Norman, Ref. 15, derive differential equations, but do not explicitly give their solution since they are mainly interested in determining whether or not solutions exist.

where $\alpha(t)$, $\beta(t)$, and $\gamma(t)$ are given by Eqs. (8.35), (8.36), and (8.31), respectively, and where

$$\psi(t) \equiv \int_0^t d\tau[u(\tau) - s\alpha(\tau)r(\tau)e^{-s\gamma(\tau)}]. \quad (8.38)$$

One may use either of these solutions to solve the problem of the driven harmonic oscillator, without the necessity of first transforming to the interaction representation as is usually done.[56]

9. SOLUTIONS OF $i\hbar\partial\rho(t)/\partial t = [\mathcal{H}(t), \rho(t)]$

Consider the equation for the time evolution of the density matrix $\rho(t)$, for a system described by a Hamiltonian $\mathcal{H}(t)$,[57]

$$i\hbar\partial\rho(t)/\partial t = [\mathcal{H}(t), \rho(t)]. \quad (9.1)$$

(Aside from a sign, the same equation of course applies to any Heisenberg operator.) In case $\mathcal{H}(t)$ is an LCLE and $\rho(t)$ is initially an LCLE, then Eqs. (9.1) and (6.1) imply that $\rho(t)$ will remain an LCLE for all time. A solution may then be readily obtained, as in Sec. 6. However, the condition that an arbitrary density matrix be expressible as an LCLE is very restrictive, so that not many cases occur in practice. A case where an arbitrary density matrix may be expressed in terms of an LCLE occurs in the problem of a spin $-\frac{1}{2}$ magnetic moment in a time-varying magnetic field.[58]

However, as is well known, even if $\rho(0)$ is not an LCLE, a solution of Eq. (9.1) is given by

$$\rho(t) = U(t)\rho(0)U^\dagger(t), \quad (9.2)$$

where $U(t)$ is a solution of

$$i\hbar\partial U(t)/\partial t = \mathcal{H}(t)U(t), \quad U(0) = I, \quad (9.3)$$

the problem discussed in Sec. 8.

A case of much interest is where the density matrix is initially in thermal equilibrium[59]:

$$\rho(0) = e^{-\beta\mathcal{H}(0)}/\mathrm{Tr}\,[e^{-\beta\mathcal{H}(0)}]. \quad (9.4)$$

It follows from the basic lemma given at the end of Sec. 2 that a solution to Eqs. (9.1) and (9.4) is given by

$$\rho(t) = e^{-\beta H(t)}/\mathrm{Tr}\,[e^{-\beta\mathcal{H}(0)}], \quad (9.5)$$

where $H(t)$ satisfies

$$i\hbar\partial H(t)/\partial t = [\mathcal{H}(t), H(t)], \quad H(0) = \mathcal{H}(0). \quad (9.6)$$

Since $H(t)$ satisfies the same differential equation as $\rho(t)$, it may appear that nothing has been gained. However, it may be easier to apply the boundary conditions to $H(0)$ than to $\rho(0)$. In particular, $H(0) \equiv \mathcal{H}(0)$ may be an LCLE although $\rho(0)$ is not. Considering the various methods available for handling exponential operators, Eqs. (9.5) and (9.6) may also be a useful starting point for dealing with practical many-body systems for which exact solutions are impossible.

Since $\mathrm{Tr}\,\rho(t) = 1$ at all times, as is implied by Eq. (9.4), Eq. (9.1), and the cyclic property of the trace, Eq. (9.5) may also be written in the form

$$\rho(t) = e^{-\beta H(t)}/\mathrm{Tr}\,[e^{-\beta H(t)}]. \quad (9.7)$$

Equation (9.7) shows that at all times the system behaves as though its temperature remains constant with a "canonical distribution" determined by the instantaneous value of the fictitious Hamiltonian $H(t)$. It may thus be possible to treat nonequilibrium situations by the methods of equilibrium statistical mechanics.

A simple illustrative example is the problem of an arbitrary spin magnetic moment in an arbitrary time-varying magnetic field with a Hamiltonian

$$\mathcal{H}(t) = \gamma\mathbf{h}(t)\cdot\mathbf{J} \equiv \gamma[h_x(t)J_x + h_y(t)J_y + h_z(t)J_z]. \quad (9.8)$$

Then $H(t)$ is of the form

$$H(t) = \gamma\mathbf{b}(t)\cdot\mathbf{J}. \quad (9.9)$$

It easily follows from Eqs. (9.6), (9.8), (9.9), and (6.31) that the fictitious field $\mathbf{b}(t)$ precesses about the instantaneous direction of $h(t)$ according to the equation

$$\dot{\mathbf{b}}(t) = \gamma\mathbf{h}(t) \times \mathbf{b}(t), \quad (9.10)$$

with $\mathbf{b}(0) \equiv \mathbf{h}(0)$.

10. NORMAL-ORDERING OF OPERATORS

Normal-ordering techniques are useful for solving operator differential equations,[60] evaluating matrix elements,[61] and finding the quantum operator corresponding to a given classical operator. We have already encountered an example of this last application in McCoy's theorem, Eq. (5.6). Using methods and formulas of previous sections, we may efficiently derive a number of related results.

