Representations of the Lorentz Group: New Integral Relations between Legendre Functions

P. PAJAS* AND P. WINTERNITZ*
Rutherford High Energy Laboratory, Chilton, Didcot, Berkshire, England

(Received 1 August 1969)

New integral relations between Legendre functions of the first and second kind are derived. These functions figure as the basis functions of the irreducible representations of the homogeneous Lorentz group, so that the derived formulas have direct applications to invariant expansions of relativistic amplitudes.

1. INTRODUCTION

In the last few years, the theory of the irreducible representations of the homogeneous Lorentz group has received a large amount of attention and has been widely applied in elementary particle scattering theory. In particular, 2-variable expansions of relativistic amplitudes, based on the homogeneous Lorentz group $O(3, 1)$, have been introduced and discussed in a series of papers.[1-4] The $O(3, 1)$ group here figures as the group of motions of the space of independent kinematic variables (a hyperboloid in the 4-dimensional relativistic velocity space) and generates expansions simultaneously in the Mandelstam s and t variables (or rather in suitable combinations of s and t). Since the $O(3, 1)$ group in this approach does not figure as an invariance group of the scattering amplitude (except for elastic scattering at $t = 0$), it is possible to apply the expansions for arbitrary values of s and t and for arbitrary masses.

On the other hand, the homogeneous Lorentz group is the little group of the Poincaré group corresponding to zero momentum transfer and, thus, is an invariance group for elastic forward scattering, which makes it possible to write quite specific 1-variable expansions at $t = 0$ in terms of the irreducible representations of $O(3, 1)$ (see Refs. 5 and 6 which also contain a complete list of references). These 1-variable expansions have also been generalized[7,8] to scattering for $t \neq 0$ and applied to various reactions among elementary particles.

The 2-variable expansions mentioned above have been developed in detail for the 2-body scattering of particles with spin zero. The scattering amplitude $f(s, t)$ is written as a function of a single point v, the components of which satisfy $v^2 = v_0^2 - v_1^2 - v_2^2 - v_3^2 = 1$, and is expanded in terms of the basis functions of $O(3, 1)$. In the simple spin-zero case considered, these are just harmonic functions on the hyperboloid $v^2 = 1$. Different expansions are obtained by considering different possible reductions of $O(3, 1)$ to its subgroups, and the reductions $O(3, 1) \supset O(3) \supset O(2)$, $O(3, 1) \supset O(2, 1) \supset O(2)$, and $O(3, 1) \supset E_2 \supset O(2)$ are of special interest since the subgroups $O(3)$, $O(2, 1)$, and E_2 also figure as the little groups of the physical Poincaré group leaving, respectively, the total energy vector, or the spacelike, or lightlike momentum transfer invariant.

In this paper, we concentrate on the special functions appearing in the reductions to the $O(3)$ and $O(2, 1)$ subgroups. The expansions of the scattering amplitude can be written as[1-3]

$$f(s,t) \equiv f(a, \vartheta) = \sum_{l=0}^{\infty} (2l+1) \int_0^{\infty} p^2\, dp A_l(p) \frac{\Gamma(ip)}{\Gamma(ip-l)} \frac{1}{(\sinh a)^{\frac{1}{2}}} \times P^{-\frac{1}{2}-l}_{-\frac{1}{2}-ip}(\cosh a) \mathbb{P}_l(\cos\vartheta) \quad (1)$$

and

$$f(s,t) \equiv f(\alpha, \beta) = \frac{1}{2i} \int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty} \frac{dl(2l+1)}{\sin \pi l} \times \int_0^{\infty} p^2\, dp \frac{\Gamma(l+1+ip)\Gamma(-l+ip)}{\Gamma(1+ip)} \frac{\cos \pi l}{\cosh \alpha} \times [B^+(l,p)\mathbb{P}_l^{-ip}(-\tanh\alpha) + B^-(l,p)\mathbb{P}_l^{-ip}(\tanh\alpha)] P_l(\cosh\beta), \quad (2)$$

respectively.

Here $A_l(p)$ and $B^{\pm}(l, p)$ are the "Lorentz amplitudes" which carry all the dynamics, $P^{\mu}_{\nu}(z)$ and $\mathbb{P}^{\mu}_{\nu}(x)$ are Legendre functions of the first kind in the complex z plane and on the $x \in [-1, 1]$ cut, respectively.[9] If we write expansion (1) in the physical region of the t channel and (2) in the s channel, then the variables are

$$z = \cos\vartheta = \cosh\beta = -[2t(s - m_1^2 - m_2^2) + (t + m_1^2 - m_3^2)(t + m_2^2 - m_4^2)] \times \{[t - (m_1 + m_3)^2][t - (m_1 - m_3)^2] \times [t - (m_2 + m_4)^2][t - (m_2 - m_4)^2]\}^{-\frac{1}{2}}, \quad (3)$$

$$\eta = \coth a = \tanh\alpha = (t + m_1^2 - m_3^2) \times \{[t - (m_1 + m_3)^2][t - (m_1 - m_3)^2]\}^{-\frac{1}{2}}. \quad (4)$$

Formulas (1) and (2) correspond to expansions in terms of the unitary representations of the principal series and their convergence implies that the amplitude $f(s, t)$ is square integrable over the whole hyperboloid (generalizations are considered, e.g., in Ref. 3).

The inverse formulas to (1) and (2) can be written as

$$A_l(p) = \frac{\Gamma(-ip)}{2\Gamma(-ip-l)} \int_0^\infty \sinh^2 a\, da \times \int_0^\pi \sin\vartheta\, d\vartheta f(a,\vartheta) \frac{1}{(\sinh a)^{\frac{1}{2}}} P^{-l-\frac{1}{2}}_{-\frac{1}{2}-ip}(\cosh a)\mathbb{P}_l(z)$$

or, making use of HTF I, 3.3 (13), as

$$A_l(p) = -\frac{\Gamma(-ip)e^{p\pi}\sin(l+ip)\pi}{\pi}\left(\frac{1}{2\pi}\right)^{\frac{1}{2}} \times \int_1^\infty \frac{d\eta}{(\eta^2-1)^{\frac{3}{2}}} Q_l^{ip}(\eta) \int_{-1}^1 dz f(\eta, z)\mathbb{P}_l(z), \quad (5)$$

[where $Q^\mu_\nu(z)$ is the Legendre function of the second type] and

$$B^{\pm}(l,p) = \frac{2\Gamma(-l-ip)\Gamma(l+1-ip)}{\Gamma(1-ip)} \times \int_{-1}^1 \frac{d\eta}{(1-\eta^2)^{\frac{3}{2}}} \mathbb{P}_l^{ip}(\mp\eta) \int_1^\infty dz f(\eta, z) P_l(z), \quad l = -\tfrac{1}{2} + iq, \quad q \text{ real}. \quad (6)$$

When considering the crossing transformation in the framework of 2-variable $O(3, 1)$ expansions[10] it is possible to relate expansions (1) and (2) to each other via a transformation of the Sommerfeld–Watson type followed by analytic continuation in η and z. $A_l(p)$ and $B^{\pm}(l, p)$ can be considered as different "pieces" of certain analytic functions $\mathcal{A}^{\pm}(l, p)$ defined in the whole complex l-plane.

To do this it is necessary to derive new integral formulas, relating the Legendre functions of the first and second kind to each other. Since we feel that these are of interest both for the representation theory of the Lorentz group and other classical noncompact groups[11] and for other purposes in mathematical physics, we have decided to publish them separately.

2. INTEGRAL RELATIONS BETWEEN LEGENDRE FUNCTIONS

In this section, we give a summary of the obtained formulas, together with their regions of applicability, while postponing the proofs to the following section.

Applying Cauchy's integral formula to the function $Q^\mu_\nu(z)(z^2-1)^a z^b$, which is holomorphic in the z plane cut from 1 to $-\infty$, we obtain

$$\begin{aligned} &Q^\mu_\nu(z)(z^2-1)^a z^b \\ &= -\frac{1}{\pi}\sin\pi(\nu-2a-b) \\ &\times \int_1^\infty \frac{Q^\mu_\nu(z')[(z')^2-1]^a (z')^b\, dz'}{z'+z} \\ &+ \frac{e^{i\mu\pi}}{2\sin\pi(\nu+\mu)}\Bigg(\int_0^1 [\sin\pi(\nu+\tfrac{1}{2}\mu-a-b)\mathbb{P}^\mu_\nu(-z') \\ &+ \sin\pi(\tfrac{1}{2}\mu+a+b)\mathbb{P}^\mu_\nu(+z')]\frac{[1-(z')^2]^a (z')^b\, dz'}{z'+z} \\ &- \int_0^1 [\sin\pi(\nu+\tfrac{1}{2}\mu-a)\mathbb{P}^\mu_\nu(z') \\ &+ \sin\pi(\tfrac{1}{2}\mu+a)\mathbb{P}^\mu_\nu(-z')]\frac{[1-(z')^2]^a (z')^b\, dz'}{z'-z}\Bigg), \end{aligned} \quad (7)$$

valid for

$$\operatorname{Re} b > -1, \quad |\operatorname{Re}\mu| < 2\operatorname{Re} a + 2, \quad \operatorname{Re}\nu > 2\operatorname{Re} a + \operatorname{Re} b - 1, \quad \operatorname{Re} a > -1, \quad \nu+\mu \neq \text{integer}. \quad (8)$$

[Conditions (8) and all following validity conditions are those under which the corresponding formulas have been derived. A possible extension of the region of applicability of the individual integral formulas must be subjected to further investigations.]

Considering the same function as a function of z^2 on a 2-sheet Riemann surface, we obtain

$$\begin{aligned} &Q^\mu_\nu(z)(z^2-1)^a z^b \\ &= -\frac{2}{\pi}\sin\pi(\nu-2a-b) \\ &\times \int_1^\infty \frac{Q^\mu_\nu(z')[(z')^2-1]^a (z')^{b+1}\, dz'}{(z')^2-z^2} \\ &+ \frac{e^{i\mu\pi}}{\sin\pi(\nu+\mu)}\Bigg(\int_0^1 [\sin\pi(\nu+\tfrac{1}{2}\mu-a-b)\mathbb{P}^\mu_\nu(-z') \\ &+ \sin\pi(\tfrac{1}{2}\mu+a+b)\mathbb{P}^\mu_\nu(+z')]\frac{[1-(z')^2]^a (z')^{b+1}\, dz'}{(z')^2-z^2} \\ &- \int_0^1 [\sin\pi(\nu+\tfrac{1}{2}\mu-a)\mathbb{P}^\mu_\nu(z') \\ &+ \sin\pi(\tfrac{1}{2}\mu+a)\mathbb{P}^\mu_\nu(-z')]\frac{[1-(z')^2]^a (z')^{b+1}\, dz'}{(z')^2-z^2}\Bigg), \end{aligned} \quad (9)$$

which is valid for

$$\operatorname{Re} a > -1, \quad |\operatorname{Re}\mu| < 2\operatorname{Re} a + 2, \quad \operatorname{Re}\nu > 2\operatorname{Re} a + \operatorname{Re} b - 1, \quad \operatorname{Re} b > -2, \quad \nu+\mu \neq \text{integer}. \quad (10)$$

Combining (7) and (9), we derive the integral formula

$$\frac{1}{\pi}\sin\pi(\nu-2a-b)\int_1^\infty \frac{Q_\nu^\mu(z')[(z')^2-1]^a(z')^b\,dz'}{z'-z}$$
$$=\frac{e^{i\mu\pi}}{2\sin\pi(\nu+\mu)}\Bigg(\int_0^1[\sin\pi(\nu+\tfrac{1}{2}\mu-a-b)\mathbb{P}_\nu^\mu(-z')$$
$$+\sin\pi(\tfrac{1}{2}\mu+a+b)\mathbb{P}_\nu^\mu(+z')]\frac{[1-(z')^2]^a(z')^b\,dz'}{z'-z}$$
$$-\int_0^1[\sin\pi(\nu+\tfrac{1}{2}\mu-a)\mathbb{P}_\nu^\mu(z')$$
$$+\sin\pi(\tfrac{1}{2}\mu+a)\mathbb{P}_\nu^\mu(-z')]\frac{[1-(z')^2]^a(z')^b\,dz'}{z'+z}\Bigg), \tag{11}$$

valid for

$$\operatorname{Re} b>-1,\quad |\operatorname{Re}\mu|<2\operatorname{Re}a+2,$$
$$\operatorname{Re}\nu>2\operatorname{Re}a+\operatorname{Re}b-1,\quad \operatorname{Re}a>-1,$$
$$\nu+\mu\neq\text{integer}. \tag{12}$$

Similarly, when we consider the Legendre function of the first type in the z plane, we obtain

$$P_\nu^\mu(z)(z^2-1)^a z^b$$
$$=-\frac{1}{\pi}\sin\pi(\nu+2a+b)\int_1^\infty\frac{P_\nu^\mu(z')[(z')^2-1]^a(z')^b\,dz'}{z'+z}$$
$$+\frac{2}{\pi^2}\sin\pi(2a+b)\sin\pi(\nu+\mu)$$
$$\times e^{-i\mu\pi}\int_1^\infty\frac{Q_\nu^\mu(z')[(z')^2-1]^a(z')^b\,dz'}{z'+z}$$
$$+\frac{1}{\pi}\sin\pi(\tfrac{1}{2}\mu-a-b)$$
$$\times\int_0^1\frac{\mathbb{P}_\nu^\mu(-z')[1-(z')^2]^a(z')^b\,dz'}{z'+z}$$
$$-\frac{1}{\pi}\sin\pi(\tfrac{1}{2}\mu-a)$$
$$\times\int_0^1\frac{\mathbb{P}_\nu^\mu(z')[1-(z')^2]^a(z')^b\,dz'}{z'-z}, \tag{13}$$

valid for

$$\operatorname{Re}b>-1,\quad |\operatorname{Re}(\nu+\tfrac{1}{2})|<\tfrac{1}{2}-2\operatorname{Re}a-\operatorname{Re}b,$$
$$|\operatorname{Re}\mu|<2\operatorname{Re}a+2,\quad \operatorname{Re}a>-1,$$
$$\tfrac{1}{2}>2\operatorname{Re}a+\operatorname{Re}b>-3. \tag{14}$$

The Cauchy integral formula for the z^2 surface gives

$$P_\nu^\mu(z)(z^2-1)^a z^b$$
$$=-\frac{2}{\pi}\sin\pi(\nu+2a+b)$$
$$\times\int_1^\infty\frac{P_\nu^\mu(z')[(z')^2-1]^a(z')^{b+1}\,dz'}{(z')^2-z^2}$$
$$+\frac{4}{\pi^2}\sin\pi(2a+b)\sin\pi(\nu+\mu)$$
$$\times e^{-i\mu\pi}\int_1^\infty\frac{Q_\nu^\mu(z')[(z')^2-1]^a(z')^{b+1}\,dz'}{(z')^2-z^2}$$
$$+\frac{2}{\pi}\sin\pi(\tfrac{1}{2}\mu-a-b)$$
$$\times\int_0^1\frac{\mathbb{P}_\nu^\mu(-z')[1-(z')^2]^a(z')^{b+1}\,dz'}{(z')^2-z^2}$$
$$-\frac{2}{\pi}\sin\pi(\tfrac{1}{2}\mu-a)$$
$$\times\int_0^1\frac{\mathbb{P}_\nu^\mu(z')[1-(z')^2]^a(z')^{b+1}\,dz'}{(z')^2-z^2}, \tag{15}$$

valid for

$$|\operatorname{Re}\nu+\tfrac{1}{2}|<-2\operatorname{Re}a-\operatorname{Re}b+\tfrac{1}{2},$$
$$|\operatorname{Re}\mu|<2\operatorname{Re}a+2,\quad \operatorname{Re}b>-2,\quad \operatorname{Re}a>-1. \tag{16}$$

Combining (13) and (15), we obtain the integral formula

$$\sin\pi(\nu+2a+b)\int_1^\infty\frac{P_\nu^\mu(z')[(z')^2-1]^a(z')^b\,dz'}{z'-z}$$
$$=\frac{2}{\pi}\sin\pi(2a+b)\sin\pi(\nu+\mu)$$
$$\times e^{-i\mu\pi}\int_1^\infty\frac{Q_\nu^\mu(z')[(z')^2-1]^a(z')^b\,dz'}{z'-z}$$
$$+\sin\pi(\tfrac{1}{2}\mu-a-b)\int_0^1\frac{\mathbb{P}_\nu^\mu(-z')[1-(z')^2]^a(z')^b\,dz'}{z'-z}$$
$$-\sin\pi(\tfrac{1}{2}\mu-a)\int_0^1\frac{\mathbb{P}_\nu^\mu(z')[1-(z')^2]^a(z')^b\,dz'}{z'+z}, \tag{17}$$

valid for

$$\operatorname{Re}b>-1,\quad |\operatorname{Re}(\nu+\tfrac{1}{2})|<\tfrac{1}{2}-2\operatorname{Re}a-\operatorname{Re}b,$$
$$|\operatorname{Re}\mu|<2\operatorname{Re}a+2,\quad \operatorname{Re}a>-1,$$
$$\tfrac{1}{2}>2\operatorname{Re}a+\operatorname{Re}b>-3. \tag{18}$$

The above formulas are written in a general and, therefore, rather complicated form, but they simplify considerably for values of the parameters a, b, ν, and

μ occurring in the $O(3, 1)$ group expansions and in other cases of interest.

For instance, putting $a = b = \mu = 0$ in (7), we obtain a well-known formula used in the complex angular-momentum theory to continue partial-wave amplitudes in l,[12] which for l integer coincides with the Neumann formula HTF I, 3.6 (29).

Of particular interest for the $O(3, 1)$ 2-variable expansions are the following special cases of (11) and (17):

$$\frac{\sin \pi l}{\pi} \int_1^\infty \frac{Q_l^{ip}(\eta')\, d\eta'}{[(\eta')^2 - 1]^{\frac{1}{2}}(\eta' - \eta)}$$

$$= -\frac{e^{-p\pi}}{2 \sin \pi(l + ip)} \int_{-1}^{1} \frac{d\eta'}{[1 - (\eta')^2]^{\frac{1}{2}}(\eta' - \eta)}$$

$$\times [\cos \pi(l + \tfrac{1}{2}ip)\mathbb{P}_l^{ip}(-\eta') - \cosh(\tfrac{1}{2}p\pi)\mathbb{P}_l^{ip}(\eta')], \tag{19}$$

valid for

$$\operatorname{Re} l > -2, \quad |\operatorname{Im} p| < 1, \tag{20}$$

and

$$\sin \pi l \int_1^\infty \frac{d\eta' P_l^{ip}(\eta')}{[(\eta')^2 - 1]^{\frac{1}{2}}(\eta' - \eta)}$$

$$= -\cosh(\tfrac{1}{2}p\pi) \int_{-1}^{1} \frac{d\eta' \mathbb{P}_l^{ip}(-\eta')}{[1 - (\eta')^2]^{\frac{1}{2}}(\eta' - \eta)}, \tag{21}$$

valid for

$$-\tfrac{3}{2} < \operatorname{Re} l < \tfrac{1}{2}, \quad |\operatorname{Im} p| < 1. \tag{22}$$

It is important that the regions of validity of (19) and (21) include $\operatorname{Im} p = 0$, corresponding to unitary irreducible representations of the principal series of $O(3, 1)$. Further, formula (19) is valid for the unitary irreducible representations of the $O(3)$ subgroup ($l =$ nonnegative integer) and for the unitary continuous representations of the $O(2, 1)$ subgroup ($\operatorname{Re} l = -\frac{1}{2}$), whereas (21) holds for those of $O(2, 1)$, but not for those of $O(3)$ (except for $l = 0$).

3. DERIVATION OF THE INTEGRAL FORMULAS

Both the $P_\nu^\mu(z)(z^2 - 1)^a z^b$ and $Q_\nu^\mu(z)(z^2 - 1)^a z^b$ functions are holomorphic in the z-plane cut from $+1$ to $-\infty$ along the real axis, the singular points being $z = +1, 0, -1$ and $|z| = \infty$. So we can use the Cauchy integral formula for an integration contour in the z plane as shown in Fig. 1.

Thus, e.g.,

$$P_\nu^\mu(z)(z^2 - 1)^a z^b = \frac{1}{2\pi i} \int_{C(z)} \frac{P_\nu^\mu(z')[(z')^2 - 1]^a (z')^b\, dz'}{z' - z}. \tag{23}$$

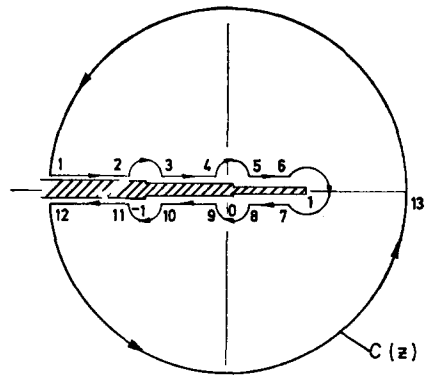


FIG. 1. The integration contour in the z plane.

Since in the limit of $|z| \to \infty$ we have [HTF I, 3.9 (19) and (20)]

$$P_\nu^\mu(z) \xrightarrow[|z| \to \infty]{} 2^{\pm(\nu+\frac{1}{2})-\frac{1}{2}} \times \frac{1}{\sqrt{\pi}} \frac{\Gamma(\pm\nu \pm \frac{1}{2})}{\Gamma(\pm\nu \pm \frac{1}{2} + \frac{1}{2} - \mu)} z^{\pm(\nu+\frac{1}{2})-\frac{1}{2}}, \tag{24}$$

for $\operatorname{Re} \nu \gtrless -\frac{1}{2}$, the integral over the arc 12–13–1, for which $z' = Re^{i\psi}$, vanishes in the limit $R \to \infty$ if

$$|\operatorname{Re} \nu + \tfrac{1}{2}| + 2 \operatorname{Re} a + \operatorname{Re} b - \tfrac{1}{2} < 0. \tag{25}$$

The behavior of $P_\nu^\mu(z)$ at $|z - 1| \to 0$, given by HTF I, 3.9 (3) for $\mu \neq 1, 2, 3, \cdots$,

$$P_\nu^\mu(z) \xrightarrow[|z-1| \to 0]{} \frac{2^{\frac{1}{2}\mu}}{\Gamma(1 - \mu)} (z - 1)^{-\frac{1}{2}\mu}, \tag{26}$$

assures the vanishing of the integral over the arc 6–7 in the limit $|z - 1| \to 0$ if

$$\operatorname{Re}(-\tfrac{1}{2}\mu + a + 1) > 0. \tag{27}$$

In the case of the two arcs around $z = -1$, namely, 2–3 and 10–11, we have to use first the relation HTF I, 3.3 (10):

$$P_\nu^\mu(-z) = e^{\mp\nu\pi i} P_\nu^\mu(z) - (2/\pi) e^{-i\mu\pi} \sin \pi(\nu + \mu) Q_\nu^\mu(z), \tag{28}$$

where the upper or lower sign is to be taken according as $\operatorname{Im} z \gtrless 0$.

Since now

$$Q_\nu^\mu(z) \xrightarrow[|z-1| \to 0]{} e^{i\mu\pi} 2^{-\frac{1}{2}\mu-1} \Gamma(-\mu) \frac{\Gamma(\nu + \mu + 1)}{\Gamma(\nu - \mu + 1)} (z - 1)^{\frac{1}{2}\mu}, \quad \text{for } \operatorname{Re} \mu < 0, \tag{29}$$

and

$$Q_\nu^\mu(z) \xrightarrow[|z-1| \to 0]{} e^{i\mu\pi} 2^{\frac{1}{2}\mu-1} \Gamma(\mu)(z - 1)^{-\frac{1}{2}\mu}, \quad \text{for } \operatorname{Re} \mu > 0, \tag{29'}$$

we find, combining Eqs. (26), (28), (29), and (29′), the condition under which the integrals over 2–3 and 10–11 vanish for $|z+1| \to 0$ to be

$$-\tfrac{1}{2}|\mathrm{Re}\,\mu| + \mathrm{Re}\,a + 1 > 0. \tag{30}$$

The remaining two arcs 4–5 and 8–9 do not contribute for $|z| \to 0$ if

$$\mathrm{Re}\,b > -1. \tag{31}$$

The formula (13) follows immediately if we calculate the discontinuities over the cuts by making use of the formula (28) for $z' \in (-\infty, -1)$ and introducing the Legendre functions on the cut [HTF I, 3.4 (5)]:

$$\mathbb{P}^{\mu}_{\nu}(x) = e^{\frac{1}{2}i\mu\pi}P^{\mu}_{\nu}(x+i0) = e^{-\frac{1}{2}i\mu\pi}P^{\mu}_{\nu}(x-i0) \tag{32}$$

for $x \in (-1, 1)$.

Using the same method for the variable in a z^2 plane, we can write a Cauchy integral formula

$$P^{\mu}_{\nu}(z)(z^2-1)^a z^b = \frac{1}{2\pi i}\int_{C(z^2)} \frac{P^{\mu}_{\nu}(z')[(z')^2-1]^a (z')^b\, d(z')^2}{(z')^2 - z^2}. \tag{33}$$

The integration contour $C(z^2)$ is shown on Fig. 2. It is a closed curve on a 2-sheet Riemann surface z^2. The two sheets of z^2 on Fig. 2 represent the $z \to z^2$ mapping of Fig. 1, the points indicated by numbers on Fig. 2 being images of the corresponding points on Fig. 1.

Similarly, we now find the conditions under which the integrals over the arcs limit to zero and calculate the discontinuities over the cuts, thus obtaining relation (15), together with conditions (16).

Needless to say, the derivation of the formulas for $Q^{\mu}_{\nu}(z)(z^2-1)^a z^b$ in both the z and z^2 plane is quite similar to that described above. In addition to the formulas (29), (29′), and (32), one needs the formula for the asymptotic behavior of $Q^{\mu}_{\nu}(z)$ for $|z| \to \infty$, given in HTF I, 3.9 (21),

$$Q^{\mu}_{\nu}(z) \xrightarrow[|z|\to\infty]{} e^{i\mu\pi}2^{-\nu-1}\sqrt{\pi}\,\frac{\Gamma(\nu+\mu+1)}{\Gamma(\nu+\frac{3}{2})}\,z^{-\nu-1}, \tag{34}$$

and the relation [HTF I, 3.3 (12)]

$$Q^{\mu}_{\nu}(-z) = -e^{\pm i\mu\nu}Q^{\mu}_{\nu}(z), \quad \mathrm{Im}\,z \gtrless 0. \tag{35}$$

The integral formulas (11) and (17) are then simply obtained by subtracting (9) from (7) and (15) from (13) with b replaced by $(b-1)$ in (9) and (15). The conditions of validity are then the intersections of the conditions (8) and (10) or (14) and (16), respectively.

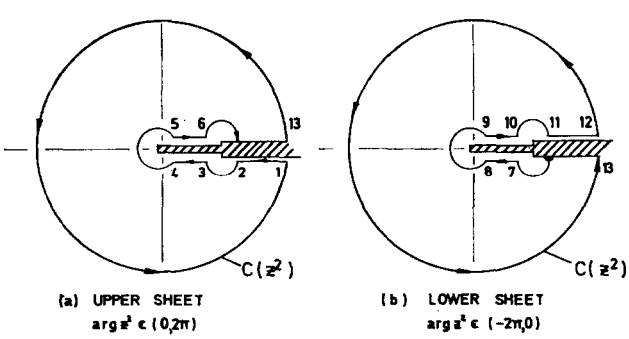


FIG. 2. The integration contour in the z^2 plane.

4. CONCLUSIONS

From the point of view of physical applications, the most immediate result is a connection between the 2-variable expansions, corresponding to the reduction $O(3,1) \supset O(2,1) \supset O(2)$ in (say) the s channel and to $O(3,1) \supset O(3) \supset O(2)$ in (say) the t channel, expressed by formulas (19) and (21).

We find it amusing that it is possible to derive new useful relations between special functions by combining Cauchy integral formulas for these functions, multiplied by certain combinations of their arguments, in different complex planes (in our case the z and z^2 planes). The idea of this method is somewhat similar to that applied in elementary particle theory when deriving superconvergence relations.

Clearly, it should be possible to derive integral relations between other special functions by applying similar techniques, which may be of some importance when considering expansions of amplitudes based on other groups than $O(3,1)$ (cf.,e.g., Refs. 13 and 14), or when generalizing the $O(3,1)$ expansions to particles with spin. In particular, it should be possible to obtain relations between the special functions which appear in the basis functions of $O(3,1)$ representations, corresponding to the reduction $O(3,1) \supset E_2 \supset O(2)$ or $O(3,1) \supset O(2) \times O(1,1)$, as well as those appearing in the corresponding D matrices of the Lorentz group.

ACKNOWLEDGMENTS

The authors thank Dr. R. J. N. Phillips, Dr. T. G. Pickavance, and Dr. G. H. Stafford for their hospitality at the Rutherford High Energy Laboratory and Dr. J. M. Charap for a discussion of related problems.

* On leave of absence from the Nuclear Research Institute of the Czechoslovak Academy of Sciences, Řež near Prague, Czechoslovakia.

[1] N. J. Vilenkin and J. A. Smorodinsky [Sov. Phys.—JETP **19**, 1209 (1964)].

[2] P. Winternitz, J. Smorodinsky, and M. Sheftel, Sov. J. Nuc. Phys. **7**, 785 (1968), and further references quoted herein.
[3] P. Winternitz, J. Smorodinsky, and M. Sheftel, Sov. J. Nucl. Phys. **8**, 489 (1968).
[4] W. Montgomery, L. O'Raifeartaigh, and P. Winternitz, Nucl. Phys. **B11**, 39 (1969).
[5] A. Sciarrino and M. Toller, J. Math. Phys. **8**, 1252 (1967).
[6] M. Toller, Nuovo Cimento **53A**, 671 (1968).
[7] R. Delbourgo, Abdus Salam, and J. Strathdee, Phys. Rev. **164**, 1981 (1967).
[8] R. J. Oakes, Phys. Letters **24B**, 154 (1967).
[9] *Higher Transcendental Functions*, A. Erdélyi, Ed. (McGraw-Hill, New York, 1953). This is referred to as HTF followed by the specific volume, chapter, and formula.
[10] P. Pajas and P. Winternitz, "Two-Variable Expansions of Relativistic Amplitudes and the Crossing Transformation," Phys. Rev. (to be published).
[11] P. Pajas, J. Math. Phys. **10**, 1777 (1969) (containing references to related work).
[12] E. J. Squires, Complex Angular Momenta in Particle Physics (W. A. Benjamin, Inc., New York, 1963).
[13] A. P. Balachandran and J. Nuyts, Phys. Rev. **172**, 1821 (1968).
[14] J. M. Charap and B. M. Minton, J. Math. Phys. **10**, 1823 (1969).

Nonexistence of Global Solutions to the Nonhomogeneous Wave Equation, Regardless of Boundary Conditions

JOHN M. BOWNDS
Department of Mathematics, University of Arizona, Tucson, Arizona 85721

(Received 12 September 1969)

The fact that the ordinary derivative y' is known to be a Darboux function implies that the right-hand side of the differential equation $y' = f(x, y)$ must also be such a function. Since it is now known that the mixed hyperbolic derivative u_{xy} is also a Darboux function, this implies that under certain mild conditions the equation $u_{xx} - u_{tt} = f$ may not have a classical global solution unless f has the proper Darboux structure. As with the ordinary differential equation, this nonexistence result does not depend on boundary conditions.

1. INTRODUCTION

In the theory of ordinary differential equations, it is known that,for the basic equation

$$y' = f(x, y) \tag{1}$$

to be well defined, it is necessary that the function $f(x, y)$ have certain properties which are known to hold for derivatives, irrespective of any boundary conditions which are imposed on the equation. More specifically, since every ordinary derivative is a Darboux function[1] [that is, a function which satisfies the following intermediate value property: $x_1 < x_2$ and $y'(x_1) < \lambda < y'(x_2)$ implies there exists a ξ with $x_1 < \xi < x_2$ such that $y_1'(\xi) = \lambda$], it must be the case that $f(x, y(x))$ also has this property for every differentiable function $y(x)$. Hence, one can say that certain ordinary differential equations with non-Darboux right-hand sides cannot have solutions at all.

As an example, consider the forced, second-order differential equation

$$y''(t) + Cy'(t) + ky(t) = F(t). \tag{2}$$

Obviously, any solution must be continuous, and since this equation is equivalent to the system

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}' = \begin{pmatrix} y_2 \\ F(t) - Cy_2 - ky_1 \end{pmatrix}, \tag{3}$$

it follows that the original equaton (2) cannot have a solution if the driving term $F(t)$ does not have the above intermediate value property. This, of course, is not to say that the equation does not have a solution in regions where $F(t)$ does have the required property.

It is quite natural to ask whether similar nonexistence results hold for partial differential equations. In particular, since the mixed partial derivative u_{xy} is known to have many properties resembling those of the ordinary derivatives,[2-8] one may expect similar behavior. Some results in recent years have established Darboux-type properties for functions of several variables and interval functions in very general settings,[9-11] and this paper will show that u_{xy}, and hence the wave operator, has a Darboux property which imposes implicit conditions on the nonhomogeneous (possibly nonlinear) wave equation. This enables the assertion of a nonexistence result which is independent of any boundary conditions.

2. THE NONHOMOGENEOUS WAVE EQUATION

Although the results in this section could be proved in greater generality, the basic statements are as follows. Consider the equation

$$\bar{u}_{\xi\xi} - \bar{u}_{\eta\eta} = f(\xi, \eta, \bar{u}, \bar{u}_\xi, \bar{u}_\eta) \tag{4}$$

defined on some region $\bar{D}$, and let

$$u_{xy} = F(x, y, u, u_x, u_y), \tag{5}$$

defined on the region D, be the transformed version of (4), under the transformation $2x = \xi + \eta$, $2y = \xi - \eta$. One recalls here that a sufficient condition for the validity of the transformation of (4) into (5) is that $\bar{u}_\xi$ and $\bar{u}_\eta$ be differentiable.[12]

Theorem 1: A necessary condition for (4) to have a solution $\bar{u}$ with the property that $\bar{u}_\xi$ and $\bar{u}_\eta$ are differentiable is that, for each rectangle R in D, with sides parallel to the axes, with diagonal points $p_1(x_1, y_1)$ and $p_2(x_2, y_2)$ such that $u_{xy}(x_1, y_1) < \lambda < u_{xy}(x_2, y_2)$, it follows that there exists a point (a, b) in R such that $u_{xy}(a, b) = \lambda$.

It is clear that this theorem could be stated without reference to any differentiability requirements using equation (5) directly. The relevance of the result to the wave operator would in that case, however, not be clear. Also, notice that the assumption of differentiability for $\bar{u}_\xi$ and $\bar{u}_\eta$ implies the differentiability and, hence, continuity of $\bar{u}$.

Proof: For each interval $[z_1, z_2]$, we define two functions. Letting $m = z_2 - z_1/2$, for $z_1 \leq z \leq z_2$, we have

$$\phi(z_1, z_2, z) = \frac{(z_2 - z_1)(z - z_1)}{(z_2 - z_1 - m)} + z_1, \quad \text{for} \quad z_1 \leq z \leq z_2 - m,$$
$$= z_2, \quad \text{for} \quad z_2 - m \leq z \leq z_2, \tag{6}$$

$$\psi(z_1, z_2, z) = z_1, \quad \text{for} \quad z_1 \leq z \leq z_1 + m,$$
$$= \frac{(z_2 - z_1)(z - z_1 - m)}{(z_2 - z_1 - m)} + z_1, \quad \text{for} \quad z_1 + m \leq z. \tag{7}$$

It follows that both functions are continuous:

$$\phi(z_1, z_2, z_1) = \psi(z_1, z_2, z_1),$$
$$\phi(z_1, z_2, z_2) = \psi(z_1, z_2, z_2),$$

and for $z_1 < z < z_2$,

$$\phi(z_1, z_2, z) > \psi(z_1, z_2, z).$$

Let R be a rectangle in D with two of its diagonal points (x_1, y_1) and (x_2, y_2) with $u_{xy}(x_1, y_1) < \lambda < u_{xy}(x_2, y_2)$. We define four functions α, β, γ, and δ as follows:

$$\alpha(x) = \phi(x_1, x_2, x), \quad x_2 > x_1,$$
$$= \phi(x_2, x_1, x), \quad x_1 > x_2; \tag{8a}$$

$$\beta(y) = \phi(y_1, y_2, y), \quad y_2 > y_1,$$
$$= \phi(y_2, y_1, y), \quad y_1 > y_2; \tag{8b}$$

$$\gamma(x) = \psi(x_1, x_2, x), \quad x_2 > x_1,$$
$$= \psi(x_2, x_1, x), \quad x_1 > x_2; \tag{8c}$$

$$\delta(y) = \psi(y_1, y_2, y), \quad y_2 > y_1,$$
$$= \psi(y_2, y_1, y), \quad y_1 > y_2. \tag{8d}$$

Next, let $\Delta(x, y)$ be defined on the interior of the rectangle R by the difference quotient

$$\Delta(x, y) = \frac{[u(\alpha(x), \beta(y)) + u(\gamma(x), \delta(y))] - [u(\alpha(x), \delta(y)) + u(\gamma(x), \beta(y))]}{[\alpha(x) - \gamma(x)][\beta(y) - \delta(y)]}. \tag{9}$$

Then, by letting

$$\Delta(x_1, y_1) = u_{xy}(x_1, y_1) \quad \text{and} \quad \Delta(x_2, y_2) = u_{xy}(x_2, y_2), \tag{10}$$

$\Delta(x, y)$ becomes a continuous function on the interior of the rectangle together with the points (x_1, y_1) and (x_2, y_2). Hence, using the intermediate value property for continuous functions, we see that there exists a point $(\bar{x}, \bar{y})$ in R such that $\Delta(\bar{x}, \bar{y}) = \lambda$. But, if we use the mean-value theorem for the operator u_{xy}, this implies, from (9), that there exists a point (a, b) with $\alpha(\bar{x}) < a < \gamma(\bar{x})$ and $\beta(\bar{y}) < b < \delta(\bar{y})$ such that $\Delta(\bar{x}, \bar{y}) = u_{xy}(a, b)$. Hence, $u_{xy}(a, b) = \lambda$, and (a, b) is in R.

As an example of the applicability of this theorem, consider the equation

$$\bar{u}_{\xi\xi} - \bar{u}_{\eta\eta} = f(\xi)f(\eta), \tag{11}$$

where

$$f(x) = 1, \quad 0 \leq x \leq 1,$$
$$= 0, \quad \text{otherwise.}$$

Then,

$$\bar{u}_{\xi\xi} - \bar{u}_{\eta\eta} = 1, \quad \text{for} \quad 0 \leq \xi \leq 1 \quad \text{and} \quad 0 \leq \eta \leq 1,$$
$$= 0, \quad \text{otherwise,}$$

and so,

$$u_{xy} = 1, \quad \text{for } (x, y) \text{ in } D,$$
$$= 0, \quad \text{otherwise,} \tag{12}$$

where D is a diamond-shaped region with vertices $(0, 0)$, $(\frac{1}{2}, \frac{1}{2})$, $(1, 0)$, and $(\frac{1}{2}, -\frac{1}{2})$. According to the theorem, if (11) has a continuous solution with the required differentiability, then u_{xy}, given by (12), must attain intermediate values on rectangles. But $u_{xy}(-1, -1) = 0 < \frac{1}{2} < u_{xy}(0, 0) = 1$, and there is no point (a, b) such that $u_{xy}(a, b) = \frac{1}{2}$, owing to the nature of the right-hand side of (11). Hence, (11) can have no such global solution.

[1] W. Rudin, *Principles of Mathematical Analysis* (McGraw-Hill Book Company, New York, 1964), p. 93.
[2] J. B. Diaz, Arch. Ratl. Mech. Anal. **1**, 357 (1958).
[3] J. B. Diaz, Annali di Mat. Pura ed Appl., "Serie IV," **52**, 163 (1960).
[4] P. Hartman and A. Wintner, Am. J. Math. **74**, 834 (1952).
[5] G. Zwirner, Rend. Sem. Mat. Univ. Padova, **20**, 329 (1951).
[6] J. B. Diaz and W. Walter, Trans. Am. Math. Soc. **96**, 90 (1960).
[7] W. Walter, *Differential- und Integral-Ungleichungen* (Springer-Verlag, Berlin, 1964).
[8] S. Marcus, Rev. Roum. Math. Pures Appl. **9**, 767 (1964).
[9] C. J. Neugebauer, Trans. Am. Math. Soc. **107**, 30 (1963).
[10] A. M. Bruckner and J. B. Bruckner, Trans. Am. Math. Soc. **128**, 103 (1967).
[11] L. Mišik, Casopis Pest. Mat. **91**, 394 (1966).
[12] A. Taylor, *Advanced Calculus* (Ginn and Company, Boston, Mass., 1955).

Lattice Dynamics of Cubic Lattices with Long-Range Interactions

JOHN A. DAVIES
Department of Physics, Clark University, Worcester, Massachusetts 01610

(Received 25 July 1969)

An analytic study is made of the dispersion relations and frequency spectra of unbounded cubic lattices in which there exist long-range pair potentials of the form r^{-p}, where r is the distance between particles. This paper is an extension of work by Davies and Yedinak in which simple cubic lattices are studied. Cubic lattices having two particles per unit cell are now considered. The analytic behavior of the optical branches near the origin of the first Brillouin zone is determined for the case of $1 \leq p \leq 3$ and the resulting contribution of the longitudinal optical branch to the frequency spectrum is obtained. A special case of such lattices is Kellermann's model for NaCl. This model is studied in detail.

I. INTRODUCTION

In an earlier paper,[1] we present a study of the analytic properties of the vibrational dispersion relations and frequency spectra of simple cubic lattices in which long-range interactions are present. In this paper, we extend our work to cubic lattices having two particles per cell. We rely heavily upon references to the earlier paper (Ref. 1).

We consider an infinite cubic lattice having either the sodium chloride or caesium chloride structure. Its particles are assumed to interact through long-range pair potentials falling off with distance r as r^{-p}, where $p \geq 1$. In addition, the particles are assumed to interact through short-range pair potentials which extend to a finite number of neighbors only.

In Sec. II, we show that the elements of the dynamical matrix for such lattices are analytic functions of the propagation vector everywhere in the first Brillouin zone except at the origin. We then determine the behavior of the dynamical matrix near the origin. In Sec. III, we determine the behavior of the optical modes of the dispersion relations near the origin, provided $p > 1$. The analysis is extended to the case of $p = 1$ (the Coulomb interaction) in Sec. IV. The contribution of the region near the origin, to the frequency spectrum when $p > 1$, is studied in Sec. V. In Sec. VI, we extend our analysis to the case of $p = 1$ for a specific model for sodium chloride.