The derivations are greatly facilitated with the aid of the "normal-ordering operator" $\mathcal{N}$ defined as

[56] See, e.g., Louisell, Ref. 3, p. 119, Sec. 3.5.

[57] See, e.g., Louisell, Ref. 3, Chap. 6. Other references are given there.

[58] See, e.g., F. A. Kaempffer, *Concepts in Quantum Mechanics* (Academic Press Inc., New York, 1965), Sec. 4.

[59] This situation has been treated by Kubo, Ref. 4.

[60] J. L. Anderson, Phys. Rev. **94**, 703 (1954); N. N. Bogoliubov and D. V. Shirkov, *Introduction to the Theory of Quantized Fields* (Interscience Publishers, Inc., New York, 1959), p. 486; see also Louisell, Ref. 3, and Heffner and Louisell, Ref. 16.

[61] G. Wick, Phys. Rev. **80**, 268 (1950); Bogoliubov and Shirkov, Ref. 60, Sec. 16; Louisell, Ref. 3; Wilcox, Ref. 50.

follows[62]: With $[P, Q] = cI$, let the function $f(P, Q)$ be defined by its formal power-series expansion,

$$f(P, Q) \equiv aI + bP + dQ + ePQ + gQP + \cdots. \tag{10.1}$$

Then the linear "superoperator" $\mathcal{N}$ acting on $f(P, Q)$ moves all P's to the right of the Q's *as though* P and Q commute; i.e.,

$$\mathcal{N}[f(P, Q)] \equiv aI + bP + dQ + (e + g)QP + \cdots. \tag{10.2}$$

Although other operator expressions may be obtained from $f(P, Q)$ by application of the commutation relation, we emphasize that their functional forms will be different. Thus, even though $f(P, Q) = g(P, Q)$ in the usual operator sense, it need not follow that $\mathcal{N}[f(P, Q)] = \mathcal{N}[g(P, Q)]$. Some useful properties of $\mathcal{N}$ are a direct consequence of its definition. If $f(P, Q)$ and $g(P, Q)$ are any two operator functions, then

$$\mathcal{N}[f(P, Q)g(P, Q)] = \mathcal{N}[g(P, Q)f(P, Q)]. \tag{10.3}$$

If $|P'\rangle$ and $|Q'\rangle$ are eigenstates of P and Q, $P|P'\rangle = P'|P'\rangle$ and $Q|Q'\rangle = Q'|Q'\rangle$, then

$$\{\mathcal{N}[f(P, Q)]\}|P'\rangle = f(P', Q)|P'\rangle, \tag{10.4a}$$

$$\langle Q'|\{\mathcal{N}[f(P, Q)]\} = \langle Q'|f(P, Q'^*), \tag{10.4b}$$

$$\langle Q'|\{\mathcal{N}[f(P, Q)]\}|P'\rangle = f(P', Q'^*)\langle Q'|P'\rangle. \tag{10.4c}$$

Equations (10.4) usually occur in practice for the case where P and Q correspond to the annihilation and creation operators a and $a^\dagger$, respectively, and the vacuum state $|0\rangle$ is involved, $a|0\rangle = 0$ or $\langle 0|a^\dagger = 0$. If the f in Eq. (10.1) also depends upon a parameter λ, $f \equiv f(P, Q, \lambda)$, then it is apparent that $\mathcal{N}$ commutes with $\partial/\partial\lambda$ since

$$\begin{aligned}\mathcal{N}[\partial f(P, Q, \lambda)/\partial\lambda] &= \mathcal{N}[a'I + b'P + d'Q + e'PQ + g'QP + \cdots] \\ &= a'I + b'P + d'Q + (e' + g')QP + \cdots \\ &= \partial\{\mathcal{N}[f(P, Q, \lambda)]\}/\partial\lambda,\end{aligned} \tag{10.5}$$

where $a' \equiv \partial a(\lambda)/\partial\lambda$, etc. From the definition of the derivative with respect to an operator, Eq. (2.0), it readily follows that $\mathcal{N}$ also commutes with $\partial/\partial P$ and $\partial/\partial Q$.

An important problem of normal-ordering is to express the product of two normal-ordered products in normal form.[3,16]

Theorem: Let $F(P, Q)$ and $G(P, Q)$ be in normal form: $F(P, Q) \equiv \mathcal{N}[F(P, Q)]$; $G(P, Q) \equiv \mathcal{N}[G(P, Q)]$. Then the normal form of the product is given by

$$\mathcal{N}[F(P + c\partial/\partial Q, Q)G(P, Q)]. \tag{10.6}$$

Proof: We prove the result first for the special case where $F(P, Q)$ and $G(P, Q)$ have the forms

$$F(P, Q) = e^{wQ}e^{xP}, \quad G(P, Q) = e^{vQ}e^{zP}.$$

Then

$$\begin{aligned}F(P, Q)G(P, Q) &= e^{wQ}e^{xP}e^{vQ}e^{zP} \\ &= e^{xvc}e^{wQ}e^{vQ}e^{xP}e^{zP} \\ &= \mathcal{N}[e^{xvc}e^{wQ}e^{vQ}e^{xP}e^{zP}] \\ &= \mathcal{N}[e^{wQ}e^{x(P+cv)}e^{vQ}e^{zP}] \\ &= \mathcal{N}[e^{wQ}e^{x(P+c\partial/\partial Q)}e^{vQ}e^{zP}] \\ &= \mathcal{N}[F(P + c\partial/\partial Q, Q)G(P, Q)].\end{aligned}$$

The general case easily follows by the method of linear superposition. The theorem may be readily generalized to the case of more than one pair of conjugate variables.

As an application of this theorem, consider the case where $F(P, Q) = P^m$, $G(P, Q) = Q^n$. Then

$$\begin{aligned}P^mQ^n &= \mathcal{N}[(P + c\partial/\partial Q)^mQ^n] \\ &= \mathcal{N}\sum_{j=0}^{m}\frac{m!\,P^{m-j}c^j}{j!\,(m-j)!}\frac{\partial^jQ^n}{\partial Q^j} \\ &= \sum_{j=0}^{m}\frac{m!\,n!\,c^jQ^{n-j}P^{m-j}}{j!\,(m-j)!\,(n-j)!}.\end{aligned} \tag{10.7}$$

To show how the theorem may be used for solving differential equations, and to compare with the treatment of Heffner and Louisell,[16] we consider their example of a spin magnetic moment in a rotating magnetic field. The Hamiltonian is then of the form[16]

$$\mathcal{H} = \tfrac{1}{2}\hbar\omega_0(b^\dagger b - c^\dagger c) + \tfrac{1}{2}\hbar\gamma H_1(b^\dagger ce^{-i\omega t} + c^\dagger be^{i\omega t}), \tag{10.8}$$

where the annihilation and creation operators b, $b^\dagger$, c, and $c^\dagger$ were defined in Eq. (6.39), and $\hbar$, ω_0, γ, H_1, and ω are costants. We assume that the solution of the equation

$$i\hbar\dot{U} \equiv i\hbar\partial U/\partial t = \mathcal{H}U \tag{10.9}$$

is of the form[63]

$$U = \mathcal{N}\{\exp[(A - 1)b^\dagger b + (B - 1)c^\dagger c + Db^\dagger c + Ec^\dagger b]\}, \tag{10.10}$$

where $\mathcal{N}$ orders b and c to the right of $b^\dagger$ and $c^\dagger$, and

[62] Our definition of the "normal-ordering operator" is not the same as that given by Louisell, Ref. 3, and Heffner and Louisell, Ref. 16, but we believe it is more convenient.

[63] For convenience, we define U such that A, B, D, and E correspond to the notation of Ref. 16, but this, of course, is not essential for the solution of the problem.

the scalar functions A, B, D, and E are to be determined. It clearly follows that

$$\dot{U} = \dot{A}b^\dagger Ub + \dot{B}c^\dagger Uc + \dot{D}b^\dagger Uc + \dot{E}c^\dagger Ub. \quad (10.11)$$

From the above theorem, $2\mathcal{H}U/\hbar$ is given by

$$\mathcal{N}[(\omega_0 b^\dagger + \gamma H_1 e^{i\omega t}c^\dagger)(b + \partial/\partial b^\dagger)U + (-\omega_0 c^\dagger + \gamma H_1 e^{-i\omega t}b^\dagger)(c + \partial/\partial c^\dagger)U]. \quad (10.12)$$

Since

$$\partial U/\partial b^\dagger = U[(A - 1)b + Dc] \quad (10.13)$$

and

$$\partial U/\partial c^\dagger = U[(B - 1)c + Eb], \quad (10.14)$$

it follows from Eqs. (10.9)–(10.14), by equating coefficients of $b^\dagger Ub$, $c^\dagger Uc$, $b^\dagger Uc$, and $c^\dagger Ub$, respectively, that

$$\begin{aligned} i\partial A/\partial t &= \tfrac{1}{2}\omega_0 A + \tfrac{1}{2}\gamma H_1 e^{-i\omega t}E, \\ i\partial B/\partial t &= -\tfrac{1}{2}\omega_0 B + \tfrac{1}{2}\gamma H_1 e^{i\omega t}D, \\ i\partial D/\partial t &= \tfrac{1}{2}\omega_0 D + \tfrac{1}{2}\gamma H_1 e^{-i\omega t}B, \\ i\partial E/\partial t &= -\tfrac{1}{2}\omega_0 E + \tfrac{1}{2}\gamma H_1 e^{i\omega t}A. \end{aligned} \quad (10.15)$$

We refer the reader to Ref. 16 for the solution of these equations. Note that the treatment given here avoids the necessity of transforming the annihilation–creation operator space onto a space of commuting algebraic variables, and back again.