II. THE MODEL

Consider an infinite diatomic lattice having either the sodium chloride or the caesium chloride structure. Let $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ be the primitive translation vectors. The equilibrium position of the kth particle in the lth cell is given by

$$\mathbf{x}(lk) = \sum_{i=1}^{3} l_i a_i + \mathbf{x}(k) = \mathbf{x}(l) + \mathbf{x}(k), \tag{1}$$

where $k = 0$ or 1 and $\mathbf{x}(k) = 0$, when $k = 0$. Cartesian coordinates are imbedded in the lattice with the origin at $\mathbf{x}(00)$ and with axes parallel to the edges of the unit cubes. The primitive translation vectors of the reciprocal lattice are $\mathbf{b}_1$, $\mathbf{b}_2$, and $\mathbf{b}_3$, and a reciprocal lattice vector is denoted by $\mathbf{y}(h) = \sum_i h_i \mathbf{b}_i$, where h_i is an integer. We employ the following notation. Let a be some characteristic dimension of a unit cell in the direct lattice. Then, we define

$$\boldsymbol{\eta}(lk) = \mathbf{x}(lk)/a = \boldsymbol{\eta}(l) + \boldsymbol{\eta}(k)$$

and

$$\boldsymbol{\xi}(h) = a\mathbf{y}(h). \tag{2}$$

It is assumed that the particles of the lattice interact through long- and short-range central pair potentials. The long-range potentials fall off with distance r as r^{-p}, where $p \geq 1$. The short-range potentials are assumed to vanish identically after a finite range. Explicitly, the potential energy of the k_0th particle in the zeroth cell is given by

$$V_{k_0} = V_{k_0}^L + V_{k_0}^S, \tag{3}$$

where $V_{k_0}^S$ is the contribution of the short-range potentials and

$$V_{k_0}^L = G \sum_{\substack{l,k \\ l,k \neq 0,k_0}} (-1)^{k+k_0} \times \left\{ \sum_i [x_i(lk) + u_i(lk) - x_i(0k_0) - u_i(0k_0)]^2 \right\}^{-p/2}. \tag{4}$$

Here G is a constant and $x_i(lk) + u_i(lk)$ is the ith Cartesian component of the position of the particle with equilibrium position $\mathbf{x}(lk)$.

In the harmonic approximation, the equations of motion for the particles in the zeroth cell are given by

$$m_{k_0} \ddot{u}(0k_0) = -\sum_{l,k,j} \frac{\partial^2 V_{k_0}}{\partial u_i(0k_0) \partial u_j(lk)}\Big)_0 u_j(lk), \tag{5}$$

where m_{k_0} is the mass of the k_0th particle. We choose

normal mode solutions of the form

$$u_j(lk) = m_k^{-\frac{1}{2}} W_{kj} \exp(2\pi i \boldsymbol{\phi} \cdot \boldsymbol{\eta}(lk) - i\omega(\boldsymbol{\phi})t). \quad (6)$$

Substitution of Eq. (6) into Eq. (5) gives us an eigenvector equation of the form

$$\sum_{k,j} D_\phi(k_0 i, kj) W_{kj} = \lambda(\boldsymbol{\phi}) W_{k_0 i}, \quad (7)$$

where $\lambda = a^{p+2} m_0 \omega^2 / G$. We regard the W_{kj} as elements of a 6-dimensional column matrix $\mathbf{W}$ and the $\mathbf{D}_\phi(k_0 i, kj)$ as elements of a 6×6 dynamical matrix $\mathbf{D}_\phi$. We can write

$$\mathbf{D}_\phi = \mathbf{D}_\phi^L + \mathbf{D}_\phi^S, \quad (8)$$

where $\mathbf{D}_\phi^L$ and $\mathbf{D}_\phi^S$ are the contributions of the long- and short-range interactions, respectively. Since the short-range interactions vanish after a finite range, we assume that the elements of $\mathbf{D}_\phi^S$ are analytic functions of $\boldsymbol{\phi}$ everywhere in the reciprocal space. The secular equation giving the dispersion relations $\lambda(\boldsymbol{\phi})$ for the lattice is

$$|\mathbf{D}_\phi - \lambda(\boldsymbol{\phi})\mathbf{I}| = 0. \quad (9)$$

There are six branches of the dispersion relations, three acoustic and three optical.

Explicit expressions for the elements of $\mathbf{D}_\phi^L$ are given in Appendix A. In order to study the analytic properties of these elements, we apply the Ewald transformation[2,3] to the sums appearing in Eqs. (A1) and (A3). For lattices of either the sodium chloride or caesium chloride structure, the transformed relations have the form given in Eqs. (B1)–(B3) of Appendix B. Referring to Eqs. (B1)–(B3) and using arguments similar to those used in Sec. II of Ref. 1, we see that the elements of $\mathbf{D}_\phi^L$ (and thus of $\mathbf{D}_\phi$) are analytic functions of $\boldsymbol{\phi}$, unless $\boldsymbol{\phi} = \boldsymbol{\xi}(h)$ for any reciprocal lattice point h. Thus, the elements of $\mathbf{D}_\phi$ are analytic everywhere in the first Brillouin zone except at $\boldsymbol{\phi} = 0$, which is a critical point. Again, using methods discussed in Sec. II of Ref. 1, we easily obtain expansions of the elements of $\mathbf{D}_\phi^L$ about $\boldsymbol{\phi} = 0$. These are given in Eqs. (C1)–(C6) of Appendix C.

We are interested in those values of p for which the nonanalytic terms appearing in Eqs. (C1)–(C6) dominate the second-order terms near $\boldsymbol{\phi} = 0$. Under these circumstances, the critical point at $\boldsymbol{\phi} = 0$ may not produce the usual square root infinite slopes in the frequency spectrum, which one expects to occur when only short-range interactions are present. Referring to Eqs. (C1)–(C6), we see that the following three cases hold true:

(a) If $p > 3$, then the second-order terms dominate the nonanalytic terms.

(b) If $1 < p \leq 3$, then the nonanalytic terms dominate the second-order terms and the behavior of $\mathbf{D}_\phi(k_0 i, kj)$ near $\boldsymbol{\phi} = 0$ is of the form

$$D_\phi^L(k_0 i, kj) = f(k_0, k)(\alpha \delta_{ij} + h(\phi)\phi_i \phi_j) + O(\phi^2), \quad (10)$$

where $f(0, 0) = 1$, $f(1, 1) = m_0/m_1$, and $f(k_0, k) = -(m_0/m_1)^{\frac{1}{2}}$, if $k_0 \neq k$. The term α is a constant. If $p = 3$, then $h(\phi)$ is defined by

$$h(\phi) = -8\pi^3 (a^3/v_a) \ln(\phi^2). \quad (11)$$

If $1 < p < 3$, then $h(\phi)$ is given by

$$h(\phi) = 2p\pi^{p+\frac{1}{2}} \left(\frac{a^3}{v_a}\right) \frac{\Gamma(\frac{3}{2} - \frac{1}{2}p)}{\Gamma(\frac{1}{2}p + 1)} \phi^{p-3}. \quad (12)$$

In the above equations, v_a is the volume of a unit cell in the direct lattice.

(c) If $p = 1$, then we have

$$h(\phi) = 4\pi (a^3/v_a) \phi^{-2}. \quad (13)$$

The nonanalytic terms dominate the second-order terms but the values the nonanalytic terms approach as $\boldsymbol{\phi}$ approaches zero depend upon the direction of approach.

The above results are similar to those previously found for the corresponding case with simple cubic lattices.[1] In the following sections, we study the behavior of the dispersion relations near $\boldsymbol{\phi} = 0$ and the resulting behavior of the vibrational frequency spectrum of the lattice. Our analysis is limited to the case of $1 \leq p \leq 3$. (The above analysis shows that the long-range interaction should produce no unusual behavior in the dispersion relations, when $p > 3$.) Because the case of $p = 1$ presents mathematical difficulties, we first assume that $1 < p \leq 3$ and then study the behavior of our results as p approaches one.

III. THE DISPERSION RELATIONS

Initially we assume that $1 < p \leq 3$. Then, using Eqs. (C1)–(C6) of Appendix C and taking the cubic symmetry of the lattice into account, we find that the elements of $\mathbf{D}_\phi$ have the following form near $\boldsymbol{\phi} = 0$:

$$D_\phi(01, 11) = -\mu^{\frac{1}{2}}[\alpha + h(\phi)\phi_1^2 + \beta\phi_1^2 + \gamma(\phi_2^2 + \phi_3^2)] + O(\phi^4), \quad (14)$$

$$D_\phi(01, 12) = -\mu^{\frac{1}{2}}(h(\phi)\phi_1\phi_2 + \delta\phi_1\phi_2) + O(\phi^4), \quad (15)$$

$$D_\phi(01, 01) = \alpha + h(\phi)\phi_1^2 + b\phi_1^2 + c(\phi_2^2 + \phi_3^2) + O(\phi^4), \quad (16)$$

$$D_\phi(01, 02) = h(\phi)\phi_1\phi_2 + d\phi_1\phi_2 + O(\phi^4), \quad (17)$$

where $\mu = m_0/m_1$. The constants α, β, γ, δ, b, c, and d are sums of long- and short-range contributions. The long-range contributions are given in Appendix C. The remaining elements of $\mathbf{D}_\phi$ are easily obtained from the cubic symmetry and from the relations

$$D_\phi(k_0 i, kj) = D_\phi(ki, k_0 j) = D_\phi(k_0 j, ki) \quad (18)$$

and

$$m_{k_0} D_\phi(k_0 i, k_0 j) = m_k D_\phi(ki, kj). \quad (19)$$

In order to determine the behavior of $\lambda(\boldsymbol{\phi})$ near $\boldsymbol{\phi} = 0$, we solve Eq. (7) using perturbation theory. We regard $\mathbf{D}_\phi$ as the sum of an unperturbed matrix $\mathbf{D}_\phi^{(0)}$ and a perturbation $\mathbf{D}_\phi^{(p)}$. The matrix $\mathbf{D}_\phi^{(0)}$ is given by

$$\mathbf{D}_\phi^{(0)} = \begin{pmatrix} \mathbf{M} & -\mu^{\frac{1}{2}}\mathbf{M} \\ -\mu^{\frac{1}{2}}\mathbf{M} & \mu\mathbf{M} \end{pmatrix}, \quad (20)$$

where $\mathbf{M}$ is a 3×3 submatrix, defined by

$$\mathbf{M} = \begin{pmatrix} \alpha + h(\phi)\phi_1^2 & h(\phi)\phi_1\phi_2 & h(\phi)\phi_1\phi_3 \\ h(\phi)\phi_1\phi_2 & \alpha + h(\phi)\phi_2^2 & h(\phi)\phi_2\phi_3 \\ h(\phi)\phi_1\phi_3 & h(\phi)\phi_2\phi_3 & \alpha + h(\phi)\phi_3^2 \end{pmatrix}. \quad (21)$$

The perturbation is given by

$$\mathbf{D}_\phi^{(p)} = \begin{pmatrix} \mathbf{N}(b, c, d) & -\mu^{\frac{1}{2}}\mathbf{N}(\beta, \gamma, \delta) \\ -\mu^{\frac{1}{2}}\mathbf{N}(\beta, \gamma, \delta) & \mu\mathbf{N}(b, c, d) \end{pmatrix} + \mathbf{O}(\phi^4), \quad (22)$$

where $\mathbf{N}(x, y, z)$ is defined by

$$\mathbf{N}(x, y, z) = \begin{pmatrix} x\phi_1^2 + y(\phi_2^2 + \phi_3^2) & z\phi_1\phi_2 & z\phi_1\phi_3 \\ z\phi_1\phi_2 & x\phi_2^2 + y(\phi_1^2 + \phi_3^2) & z\phi_2\phi_3 \\ z\phi_1\phi_3 & z\phi_2\phi_3 & x\phi_3^2 + y(\phi_1^2 + \phi_2^2) \end{pmatrix}. \quad (23)$$

First-order perturbation theory is applied in the usual manner.[4] First, the eigenvalues of $\mathbf{D}_\phi^{(0)}$ are determined. For each such eigenvalue $\lambda^{(0)}$ (which may be degenerate), we form a complete set of orthonormal eigenvectors. The matrix elements of $\mathbf{D}_\phi^{(p)}$, with respect to this basis, are calculated and the eigenvalues of the resulting matrix determined. Each such eigenvalue, when added to $\lambda^{(0)}$, gives the behavior of λ near $\boldsymbol{\phi} = 0$ for one branch of the dispersion relations.

The eigenvalues of $\mathbf{D}_\phi^{(0)}$ are determined by solving

$$|\mathbf{D}_\phi^{(0)} - \lambda^{(0)}\mathbf{I}| = 0 \quad (24)$$

for $\lambda^{(0)}$. Although the determinant is 6×6, the roots are easily calculated after a sequence of row and column reductions is performed. The following eigenvalues are obtained:

$$\lambda^{(0)} = (1 + \mu)[\alpha + h(\phi)\phi^2], \quad \text{one eigenvalue,} \quad (25)$$

$$\lambda^{(0)} = (1 + \mu)\alpha, \quad \text{two eigenvalues,} \quad (26)$$

$$\lambda^{(0)} = 0, \quad \text{three eigenvalues.} \quad (27)$$

Clearly, the first three eigenvalues correspond to optical modes and the last three to acoustic modes.

The dispersion relations for the optical branches near $\boldsymbol{\phi} = 0$ are easily determined. A normalized eigenvector corresponding to the eigenvalue given in Eq. (25) is

$$\mathbf{W}_1 = [(1 + \mu)\phi^2]^{-\frac{1}{2}} \begin{pmatrix} \boldsymbol{\psi}_1 \\ -\mu^{\frac{1}{2}}\boldsymbol{\psi}_1 \end{pmatrix}, \quad (28)$$

where

$$\boldsymbol{\psi}_1 = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{pmatrix}. \quad (29)$$

Applying perturbation theory as outlined above, we find that the behavior of the corresponding optical branch is given by

$$\lambda_1(\boldsymbol{\phi}) = (1 + \mu)[\alpha + h(\phi)\phi^2] + (1 + \mu)^{-1}\phi^{-2} \times [g(b, \beta)\langle 4\rangle + 2g(c + d, \gamma + \delta)\langle 2, 2\rangle] + O[\phi^2/h(\phi)], \quad (30)$$

where $g(x, y) = (1 + \mu^2)x + 2\mu y$. The bracketed terms ($\langle\ \rangle$) are polynomials in the components of $\boldsymbol{\phi}$, which are defined in Appendix D.

Two (nonunique) orthonormal eigenvectors corresponding to the doubly degenerate eigenvalue given in Eq. (26) are

$$\mathbf{W}_2 = [(1 + \mu)(\phi_2^2 + \phi_3^2)]^{-\frac{1}{2}} \begin{pmatrix} \boldsymbol{\psi}_2 \\ -\mu^{\frac{1}{2}}\boldsymbol{\psi}_2 \end{pmatrix} \quad (31)$$

and

$$\mathbf{W}_2' = [(1 + \mu)(\phi_2^2 + \phi_3^2)\phi^2]^{-\frac{1}{2}} \begin{pmatrix} \boldsymbol{\psi}_2' \\ -\mu^{\frac{1}{2}}\boldsymbol{\psi}_2' \end{pmatrix}, \quad (32)$$

where

$$\boldsymbol{\psi}_2 = \begin{pmatrix} 0 \\ \phi_3 \\ -\phi_2 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\psi}_2' = \begin{pmatrix} -(\phi_2^2 + \phi_3^2) \\ \phi_1\phi_2 \\ \phi_1\phi_3 \end{pmatrix}. \quad (33)$$

The behavior of the two corresponding optical branches is then given by

$$\left.\begin{matrix} \lambda_2 \\ \lambda_3 \end{matrix}\right\} = (1 + \mu)\alpha + [(1 + \mu)\phi^2]^{-1} \times \{g(c, \gamma)\langle 4\rangle + g(b - d + c, \beta - \delta + \gamma) \times \langle 2, 2\rangle \pm |g(b - d - c, \beta - \delta - \gamma)| \times (\langle 4, 4\rangle - \langle 4, 2, 2\rangle)^{\frac{1}{2}}\} + O[\phi^2/h(\phi)]. \quad (34)$$

In directions of high symmetry, one of the optical modes must be pure longitudinal and the remaining two must be transverse. Referring to Eqs. (6), (7), and (28), we see that to zero order the λ_1 mode is longitudinal in any direction near $\boldsymbol{\phi} = 0$ and should be identified as the longitudinal branch. The zero-order eigenvectors for the second and third branches (λ_2 and λ_3) are some linear combinations of $\mathbf{W}_2$ and $\mathbf{W}_2'$. Referring to Eqs. (32) and (33), we see that $\mathbf{W}_2$ and $\mathbf{W}_2'$ separately represent transverse vibrations. Thus, we identify λ_2 and λ_3 as the transverse modes.

Notice that $h(\phi)$ does not contribute to the behavior of λ_2 and λ_3 near $\boldsymbol{\phi} = 0$. However, it is the leading term in the expression for λ_1 near $\boldsymbol{\phi} = 0$, for all p such that $1 < p \leq 3$. These results very closely resemble those found in work on simple cubic lattices.[1]

We have also carried out the perturbation calculation for the acoustic modes using Eq. (27). In this case, $\lambda^{(0)}$ is three-fold degenerate, and we found it necessary to solve a cubic equation in order to determine the corrections given by first-order perturbation theory. We were unable to express our final results in any simple, neat form, and as a result we do not include them in this paper. However, one important result of the calculation was that the function $h(\phi)$ does not contribute to the leading behavior of any of the acoustic branches near $\boldsymbol{\phi} = 0$. That this is true can easily be seen. A (nonunique) orthonormal set of eigenvectors of $\mathbf{D}_\phi^{(0)}$ corresponding to the eigenvalue $\lambda^{(0)} = 0$ is given by

$$W_3 = (1+\mu)^{-\frac{1}{2}} \begin{pmatrix} \sqrt{\mu} \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix},$$

$$W_3' = (1+\mu)^{-\frac{1}{2}} \begin{pmatrix} 0 \\ \sqrt{\mu} \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix},$$

$$W_3'' = (1+\mu)^{-\frac{1}{2}} \begin{pmatrix} 0 \\ 0 \\ \sqrt{\mu} \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

None of these vectors contains the function $h(\phi)$. Further, $\lambda^{(0)}$ and $\mathbf{D}_\phi^{(p)}$ are independent of $h(\phi)$. Thus, any corrections to $\lambda^{(0)}$ given by first-order perturbation theory are also independent of $h(\phi)$, and the leading terms for the acoustic branches near $\boldsymbol{\phi} = 0$ should be independent of $h(\phi)$. Therefore, we do not expect the presence of the long-range interaction to produce any unusual behavior in the acoustic branches.

IV. DISPERSION RELATIONS WHEN $p = 1$

We now consider the behavior of the optical branches for the case of $p = 1$ (the Coulomb interaction). The behavior of the second and third branches for $p > 1$ is given by Eq. (34). The form of this equation does not change as $p \to 1$. Referring to Eq. (13), we see that the remainder $O[\phi^2/h(\phi)] \to O(\phi^4)$ as $p \to 1$. Further, a study of Appendix C shows that the constants appearing in Eq. (34) remain well defined when $p = 1$. Thus, the second and third branches obey Eq. (34) even when the long-range interaction is the Coulomb interaction.

Next, we consider the first optical branch. From Eq. (30), we see that the leading term in λ_1 near $\boldsymbol{\phi} = 0$ is given by

$$\lambda_1(\boldsymbol{\phi}) = (1+\mu)\alpha + 2(1+\mu)a^3 p \pi^{p+\frac{1}{2}} \times \Gamma(\tfrac{3}{2} - \tfrac{1}{2}p)\phi^{p-1}[v_a \Gamma(\tfrac{1}{2}p + 1)]^{-1} + O(\phi^2). \quad (35)$$

The above equation is similar to Eq. (27) of Ref. 1. Using arguments similar to those used in Sec. III of Ref. 1, we see that the value approached by λ_1, as both $\boldsymbol{\phi} \to 0$ and $p \to 1$, is not well defined. However, consider any neighborhood of $\boldsymbol{\phi} = 0$ sufficiently small that the remainder $O(\phi^2)$ in Eq. (35) may be neglected. As $p \to 1$, $\lambda_1 \to (1+\mu)\alpha + 4(1+\mu)\pi a^3/v_a$ unambiguously at any point in this neighborhood except at $\boldsymbol{\phi} = 0$ itself. Referring to Eq. (30), we see that, when $p = 1$, the behavior of λ_1 near (but not exactly at) $\boldsymbol{\phi} = 0$ is given by

$$\lambda_1 = (1+\mu)(\alpha + 4\pi a^3/v_a) + (1+\mu)^{-1}\phi^{-2} \times [g(b, \beta)\langle 4\rangle + 2g(c+d, \gamma+\delta)\langle 2, 2\rangle] + O(\phi^4), \quad (36)$$

where $g(x, y)$ is defined after Eq. (30).

Notice that a three-fold degeneracy between the optical branches no longer occurs at $\boldsymbol{\phi} = 0$, in the sense that λ_1 does not approach the same value as λ_2 and λ_3, when $\boldsymbol{\phi} \to 0$.

V. DISCUSSION OF THE FREQUENCY SPECTRUM WHEN $1 < p \leq 3$

Let $G_1(\lambda)$ be the contribution of the longitudinal optical branch to the vibrational-frequency spectrum

of the lattice. Here $G_1(\lambda)$ is normalized such that

$$\int_0^\infty G_1(\lambda)\,d\lambda = 1. \tag{37}$$

We consider the contribution $\Delta G_1(\lambda)$ of the region near $\phi = 0$ to $G_1(\lambda)$ when λ is near to $(1+\mu)\alpha$. Our discussion in this section is limited to the case of $1 < p \leq 3$. The contribution of the region about $\phi = 0$ to $G_1(\lambda)$, when $p = 1$, depends upon the particular model chosen for the lattice and is discussed in the following section.

The function $\lambda_1(\phi)$ is a minimum at $\phi = 0$ (if $1 < p \leq 3$). Thus, we find $\Delta G_1(\lambda) = 0$, when $\lambda < (1+\mu)\alpha$. The calculation of $\Delta G_1(\lambda)$, for $\lambda > (1+\mu)\alpha$ and $\lambda \approx (1+\mu)\alpha$, is similar to the calculation discussed in Sec. IV of Ref. 1. Therefore, we simply list the results of the calculation below:

(a) If $p = 3$, then

$$\Delta G_1(\lambda) = 2\pi\left(\frac{v_a}{a^3}\right)\frac{K\{K[\lambda-(1+\mu)\alpha]\}^{\frac{1}{2}}}{(-\ln\{K[\lambda-(1+\mu)\alpha]\})^{\frac{3}{2}}}, \tag{38}$$

where $K = v_a[8\pi^3 a^3(1+\mu)]^{-1}$. The leading term in the derivative of this function, near $\lambda = (1+\mu)\alpha$, is given by

$$\frac{d\Delta G_1(\lambda)}{d\lambda} = \left(\frac{\pi v_a K^2}{a^3}\right)\frac{\{K[\lambda-(1+\mu)\alpha]\}^{-\frac{1}{2}}}{(-\ln\{K[\lambda-(1+\mu)\alpha]\})^{\frac{3}{2}}}. \tag{39}$$

(b) If $1 < p < 3$, then

$$\Delta G_1(\lambda) = \left(\frac{4\pi}{(p-1)}\right)\left(\frac{v_a}{a^3}\right) \times \epsilon(p)[\lambda-(1+\mu)\alpha]^{(4-p)/(p-1)}, \tag{40}$$

where

$$\epsilon(p) = \left(\frac{v_a\Gamma(\frac{1}{2}p+1)}{2p\pi^{p+\frac{1}{2}}a^3(1+\mu)\Gamma(\frac{3}{2}-\frac{1}{2}p)}\right)^{3/(p-1)}$$

The derivative is given by

$$\frac{d\Delta G_1(\lambda)}{d\lambda} = \left(\frac{4\pi(4-p)}{(p-1)^2}\right)\left(\frac{v_a}{a^3}\right)\epsilon(p) \times [\lambda-(1+\mu)\alpha]^{(5-2p)/(p-1)}. \tag{41}$$

The above behavior of $\Delta G_1(\lambda)$ and its derivative, near $\lambda = (1+\mu)\alpha$, is similar to that found for the corresponding case in simple cubic lattices.[1] If no long-range interactions were present, we would expect to find an inverse square root infinity in the slope of $\Delta G_1(\lambda)$ as we approach $\lambda = (1+\mu)\alpha$ from above. Here, we find that when $p = 3$ the infinity in the slope is slightly weaker than an inverse square root. In the case of $1 < p < 3$, the derivative of $G_1(\lambda)$ behaves as $[\lambda-(1+\mu)\alpha]^{(5-2p)/(p-1)}$. If $p > \frac{5}{2}$, the derivative approaches infinity as $\lambda \to (1+\mu)\alpha$ from above, but more weakly than as an inverse square root. On the other hand, if $p < \frac{5}{2}$, the derivative approaches zero.

When $p = 1$, we do not expect the region near $\phi = 0$ to contribute to the frequency spectrum near $\lambda = (1+\mu)\alpha$. Instead, this region should contribute to the frequency spectrum near $\lambda = (1+\mu)\alpha + 4(1+\mu)\pi a^3/v_a$. It is easily shown that, as given in Eq. (40), $\Delta G_1(\lambda) \to 0$ as $p \to 1$.

The contribution of the transverse optical branches to the frequency spectrum is not discussed in this paper. However, arguments similar to those given in Sec. IV of Ref. 1 indicate that each of these branches contributes an inverse square root infinity to the slope of the frequency spectrum at $\lambda = (1+\mu)\alpha$.

VI. THE COULOMB INTERACTION

We choose a specific crystal model for the case of $p = 1$ (the Coulomb interaction). The model chosen is Kellermann's model for sodium chloride.[5] Kellermann assumed that, in addition to the long-range Coulomb interaction, a short-range repulsive potential existed in the lattice which extended between nearest neighbors only. The contribution of the short-range potential to the dynamical matrix was determined from the condition that the lattice be in equilibrium and by assuming a compressibility of 4.16×10^{-12} cm^2/dyn for sodium chloride. The above is equivalent to assuming the existence of a short-range potential energy between nearest neighbors of the form H/r^q with $q = 7.738$ and $H/G = 0.0376 r_0^{q-1}$, where r_0 is the nearest-neighbor distance, assumed by Kellermann to be 2.814×10^{-8} cm. [G is defined in Eq. (4) and is now the electronic charge squared.]

The constants appearing in Eqs. (34) and (36) are easily computed. Equations analogous to those appearing in Appendix A exist for the short-range contributions to $\mathbf{D}_\phi$. From these equations, the short-range contributions to the constants are easily obtained. The long-range contributions are determined from Eqs. (C7)–(C24). Our results are given in Table I. To obtain these values, we chose $a = \sqrt{2}(r_0)$, where a is defined by Eq. (2).

The behavior of the dispersion relations for the optical branches near $\phi = 0$ are then given by

$$\lambda_1 = 37.834 - (104.820\langle 4\rangle + 101.398\langle 2,2\rangle)\phi^{-2} + O(\phi^4) \tag{42}$$

TABLE I. Constants for Kellermann's model.

Short-range contribution	Long-range contribution	Total
$\alpha^S = 11.101$	$\alpha^L = -5.924$	$\alpha = 5.178$
$\beta^S = -142.089$	$\beta^L = 36.718$	$\beta = -105.371$
$\gamma^S = 16.262$	$\gamma^L = -22.440$	$\gamma = -6.178$
$\delta^S = 0.000$	$\delta^L = -55.020$	$\delta = -55.020$
$b^S = 0.000$	$b^L = -21.909$	$b = -21.909$
$c^S = 0.000$	$c^L = 13.236$	$c = 13.236$
$d^S = 0.000$	$d^L = -16.190$	$d = -16.190$

and

$$\left.\begin{matrix}\lambda_2\\ \lambda_3\end{matrix}\right\} = 8.536 + [6.545\langle 4\rangle - 38.002\langle 2,2\rangle \pm 51.092(\langle 4,4\rangle - \langle 4,2,2\rangle)^{\frac{1}{2}}]\phi^{-2} + O(\phi^4). \quad (43)$$

In Figs. 1–3, we compare our expansions to second order with Kellermann's computation of the dispersion relations in various directions. Here we have converted from Kellermann's values for the frequency into corresponding values of λ.

The behavior of λ_1 in the region near $\phi = 0$ is similar to that found for the corresponding case in simple cubic lattices.[1] When $p = 1$, a fluted maximum occurs at $\phi = 0$ instead of the minimum which was previously found when $3 \geq p > 1$. The contribution ΔG_1 of the region near $\phi = 0$ to the frequency spectrum for the longitudinal branch is determined by a calculation similar to that discussed in Sec. VI of Ref. 1. We first rewrite the leading terms in Eq. (42)

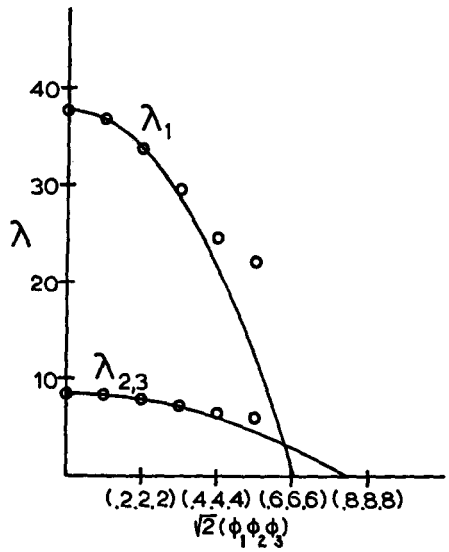


FIG. 2. Comparison of the dispersion relations calculated by Kellermann (circles) with our second-order expansions (solid lines) in the direction in reciprocal space indicated on the horizontal axis.

in terms of spherical coordinates:

$$\lambda_1 = 37.834 - 104.820\phi^2 + 108.242\phi^2 \sin^2\theta \times \{1 - (\tfrac{1}{8})[7 + \cos(4\zeta)] \sin^2\theta\}. \quad (44)$$

Here θ is the polar angle and ζ is the azimuthal angle. Using Eq. (44), one can easily show that the contribution of the region about $\phi = 0$ to the frequency spectrum, when $\lambda_1 \approx 37.834$, is given by

$$\Delta G_1(\lambda) = 0, \quad \lambda > 37.834,$$

and

$$\Delta G_1(\lambda) = 2^{-\frac{1}{2}}Q(37.834 - \lambda)^{\frac{1}{2}}, \quad \lambda < 37.834. \quad (45)$$

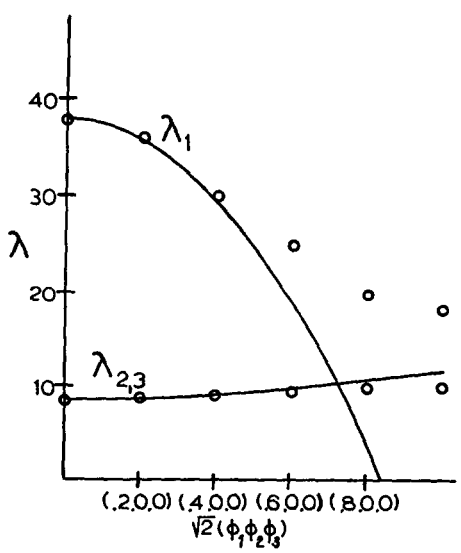


FIG. 1. Comparison of the dispersion relations calculated by Kellermann (circles) with our second-order expansions (solid lines) in the direction in reciprocal space indicated on the horizontal axis.

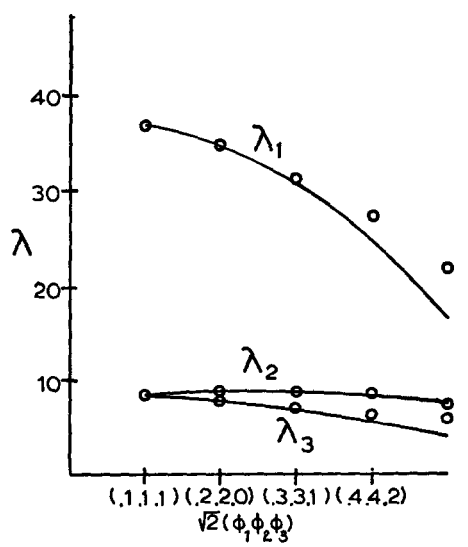


FIG. 3. Comparison of the dispersion relations calculated by Kellermann (circles) with our second-order expansions (solid lines) in the direction in reciprocal space indicated on the horizontal axis.

In the above equation, the term Q is a constant defined by

$$Q = \tfrac{1}{2}\int_0^{2\pi}\int_0^{\pi}\{104.820 - 108.242\sin^2\theta \times [1 - \tfrac{1}{8}(7 + \cos 4\zeta)\sin^2\theta]\}^{-\frac{3}{2}}\sin\theta\, d\theta\, d\zeta. \quad (46)$$

We have not evaluated this integral, but one can easily show that $0.00586 < Q < 0.01103$. From Eq. (45), it is clear that the maximum in λ_1 at $\boldsymbol{\phi} = 0$ produces a square root infinity in the slope of the frequency spectrum at $\lambda_1 = 37.834$. Thus, as we found earlier for simple cubic lattices,[1] the analytic behavior of the frequency spectrum for the longitudinal branch, when $p = 1$, is quite different from its behavior when $1 < p < 3$.

VII. CONCLUDING REMARKS

Two important limitations of the above calculation should be pointed out. First, we have restricted our calculation to a point-ion type of model. Although we suspect that a similar analysis could be carried out for the shell model,[6] we have not attempted to do so. Second, it must be noted that we work with unbounded lattices. Such results as the lack of degeneracy between the longitudinal and transverse optical branches at $\boldsymbol{\phi} = 0$ when $p = 1$ are not necessarily true for finite lattices.[7]

APPENDIX A

Using Eqs. (4)–(7), we obtain the following expressions for the elements $D_\phi^L(k_0 j, ks)$ of $\mathbf{D}_\phi^L$. If $k \neq k_0$, then

$$\mathbf{D}_\phi^L(k_0 j, ks) = -\left(\frac{m_0}{m_k}\times\frac{m_0}{m_{k_0}}\right)^{\frac{1}{2}}\exp\{2\pi i\boldsymbol{\phi}\cdot[\boldsymbol{\eta}(k) - \boldsymbol{\eta}(k_0)]\} \times\left(-p(p+2)\sum_l \frac{[\eta_j(lk) - \eta_j(k_0)][\eta_s(lk) - \eta_s(k_0)]}{\{[\boldsymbol{\eta}(lk) - \boldsymbol{\eta}(k_0)]^2\}^{\frac{1}{2}p+2}} \times \exp[2\pi i\boldsymbol{\phi}\cdot\boldsymbol{\eta}(l)] + \delta_{js}p\sum_l \frac{\exp[2\pi i\boldsymbol{\phi}\cdot\boldsymbol{\eta}(l)]}{\{(\boldsymbol{\eta}(lk) - \boldsymbol{\eta}(k_0)]^2\}^{\frac{1}{2}p+1}}\right). \quad (A1)$$

For the case of $k = k_0$, first define

$$C_\phi^L(k_0 j, k_0 s) = \frac{m_0}{m_{k_0}}\left(-p(p+2)\sum_{l\neq 0}\frac{\eta_j(l)\eta_s(l)}{\{[\boldsymbol{\eta}(l)]^2\}^{\frac{1}{2}p+2}} \times \exp[2\pi i\boldsymbol{\phi}\cdot\boldsymbol{\eta}(l)] + \delta_{js}p\sum_{l\neq 0}\frac{\exp[2\pi i\boldsymbol{\phi}\cdot\boldsymbol{\eta}(l)]}{\{[\boldsymbol{\eta}(l)]^2\}^{\frac{1}{2}p+1}}\right). \quad (A2)$$

Then, we have

$$D_\phi^L(k_0 j, k_0 s) = C_\phi^L(k_0 j, k_0 s) - \sum_{k\neq k_0}\left(\frac{m_k}{m_{k_0}}\right)^{\frac{1}{2}} \times D_{\phi=0}^L(k_0 j, ks) - C_{\phi=0}^L(k_0 j, k_0 s). \quad (A3)$$

APPENDIX B

Applying the Ewald transformation to the sums appearing in Appendix A, we obtain the following expressions for the elements of $\mathbf{D}_\phi^L$ for lattices of either the sodium chloride or caesium chloride structure. Only the real parts of the following expressions are nonvanishing.

If $k \neq k_0$, then

$$D_\phi^L(k_0 j, ks) = -\frac{p\pi^{\frac{1}{2}p+1}}{\Gamma(\frac{1}{2}p+1)}\left(\frac{m_0}{m_1}\right)^{\frac{1}{2}} \times\left(\frac{2\pi a^3}{v_a}\sum_h(-1)^{h_1+h_2+h_3}\Phi_{-\frac{1}{2}p+\frac{1}{2}}[\pi\sigma^2(h)]\sigma_j(h)\sigma_s(h) - 2\pi\sum_l\exp[2\pi i\boldsymbol{\phi}\cdot\boldsymbol{\eta}(lkk_0)]\Phi_{\frac{1}{2}p+1}[\pi\eta^2(lkk_0)] \times\eta_j(lkk_0)\eta_s(lkk_0) + \delta_{js}\sum_l\exp[2\pi i\boldsymbol{\phi}\cdot\boldsymbol{\eta}(lkk_0)]\Phi_{\frac{1}{2}p}[\pi\eta^2(lkk_0)]\right), \quad (B1)$$

where $\sigma_j(h) = \phi_j + \xi_j(h)$ and $\eta_j(lkk_0) = \eta_j(lk) - \eta_j(k_0)$. The function $\Phi_m(x)$ is the incomplete gamma function defined by

$$\Phi_m(x) = \int_1^\infty t^m e^{-xt}\, dt. \quad (B2)$$

For the case of $k = k_0$, we have

$$C_\phi^L(k_0 j, k_0 s) = \frac{p\pi^{\frac{1}{2}p+1}}{\Gamma(\frac{1}{2}p+1)}\frac{m_0}{m_{k_0}} \times\left(\frac{2\pi a^3}{v_a}\sum_h\Phi_{-\frac{1}{2}p+\frac{1}{2}}[\pi\sigma^2(h)]\sigma_j(h)\sigma_s(h) - 2\pi\sum_{l\neq 0}\eta_j(l)\eta_s(l)\Phi_{\frac{1}{2}p+1}[\pi\eta^2(l)]\exp[2\pi i\boldsymbol{\phi}\cdot\boldsymbol{\eta}(l)] - \frac{2\delta_{js}}{p+2} + \delta_{js}\sum_{l\neq 0}\Phi_{\frac{1}{2}p}[\pi\eta^2(l)]\exp[2\pi i\boldsymbol{\phi}\cdot\boldsymbol{\eta}(l)]\right). \quad (B3)$$

The element $D_\phi^L(k_0 j, k_0 s)$ is then given by Eq. (A3).

APPENDIX C

Using properties of the incomplete gamma functions given in Appendix B of Ref. 1, we obtain the following expansions of the elements of $\mathbf{D}_\phi^L$ about $\boldsymbol{\phi} = 0$:

$$D_\phi^L(01, 11) = -\mu^{\frac{1}{2}}[\alpha^L + h(\phi)\phi_1^2 + \beta^L\phi_1^2 + \gamma^L(\phi_2^2 + \phi_3^2)] + O(\phi^4), \quad (C1)$$

$$D_\phi^L(01, 12) = -\mu^{\frac{1}{2}}[h(\phi)\phi_1\phi_2 + \delta^L\phi_1\phi_2] + O(\phi^4), \quad (C2)$$

$$D_\phi^L(01, 01) = \alpha^L + h(\phi)\phi_1^2 + b^L\phi_1^2 + c^L(\phi_2^2 + \phi_3^2) + O(\phi^4), \quad (C3)$$

and

$$D_\phi^L(01, 02) = h(\phi)\phi_1\phi_2 + d^L\phi_1\phi_2 + O(\phi^4), \quad (C4)$$

where $\mu = m_0/m_1$. The remaining elements follow from the cubic symmetry and the relation

$$D_\phi^L(1i, 1j) = \mu D_\phi^L(0i, 0j).$$

In the above equations, the function $h(\phi)$ is defined as follows:

(a) If $-\frac{1}{2}p + \frac{1}{2}$ is not a negative integer, then

$$h(\phi) = 2p\pi^{p+\frac{1}{2}} \frac{a^3}{v_a} \frac{\Gamma(\frac{3}{2} - \frac{1}{2}p)}{\Gamma(\frac{1}{2}p + 1)} \phi^{p-3}. \quad \text{(C5)}$$

(b) If $-\frac{1}{2}p + \frac{1}{2}$ is a negative integer, then

$$h(\phi) = \frac{2p\pi^{p+\frac{1}{2}}(-1)^{\frac{1}{2}p-\frac{1}{2}}}{\Gamma(\frac{1}{2}p + 1)\Gamma(\frac{1}{2}p - \frac{1}{2})} \left(\frac{a^3}{v_a}\right) \phi^{p-3} \ln(\phi^2). \quad \text{(C6)}$$

The quantities α^L, β^L, γ^L, δ^L, b^L, c^L, and d^L are constants. Using methods discussed in Ref. 1, we can easily determine their values. For the case of $1 \leq p \leq 3$, these values are given by the following set of equations:

$$\alpha^L = C(p)[(2\pi a^3/v_a)(H_1) - 2\pi L_1 + L_2], \quad \text{(C7)}$$

$$\beta^L = C(p)[(2\pi a^3/v_a)(2\pi^2 H_4 - 5\pi H_3 + H_2) + 4\pi^3 L_3 - 2\pi^2 L_4 - S(p)], \quad \text{(C8)}$$

$$\gamma^L = C(p)[(2\pi a^3/v_a)(2\pi^2 H_5 - \pi H_3) + 4\pi^3 L_5 - 2\pi^2 L_4], \quad \text{(C9)}$$

$$\delta^L = C(p)\{(2\pi a^3/v_a)[H_2 - 4\pi H_3 + (2\pi)^2 H_5] + (2\pi)^3 L_5 - S(p)\}, \quad \text{(C10)}$$

$$b^L = C(p)[(2\pi a^3/v_a)(2\pi^2 H_4' - 5\pi H_3' + H_2') + 4\pi^3 L_3' - 2\pi^2 L_4' - S(p)], \quad \text{(C11)}$$

$$c^L = C(p)[(2\pi a^3/v_a)(2\pi^2 H_5' - \pi H_3') + 4\pi^3 L_5' - 2\pi^2 L_4'], \quad \text{(C12)}$$

and

$$d^L = C(p)\{(2\pi a^3/v_a)[H_2' - 4\pi H_3' + (2\pi)^2 H_5'] + (2\pi)^3 L_5' - S(p)\}. \quad \text{(C13)}$$

If $1 \leq p < 3$, then

$$S(p) = 4\pi(3 - p)^{-1}\left(\frac{a^3}{v_a}\right).$$

If $p = 3$, then

$$S(p) = 2\pi\left(\frac{a^3}{v_a}\right)(\ln \pi - N),$$

where $N = 0.577215\cdots$.