We next derive a general result which is primarily of interest for cases where P and Q are annihilation and creation operators. Let

$$e^{zP}e^{xQP}e^{yQ} = \mathcal{N}[\exp(fQP + gP + hQ + sI)]. \quad (10.16)$$

In Eq. (10.16), we regard f, g, h, and s to be scalar functions of x which depend parametrically upon z and y. Differentiating Eq. (10.16) with respect to x and substituting Eq. (10.16) into the result, one obtains

$$e^{zP}QPe^{xQP}e^{yQ} = Qe^{zP}e^{xQP}e^{yQ}(f'P + h') + e^{zP}e^{xQP}e^{yQ}(g'P + s'I). \quad (10.17)$$

Multiplying Eq. (10.17) from the right by $e^{-yQ}e^{-xQP} \times e^{-zP}$, carrying out the similarity transformations, and equating coefficients of QP, P, Q, and I, respectively, on the two sides of the equation, one obtains

$$\begin{aligned} f' &= e^{cx}, \\ g' &= cze^{cx}, \\ h' &= cyf' = cye^{cx}, \\ s' &= cyg' = c^2yze^{cx}. \end{aligned} \quad (10.18)$$

Integrating Eqs. (10.18) subject to the conditions $f(0) = 0$, $g(0) = z$, $h(0) = y$, and $s(0) = cyz$ and substituting the results into Eq. (10.16), one obtains

$$e^{zP}e^{xQP}e^{yQ} = \mathcal{N}\{\exp[c^{-1}(e^{cx} - 1)QP + e^{cx}(zP + yQ + cyzI)]\}. \quad (10.19)$$

A special case of interest is obtained by setting $z = 0$, $P = a$, $Q = a^\dagger$, and $c = 1$ in Eq. (10.19):

$$e^{xa^\dagger a}e^{ya^\dagger} = \exp(ye^x a^\dagger)\mathcal{N}\{\exp[(e^x - 1)a^\dagger a]\}. \quad (10.20)$$

By the method of linear superposition on y, it follows from Eq. (10.20) that

$$e^{xa^\dagger a}f(a^\dagger) = f(e^x a^\dagger)\mathcal{N}\{\exp[(e^x - 1)a^\dagger a]\}. \quad (10.21)$$

A special case of Eq. (10.20) or (10.21) has been derived in Ref. 3 by a different method:

$$e^{xa^\dagger a} = \mathcal{N}\{\exp[(e^x - 1)a^\dagger a]\} \quad (10.22a)$$

$$= \sum_{r=0}^{\infty}(e^x - 1)^r(r!)^{-1}a^{\dagger r}a^r. \quad (10.22b)$$

This result is also a special case of a theorem due to McCoy. [See Eqs. (10.38) and (10.41).] Another result given in Ref. 3 is easily obtained from Eq. (10.22b) by using the binomial expansion,

$$(e^x - 1)^r = \sum_{s=0}^{r}\frac{(-)^{r-s}r!\,e^{sx}}{s!\,(r-s)!},$$

and employing the method of linear superposition on x:

$$f(a^\dagger a) = \sum_{r=0}^{\infty}\sum_{s=0}^{r}\frac{(-)^{r-s}f(s)}{s!\,(r-s)!}a^{\dagger r}a^r. \quad (10.23)$$

It is sometimes useful to put expressions in antinormal form. Let $\bar{N}$ be the antinormal-ordering operator defined similarly to $\mathcal{N}$ except that it puts the a's to the left of the $a^\dagger$'s. Thus the antiequation of (10.22a) is given by

$$e^{xa^\dagger a} = e^{-x}\bar{N}\{\exp[(1 - e^{-x})a^\dagger a]\}. \quad (10.24)$$

Substituting Eq. (10.24) into Eq. (10.21) and letting $(1 - e^{-x}) = z$, one obtains a result derived by Schwinger in a different manner[17]:

$$\bar{N}[e^{za^\dagger a}]f(a^\dagger) = \frac{1}{1-z}f\left(\frac{a^\dagger}{1-z}\right)\mathcal{N}\left[\exp\left(\frac{za^\dagger a}{1-z}\right)\right]. \quad (10.25a)$$

Another formula given by Schwinger is similarly derived:

$$f(a)\bar{N}e^{za^\dagger a} = \mathcal{N}\left[\exp\left(\frac{za^\dagger a}{1-z}\right)\right]\frac{1}{1-z}f\left(\frac{a}{1-z}\right). \quad (10.25b)$$

Using these formulas with $a^\dagger = x$ and $a = \partial/\partial x$, Schwinger has derived some interesting operator identities and classical formulas involving cylinder

functions and associated Laguerre polynomials. A corollary of one of these is the formula

$$a^m(a^\dagger)^{m+r} = m!\,(a^\dagger)^r \mathcal{N}[L_m^r(-a^\dagger a)], \quad (10.26)$$

a result which one may also obtain directly from Eq. (10.7) by using the definition of the associated Laguerre polynomial.