The remaining symbols appearing above are defined as follows:

$$C(p) = p\pi^{\frac{1}{2}p+1}/\Gamma(\tfrac{1}{2}p + 1), \quad \text{(C14)}$$

$$H_1 = \tfrac{1}{3}\sum_{h\neq 0}(-1)^{h_1+h_2+h_3}\Phi_{-\frac{1}{2}p+\frac{1}{2}}[\pi\xi^2(h)]\xi^2(h), \quad \text{(C15)}$$

$$H_2 = \sum_{h\neq 0}(-1)^{h_1+h_2+h_3}\Phi_{-\frac{1}{2}p+\frac{1}{2}}[\pi\xi^2(h)], \quad \text{(C16)}$$

$$H_3 = \tfrac{1}{3}\sum_{h\neq 0}(-1)^{h_1+h_2+h_3}\Phi_{-\frac{1}{2}p+\frac{3}{2}}[\pi\xi^2(h)]\xi^2(h), \quad \text{(C17)}$$

$$H_4 = \sum_{h\neq 0}(-1)^{h_1+h_2+h_3}\Phi_{-\frac{1}{2}p+\frac{5}{2}}[\pi\xi^2(h)]\xi_1^4(h), \quad \text{(C18)}$$

$$H_5 = \sum_{h\neq 0}(-1)^{h_1+h_2+h_3}\xi_1^2(h)\xi_2^2(h)\Phi_{-\frac{1}{2}p+\frac{5}{2}}[\pi\xi^2(h)], \quad \text{(C19)}$$

$$L_1 = \tfrac{1}{3}\sum_l \eta^2(lkk_0)\Phi_{\frac{1}{2}p+1}[\pi\eta^2(lkk_0)], \quad \text{(C20)}$$

$$L_2 = \sum_l \Phi_{\frac{1}{2}p}[\pi\eta^2(lkk_0)], \quad \text{(C21)}$$

$$L_3 = \sum_l \eta_1^4(lkk_0)\Phi_{\frac{1}{2}p+1}[\pi\eta^2(lkk_0)], \quad \text{(C22)}$$

$$L_4 = \tfrac{1}{3}\sum_l \eta^2(lkk_0)\Phi_{\frac{1}{2}p}[\pi\eta^2(lkk_0)], \quad \text{(C23)}$$

and

$$L_5 = \sum \eta_1^2(lkk_0)\eta_2^2(lkk_0)\Phi_{\frac{1}{2}p+1}[\pi\eta^2(lkk_0)]. \quad \text{(C24)}$$

In the above equations, $\eta_i(lkk_0) = \eta_i(lk) - \eta_i(k_0)$. The H_i' are obtained by replacing $(-1)^{h_1+h_2+h_3}$ with one in Eqs. (C15)–(C19). The L_i' are obtained by replacing $\eta_i(lkk_0)$ with $\eta_i(l)$ in Eqs. (C20)–(C24) and not including $l = 0$ in the summations.

APPENDIX D

The bracketed terms ($\langle\ \rangle$), first appearing in Eqs. (30) and (34), are defined as follows: $\langle m, n, r\rangle$ is the sum of all distinct terms of the form $\phi_i^m\phi_j^n\phi_k^r$, where i, j, and k are distinct. For example, $\langle 4, 2, 2\rangle = \phi_1^4\phi_2^2\phi_3^2 + \phi_2^4\phi_1^2\phi_3^2 + \phi_3^4\phi_1^2\phi_2^2$. Finally, we define $\langle m, n\rangle \equiv \langle m, n, 0\rangle$ and $\langle m\rangle \equiv \langle m, 0, 0\rangle$. For example, $\langle 2\rangle = \phi^2$.

[1] J. A. Davies and P. D. Yedinak, J. Math. Phys. **10**, 1090 (1969).
[2] P. P. Ewald, Ann. Physik (Leipzig) **64**, 253 (1921).
[3] M. Born and K. Huang, *Dynamical Theory of Crystal Lattices* (Oxford Univ. Press, London, 1954).
[4] See, for example, E. Merzbacher, *Quantum Mechanics* (John Wiley & Sons, Inc., New York, 1961).
[5] E. W. Kellermann, Phil. Trans. Roy. Soc. London **A238**, 513 (1940).
[6] A. D. B. Woods, W. Cochran, and B. N. Brockhouse, Phys. Rev. **119**, 980 (1960).
[7] A. A. Maradudin and G. H. Weiss, Phys. Rev. **123**, 1968 (1961).

Derivation of a Lower Bound on the Free Energy*

ROBERT H. T. YEH
Department of Physics and Astronomy, State University of New York
Buffalo, New York 14214

(Received 10 December 1969)

For a Hamiltonian with a ground-state energy $E_0 = 0$, we show that for any normalized state $|\psi\rangle$, $\langle\psi| e^{-\beta H} |\psi\rangle \leq e^{-\beta \langle\psi| H |\psi\rangle a}$, where $a = \exp[-\beta \langle\psi| H^2 |\psi\rangle / \langle\psi| H |\psi\rangle]$. From this inequality, a general formula for a lower bound on the free energy is derived.

Bounds of free energy and canonical ensemble averages are of considerable interest.[1] For many complex systems, such as Ising ferromagnets[2] or composite materials,[3] methods of obtaining bounds are the only useful tools which are both tractable and informative. Free energy in the canonical ensemble is given by $F = -\beta^{-1} \ln \sum_s \langle s| e^{-\beta H} |s\rangle$, where $\{|s\rangle\}$ is any complete set of orthonormal states. Bounds of F can be obtained from bounds of $\langle s| e^{-\beta H} |s\rangle$. A very simple upper bound of F was given by Peierls,[4] and one way[5] to prove his theorem is by showing that $\langle\psi| e^{-\beta H} |\psi\rangle \geq e^{-\beta \langle\psi| H |\psi\rangle}$. We derive a rather simple lower bound of F by similar method. For simplicity, we set $E_0 = 0$. This is not a severe restriction, as E_0 is usually known, and merely shifting the energy reference does not alter the thermal properties of the system.[6]

We first note that, for a real function $f(E) = e^{-\beta E}$, $\beta > 0$, Taylor's theorem gives

$$f(\epsilon) - \beta(E - \epsilon) f(\epsilon) \leq f(E) \leq f(\epsilon') - \beta(E - \epsilon') f(E) \tag{1}$$

for arbitrary ϵ, ϵ'.

Next, let H be any Hermitian operator with orthonormal eigenstates $|\psi_n\rangle$, eigenvalues E_n, and $E_n \geq E_0 = 0$. Let $|\psi\rangle$ be any normalized state with $|\psi\rangle = \sum_n a_n |\psi_n\rangle$, $\langle\psi | \psi\rangle = \sum_n \rho_n = 1$, $\rho_n = a_n a_n^*$. Then the expectation value of any function of H in the state $|\psi\rangle$ is simply given by $\overline{O(H)} = \langle\psi| O |\psi\rangle = \sum_n \rho_n O(E_n)$, e.g., $\bar{E} = \langle\psi| H |\psi\rangle = \sum_n \rho_n E_n$, $\bar{f} = \langle\psi| e^{-\beta H} |\psi\rangle = \sum_n \rho_n e^{-\beta E_n}$. Since $E_n \geq 0$, we can multiply Eq. (1) by $\rho_n E_n^i$, and then sum over n.

The choice

$$\epsilon \equiv \epsilon_i = \overline{E^{i+1}} / \overline{E^i}, \quad \epsilon' \equiv \epsilon_i' = \overline{E^{i+1} f} / \overline{E^i f}$$

causes the term with the coefficient β to vanish and we find that

$$f(\epsilon_i') \overline{E^i} \geq \overline{E^i f} \geq f(\epsilon_i) \overline{E^i}. \tag{2}$$

If we let $i = 0$, we obtain $f(\overline{Ef}/\bar{f}) \geq \bar{f} \geq f(\bar{E})$. Since $E_0 = 0$, so $\bar{f} \leq 1$. This, together with Eq. (2) for the case $i = 1$, yields a weaker but simpler upper bound for $\bar{f}$ as

$$e^{-\beta \langle\psi| H |\psi\rangle a} \geq \langle\psi| e^{-\beta H} |\psi\rangle \geq e^{-\beta \langle\psi| H |\psi\rangle}, \tag{3}$$

where $a = \exp[-\beta \langle\psi| H^2 |\psi\rangle / \langle\psi| H |\psi\rangle]$.[7] Identifying $\beta = (kT)^{-1}$ and H as Hamiltonian, we obtain a lower bound of free energy from Eq. (3) as

$$F \geq -\beta^{-1} \ln \sum_\psi e^{-\beta \langle\psi| H |\psi\rangle a},$$

where $\{|\psi\rangle\}$ is any complete orthonormal set of states.

Upper and lower bounds of the canonical ensemble average of any operator O can be written down in terms of $\langle O_n| H |O_n\rangle$, where the $|O_n\rangle$ are eigenstates of O. Furthermore, bounds of thermodynamic derivatives can be obtained by noting that the bounds of $\partial^i \bar{f} / \partial \beta^i$ $[= (-1)^i \overline{E^i f}]$ are given in Eq. (2). From Eq. (3), it is clear that all our bounds are more accurate at higher temperatures. These bounds have been useful in determining the properties of Heisenberg ferromagnets.[8]

* This work was supported by a Faculty Research Fellowship from the Research Foundation of State University of New York.

[1] H. Falk, J. Math. Phys. **7**, 977 (1966), and references cited there.

[2] R. B. Griffiths, Commun. Math. Phys. **6**, 121 (1967).

[3] Z. Hashin and S. Shtrikman, J. Appl. Phys. **33**, 3125 (1962); R. Yeh, *ibid.* **41**, 224 (1970).

[4] R. Peierls, Phys. Rev. **54**, 918 (1963).

[5] H. Falk, Physica **29**, 1114 (1938).

[6] For arbitrary E_0, we simply replace E^i of Eq. (2) by $(E - E_0)^i$. We obtain instead $e^{-\beta \langle\psi| H' |\psi\rangle a'} \geq \langle\psi| e^{-\beta H} |\psi\rangle$, with $H' \equiv H - E_0$ and $a' \equiv \exp[-\beta \langle\psi| H'^2 |\psi\rangle / \langle\psi| H' |\psi\rangle]$. Also,

$$F \geq -\beta^{-1} \ln \sum_\psi e^{-\beta \langle\psi| H' |\psi\rangle a'}.$$

[7] Better upper bounds can be obtained by iteration, e.g.,

$$e^{-\beta b} \geq \exp(-\beta b e^{+\beta b}) \geq \langle\psi| e^{-\beta H} |\psi\rangle,$$

where $b = \langle\psi| H |\psi\rangle a$.

[8] R. Yeh, Phys. Rev. Letters **23**, 1241 (1969).

Diffraction of Electromagnetic Waves by Planar Dielectric Structures. I. Transverse Electric Excitation

G. W. LEHMAN
Science Center, North American Rockwell Corporation, Thousand Oaks, California 91360

(Received 6 November 1969)

A formalism is developed for solving a class of electromagnetic diffraction problems involving heterogeneous planar dielectric structures. Using analytic properties of finite Fourier transforms, we show that the electromagntic field distributions can be obtained by solving a single-variable Fredholm integral equation of the second kind. The special case of TE-wave propagation through a finite metallic guide with a dielectric insert is discussed in detail. Our numerical approach spans the Rayleigh to geometric optics range. Field distributions are presented for typical configurations involving positively and negatively absorbing media. Excitation of bound-mode resonances is discussed.

I. INTRODUCTION

The purpose of this paper is to present a formalism for solving a class of electromagnetic diffraction problems for which the scattering structure is a plane dielectric slab of infinite extent, inside of which the dielectric properties are allowed to vary along directions parallel to the slab boundaries (cf. Fig. 1). Propagation of electromagnetic waves through multilayered dielectric media has received considerable attention in the literature.[1] On the other hand, the excitation of bound electromagnetic modes propagating *parallel* to the layer structure involves a vacuum-dielectric interface or some dielectric inhomogeneity in order to couple with the external field. No previous exact theoretical treatment has been given of this very important class of diffraction problems. The class of exactly solvable problems in diffraction theory, which exploit Wiener–Hopf or similar techniques, are mainly restricted to one or more parallel δ-function discontinuities of the Sommerfeld half-plane type.[2]

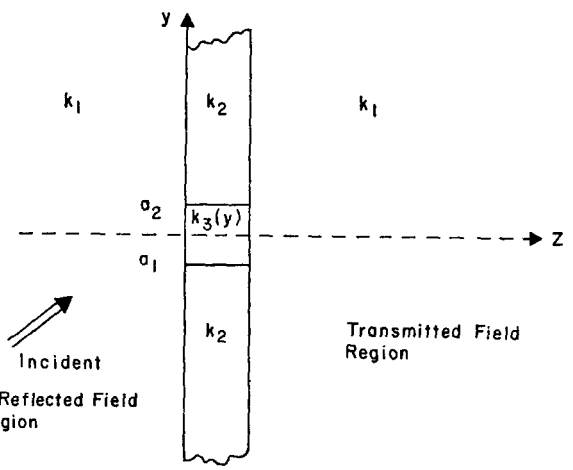


FIG. 1. Geometry of heterogeneous-slab-diffraction problem. Electric field polarized in x direction perpendicular to paper. Propagation constants are denoted by k_n, $\mathrm{Im}\, k_n > 0$ for $n = 1, 2$. Geometry and incident field infinite in x direction.

The mathematical approach taken in this paper is based on quite elementary ideas involving analytical properties of finite Fourier transforms. For our initial exposition, the electric field vector **E** will be taken as polarized parallel to the axis of translational invariance of the scattering structure (cf. Fig. 1). This mode is commonly referred to as TE (transverse electric).

In Sec. II of this paper, a procedure will be developed which allows one to determine field distributions from a single-variable Fredholm integral equation of the second kind. The utility of our approach is established in Sec. IV by studying diffraction of a TE wave from an infinitely conducting screen of finite thickness with a dielectric insert, for the entire wavelength range from the Rayleigh to geometric optics limit.

Numerical results are presented in Sec. V for a variety of slab geometries. The special case of the aperture $L = 0$ is discussed and contact is made with previous work.[3–8]

II. THEORETICAL FORMALISM

The geometry for our diffraction problem is given in Fig. 1. An external, monochromatic field of angular frequency ω and of finite extent in the y direction is incident on the entrance surface $z = 0$. The dielectric regions are characterized by complex propagation constants k_1 and k_2 whose imaginary parts $\mathrm{Im}\, k_n > 0$ refer to an absorptive medium. The region characterized by $k_3(y)$ may have a positive or negative region of absorption.

Maxwell's equations describing the x component E_x of the electric field vector **E** for a TE mode, for a harmonic time-varying field $e^{-i\omega t}$, reduce to

$$\left\{\frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} + [k(y, z)]^2\right\} E_x(y, z) = 0, \qquad (1)$$

with the magnetic field components H_y and H_z being proportional to $\partial E_x/\partial z$ and $\partial E_x/\partial y$, respectively.

Continuity of the tangential components of the electric field vector **E** and magnetic field vector **H**, at the entrance surface $z = 0$, $y = a_1$, and $y = a_2$ and the exit surface $z = L$, requires that

$$E_x(y, -0) = E_x(y, +0),$$

$$\frac{\partial}{\partial z} E_x(y, z)|_{z=-0} = \frac{\partial}{\partial z} E_x(y, z)|_{z=+0} \tag{2}$$

at $z = 0$, with an identical equation for $z = L$. Similarly, at the $y = a_n$ and $n = 1, 2$ surfaces,

$$E_x(a_n - 0, z) = E_x(a_n + 0, z),$$

$$\frac{\partial}{\partial y} E_x(y, z)|_{y=a_n-0} = \frac{\partial}{\partial y} E_x(y, z)|_{y=a_n+0}. \tag{3}$$

A. Reflected Wave Field $z \leq 0$ and Transmitted Wave Field $z \geq L$

If $E_{0x}(y, z)$ represents the incident wave field falling on the medium $z > 0$, then the solution of Eq. (1) for $z < 0$ can be written as

$$E_x(y, z) = E_{0x}(y, z) - \frac{1}{4\pi i} \int_{-\infty}^{\infty} dk_y e^{-ik_y y} e^{-i\alpha_1 z} \frac{C(k_y)}{\alpha_1}, \tag{4}$$

where

$$\alpha_n = (k_n^2 - k_y^2)^{\frac{1}{2}}, \quad \text{Im } \alpha_n > 0, \tag{5}$$

for $n = 1$ and 2, will be used throughout this paper. The second term in Eq. (4) refers to the reflected wave field at the entrance surface $z = 0$.

The wave field for $z > L$ of Fig. 1, characterized by k_1, is given by

$$E_x(y, z) = \frac{-1}{4\pi i} \int_{-\infty}^{\infty} dk_y e^{-ik_y y} e^{i\alpha_1 (z-L)} \frac{D(k_y)}{\alpha_1}, \tag{6}$$

where α_1 is defined by Eq. (5).

B. Procedure for Fixing C and D

In order to obtain $E_x(y, z)$ for $0 < z < L$, we employ a finite Fourier transform representation[9] defining

$$E_+(y, k) = \int_0^L dz e^{ikz} E_x(y, z), \tag{7}$$

so that, by MacRobert's theorem,[9]

$$E_x(y, z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{-ikz} E_+(y, k), \tag{8}$$

for $0 < z < L$. We note, in particular, that $E_+(y, k)$ is an entire function which tends to zero in the upper half-k-plane.

We now multiply Eq. (1) by e^{ikz}, integrate over z from 0 to L, and make use of Eq. (2) and the boundary conditions at the entrance and exist surfaces. One obtains an inhomogeneous differential equation for $E_+(y, k)$ containing C and D which can be written, for the nth region, as

$$\left\{\frac{d^2}{dy^2} - k^2 + [k_n(y)]^2\right\} E_+(y, k)$$

$$= f_0(y, k) + \frac{1}{4\pi} \int_{-\infty}^{\infty} dk_y \frac{e^{-ik_y y}}{\alpha_1} \{(k + \alpha_1) C(k_y) + e^{ikL}(\alpha_1 - k) D(k_y)\}. \tag{9}$$

In Eq. (9),

$$f_0(y, k) = \frac{\partial E_{0x}(y, z)}{\partial z}\bigg|_{z=0} - ikE_{0x}(y, 0) \tag{10}$$

is a source term. The second term on the right-hand side of Eq. (9) accounts for reflection at the exit and entrance surfaces.

The functions $E_+(y, k)$ and $e^{ikL} E_+(y, -k)$ tend to zero as $\text{Im } k \to \infty$ and exist for all k in the upper half-plane. Consequently, it will be possible to determine C and D as a linear combination of the two previously mentioned functions at $k = \alpha_2$.

Let

$$B_\pm(y, k) = E_+(y, k) \pm e^{ikL} E_+(y, -k), \tag{11}$$

$$Q_\pm(k_y) = C(k_y) \pm D(k_y), \tag{12}$$

and

$$f_\pm(y, k) = f_0(y, k) \pm e^{ikL} f_0(y, -k); \tag{13}$$

then, from Eq. (9), one has

$$\left\{\frac{d^2}{dy^2} - k^2 + [k_n(y)]^2\right\} B_\pm(y, k)$$

$$= f_\pm(y, k) + \frac{1}{4\pi} \int_{-\infty}^{\infty} dk_y \frac{1}{\alpha_1} \times e^{-ik_y y}[k + \alpha_1 \pm e^{ikL}(\alpha_1 - k)] Q_\pm(k_y). \tag{14}$$

One should note that the $B_\pm(y, k)$ functions are also entire functions which tend to zero as $\text{Im } k \to +\infty$.

Since $f_0(y, k)$ is assumed to vanish outside the range $|y| > c$, we can Fourier-transform Eq. (14) with respect to $-\infty < y < \infty$ and obtain

$$\hat{B}_\pm(k_y, k) = \frac{1}{k^2 - \alpha_2^2} \left\{-\hat{f}_\pm(k_y, k) - \frac{1}{2\alpha_1} [k + \alpha_1 \pm e^{ikL}(\alpha_1 - k)] Q_\pm(k_y) + \int_{a_1}^{a_2} dy e^{ik_y y} U(y) B_\pm(y, k)\right\}, \tag{15}$$

where

$$U(y) = [k_3(y)]^2 - k_2^2 \tag{16}$$

and $\hat{B}_\pm(k_y, k)$ and $\hat{f}_\pm(k_y, k)$ denote the respective Fourier transforms of $B_\pm(y, k)$ and $f_\pm(k_y, k)$.

The last term in Eq. (15) arises from the fact that a term $k_2^2 B_\pm(y, k)$ was added and subtracted from the left side of Eq. (14) prior to taking the Fourier transform on y.

As previously noted, the $B_\pm(y, k)$ functions are entire functions [cf. Eqs. (7) and (11)] which tend to zero as $\text{Im}\, k \to \infty$. According to Eq. (5), $\alpha_2 = (k_2^2 - k_y^2)^{\frac{1}{2}}$, $-\infty < k_y < \infty$, $\text{Im}\, \alpha_2 > 0$, lies in the upper right half of the complex plane since $\text{Im}\,(k_2) > 0$. In particular, $\alpha_2 = k_2$ for $k_y = 0$ and $\alpha_2 \to i\infty$ as $k_y \to \pm\infty$. In order for Eq. (15) to hold for all k such that $\text{Im}\, k > -k_0$ and $k_0 > 0$, the function in the braces must vanish at $k = \alpha_2$, $\text{Im}\, \alpha_2 > 0$, so that

$$Q_\pm(k_y) = \frac{2\alpha_1}{\alpha_1 + \alpha_2 \pm (\alpha_1 - \alpha_2)e^{i\alpha_2 L}} \times \{-\hat{f}_\pm(k_y, \alpha_2) + \Psi_\pm(k_y)\}, \quad (17)$$

where

$$\Psi_\pm(k_y) = \int_{a_1}^{a_2} dy e^{ik_y y} U(y) B_\pm(y, \alpha_2). \quad (18)$$

The function $B_\pm(y, \alpha_2)$, $\text{Im}\, \alpha_2 > 0$, can be obtained by solving Eq. (14), subject to the boundary condition (3) for $k = \zeta$, $-\infty < \zeta < \infty$, and by using Cauchy's integral formula, i.e.,

$$B_\pm(y, \alpha_2) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} d\zeta \frac{1}{\zeta - \alpha_2} B_\pm(y, \zeta). \quad (19)$$

Equation (14) for $B_\pm$, with $k = \zeta$, $-\infty < \zeta < \infty$, subject to the boundary condition (3), can be solved in terms of elementary functions if the inset in Fig. 1 consists of dielectric layers running parallel to the z axis so that $k_3(y)$ is piecewise constant. After $B_\pm(y, \zeta)$ has been constructed, one can obtain an integral equation for $\Psi_\pm(k_y)$, using Eqs. (18) and (19). In general, the integrals involving ζ in Eq. (19) cannot be carried out explicitly, and one must resort to numerical evaluation. If a_1, a_2, k_2, and $k_3(y)$ are *finite*, $\Psi_\pm(k_y)$ will obey a Fredholm equation of the second kind whose kernel is Hilbert–Schmidt type. When $a_2 = \infty$, one has an exceptional case. If our dielectric structure is characterized by the parameters $a_1 = 0$ and $a_2 = \infty$, with $k_3(y) = k_T(y)$ for $0 \leq y \leq a$ and $k_3(y) = k_3$ (const) for $a \leq y < \infty$, the integral equation for $\Psi_\pm$ is singular. It can, however, be converted to Fredholm by the well-known Carleman technique. On the other hand, if $k_3 = k_1$ (vacuum), $a_1 = 0$, and a_2 and $L = \infty$, i.e., a dielectric overlay on a right-angle prism, one *still* obtains a Fredholm integral equation with a Hilbert–Schmidt kernel.

The case $k_3(y) = k_3$ (const) and $a_2 = -a_1 = b$ is investigated in Sec. III of this paper, where $B_\pm(y, \zeta)$ is explicitly exhibited [Eq. (40)]. In Sec. IV, we derive an integral equation for $(k_2)^{-1}\Psi_\pm(k_y)$ for the limiting case of a dielectric insert imbedded in a medium having infinite conductivity, i.e., $k_2 \to i\infty$. The field distributions for this case are considered in complete detail.

At this point, it appears worthwhile to summarize our formal procedure for finding $B_\pm(y, \zeta)$, $B_\pm(y, \alpha_2)$, $\Psi_\pm(k_y)$, $C(k_y)$, $D(k_y)$, and $E_x(y, z)$.

C. Summary of Formal Expressions

In general, one is interested in physical situations where a beam of finite width is used. For this case,

$$E_{0x}(y, z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk_y e^{-ik_y y} e^{i\alpha_1 z} E_I(k_y). \quad (20)$$

Equation (13) becomes, by means of Eq. (10),

$$f_\pm(y, k) = \frac{i}{2\pi} \int_{-\infty}^{\infty} dk_y e^{-ik_y y} M_\pm^{(1)}(k_y, k) E_I(k_y), \quad (21)$$

where

$$M_\pm^{(1)}(k_y, k) = \alpha_1 - k \pm e^{ikL}(\alpha_1 + k). \quad (22)$$

From Eqs. (17) and (21), one can write Eq. (14) for $k = \zeta$, $-\infty < \zeta < \infty$, as

$$\left\{\frac{d^2}{dy^2} - \zeta^2 + [k_n(y)]^2\right\} B_\pm(y, \zeta) = H_\pm(y, \zeta) + \frac{1}{2\pi} \int_{-\infty}^{\infty} dk_y e^{-ik_y y} M_\pm^{(2)}(k_y, \zeta) \Psi_\pm(k_y), \quad (23)$$

where

$$H_\pm(y, \zeta) = \frac{i}{2\pi} \int_{-\infty}^{\infty} dk_y e^{-ik_y y} E_I(k_y) [M_\pm^{(1)}(k_y, \zeta) - M_\pm^{(2)}(k_y, \zeta) M_\pm^{(1)}(k_y, \alpha_2)] \quad (24)$$

and

$$M_\pm^{(2)}(k_y, \zeta) = \frac{\zeta + \alpha_1 \pm e^{i\zeta L}(\alpha_1 - \zeta)}{\alpha_2 + \alpha_1 \pm e^{i\alpha_2 L}(\alpha_1 - \alpha_2)}. \quad (25)$$

It is also instructive to convert Eq. (23) to an integral equation. The simplest way to do this is to add the term $(k_2)^2 B_\pm(y, \zeta)$ to both sides of Eq. (23), take the Fourier transform on y, solve for $\hat{B}_\pm(k_y, \zeta)$, as was done in constructing Eq. (15), and then re-invert. [Inversion of Eq. (15) with $k = \zeta$ leads to the same result when the appropriate substitutions are made.] One finds that

$$\begin{aligned} B_\pm(y, \zeta) &= \frac{i}{2\pi} \int_{-\infty}^{\infty} dk_y e^{-ik_y y} E_I(k_y) [M_\pm^{(1)}(k_y, \zeta) \\ &\quad - M_\pm^{(2)}(k_y, \zeta) M_\pm^{(1)}(k_y, \alpha_2)](\alpha_2^2 - \zeta^2)^{-1} \\ &\quad + \frac{1}{2\pi} \int_{-\infty}^{\infty} dk_y e^{-ik_y y} M_\pm^{(2)}(k_y, \zeta)(\alpha_2^2 - \zeta^2)^{-1} \Psi_\pm(k_y) \\ &\quad + \int_{a_1}^{a_2} dy' \frac{e^{-\gamma|y-y'|}}{2\gamma} U(y') B_\pm(y', \zeta) \end{aligned} \quad (26)$$

and

$$\gamma = (\zeta^2 - k_2^2)^{\frac{1}{2}}, \quad \text{Re}\,\gamma > 0. \tag{27}$$

Finally, one fixes $\Psi_{\pm}(k_y)$ by using Eqs. (18) and (19), i.e.,

$$\Psi_{\pm}(k_y) = \frac{1}{2\pi i}\int_{-\infty}^{\infty} d\zeta \frac{1}{\zeta - \alpha_2}\int_{a_1}^{a_2} dy e^{ik_y y} U(y) B_{\pm}(y, \zeta). \tag{28}$$

The field distribution for $z > L$ is given by Eq. (6) via Eqs. (12) and

$$E_x(y, z > L) = \frac{2}{\pi}\int_{-\infty}^{\infty} dk_y e^{-ik_y y} e^{i\alpha_1 z} \times E_I(k_y) \frac{\alpha_1\alpha_2 e^{i(\alpha_2 - \alpha_1)L}}{(\alpha_1 + \alpha_2)^2 - e^{2i\alpha_2 L}(\alpha_1 - \alpha_2)^2} - \frac{1}{4\pi i}\int_{-\infty}^{\infty} dk_y e^{-ik_y y} e^{i\alpha_1(z-L)}\left(\frac{\Psi_+(k_y)}{N_+} - \frac{\Psi_-(k_y)}{N_-}\right), \tag{29}$$

where

$$N_{\pm} = \alpha_1 + \alpha_2 \pm e^{i\alpha_2 L}(\alpha_1 - \alpha_2). \tag{30}$$

The field distributions for $0 < z < L$ are obtained from

$$E_x(y, 0 < z < L) = \frac{1}{4\pi}\int_{-\infty}^{\infty} d\zeta e^{-i\zeta z}[B_+(y, \zeta) + B_-(y, \zeta)]. \tag{31}$$

One should note that Eqs. (29) and (31) give the field distributions in analytic form for an infinite slab of thickness L, with propagation constant k_2 imbedded in vacuum $[k_1]$ since $a_1 = a_2 = \Psi_{\pm} = 0$ for this case.

III. DIFFRACTION BY A HOMOGENEOUS SLAB IMBEDDED IN AN INFINITE SLAB

In this section, the diffraction of a TE wave by a homogeneous dielectric slab, which has $k_3(y) = k_3$, thickness $2b$, and length L, imbedded in an infinite slab of width L (shown in Fig. 1), with $a_2 = -a_1 = b$, will be discussed. If $k_2 = k_1$, one has the isolated slab problem. If $L = \infty$, one has a semi-infinite diffraction problem. If $\text{Im}\,k_2 b \gg 1$, one has a dielectric slab, characterized by $k_3(y)$ imbedded in a highly absorbing medium. In the limit of infinite conductivity, $k_2 \to i\infty$, our diffraction problem takes its simplest form. Even this latter problem cannot be solved in closed form, but our technique leads to a tractable procedure for obtaining the field distributions, as will be shown below.

It is of interest to consider the case where $\text{Im}\,k_3 < 0$, corresponding to negative absorption as well as the normal $\text{Im}\,k_3 > 0$. For this reason, quantity α_3 appearing below will be defined as

$$\alpha_3 = \text{sgn}\,(\text{Im}\,k_3)(k_3^2 - k_y^2)^{\frac{1}{2}}, \quad \text{Im}\,\alpha_3 > 0, \tag{32}$$

for $-\infty \leq k_y \leq \infty$.

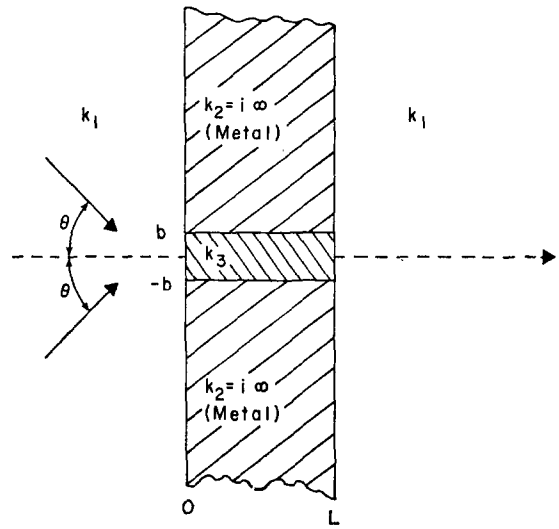


FIG. 2. Metal–dielectric-guide diffraction geometry [cf. Fig. 1]. Incident wave field is taken to be symmetric with respect to reflection in (x, z) plane. $\text{Im}\,k_3 > 0$ corresponds to passive absorbing dielectric material. $\text{Im}\,k_3 < 0$ corresponds to active (optically pumped) material.

The solutions of Eq. (23) for $B_{\pm}(y, \zeta)$ can be decomposed into even and odd parts, with respect to reflection in the $y = 0$ plane of Fig. 2. The letters g (gerade) and u (ungerade) will be used to denote evenness and oddness. The quantities a_1 and a_2 are given by $-b$ and b, respectively. One can write

$$B_{\pm}(y, \zeta) = B_{\pm}^{g}(y, \zeta) + iB_{\pm}^{u}(y, \zeta). \tag{33}$$

In the remainder of this paper, we assume that the excitation field is an even function of y. Experimentally, one can easily achieve this situation by using a beam splitter and adjusting the phases to give a symmetric transmission pattern. For example, a symmetry plane wave excitation corresponds to

$$E_{0I}^{g}(k_y) = 2\pi[\delta(k_y - k_1 \sin\theta) + \delta(k_y + k_1 \sin\theta)], \tag{34}$$

where θ is the angle of incidence measured from the z axis and δ is the Dirac δ function.

A. Gerade Field Distributions $B_{\pm}^{g}(y, \zeta)$

Equation (23), subject to the boundary conditions at $y = b$, must be solved for $B_{\pm}^{g}(y, \zeta)$. Standard procedures exist for doing this. For our purposes, a particularly compact result can be obtained. Since

$B^g_\pm(y, \zeta) = B^g_\pm(-y, \zeta)$, one can show that

$$B^g_\pm(y, \zeta) = G_3(y) + \cos(\kappa y)\{\gamma[G_2(b) - G_3(b)] + G_2'(b) - G_3'(b)\}/\Delta \quad (35)$$

for $-b \leq y \leq b$. The primes on G_n denote differentiation with respect to y and

$$G_n(y) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dk_y W(k_y, \zeta)(\alpha_n^2 - \zeta^2)^{-1}\cos(k_y y), \quad (36)$$

where

$$W(k_y, \zeta) = M^{(2)}_\pm(k_y, \zeta)\Psi^g_\pm(k_y) + iE^g_I(k_y)[M^{(1)}_\pm(k_y, \zeta) - M^{(2)}_\pm(k_y, \zeta)M^{(1)}_\pm(k_y, \alpha_2)], \quad (37)$$

$$\Delta = \gamma\cos\kappa b - \kappa\sin\kappa b, \quad \kappa = (k_3^2 - \zeta^2)^{\frac{1}{2}}. \quad (38)$$

For $y \geq b$,

$$B^g_\pm(y, \zeta) = G_2(y) + e^{-\gamma(y-b)}\int_0^b dy' \frac{\cos(\kappa y')G_2(y')(k_3^2 - k_2^2)}{\Delta}. \quad (39)$$

In order to construct the field distribution for $0 < z < L$ from $B^g_\pm(y, \zeta)$, one must invert Eq. (35) with respect to ζ after $\Psi^g_\pm$ has been obtained. For this reason, it is convenient to rewrite Eq. 35 as

$$B^g_\pm(y, \zeta) = G_2(y) + \frac{k_3^2 - k_2^2}{2\pi}\int_{-\infty}^{\infty} dk_y' \times W(k_y', \zeta)(\alpha_2'^2 - \zeta^2)^{-1}(\alpha_3'^2 - \zeta^2)^{-1} \times \left[\left(\frac{\gamma\cos(k_y'b) - k_y'\sin(k_y'b)}{\gamma\cos(\kappa b) - \kappa\sin(\kappa b)}\right)\cos\kappa y - \cos k_y' y\right], \quad (40)$$

where the primes on α_n denote k_y' rather than k_y.

An integral equation for $\Psi^g_\pm(k_y)$ is now obtained by inserting Eq. (40) into Eq. (28). The kernel for this integral equation can be shown to be Hilbert–Schmidt for finite k_1, k_2, and k_3. This ends our general discussion showing the existence of an integral equation for $\Psi^g_\pm$. Further discussion on this general problem will be deferred to a future publication. A detailed discussion of the case $k_2 \to -i\infty$ is presented in the next section.

IV. DIFFRACTION BY A DIELECTRIC SLAB BOUNDED BY STRONGLY ABSORBING MEDIA

From an experimental as well as theoretical point of view, it is of considerable interest to examine a special, but simple,diffraction problem employing the geometry of Fig. 2. Here we take $k_2 \to i\infty$. The factor $\gamma \to \infty$; α_2' and $\alpha_2 \to k_2$ and

$$W(k_y', \zeta) \to [\zeta + \alpha_1' \pm e^{i\zeta L}(\alpha_1' - \zeta)]I^g_\pm(k_y') + 2i\alpha_1' E^g_I(k_y')[1 \pm e^{i\zeta L}], \quad (41)$$

where

$$I^g_\pm(k_y') = \lim_{k_2 \to i\infty} k_2^{-1}\Psi^g_\pm(k_y'). \quad (42)$$

The quantity in the large brackets of Eq. (40) reduces to $\cos(k_y'b)\sec(\kappa b)\cos(\kappa y) = \cos(k_y' y)$, which vanishes for $\zeta = \pm\alpha_3'$ and has poles at $\kappa b = (r + \frac{1}{2})\pi$, $r = 0, \pm 1, \cdots, \infty$. When this bracket expression is divided by $(\alpha_3'^2 - \zeta^2)$, one has a meromorphic function of ζ which can be expanded in partial fractions, with the result that

$$b^2(\tau^2 - \sigma'^2)^{-1}[\cos(\sigma')\sec(\tau)\cos(\tau\xi) - \cos(\sigma'\xi)] = 2\sum_{r=1}^{\infty}\tau_r(-1)^{r+1}\cos\left(\frac{\tau_r y}{b}\right)(\tau_r^2 - \sigma'^2)^{-1} \times \cos(\sigma')(\zeta^2 - \zeta_r^2)^{-1}, \quad (43)$$

where

$$\xi = y/b, \quad \sigma' = k_y' b, \quad \tau = (k_3^2 - \zeta^2)^{\frac{1}{2}}b, \quad \tau_r = (r - \tfrac{1}{2})\pi, \quad (44)$$

and

$$\zeta_r = \operatorname{sgn}(\operatorname{Im} k_3)[k_3^2 - (\tau_r/b)^2]^{\frac{1}{2}}, \quad \operatorname{Im}\zeta_r > 0. \quad (45)$$

The first term in Eq. (40) drops out as $k_2 \to i\infty$ and one is left with

$$B^g_\pm(y, \zeta) = \sum_{r=1}^{\infty}(-1)^r\tau_r\cos\left(\frac{\tau_r y}{b}\right)(\zeta^2 - \zeta_r^2)^{-1} \times [(1 \pm e^{i\zeta L})(S_r + p_{1r}) + (1 \mp e^{i\zeta L})\zeta p_{2r}], \quad (46)$$

where

$$S_r = \frac{2i}{\pi}\int_{-\infty}^{\infty} dk_y'\alpha_1'\cos(\sigma')E^g_I(k_y')(\tau_r^2 - \sigma'^2)^{-1} \quad (47)$$

and

$$p_{nr} = \frac{1}{\pi}\int_{-\infty}^{\infty} dk_y'(\alpha_1')^{2-n}I^g_\pm(k_y')(\tau_r^2 - \sigma'^2)^{-1}\cos(\sigma'). \quad (48)$$

Field Expansions in the Slab

The field distribution in the slab is obtained by inverting Eq. (46). [Eq. (31) relates E_x and the inverse of $B_\pm$.] One finds that

$$B^g_\pm(y, z) = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\zeta e^{-i\zeta z}B^g_\pm(y, \zeta) = i\sum_{r=1}^{\infty}(-1)^r\tau_r(2\zeta_r)^{-1}\cos\left(\frac{\tau_r y}{b}\right)[e^{i\zeta_r z} \pm e^{i\zeta_r(L-z)}]\hat{g}_\pm(\tau_r), \quad (49)$$

where

$$\hat{g}_\pm(\tau_r) = S_r + p_{1r} - \zeta_r p_{2r}. \quad (50)$$

Equation (49) shows that one can determine $\hat{g}_\pm(\tau_r)$, rather than p_{1r} and p_{2r} separately, to fix the field distribution in the slab. Using Eqs. (42), (28), and (46), one further finds that

$$I_\pm^g(k_y) = -b\sum_{r=1}^{\infty}\tau_r^2(\tau_r^2 - \sigma^2)^{-1} \times (1 \pm e^{i\zeta_r L})\zeta_r^{-1}\hat{g}_\pm(\tau_r)\cos(\sigma), \quad (51)$$

where $\sigma = k_y b$.

From Eqs. (48), (50), and (51), the $\hat{g}_\pm$'s are fixed by

$$\hat{g}_\pm(\tau_n) = S_n - \frac{1}{\pi}\int_{-\infty}^{\infty} dk_y'(\zeta_n - \alpha_1')(\tau_n^2 - \sigma'^2)^{-1} \times \cos(\sigma')I_\pm^g(k_y'). \quad (52)$$

In the remainder of this paper, it will be assumed that the incident field corresponds to a symmetric plane wave excitation given by $\exp(ik_1 z\cos\theta) \times \cos(k_1 y\sin\theta)$ so that Eq. (47) reduces to

$$bS_r = 4i\sigma_1\cos(\sigma_1\sin\theta)[\tau_r^2 - (\sigma_1\sin\theta)^2]^{-1}, \quad (53)$$

where θ is the angle of incident measured from the z axis.

The field in the guide at (y, z) is obtained from Eq. (49), where $\hat{g}_\pm(\tau_r)$ is determined by solving Eqs. (51) and (52). Similarly, to determine the transmitted or reflected wave fields, one needs $I_\pm^g(k_y)$, fixed by Eq. (51).

A. Transformed Equations

From a numerical standpoint, it is judicious to eliminate the $\cos\sigma$ factors in our equations. To do this, one makes use of the identity

$$\frac{2}{\pi}\int_{-\infty}^{\infty} d\sigma\cos^2(\sigma)(\tau_r^2 - \sigma^2)^{-1}(\tau_n^2 - \sigma^2)^{-1} = \tau_n^{-2}\delta_{nr} \quad (54)$$

to rewrite Eq. (53) in the form

$$g_\pm(\tau_n) = bS_n + \frac{1}{2\pi}\int_{-\infty}^{\infty} d\sigma\,\frac{(1 + e^{2i\sigma})[(\sigma_1^2 - \tau_n^2)^{\frac{1}{2}} - (\sigma_1^2 - \sigma^2)^{\frac{1}{2}}]}{\tau_n^2 - \sigma^2}\hat{\phi}_\pm(\sigma), \quad (55)$$

where $2\cos^2\sigma$ is replaced by $1 + e^{2i\sigma}$ and $b(\zeta_n - \alpha_1')$ is rewritten as

$$b\zeta_n - (\sigma_1^2 - \tau_n^2)^{\frac{1}{2}} + (\sigma_1^2 - \tau_n^2)^{\frac{1}{2}} - (\sigma_1^2 - \sigma^2)^{\frac{1}{2}}$$

since $I_\pm^g(k_y)$ is an even function of $k_y b = \sigma$. In Eq. (55),

$$\hat{\phi}_\pm(\sigma) = \sum_{r=1}^{\infty}\beta_\pm(\tau_r)(\tau_r^2 - \sigma^2)^{-1}g_\pm(\tau_r), \quad (56)$$

with

$$g_\pm(\tau_r) = (2\zeta_r)^{-1}\hat{g}_\pm(\tau_r)\{b\zeta_r + (\sigma_1^2 - \tau_r^2)^{\frac{1}{2}} \pm e^{i\zeta_r L}[-b\zeta_r + (\sigma_1^2 - \tau_r^2)^{\frac{1}{2}}]\} \quad (57)$$

and

$$\beta_\pm(\tau_r) = 2\tau_r^2(1 \pm e^{i\zeta_r L})\{b\zeta_r + (\sigma_1^2 - \tau_r^2)^{\frac{1}{2}} \pm e^{i\zeta_r L}[-b\zeta_r + (\sigma_1^2 - \tau_r^2)^{\frac{1}{2}}]\}^{-1}. \quad (58)$$

The relationship between $I_\pm^g(k_y)$ [Eq. (51)] and $\hat{\phi}_\pm(\sigma)$ is simply

$$I_\pm^g(k_y) = -b\cos(\sigma)\hat{\phi}_\pm(\sigma). \quad (59)$$

The function $\hat{\phi}_\pm(\sigma)$ defined by Eq. (56) has poles at $\sigma = \pm\tau_r$, $r = 1, 2, \cdots$. Equation (55), relating $g_\pm(\tau_n)$ to the other $g_\pm(\tau_r)$'s, was designed to cancel out the pole of *second* order associated with the diagonal term $r = n$; i.e., the numerator of the integrand in Eq. (55) vanishes like $(\sigma - \tau_n)^2$ for $\sigma \to \tau_n$.