The next derivation shows how the exponential of an arbitrary second-degree polynomial may be put into normal form. This result generalizes a theorem of McCoy.[18]

Theorem: Let

$$e^Z = \mathcal{N}(e^W), \quad (10.27)$$

where

$$Z \equiv \alpha P^2 + \beta Q^2 + \gamma QP + \delta P + \epsilon Q, \quad (10.28)$$

$$W \equiv AP^2 + BQ^2 + GQP + DP + EQ + FI. \quad (10.29)$$

Then Eq. (10.27) is identically satisfied provided

$$\begin{aligned} &\alpha^{-1}A = \beta^{-1}B = H \equiv (\lambda J)^{-1}\sinh\lambda, \\ &G = c^{-1}(J^{-1} - 1), \\ &D = \rho^2(\tau G + 2\mu A), \\ &E = \rho^2(\mu G + 2\tau B), \\ &F = -\tfrac{1}{2}\ln J - \tfrac{1}{2}c\gamma + \rho^2(\varphi - \gamma\delta\epsilon) + \rho^4\tau\mu G \\ &\qquad + \rho^4(4\alpha\beta\varphi + \gamma^2\varphi - 8\alpha\beta\gamma\delta\epsilon)H, \end{aligned} \quad (10.30)$$

where

$$\begin{aligned} &J \equiv \cosh\lambda - \rho\gamma\sinh\lambda, \quad \rho \equiv \lambda^{-1}c, \\ &\lambda = c[\gamma^2 - 4\alpha\beta]^{\frac{1}{2}}, \quad \varphi \equiv \alpha\epsilon^2 + \beta\delta^2, \\ &\tau \equiv \gamma\delta - 2\alpha\epsilon, \quad \mu \equiv \gamma\epsilon - 2\beta\delta. \end{aligned} \quad (10.31)$$

Proof: Instead of Eq. (10.27), consider

$$e^{xZ} = \mathcal{N}[e^{W(x)}], \quad (10.32)$$

where $W(0) = 0$. In Eq. (10.29), W, A, B, G, D, E, and F are now considered to be functions of x. Differentiating Eq. (10.32) with respect to x, one obtains

$$\begin{aligned} Ze^{xZ} &= \mathcal{N}[W'(x)e^{W(x)}] \\ &= A'e^{xZ}P^2 + B'Q^2e^{xZ} + G'Qe^{xZ}P \\ &\quad + D'e^{xZ}P + E'Qe^{xZ} + F'e^{xZ}, \end{aligned} \quad (10.33)$$

where $A' \equiv \partial A/\partial x$, etc. Multiplying Eq. (10.33) from the right by e^{-xZ}, one obtains

$$\begin{aligned} Z = A'e^{xZ}P^2e^{-xZ} + (G'Q + D')e^{xZ}Pe^{-xZ} \\ + B'Q^2 + E'Q + F'I. \end{aligned} \quad (10.34)$$

The quantities $e^{xZ}Pe^{-xZ}$ and

$$e^{xZ}P^2e^{-xZ} = [e^{xZ}Pe^{-xz}]^2$$

are obtained from Eqs. (6.48)–(6.51). Equating coefficients of P^2, QP, Q^2, P, Q, and I, respectively, on the two sides of Eq. (10.34) leads to the differential equations

$$\begin{aligned} &\alpha = A'd^2, \quad \gamma = 2A'de + G'd, \\ &\beta = A'e^2 + B' + G'e, \quad \delta = 2A'df + D'd, \\ &\epsilon = 2A'ef + G'f + D'e + E', \\ &0 = A'f^2 + cA'de + D'f + F', \end{aligned} \quad (10.35)$$

where $d \equiv d(x)$, $e \equiv e(x)$, and $f \equiv f(x)$ are defined by Eqs. (6.50). Equations (10.35) simplify to

$$\begin{aligned} &A'/\alpha = B'/\beta = d^{-2}, \quad G' = -c^{-1}d^{-2}d', \\ &D' = \rho^2(\tau G' + 2\mu A'), \quad E' = \rho^2(\mu G' + 2\tau B'), \\ &F' = -\tfrac{1}{2}d'd^{-1} + \rho^2(\varphi - \gamma\delta\epsilon) + \rho^4\tau\mu G' \\ &\qquad -\tfrac{1}{2}c\gamma + \rho^4(4\alpha\beta\varphi + \gamma^2\varphi - 8\alpha\beta\gamma\delta\epsilon)d^{-2}. \end{aligned} \quad (10.36)$$

Integrating Eqs. (10.36) subject to $A(0) = \cdots = F(0) = 0$ and setting $x = 1$, one obtains Eqs. (10.30) for $A \equiv A(1)$, $B \equiv B(1)$, etc. Q.E.D.

Using Eq. (2.8), one may verify that both sides of Eq. (10.27) satisfy the pair of partial differential equations

$$\partial U/\partial P = 2AUP + GQU + DU, \quad (10.37a)$$

$$\partial U/\partial Q = GUP + 2BQU + EU. \quad (10.37b)$$

Some special cases of Eqs. (10.27)–(10.31) are of interest. If $\delta = \epsilon = 0$, then $D = E = 0$ and

$$\begin{aligned} &\exp(\alpha P^2 + \beta Q^2 + \gamma QP) \\ &\quad = [Je^{c\gamma}]^{-\frac{1}{2}}\mathcal{N}[\exp(AP^2 + BQ^2 + GQP)], \end{aligned} \quad (10.38)$$

where A, B, G, and J are given by Eqs. (10.30) and (10.31). This result may be shown to be equivalent to McCoy's theorem.[18] We note that McCoy gave an ingenious derivation based upon a pair of partial differential equations like Eqs. (10.37) with $D = E = 0$. If we set $\gamma = 0$ in Eq. (10.28) and define

$$S \equiv P + \theta I, \quad R \equiv Q + \sigma I, \quad y \equiv 2c(\alpha\beta)^{\frac{1}{2}}, \quad (10.39)$$

then

$$\begin{aligned} &\exp(\alpha S^2 + \beta R^2) \\ &\quad = (\sec y)^{\frac{1}{2}}\mathcal{N}\{\exp[(y^{-1}\tan y)(\alpha S^2 + \beta R^2) \\ &\qquad + c^{-1}(\sec y - 1)RS]\}. \end{aligned} \quad (10.40)$$