Next, it is convenient to transform Eq. (55) into a more useful one by putting $\sigma = -\sigma_1\cosh t_1$ and $(\sigma_1^2 - \sigma^2)^{\frac{1}{2}} = -i\sigma_1\sinh t_1$, where t_1 runs from $-\infty$ to 0 to $i\pi$ to $i\pi + \infty$ in straight-line segments. As previously mentioned, the poles in Eq. (55) were eliminated. Now, we shift the entire contour upward by $\frac{1}{2}i\pi$, i.e., we let $t_1' = \frac{1}{2}i\pi + t_1$ and then

$$\sigma = -i\sigma_1\sinh t_1'$$

for the segment $-\infty < t_1 < 0$. Similar remarks hold for the other two parts of the contour. The terms involving the factor $(\sigma_1^2 - \tau_n^2)^{\frac{1}{2}}$ cancel out. Finally, setting $v = \sigma_1\sinh(t')$, we can rewrite Eqs. (55) and (56) as

$$g_\pm(\tau_n) = bS_n + \int_{-i\sigma_1}^{\infty} dv\,\frac{(\sigma_1^2 + v^2)^{\frac{1}{2}}}{\tau_n^2 + v^2}(1 + e^{-2v})\phi_\pm(v) \quad (60)$$

and

$$\phi_\pm(v) = \frac{i}{\pi}\sum_{r=1}^{\infty}\beta_\pm(\tau_r)(\tau_r^2 + v^2)^{-1}g_\pm(\tau_r). \quad (61)$$

Equations (56) and (61) show the relationship between $\phi_\pm(v)$ and $\hat{\phi}_\pm(\sigma)$. The contour in Eq. (60) runs from $v = -i\sigma_1$ to ∞, avoiding the poles at $v = \pm i\tau_n$. On the other hand, the factor $(\sigma_1^2 + v^2)^{\frac{1}{2}}(1 + e^{-2v})$ assures the existence of the integrals in the limiting case $\sigma_1 \to \tau_n$ for off-diagonal as well as diagonal terms, in case Eq. (61) is substituted into Eq. (60), and the resulting system solved for $g_\pm(\tau_n)$.

B. Asymptotic Behavior of $\phi_\pm(v)$ and $g_\pm(\tau_r)$ for $L > 0$

For cases of physical interest $S_n = O(n^{-2})$ as $n \to \infty$. Similarly, any finite portion of the line integral in Eq. (60) leads to the same result, i.e., $O(n^{-2})$. These conclusions also hold for $\phi_\pm(v)$ with respect to any finite portion of the τ_r spectrum in Eq. (61). Hence, one readily sees that the behavior of $\phi_\pm(v)$ for $v \to \infty$ is controlled by $g_\pm(\tau_r)$, for $r \gg 1$, since $\beta_\pm(\tau_r) \to -i\tau_r + O(r^{-1})$ as $r \to \infty$.

Using the Euler summation formula[10] and the above remarks, one can write Eq. (61), for $L > 0$, as

$$\phi_\pm(v) = O(v^{-2}) + \frac{1}{\pi^2}\int_{v_0}^{\infty} d\tau\,\frac{\tau}{\tau^2 + v^2}g_\pm(\tau), \quad (62)$$

where

$$g_{\pm}(v) = O(v^{-2}) + \int_{v_0}^{\infty} d\tau \frac{\tau}{\tau^2 + v^2} \phi_{\pm}(\tau), \quad (63)$$

for $1 \ll v_0 \leq v < \infty$. These equations are identical in structure except for the multiplying factor π^{-2}. This structure suggests that

$$\phi_{\pm}(v) \to A_1 v^{-p}, \quad g_{\pm}(v) \to \pi A_1 v^{-p}, \quad v \to \infty. \quad (64)$$

The factor π simply scales $\phi_{\pm}$ and $g_{\pm}$ properly. If one multiplies Eq. (62) by $2\pi v^p$ and takes the limit as $v \to \infty$, using Eq. (64), then the constant p is fixed by

$$2\pi A_1 = 2\int_0^{\infty} dx \frac{x^{1-p}}{1 + x^2} A_1 \equiv \pi \csc(\tfrac{1}{2}\pi p) A_1. \quad (65)$$

Similarly, one obtains the same equation for p from Eq. (63). In deriving Eq. (65), the term involving $O(v^{-2})$ drops out provided that $p < 2$. In fact, the only *permissible* solution of Eq. (65) for $A_1 \neq 0$ is $p = \frac{5}{3}$, consistent with the above assumption, $p < 2$. These arguments indicate that the series defining $B_{\pm}(y, z)$ in Eq. (49) are absolutely convergent.

When $L = 0$, we have $g_- = I_-^g \equiv 0$ and $\beta_+(\tau_r) \to -2i\tau_r$, as $r \to \infty$. One can show that $p_1 = p_2 = \frac{3}{2}$ in the previous $L > 0$ analysis so that $\phi_+(v) \to O(v^{-\frac{3}{2}})$ and $g_+(\tau_r) \to \tau_r^{-\frac{3}{2}}$ as $|v|$ or $r \to \infty$. This latter result is in agreement with the well-known result that

$$E_x(y, 0) \to O[(1 - (y/b)^2)^{\frac{1}{2}}]$$

as $|y| \to b$, established by Sommerfeld and others (see Ref. 3, p. 73, for a discussion).

C. Integral Equation for $\phi_{\pm}(v)$

In this section, an integral equation will be constructed for $\phi_{\pm}(v)$ by substituting Eq. (60) into Eq. (61). To achieve a compact notation, it is convenient to specify the precise contour for v in Eq. (60) so that a single continuous kernel can be used for the entire path. To do this, let v_0 be a real change of scale parameter[11] and define

$$\begin{aligned} v &= -i\sigma_1 + v_0(1 + 2t'), \quad 0 \leq t' < \infty, \\ &= -i\sigma_1 + v_0(1 + t')^2, \quad -1 \leq t' \leq 0, \end{aligned} \quad (66)$$

for the path of integration in Eq. (60). Then, for $-1 \leq t' \leq 0$,

$$\begin{aligned} \frac{dv}{dt'} &= 2v_0(1 + t'), \\ (\sigma_1^2 + v^2)^{\frac{1}{2}} &= (1 + t')v_0^{\frac{1}{2}}(v + i\sigma_1)^{\frac{1}{2}}, \end{aligned} \quad (67)$$

where the real part of the square-root factor is to be taken as positive. (We note that dv/dt' is continuous at $t' = 0$.) Noting that

$$1 + e^{-2v} = O[(1 + t)^2], \quad \sigma_1 \equiv (N - \tfrac{1}{2})\pi \equiv \tau_N, \quad (68)$$

when $t' \cong -1$, one sees that $1 + e^{-2v}$ cancels out the pole in Eq. (61) at $r = N$ exactly. Similarly, the factor $(\sigma_1^2 + v^2)^{\frac{1}{2}}(dv/dt')$ in the integrand of Eq. (60) in t' space cancels out the pole at $n = N$. When $\sigma_1 \neq \tau_N$, $N = 1, 2, \cdots$, no pole occurs along the $-1 \leq t' \leq 0$ path.

Now, we put

$$\begin{aligned} u &= -i\sigma_1 + v_0(1 + 2t), \quad t \geq 0, \\ &= -i\sigma_1 + v_0(1 + t)^2, \quad -t \leq t \leq 0, \end{aligned} \quad (69)$$

covering the same path as v. Next, we let

$$\begin{aligned} A(t) &= 1 + (u/\sigma_1)^2, \quad t \leq 0, \\ &= 1 + (v_a/\sigma_1)^2, \quad t \geq 0, \end{aligned} \quad (70)$$

$$\begin{aligned} B(t') &= \sigma_1^2(\sigma_1^2 + v^2)^{\frac{1}{2}} \frac{dv}{dt'}, \quad t' \leq 0, \\ &= 2(\sigma_1^2 + v^2)^{\frac{1}{2}} v_0[1 + (v_a/\sigma_1)^2]^{-1}, \quad t' \geq 0, \end{aligned} \quad (71)$$

where

$$v_a = v_0 - i\sigma_1. \quad (72)$$

Similarly, we define a new function

$$\chi_{\pm}(t) = \phi_{\pm}(u)A(t) \quad (73)$$

and find that

$$\chi_{\pm}(t) = S_{\pm}(t) + \int_{-1}^{\infty} dt' \Gamma_{\pm}(t, t')\chi_{\pm}(t'), \quad (74)$$

where

$$S_{\pm}(t) = 4i\sigma_1 \cos(\sigma_1 \sin\theta)\Lambda_{\pm}(u, i\sigma_1 \sin\theta)A(t), \quad (75)$$

and

$$\Gamma_{\pm}(t, t') = A(t)B(t')(1 + e^{-2v})\Lambda_{\pm}(u, v). \quad (76)$$

In Eqs. (23) and (24), use was made of the function

$$\Lambda_{\pm}(u, v) = \frac{i}{\pi} \sum_{r=1}^{\infty} \beta_{\pm}(\tau_r)(\tau_r^2 + u^2)^{-1}(\tau_r^2 + v^2)^{-1}. \quad (77)$$

Again, in case $\sigma_1 = \tau_N$, the pole at $r = N$ in the source function $S_{\pm}(t)$ is canceled out by the factor $A(t)$.

Our stated goal of transforming out the oscillatory behavior in the integral equation caused by $\cos k_y b$ has been achieved. In terms of $\chi_{\pm}(t')$, one finds from Eq. (60) that

$$g_{\pm}(\tau_n) = bS_n + \int_{-1}^{\infty} dt' \left(\frac{1 + e^{-2v}}{\tau_n^2 + v^2}\right) B(t')\chi_{\pm}(t'). \quad (78)$$

Similarly, one finds, from Eqs. (57), (59), (61), and (78), that

$$I_{\pm}^g(k_y) = i\pi b \cos(k_y b)\phi_{\pm}(ik_y b), \quad (79)$$

with

$$\phi_{\pm}(u) = 4i\sigma_1 \cos(\sigma_1 \sin\theta)\Lambda_{\pm}(u, i\sigma_1 \sin\theta) + \int_{-1}^{\infty} dt' \Lambda_{\pm}(u, v)(1 + e^{-2v})B(t')\chi_{\pm}(t'), \quad (80)$$

defining $\phi_{\pm}(u)$ at all points in the complex u plane. The simple poles at $u = \pm i\tau_r$ are canceled out in $I^g_{\pm}(k_y)$ by the cos $(k_y b)$ factor.

D. Evaluation of $\Lambda_{\pm}(u, v)$

For fixed (u, v), the terms in Eq. (77) decrease like r^{-3} as $r \to \infty$. For large $|u|$ and $|v|$ lying on the contour of integration, $\Lambda_{\pm}(u, v)$ approaches

$$\Lambda^s_{\pm}(u, v) = \frac{L^c_{\pm}}{2\pi^2}(u^2 - v^2)^{-1} \ln\left(\frac{u^2 + U_b}{v^2 + U_b}\right), \quad (81)$$

where

$$L^c_{\pm} = 1, \quad L > 0, \qquad = 1 \pm 1, \quad L = 0, \quad (82)$$

owing to the fact that $\beta_{\pm} \to -(1 \pm 1)i\tau_r$ or $-i\tau_r$ for $L = 0$ or $L > 0$, respectively. The constant $U_b >$ Re σ_1^2 is introduced for numerical convenience.

One can now write

$$\Lambda_{\pm}(u, v) = \Lambda^s_{\pm}(u, v) + \Lambda^F_{\pm}(u, v), \quad (83)$$

where

$$\Lambda^F_{\pm}(u, v) = \frac{i}{\pi}\beta_{\pm}(\tau_p)(\tau_p^2 + u^2)^{-1}(\tau_p^2 + v^2)^{-1} + (u^2 - v^2)^{-1}[G_{\pm}(u) - G_{\pm}(v)] \quad (84)$$

and

$$G_{\pm}(u) = \frac{iu^2}{\pi}\sum_{r=1}^{N-1}{}'\beta_{\pm}(\tau_r)[\tau_r^2(\tau_r^2 + u^2)]^{-1} + L^c_{\pm}\left\{G_N(u) - \frac{1}{2\pi^2}\ln\left[1 + \left(\frac{u^2}{U_b}\right)\right]\right\}. \quad (85)$$

In Eq. (85), the prime denotes exclusion of the term $r = p$ corresponding to the first term of Eq. (84) and deals effectively with the situation where $\sigma_1 \cong \tau_p$. The integer N is chosen so that max $(|\sigma_1|, |\sigma_3|) \ll \tau_N$ simultaneous with

$$|e^{i\zeta_N L}| \ll 1, \quad L > 0, \quad (86)$$

unless $L \equiv 0$, in which case

$$\beta_{\pm}(\tau_r) = (1 \pm 1)\tau_r^2(\sigma_1^2 - \tau_r^2)^{-\frac{1}{2}}. \quad (87)$$

Finally,

$$G_N(u) = (2u^2/\pi)\sum_{r=N}^{\infty}[(\tau_r^2 - \sigma_1^2)^{\frac{1}{2}} + (\tau_r^2 - \sigma_3^2)^{\frac{1}{2}}]^{-1} \times (\tau_r^2 + u^2)^{-1}. \quad (88)$$

If $L \equiv 0$, we take $\sigma_1 = \sigma_3$ in Eq. (88) for the aperture problem.

The last step is to evaluate $G_N(u)$ and the most efficient way to do this is to use the Euler summation formula,[10] since the integration is elementary and the remainder terms can be evaluated to any desired degree of accuracy. The results of this operation are listed below and depend upon $\sigma_1 = \sigma_3$ or $\neq \sigma_3$:

(1) $\sigma_1 = \sigma_3$: vacuum insert:

$$G_N(u) = \frac{u}{2\pi^2 T_1}\left\{\ln\left[1 + \left(\frac{u}{\tau_N}\right)^2\right] + 2\ln\left(\frac{u + T_1}{y_1 u + T_1}\right)\right\} + \Delta G_N, \quad (89)$$

where

$$\Delta G_N(u) = 2(\tau_N y_1 \pi)^{-2}\mathfrak{L}_1(C, u); \quad (90)$$

(2) $\sigma_1 \neq \sigma_3$:

$$G_N(u) = [\pi^2(\sigma_1^2 - \sigma_3^2)]^{-1}[Q_1(u) - Q_3(u)] + \Delta G_N, \quad (91)$$

where

$$Q_r = 2u^2 \ln[(1 + y_r)/2] + uT_r\left\{\ln\left[1 + \left(\frac{u}{\tau_N}\right)^2\right] + 2\ln\left(\frac{u + T_1}{y_1 u + T_1}\right)\right\} \quad (92)$$

and

$$\Delta G_N(u) = u^2(\sigma_1^2 - \sigma_3^2)^{-1}[\mathfrak{L}_1(C, u) - \mathfrak{L}_3(C, u)]. \quad (93)$$

In the above equations,

$$y_r = [1 - (\sigma_r/\tau_N)^2]^{\frac{1}{2}}, \quad T_r = (\sigma_r^2 + u^2)^{\frac{1}{2}}. \quad (94)$$

The functions $\mathfrak{L}_r(C, u)$ in Eqs. (90) and (93) arise from the Euler-summation remainder terms given by

$$\mathfrak{L}_r(C, u) = \frac{F_{1r}}{N_1} - \frac{F_{2r}}{3N_1^2} + \frac{3F_{3r} + 2F_{4r}}{90N_1^4} - \frac{15F_{4r} + 20F_{5r} + 4F_{6r}}{1920N_1^6} + O(N_1^{-8}), \quad (95)$$

with

$$N_1 = N - \tfrac{1}{2} \quad (96)$$

and

$$F_{sr} = \sum_{v=1}^{s} C_{sv}(y_r)^{1+2(v-s)}\left[1 + \left(\frac{u}{\tau_N}\right)^2\right]^{-v}. \quad (97)$$

The coefficients C_{sv} appearing in Eq. (93) are given in Table I for the two cases $\sigma_1 = \sigma_3$ and $\sigma_1 \neq \sigma_3$.

TABLE I. Coefficients C_{sv} appearing in Eq. (97).

$\sigma_1 \neq \sigma_3$	$C_{11} = 1$; $C_{21} = \frac{1}{2}$, $C_{22} = -1$; $C_{31} = -\frac{1}{4}$, $C_{32} = -1$; $C_{33} = 2$; $C_{41} = \frac{3}{8}$, $C_{42} = \frac{3}{4}$, $C_{43} = 3$, $C_{44} = -6$; $C_{51} = -\frac{15}{16}$, $C_{52} = -\frac{3}{2}$, $C_{53} = -3$, $C_{54} = -12$, $C_{55} = 24$; $C_{61} = \frac{105}{32}$, $C_{62} = \frac{75}{16}$, $C_{63} = \frac{15}{2}$, $C_{64} = 15$, $C_{65} = 60$, $C_{66} = -120$.
$\sigma_1 = \sigma_3$	$C_{11} = 1$; $C_{21} = -\frac{1}{2}$, $C_{22} = -1$; $C_{31} = \frac{3}{4}$, $C_{32} = 1$, $C_{33} = 2$; $C_{41} = -\frac{9}{8}$, $C_{42} = -\frac{9}{4}$, $C_{43} = -3$, $C_{44} = -6$; $C_{51} = \frac{105}{16}$, $C_{52} = \frac{15}{2}$, $C_{53} = 9$, $C_{54} = 12$, $C_{55} = 24$; $C_{61} = -\frac{945}{32}$, $C_{62} = -\frac{525}{16}$, $C_{63} = -\frac{75}{2}$, $C_{64} = -45$, $C_{65} = -60$, $C_{66} = -120$.

E. Behavior of $\Gamma_{\pm}(t, t')$

The kernel $\Gamma_{\pm}$ in our integral equation for $\chi_{\pm}(t)$ [Eq. (74)] can now be examined from the analytic representations given in the preceding section. With the behavior for $|u| \gg \tau_N$ in mind, one can rewrite

$$\Gamma_{\pm} = \Gamma_{\pm}^{s} + \Gamma_{\pm}^{F}, \tag{98}$$

where

$$\Gamma_{\pm}^{s}(t, t') = \Omega(t')\Lambda_{\pm}^{s}(u, v), \tag{99}$$

$$\Gamma_{\pm}^{F}(t, t') = [A(t)B(t')(1 + e^{-2v}) - \Omega(t')]\Lambda_{\pm}^{s}(u, v) + A(t)B(t')(1 + e^{-2v})\Lambda_{\pm}^{F}(u, v), \tag{100}$$

and

$$\Omega(t') = 2v_0(\sigma_1^2 + v_a^2)^{\frac{1}{2}}, \quad t' \leq 0, \qquad = 2v_0(\sigma_1^2 + v^2)^{\frac{1}{2}}, \quad t' \geq 0. \tag{101}$$

From Eqs. (85), (89), and (91), one notes that

$$G_{\pm}(u) = G_{\pm}^{0} - L_{\pm}^{c}[\tau_N/(2\pi u)]^{-2} \ln [1 + (u/\tau_N)^2] + O(u^{-2}), \quad |u| \gg \tau_N, \tag{102}$$

where $G_{\pm}^{0}$ is some constant. For $(t, t') \geq 0$, $A(t)B(t') = \Omega(t')$ and it follows from Eqs. (84), (100), and (102) that $\Gamma_{\pm}^{F}(t, t')$ is a Hilbert–Schmidt kernel, i.e.,

$$\int_{-1}^{\infty} dt \int_{-1}^{\infty} dt' \, |\Gamma_{\pm}^{F}(t, t')|^2 < \infty. \tag{103}$$

On the other hand, $\Gamma_{\pm}^{s}(t, t')$ is not Hilbert–Schmidt, but possesses an L_1 norm

$$\int_{-1}^{\infty} dt' \, |\Gamma_{\pm}^{s}(t, t')| < \tfrac{1}{4} L_{\pm}^{c}, \tag{104}$$

a result substantiated by numerical studies. Our procedure for solving Eq. (74) numerically is discussed in the Appendix.

F. Field Distributions and Conservation Laws

Using $g_{\pm}(\tau_n)$ determined by using the previously described Gaussian quadrature to evaluate the integral in Eq. (78), one obtains the field distribution in the guide from Eqs. (49) and (57), i.e.,

$$E_x(y, z) = \tfrac{1}{2}[\tilde{B}_{+}^{g}(y, z) + \tilde{B}_{-}^{g}(y, z)], \tag{105}$$

where

$$\tilde{B}_{\pm}^{g} = \sum_{r=1}^{\infty} J_{\pm}(\tau_r) \cos\left(\frac{\tau_r y}{b}\right)(e^{i\zeta_r z} \pm e^{i\zeta_r(L-z)}), \tag{106}$$

with

$$J_{\pm}(\tau_r) = ig_{\pm}(\tau_r)(-1)^r \tau_r \{b\zeta_r + (\sigma_1^2 - \tau_r^2)^{\frac{1}{2}} \pm e^{i\zeta_r L}[-b\zeta_r + (\sigma_1^2 - \tau_r^2)^{\frac{1}{2}}]\}^{-1} \tag{107}$$

and

$$b\zeta_r = (\sigma_3^2 - \tau_r^2)^{\frac{1}{2}} \operatorname{sgn} (\operatorname{Im} k_3), \tag{108}$$

so that

$$\operatorname{Im} b\zeta_r > 0. \tag{109}$$

Similarly, from Eqs. (79), (80), (29), and (42), one can evaluate the transmitted field

$$E_x(y, z \geq L) = -\frac{1}{4\pi i} \int_{-\infty}^{\infty} dk_y e^{-ik_y y} e^{i\alpha_1(z-L)}[I_{+}^{g}(k_y) - I_{-}^{g}(k_y)]. \tag{110}$$

Finally, it is also possible to show that the field for $z < 0$ is given by

$$E_x(y, z \leq 0) = [\exp(ik_1 z \cos\theta) - \exp(-ik_1 z \cos\theta)] \cos(k_1 y \sin\theta) - \frac{1}{4\pi i} \int_{-\infty}^{\infty} dk_y e^{-ik_y y} e^{-i\alpha_1 z}[I_{+}^{g}(k_y) + I_{-}^{g}(k_y)]. \tag{111}$$

The first term in Eq. (111) is the incident field, the second term arises from specular reflection from the infinitely conducting surface at $z = 0$, and the last term is a cylindrically diffracted wave, propagating away from the entrance surface in the negative z direction.

One should note that Eqs. (110) and (111) give expressions for the fields at the entrance and exit surfaces $z = 0$ and L, respectively, that are identical to the ones we obtain from Eqs. (105) and (106). (This feature was built in the original formalism.) Moreover, Eqs. (110) and (111) give $E_x(y, 0) = E_x(y, L) \equiv 0$, for $|y| \geq b$, since $I_{\pm}^{g}(k_y)$ is analytic and one can close the contours in the lower half k_y plane since $\exp(-ik_y y)$ dominates $\cos k_y b$ and $|\phi_{\pm}(v)|$ [cf. Eqs. (56), (59), and (61)] falls off at least as $|v|^{-\frac{3}{2}}$ for large v.

1. Conservation of Flux

The total electromagnetic flux crossing any plane $z < 0$ per unit incident flux (at the entrance surface) is given, following Levine and Schwinger,[4] by

$$\mathcal{F}_{-} = -(2\sigma_1)^{-1} \operatorname{Im} \int_{-\infty}^{\infty} dy E_x(y, z) \frac{\partial}{\partial z} \bar{E}_x(y, z). \tag{112}$$

Using Eq. (111), one can show that

$$\mathcal{F}_{-} = \cos(\theta) \operatorname{Im} \frac{[I_{+}^{g}(k_1 \sin\theta) + I_{-}^{g}(k_1 \sin\theta)]}{4b} - \frac{k_1^2}{8\pi\sigma_1} \int_0^{\frac{1}{2}\pi} d\theta' \cos^2(\theta') \times |I_{+}^{g}(k_1 \sin\theta') + I_{-}^{g}(k_1 \sin\theta')|^2. \tag{113}$$

Similarly, from Eq. (110), the total flux crossing any plane $z > L$ per unit incident flux is

$$\mathcal{F}_{+} = \frac{k_1^2}{8\pi\sigma_1} \int_0^{\frac{1}{2}\pi} d\theta' \cos^2(\theta') \times |I_{+}^{g}(k_1 \sin\theta') - I_{-}^{g}(k_1 \sin\theta')|^2. \tag{114}$$

From Maxwell's equation [Eq. (1)], it follows that

$$\mathcal{F}_+ - \mathcal{F}_- \equiv -\text{Im}\,(k_3^2)\sigma_1^{-1}\int_0^b dy\int_0^L dz\,|E_x(y,z)|^2, \tag{115}$$

that is, the total flux balance must be compensated by absorption or generation of flux (photons) in the dielectric insert if Im $k_3 >$ or < 0, respectively. Using Eqs. (106) and (107), one can also show that

$$\mathcal{F}_+ - \mathcal{F}_- = -\text{Im}\,(\sigma_3^2)4\sigma_1^{-1} \times \frac{L}{b}\sum_{r=1}^{\infty} e^{-c_r}\sum_{\nu=-1}^{1}\left[|J_\nu(\tau_r)|^2\left(\frac{\sinh c_r}{c_r} + \nu\frac{\sin d_r}{d_r}\right)\right], \tag{116}$$

where

$$c_r = \text{Im}\,\zeta_r L, \quad d_r = \text{Re}\,\zeta_r L. \tag{117}$$

Equation (116) provides an excellent check on our numerical calculation and will be used as such. The series in Eq. (116) converges fairly rapidly since $d_r = O(r^{-1}), |J_\pm|^2 = O(r^{-\frac{1}{2}})$, and $c_r = O(r)$ as $r \to \infty$.

2. Far Field Radiation Pattern

To evaluate Eq. (110), it is convenient to use cylindrical coordinates

$$y = \rho\sin\theta_0, \quad z = \rho\cos\theta_0 + L, \tag{118}$$

make use of the analyticity of $I_\pm^g(k_y)$, and deform the k_y contour by introducing[12]

$$k_y = -k_1[\sin\theta_0(1 + i\alpha^2) - 2^{\frac{1}{2}}\alpha\cos\theta_0 e^{-\frac{1}{4}i\pi}(1 + \tfrac{1}{2}i\alpha^2)^{\frac{1}{2}}], \tag{119}$$

with $-\infty < \alpha < \infty$. When $-\frac{1}{2}\pi < \theta_0 < \frac{1}{2}\pi$, this contour avoids the branch cuts in α_1. In fact,

$$\alpha_1 = k_1[\cos\theta_0(1 + i\alpha^2) + 2^{\frac{1}{2}}\alpha\sin\theta_0 e^{-\frac{1}{4}i\pi}(1 + \tfrac{1}{2}i\alpha^2)^{\frac{1}{2}}], \tag{120}$$

since $k_1^2 - k_y^2$ is a perfect square and Eq. (110) becomes

$$E_x(y, z \geq L) = -\frac{e^{ik_1\rho}}{4\pi i}\int_{-\infty}^{\infty} d\alpha\,\frac{dk_y}{d\alpha}\,e^{-k_1\rho\alpha^2}[I_+^g(k_y) - I_-^g(k_y)]. \tag{121}$$

Equation (121) is valid for all $\rho > 2b\sin\theta_0 > 0$, since $I_\pm^g(k_y)$ possesses exponential growth for large α like $\exp(2k_1 b\sin\theta_0\alpha^2)$.

If one desires to evaluate the field near the exit surface (with $\rho > 2b$ in θ_0), a Hermite quadrature can be employed. On the other hand, for $k_1\rho \gg 1$, an appeal can be made to Watson's lemma,[13,14] allowing one to obtain an asymptotic expansion in powers of $(k_1\rho)^{-\frac{1}{2}n}$, $n = 1, 2, \cdots$. The first term gives good approximation for the far field limit

$$E_x(y, z > L) \cong -(k_1/2\pi\rho)^{\frac{1}{2}}e^{ik_1\rho}e^{\frac{1}{4}i\pi}f_T(\theta_0), \tag{122}$$

where

$$f_T(\theta_0) = -\tfrac{1}{2}\cos(\theta_0)[I_+^g(k_1\sin\theta_0) - I_-^g(k_1\sin\theta_0)] \tag{123}$$

will be referred to as the forward scattering amplitude from the *exit* surface. Our definition of $f_T(\theta_0)$ coincides with Keller's[5] for the case $L = 0$. For this case, $\mathcal{F}_+ \equiv \text{Im}\,f_T(\theta_0)/2b$ according to Eqs. (112)–(115) in agreement with the well-known cross-section theorem of Levine and Schwinger.[3–5]

Similarly, the field distribution reflected from the extrance surface at a point

$$y = \rho\sin\theta, \quad z = -\rho\cos\theta, \tag{124}$$

far from the entrance surface, is given by

$$E_x(y, z < 0) \cong -(k_1/2\pi\rho)^{\frac{1}{2}}e^{ik_1\rho}e^{\frac{1}{4}i\pi}f_R(\theta) - \exp(-ik_1 z\cos\theta)\cos(k_1 y\sin\theta), \tag{125}$$

where

$$f_R(\theta) = -\tfrac{1}{2}\cos(\theta)[I_+^g(k_1\sin\theta) + I_-^g(k_1\sin\theta)]. \tag{126}$$

V. NUMERICAL FIELD DISTRIBUTIONS

The equations in the previous section and the procedure outlined in the Appendix for solving Eq. (74) gives $\chi_\pm(t)$ numerically. These results allow us, in turn, to compute $E_x(y, z)$ in the dielectric insert [Eqs. (105), (106), (107), and (78)]. The reflected and transmitted field amplitudes far from the entrance and exit surfaces are also found in a relatively simple manner using Eqs. (122), (123), (125), (126), (79), and (80).

A. $L = 0$: Aperture

A comparison of $E_x(y, 0)$ for the aperture is made in Table II when $\sigma_1 = 2^{\frac{1}{2}}$. Our results are listed under present work obtained by replacing Eq. (74) with an 18×18 matrix as discussed in Appendix A. The analytic result for the field in the aperture, calculated by Hsu[6] using the Mathieu function expansion, is given in the 4th and 5th columns. An approximate field distribution derived by Tan[7] for this case is quite good, but the approximation introduced by Karp and Russek[8] is rather poor near the aperture $y = b$ for long wavelengths.

When $\sigma_1 \gg 1$, our numerical calculation for the far field $f_T(\theta)$ reinforce Keller's estimate[5] of the validity of his approximate approach. For example, the f_T's calculated by both methods disagree in the third significant figure for $\sigma_1 = 10\pi$.

TABLE II. Comparison of $E_x(y, 0)$ for $L = 0$ [aperture] with other results when $\sigma_1 = 2^{\frac{1}{2}}$. Equation (74) replaced by 18 × 18 matrix. [Re and Im E_x listed as a pair below.]

y/b	Present work		Hsu (Mathieu)[6]		Tan's Eq. 19[7]		Karp and Russek[8]	
0	1.5222	−0.9367	1.522	−0.938	1.524	−0.947	1.825	−0.861
0.1736	1.4886	−0.9196	1.489	−0.921	1.490	−0.929	1.791	−0.863
0.3417	1.3918	−0.8696	1.392	−0.871	1.392	−0.878	1.691	−0.870
0.5	1.2425	−0.7901	1.243	−0.790	1.242	−0.798	1.533	−0.890
0.6431	1.0562	−0.6866	1.057	−0.688	1.054	−0.693	1.319	−0.940
0.7667	0.8486	−0.5652	0.850	−0.566	0.845	−0.569	1.045	−1.048
0.8667	0.6334	−0.4318	0.635	−0.433	0.627	−0.433	0.675	−1.282
0.9028	0.5374	−0.3697						
0.9847	0.2107	−0.1483	0.209	−0.146	0.199	−0.136	−1.490	−3.680

B. $L > 0$: Vacuum Guide

When $\sigma_1 = \sigma_3$ and $L > 0$, one has an extended slit. For purposes of illustration, we tabulate $|E_x(y, 0)|^2$ for $\sigma_1 = \sigma_3 = 10\pi$, with $b = 5\lambda$, λ being the wavelength, for $L = 0$, $0.09327b$, and $10b$ in Table III. Equation (74) is replaced by a 30 × 30 matrix, giving the results shown in the second column. The remaining three cases were obtained by replacing Eq. (74) with an 18 × 18 matrix. A comparison of the 2nd and 3rd columns shows a difference appearing in the 4th decimal place. The last two columns indicate that $|E_x(y, 0)|^2$ is nearly independent of L for the cases under consideration.

C. $L > 0$: Dielectric Insert

Unlike the vacuum guide, a guide with a dielectric insert can be made to exhibit a variety of phenomena depending upon the values of $k_3 b = \sigma_3$ as well as σ_1 and L. In Fig. 3, $|E_x, (y, z)|^2$ is plotted for $z = 0$ and L as a function of y/b when $\sigma_1 = \pi$, $\sigma_3 = 2\pi + 0.001i$, and $L = 20b$, with $b = \frac{1}{2}\lambda_{\text{vac}}$. The angle of incidence for the symmetrized incoming plane wave is $\theta_I = 31°$. When the vacuum wavelength is halved, one notes a striking increase and localization of $|E_x(y, L)|^2$ near $y = 0.8b$ as shown in Fig. 4. The presence of two nodes corresponds to the fact that the propagating eigenfunctions $\exp(i\zeta_r z) \pm \exp[i\zeta_r(L - z)]$ in Eq. (106) are strongly damped for $0 < z < L$ when $r \geq 3$.

TABLE III. Comparison of $|E_x(y, 0)|^2$ for $\sigma_1 = \sigma_3 = 10\pi$, $b = 5\lambda$, for $L = 0$ (aperture), $0.09327b$, and $10b$.

y/b	$L = 0$[a]	$L = 0$	$L = 0.09327b$	$L = 10b$
0	0.7499	0.7502	0.7488	0.7338
0.1	1.3269	1.3269	1.3568	1.3519
0.2	0.7464	0.7467	0.7454	0.7292
0.3	1.3392	1.3392	1.3702	1.3643
0.4	0.7344	0.7348	0.7340	0.7138
0.5	1.3699	1.3699	1.4035	1.4022
0.6	0.7078	0.7083	0.7083	0.6875
0.7	1.4425	1.4425	1.4824	1.4697
0.8	0.6406	0.6413	0.6398	0.6205
0.9	1.7195	1.7187	1.7985	1.8139

[a] Gaussian quadrature of order $M = 10$, i.e., Eq. (74) replaced by 30 × 30 matrix. $M = 6$ for the other three cases.

The far field intensities for the above-mentioned cases are given in terms of $|R|^2 = |2f_R(\theta)/\pi b|^2$ [Eqs. (125) and (126)] and $|T|^2 = |2f_T(\theta)/\pi b|^2$ [Eqs. (122) and (123)]. If one considers the case of a *negatively* absorbing media with $\sigma_3 = 4\pi - 0.01i$, then a striking increase in both $|R|^2$ and $|T|^2$ are found, as Fig. 5 indicates. The peaks in $|R|^2$ and $|T|^2$ near $\theta = 39°$ corresponds to a slightly shifted maximum expected on the basis of geometric optics since $\theta_I = 30°$ here. The maximum in the transmitted field at $\theta = 0$ is anomalous from a geometric optics point of view and results from two overlapping sidelobes. A very small value of $|R|^2$ obtains near $\theta = 8°$. When the wavelength is again halved, more nodes result in $|R|^2$ and $|T|^2$, as Fig. 6 indicates.

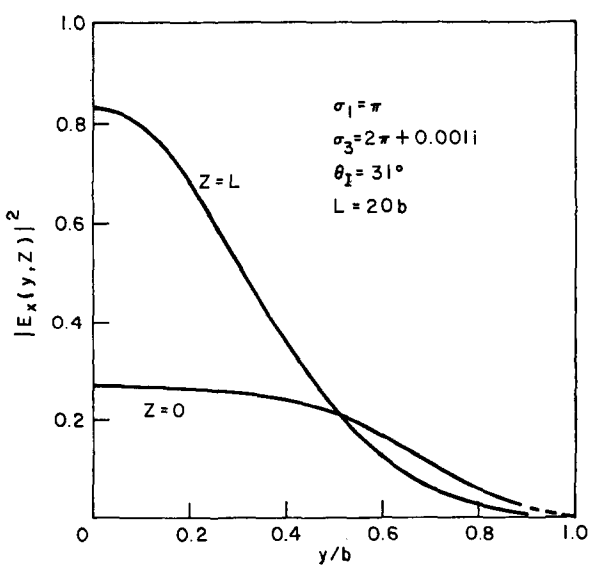


FIG. 3. Electric field intensity $|E_x(y, z)|^2$, for $z = 0$, L as a function of y/b for parameters indicated on figure. Symmetrized-plane-wave input, with angle of incidence given by θ_I.

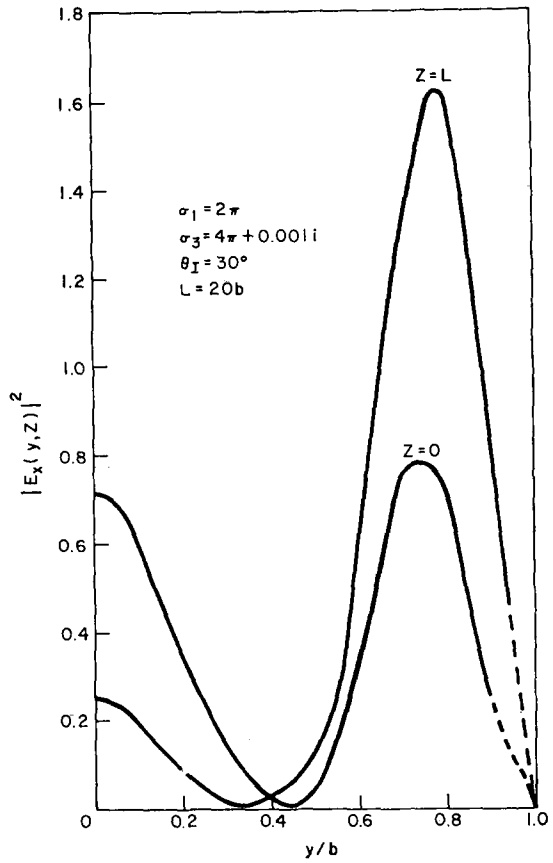


FIG. 4. Electric field intensity $|E_x(y, z)|^2$, for $z = 0, L$ as a function of y/b for parameters indicated on figure. Symmetrized-plane-wave input with angle of incidence given by θ_I.

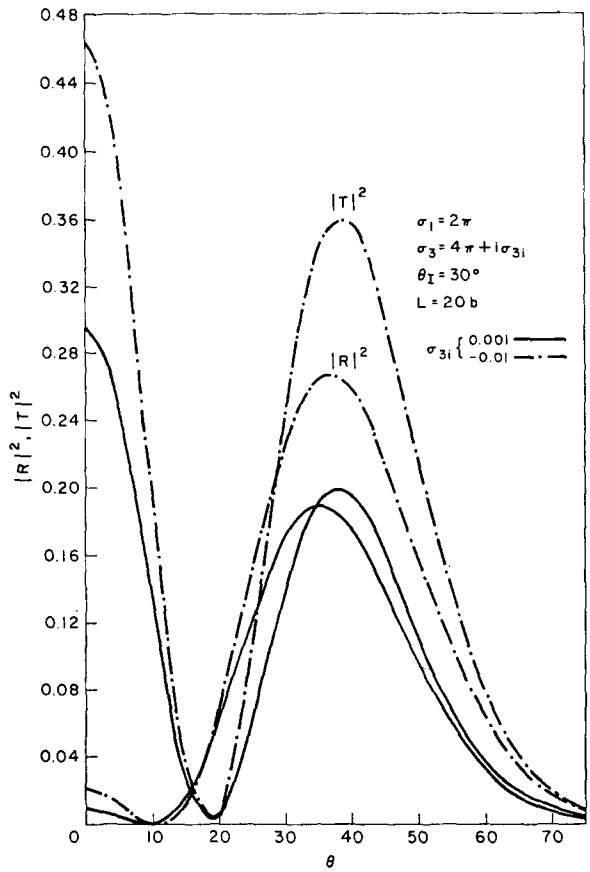


FIG. 5. Spectral-reflection- and transmission-intensity coefficients. $|A|^2 = |2f_A(\theta)/\pi b|^2$, where $A = R$ or T for reflection or transmission; cf. Eqs. (125) and (126) or (122) and (123) of text. Symmetrized-plane-wave input with angle of incidence given by θ_I.

Figures 7 and 8 depict the field distributions for a case where the denominator of Eq. (107) is small for $r = 1$, i.e.,

$$b\zeta_{s1} + (\sigma_1^2 - \tau_1^2)^{\frac{1}{2}} - e^{i\zeta_{s1}L}[(\sigma_1^2 - \tau_1^2)^{\frac{1}{2}} - b\zeta_{s1}] \cong 0. \tag{127}$$

If Eq. (127) above is identically zero, then $g_-(\tau_1) \equiv 0$ in Eq. (107) and one has an exact resonant excitation. As Fig. 8 shows, the transmitted-beam spectral intensity for a material characterized by a homogeneous negative absorbtion, Im $\sigma_3 = -0.1$, is only 2% of that corresponding to a nonabsorbing media. This situation is in marked contrast to that depicted in Fig. 5 where an enhancement in $|T|^2$ was found for Im $\sigma_3 = -0.01$.

VI. DISCUSSION AND CONCLUSIONS

A theoretical formalism has been developed in this paper for solving a class of electromagnetic diffraction problems involving heterogeneous planar dielectric structures shown in Figs. 1 and 2. From a numerical

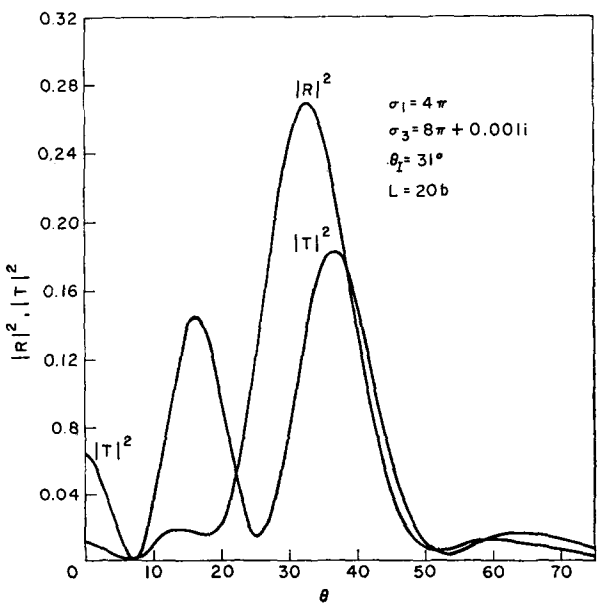


FIG. 6. Spectral-reflection- and transmission-intensity coefficients. $|A|^2 = |2f_A(\theta)/\pi b|^2$, where $A = R$ or T for reflection or transmission; cf. Eqs. (125) and (126) or (122) and (123) of text. Symmetrized-plane-wave input with angle of incidence given by θ_I.