If we set $\alpha = \beta = 0$ in Eq. (10.28) and define S and R by Eqs. (10.39), then

$$e^{\gamma RS} = \mathcal{N}\{\exp[c^{-1}(e^{c\gamma} - 1)RS]\}. \quad (10.41)$$

Equations (10.40) and (10.41) may also be derived from Eq. (10.38) by using the fact that S and R satisfy the same commutation relation as P and Q.

An expansion of Kermack and McCrea, which

provided the starting point for Sach's derivation of his "Taylor's theorem for shift operators,"[11] is as follows[19]:

$$F[P + \varphi(Q)]$$

$$= \mathcal{N}\left\{\left[\exp\int_0^{\partial/\partial P} dx\varphi(Q + cx)\right]F(P)\right\} \quad (10.42a)$$

$$= \mathcal{N}\{\exp[\varphi(Q)\partial/\partial P + \tfrac{1}{2}c\varphi'(Q)\partial^2/\partial P^2 + \tfrac{1}{6}c^2\varphi''(Q)\partial^3/\partial P^3 + \cdots]\}F(P). \quad (10.42b)$$

For the special case where $F(P) = e^{\lambda P}$, Eq. (10.42a) easily follows from Eq. (7.33), while the general case follows by the method of linear superposition. Equation (10.42b) follows from Eq. (10.42a) by Taylor-expanding $\varphi(Q + cx)$ about $x = 0$ and carrying out the integration over x.

A result recently obtained by Cohen is readily obtained by setting $\varphi(Q) = Q$ and $F(P) = P^n$ in Eq. (10.42b).

$$\begin{aligned}(P + Q)^n &= \mathcal{N}[\exp(Q\partial/\partial P + \tfrac{1}{2}c\partial^2/\partial P^2)]P^n \\ &= \mathcal{N}\left[\exp\left(\frac{Q\partial}{\partial P}\right)\sum_{k=0}^{\infty}\frac{(\frac{1}{2}c)^k\partial^{2k}P^n/\partial P^{2k}}{k!}\right. \\ &= \sum_{k=0}^{[\frac{1}{2}n]}\frac{(\frac{1}{2}c)^k n!}{k!\,(n-2k)!}\sum_{s=0}^{\infty}\frac{Q^s(\partial^s P^{n-2k}/\partial P^s)}{s!} \\ &= \sum_{k=0}^{[\frac{1}{2}n]}\sum_{s=0}^{n-2k}\frac{(\frac{1}{2}c)^k n!\,Q^s P^{n-2k-s}}{k!\,s!\,(n-2k-s)!}. \end{aligned} \quad (10.43)$$

Equation (10.43) may be easily shown to be equivalent to Cohen's result which he obtained as an application of his "Expansion Theorem for Functions of Operators."[20] In contrast to the normal-ordering theorems considered above, this theorem requires for its application that one knows the solution to an eigenvalue problem. A slightly more general statement of Cohen's theorem and a briefer proof of it follow.

Theorem: Let $F(q, p)$ be a Hermitian operator function of q and p, with $[q, p] = i\hbar$, which satisfies the eigenvalue equation

$$F|k\rangle = \alpha_k|k\rangle. \quad (10.44)$$

Then any operator function $g(F)$ may be represented by

$$g(F) = \sum_{k=0}^{\infty} g(\alpha_k)\psi_k(q)\int_{-\infty}^{\infty}\psi_k^*(q + \theta)e^{i\theta p/\hbar}\,d\theta, \quad (10.45)$$

where $\psi_k(q') \equiv \langle q'|k\rangle$. If the eigenvalue spectrum of F is continuous, the summation is to be replaced by an integration.

Proof: Equation (10.45) may be verified by showing that the matrix elements of both sides are the same between states $\langle q'|$ and $|q''\rangle$. Since $g(F)$ has the same eigenstates as F, the left-hand side of Eq. (10.45) may be written

$$g(F) = \sum_{k=0}^{\infty} g(\alpha_k)|k\rangle\langle k|. \quad (10.46)$$

Hence

$$\langle q'|\,g(F)\,|q''\rangle = \sum_{k=0}^{\infty} g(\alpha_k)\psi_k(q')\psi_k^*(q''). \quad (10.47)$$

The corresponding matrix element of the right-hand side of Eq. (10.45) is given by

$$\sum_{k=0}^{\infty} g(\alpha_k)\psi_k(q')\int_{-\infty}^{\infty}\psi_k^*(q' + \theta)\langle q'|\,e^{i\theta p/\hbar}\,|q''\rangle\,d\theta. \quad (10.48)$$

From Eq. (2.11), we have

$$\begin{aligned}\langle q'|\,e^{i\theta p/\hbar}\,|q''\rangle &= \langle q'\,|\,q'' - \theta\rangle \\ &= \delta(q' - q'' + \theta). \end{aligned} \quad (10.49)$$

Substituting Eq. (10.49) into Eq. (10.48), one obtains Eq. (10.47). Q.E.D.