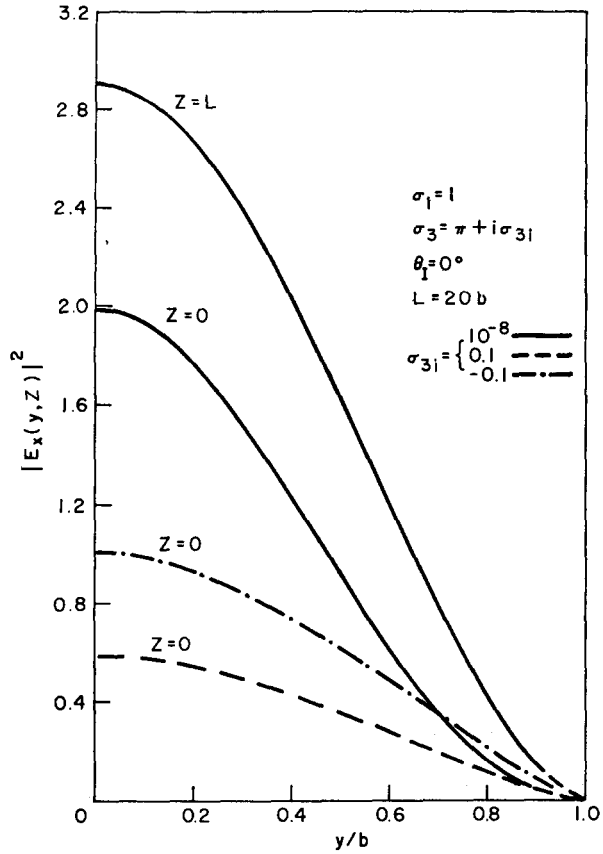


FIG. 7. Electric field intensity $|E_x(y, z)|^2$, for $z = 0, L$ as a function of y/b for parameters indicated on figure. Symmetrized-plane-wave input with angle of incidence given by θ_I.

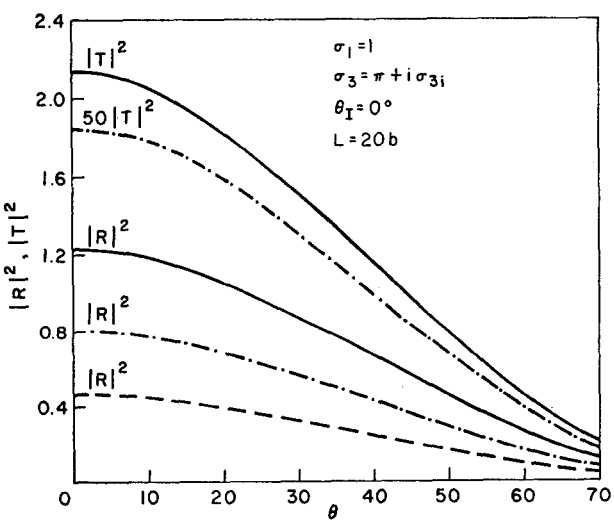


FIG. 8. Spectral-reflection- and transmission-intensity coefficients. $|A|^2 = |2f_A(\theta)/\pi b|^2$, where $A = R$ or T for reflection or transmission; cf. Eqs. (125) and (126) or (122) and (123) of text. Symmetrized-plane-wave input with angle of incidence given by θ_I.

viewpoint, it was possible to calculate field distributions in a metal–dielectric guide assembly covering the Rayleigh to geometric optics limit when the exciting electric field was polarized in a transverse fashion $[E_x(y, z)]$. Typical numerical results were presented for several geometries.

It appears possible to extend the approach used here to treat a variety of metal–dielectric guide problems. In particular, when k_3 corresponds to a layered rather than a homogeneous medium as shown in Fig. 2, one may find radiating eigenmodes, provided that a portion of the dielectric media is optically pumped to provide a negatively absorbing media. At this writing, it is unclear whether a negatively absorbing homogeneous k_3 region will give such eigenmodes. These points are currently being investigated.

ACKNOWLEDGMENTS

The author is indebted to W. F. Hall, E. A. Kraut, and J. Spanier for several discussions concerning this paper.

APPENDIX: NUMERICAL SOLUTION FOR $\chi_\pm(t)$

The results of the analysis in Sec. IV establish a procedure to follow in solving Eq. (74). For numerical convenience, one can divide $-1 \leq t < \infty$ into three parts:

$$\begin{aligned} t &= \tfrac{1}{2}(\mu - 1), \qquad -1 \leq t \leq 0, \\ &= \tfrac{1}{2}U_a(1 + \mu), \qquad 0 \leq t \leq U_a, \\ &= 2U_a(1 - \mu)^{-1}, \quad U_a \leq t < \infty, \end{aligned} \tag{A1}$$

with

$$-1 \leq \mu \leq 1. \tag{A2}$$

The above division of the t interval allows us to use a single Gaussian quadrature for each of three regions. Hence, Eq. (74) is approximated by a set of algebraic equations

$$\chi_j = S_j + \sum_p w_p(\Gamma^s_{jp} + \Gamma^F_{jp})\chi_p, \tag{A3}$$

where $\chi_j = \chi_\pm(t_j)$, etc. The w_p are standard Gaussian weight factors times $(dt/d\mu)_{\mu_p}$ and t_j are defined by Eq. (A1) in terms of the μ_j, the zeros of the Mth-order Legendre polynomial. If

$$\chi_j = \omega_j \eta_j, \tag{A4}$$

$$\omega_j = (w_j \Omega_j)^{\frac{1}{2}}, \tag{A5}$$

$$G^s_{jp} = \omega_j \omega_p A^s_\pm(u_j, u_p), \tag{A6}$$

$$G^F_{jp} = \omega_j \omega_p \Gamma^F_\pm(t_j, t_p)/\Omega_p, \tag{A7}$$

and

$$f_j = \omega_j S_j, \tag{A8}$$

then, in matrix notation,

$$\boldsymbol{\eta} = \mathbf{f} + \mathbf{G}^s\boldsymbol{\eta} + \mathbf{G}^F\boldsymbol{\eta}, \tag{A9}$$

where $\boldsymbol{\eta}$ and $\mathbf{f}$ are column vectors whose $3M$ components are η_j and f_j, respectively. The elements of the $\mathbf{G}^s$ and $\mathbf{G}^F$ matrices are listed in Eq. (A6) and (A7).

Equation (104) guarantees the existence of

$$\mathcal{R}^s = (\mathbf{I} - \mathbf{G}^s)^{-1}\mathbf{G}^s \tag{A10}$$

so that

$$\boldsymbol{\eta} = \mathbf{f} + \mathcal{R}^s\mathbf{f} + \tilde{G}^F\boldsymbol{\eta}, \tag{A11}$$

where

$$\tilde{\mathbf{G}}^F = \mathbf{G}^F + \mathcal{R}^s\mathbf{G}^F. \tag{A12}$$

Since the norm of $\mathbf{G}^s$ is less than 0.5, one can construct

$$\mathcal{R}^s = \lim_{r\to\infty} \mathcal{R}^s_r, \tag{A13}$$

where

$$\mathcal{R}^s_{r+1} = \mathcal{R}^s_r + (\mathbf{G}^s)^{2^r}\mathcal{R}^s_r. \tag{A14}$$

Starting with

$$\mathcal{R}^s_1 = \mathbf{G}^s + [\mathbf{G}^s]^2, \tag{A15}$$

one readily sees that

$$\mathcal{R}^s_r \equiv \sum_{\nu=1}^{2^r} [\mathbf{G}^s]^\nu \tag{A16}$$

which is the same as the power series expansion of Eq. (A10) in the limit $r \to \infty$. The error in generating an approximation to $\mathcal{R}^s$ by $\mathcal{R}^s_6$ as measured by its norm is of the order of or less than $(16)^{-16}$. Obviously, in less than 6 cycles, one has an excellent approximation to $\mathcal{R}^s$ with trivial rounded off errors. To achieve this, one must, of course, replace $[\mathbf{G}^s]^{2^r}$ by its square $[\mathbf{G}^s]^{2^{r+1}}$ at each step.

Our final numerical step is to solve Eq. (A11) by use of the Fredholm construction[15] which is equivalent to the so-called Frame's method[16] of matrix inversion. One finds that

$$(\mathbf{I} - \mathbf{G}^F)^{-1} = \mathbf{I} - \mathbf{D}/\Delta_0, \tag{A17}$$

where

$$\mathbf{D} = \sum_{r=1}^{3M-1} \mathbf{D}_r, \tag{A18}$$

$$\Delta_0 = 1 + \sum_{r=1}^{3M-1} d_r, \tag{A19}$$

$$\mathbf{D}_1 = -\mathbf{G}^F, \tag{A20}$$

$$\mathbf{D}_{r+1} = \mathbf{G}^F\mathbf{D}_r - d_r\mathbf{G}^F, \tag{A21}$$

and

$$rd_r = \mathrm{Tr}\,(\mathbf{D}_r). \tag{A22}$$

In Eq. (A22), Tr means trace or diagonal sum.

Numerous machine calculations using the CDC 6600 have been carried out for $\sigma_1 = k_1 b$ ranging from 0.1 to 10π for $L = 0$, $0.0935b$, b, $10b$, and $20b$. In all cases investigated to date, for $M = 6$ and 10, i.e., Gaussian quadratures in each of the three regions of order 6 and 10, we find that $|d_8| < 10^{-9}$ with similar bounds on the elements of $\mathbf{D}_8$. This means that the Fredholm construction converges extremely rapidly, living up to its theoretical expectations.

In our machine calculations, we took the solution for $\boldsymbol{\eta}$ of Eq. (A11) obtained by using Eq. (A17) and substituted it into the right side of Eq. (A9). With an 8-place floating-point output, the *equality* in Eq. (A9) held exactly, attesting to the accuracy of our $\mathcal{R}^s$ and $(\mathbf{I} - \mathbf{G}^F)^{-1}$ construction and, of course, the high single precision accuracy of the CDC 6600.

[1] M. Born and E. Wolf, *Principles of Optics* (Pergamon Press, New York, 1959).

[2] B. Noble, *Wiener–Hopf Techniques* (Pergamon Press, New York, 1958).

[3] For a review prior to 1954, see C. J. Bouwkamp, Rept. Progr. Phys. **17**, 35 (1954).

[4] H. Levine and J. Schwinger, Phys. Rev. **74**, 958 (1948).

[5] For an important approximate treatment of diffraction by an aperture see J. B. Keller, J. Appl. Phys. **28**, 426 (1957).

[6] H. P. Hsu, Case Institute of Technology, Cleveland, Ohio, Scientific Report No. 5, AF/19/604/3887, 1959.

[7] H. S. Tan, Aust. J. Phys. **21**, 35 (1968).

[8] S. N. Karp and A. Russek, J. Appl. Phys. **27**, 886 (1956).

[9] I. N. Sneddon, *Fourier Transforms* (McGraw-Hill Book Co., New York, 1951).

[10] E. T. Whittaker and G. N. Watson, *A Course in Modern Analysis* (Cambridge University Press, Cambridge, 1963), 4th ed.

[11] In the numerical calculations reported in this paper, $v_0 = 1$.

[12] Our contour is a generalization of that used by P. C. Clemmow, *Plane Wave Spectrum of Electromagnetic Fields* (Pergamon Press, New York, 1966), p. 47.

[13] G. N. Watson, *Bessel Functions* (Cambridge University Press, Cambridge, 1944), 2nd ed., p. 236.

[14] See Ref. 12, p. 49.

[15] R. G. Newton, *Scattering Theory of Waves and Particles* (McGraw-Hill Book Co., New York, 1966), p. 253.

[16] A. E. Householder, *Principles of Numerical Analysis* (McGraw-Hill Book Co., New York, 1953), Secs. 4.2 and 4.3.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 5 MAY 1970

Kronecker Products and Symmetrized Squares of Irreducible Representations of Space Groups

C. J. BRADLEY AND B. L. DAVIES*
Department of Metallurgy, University of Oxford, Oxford, England

(Received 21 May 1969)

Application is made to space groups of a theorem by Mackey on the symmetrized and antisymmetrized squares of induced representations. A new and constructive proof is given of Mackey's theorem which enables the representations appearing in the decomposition of the symmetrized squares to be easily identified. The method is used to extend Birman's tables for T_d^2, the zinc-blende structure, to include products involving its double-valued representations

1. INTRODUCTION

This paper is a development of work by one of the authors,[1] and has similar aims in mind: to make more accessible to theoretical physicists the results of Mackey[2,3] on induced representations, and to show how to adapt his results so that they can be developed constructively in application to the theory of the crystallographic space groups.

The previous paper,[1] referred to hereafter as I, covered the problem of the Clebsch–Gordan decomposition of the Kronecker products of irreducible representations of space groups, and in so doing made considerable use of results in the first of the above papers by Mackey.[2] The present paper is devoted mainly to the problem of separating the decomposition of the Kronecker square of an irreducible representation of a space group into its symmetrized and antisymmetrized parts, and the theorem for solving this, for which we give a new and simplified proof, originated from the second of the above papers by Mackey.[3] As in I, the style of presentation and mode of proof is organized so as to be easily understood by theoretical physicists who have attended nothing more advanced than an introductory course on the representation theory of finite groups. However, it is advisable for the reader who wishes to follow in detail the mathematics of Secs. 3 and 4 to assimilate first of all the results contained in I. Someone who is interested only in the tabulations for zinc-blende and not in how they have been obtained mathematically will wish to turn immediately to Sec. 5. In Sec. 2, we show briefly how the physical problem we are treating leads inevitably to the mathematical problems described above, and we give a short review of the background work which it is necessary to know before tackling the main problem. The following section contains a new and simplified proof of Mackey's theorem on the symmetrized and antisymmetrized squares of induced representations. This part of the work is framed in terms of abstract finite groups, and it is hoped thereby to provide useful results for physicists working in areas other than in the dynamics of crystals. The advantage of the proof we present in Sec. 3 is its constructive nature: in particular, a precise formula is obtained for the matrix P of a certain unitary transformation and, whereas the main theorem rests for its validity only on the existence of this transformation, any calculation using the theorem requires its precise form, which we obtain. The application to the crystallographic space groups is covered in Sec. 4. Finally, after giving an example, we tabulate in Sec. 5 the results for the zinc-blende structure. For completeness we include expansions for the Kronecker products of distinct irreducible representations as well as those for the symmetrized squares. We include previous results obtained by Birman[4,5] (with whose results we agree, except for some minor discrepancies involving an ambiguity over the labeling of the representations associated with the point Δ), and we extend his results to cover products involving the double-valued representations. Our reason for including previous results is that in making a complete tabulation we can present the results in a highly abbreviated form which would not be open to us in a partial tabulation. It should also be mentioned that Karavaev[6] has obtained certain of our results; but his paper is untranslated and appears to us to contain too many misprints to be safe to use as an accurate reference.

The method used by Birman[4,5] is very different in approach from the one we employ, being essentially a full-group rather than a little-group procedure. The two methods, since both are correct, lead to identical results, and it seems to be very much a matter of taste which method to use as far as space groups are concerned. However, induced representations now appear in a variety of physical contexts, and it is

because of this possible wide range of application that we are prompted to present this alternative method of solution.

2. BACKGROUND THEORY

Let **G** be a finite group and Δ a UR (unitary representation) of **G** with basis $\langle\psi| = \langle\psi_1, \psi_2, \cdots, \psi_f|$, so that for all $g \in \mathbf{G}$,

$$g\psi_i = \sum_{j=1}^{f} \psi_j \Delta(g)_{ji}. \tag{2.1}$$

Denoting the carrier space of Δ by V, we can consider the tensor product of V with itself, $V \otimes V$, to be the vector space spanned by linear combinations of ordered pairs of functions (ψ_i, ψ_j), $i = 1$ to f, $j = 1$ to f. The space $V \otimes V$ is of dimension f^2 and is invariant under the outer direct-product group $\mathbf{G} \otimes \mathbf{G}$. The diagonal subgroup of $\mathbf{G} \otimes \mathbf{G}$, which we call the inner direct-product $\mathbf{G} \boxtimes \mathbf{G}$, is isomorphic with **G** under the natural mapping $(g, g) \leftrightarrow g$, and by virtue of this we may regard $V \otimes V$ as being invariant under **G** itself; and the UR of **G** so defined is called the Kronecker square of Δ, which we denote by $(\Delta \otimes \Delta)$. With respect to the basis (ψ_i, ψ_j), the explicit form for $(\Delta \otimes \Delta)$ is given by the Kronecker square of the matrix Δ,

$$(\Delta \otimes \Delta)(g)_{kl,ij} = \Delta(g)_{ki}\Delta(g)_{lj}. \tag{2.2}$$

The character of $(\Delta \otimes \Delta)$ in terms of the character of Δ is given by

$$\chi_{(\Delta\otimes\Delta)}(g) = \chi_\Delta^2(g). \tag{2.3}$$

When $f > 1$, the UR $(\Delta \otimes \Delta)$ is reducible over **G**, but a partial reduction can be achieved immediately by separating $V \otimes V$ into its symmetric and antisymmetric parts. These subspaces of $V \otimes V$, which we denote by $[V \otimes V]$ and $\{V \otimes V\}$ respectively, are defined as follows: $[V \otimes V]$ is the vector space of dimension $\frac{1}{2}f(f+1)$ spanned by linear combinations of the symmetrized pairs $(\psi_i, \psi_j) + (\psi_j, \psi_i)$, and $\{V \otimes V\}$ is the vector space of dimension $\frac{1}{2}f(f-1)$ spanned by linear combinations of the antisymmetrized pairs $(\psi_i, \psi_j) - (\psi_j, \psi_i)$. These spaces are invariant under **G**, and the UR's defined on these spaces we denote by $[\Delta \otimes \Delta]$ and $\{\Delta \otimes \Delta\}$, respectively. Expressions analogous to Eqs. (2.2) and (2.3) for the UR $[\Delta \otimes \Delta]$ are

$$[\Delta \otimes \Delta](g)_{kl,ij} = \tfrac{1}{2}(\Delta(g)_{ki}\Delta(g)_{lj} + \Delta(g)_{kj}\Delta(g)_{li}) \tag{2.4}$$

and

$$\chi_{[\Delta\otimes\Delta]}(g) = \tfrac{1}{2}(\chi_\Delta^2(g) + \chi_\Delta(g^2)). \tag{2.5}$$

Expressions for $\{\Delta \otimes \Delta\}$ are obtained by replacing the plus signs on the right of Eqs. (2.4) and (2.5) by minus signs, as, of course, must be the case since

$$(\Delta \otimes \Delta) \equiv [\Delta \otimes \Delta] + \{\Delta \otimes \Delta\}. \tag{2.6}$$

For proofs of the assertions so far made in this section, we refer to Hamermesh,[7] Chap. 5. The UR's $[\Delta \otimes \Delta]$ and $\{\Delta \otimes \Delta\}$ are commonly referred to as the *symmetrized and antisymmetrized squares of* Δ.

Now let **K** be a subgroup of **G** and D a UR of **K** of dimension d, with basis $\langle\phi| = \langle\phi_1, \phi_2, \cdots, \phi_d|$. The carrier space of D we denote by V_1. We are using the same notation as in I, and reproduce the following equations, definitions, and theorems from that paper which are relevant to the present discussion.

The left coset decomposition of **G** relative to **K** is written as

$$\mathbf{G} = \sum_\sigma p_\sigma \mathbf{K}. \tag{2.7}$$

Here we take p_1 to be the identity, and the sum over σ runs from 1 to $|G|/|K|$, where $|G|$ is the order **G** and $|K|$ is the order of **K**.

The vector space spanned by the linear combination of functions $p_\sigma\phi_i$, $\sigma = 1$ to $|G|/|K|$, $i = 1$ to d, is invariant under **G**, and defines the induced representation of D in **G**. We denote this induced representation by $\Delta = D \uparrow \mathbf{G}$. Its precise matrix form is given by Eq. (2.4) of I. For fixed σ, we denote the vector space spanned by the d functions $p_\sigma\phi_i$, $i = 1$ to d, by V_σ; the carrier space V of Δ of dimension $f = d\,|G|/|K|$ may be written symbolically as

$$V = \sum_\sigma V_\sigma.$$

The mathematical problem considered in this paper is the simplification of the expressions for $[\Delta \otimes \Delta]$ and $\{\Delta \otimes \Delta\}$ in the case when $\Delta = D \uparrow \mathbf{G}$. The problem of the simplification of $(\Delta \otimes \Delta)$ was dealt with in I, and the results are now given in summary form.

The double coset decomposition of **G** relative to **K** is written as

$$\mathbf{G} = \sum_\alpha \mathbf{K} d_\alpha \mathbf{K}. \tag{2.8}$$

Here we take d_1 to be the identity, and in the complex $\mathbf{K}d_\alpha\mathbf{K}$ we count any element that appears only once. We define the subgroups $\mathbf{K}_\alpha = d_\alpha\mathbf{K}d_\alpha^{-1}$ and $\mathbf{L}_\alpha = \mathbf{K} \cap \mathbf{K}_\alpha$. We write

$$\mathbf{K} = \sum_\gamma q_{\alpha\gamma}\mathbf{L}_\alpha. \tag{2.9}$$

We also define the UR D_α of $\mathbf{K}_\alpha$ by the equation

$$D_\alpha(d_\alpha k d_\alpha^{-1}) = D(k), \quad \text{for all} \quad k \in \mathbf{K}, \tag{2.10}$$

and we note the important fact that this is the UR of $\mathbf{K}_\alpha$ determined by the vector space V_α. This follows since

$$\begin{aligned}(d_\alpha k d_\alpha^{-1})\phi_{\alpha i} &= d_\alpha k \phi_i \\ &= d_\alpha \sum_{j=1}^{d} \phi_j D(k)_{ji} \\ &= \sum_{j=1}^{d} \phi_{\alpha j} D(k)_{ji}.\end{aligned} \tag{2.11}$$

Since D_α is defined on $\mathbf{K}_\alpha$ and D on $\mathbf{K}$, it is clear that the Kronecker product $(D_\alpha \otimes D)$ is defined on $\mathbf{L}_\alpha$. Bearing this in mind, we can now state one of the main results of I:

$$(D \uparrow G) \otimes (D \uparrow G) \equiv \sum_\alpha ((D_\alpha \otimes D) \downarrow \mathbf{L}_\alpha) \uparrow G. \quad (2.12)$$

In Eq. (2.12) the downward arrow denotes the restriction of $(D_\alpha \otimes D)$ to elements of $\mathbf{L}_\alpha$, and the sum over α runs over the same suffices as the sum in Eq. (2.8).

We turn briefly now to the physical application of these ideas to the study of quantum-mechanical selection rules. Suppose that the Schrödinger group of the system under consideration is a group $\mathbf{G}$ and that its UIR's (unitary irreducible representations) are denoted by Δ^i. Let ψ_r^i denote a wavefunction belonging to row r of the UIR Δ^i, and suppose that W is a self-adjoint operator belonging to the UR Δ. The transition between an initial state ψ_t^k and a final state ψ_r^i under the operation W is governed by the absolute value of the matrix element

$$W_{rt}^{ik} = \langle \psi_r^i, W\psi_t^k \rangle. \quad (2.13)$$

The transition is said to be forbidden by symmetry if the triple Kronecker product $(\Delta^{i*} \otimes \Delta^k \otimes \Delta)$ does not contain the identity representation of $\mathbf{G}$, for then W_{rt}^{ik} vanishes identically. Here the asterisk denotes complex conjugation. For proof of this result see Hamermesh,[7] Chap. 6.

If the group $\mathbf{G}$ is augmented by the operation θ of time reversal as a symmetry element of the system, it is sometimes possible to sharpen the above criterion. This happens when the final state is related to the initial state by time reversal, and the effect is that additional selection rules may result. The original papers[8,9] demonstrating this were on the subject of the Jahn–Teller effect, but the following piece of theory due to Lax[10] seems to be all that is really necessary. We consider the matrix element

$$W_{rt}^i = \langle \theta\psi_r^i, W\psi_t^i \rangle \quad (2.14)$$

$$= \langle \theta W\psi_t^i, \theta^2\psi_r^i \rangle, \quad (2.15)$$

since θ is antiunitary. Now $\theta^2 = \omega = \pm 1$, and making the not very restrictive assumption on W that $(\theta W\theta^{-1})^\dagger = \beta W$, $\beta = \pm 1$, where the dagger denotes the adjoint operator, we find that Eq. (2.15) reduces to

$$W_{rt}^i = \beta\omega W_{tr}^i = \pm W_{tr}^i. \quad (2.16)$$

Since both ψ_t^i and $\phi_r^i = (\theta\psi_r^i)^*$ belong to Δ^i, we conclude from Eq. (2.16) that the existence of a selection rule depends in this case on the behavior of the integral

$$\int (\phi_r^i W\psi_t^i \pm \phi_t^i W\psi_r^i)\, d\tau$$

under the operations of $\mathbf{G}$. The invariance under $\mathbf{G}$ of the subspaces spanned by the functions $(\phi_r^i\psi_t^i \pm \phi_t^i\psi_r^i)$, taken together with the usual theory on selection rules, shows that in this case the transition is forbidden if the product $([\Delta^i \otimes \Delta^i] \otimes \Delta)$, in the case of the plus sign in Eq. (2.16), or if the product $(\{\Delta^i \otimes \Delta^i\} \otimes \Delta)$, in the case of the minus sign, does not contain the identity representation of $\mathbf{G}$. An equivalent criterion is whether the symmetrized (plus sign) or antisymmetrized (minus sign) square of Δ^i contains any of the irreducible components of Δ^*.

Another physical application which requires for its analysis the symmetrized powers of a given UR is the Landau and Lifshitz theory of second-order phase transitions—see, for example, Lyubarskii,[11] Chap. VII.

3. SYMMETRIZED SQUARES OF INDUCED REPRESENTATIONS

The task ahead is to rearrange or to modify the right-hand side of Eq. (2.12) so that one part can be identified with the symmetrized square $[(D \uparrow \mathbf{G}) \otimes (D \uparrow \mathbf{G})]$ and the other part with the antisymmetrized square $\{(D \uparrow \mathbf{G}) \otimes (D \uparrow \mathbf{G})\}$. In order to do this, we need to study the underlying carrier spaces of the various UR's under consideration.

Since D has as basis the set of functions ϕ_i, $i = 1$ to d, the induced UR $(D \uparrow \mathbf{G})$ has as basis the set $p_\sigma\phi_i$, $i = 1$ to d, $\sigma = 1$ to $|G|/|K|$, where the p_σ are the coset representatives in Eq. (2.7). From this coset decomposition and that of Eq. (2.9), we also have the result that, for fixed α, the basis for $[(D_\alpha \otimes D) \uparrow \mathbf{L}_\alpha] \uparrow \mathbf{G}$ is the set of functions $p_\sigma q_{\alpha\gamma}(\phi_{\alpha i}, \phi_j)$, $i = 1$ to d, $j = 1$ to d, $\gamma = 1$ to $|K|/|L_\alpha|$, $\sigma = 1$ to $|G|/|K|$. By virtue of the theorem embodied by Eq. (2.12), we know that the vector space Ω_α spanned by these $d^2\,|G|/|L_\alpha|$ functions is invariant under $\mathbf{G} \boxtimes \mathbf{G}$. We may write symbolically

$$\Omega_\alpha = \sum_{\sigma,\gamma,i,j} p_\sigma q_{\alpha\gamma}(\phi_{\alpha i}, \phi_j), \quad (3.1)$$

where the right-hand side means the set of all linear combinations of the functions $p_\sigma q_{\alpha\gamma}(\phi_{\alpha i}, \phi_j)$. The space Ω_α is independent of the particular double coset representative chosen, in the sense that, if instead of d_α we use $d_\beta = k_a d_\alpha k_b, k_a, k_b \in \mathbf{K}$, then Ω_β coincides with Ω_α. This fact was tacitly assumed in I, but is important, since we must be sure before proceeding that decompositions of the sort that appear in Eq. (2.12) and later again in this section are, term by term, unique up to equivalence. We prove this by showing

that the basis functions of Ω_α are obtained by a non-singular transformation from those of Ω_β.

Theorem 1: If $\mathbf{K}d_\alpha\mathbf{K} = \mathbf{K}d_\beta\mathbf{K}$, then $\Omega_\alpha = \Omega_\beta$.

Proof:

$$\begin{aligned}\mathbf{L}_\beta &= \mathbf{K} \cap \mathbf{K}_\beta = K \cap d_\beta\mathbf{K}d_\beta^{-1} \\ &= \mathbf{K} \cap k_a d_\alpha k_b \mathbf{K} k_b^{-1} d_\alpha^{-1} k_a^{-1} = \mathbf{K} \cap k_a\mathbf{K}_\alpha k_a^{-1}\end{aligned}$$

and, hence,

$$k_a^{-1}\mathbf{L}_\beta k_a = \mathbf{K} \cap \mathbf{K}_\alpha = \mathbf{L}_\alpha.$$

Thus, if

$$\mathbf{K} = \sum_\gamma q_{\alpha\gamma}\mathbf{L}_\alpha,$$

then we also have

$$\mathbf{K} = \mathbf{K}k_a^{-1} = \sum_\gamma q_{\alpha\gamma}k_a^{-1}\mathbf{L}_\beta,$$

and so, for each γ we may choose, $q_{\beta\gamma} = q_{\alpha\gamma}k_a^{-1}$ and

$$\mathbf{K} = \sum_\gamma q_{\beta\gamma}\mathbf{L}_\beta.$$

Then, from Eq. (3.1) and the expressions just obtained, we have

$$\begin{aligned}\Omega_\beta &= \sum_{\sigma,\gamma,i,j} p_\sigma q_{\beta\gamma}(\phi_{\beta i}, \phi_j) \\ &= \sum_{\sigma,\gamma,i,j} p_\sigma q_{\alpha\gamma}k_a^{-1}(d_\beta\phi_i, \phi_j) \\ &= \sum_{\sigma,\gamma,i,j} p_\sigma q_{\alpha\gamma}(d_\alpha k_b\phi_i, k_a^{-1}\phi_j) \\ &= \sum_{\sigma,\gamma,k,l} p_\sigma q_{\alpha\gamma}(\phi_{\alpha k}, \phi_l) = \Omega_\alpha. \qquad (3.2)\end{aligned}$$

The vector space $\Omega = V \otimes V$, which is the carrier space for the UR $(D \uparrow \mathbf{G}) \otimes (D \uparrow \mathbf{G})$ is, of course, already proved in I to be the direct sum $\sum_\alpha \Omega_\alpha$, where the sum over α is taken over the distinct double cosets of Eq. (2.8). Together with Theorem 1 this implies that no two vector spaces Ω_α and Ω_β coincide unless d_α and d_β belong to the same double coset.

For reasons which will soon emerge, given a particular double coset representative d_α, we find it useful to define $\phi_{\tilde{\alpha}i} = d_\alpha^{-1}\phi_i$ and to write

$$\mathbf{L}_{\tilde{\alpha}} = \mathbf{K} \cap \mathbf{K}_{\tilde{\alpha}} = \mathbf{K} \cap d_\alpha^{-1}\mathbf{K}d_\alpha.$$

Note that $\mathbf{L}_{\tilde{\alpha}} = d_\alpha^{-1}\mathbf{L}_\alpha d_\alpha$ and that, from Eqs. (2.7) and (2.9), we may write

$$\mathbf{G} = \mathbf{G}d_\alpha = \sum_{\sigma,\gamma} p_\sigma q_{\alpha\gamma}\mathbf{L}_\alpha d_\alpha = \sum_{\sigma,\gamma} p_\sigma q_{\alpha\gamma} d_\alpha \mathbf{L}_{\tilde{\alpha}}.$$

Hence, as σ and γ run over all possible values, the set $p_\sigma q_{\alpha\gamma} d_\alpha$ forms a complete set of coset representatives of $\mathbf{L}_{\tilde{\alpha}}$ in $\mathbf{G}$. Thus, if we define the vector space $\Phi_{\tilde{\alpha}}$ as the carrier space of $((D \otimes D_{\tilde{\alpha}}) \downarrow \mathbf{L}_{\tilde{\alpha}}) \uparrow \mathbf{G}$, we find

$$\begin{aligned}\Phi_{\tilde{\alpha}} &= \sum_{\sigma,\gamma,i,j} p_\sigma q_{\alpha\gamma} d_\alpha(\phi_i, \phi_{\tilde{\alpha}j}) \\ &= \sum_{\sigma,\gamma,i,j} p_\sigma q_{\alpha\gamma}(\phi_{\alpha i}, \phi_j) = \Omega_\alpha. \qquad (3.3)\end{aligned}$$

Hence, the vector spaces $\Phi_{\tilde{\alpha}}$ and Ω_α coincide. Similarly, Φ_α and $\Omega_{\tilde{\alpha}}$ coincide. Now the character of

$$((D \otimes D_\alpha) \downarrow \mathbf{L}_\alpha) \uparrow \mathbf{G}$$

is clearly the same as that of $((D_\alpha \otimes D) \downarrow \mathbf{L}_\alpha) \uparrow \mathbf{G}$, and since the first of these is defined on $\Phi_\alpha = \Omega_{\tilde{\alpha}}$ and the second is defined on Ω_α, it follows that the UR's defined on Ω_α and $\Omega_{\tilde{\alpha}}$ are equivalent. There are now two cases that can occur: either Ω_α and $\Omega_{\tilde{\alpha}}$ coincide (and trivially define equivalent UR's by some unitary transformation of the space onto itself), or Ω_α and $\Omega_{\tilde{\alpha}}$ are distinct (but, nevertheless, define equivalent UR's). Now, from what has gone previously we know that Ω_α and $\Omega_{\tilde{\alpha}}$ coincide if and only if $\mathbf{K}d_\alpha\mathbf{K} = \mathbf{K}d_\alpha^{-1}\mathbf{K}$, so that which of the two cases occurs depends critically upon whether the double coset is self-inverse or not [for, of course, $(\mathbf{K}d_\alpha\mathbf{K})^{-1} = \mathbf{K}d_\alpha^{-1}\mathbf{K}$, since $\mathbf{K} = \mathbf{K}^{-1}$)]. We have proved the following theorem.

Theorem 2: Ω_α and $\Omega_{\tilde{\alpha}}$ coincide if and only if α is a self-inverse double coset. If α is not self-inverse, Ω_α and $\Omega_{\tilde{\alpha}}$ are distinct, but define equivalent representations in the expansion (2.12).

Consider first the case in which Ω_α and $\Omega_{\tilde{\alpha}}$ are distinct. Now,

$$\Omega_\alpha = \sum_{\sigma,\gamma,i,j} p_\sigma q_{\alpha\gamma}(\phi_{\alpha i}, \phi_j)$$

and

$$\Omega_{\tilde{\alpha}} = \Phi_\alpha = \sum_{\sigma,\gamma,i,j} p_\sigma q_{\alpha\gamma}(\phi_i, \phi_{\alpha j}).$$

These are distinct spaces and from Theorem (3.2) define equivalent UR's. Clearly, it is possible to form the direct sum $(\Omega_\alpha + \Omega_{\tilde{\alpha}})$ and to decompose this direct sum into an alternative one $(\Omega_\alpha^+ + \Omega_\alpha^-)$, in which $\Omega_\alpha^\pm$ are given by

$$\Omega_\alpha^\pm = \sum_{\sigma,\gamma,i,j} p_\sigma q_{\alpha\gamma}((\phi_{\alpha i}, \phi_j) \pm (\phi_j, \phi_{\alpha i})), \qquad (3.4)$$

with upper and lower signs corresponding. From the form of the right-hand side of Eq. (3.4), it is also clear that Ω_α^+ is a subspace of Ω^+ and Ω_α^- is a subspace of Ω^-, where Ω^+ and Ω^- are the symmetrized and antisymmetrized subspaces of Ω itself. Moreover, the UR's on Ω_α^+, Ω_α^-, Ω_α, and $\Omega_{\tilde{\alpha}}$ are all equivalent, the first two being derived from the last two by the simple equivalence transformation expressed by Eq. (3.4). We have proved the next theorem.

Theorem 3: If Ω_α and $\Omega_{\tilde{\alpha}}$ are distinct, that is, if d_α and d_α^{-1} belong to different double cosets, then $[(D \uparrow \mathbf{G}) \otimes (D \uparrow \mathbf{G})]$ and $\{(D \uparrow \mathbf{G}) \otimes (D \uparrow \mathbf{G})\}$ will both contain a representation equivalent to

$$((D_\alpha \otimes D) \downarrow \mathbf{L}_\alpha) \uparrow \mathbf{G}.$$

Next, consider the case $\Omega_\alpha = \Omega_{\bar{\alpha}}$. There are two possibilities which require separate treatment, since the decomposition depends upon whether the carrier spaces of D_α and D are identical or distinct. These spaces are identical if and only if $d_\alpha \in \mathbf{K}$, in which case d_α can be chosen to be the identity d_1. Then, $\mathbf{K}_1 = \mathbf{K}$ and $\mathbf{L}_1 = \mathbf{K} \cap \mathbf{K}_1 = \mathbf{K}$. The term in Eq. (2.12) now under consideration is just $(D \otimes D) \uparrow \mathbf{G}$ and

$$\Omega_1 = \sum_{\sigma,i,j} p_\sigma(\phi_i, \phi_j).$$

It is immediately clear that Ω_1 decomposes into the direct sum $(\Omega_1^+ + \Omega_1^-)$, where

$$\Omega_1^\pm = \sum_{\sigma,i,j} p_\sigma((\phi_i, \phi_j) \pm (\phi_j, \phi_i)), \tag{3.5}$$

and that Ω_1^+ is a subspace of Ω^+ and Ω_1^- is a subspace of Ω^-. From Eq. (3.5), we see that Ω_1^+ defines the UR $[D \otimes D] \uparrow \mathbf{G}$ and Ω_1^- defines the UR $\{D \otimes D\} \uparrow \mathbf{G}$, the dimensions of these UR's being $\frac{1}{2}d(d+1)h$ and $\frac{1}{2}d(d-1)h$, respectively, where $h = |G|/|K|$. The difference in dimension between these two UR's is, therefore, equal to dh. Since Ω^+ and Ω^- are of dimension $\frac{1}{2}dh(dh+1)$ and $\frac{1}{2}dh(dh-1)$, respectively, and therefore have a difference of dimension also equal to dh, it follows that this difference is entirely accounted for by the decomposition of Ω_1. Taken with Theorem 3, this makes it reasonable to suppose that any other space Ω_α for which $\Omega_\alpha = \Omega_{\bar{\alpha}}$ will somehow yield spaces of equal dimensionality in the final decomposition of Ω. We now embark upon some fairly elaborate analysis to show how this, in fact, occurs.

We consider, therefore, the case in which

$$\mathbf{K}d_\alpha\mathbf{K} = \mathbf{K}d_\alpha^{-1}\mathbf{K} \neq \mathbf{K}.$$

There must exist elements k_a, k_b, k_c, and k_d such that $k_a d_\alpha k_b = k_c d_\alpha^{-1} k_d$; that is, $d_\alpha k_b k_d^{-1} = k_a^{-1} k_c d_\alpha^{-1}$. Thus, $d_\alpha\mathbf{K} \cap \mathbf{K}d_\alpha^{-1}$ is nonempty. Let a be any element of this intersection; then, a is not in $\mathbf{K}$, and we may write $a = d_\alpha k = \bar{k} d_\alpha^{-1}$, where $k, \bar{k} \in \mathbf{K}$. Now

$$a^2 = d_\alpha k \bar{k} d_\alpha^{-1} = k\bar{k} \in d_\alpha\mathbf{K}d_\alpha^{-1} \cap \mathbf{K} = \mathbf{L}_\alpha$$

and, furthermore, $a\mathbf{K}a^{-1} = d_\alpha k \mathbf{K} k^{-1} d_\alpha^{-1} = \mathbf{K}_\alpha$, and

$$a\mathbf{K}_\alpha a^{-1} = \bar{k} d_\alpha^{-1} \mathbf{K}_\alpha d_\alpha \bar{k}^{-1} = \mathbf{K}.$$

Thus, we may construct the group

$$\mathbf{M}_\alpha = \mathbf{L}_\alpha + a\mathbf{L}_\alpha, \tag{3.6}$$

and then $\mathbf{L}_\alpha$ is an invariant subgroup of $\mathbf{M}_\alpha$ of index 2. Starting from the UR $(D_\alpha \otimes D)$ on $\mathbf{L}_\alpha$ with basis $\langle\phi_\alpha| \langle\phi|$ [i.e., the set of functions $(\phi_{\alpha i}, \phi_j)$], we can construct the induced representation $C_\alpha = (D_\alpha \otimes D) \uparrow \mathbf{M}_\alpha$ with basis $(\langle\phi_\alpha| \langle\phi|, a \langle\phi_\alpha| \langle\phi|)$. The analysis follows Sec. 2 of I and the result is that, if $l \in \mathbf{L}_\alpha$, then the induced representation C_α is given in block-matrix form by the expressions

$$C_\alpha(l) = \begin{pmatrix} D(d_\alpha^{-1} l d_\alpha) \otimes D(l) & 0 \\ 0 & D(d_\alpha^{-1} a^{-1} l a d_\alpha) \otimes D(a^{-1} l a) \end{pmatrix} \tag{3.7}$$

and

$$C_\alpha(a) = \begin{pmatrix} 0 & D(d_\alpha^{-1} a^2 d_\alpha) \otimes D(a^2) \\ 1 & 0 \end{pmatrix}. \tag{3.8}$$

For convenience, we write

$$N_\alpha(l) = D(d_\alpha^{-1} l d_\alpha) \otimes D(l) \tag{3.9}$$

and

$${}^aN_\alpha(l) = N_\alpha(a^{-1} l a), \tag{3.10}$$

so that

$$C_\alpha(l) = \begin{pmatrix} N_\alpha(l) & 0 \\ 0 & {}^aN_\alpha(l) \end{pmatrix} \tag{3.11}$$

and

$$C_\alpha(a) = \begin{pmatrix} 0 & N_\alpha(a^2) \\ 1 & 0 \end{pmatrix}. \tag{3.12}$$

A short analysis shows that, if we can find a unitary matrix P such that

$$P^2 = N_\alpha(a^2) \tag{3.13}$$

and

$$P\,{}^aN_\alpha(l)P^{-1} = N_\alpha(l), \tag{3.14}$$

for all $l \in \mathbf{L}_\alpha$, then ${}^aN_\alpha$ and N_α are equivalent; and with respect to the transformed basis $(\langle\phi_\alpha| \langle\phi|, a \langle\phi_\alpha| \langle\phi| P^{-1})$, the UR C_α becomes transformed into the equivalent UR C'_α given by

$$C'_\alpha(l) = \begin{pmatrix} N_\alpha(l) & 0 \\ 0 & N_\alpha(l) \end{pmatrix} \tag{3.15}$$

and

$$C'_\alpha(a) = \begin{pmatrix} 0 & P \\ P & 0 \end{pmatrix}. \tag{3.16}$$

A matrix P which satisfies Eqs. (3.13) and (3.14) can, in fact, be found and is given by

$$P_{ts,uv} = D(\bar{k})_{su} D(k)_{tv} = D(\bar{k}) \otimes D(k)_{st,uv}. \tag{3.17}$$

For example, with this choice of P, we can soon check that Eq. (3.13) holds:

$$\begin{aligned} P^2_{ab,cd} &= \sum_{e,f} P_{ab,ef} P_{ef,cd} \\ &= \sum_{e,f} D(\bar{k})_{be} D(k)_{af} D(\bar{k})_{fc} D(k)_{ed} \\ &= D(\bar{k}k)_{bd} D(k\bar{k})_{ac} \\ &= D(a^2)_{bd} D(d_\alpha^{-1} a^2 d_\alpha)_{ac} \\ &= D(d_\alpha^{-1} a^2 d_\alpha) \otimes D(a^2)_{ab,cd} \\ &= N_\alpha(a^2)_{ab,cd}. \end{aligned}$$

A similar straightforward calculation soon verifies Eq. (3.14) in the form $(P^{\alpha}N_{\alpha}(l))_{ab,cd} = (N_{\alpha}(l)P)_{ab,cd}$, and we omit the details.

The motivation for the above analysis is that the basis for C'_{α} now takes on a particularly simple form. To see this, let us compute the (kl) element of the second member of the basis

$$\begin{aligned}(a\langle\phi_{\alpha}|\langle\phi|P^{-1})_{kl} &= \sum_{i,j} a(\phi_{\alpha i},\phi_j)P^{-1}_{ij,kl}\\ &= \sum_{i,j}(\bar{k}d_{\alpha}^{-1}\phi_{\alpha i}, d_{\alpha}k\phi_j)P^{-1}_{ij,kl}\\ &= \sum_{i,j,t,s}(\phi_s,\phi_{\alpha t})D(\bar{k})_{si}D(k)_{tj}P^{-1}_{ij,kl}\\ &= \sum_{i,j,t,s}(\phi_s,\phi_{\alpha t})P_{ts,ij}P^{-1}_{ij,kl}\\ &= (\phi_l,\phi_{\alpha k}).\end{aligned} \tag{3.18}$$

Also,

$$(\langle\phi_{\alpha}|\langle\phi|)_{kl} = (\phi_{\alpha k},\phi_l). \tag{3.19}$$

Hence, combining Eqs. (3.18) and (3.19), we obtain

$$(\langle\phi_{\alpha}|\langle\phi| \pm a\langle\phi_{\alpha}|\langle\phi|P^{-1})_{kl} = (\phi_{\alpha k},\phi_l) \pm (\phi_l,\phi_{\alpha k}). \tag{3.20}$$

Using as a new basis the two functions on the left-hand side of Eq. (3.20), we find that the UR C'_{α} becomes transformed into the equivalent UR C''_{α} given by

$$C''_{\alpha}(l) = \begin{pmatrix} N_{\alpha}(l) & 0\\ 0 & N_{\alpha}(l)\end{pmatrix} \tag{3.21}$$

and

$$C''_{\alpha} = \begin{pmatrix} P & 0\\ 0 & -P\end{pmatrix}. \tag{3.22}$$

Thus, the symmetrized basis $(\phi_{\alpha k},\phi_l) + (\phi_l,\phi_{\alpha k})$ yields a UR N^{+}_{α} of $\mathbf{M}_{\alpha}$ given by

$$N^{+}_{\alpha}(l) = N_{\alpha}(l), \quad N^{+}_{\alpha}(a) = P, \tag{3.23}$$

and the antisymmetrized basis $(\phi_{\alpha k},\phi_l) - (\phi_l,\phi_{\alpha k})$ yields a UR N^{-}_{α} of $\mathbf{M}_{\alpha}$ given by

$$N^{-}_{\alpha}(l) = N_{\alpha}(l), \quad N^{-}_{\alpha}(a) = -P. \tag{3.24}$$

Furthermore, when the UR's N^{+}_{α} and N^{-}_{α} are induced from $\mathbf{M}_{\alpha}$ into $\mathbf{G}$, they are clearly defined on the spaces Ω^{+}_{α} and Ω^{-}_{α}, where $\Omega_{\alpha} = (\Omega^{+}_{\alpha} + \Omega^{-}_{\alpha})$ is a direct-sum decomposition of Ω_{α} and where Ω^{+}_{α} is a subspace of Ω^{+} and Ω^{-}_{α} is a subspace of Ω^{-}. Thus, $[(D\uparrow\mathbf{G})\otimes(D\uparrow\mathbf{G})]$ contains $N^{+}_{\alpha}\uparrow\mathbf{G}$ and $\{(D\uparrow\mathbf{G})\otimes(D\uparrow\mathbf{G})\}$ contains $N^{-}_{\alpha}\uparrow\mathbf{G}$.

Also, from Eqs. (3.9), (3.17), (3.23), and (3.24), the characters of N^{+}_{α} and N^{-}_{α} are easily calculated and are, for all $l\in\mathbf{L}_{\alpha}$,

$$\chi_{N\alpha^{\pm}}(l) = \chi_D(d_{\alpha}^{-1}ld_{\alpha})\chi_D(l), \tag{3.25}$$

$$\chi_{N\alpha^{\pm}}(al) = \pm\chi_D(alal), \tag{3.26}$$

where χ_D is the character of D.

For example,

$$\begin{aligned}\chi_{N\alpha^{+}}(al) &= \sum_{i,j} N^{+}_{\alpha}(al)_{ij,ij}\\ &= \sum_{i,j,k,l} P_{ij,kl}N_{\alpha}(l)_{kl,ij}\\ &= \sum_{i,j,k,l}(D(\bar{k})\otimes D(k))_{ji,kl}(D(d_{\alpha}^{-1}ld_{\alpha})\otimes D(l))_{kl,ij}\\ &= \sum_{i,j}(D(\bar{k}d_{\alpha}^{-1}ld_{\alpha})\otimes D(kl))_{ji,ij}\\ &= \sum_{i,j} D(\bar{k}d_{\alpha}^{-1}ld_{\alpha})_{ji}D(kl)_{ij}\\ &= \sum_{j} D(\bar{k}d_{\alpha}^{-1}ld_{\alpha}kl)_{jj}\\ &= \chi_D(alal).\end{aligned}$$

We have now proved the following theorems.

Theorem 4: Corresponding to $\Omega_{\alpha} = \Omega_{\bar{\alpha}} = \Omega_1$, the $\alpha = 1$ term in the expansion (2.12), $(D\otimes D)\uparrow\mathbf{G}$, decomposes so that $[D\otimes D]\uparrow\mathbf{G}$ appears in the reduction of $[(D\uparrow\mathbf{G})\otimes(D\uparrow\mathbf{G})]$ and $\{D\otimes D\}\uparrow\mathbf{G}$ appears in the reduction of $\{(D\uparrow\mathbf{G})\otimes(D\uparrow\mathbf{G})\}$.

Theorem 5: Corresponding to $\Omega_{\alpha} = \Omega_{\bar{\alpha}} \neq \Omega_1$, the α term in the expansion (2.12), which we write in the form $N_{\alpha}\uparrow\mathbf{G}$, decomposes so that $N^{+}_{\alpha}\uparrow\mathbf{G}$ appears in the reduction of $[(D\uparrow\mathbf{G})\otimes(D\uparrow\mathbf{G})]$ and $N^{-}_{\alpha}\uparrow\mathbf{G}$ appears in the reduction of $\{(D\uparrow\mathbf{G})\otimes(D\uparrow\mathbf{G})\}$. Here $N^{\pm}_{\alpha}$ are defined by Eqs. (3.23) and (3.24).

Theorems 3–5 are now combined and presented in full detail as

Theorem 6: Let D be a representation of the subgroup $\mathbf{K}$ of the finite group $\mathbf{G}$ and suppose that

$$\mathbf{G} = \sum_{\alpha}\mathbf{K}d_{\alpha}\mathbf{K}$$

is the double coset decomposition of $\mathbf{G}$ relative to $\mathbf{K}$. Let $\mathbf{L}_{\alpha} = \mathbf{K}\cap d_{\alpha}\mathbf{K}d_{\alpha}^{-1}$ and denote by $N_{\alpha}(l)$ the representation of $\mathbf{L}_{\alpha}$ defined by $D(d_{\alpha}^{-1}l\,d_{\alpha})\otimes D(l)$, for all $l\in\mathbf{L}_{\alpha}$. Let A be the set of all self-inverse double cosets $\mathbf{K}d_{\alpha}\mathbf{K} = \mathbf{K}d_{\alpha}^{-1}\mathbf{K}$ except $\mathbf{K}$ itself. For each $\alpha\in A$, $d_{\alpha}\mathbf{K}\cap\mathbf{K}d_{\alpha}^{-1}$ is nonempty. Let a be any one of its members, say $a = d_{\alpha}k = \bar{k}\,d_{\alpha}^{-1}$ $(k,\bar{k}\in\mathbf{K})$, and let $\mathbf{M}_{\alpha}$ be the subgroup generated by $\mathbf{L}_{\alpha}$ and a. Then, $\mathbf{L}_{\alpha}$ is an invariant subgroup of $\mathbf{M}_{\alpha}$ of index 2. Let P be the matrix given by $P_{ab,cd} = D(\bar{k})_{bc}D(k)_{ad}$. Then, there exist extensions N^{+}_{α} and N^{-}_{α} of N_{α} into the group $\mathbf{M}_{\alpha}$, such that $N^{+}_{\alpha}(a) = P$ and $N^{-}_{\alpha}(a) = -P$. Let B be the set of all distinct sets of the form $\mathbf{K}d_{\beta}\mathbf{K}\cup\mathbf{K}d_{\beta}^{-1}\mathbf{K}$, where $\mathbf{K}d_{\beta}\mathbf{K}$ is not a self-inverse double coset. Then,

finally, we have

$$[(D \uparrow \mathbf{G}) \otimes (D \uparrow \mathbf{G})] = [D \otimes D] \uparrow \mathbf{G} + \sum_{\alpha \in A} N_\alpha^+ \uparrow \mathbf{G} + \sum_{\beta \in B} N_\beta \uparrow \mathbf{G}$$

and

$$\{(D \uparrow \mathbf{G}) \otimes (D \uparrow \mathbf{G})\} = \{D \otimes D\} \uparrow \mathbf{G} + \sum_{\alpha \in A} N_\alpha^- \uparrow \mathbf{G} + \sum_{\beta \in B} N_\beta \uparrow \mathbf{G}.$$

A closing remark that ought to be made about Theorem 6 is that its form as it appears here is not quite as general as that established by Mackey,[3] but is sufficiently general for all computational needs. Finally, since the expansions derived are in the form of sums of induced representations, it follows that repeated application of the main theorem in I allows computation of triple products of the form

$$[(D \uparrow \mathbf{G}) \otimes (D \uparrow \mathbf{G})] \otimes (C \uparrow \mathbf{G})$$

and $\{(D \uparrow \mathbf{G}) \otimes (D \uparrow \mathbf{G})\} \otimes (C \uparrow \mathbf{G})$, where C is a representation of $\mathbf{H}$, another subgroup of $\mathbf{K}$. This is important in view of the criterion established at the end of Sec. 2 for the existence of extra selection rules.

4. APPLICATION TO SPACE GROUPS

In the case of the crystallographic space groups, $\mathbf{G}$ is to be identified with the space group, $\mathbf{K}$ with the little group $\mathbf{G}^{\mathbf{k}}$ of a wave vector $\mathbf{k}$ from the Brillouin zone, D with a small UR (or allowed UR) $D_p^{\mathbf{k}}$ of $\mathbf{G}^{\mathbf{k}}$, and both $D_p^{\mathbf{k}}$ and $\Delta_p^{\mathbf{k}} = D_p^{\mathbf{k}} \uparrow \mathbf{G}$ are irreducible. Indeed, every UIR of a space group arises by inducing a small UR of some little group: see Sec. 3 of I for a review of the representation theory of space groups.

The calculation to be performed is the evaluation of the coefficients $C_{pq}^{\mathbf{kh}}$, $[C_{pq}^{\mathbf{kh}}]$, and $\{C_{pq}^{\mathbf{kh}}\}$ in the expansions:

$$(D_p^{\mathbf{k}} \uparrow \mathbf{G}) \otimes (D_p^{\mathbf{k}} \uparrow \mathbf{G}) \equiv \sum_{\mathbf{h},q} C_{pq}^{\mathbf{kh}} (D_q^{\mathbf{h}} \uparrow \mathbf{G}), \quad (4.1)$$

$$[(D_p^{\mathbf{k}} \uparrow \mathbf{G}) \otimes (D_p^{\mathbf{k}} \uparrow \mathbf{G})] = \sum_{\mathbf{h},q} [C_{pq}^{\mathbf{kh}}] (D_q^{\mathbf{h}} \uparrow \mathbf{G}), \quad (4.2)$$

$$\{(D_p^{\mathbf{k}} \uparrow \mathbf{G}) \otimes (D_p^{\mathbf{k}} \uparrow \mathbf{G})\} = \sum_{\mathbf{h},q} \{C_{pq}^{\mathbf{kh}}\} (D_q^{\mathbf{h}} \uparrow \mathbf{G}). \quad (4.3)$$

In Eqs. (4.1)-(4.3), the sum over $\mathbf{h}$ is over all wave vectors in the basic zone (see Sec. 3 of I for definition), and the sum over q is over all the small UR's of $\mathbf{G}^{\mathbf{h}}$, the little group of $\mathbf{h}$. From Eq. (2.6), it is clear that

$$C_{pq}^{\mathbf{kh}} = [C_{pq}^{\mathbf{kh}}] + \{C_{pq}^{\mathbf{kh}}\}, \quad (4.4)$$

so that it is sufficient to evaluate just two of these coefficients (there is probably least labor involved in evaluating $C_{pq}^{\mathbf{kh}}$ and then $\{C_{qp}^{\mathbf{kh}}\}$).

The first step is the use of the completeness and orthogonality of the characters, whereby $C_{pq}^{\mathbf{kh}}$ is seen to be equal to the frequency of the identity representation in the triple product $(D_p^{\mathbf{k}} \uparrow \mathbf{G}) \otimes (D_p^{\mathbf{k}} \uparrow \mathbf{G}) \otimes (D_q^{\mathbf{h}^*} \uparrow \mathbf{G})$. Similarly, $[C_{pq}^{\mathbf{kh}}]$ is the frequency of the identity representation in the triple product

$$[(D_p^{\mathbf{k}} \uparrow \mathbf{G}) \otimes (D_p^{\mathbf{k}} \uparrow \mathbf{G})] \otimes (D_q^{\mathbf{h}^*} \uparrow \mathbf{G})$$

and, similarly, for $\{C_{pq}^{\mathbf{kh}}\}$. The evaluation of the first of these triple products is covered by the analysis of Sec. 4 of I. Following the prescription there, we form the double-coset decomposition

$$\mathbf{G} = \sum_\alpha \mathbf{G}^{\mathbf{k}} \{R_\alpha \mid \mathbf{v}_\alpha\} \mathbf{G}^{\mathbf{k}}, \quad (4.5)$$

write $\mathbf{G}_\alpha^{\mathbf{k}} = \{R_\alpha \mid \mathbf{v}_\alpha\} \mathbf{G}^{\mathbf{k}} \{R_\alpha \mid \mathbf{v}_\alpha\}^{-1}$ (the little group of $\mathbf{k}_\alpha = R_\alpha \mathbf{k}$), and define $\mathbf{L}_\alpha^{\mathbf{k}} = \mathbf{G}^{\mathbf{k}} \cap \mathbf{G}_\alpha^{\mathbf{k}}$. $D_p^{\mathbf{k}}$ is a small UR of $\mathbf{G}^{\mathbf{k}}$, and $D_{\alpha p}^{\mathbf{k}}$ is the small UR of $\mathbf{G}_\alpha^{\mathbf{k}}$ for which, if $\{R \mid \mathbf{v}\} \in \mathbf{G}^{\mathbf{k}}$,

$$D_{\alpha p}^{\mathbf{k}}(\{R_\alpha \mid \mathbf{v}_\alpha\}\{R \mid \mathbf{v}\}\{R_\alpha \mid \mathbf{v}_\alpha\}^{-1}) = D_p^{\mathbf{k}}(\{R \mid \mathbf{v}\}). \quad (4.6)$$

We next form the decomposition

$$\mathbf{G} = \sum_\gamma \mathbf{L}_\alpha^{\mathbf{k}} \{R_\gamma \mid \mathbf{v}_\gamma\} \mathbf{G}^{\mathbf{h}}, \quad (4.7)$$

and write

$$\mathbf{G}_\gamma^{\mathbf{h}} = \{R_\gamma \mid \mathbf{v}_\gamma\} \mathbf{G}^{\mathbf{h}} \{R_\gamma \mid \mathbf{v}_\gamma\}^{-1}$$

(the little group of $\mathbf{h}_\gamma = R_\gamma \mathbf{h}$) and define

$$\mathbf{N}_{\alpha\gamma}^{\mathbf{kh}} = \mathbf{L}_\alpha^{\mathbf{k}} \cap \mathbf{G}_\gamma^{\mathbf{h}} = \mathbf{G}^{\mathbf{k}} \cap \mathbf{G}_\alpha^{\mathbf{k}} \cap \mathbf{G}_\gamma^{\mathbf{h}}.$$

The UR $D_{\gamma q}^{\mathbf{h}}$ is defined in terms of $D_q^{\mathbf{h}}$ by an equation analogous to (4.6).

The results of Sec. 4 of I show that $C_{pq}^{\mathbf{kh}}$ is equal to the sum over allowed pairs of values of α and γ of the frequency of the identity representation of $\mathbf{N}_{\alpha\gamma}^{\mathbf{kh}}$ in the triple product $(D_p^{\mathbf{k}} \otimes D_{\alpha p}^{\mathbf{k}} \otimes D_{\gamma q}^{\mathbf{h}^*}) \downarrow \mathbf{N}_{\alpha\gamma}^{\mathbf{kh}}$; the allowed pairs of α and γ are determined from Eqs. (4.5) and (4.7). For a particular (α, γ) pair, this frequency is zero because of a sum over the translations, unless

$$\mathbf{k} + \mathbf{k}_\alpha \equiv \mathbf{h}_\gamma, \quad (4.8)$$

and, when Eq. (4.8) holds, the value of this frequency is

$$\frac{|T|}{|N|} \sum_{\{S \mid \mathbf{w}\}} \chi_p^{\mathbf{k}}(\{S \mid \mathbf{w}\}) \chi_p^{\mathbf{k}}(\{R_\alpha \mid \mathbf{v}_\alpha\})^{-1} \{S \mid \mathbf{w}\}\{R_\alpha \mid \mathbf{v}_\alpha\}) \times \chi_q^{\mathbf{h}^*}(\{R_\gamma \mid \mathbf{v}_\gamma\})^{-1} \{S \mid \mathbf{w}\}\{R_\gamma \mid \mathbf{v}_\gamma\}), \quad (4.9)$$

where $|N|/|T|$ is the index of the translational subgroup $\mathbf{T}$ in the group $N_{\alpha\gamma}^{\mathbf{kh}}$ and the $\{S \mid \mathbf{w}\}$, over which the summation is taken, are the coset representatives of $\mathbf{N}_{\alpha\gamma}^{\mathbf{kh}}$ with respect to $\mathbf{T}$. Sums of the form (4.9) are over very few elements and are, therefore, quite easy to evaluate. It is important to realize that, for a particular (α, γ) pair allowed by Eqs. (4.5) and (4.7), Eq. (4.8) may lead to no solution for $\mathbf{h}$—in fact, it leads to a solution only if $R_\gamma^{-1}(\mathbf{k} + R_\alpha \mathbf{k})$ is equivalent to a

vector in the basic zone (the symmetrically chosen part of the Brillouin zone being used to label the space group UR's). This means that out of all possible (α, γ) pairs very few survive, so that the number of summations of the form (4.9) to be carried out is also very few. In our opinion, the advantage of the method lies in the strictly limited amount of numerical computation required to obtain the results and in the exact specification of just which computations have to be performed.

We now look again at Eq. (2.12) which reads, when translated into the language of this section,

$$(D_p^{\mathbf{k}} \uparrow \mathbf{G}) \otimes (D_\nu^{\mathbf{k}} \uparrow \mathbf{G}) \equiv \sum_\alpha ((D_{\alpha p}^{\mathbf{k}} \otimes D_\nu^{\mathbf{k}}) \downarrow \mathbf{L}_\alpha^{\mathbf{k}}) \uparrow \mathbf{G}. \tag{4.10}$$

We see from Eq. (4.1) that what we have done so far in this section is to analyze the right-hand side of Eq. (4.10) into a sum of irreducibles

$$\sum_{\mathbf{h},q} C_{pq}^{\mathbf{kh}} (D_q^{\mathbf{h}} \uparrow \mathbf{G}). \tag{4.11}$$

But, more than this, we are in a position to say which terms in the sum (4.11) arise from each particular value of α on the right-hand side of Eq. (4.10). For a fixed α, they are just those arising from the nonzero frequencies calculated from Eqs. (4.8) and (4.9) for that particular α. For reasons which will soon emerge it is an important labor-saving device to keep a record of the terms that appear in (4.11) for each value of α.

The next step is to use Theorem 6 to distribute the UIR's that appear in the sum (4.11), some of them to the expansion for the symmetrized square and the remainder to the expansion for the antisymmetrized square. The task will then be complete because only those UIR's which appear in the expansion for the square can possibly appear in the expansions for the symmetrized or antisymmetrized square. Furthermore, the strong link between the three expansions means that this distribution must take place, term by term, separately for each value of α. Hence, the importance of keeping a record of the terms that appear in (4.11) for each value of α is that it provides us with a short list of suitable candidates for each term in the expansions resulting from Theorem 6. Using an obvious notation adapted for the purposes of this section, we find that these expansions are

$$[(D_p^{\mathbf{k}} \uparrow \mathbf{G}) \otimes (D_\nu^{\mathbf{k}} \uparrow \mathbf{G})]$$
$$\equiv [D_p^{\mathbf{k}} \otimes D_\nu^{\mathbf{k}}] \uparrow \mathbf{G} + \sum_{\alpha \in A} N_{\alpha p}^{\mathbf{k}+} \uparrow \mathbf{G} + \sum_{\beta \in B} N_{\beta p}^{\mathbf{k}} \uparrow \mathbf{G} \tag{4.12}$$

and

$$\{(D_p^{\mathbf{k}} \uparrow \mathbf{G}) \otimes (D_\nu^{\mathbf{k}} \uparrow \mathbf{G})\}$$
$$\equiv \{D_p^{\mathbf{k}} \otimes D_\nu^{\mathbf{k}}\} \uparrow \mathbf{G} + \sum_{\alpha \in A} N_{\alpha p}^{\mathbf{k}-} \uparrow \mathbf{G} + \sum_{\beta \in B} N_{\beta p}^{\mathbf{k}} \uparrow \mathbf{G}. \tag{4.13}$$

In Eqs. (4.12) and (4.13), the first term corresponds to $\alpha = 1$ [that is, $\{R_\alpha \mid \mathbf{v}_\alpha\} = \{E \mid \mathbf{0}\}$ in Eq. (4.5)], the second term to self-inverse double cosets, $\alpha \neq 1$, and the third term to the remaining double cosets grouped in pairs, members of each pair being inverse to each other. Also, $N_{\beta p}^{\mathbf{k}} = (D_{\beta p}^{\mathbf{k}} \otimes D_p^{\mathbf{k}})$ and $N_{\alpha p}^{\mathbf{k}\pm}$ derive from $N_{\alpha p}^{\mathbf{k}}$ by the processes described in Sec. 3. It is, therefore, appropriate at this stage to allocate the double cosets appearing in Eq. (4.5) accordingly. Having done this, we can immediately dispose of the set B. Take any $\beta \in B$. We are required to evaluate $N_{\beta p}^{\mathbf{k}} \uparrow \mathbf{G}$. But this has already been evaluated in dealing with the square [see Eq. (4.10)]. A record having been kept of the UIR's that appear for this value of β, they can then be marked down as belonging to both the symmetrized and antisymmetrized square (this, of course, uses up not only the set of UIR's that appear for the square corresponding to the double coset $\mathbf{G}^{\mathbf{k}}\{R_\beta \mid \mathbf{v}_\beta\}\mathbf{G}^{\mathbf{k}}$, but also its replica corresponding to the inverse double coset $\mathbf{G}^{\mathbf{k}}\{R_\beta \mid \mathbf{v}_\beta\}^{-1}\mathbf{G}^{\mathbf{k}}$).

Consider next $\alpha = 1$. For the square we have already evaluated the corresponding term $(D_p^{\mathbf{k}} \otimes D_\nu^{\mathbf{k}}) \uparrow \mathbf{G}$, using Eq. (4.9) with $\{R_\alpha \mid \mathbf{v}_\alpha\} = \{E \mid \mathbf{0}\}$. Inspection of Eqs. (2.5), (4.12), and (4.13) shows that all that is necessary is the additional computation of the following expressions:

$$\frac{|T|}{2\,|N|} \sum_{\{S|\mathbf{w}\}} (\chi_p^{\mathbf{k}^2}(\{S \mid \mathbf{w}\}) \pm \chi_p^{\mathbf{k}}(\{S \mid \mathbf{w}\}^2)) \times \chi_q^{\mathbf{h}^*}(\{R_\gamma \mid \mathbf{v}_\gamma\}^{-1}\{S \mid \mathbf{w}\}\{R_\gamma \mid \mathbf{v}_\gamma\}). \tag{4.14}$$

A nonzero value of the sum (4.14) with the plus sign contributes to the value of $[C_{pq}^{\mathbf{kh}}]$, and similarly a nonzero value with the minus sign contributes to the value of $\{C_{pq}^{\mathbf{kh}}\}$. Note that, if $D_p^{\mathbf{k}}$ is a 1-dimensional small UR, there is zero contribution to the antisymmetrized square, since in this case $\{D_p^{\mathbf{k}} \otimes D_p^{\mathbf{k}}\}$ vanishes identically.

Finally, we consider the self-inverse double cosets $\mathbf{G}^{\mathbf{k}}\{R_\alpha \mid \mathbf{v}_\alpha\}\mathbf{G}^{\mathbf{k}}$ in the set A. Following the prescription of Theorem 6 we choose an element $\{A \mid \mathbf{a}\} \in \{R_\alpha \mid \mathbf{v}_\alpha\}\mathbf{G}^{\mathbf{k}} \cap \mathbf{G}^{\mathbf{k}}\{R_\alpha \mid \mathbf{v}_\alpha\}^{-1}$. The group generated by $\mathbf{L}_\alpha^{\mathbf{k}}$ and $\{A \mid \mathbf{a}\}$ we denote by $\mathbf{M}_\alpha^{\mathbf{k}}$. Then, the UR's $N_{\alpha p}^{\mathbf{k}\pm}$ of $\mathbf{M}_\alpha^{\mathbf{k}}$ have the following characters [see Eqs. (3.25) and (3.36)]: for all $\{S \mid \mathbf{w}\} \in \mathbf{L}_\alpha^{\mathbf{k}}$,

$$\chi_{\alpha p}^{\mathbf{k}\pm}(\{S \mid \mathbf{w}\}) = \chi_p^{\mathbf{k}}(\{R_\alpha \mid \mathbf{v}_\alpha\}^{-1}\{S \mid \mathbf{w}\}\{R_\alpha \mid \mathbf{v}_\alpha\}) \times \chi_p^{\mathbf{k}}(\{S \mid \mathbf{w}\}), \tag{4.15}$$

$$\chi_{\alpha p}^{\mathbf{k}\pm}(\{A \mid \mathbf{a}\}\{S \mid \mathbf{w}\}) = \pm\chi_p^{\mathbf{k}}(\{A \mid \mathbf{a}\}\{S \mid \mathbf{w}\}\{A \mid \mathbf{a}\}\{S \mid \mathbf{w}\}). \tag{4.16}$$

Since for these α we are inducing from $\mathbf{M}_\alpha^{\mathbf{k}}$ rather than $\mathbf{L}_\alpha^{\mathbf{k}}$, comparison with the case of the Kronecker square

indicates that we now have to modify the analysis somewhat. We have to form the decomposition

$$\mathbf{G} = \sum_{\delta} \mathbf{M}_{\alpha}^{\mathbf{k}} \{R_{\delta} \mid \mathbf{v}_{\delta}\} \mathbf{G}^{\mathbf{h}} \tag{4.17}$$

and write $\mathbf{G}_{\delta}^{\mathbf{h}} = \{R_{\delta} \mid \mathbf{v}_{\delta}\} \mathbf{G}^{\mathbf{h}} \{R_{\delta} \mid \mathbf{v}_{\delta}\}^{-1}$. We also define $\mathbf{P}_{\alpha\delta}^{\mathbf{kh}} = \mathbf{M}_{\alpha}^{\mathbf{k}} \cap \mathbf{G}_{\delta}^{\mathbf{h}}$. For the fixed α under consideration, the allowed δ must satisfy Eq. (4.17) and

$$\mathbf{k} + \mathbf{k}_{\alpha} \equiv \mathbf{h}_{\delta} \tag{4.18}$$

and contributions to the frequency of $[C_{pq}^{\mathbf{kh}}]$ and $\{C_{pq}^{\mathbf{kh}}\}$ follow from evaluation of the sums

$$\frac{|T|}{|P|} \sum_{\{Q \mid \mathbf{x}\}} \chi_{\alpha p}^{\mathbf{k}\pm}(\{Q \mid \mathbf{x}\}) \chi_{q}^{\mathbf{h}^*}(\{R_{\delta} \mid \mathbf{v}_{\delta}\}^{-1} \{Q \mid \mathbf{x}\} \{R_{\delta} \mid \mathbf{v}_{\delta}\}), \tag{4.19}$$

where $|P|/|T|$ is the index of the subgroup $\mathbf{T}$ in the group $\mathbf{P}_{\alpha\delta}^{\mathbf{kh}}$ and the $\{Q \mid \mathbf{x}\}$, over which the summation is taken, are the coset representatives of $P_{\alpha\delta}^{\mathbf{kh}}$ with respect to $\mathbf{T}$. If $\{Q \mid \mathbf{x}\}$ belongs to $\mathbf{L}_{\alpha}^{\mathbf{k}}$, the character $\chi_{\alpha p}^{\mathbf{k}\pm}$ is calculated from Eq. (4.15); if not, from Eq. (4.16). Once again, only those values of $\mathbf{h}$ and q, for which $(D_q^{\mathbf{h}} \uparrow \mathbf{G})$ has already appeared in the evaluation $(N_{\alpha p}^{\mathbf{k}} \uparrow \mathbf{G})$ for the Kronecker square, need be tested by means of Eqs. (4.18) and (4.19).

Finally we relate some of our results in this section with those of Lax,[10] Sec. 6. The identification is made by noting that our $\mathbf{k}$, $\mathbf{k}_{\alpha}$, and $\mathbf{h}_{\gamma}$ correspond to his $\mathbf{k}$, $\mathbf{k}'$, and $\mathbf{k}''$. Also our $\{A \mid \mathbf{a}\}$ corresponds to his Q: this can be seen by remembering that $\{A \mid \mathbf{a}\} \in \{R_{\alpha} \mid \mathbf{v}_{\alpha}\} \mathbf{G}^{\mathbf{k}} \cap \mathbf{G}^{\mathbf{k}} \{R_{\alpha} \mid \mathbf{v}_{\alpha}\}^{-1}$, so that both $A\mathbf{k} \equiv \mathbf{k}_{\alpha}$ and $A\mathbf{k}_{\alpha} \equiv \mathbf{k}$. Hence $\{A \mid \mathbf{a}\}^2 \in \mathbf{G}^{\mathbf{k}} \cap \mathbf{G}_{\alpha}^{\mathbf{k}} = \mathbf{L}_{\alpha}^{\mathbf{k}}$. Also, there is no case ever to be considered for which $\{A \mid \mathbf{a}\}$ does not belong to $\mathbf{G}_{\gamma}^{\mathbf{h}}$. For it follows from Eq. (4.8) that

$$A\mathbf{h}_{\gamma} \equiv A(\mathbf{k} + \mathbf{k}_{\alpha}) \equiv \mathbf{k}_{\alpha} + \mathbf{k} \equiv \mathbf{h}_{\gamma}, \tag{4.20}$$

that is, $\{A \mid \mathbf{a}\} \in \mathbf{G}_{\gamma}^{\mathbf{h}}$. We have *proved* the statement leading to Lax's Eq. (6.2), that only such Q need to be considered, whereas in Lax's paper this statement rested only on certain typical examples given earlier. Note further from Eqs. (4.8) and (4.18) that $\mathbf{h}_{\gamma} \equiv \mathbf{h}_{\delta}$ and, hence, $\{A \mid \mathbf{a}\}$ belong to both $\mathbf{M}_{\alpha}^{\mathbf{k}}$ and $\mathbf{G}_{\delta}^{\mathbf{h}}$. It follows that

$$\mathbf{P}_{\alpha\delta}^{\mathbf{kh}} = \mathbf{N}_{\alpha\gamma}^{\mathbf{kh}} + \{A \mid \mathbf{a}\} \mathbf{N}_{\alpha\gamma}^{\mathbf{kh}}. \tag{4.21}$$

This corresponds to Lax's considering not the triple intersection group $\mathbf{G}_S$ but the group $(\mathbf{G}_S + Q\mathbf{G}_S)$. Finally, after some more identification of symbols, it can be seen that our formula (4.19) corresponds to Lax's formula (6.18). Of course, Lax was only considering transitions involving one particular triple of vectors $\mathbf{k}$, $\mathbf{k}'$, and $\mathbf{k}''$ and not the evaluation of the Clebsch–Gordan coefficients; but it is instructive to see how our work incorporates his as a by-product and, furthermore, fills in the slight gap noted above in the subgroup argument, by obtaining the relevant formulas, starting from considerations involving the representations of the full space group.

5. EXAMPLE AND APPLICATION TO ZINC-BLENDE

In dealing with the zinc-blende structure $\mathbf{T}_d^2$ we have taken care to conform as much as possible with the notation and conventions of previous authors who have treated face-centered-cubic space groups: Bouckaert, Smoluchowski, and Wigner,[12] Parmenter,[13] Dimmock and Wheeler,[14] and Altmann and Cracknell.[15] Since Birman[4] himself used the tables of Parmenter[13] in presenting results for the products of the single-valued UR's of $\mathbf{T}_d^2$, our results coincide with his wherever comparison is possible (Birman uses superscripts rather than suffixes, but this does not prevent comparison). However, Parmenter[13] does not distinguish between the two reflection planes in the group of Δ, leaving an ambiguity between the two UR's Δ_3 and Δ_4. Although it is true that these representations are time-reversal degenerate, this does not mean that they can be interchanged at will (every product has a unique decomposition). Similar ambiguities will always occur, unless authors give a label with a definite meaning to each of the symmetry elements under consideration and state explicitly whether they use an active or a passive convention for the action of the group elements (active operators move the field and leave the axes fixed; passive operators move the axes and leave the field fixed). In order to avoid such ambiguities, we adopt the notation for rotations and reflections defined by Altmann and Cracknell,[15] and like them we use the active convention. It is also necessary to distinguish carefully between the two elements in the double groups that correspond to each rotation or reflection; otherwise, similar ambiguities occur in the labeling of the double-valued UR's. This we do below by means of Eqs. (5.4)–(5.6).

We now give a summary of the meaning of the various notations we use.

Direct Lattice

This is the face-centered-cubic lattice based on the following lattice translations with coordinates referred to the x, y, z axes, a being the length of a cube edge:

$$\begin{aligned} \mathbf{t}_1 &= \tfrac{1}{2}a(0, 1, 1), \\ \mathbf{t}_2 &= \tfrac{1}{2}a(1, 0, 1), \\ \mathbf{t}_3 &= \tfrac{1}{2}a(1, 1, 0). \end{aligned} \tag{5.1}$$

Reciprocal Lattice

We use the convention

$$\mathbf{g}_i \cdot \mathbf{t}_j = 2\pi\delta_{ij}, \quad i,j = 1,2,3, \tag{5.2}$$

so that, with coordinates referred to the k_x, k_y, k_z axes, the reciprocal lattice vectors are

$$\begin{aligned} \mathbf{g}_1 &= (2\pi/a)(-1,1,1), \\ \mathbf{g}_2 &= (2\pi/a)(1,-1,1), \\ \mathbf{g}_3 &= (2\pi/a)(1,1,-1). \end{aligned} \tag{5.3}$$

Brillouin Zone

If we denote the vector $\mathbf{k} = \lambda\mathbf{g}_1 + \mu\mathbf{g}_2 + \nu\mathbf{g}_3$ by triple $(\lambda\mu\nu)$, then the points and lines of symmetry Γ, X, L, W, Λ, Δ, and Σ of the Brillouin zone (which are the only points and lines we treat) are as follows: $\Gamma(000)$, $X(\frac{1}{2}\,0\,\frac{1}{2})$, $L(\frac{1}{2}\,\frac{1}{2}\,\frac{1}{2})$, $W(\frac{1}{2}\,\frac{1}{4}\,\frac{3}{4})$, $\Lambda(p\,p\,p)$ $0 < p < \frac{1}{2}$, $\Delta(p\,0\,p)$ $0 < p < \frac{1}{2}$, and $\Sigma(p\,p\,2p)$ $0 < p \leq \frac{3}{8}$. See Fig. 4 of Bouckaert, Smoluchowski, and Wigner[12] with Δ on ΓX, Λ on ΓL, and Σ on ΓK (the point K has no more symmetry than Σ and corresponds to Σ with $p = \frac{3}{8}$); however, we use right-handed axes k_x, k_y, k_z, so that X is on the k_y axis (rather than on the k_x axis, as in the figure).

Point Group Elements

We use the same notation for the elements of the cubic group as Altmann and Cracknell.[15] We also use the active convention whereby the group elements move the field and leave the axes fixed.

Little Groups

The little groups (or groups of $\mathbf{k}$) are the products of the translations with the following point groups (little cogroups): for Γ, $\mathbf{T}_d(\bar{4}3m)$; for X, $\mathbf{D}_{2d}(\bar{4}2m)$; for L and Λ, $\mathbf{C}_{3v}(3m)$; for W, $\mathbf{S}_4(\bar{4})$; for Δ, $\mathbf{C}_{2v}(mm2)$; and for Σ, $\mathbf{C}_s(m)$. The elements appearing for the segment of the Brillouin zone defined above are as follows: for Γ, all elements of $\mathbf{T}_d$; for X, E, C_{2x}, C_{2y}, C_{2z}, S_{4y}^{+}, S_{4y}^{-}, σ_{dc}, σ_{de}; for L and Λ, E, C_{31}^{+}, C_{31}^{-}, σ_{db}, σ_{de}, σ_{df}; for W, E, C_{2x}, S_{4x}^{+}, S_{4x}^{-}; for Δ, E, C_{2y}, σ_{dc}, σ_{de}; and for Σ, E, σ_{db}.

Double Group Elements

If R is the proper rotation through a positive angle θ about an axis with direction cosines (l, m, n) (we use the right-hand-screw convention so that positive rotations are anticlockwise; for example, a unit vector along the x axis is moved into a position along the y axis by the operator C_{4z}^{+}, a positive rotation of $\frac{1}{2}\pi$ about the z axis), then we write

$$M(R) = \begin{pmatrix} \cos\frac{1}{2}\theta + in\sin\frac{1}{2}\theta & (-m+il)\sin\frac{1}{2}\theta \\ (m+il)\sin\frac{1}{2}\theta & \cos\frac{1}{2}\theta - in\sin\frac{1}{2}\theta \end{pmatrix} \tag{5.4}$$

and define

$$\bar{M}(R) = -M(R). \tag{5.5}$$

It has become customary to abbreviate $M(R)$ by R, and $\bar{M}(R)$ by $\bar{R}$, and we conform to this custom.

As an example, C_{31}^{+} has $\theta = \frac{2}{3}\pi$ and $l = m = n = (\frac{1}{3})^{\frac{1}{2}}$ and is, therefore, represented by the matrix

$$\frac{1}{2}\begin{pmatrix} 1+i & -1+i \\ 1+i & 1-i \end{pmatrix}.$$

$\bar{C}_{31}^{+}$ is represented by the negative of this matrix.

If Q is an improper notation $Q = IR$, where I is the inversion, so that R is a proper rotation, then we define

$$\begin{aligned} M(Q) &= -M(R), \\ \bar{M}(Q) &= M(R), \end{aligned} \tag{5.6}$$

where $M(R)$ is given by Eq. (5.4). $M(Q)$ is abbreviated by Q and $\bar{M}(Q)$ by $\bar{Q}$. All the double point groups can be constructed from the ordinary point groups in this way, and, of course, the elements of the matrices $M(R)$ are chosen to preserve the 2-to-1 homomorphism between the groups $SU(2)$ and $SO(3)$. In dealing with the double-valued UR's of $\mathbf{T}_d^2$, we need to use double-valued operators, so that, for example, the little cogroup of Σ contains E, $\bar{E}$, σ_{db}, $\bar{\sigma}_{db}$.