11. HIGHER DERIVATIVES

Higher derivatives of exponential operators may be obtained straightforwardly by repeated application of Eq. (2.1). The results obtained, however, are not in the most concise and symmetrical form possible. Although they may be put into a symmetrical "time-ordered" form by a change of integration variables, we prefer to proceed in a different manner which makes use of an integral representation due to Poincaré[21]:

$$f(H) = (2\pi i)^{-1}\int_c f(z)(zI - H)^{-1}\,dz. \quad (11.1)$$

In Eq. (11.1), it is assumed that the contour of integration encloses a region of the complex z plane where all the eigenvalues of H lie and throughout which $f(z)$ is analytic. Equation (11.1) may be verified by taking matrix elements of both sides in a representation in which H is diagonal, and employing the Cauchy integral formula. For convenience, we assume that H is a positive-definite Hermitian operator with a discrete eigenvalue spectrum although the results obtained are often valid under less restrictive conditions.

We consider first the case where $f(z) = e^{-\beta z}$, and rotate 90° to the x plane defined by $x \equiv iz$. From Eq. (11.1), it then follows that

$$E(\beta) \equiv (2\pi i)^{-1}\int_{-\infty}^{\infty} e^{i\beta x}(x - iH)^{-1}\,dx \quad (11.2a)$$

$$= \begin{cases} e^{-\beta H}, & \beta > 0, \\ 0, & \beta < 0. \end{cases} \quad (11.2b)$$

The cases $\beta > 0$ and $\beta < 0$, respectively, are obtained by choosing the contours of integration to be infinite semicircles in the upper- and lower-half x plane. For the special case where $H \equiv H(\lambda)$ is of the form $H = \mathcal{H} + \lambda V$, one finds that the nth derivative of $E(\beta)$ with respect to λ, $\partial^n E(\beta)/\partial\lambda^n \equiv E^{(n)}(\beta)$, is given by

$$\frac{n!\, i^n}{2\pi i}\int_{-\infty}^{\infty} dx e^{i\beta x}(x - iH)^{-1}[V(x - ih)^{-1}]^n. \quad (11.3)$$

Equation (11.3) follows directly from Eq. (11.2a) by using the formula for the derivative of the inverse,

$$\partial[(x - iH)^{-1}]/\partial\lambda = i(x - iH)^{-1}V(x - iH)^{-1}.$$

A recursive formula for $E^{(n)}(\beta)$ which shows its "time-ordered" form is given by

$$E^{(n)}(\beta) = -n\int_0^{\beta} du E(\beta - u)VE^{(n-1)}(u). \quad (11.4)$$

Equation (11.4) follows from Eq. (11.3) by writing it in the form

$$\begin{aligned}\frac{n!\, i^n}{2\pi i}\iint_{-\infty}^{\infty} dx\, dy \delta(x - y)e^{i\beta x}(x - iH)^{-1}[V(y - iH)^{-1}]^n \\ = \frac{n!\, i^n}{4\pi^2 i}\iiint_{-\infty}^{\infty} du\, dx\, dy e^{iu(y-x)}e^{i\beta x} \\ \times (x - iH)^{-1}[V(y - iH)^{-1}]^n \\ = -n\int_{-\infty}^{\infty} du E(\beta - u)VE^{(n-1)}(u). \quad (11.5)\end{aligned}$$

In the integrations over x and y, we have used Eqs. (11.2) and (11.3), respectively. Equation (11.5) is the same as Eq. (11.4), since $E(\beta - u) = 0$ if $u > \beta$, and $E^{(n-1)}(u) = 0$ if $u < 0$. For the case where $n = 1$, since $E^{(0)}(u) \equiv E(u) = e^{-uH}$, Eq. (11.4) constitutes another derivation of Eq. (2.1). An equation like Eq. (11.4) has been derived by Kumar in a different manner, while the derivatives of the exponential evaluated at $\lambda = 0$ occur in a well-known expansion of $e^{\mathcal{H}+\lambda V}$. Of course, higher derivatives may also be calculated from Eqs. (11.2) in the same manner when H does not just depend linearly upon λ, but the results will be more complicated. For example, the second derivative, $E''(\beta) = \partial^2 e^{-\beta H}/\partial\lambda^2$, is then given by[64]

$$\begin{aligned}E''(\beta) = -\int_0^{\beta} du E(\beta - u)H''E(u) \\ + 2\int_0^{\beta} du\int_0^{u} dv E(\beta - u)H'E(u - v)H'E(v),\end{aligned} \quad (11.6)$$

where $H' \equiv \partial H/\partial\lambda$ and $H'' \equiv \partial^2 H/\partial\lambda^2$.