Character Tables

We need character tables for the double groups of $\mathbf{T}_d$, $\mathbf{D}_{2d}$, $\mathbf{C}_{3v}$, $\mathbf{S}_4$, $\mathbf{C}_{2v}$, and $\mathbf{C}_s$. These are given by Dimmock and Wheeler,[14] and we use their notation with the following elaborations to avoid the sort of ambiguities referred to above: For $\mathbf{C}_s$, their σ is our σ_{db}; for $\mathbf{C}_{2v}$, their σ_v is our σ_{dc}, and their σ_v' is our σ_{de} (this removes the ambiguity for Δ between the UR's Δ_3 and Δ_4); for $\mathbf{S}_4$, their S_4 is our S_{4x}^{+} and their S_4^{-1} is our S_{4x}^{-}; for $\mathbf{D}_{2d}$, their C_2 is our C_{2y} and their $2C_2'$ corresponds to our C_{2z} and C_{2x}. Their unbarred operators correspond to our unbarred operators in all cases, and similarly for the barred operators. As far as the labeling of the UR's is concerned we use the appropriate Brillouin zone point (instead of the arbitrary label Γ), together with the same distinguishing suffix as Dimmock and Wheeler[14]: thus, Δ_1, Δ_2, Δ_3, Δ_4, Δ_5 is our labeling for the UR's of the little group of Δ, and these correspond respectively to the UR's Γ_1, Γ_2, Γ_3, Γ_4, Γ_5 for $\mathbf{C}_{2v}$ listed by Dimmock and Wheeler.[14] To enable a speedy comparison with the work of Birman[4,5] we list in Table I the correspondence between our labeling as taken from Dimmock and Wheeler[14] and that used by Parmenter.[13]

As an example, we give intermediate results for the evaluation of the symmetrized and antisymmetrized

TABLE I. Comparison of notations for the labeling of the representations of the little groups of Γ, X, W, and L.[a,b,c]

P[d]	DW[e]	P[d]	DW[e]	P[e]	DW[e]	P[d]	DW[e]
Γ_1	Γ_1	X_1	X_1	W_1	W_1	L_1	L_1
Γ_2	Γ_2	X_4	X_2	W_2	W_2	L_2	L_2
Γ_{12}	Γ_3	X_2	X_3	W_3	W_3	L_3	L_3
Γ_{25}	Γ_4	X_3	X_4	W_4	W_4	L_6	L_4
Γ_{15}	Γ_5	X_5	X_5	W_5	W_5	L_4	L_5
Γ_6	Γ_6	X_6	X_6	W_7	W_6	L_5	L_6
Γ_7	Γ_7	X_7	X_7	W_6	W_7		
Γ_8	Γ_8			W_8	W_8		

[a] The notations of the authors coincide for Δ and Σ.
[b] Λ is not listed. Since the little group of Λ coincides with that of L, corresponding analogous differences in labeling occur for Λ exactly as for L.
[c] Where ambiguities occur in Parmenter's labeling we have listed only one of the various possibilities.
[d] The first column in each block indicates the notation used by Parmenter[13] and also by Birman.[4,5]
[e] The second column in each block indicates the notation used by Dimmock and Wheeler[14] and also by us.

squares of the UR's $X_i \uparrow \mathbf{T}_d^2$, $i = 1, 2, 3, 4, 5$. We restrict consideration to the single-valued UR's belonging to the point X, because an example involving the double-valued UR's would give no further insight into the methods involved, while at the same time being more complicated to present in the sense that all the groups involved would have twice as many elements. We use the notation of Sec. 4 without further introduction. $\mathbf{k} = X = (\frac{1}{2}\ 0\ \frac{1}{2})$ and $\mathbf{G}^X$ is the group consisting of the translations and their products with the eight elements E, $S_{4y}^{\pm}$, C_{2y}, C_{2x}, C_{2z}, σ_{dc}, σ_{de}. The character table of its single-valued small UR's is shown in Table II.

The star of X consists of three vectors: $X = (\frac{1}{2}\ 0\ \frac{1}{2})$, $C_{31}^{+}X = (\frac{1}{2}\ \frac{1}{2}\ 0)$ and $C_{31}^{-}X = (0\ \frac{1}{2}\ \frac{1}{2})$.

The expansion (4.5) has two terms with $\alpha_1 = \{E \mid \mathbf{0}\}$ and $\alpha_2 = \{C_{31}^{+} \mid \mathbf{0}\}$. $\mathbf{G}_E^X = \mathbf{G}^X$ and $\mathbf{G}_{C_{31}^{+}}^X$, being the little group of $C_{31}^{+}X$, is the group consisting of the translations and their products with the eight elements E, $S_{4z}^{\pm}$, C_{2z}, C_{2x}, C_{2y}, σ_{da}, σ_{db}. Thus, $\mathbf{L}_E^X = \mathbf{G}^X$ and $\mathbf{L}_{C_{31}^{+}}^X$ is the group consisting of the translations and their products with the four elements E, C_{2x}, C_{2y}, C_{2z}.

Consider first $\alpha_1 = \{E \mid \mathbf{0}\}$. From Eq. (4.8), we see that there can be no contribution unless $\mathbf{h}_\gamma = X + X = (101) \equiv \Gamma$. Since $\mathbf{G}^\Gamma = \mathbf{T}_d^2$, the expansion (4.7) has only one term with $\gamma_1 = \{E \mid \mathbf{0}\}$ corresponding to an allowed solution $\mathbf{h} = \Gamma$. Also $\mathbf{N}_{EE}^{X\Gamma} = \mathbf{G}^X$.

Consider next $\alpha_2 = \{C_{31}^{+} \mid \mathbf{0}\}$. We must now have $\mathbf{h}_\gamma = X + C_{31}^{+}X = (1\ \frac{1}{2}\ \frac{1}{2}) \equiv C_{31}^{-}X$, so we hope for an allowed solution $\mathbf{h} = X$. The expansion (4.7) now has three terms with $\gamma_1 = \{E \mid \mathbf{0}\}$, $\gamma_2 = \{C_{31}^{-} \mid \mathbf{0}\}$, and $\gamma_3 = \{C_{31}^{+} \mid \mathbf{0}\}$. From Eq. (4.8), the only allowed contribution arises from $\gamma_2 = \{C_{31}^{-} \mid 0\}$, and

$$\mathbf{N}_{C_{31}^{+}C_{31}^{-}}^{XX} = \mathbf{L}_{C_{31}^{+}}^X \cap \mathbf{G}_{C_{31}^{+}}^X = \mathbf{L}_{C_{31}^{+}}^X.$$

Using the sum (4.9) with the allowed pair (α_1, γ_1) and $\mathbf{h} = \Gamma$, we find that for $p = 1, 2, 3, 4$ we get contribution 1 for $q = 1, 3$, and 0 for $q = 2, 4, 5$, and for $p = 5$ we get contribution 1 for $q = 1, 2, 4, 5$ and 2 for $q = 3$. Then, with the allowed pair

$$(\alpha_2, \gamma_2) = (\{C_{31}^{+} \mid \mathbf{0}\}, \{C_{31}^{-} \mid \mathbf{0}\}),$$

the sum (4.9) becomes

$$\tfrac{1}{4} \sum_R \chi_p^X(R)\chi_p^X(C_{31}^{-}RC_{31}^{+})\chi_q^X(C_{31}^{+}RC_{31}^{-}), \qquad (5.7)$$

in which the sum is over $R = E$, C_{2x}, C_{2y}, C_{2z}. For $p = 5$, for example, the sum (5.7) reduces to $\chi_q^X(E)$, thereby yielding contribution 1 for $q = 1, 2, 3, 4$, and 2 for $q = 5$. Similarly, for $p = 1, 3$ there is contribution 1 for $q = 1, 3$, and 0 for $q = 2, 4, 5$, and for $p = 2, 4$ there is contribution 1 for $q = 2, 4$, and 0 for $q = 1, 3, 5$.

Putting all these results together, we obtain for the Kronecker squares (writing for brevity $X_i \uparrow \mathbf{G}$ as X_i, etc.)

$$X_1 \otimes X_1 = \Gamma_1 + \Gamma_3 + X_1 + X_3 = X_3 \otimes X_3, \qquad (5.8)$$

$$X_2 \otimes X_2 = \Gamma_1 + \Gamma_3 + X_2 + X_4 = X_4 \otimes X_4, \qquad (5.9)$$

$$X_5 \otimes X_5 = \Gamma_1 + \Gamma_2 + 2\Gamma_3 + \Gamma_4 + \Gamma_5 + X_1 + X_2 + X_3 + X_4 + 2X_5. \qquad (5.10)$$

Turning now to the symmetrized and antisymmetrized squares, we first note that α_1 corresponds to the identity double coset and that α_2 corresponds to a self-inverse double coset and, therefore, belongs to the set A. For α_1 we must, therefore, evaluate the sum (4.14), with $\gamma = \gamma_1 = \{E \mid \mathbf{0}\}$ and $\mathbf{h} = \Gamma$. Also, we can restrict ourselves to the values of q appearing as subscripts to Γ in Eqs. (5.8)–(5.10). We find that there is a contribution to the antisymmetrized square only for $p = 5$ and $q = 4$, so that all other Γ terms in Eqs. (5.8)–(5.10) must belong to the symmetrized squares. In this example, the set B is void (double cosets belonging to B are so straightforward to deal with that the instructiveness of the example is not lost). It remains to deal with $\alpha_2 = \{C_{31}^{+} \mid \mathbf{0}\} \in A$. We form the intersection $C_{31}^{+}\mathbf{G}^X \cap \mathbf{G}^X C_{31}^{-}$. It contains products of

TABLE II. The small UR's of X.[a,b]

X	E	$S_{4y}^{\pm}$	C_{2y}	C_{2x}, C_{2z}	σ_{dc}, σ_{de}
X_1	1	1	1	1	1
X_2	1	1	1	−1	−1
X_3	1	−1	1	1	−1
X_4	1	−1	1	−1	1
X_5	2	0	−2	0	0

[a] $\{E \mid \mathbf{t}_1\}$ and $\{E \mid \mathbf{t}_3\}$ are represented by −1 and $\{E \mid \mathbf{t}_2\}$ by +1.
[b] The symmetry elements are to be identified from Fig. 4 of Altmann and Cracknell.[15]

the translations with the four elements S_{4x}^{+}, S_{4x}^{-}, σ_{dd}, σ_{df}. We fix on σ_{dd} for the element $\{A \mid \mathbf{a}\}$. Then,

$$\mathbf{M}_{C_{31}^{+}}^{X} = (\mathbf{L}_{C_{31}^{+}}^{X} + \sigma_{dd}\mathbf{L}_{C_{31}^{+}}^{X})$$

consists of the translations and their products with the eight elements E, C_{2x}, C_{2y}, C_{2z}, S_{4x}^{+}, S_{4x}^{-}, σ_{dd}, σ_{df}. The UR's $N_{C_{31}^{+}p}^{X+}$, $p = 1$ to 5, of the group $\mathbf{M}_{C_{31}}^{X+}$ are found from Eqs. (4.15) and (4.16) and are shown in Table III.

The decomposition (4.17) yields two double cosets $\delta_1 = \{E \mid \mathbf{0}\}$ and $\delta_2 = \{C_{31}^{-} \mid \mathbf{0}\}$ [note that the corresponding decomposition (4.7) had three terms]. Of these only $\delta_2 = \{C_{31}^{-} \mid \mathbf{0}\}$ is compatible with the solution $\mathbf{h} = X$ of Eq. (4.18). We must, therefore, evaluate the sum (4.19), which reduces to

$$\sum_{Q} \chi_{C_{31}^{+}p}^{X+}(Q)\chi_{q}^{X}(C_{31}^{+}QC_{31}^{-}), \tag{5.11}$$

where the sum is over $Q = E$, $S_{4x}^{\pm}$, C_{2x}, C_{2y}, C_{2z}, σ_{dd}, σ_{df}. Using Tables II and III, we find that we obtain from (5.11) a contribution 1 for the following pairs of values: $p = 1$, $q = 1$; $p = 2$, $q = 4$; $p = 3$, $q = 1$; $p = 4$, $q = 4$; $p = 5$, $q = 1, 4, 5$; and zero otherwise. Other pairs of values of p and q appearing as subscripts to X in Eqs. (5.8)–(5.10) must belong to the antisymmetrized squares.

Finally (using once more the abbreviation X_i for $X_i \uparrow \mathbf{G}$, etc.) we obtain, for the symmetrized squares,

$$[X_1 \otimes X_1] = \Gamma_1 + \Gamma_3 + X_1 = [X_3 \otimes X_3], \tag{5.12}$$

$$[X_2 \otimes X_2] = \Gamma_1 + \Gamma_3 + X_4 = [X_4 \otimes X_4], \tag{5.13}$$

$$[X_5 \otimes X_5] = \Gamma_1 + \Gamma_2 + 2\Gamma_3 + \Gamma_5 + X_1 + X_4 + X_5, \tag{5.14}$$

and, for the antisymmetrized squares,

$$\{X_1 \otimes X_1\} = X_3 = \{X_3 \otimes X_3\}, \tag{5.15}$$

$$\{X_2 \otimes X_2\} = X_2 = \{X_4 \otimes X_4\}, \tag{5.16}$$

$$\{X_5 \otimes X_5\} = \Gamma_4 + X_2 + X_3 + X_5. \tag{5.17}$$

TABLE III. The UR's of $\mathbf{M}_{C_{31}^{+}}^{X}$.[a]

p	E	S_{4x}^{+}, S_{4x}^{-}	C_{2x}	C_{2x}, C_{2z}	σ_{dd}, σ_{df}
1	1	1	1	1	1
2	1	−1	1	−1	1
3	1	1	1	1	1
4	1	−1	1	−1	1
5	4	0	0	0	2

[a] $\{E \mid \mathbf{t}_2\}$ and $\{E \mid \mathbf{t}_3\}$ are represented by −1 and $\{E \mid \mathbf{t}_1\}$ by 1. See also footnote b to Table II.

We have derived expansions for all the important Kronecker products and for all the symmetrized and antisymmetrized squares of the UIR's of the zinc-blende structure associated with the Brillouin zone points Γ, X, L, W, Λ, and Δ, and these results are presented in Appendix A in coded form. Before the table is used, reference should be made to the paragraphs at the beginning of Sec. 5 headed direct lattice, reciprocal lattice, Brillouin zone, point group elements, little groups, double group elements, and character tables. These paragraphs permit a completely unambiguous identification of the UIR's appearing in the table. The notes preceding the table should also be consulted, as they contain the rules for decoding the entries that appear in the table.

ACKNOWLEDGMENTS

Dr. B. L. Davies is indebted to the United Kingdom Science Research Council for a grant during the tenure of which part of the above work was carried out. Dr. C. J. Bradley wishes to thank Mr. N. B. Backhouse for some useful discussions on the applications of group theory. The authors also wish to thank Dr. Ll. G. Chambers for his help in translating the paper by Karavaev.[6]

APPENDIX: KRONECKER PRODUCTS, SYMMETRIZED SQUARES AND ANTISYMMETRIZED SQUARES FOR THE ZINC-BLENDE STRUCTURE $\mathbf{T}_d^2$

(i) The entries *ai*, *xi*, *wi*, *bi*, *pi*, *di*, *si* stand, respectively, for $(\Gamma_i \uparrow \mathbf{G})$, $(X_i \uparrow \mathbf{G})$, $(W_i \uparrow \mathbf{G})$, $(L_i \uparrow \mathbf{G})$, $(\Lambda_i \uparrow \mathbf{G})$, $(\Delta_i \uparrow \mathbf{G})$, $(\Sigma_i \uparrow \mathbf{G})$, where $\mathbf{G}$ is the space group $\mathbf{T}_d^2$ and Γ_i, X_i, $\cdots$, etc., are the labels of small UR's of the Brillouin zone points, Γ, X, $\cdots$, etc. An entry such as 3*d*5 will, therefore, mean $3(\Delta_5 \uparrow \mathbf{G})$, the direct sum of three copies of $(\Delta_5 \uparrow \mathbf{G})$.

(ii) The table is to be read as if in one continuous column, though for typographical reasons it is printed in several columns over a number of pages. On each page the left-hand column comes first and is to be read from top to bottom before proceeding to the next column. The right-hand column on a page is followed by the left-hand column on the next page.

(iii) The table is separated into 19 sections each section headed by symbols such as $X \otimes X$, $[X \otimes X]$, $X \otimes L$. In the section headed $X \otimes X$, we list the 28 expansions of the inner Kronecker products $(X_i \uparrow \mathbf{G}) \otimes (X_j \uparrow \mathbf{G})$, $1 \leq i \leq j \leq 7$. In the section headed $[X \otimes X]$, we list the seven expansions of the symmetrized squares $[(X_i \uparrow \mathbf{G}) \otimes (X_i \uparrow \mathbf{G})]$, $i = 1$ to 7. And in the section headed $X \otimes L$, we list the 42 expansions of the inner Kronecker products of $(X_i \uparrow \mathbf{G}) \otimes (L_j \uparrow \mathbf{G})$,

$\Gamma \otimes \Gamma$
$a1$
$a2$
$a3$
$a4$
$a5$
$a6$
$a7$
$a8$
$a1$
$a3$
$a5$
$a4$
$a7$
$a6$
$a8$
$a1$
$a2$
$a3$
$a4$
$a5$
$a4$
$a5$
$a8$
$a1$
$a3$
$a6$
$a7$
$a8$
$a1$
$a3$
$a4$
$a5$
$a2$
$a3$
$a4$
$a5$
$a6$
$a8$
$a7$
$a8$
$a6$
$a7$
$2a8$

$a1$
$a3$
$a4$
$a5$
$a7$
$a8$
$a6$
$a8$
$a6$
$a7$
$2a8$
$a1$
$a4$
$a2$
$a5$
$a3$
$a4$
$a5$
$a1$
$a4$
$a3$
$a4$
$a5$
$a1$
$a2$
$a3$
$2a4$
$2a5$
$[\Gamma \otimes \Gamma]$
$a1$
$a1$
$a1$
$a3$
$a1$
$a3$
$a5$
$a1$
$a3$
$a5$
$a4$
$a4$
$a2$
$2a4$
$a5$
$\Gamma \otimes X$
$x1$
$x2$

$x3$
$x4$
$x5$
$x6$
$x7$
$x3$
$x4$
$x1$
$x2$
$x5$
$x7$
$x6$
$x1$
$x3$
$x2$
$x4$
$x1$
$x3$
$x2$
$x4$
$2x5$
$x6$
$x7$
$x6$
$x7$
$x2$
$x5$
$x1$
$x5$
$x4$
$x5$
$x3$
$x5$
$x1$
$x2$
$x3$
$x4$
$x5$
$2x6$
$x7$
$x6$
$2x7$

$x4$
$x5$
$x3$
$x5$
$x2$
$x5$
$x1$
$x5$
$x1$
$x2$
$x3$
$x4$
$x5$
$x6$
$2x7$
$2x6$
$x7$
$x6$
$x6$
$x7$
$x7$
$x6$
$x7$
$x1$
$x2$
$x5$
$x3$
$x4$
$x5$
$x7$
$x7$
$x6$
$x6$
$x6$
$x7$
$x6$
$x7$
$x6$

$x7$
$x6$
$x7$
$2x6$
$2x7$
$x1$
$x2$
$x3$
$x4$
$2x5$
$x1$
$x2$
$x3$
$x4$
$2x5$
$\Gamma \otimes L$
$b1$
$b2$
$b3$
$b4$
$b5$
$b6$
$b2$
$b1$
$b3$
$b4$
$b6$
$b2$
$b3$
$b5$
$b3$
$b4$
$b1$
$b2$
$b3$
$b4$
$b5$
$b6$
$b1$
$b3$
$b1$

$b2$
$2b3$
$2b4$
$b5$
$b6$
$b4$
$b6$
$b4$
$b5$
$b1$
$b3$
$b2$
$b3$
$b1$
$b2$
$2b3$
$2b4$
$b5$
$b6$
$b4$
$b5$
$b4$
$b6$
$b4$
$b2$
$b4$
$b1$
$b4$
$b5$
$b6$
$b1$
$b2$
$b3$
$b3$
$b3$
$b4$
$b4$
$b1$
$b2$
$b3$
$b4$
$b5$
$b6$

$b4$
$b5$
$b6$
$3b4$
$b5$
$b6$
$b1$
$b2$
$3b3$
$b1$
$b2$
$b3$
$b1$
$b2$
$b3$
$\Gamma \otimes W$
$w1$
$w2$
$w3$
$w4$
$w5$
$w6$
$w7$
$w8$
$w1$
$w2$
$w4$
$w3$
$w7$
$w8$
$w2$
$w4$
$w3$
$w7$
$w8$
$w1$
$w5$
$w6$
$w8$
$w4$
$w5$
$w7$
$w8$
$w1$
$w2$
$w3$
$w4$
$w5$
$w6$
$w7$
$w8$

$w6$
$w8$
$w5$
$w7$
$w6$
$w8$
$w1$
$w3$
$w4$
$w2$
$w3$
$w4$
$w1$
$w2$
$w3$
$w1$
$w2$
$w4$
$w5$
$w6$
$w8$
$w5$
$w6$
$w7$
$w6$
$w7$
$w8$
$w5$
$w7$
$w8$
$w2$
$w3$
$w4$
$w1$
$w3$
$w4$
$w1$
$w2$
$w4$
$w1$
$w2$
$w3$
$w6$
$w7$
$w8$
$w5$
$w7$
$w8$
$w5$
$w6$
$w8$

(*continued*)

$w5$	$w7$	$d4$	$x1$	$x2$	$a3$	$2b4$	
$w6$	$w8$		$x3$	$x4$	$a4$	$b5$	$b1$
$w7$		$d1$		$x5$	$2a5$	$b6$	$b2$
	$w1$	$d2$	$a4$		$x1$		$2b3$
$w5$	$w2$	$d3$	$x5$	$a7$	$x2$	$b4$	
$w6$	$w3$			$a8$	$x3$	$b6$	$b1$
	$w4$	$3d5$	$a2$	$x6$	$x4$		$b2$
$w7$			$a3$	$x7$	$2x5$	$b4$	$2b3$
$w8$	$w1$	$d1$	$x1$			$b5$	
	$w2$	$d3$	$x3$	$a6$	$a1$		$2b1$
$w5$	$w3$	$d4$		$a8$	$a3$	$b2$	$2b2$
$w8$	$w4$		$a5$	$x6$	$2a4$	$b3$	$4b3$
		$d2$	$x5$	$x7$	$a5$		
$w6$	$w1$	$d3$			$x1$	$b1$	$4b4$
$w7$	$w2$	$d4$	$a4$	$a1$	$x2$	$b3$	$2b5$
	$w3$		$a5$	$a3$	$x3$		$2b6$
$w1$	$w4$	$d1$	$x2$	$x2$	$x4$	$b1$	
$w3$		$d2$	$x4$	$x4$	$2x5$	$b2$	$2b4$
	$w1$	$d3$	$x5$			$2b3$	$b5$
$w1$	$w2$			$a4$	$[X \otimes X]$		$b6$
$w4$	$w3$	$d1$	$a6$	$a5$		$2b4$	
	$w4$	$d2$	$a8$	$x1$	$a1$	$b5$	$2b4$
$w2$		$d4$	$x6$	$x3$	$a3$	$b6$	$b5$
$w4$	$\Gamma \otimes \Delta$		$x7$	$x5$	$x1$		$b6$
		$3d5$				$b4$	
$w2$	$d1$		$a7$	$a7$	$a1$	$b5$	$2b4$
$w3$		$d5$	$a8$	$a8$	$a3$		$b5$
	$d2$		$x6$	$x6$	$x4$	$b4$	$b6$
$w7$		$d5$	$x7$	$x7$		$b6$	
$w8$	$d3$				$a1$		$2b4$
		$d5$	$a1$	$a6$	$a3$	$b2$	$b5$
$w5$	$d4$		$a3$	$a8$	$x1$	$b3$	$b6$
$w6$		$d5$	$x2$	$x6$			
	$d5$		$x4$	$x7$	$a1$	$b1$	$4b4$
$w6$		$d1$			$a3$	$b3$	$2b5$
$w7$	$d2$	$d2$	$a5$	$a1$	$x4$		$2b6$
		$d3$	$x5$	$a2$		$b1$	
$w5$	$d1$	$d4$		$2a3$	$a1$	$b2$	$2b1$
$w8$			$a2$	$a4$	$a2$	$2b3$	$2b2$
	$d4$	$d5$	$a3$	$a5$	$2a3$		$4b3$
$w2$			$x2$	$x1$	$a5$	$2b4$	
$w4$	$d3$	$d5$	$x4$	$x2$	$x1$	$b5$	$b1$
				$x3$	$x4$	$b6$	$b2$
$w2$	$d5$	$d5$	$a4$	$x4$	$x5$		$2b3$
$w3$			$a5$	$2x5$		$b4$	
	$d1$	$d5$	$x1$		$2a4$	$b5$	$b1$
$w1$	$d2$		$x3$	$a6$	$a5$		$b2$
$w3$		$d1$	$x5$	$a7$	$x2$	$b4$	$2b3$
	$d1$	$d2$		$2a8$	$x3$	$b6$	
$w1$	$d2$	$d3$	$a6$	$2x6$	$x5$		$2b4$
$w4$		$d4$	$a8$	$2x7$		$b1$	$b5$
	$d3$		$x6$		$2a4$	$b3$	$b6$
$w5$	$d4$	$2d5$	$x7$	$a6$	$a5$		
$w6$				$a7$	$x2$	$b2$	$2b4$
$w7$	$d3$	$2d5$	$a7$	$2a8$	$x3$	$b3$	$b5$
$w8$	$d4$		$a8$	$2x6$	$x5$		$b6$
		$2d5$	$x6$	$2x7$		$b1$	
$w5$	$2d5$		$x7$		$X \otimes L$	$b2$	$4b4$
$w6$		$2d5$		$a1$		$2b3$	$2b5$
$w7$	$d2$		$a1$	$a3$	$b1$		$2b6$
$w8$	$d3$	$2d1$	$a3$	$2a4$	$b3$	$2b4$	
	$d4$	$2d2$	$x1$	$a5$		$b5$	$2b1$
$w5$		$2d3$	$x3$	$x1$	$b2$	$b6$	$2b2$
$w6$	$d1$	$2d4$		$x2$	$b3$		$4b3$
$w7$	$d3$		$a4$	$x3$		$b4$	
$w8$	$d4$	$X \otimes X$	$x5$	$x4$	$b1$	$b6$	$b1$
				$2x5$	$b2$		$b2$
$w5$	$d1$	$a1$	$a4$		$2b3$	$b4$	$2b3$
$w6$	$d2$	$a3$	$a5$	$a2$		$b5$	

(*continued*)

$b1$
$b2$
$2b3$
$X \otimes W$
$w1$
$d1$
$d2$
$w2$
$d1$
$d2$
$w4$
$d3$
$d4$
$w3$
$d3$
$d4$
$w6$
$d5$
$w5$
$d5$
$w8$
$d5$
$w7$
$d5$
$w1$
$d3$
$d4$
$w2$
$d3$
$d4$
$w4$
$d1$
$d2$
$w3$
$d1$
$d2$
$w6$
$d5$
$w5$
$d5$
$w8$
$d5$
$w7$
$d5$
$w2$
$d1$
$d2$
$w1$
$d1$
$d2$

$w3$
$d3$
$d4$
$w4$
$d3$
$d4$
$w8$
$d5$
$w7$
$d5$
$w6$
$d5$
$w5$
$d5$
$w2$
$d3$
$d4$
$w1$
$d3$
$d4$
$w3$
$d1$
$d2$
$w4$
$d1$
$d2$
$w8$
$d5$
$w7$
$d5$
$w6$
$d5$
$w5$
$d5$
$w3$
$w4$
$d1$
$d2$
$d3$
$d4$
$w3$
$w4$
$d1$
$d2$
$d3$
$d4$
$w1$
$w2$
$d1$
$d2$
$d3$
$d4$

$w1$
$w2$
$d1$
$d2$
$d3$
$d4$
$w5$
$w7$
$2d5$
$w6$
$w8$
$2d5$
$w5$
$w7$
$2d5$
$w6$
$w8$
$2d5$
$w5$
$w6$
$2d5$
$w7$
$w8$
$2d5$
$w6$
$w7$
$2d5$
$w5$
$w8$
$2d5$
$w1$
$w4$
$d1$
$d2$
$d3$
$d4$
$w1$
$w3$
$d1$
$d2$
$d3$
$d4$
$w2$
$w4$
$d1$
$d2$
$d3$
$d4$
$w7$

$w8$
$2d5$
$w5$
$w6$
$2d5$
$w5$
$w8$
$2d5$
$w6$
$w7$
$2d5$
$w2$
$w3$
$d1$
$d2$
$d3$
$d4$
$w2$
$w4$
$d1$
$d2$
$d3$
$d4$
$w1$
$w4$
$d1$
$d2$
$d3$
$d4$
$w1$
$w3$
$d1$
$d2$
$d3$
$d4$
$L \otimes L$
$a1$
$a5$
$x1$
$x4$
$x5$
$a2$
$a4$
$x2$
$x3$
$x5$
$a3$
$a4$
$a5$
$x1$
$x2$
$x3$
$x4$
$2x5$
$a6$
$a7$
$a8$

$2x6$
$2x7$
$a8$
$x6$
$x7$
$a8$
$x6$
$x7$
$a1$
$a5$
$x1$
$x4$
$x5$
$a3$
$a4$
$a5$
$x1$
$x2$
$x3$
$x4$
$2x5$
$a6$
$a7$
$a8$
$2x6$
$2x7$
$a8$
$x6$
$x7$
$a8$
$x6$
$x7$
$a1$
$a2$
$a3$
$2a4$
$2a5$
$2x1$
$2x2$
$2x3$
$2x4$
$4x5$
$a6$
$a7$
$3a8$
$4x6$
$4x7$
$a6$
$a7$
$a8$
$2x6$
$2x7$
$a6$
$a7$
$a8$
$2x6$
$2x7$

$a1$
$a2$
$a3$
$2a4$
$2a5$
$2x1$
$2x2$
$2x3$
$2x4$
$4x5$
$a3$
$a4$
$a5$
$x1$
$x2$
$x3$
$x4$
$2x5$
$a3$
$a4$
$a5$
$x1$
$x2$
$x3$
$x4$
$2x5$
$a2$
$a4$
$x1$
$x4$
$x5$
$a1$
$a5$
$x2$
$x3$
$x5$
$a2$
$a4$
$x1$
$x4$
$x5$
$[L \otimes L]$
$a1$
$a5$
$x1$
$x4$
$a1$
$a5$
$x1$
$x4$
$a1$
$a3$
$a4$
$2a5$
$2x1$
$x2$
$x3$
$2x4$
$x5$

$a2$
$a3$
$2a4$
$a5$
$x2$
$x3$
$3x5$
$a2$
$a4$
$x5$
$a2$
$a4$
$x5$
$L \otimes W$
$s1$
$s2$
$s1$
$s2$
$s1$
$s2$
$s1$
$s2$
$s3$
$s4$
$s3$
$s4$
$s3$
$s4$
$s3$
$s4$
$s1$
$s2$
$s1$
$s2$
$s1$
$s2$
$s1$
$s2$
$s3$
$s4$
$s3$
$s4$
$s3$
$s4$
$s3$
$s4$
$s3$
$s4$
$2s1$
$2s2$

$2s1$
$2s2$
$2s1$
$2s2$
$2s1$
$2s2$
$2s3$
$2s4$
$2s3$
$2s4$
$2s3$
$2s4$
$2s3$
$2s4$
$2s3$
$2s4$
$2s3$
$2s4$
$2s3$
$2s4$
$2s3$
$2s4$
$2s1$
$2s2$
$2s1$
$2s2$
$2s1$
$2s2$
$2s1$
$2s2$
$s3$
$s4$
$s3$
$s4$
$s3$
$s4$
$s3$
$s4$
$s1$
$s2$
$s1$
$s2$
$s1$
$s2$

(*continued*)

$s3$
$s4$
$s3$
$s4$
$s3$
$s4$
$s3$
$s4$
$s1$
$s2$
$s1$
$s2$
$s1$
$s2$
$s1$
$s2$
$W \otimes W$
$a1$
$a3$
$a4$
$x1$
$x2$
$s1$
$s2$
$a2$
$a3$
$a5$
$x3$
$x4$
$s1$
$s2$
$a4$
$a5$
$x5$
$s1$
$s2$
$a4$
$a5$
$x5$
$s1$
$s2$
$a6$
$a8$
$x6$
$s3$
$s4$
$a6$
$a8$
$x6$
$s3$
$s4$
$a6$
$a8$
$x6$
$s3$
$s4$
$a7$
$a8$
$x7$
$s3$
$s4$

$a7$
$a8$
$x7$
$s3$
$s4$
$a1$
$a3$
$a4$
$x1$
$x2$
$s1$
$s2$
$a4$
$a5$
$x5$
$s1$
$s2$
$a4$
$a5$
$x5$
$s1$
$s2$
$a7$
$a8$
$x7$
$s3$
$s4$
$a7$
$a8$
$x7$
$s3$
$s4$
$a6$
$a8$
$x6$
$s3$
$s4$
$a6$
$a8$
$x6$
$s3$
$s4$
$a1$
$a3$
$a4$
$x3$
$x4$
$s1$
$s2$
$a2$
$a3$
$a5$
$x1$
$x2$
$s1$
$s2$
$a6$
$a8$
$x7$

$s3$
$s4$
$a7$
$a8$
$x6$
$s3$
$s4$
$a7$
$a8$
$x6$
$s3$
$s4$
$a6$
$a8$
$x7$
$s3$
$s4$
$a1$
$a3$
$a4$
$x3$
$x4$
$s1$
$s2$
$a7$
$a8$
$x6$
$s3$
$s4$
$a6$
$a8$
$x7$
$s3$
$s4$
$a6$
$a8$
$x7$
$s3$
$s4$
$a7$
$a8$
$x6$
$s3$
$s4$
$a1$
$a3$
$a4$
$x5$
$s1$
$s2$
$a4$
$a5$
$x1$
$x2$
$s1$
$s2$
$a2$
$a3$
$a5$
$x5$

$s1$
$s2$
$a4$
$a5$
$x3$
$x4$
$s1$
$s2$
$a1$
$a3$
$a4$
$x5$
$s1$
$s2$
$a4$
$a5$
$x3$
$x4$
$s1$
$s2$
$a2$
$a3$
$a5$
$x5$
$s1$
$s2$
$a1$
$a3$
$a4$
$x5$
$s1$
$s2$
$a4$
$a5$
$x1$
$x2$
$s1$
$s2$
$a1$
$a3$
$a4$
$x5$
$s1$
$s2$
$[W \otimes W]$
$a1$
$a3$
$x1$
$x2$
$s1$
$a1$
$a3$
$x1$
$x2$
$s1$
$a1$
$a3$
$x3$
$x4$
$s1$

$a1$
$a3$
$x3$
$x4$
$s1$
$a4$
$x5$
$s2$
$a4$
$x5$
$s2$
$a4$
$x5$
$s2$
$a4$
$x5$
$s2$
$\Lambda \otimes \Lambda$
$p1$
$d1$
$d4$
$p2$
$d2$
$d3$
$p3$
$d1$
$d2$
$d3$
$d4$
$p4$
$2d5$
$p5$
$d5$
$p6$
$d5$
$p1$
$d1$
$d4$
$p3$
$d1$
$d2$
$d3$
$d4$
$p4$
$2d5$
$p6$
$d5$
$p5$
$d5$
$p1$
$p2$
$p3$
$2d1$

$2d2$
$2d3$
$2d4$
$p4$
$p5$
$p6$
$4d5$
$p4$
$2d5$
$p4$
$2d5$
$p1$
$p2$
$p3$
$2d1$
$2d2$
$2d3$
$2d4$
$p3$
$d1$
$d2$
$d3$
$d4$
$p3$
$d1$
$d2$
$d3$
$d4$
$p2$
$d1$
$d4$
$p1$
$d2$
$d3$
$p2$
$d1$
$d4$
$[\Lambda \otimes \Lambda]$
$p1$
$d1$
$p1$
$d1$
$p1$
$p3$
$2d1$
$d2$
$d3$
$p2$
$p3$
$d2$
$d3$
$2d4$
$p2$
$d4$
$p2$

$d4$
$\Delta \otimes \Delta$
$d1$
$a1$
$a3$
$a5$
$s1$
$s2$
$d2$
$a2$
$a3$
$a4$
$s1$
$s2$
$d3$
$a4$
$a5$
$s1$
$s2$
$d4$
$a4$
$a5$
$s1$
$s2$
$d5$
$a6$
$a7$
$2a8$
$2s3$
$2s4$
$d1$
$a1$
$a3$
$a5$
$s1$
$s2$
$d4$
$a4$
$a5$
$s1$
$s2$
$d3$
$a4$
$a5$
$s1$
$s2$
$d2$
$a2$
$a3$
$a4$
$s1$
$s2$
$d1$
$a1$
$a3$
$a5$
$s1$
$s2$
$d4$
$a4$
$a5$
$s1$
$s2$
$d3$
$a4$
$a5$
$s1$
$s2$
$d5$
$a6$
$a7$
$2a8$
$2s3$
$2s4$
$d1$
$a2$
$a3$
$a4$
$s1$
$s2$
$d2$

$a1$
$a3$
$a5$
$s1$
$s2$
$d5$
$a6$
$a7$
$2a8$
$2s3$
$2s4$
$d1$
$a2$
$a3$
$a4$
$s1$
$s2$
$d5$
$a6$
$a7$
$2a8$
$2s3$
$2s4$
$d1$
$d2$
$d3$
$d4$
$a1$
$a2$
$2a3$
$3a4$
$3a5$
$4s1$
$4s1$
$[\Delta \otimes \Delta]$
$a1$
$a3$
$d1$
$s1$
$a1$
$a3$
$d1$
$s1$
$a2$
$a3$
$d1$
$s1$
$a2$
$a3$
$d1$
$s1$
$2a4$
$2a5$
$d2$
$d3$
$d4$
$s1$
$3s2$

$1 \leq i \leq 7$, $1 \leq j \leq 6$. And similarly for each other section according to its heading.

(iv) Each expansion has a block to itself bounded by horizontal lines. Extra horizontal lines have been inserted if the end of an expansion occurs at the foot of a column. Some horizontal lines are thicker than others; this is for ease of reference—see (v), below.

(v) The expansions are given within each section in dictionary order. For the section headed $X \otimes X$ the 28 expansions, therefore, occur in the following order: $(ij) = (11), (12), \cdots, (17), (22), \cdots, (27), \cdots, (77)$. And similarly for each section with a heading of the form $P \otimes P$, $P = \Gamma, X, L, W, \Lambda$, and Δ. A horizontal line occurs whenever j changes its value by a unit. A bold horizontal line occurs whenever i changes its value by a unit. For the section headed $[X \otimes X]$, the seven expansions occur in the following order: $(ii) = (11), (22), \cdots, (77)$. And similarly for each section with a heading of the form $[P \otimes P]$, $P = \Gamma, X, L, W, \Lambda$, and Δ. A horizontal line occurs whenever i changes by a unit. For the section headed $X \otimes L$, the 42 expansions appear in the following order: $(ij) = (11), (12), \cdots, (16), (21), (22), \cdots, (26), (31), \cdots, (76)$. And similarly for each section with a heading of the form $P \otimes Q$ ($P, Q = \Gamma, X$; Γ, L; Γ, W; Γ, Δ; X, L; X, W; and L, W). A horizontal line occurs whenever j changes its value by a unit. A bold horizontal line occurs whenever i changes its value by a unit.

(vi) The expansions that would have fallen under the heading $\Gamma \otimes \Lambda$ can be obtained from the section headed by $\Gamma \otimes L$ by reading p for b whenever b occurs in that section.

(vii) For each expansion required, one or more entries appear in the block appropriate to that expansion, the rule being that the product required has as its expansion the direct sum of all the entries appearing in that block.

(viii) *Examples:*

$$(X_4 \uparrow \mathbf{G}) \otimes (L_2 \uparrow \mathbf{G}) = (L_2 \uparrow \mathbf{G}) + (L_3 \uparrow \mathbf{G}),$$

$$(L_2 \uparrow \mathbf{G}) \otimes (L_6 \uparrow \mathbf{G}) = (\Gamma_8 \uparrow \mathbf{G}) + (X_6 \uparrow \mathbf{G}) + (X_7 \uparrow \mathbf{G}),$$

$$[(L_5 \uparrow \mathbf{G}) \otimes (L_5 \uparrow \mathbf{G})] = (\Gamma_2 \uparrow \mathbf{G}) + (\Gamma_4 \uparrow \mathbf{G}) + (X_5 \uparrow \mathbf{G}).$$

(ix) Antisymmetrized squares can be obtained by subtraction and are not separately tabulated. Thus, from Table IV and Eq. (4.4),

$$\{(L_5 \uparrow \mathbf{G}) \otimes (L_5 \uparrow \mathbf{G})\} = (X_1 \uparrow \mathbf{G}) + (X_4 \uparrow \mathbf{G}).$$

[1] C. J. Bradley, J. Math. Phys. **7**, 1145 (1966).
[2] G. W. Mackey, Am. J. Math. **73**, 576 (1951).
[3] G. W. Mackey, Am. J. Math. **75**, 387 (1953).
[4] J. L. Birman, Phys. Rev. **127**, 1093 (1962).
[5] J. L. Birman, Phys. Rev. **131**, 1489 (1963).
[6] G. F. Karavaev, Izv. Vysshikh Uchebn. Zavedenii, Fiz. **9**(3), 64 (1966).
[7] M. Hamermesh, *Group Theory and its Application to Physical Problems* (Pergamon Press, New York, 1962), Chaps. 5 and 6.
[8] H. A. Jahn and E. Teller, Proc. Roy. Soc. (London) **A161**, 220 (1937).
[9] H. A. Jahn, Proc. Roy. Soc. (London) **A164**, 117 (1938).
[10] M. Lax, *International Conference on the Physics of Semiconductors, Exeter* (Institute of Physics and the Physical Society, London, 1962), p. 395.
[11] G. Ya. Lyubarskii, *The Application of Group Theory in Physics* (Pergamon Press, New York, 1960), Chap. VII.
[12] L. P. Bouckaert, T. Smoluchowski, and E. P. Wigner, Phys. Rev. **50**, 58 (1936).
[13] R. H. Parmenter, Phys. Rev. **100**, 573 (1955).
[14] J. O. Dimmock and R. G. Wheeler, *The Mathematics of Physics and Chemistry*, H. Margenau and G. M. Murphy, Eds. (D. Van Nostrand Co., Princeton, N.J., 1964), Vol. II, Chap. 12.
[15] S. L. Altmann and A. P. Cracknell, Rev. Mod. Phys. **37**, 19 (1965).

Neutron Transport in a Nonuniform Slab with Generalized Boundary Conditions*

A. BELLENI-MORANTE†
Department of Physics, University of Illinois, Urbana, Illinois

(Received 24 July 1969)

We prove the existence and the uniqueness of the solution of the initial-value problem for neutron transport in a nonuniform slab with generalized boundary conditions, which include the vacuum and the perfect reflection boundary conditions as particular cases. Moreover, we show that the position-dependent transport operator has at least one real eigenvalue and we indicate the asymptotic behavior of the neutron density as $t \to +\infty$.

1. INTRODUCTION

In a recent paper,[1] Mika considered the initial-value problem for mono-energetic neutron transport in a nonuniform slab surrounded either by vacuum or by a perfect absorber. By using the theory of semi-groups of linear bounded transformations, he was able to prove the existence and the uniqueness of the solution. The problem of the point spectrum of the position-dependent transport operator was not considered in Mika's paper. Moreover, as pointed out elsewhere,[2] the vacuum boundary condition is not the only one that needs investigation. In many cases conditions may be of interest which express that each neutron suffers a perfect reflection at the boundary.[3]

In this paper we extend Mika's investigations to the case of a nonuniform slab with generalized boundary conditions, which include the vacuum and the perfect reflection boundary conditions as particular cases, i.e., we study a fairly general mono-energetic neutron-transport problem with plane symmetry. By using some results of perturbation theory for linear operators, we are able to avoid long and tedious calculations and at the same time display some spectral properties of the position-dependent transport operator. Finally, we prove the existence of at least one real eigenvalue and we indicate the asymptotic behavior of the neutron density as $t \to +\infty$.