[64] An equivalent formula given by Snider, Ref. 23, Eq. (B7), is not as concise and symmetrical.

The Poincaré formula may also be used to efficiently obtain results of the type derived by Aizu for general operator functions.[2] A formula from which a number of sum rules can be derived is obtained by considering H to be a function of two parameters, λ and μ, and taking matrix elements of $\partial^2 f(H)/\partial\lambda\partial\mu$. Then from Eq. (11.1), one obtains

$$\langle m|\frac{\partial^2 f(H)}{\partial\lambda\partial\mu}|r\rangle = \langle m|A_{\lambda\mu} + A_{\mu\lambda} + B_{\lambda\mu}|r\rangle, \quad (11.7)$$

where

$$\langle m|A_{\lambda\mu}|r\rangle \equiv \sum_n \langle m|\frac{\partial H}{\partial\lambda}|n\rangle\langle n|\frac{\partial H}{\partial\mu}|r\rangle S_{mnr} \quad (11.8)$$

and

$$\langle m|B_{\lambda\mu}|r\rangle \equiv \langle m|\frac{\partial^2 H}{\partial\lambda\partial\mu}|r\rangle T_{mr}. \quad (11.9)$$

In Eqs. (11.8) and (11.9), S_{mnr} and T_{mr} are defined by

$$S_{mnr} \equiv \frac{1}{2\pi i}\int_c \frac{f(z)\,dz}{(z - H_m)(z - H_n)(z - H_r)}, \quad (11.10)$$

$$T_{mr} \equiv \frac{1}{2\pi i}\int_c \frac{f(z)\,dz}{(z - H_m)(z - H_r)}. \quad (11.11)$$

Performing the contour integrations, one obtains

$$T_{mr} = \Delta_{mr}f'(H_m) + \rho_{mr}\frac{f(H_m) - f(H_r)}{H_m - H_r} \quad (11.12)$$

and

$$S_{mnr} = \tfrac{1}{2}\Delta_{mn}\Delta_{nr}f''(H_m) + U_{mnr} + U_{rmn} + U_{nrm}, \quad (11.13)$$

where

$$\Delta_{mn} \equiv 1 - \rho_{mn} \equiv \begin{cases}1, & H_m = H_n, \\ 0, & H_m \neq H_n,\end{cases} \quad (11.14)$$

$$\begin{aligned}U_{mnr} \equiv \frac{\rho_{mn}\rho_{nr}\rho_{rm}f(H_m)}{(H_m - H_n)(H_m - H_r)} \\ + \Delta_{mn}\rho_{nr}\left[\frac{f(H_r) - f(H_m)}{(H_r - H_m)^2} - \frac{f'(H_m)}{H_r - H_m}\right].\end{aligned} \quad (11.15)$$

Equations (11.7)–(11.9) and (11.12)–(11.15) may be shown to be equivalent to [25] of Aizu.[2]

It is straightforward to obtain matrix elements of higher derivatives by the same method, but the results obtained are complicated. If one specializes the results so obtained to the case where $f(H) = f(\mathcal{H} + \lambda V) = \exp(\mathcal{H} + \lambda V)$ and compares with the matrix elements of Eq. (11.4), one finds that the latter results do not immediately have such a simple form. Comparison of the two forms reveals the existence of the curious identity

$$\sum_{j=0}^{N}\prod_{\substack{k=1\\k\neq j}}^{N}(x_j - x_k)^{-1} = 0, \quad (11.16)$$

where the variables $x_1, x_2, \cdots, x_N$, stand for H_m, H_n, H_l, etc., and it is assumed that no two are equal. Equation (11.6) may also be derived from the Lagrangian interpolation formula.[65]

Poincaré's formula is also well suited to derive [27] and [28] of Aizu, which depend upon the cyclic property of the trace. We indicate the proof for [27].

Theorem[2]*:*

$$\mathrm{Tr}\left[\varphi(H)\frac{\partial f(H)}{\partial\lambda}\frac{\partial g(H)}{\partial\mu}\right] = \mathrm{Tr}\left[\varphi(H)\frac{\partial g(H)}{\partial\lambda}\frac{\partial f(H)}{\partial\mu}\right],$$

where φ, f, and g are arbitrary functions.

[65] The author is indebted to Dr. P. F. Wacker for suggesting to him that this might be the case.

Proof:

$$\mathrm{Tr}\left[\varphi(H)\frac{\partial f(H)}{\partial\lambda}\frac{\partial g(H)}{\partial\mu}\right] = \frac{1}{(2\pi)^2}\int_c f(z)\,dz\int_c g(w)\,dw \times \mathrm{Tr}\left[\varphi(H)(z-H)^{-1}\frac{\partial H}{\partial\lambda} \times (z-H)^{-1}(w-H)^{-1}\frac{\partial H}{\partial\mu}(w-H)^{-1}\right].$$

Since $(z-H)^{-1}$, $(w-H)^{-1}$, and $\varphi(H)$ commute, z and w can effectively be interchanged in the trace, so that the theorem follows.

ACKNOWLEDGMENT

The author wishes to thank Dr. J. H. Shirley for checking many of the results.