2. PRELIMINARY REMARKS

The neutron-transport equation, the initial and generalized boundary conditions have the form

$$\frac{\partial N}{\partial t} = AN, \quad |x| < a, \quad |\mu| \leq 1, \quad t > 0, \tag{1}$$

$$N(x, \mu, 0) = N_0(x, \mu), \quad |x| < a, \quad |\mu| \leq 1, \tag{2}$$

$$\delta(\mu)N(-a, \mu, t) = \eta(\mu)N(a, \mu, t), \quad |\mu| \leq 1, \quad t > 0, \tag{3}$$

where $N = N(x, \mu, t)$ is the neutron distribution function within the nonuniform slab of thickness $2a$ and A is the transport operator[4]

$$A = B + v\gamma(x)J, \quad B = -v\mu\frac{\partial}{\partial x} - v\Sigma(x),$$
$$J = \tfrac{1}{2}\int_{-1}^{+1} \cdots d\mu'. \tag{4}$$

Moreover, v is the neutron speed, $\Sigma(x)$ is the total cross section, $\gamma(x) = c(x)\Sigma(x)$ where $c(x)$ is the mean number of secondaries per collision,[4] and $\delta = \delta(\mu)$ and $\eta = \eta(\mu)$ are piecewise-continuous functions such that

$$0 \leq \eta_0 \leq \eta(\mu) \leq \delta(\mu) \leq \eta_1 \leq 1, \quad \mu \in (0, 1],$$
$$0 \leq \eta_0 \leq \delta(\mu) \leq \eta(\mu) \leq \eta_1 \leq 1, \quad \mu \in [-1, 0). \tag{5}$$

We note that, in particular, if $\delta(\mu) = \eta(\mu) \equiv 1$, Eq. (3) becomes a perfect reflection boundary condition.[2] Furthermore, if $\eta(\mu) \equiv 0$ and $\delta(\mu) \equiv 1$ at any $\mu \in (0, 1]$ and $\eta(\mu) \equiv 1$ and $\delta(\mu) \equiv 0$ at any $\mu \in [-1, 0)$, Eq. (3) becomes the usual vacuum boundary condition.[1,5-7]

Finally, we assume that $\Sigma(x)$ and $\gamma(x)$ are continuous a.e. in $[-a, a]$ such that

$$0 < \Sigma_m = \{\inf[\Sigma(x)], |x| \leq a\}$$
$$\leq \Sigma(x) \leq \Sigma_M = \{\sup[\Sigma(x)], |x| \leq a\},$$
$$0 < \gamma_m = \{\inf[\gamma(x)], |x| \leq a\}$$
$$\leq \gamma(x) \leq \gamma_M = \{\sup[\gamma(x)], |x| \leq a\}. \tag{6}$$

The existence and the uniqueness of the solution of the system (1)–(3) are related to the property of A of being the generator of a semigroup of linear bounded transformations $\{S(t), t \geq 0\} \in C_0$.[8,9] In turn, such a property depends on the nature of the spectrum of the transport operator A over a certain Hilbert space X.

Let X be the space of all functions $f = f(x, \mu)$ defined and square integrable over the rectangle $R \equiv \{(x, \mu): |x| \leq a, |\mu| \leq 1\}$ with scalar product

and norm given by the relations

$$(f, g) = \iint_R f(x, \mu)\bar{g}(x, \mu)\, dx\, d\mu, \quad \|f\| = \{(f,f)\}^{\frac{1}{2}},$$

where $\bar{g}(x, \mu)$ denotes the complex conjugate.

The domain $D(A)$ of A is the linear manifold of all functions $f \in X$, such that $Af \in X$ and condition (3) is satisfied. Following Jorgens,[7] it can be proved that $D(A)$ is dense in X, i.e., $\overline{D(A)} \equiv X$.

We now recall that the property of being the generator of a (C_0) semigroup is stable under bounded perturbations.[8] Hence, it will be sufficient to prove that B is the generator of a semigroup $\{T(t), t \geq 0\} \in C_0$. In fact, from (4) we see that A may be regarded as a perturbation of B by means of the bounded operator $v\gamma(x)J$ whose norm is smaller than $v\gamma_M$.

3. SPECTRAL PROPERTIES OF B

In order to prove that B is the infinitesimal generator of a C_0 semigroup we will consider the following equation:

$$(\lambda I - B)f = g, \quad f \in D(A) = D(B), \quad g \in X, \quad (7)$$

where $\lambda = \beta + i\tau$. We prove the following theorem.

Theorem 1: The region $\{\lambda : \beta = \operatorname{Re}\lambda > -v\Sigma_m\}$ of the complex λ plane belongs to the resolvent set $\rho(B)$ of operator B and the following inequality is valid:

$$\|R(\lambda, B)\| = \|(\lambda I - B)^{-1}\| \leq (\beta + v\Sigma_m)^{-1}, \quad \beta = \operatorname{Re}\lambda > -v\Sigma_m. \quad (8)$$

Proof: If we put

$$\Sigma(x, y) = \Sigma(y, x) = \frac{1}{x - y}\int_y^x \Sigma(s)\, ds, \quad \Sigma_m \leq \Sigma(x, y) \leq \Sigma_M,$$

$$\Sigma(a, -a) = \Sigma_0 = \frac{1}{2a}\int_{-a}^{a} (s)\, ds, \quad \Sigma_m \leq \Sigma_0 \leq \Sigma_M, \quad (9)$$

we obtain, after some manipulations from (7),

$$f(x,\mu) = \frac{\delta(\mu)}{\delta(\mu) - \eta(\mu)\exp(-2a\Delta_0/\mu)} \times \int_{-a}^{x} \exp\left[-\Delta(x, y)\frac{x - y}{\mu}\right]\frac{g(y, \mu)}{v\mu}\, dy + \frac{\eta(\mu)}{\delta(\mu)\exp(2a\Delta_0/\mu) - \eta(\mu)} \times \int_{x}^{a} \exp\left[-\Delta(x, y)\frac{x - y}{\mu}\right]\frac{g(y, \mu)}{v\mu}\, dy, \quad (10)$$

where $\Delta(x, y) = \Sigma(x, y) + \lambda/v$, $\Delta_0 = \Sigma_0 + \lambda/v$.

By taking into account that

$$|\exp[-\Delta(x, y)(x - y)/\mu]| \leq \exp[-(\Sigma_m + \beta/v)(x - y)/\mu], \quad y < x, \quad \mu > 0, \quad \beta > -v\Sigma_m,$$

$$|\delta(\mu) - \eta(\mu)\exp(-2a\Delta_0/\mu)| \geq \delta(\mu)\{1 - \exp[-(\Sigma_0 + \beta/v)2a/\mu]\}, \quad \mu > 0, \quad \beta > -v\Sigma_m$$

and proceeding just as in Ref. 2, we have, from (10),

$$\|f(x, \mu)\| \leq 2[p_0(\beta + v\Sigma_m)]^{-1}\|g\|, \quad \beta > -v\Sigma_m, \quad g \in X,$$

where $p_0 = 1 - \exp[-2a(\Sigma_0 + \beta/v)]$. Hence, we have

$$\|R(\lambda, B)\| \leq 2[p_0(\beta + v\Sigma_m)]^{-1}, \quad \beta > -v\Sigma_m. \quad (11)$$

Thus, we have that the region $\{\lambda : \beta = \operatorname{Re}\lambda > -v\Sigma_m\}$ belongs to the resolvent set of operator B.

Inequality (11) may now be improved. In fact, we have

$$\operatorname{Re}((\lambda I - B)\psi, \psi) \geq (\beta + v\Sigma_m)\|\psi\|^2 + \tfrac{1}{2}\int_{-1}^{+1}\{\mu[|\psi(x, \mu)|^2]_{x=-a}^{x=a}\}\, d\mu \geq (\beta + v\Sigma_m)\|\psi\|^2$$

owing to conditions (3) and (5). Now let $(\lambda I - B)\psi = g$, $\psi = R(\lambda, B)g$. From the preceding inequality we obtain

$$\|\psi\|\,\|g\| \geq |(g, \psi)| \geq \operatorname{Re}(g, \psi) \geq (\beta + v\Sigma_m)\|\psi\|^2, \quad \|\psi\| \leq (\beta + v\Sigma_m)^{-1}\|g\|,$$

and finally

$$\|R(\lambda, B)\| \leq (\beta + v\Sigma_m)^{-1}.$$

Theorem 1 is thus completely proved.

4. THE INITIAL-VALUE PROBLEM

By the Hille–Yosida theorem,[8,9] B generates a semigroup $\{T(t), t \geq 0\} \in (C_0)$ provided that (a) B is a closed operator, (b) $D(B)$ is dense in X, and (c) $\|R(\lambda, B)\| \leq (\beta - \beta_0)^{-1}$, $\beta > \beta_0$, for some real β_0. Conditions (b) and (c) have already been proved. As far as (a) is concerned, we observe that if $\operatorname{Re}\lambda > -v\Sigma_m$, then $(\lambda I - B)^{-1}$ is bounded and its domain is the whole X. Thus, $(\lambda I - B)^{-1}$, $\lambda I - B$, and also B are closed operators.[9] We conclude that B is the generator of a semigroup $\{T(t), t \geq 0\} \in C_0$, such that $\|T(t)\| \leq \exp(-v\Sigma_m t)$ at any $t \geq 0$. Moreover, because of some results of perturbation theory for linear operators,[9] we may state the following theorem.

Theorem 2: The region $\{\lambda: \mathrm{Re}\,\lambda > v(\gamma_M - \Sigma_m)\}$ belongs to the resolvent set $\rho(A)$ of A, $\|R(\lambda, A)\| \leq [\beta - v(\gamma_M - \Sigma_m)]^{-1}$, and A is the generator of a semigroup $\{S(t), t \geq 0\} \in C_0$, such that $\|S(t)\| \leq \exp[v(\gamma_M - \Sigma_m)t]$ at any $t \geq 0$.

From the preceding theorem it follows that the unique solution of the initial-value problem (1)–(3) has the form[8,9]

$$N(x, \mu, t) = S(t)N_0, \quad N_0 \in D(A), \quad t \geq 0. \tag{12}$$

Our next task will be to transform the right-hand side of (12) in order to obtain a more explicit form of the neutron density.

5. THE INTEGRAL EQUATION

Let us indicate by Γ the following region of the λ plane:

$$\Gamma \equiv \{\lambda: -v\Sigma_m < \mathrm{Re}\,\lambda < v(\gamma_M - \Sigma_m),\ \lambda \notin P_\sigma(A)\},$$

where $P_\sigma(A)$ is the point spectrum of A. We will prove that Γ belongs to the resolvent set $\rho(A)$ of operator A. With this aim in mind, we observe that the equation

$$(\lambda I - A)\psi = g, \quad g \in X, \tag{13}$$

may be written in the form

$$(\lambda I - B)\psi = g + v\gamma(x)J\psi = g + \tfrac{1}{2}v\gamma(x)\varphi = g + \tfrac{1}{2}v\alpha(x)\theta, \tag{14}$$

where

$$\varphi = \varphi(x) = 2J\psi = \int_{-1}^{+1} \psi(x, \mu')\, d\mu', \quad \|\varphi\| \leq 2\,\|\psi\|, \tag{15a}$$

$$\theta(x) = \alpha(x)\varphi(x), \quad \alpha(x) = [\gamma(x)]^{\frac{1}{2}}. \tag{15b}$$

From (14) we now obtain

$$\psi = f + \tfrac{1}{2}vR(\lambda, B)[\gamma\varphi] = f + \tfrac{1}{2}vR(\lambda, B)[\alpha\theta], \quad \mathrm{Re}\,\lambda > -v\Sigma_m, \tag{16}$$

where f is given by (10), provided $\mathrm{Re}\,\lambda > -v\Sigma_m$ and hence $\lambda \in \rho(B)$. Conversely, we get (14) from (16) by applying operator $\lambda I - B$ to both sides of (16). Finally, by applying the operator $2\alpha(x)J$ to both sides of (16),we have

$$\theta = G + K_\lambda\theta, \tag{17}$$

where

$$G = 2\alpha(x)Jf = 2\alpha(x)[JR(\lambda, B)]g, \tag{18}$$

$$K_\lambda\theta = v[JR(\lambda, B)][\alpha(x)\alpha(y)\theta(y)] = \int_{-a}^{+a} K(\lambda, x, y)\theta(y)\, dy, \tag{19}$$

where the kernel $K(\lambda, x, y)$ is given by the relation

$$K(\lambda, x, y) = K(\lambda, y, x) = \tfrac{1}{2}\alpha(x)\alpha(y) \times \int_1^{+\infty} \frac{\delta(1/t)\exp[-\Delta(x, y)\,|x - y|\,t] + \eta(1/t)\exp[-2a\Delta_0 t + \Delta(x, y)\,|x - y|\,t]}{\delta(1/t) - \eta(1/t)\exp(-2a\Delta_0 t)}\,\frac{dt}{t}, \tag{20}$$

provided that $\eta(-\mu) = \delta(\mu)$.

If, in particular, $\eta(\mu) \equiv 0$ at any $\mu \in (0, 1]$, then $K(\lambda, x, y)$ becomes the kernel used by Mika,[1] whereas, if $\eta(\mu) = \delta(\mu) \equiv 1$, then $K(\lambda, x, y)$ is similar to the kernel used in Ref. 2.

With a procedure similar to that used in Ref. 2, let us now introduce the Hilbert space L_2 of all functions $\varphi = \varphi(x)$ defined and square integrable over $[-a, a]$, with scalar product and norm given by the relations

$$(\varphi, \varphi')_0 = \int_{-a}^{+a} \varphi(x)\bar{\varphi}'(x)\, dx, \quad \|\varphi\|_0 = [(\varphi, \varphi)_0]^{\frac{1}{2}}, \quad \|\varphi\| = (2)^{\frac{1}{2}}\,\|\varphi\|_0, \quad \varphi \in L_2. \tag{21}$$

From (15) and (18) we have

$$\|G\|_0 = (2)^{-\frac{1}{2}}\,\|G\| \leq (2\gamma_M)^{\frac{1}{2}}\,\|R(\lambda, B)\|\,\|g\| \leq [(2\gamma_M)^{\frac{1}{2}}/(\beta + v\Sigma_m)]\,\|g\|,$$
$$\|\theta\|_0 = (2)^{-\frac{1}{2}}\,\|\theta\| \leq (\tfrac{1}{2}\gamma_M)^{\frac{1}{2}}\,\|\varphi\| \leq (2\gamma_M)^{\frac{1}{2}}\,\|\psi\|, \tag{22}$$

where we used (21). Hence, $\theta \in L_2$ and $G \in L_2$ provided that $\psi \in X$, $g \in X$, and $\beta > -v\Sigma_m$. This justifies the introduction of the Hilbert space L_2. Moreover, it may be easily proved that the kernel $K(\lambda, x, y)$ is square integrable over the square

$$\{(x, y): |x| \leq a, |y| \leq a\},$$

provided that $\beta > -v\Sigma_m$. Hence, if $\beta > -v\Sigma_m$, K_λ is a compact operator whose norm satisfies the following relation:

$$\|K_\lambda\|_0 \leq \left\{\int_{-a}^{+a} dx \int_{-a}^{+a} dy\,|K(\lambda, x, y)|^2\right\}^{\frac{1}{2}}; \tag{23}$$

K_λ is also self-adjoint if $\lambda = \beta$ is a real number.

We summarize the relationships between the integro-differential equation (13) and the integral equation (17) by means of the following two theorems.

Theorem 3: If $\mathrm{Re}\,\lambda > -v\Sigma_m$, then any solution $\psi \in D(A)$ of Eq. (13) is such that $\theta = 2\alpha J\psi$ is a solution of (17) belonging to L_2. Conversely, if a solution $\theta \in L_2$ of Eq. (17) is known, then the corresponding

solution $\psi \in D(A)$ of (13) may be found by means of (16).

Theorem 4: If $\operatorname{Re}\lambda > -v\Sigma_m$, if $\|g\| = 0$ (hence $\|G\|_0 = 0$), and if λ is an eigenvalue of A, then 1 is an eigenvalue of K_λ, and conversely. Moreover, the corresponding eigenfunctions are such that $\psi = \frac{1}{2}vR(\lambda, B)\alpha\theta$.

Let us now return to the region Γ defined at the beginning of this section. From (16) we have

$$\begin{aligned}\|\psi\| &\le \|R(\lambda, B)\|\,[\|g\| + v(2\gamma_M)^{\frac{1}{2}}\,\|\theta\|_0] \\ &\le (\beta + v\Sigma_m)^{-1} \\ &\quad \times [\|g\| + v(2\gamma_M)^{\frac{1}{2}}\,\|(I - K_\lambda)^{-1}\|_0\,\|G\|_0] \\ &\le C(\lambda)(\beta + v\Sigma_m)^{-1}\,\|g\|, \quad \lambda \in \Gamma, \quad g \in X,\end{aligned}$$

where, because of (22),

$$C(\lambda) = 1 + 2v\gamma_M(\beta + v\Sigma_m)^{-1}\,\|(I - K_\lambda)^{-1}\|_0. \quad (24)$$

We took into account that $(I - K_\lambda)^{-1}$ is a bounded operator provided $\lambda \in \Gamma$ because of Theorem 4. We conclude that

$$\|R(\lambda, B)\| \le C(\lambda)(\beta + v\Sigma_m)^{-1}, \quad \beta > -v\Sigma_m. \quad (25)$$

Hence, we have the following theorem.

Theorem 5: The region $\Gamma \equiv \{\lambda: -v\Sigma_m < \operatorname{Re}\lambda < v(\gamma_M - \Sigma_m),\ \lambda \notin P_\sigma(A)\}$ belongs to the resolvent set $\rho(A)$ of operator A.

6. EXISTENCE OF AT LEAST ONE EIGENVALUE OF A

The remaining question is the nature of $P_{\sigma,r}(A) \equiv P_\sigma(A) \cap \{\lambda: -v\Sigma_m < \operatorname{Re}\lambda < v(\gamma_M - \Sigma_m)\}$. Because of Theorem 4, if $\lambda \in P_{\sigma,r}(A)$, then $1 \in P(K_\lambda)$ and conversely. Hence, we may look for those λ's, such that 1 is an eigenvalue of the compact operator K_λ.

Let us first limit ourselves to the case in which λ is real, $\lambda = \beta$. Then $K_\lambda = K_\beta$ has all the properties listed in the Appendix. Consequently, if $v_1(\beta)$ is the largest positive eigenvalue of K_β and $\theta_1(x)$ is the corresponding eigenfunction

$$v_1(\beta)\theta_1 = K_\beta\theta_1, \quad \|\theta_1\|_0 = 1, \quad (26)$$

we have that $v_1(\beta)$ is a simple eigenvalue, $\theta_1(x) > 0$, at any $x \in [-a, a]$. Then the following relation is valid:

$$\begin{aligned}v_1(\beta) &= \{\sup\,[(K_\beta\theta, \theta)_0],\ \|\theta\|_0 = 1\} \\ &= (K_\beta\theta_1, \theta_1)_0 = \|K_\beta\|_0. \quad (27)\end{aligned}$$

By using (27), we can prove the following theorem.

Theorem 6: (a) $v_1(\beta)$ is a continuous function of β; (b) $v_1(\beta)$ is a decreasing function of β; (c) $\lim v_1(\beta) = 0$ as $\beta \to \infty$; (d) $\lim v_1(\beta) = +\infty$, as $\beta \to -v\Sigma_m + 0$, provided an interval of positive measure $[x_0, x_1] \subseteq [-a, a]$ exists, such that $\Sigma(x) = \Sigma_m$ at any $x \in [x_0, x_1]$.

Proof: Part (a) is a straightforward consequence of the continuity of $\|K_\beta\|_0$ with respect to β, whereas Part (b) follows from the fact that

$$v_1(\beta) = (K_\beta\theta_1, \theta_1)_0 < (K_{\beta'}\theta_1, \theta_1)_0 \le v_1(\beta'),$$

provided $\beta > \beta'$. As far as Part (c) is concerned, from (8), (19), (21), and (27) we have

$$\begin{aligned}\|K_\beta\theta\|_0 &= (2)^{-\frac{1}{2}}\,\|K_\beta\theta\| \le v\gamma_M(2)^{-\frac{1}{2}}\,\|R(\lambda, B)\|\,\|\theta\| \\ &\le v\gamma_M(\beta + v\Sigma_m)^{-1}\,\|\theta\|_0.\end{aligned}$$

Finally, let us consider Part (d). From (27) with $\theta = (2a)^{-\frac{1}{2}}$, it follows that

$$\begin{aligned}v_1(\beta) &= (2a)^{-1}\int_{-a}^{+a} dx \int_{-a}^{+a} K(\beta, x, y)\,dy \\ &\ge (\gamma_m/4a)\int_{x_0}^{x_1} dx \int_{x_0}^{x_1} E[(\Sigma_m + \beta/v)\,|x - y|]\,dy \\ &\ge (\gamma_m/4a)(x_1 - x_0)^2 E[(\Sigma_m + \beta/v)(x_1 - x_0)], \quad (28)\end{aligned}$$

where E is the exponential integral function. Hence, (d) is proved because $\lim E[(\Sigma_m + \beta/v)(x_1 - x_0)] = +\infty$, as $\beta \to -v\Sigma_m + 0$.

From Theorem 6 we infer the existence of one and only one value of β, say β_1, such that $v_1(\beta_1) = 1$. Thus we may state the following theorem.

Theorem 7: β_1 is the greatest real eigenvalue of A, it is simple and satisfies the relation

$$1 = \|K_{\beta_1}\|_0.$$

Proof: That β_1 is a simple eigenvalue of A follows from the fact that $1 = v_1(\beta_1)$ is a simple eigenvalue of K_{β_1} and from Theorem 4.

Finally, by means of procedure similar to those used in Ref. 2, the following properties can be proved without difficulty.

Theorem 8: β_1 is not an accumulation point of $P_\sigma(A)$; any other eigenvalue of A has its real part smaller than β_1; if $A\psi_1 = \beta_1\psi_1$, $\|\psi_1\| \ne 0$, and if $\psi_1^*(x, \mu) = \psi_1(x, -\mu)$, then $(\psi_1, \psi_1^*) = 1$.

We note that Theorem 8 also indicates a relationship between β_1 and the other real or complex eigenvalues of A, if any.

7. ASYMPTOTIC FORM OF THE NEUTRON DENSITY

We will now transform the right-hand side of (12) by using the results of the preceding section. With this aim in mind, we observe that if we put $\chi_1 = \exp(\beta_1 t)$, then χ_1 is an eigenvalue of the semigroup $S(t)$ with eigenfunction ψ_1.[8] Moreover, ψ_1 is the only eigenfunction corresponding to χ_1 since no eigenvalues $\lambda \neq \beta_1$ of A exist, such that $\exp(\lambda t) = \chi_1$ (see Theorem 8).

If we now indicate by P the projection operator into the linear manifold generated by ψ_1, then the operator $(I - P)S(t)$ does not possess the eigenvalue $\exp(\beta_1 t)$ and, as a consequence, its spectral radius is smaller than $\exp(b_1 t)$, where $b_1 < \beta_1$, and is greater than the real part of any other eigenvalue of A. This last condition may be fulfilled because of Theorem 8. It follows that, given any $\epsilon > 0$, a positive integer n_ϵ can be found, such that

$$\|S(t)(I - P)\| \leq c(\epsilon) \exp(b_1 t), \quad t \geq n_\epsilon, \quad (29)$$

where, by using a method similar to that of Ref. 7, we have

$$c(\epsilon) = \max_{0 \leq \tau \leq 1} \{\exp[v(\gamma_M - \Sigma_m)\tau - b\tau + \epsilon]\}.$$

From (12), we now have

$$N = S(t)N_0 = S(t)PN_0 + N_1$$
$$= S(t)[(N_0, \psi_1^*)\psi_1] + N_1,$$

where we took into account that $PN_0 = (N_0, \psi_1^*)\psi_1$ because $(\psi_1, \psi_1^*) = 1$, (see Theorem 8), and where

$$N_1 = S(t)(I - P)N_0. \quad (30)$$

Hence,

$$N = (N_0, \psi_1^*)\psi_1 \exp(\beta_1 t) + N_1, \quad (31)$$

because $S(t)\psi_1 = \exp(\beta_1 t)\psi_1$. As far as N_1 is concerned, we have

$$\|N_1\| \leq c(\epsilon) \exp(b_1 t) \|N_0\|,$$

provided t is large enough, because of (29).

Relation (31) is an asymptotic expression of the neutron density N, solution of system (1)–(3). We note that the second term on the right-hand side of (31) becomes negligible in norm with respect to the first as $t \to +\infty$.

APPENDIX

In Sec. 6 we need the following lemma.

Lemma: Let H be an integral operator in L_2 whose kernel $h(x, y)$ has the following properties: (a) $h(x, y)$ is square integrable over the square $|x| \leq a$, $|y| \leq a$; (b) $h(x, y)$ is real and such that $h(x, y) = h(y, x)$; (c) $h(x, y) > 0$. Then H is a compact self-adjoint operator, its eigenfunctions may be taken real, and $\|H\|_0 = \nu_1$, where ν_1 is the greatest positive eigenvalue. Moreover, ν_1 is simple and the corresponding eigenfunction $\varphi_1(x)$ is such that $\varphi_1(x) > 0$ at any $x \in [-a, a]$.

Proof: The compactness and the self-adjointness of H follow from the general theory of integral operators with kernels which are real, symmetric, and square integrable.[10] Moreover, let $\varphi = \varphi_r + i\varphi_i$ be an eigenfunction of H corresponding to the eigenvalue ν, where φ_r and φ_i are real functions and ν is a real number because H is self-adjoint. We have

$$H(\varphi_r + i\varphi_i) = H\varphi_r + iH\varphi_i = \nu\varphi_r + i\nu\varphi_i$$

and also

$$H\varphi_r = \nu\varphi_r, \quad H\varphi_i = \nu\varphi_i,$$

because H is a real operator. From the preceding relations it readily follows that all the eigenfunctions may be taken real.

We now observe that H must have at least one positive eigenvalue; otherwise it would be a negative operator.[10] This is obviously false since the kernel of H is positive and $(H\varphi, \varphi)_0 > 0$ if, for instance, $\varphi \equiv 1$. Let us then indicate by ν_1 the largest positive eigenvalue of H and by $\varphi_1(x)$ a corresponding real eigenfunction

$$H\varphi_1 = \nu_1\varphi_1, \quad \|\varphi_1\|_0 = 1.$$

From the theory of compact selfadjoint operators,[10] it is known that ν_1 is given by the following relation:

$$\nu_1 = \{\sup[(H\varphi, \varphi)_0], \|\varphi\|_0 = 1\} = (H\varphi_1, \varphi_1)_0. \quad (A1)$$

Assume now that $\varphi_1(x)$ is positive on a set $S^+ \subset [-a, a]$ and negative on a set $S^- \subset [-a, a]$, where S^+ and S^- are both of positive measure. We have

$$\nu_1 = (H\varphi_1, \varphi_1)_0 < (H|\varphi_1|, |\varphi_1|)_0.$$

Since $|\varphi_1|$ has also norm 1, this contradicts (A1) and hence we may state that $\varphi_1(x) \geq 0$. Moreover, since $\varphi_1 \not\equiv 0$, from the relation $\nu_1\varphi_1 = H\varphi_1$, it follows that $\varphi_1 > 0$ a.e. on $[-a, a]$.

We now prove that ν_1 is simple. Suppose in fact that there is another eigenfunction $\hat{\varphi}_1(x)$ corresponding to ν_1; we can always assume that $\hat{\varphi}_1$ is real and orthogonal to φ_1, $(\varphi_1, \hat{\varphi}_1)_0 = 0$. Therefore, $\hat{\varphi}_1$ is negative on a set $\hat{S}^- \subset [-a, a]$ and positive on a set $\hat{S}^+ \subset [-a, a]$, where $\hat{S}^-$ and $\hat{S}^+$ are both of positive measure. Assuming $\|\hat{\varphi}_1\|_0 = 1$, we have

$$\nu_1 = (H\hat{\varphi}_1, \hat{\varphi}_1)_0 < (H|\hat{\varphi}_1|, |\hat{\varphi}_1|)_0,$$

which contradicts (A1). Hence ν_1 is a simple eigenvalue.

Finally, let ν^- be the smallest (negative) eigenvalue of H with eigenfunction $\varphi^-(x)$ which may be assumed real and normalized:

$$\nu^-\varphi^- = H\varphi^-, \quad \|\varphi^-\|_0 = 1. \tag{A2}$$

We have

$$|\nu^-| = |(H\varphi^-, \varphi^-)_0| < (H|\varphi^-|, |\varphi^-|)_0 \leq (H\varphi_1, \varphi_1)_0 = \nu_1,$$

where we used (A1) and we took into account that φ^- cannot be always positive because of (A2). Since $\|H\|_0 = \max\{\nu_1; |\nu^-|\}$, we conclude that $\|H\|_0 = \nu_1$. This completes the proof of the lemma.

* Work supported in part by a NATO fellowship.

† Permanent address: Istituto Matematico, Universita' di Firenze, Firenze, Italy.

[1] J. Mika, J. Math. Phys. 7, 833 (1966).
[2] A. Belleni-Morante, J. Math. Anal. Appl. 30, 666 (1970).
[3] E. E. Lewis and F. T. Adler, Nucl. Sci. Eng. 31, 117 (1968).
[4] B. Davison, *Neutron Transport Theory* (Oxford University Press, London, 1957).
[5] J. Lehner and M. G. Wing, Commun. Pure Appl. Math. 8, 217 (1955).
[6] J. Lehner and M. G. Wing, Duke Math. J. 23, 125 (1955).
[7] K. Jorgens, Commun. Pure Appl. Math. 11, 219 (1959).
[8] E. Hille and R. S. Phillips, *Colloquia of the American Mathematical Society, Vol. 31: Functional Analysis and Semigroups* (American Mathematical Society, Providence, R.I., 1957).
[9] T. Kato, *Perturbation Theory for Linear Operators* (Springer, New York, 1966).
[10] F. Riesz and B. Sz.-Nagy, *Leçons d'analyse fonctionelle* (Gauthier-Villars, Paris, 1955).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 5 MAY 1970

Localized States on a Hyperplane*

S. C. McDONALD†
Institute for Atomic Research and Department of Physics
Iowa State University, Ames, Iowa 50010

(Received 8 August 1969)

The localized states of Newton and Wigner are reconstructed as a superposition of the canonical states of Foldy, and these states are then generalized to an arbitrary hyperplane via a procedure similar to that used in a previous work for the hyperplane generalization of helicity states. The corresponding hyperplane position operator, with mutually commuting components, is constructed and is seen to be equivalent to that local position operator given by Fleming.

1. INTRODUCTION

In 1949, Newton and Wigner[1] provided an elegant treatment of a position representation for elementary systems; such systems defined therein are describable by states which are invariant under a given irreducible representation of the Poincaré group. Starting from a set of postulates, Newton and Wigner[1] were able to construct basis states for such a position representation for the $m^2 > 0$ irreducible representations, which corresponded to a unique position operator with mutually commuting components, previously discussed by Pryce.[2] Unlike such momentum representation basis states as the canonical states due to Foldy[3] and the Jacob and Wick helicity states,[4] however, the localized states and position operator constructed by Newton and Wigner do not transform covariantly under the action of the Lorentz group. This, of course, is seen as a consequence of the frame-dependent localization criterion postulated by Newton and Wigner,[1] and is sometimes noted as an undesirable feature of their development.

On the other hand, several other nonequivalent position operators studied by Pryce[2] and Møller[5] are defined in a covariant way, but as a consequence do not have mutually commuting components. Hence, such operators do not immediately provide for a corresponding position representation of relativistic quantum mechanics, although, as Fleming[6] has pointed out, may nonetheless be interpreted so as to constitute "position" observables. The purpose of this paper is to develop a manifestly covariant realization of the Newton–Wigner[1] localized states, by means of the hyperplane formalism recently introduced by Fleming.[6,7] As such, this paper may be regarded as a sequel to an earlier work,[8] herein referred to as I, in which the helicity states of Jacob and Wick[4] are generalized to an arbitrary spacelike hyperplane.

Toward this end, Sec. 2 consists of a review of the hyperplane formalism as given by Fleming[6,7] and in I,[8] with special emphasis being given the hyperplane helicity states developed in I.[8] The hyperplane

Finally, let ν^- be the smallest (negative) eigenvalue of H with eigenfunction $\varphi^-(x)$ which may be assumed real and normalized:

$$\nu^-\varphi^- = H\varphi^-, \quad \|\varphi^-\|_0 = 1. \tag{A2}$$

We have

$$|\nu^-| = |(H\varphi^-, \varphi^-)_0| < (H|\varphi^-|, |\varphi^-|)_0 \leq (H\varphi_1, \varphi_1)_0 = \nu_1,$$

where we used (A1) and we took into account that φ^- cannot be always positive because of (A2). Since $\|H\|_0 = \max\{\nu_1; |\nu^-|\}$, we conclude that $\|H\|_0 = \nu_1$. This completes the proof of the lemma.

* Work supported in part by a NATO fellowship.

† Permanent address: Istituto Matematico, Universita' di Firenze, Firenze, Italy.

[1] J. Mika, J. Math. Phys. 7, 833 (1966).
[2] A. Belleni-Morante, J. Math. Anal. Appl. 30, 666 (1970).
[3] E. E. Lewis and F. T. Adler, Nucl. Sci. Eng. 31, 117 (1968).
[4] B. Davison, *Neutron Transport Theory* (Oxford University Press, London, 1957).
[5] J. Lehner and M. G. Wing, Commun. Pure Appl. Math. 8, 217 (1955).
[6] J. Lehner and M. G. Wing, Duke Math. J. 23, 125 (1955).
[7] K. Jorgens, Commun. Pure Appl. Math. 11, 219 (1959).
[8] E. Hille and R. S. Phillips, *Colloquia of the American Mathematical Society, Vol. 31: Functional Analysis and Semigroups* (American Mathematical Society, Providence, R.I., 1957).
[9] T. Kato, *Perturbation Theory for Linear Operators* (Springer, New York, 1966).
[10] F. Riesz and B. Sz.-Nagy, *Leçons d'analyse fonctionelle* (Gauthier-Villars, Paris, 1955).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 5 MAY 1970

Localized States on a Hyperplane*

S. C. McDONALD†
Institute for Atomic Research and Department of Physics
Iowa State University, Ames, Iowa 50010

(Received 8 August 1969)

The localized states of Newton and Wigner are reconstructed as a superposition of the canonical states of Foldy, and these states are then generalized to an arbitrary hyperplane via a procedure similar to that used in a previous work for the hyperplane generalization of helicity states. The corresponding hyperplane position operator, with mutually commuting components, is constructed and is seen to be equivalent to that local position operator given by Fleming.

1. INTRODUCTION

In 1949, Newton and Wigner[1] provided an elegant treatment of a position representation for elementary systems; such systems defined therein are describable by states which are invariant under a given irreducible representation of the Poincaré group. Starting from a set of postulates, Newton and Wigner[1] were able to construct basis states for such a position representation for the $m^2 > 0$ irreducible representations, which corresponded to a unique position operator with mutually commuting components, previously discussed by Pryce.[2] Unlike such momentum representation basis states as the canonical states due to Foldy[3] and the Jacob and Wick helicity states,[4] however, the localized states and position operator constructed by Newton and Wigner do not transform covariantly under the action of the Lorentz group. This, of course, is seen as a consequence of the frame-dependent localization criterion postulated by Newton and Wigner,[1] and is sometimes noted as an undesirable feature of their development.

On the other hand, several other nonequivalent position operators studied by Pryce[2] and Møller[5] are defined in a covariant way, but as a consequence do not have mutually commuting components. Hence, such operators do not immediately provide for a corresponding position representation of relativistic quantum mechanics, although, as Fleming[6] has pointed out, may nonetheless be interpreted so as to constitute "position" observables. The purpose of this paper is to develop a manifestly covariant realization of the Newton–Wigner[1] localized states, by means of the hyperplane formalism recently introduced by Fleming.[6,7] As such, this paper may be regarded as a sequel to an earlier work,[8] herein referred to as I, in which the helicity states of Jacob and Wick[4] are generalized to an arbitrary spacelike hyperplane.

Toward this end, Sec. 2 consists of a review of the hyperplane formalism as given by Fleming[6,7] and in I,[8] with special emphasis being given the hyperplane helicity states developed in I.[8] The hyperplane

generalization of the essentially equivalent canonical representation of Foldy[3] is also noted. In Sec. 3, the localized states first constructed by Newton and Wigner[1] are redeveloped in terms of helicity and canonical states, and the canonical position operator corresponding to that developed by Newton and Wigner[1] is derived. Although this development follows in the spirit of the Newton–Wigner[1] treatment, it is believed that the use of the canonical formalism provides for a less tedious derivation and a more convenient representation of the localized states and position operator. In Sec. 4 the canonical forms of the localized states and corresponding position operator are generalized to an arbitrary hyperplane, and the relationship between localized states on two different hyperplanes are given.

2. THE HYPERPLANE FORMALISM

Fleming[6] was first led to develop the hyperplane formalism by the desire to find a way by which the various nonequivalent position operators given by Newton and Wigner,[1] Pryce,[2] and Møller[5] could be written in a manifestly covariant way. In his development, Fleming[6] has associated one of the inertial observers with the standard basis tetrad $\{e^\mu: \mu = 0, 1, 2, 3\}$ in the Minkowski space M_x, and has (partially) specified all other inertial observers by associating with each a spacelike hyperplane, defined with respect to the standard tetrad by a unit normal $\eta = (\eta^0, \boldsymbol{\eta})$, where $\eta_\mu\eta^\mu = 1$, and by its orthogonal displacement τ from the origin.

Fleming's[6] manifestly covariant representations of the various position operators are given by 4-component operators $X^\mu(\eta, \tau)$, the components being labeled with respect to the standard observer's tetrad. However, owing to the constraint that the operator refers to a "point" on the hyperplane defined by η and τ,

$$X^\mu(\eta, \tau)\eta_\mu = \tau, \tag{2.1}$$

only three of the components may be independent. The parameter τ is interpreted by Fleming[6] to be the "time" as used by a hyperplane observer. The manifest covariance of the $X^\mu(\eta, \tau)$ is seen by the fact that, whereas the operator is equivalent to the position operator used by an arbitrary hyperplane observer, it is written in terms of parameters used by the standard observer.

As a matter of notation, Hammer, McDonald, and Pursey[9] have called Fleming's[6] standard observer the "superobserver," all other observers correspondingly being called hyperplane observers. The particular family of hyperplanes with normal $\tilde{\eta} = (1, \mathbf{0})$ are called by Fleming[6] the "instantaneous" hyperplanes, the particular member with hyperplane "time" $\tilde{\tau} = t = 0$ being spanned by the spacelike part of the superobserver's basis tetrad.

In a later article, Fleming[7] has developed his hyperplane formalism to include a discussion of the Poincaré group. The hyperplane concept is first extended to the momentum space M_p such that, for each family of hyperplanes η in M_x, there corresponds a unique hyperplane with unit normal η which passes through the superobserver's origin in M_p. As usual, the superobserver's basis tetrad in M_p is taken to be "parallel" to his tetrad in M_x by defining M_x as the parameter space of all possible space–time translations acting on the quantum-mechanical Hilbert space. Fleming[7] has then defined operators which generate the infinitesimal transformations of the Poincaré group as seen by an arbitrary hyperplane observer.

In particular, the hyperplane operators which generate infinitesimal translations parallel to and orthogonal to the hyperplane with normal η are defined[7] by

$$\begin{aligned} M^\mu(\eta) &= P^\mu - \eta^\mu(\eta P), \\ H(\eta) &= (\eta P), \end{aligned} \tag{2.2}$$

where $(\eta P) \equiv \eta_\mu P^\mu$, with the P^μ being the infinitesimal generators for 4-translations as used by the superobserver. Also, the operators which generate homogeneous transformations in the hyperplane η are defined[7] by

$$J^\mu(\eta) = \tfrac{1}{2}\epsilon^{\mu\nu\rho\sigma}M_{\nu\rho}\eta_\sigma, \tag{2.3}$$

where the $M^{\mu\nu}$ are the superobserver's generators for all homogeneous transformations and the completely antisymmetric quantities $\epsilon^{\mu\nu\rho\sigma}$ are defined by

$$\epsilon^{0123} = \epsilon^{123} = +1.$$

Furthermore, the operators which generate infinitesimal "boosts" orthogonal to the hyperplane are given by[7]

$$N^\mu(\eta) = M^{\mu\nu}\eta_\nu. \tag{2.4}$$

The component labels of these operators refer to the superobserver's basis tetrad $\{e^\mu\}$. Although Fleming's[7] operators do not transform covariantly in the ordinary sense of transformations used by the superobserver, they are manifestly covariant in the sense that they are defined, with the same form, for all hyperplanes.

In I,[8] a special basis tetrad $\{\xi^{\mu'}: \mu' = 0, 1, 2, 3\}$ for an observer on an arbitrary hyperplane is defined with respect to the superobserver tetrad $\{e^\mu\}$ by

$$\begin{aligned} (\xi^{0'})^\mu &= l^\mu{}_\nu(\eta)(e^0)^\nu = l^\mu{}_0(\eta), \\ (\xi^{i'})^\mu &= l^\mu{}_\nu(\eta)(e^i)^\nu = l^\mu{}_i(\eta), \end{aligned} \tag{2.5}$$

the quantities $l^{\mu}{}_{\nu}(\eta)$ being identified with the components of the Lorentz transformation without rotation which takes the instantaneous hyperplane normal $\tilde{\eta} = (1, \mathbf{0})$ into η. All other tetrads for a given hyperplane may then be related to that tetrad defined by Eq. (2.5) through a homogeneous transformation in the hyperplane, or "hyperplane rotation." Here, the notation of I[8] is retained, in which the "primed" indices refer to components to the hyperplane tetrad of Eq. (2.5), and a vector as seen by the hyperplane observer is written as

$$x^{\mu'} = [x^{0'}, x^{1'}, x^{2'}, x^{3'}] = [x^{0'}, \mathbf{x}]. \tag{2.6}$$

The hyperplane observer's canonical Poincaré generators $P^{\mu'}, J^{i'}$, and $N^{i'}$, which are written with respect to the hyperplane tetrad of Eq. (2.5) are seen[8] to be related to the hyperplane operators of Eqs. (2.2)–(2.4) by

$$\begin{aligned} P^{0'} &= H(\eta), \\ P^{i'} &= l_{\mu}{}^{i}(\eta) K^{\mu}(\eta), \\ J^{i'} &= l_{\mu}{}^{i}(\eta) J^{\mu}(\eta), \\ N^{i'} &= l_{\mu}{}^{i}(\eta) N^{\mu}(\eta). \end{aligned} \tag{2.7}$$

With the specification Eq. (2.5) of the hyperplane tetrad and the hyperplane Poincaré generators of Eqs. (2.2), (2.4) and (2.7), the translation, rotation, and special Lorentz transformation operators which act on the quantum-mechanical state space as used by the hyperplane observer are given respectively by[8]

$$\begin{aligned} T[\eta; t(a)] &= e^{-ia_{\mu'}P^{\mu'}} = e^{i(\alpha_{\mu}K^{\mu} - \tau H)}, \\ R[\eta; r(\eta; \alpha\beta\gamma)] &= e^{i\alpha J^{3'}} e^{i\beta J^{2'}} e^{i\gamma J^{3'}} \\ &= e^{i\alpha_{\mu}J^{\mu}} e^{i\beta_{\mu}J^{\mu}} e^{i\gamma_{\mu}J^{\mu}}, \\ L[\eta; \omega(\mathbf{u})] &= e^{iu^{i'}N^{i'}} = e^{iv_{\mu}N^{\mu}}. \end{aligned} \tag{2.8}$$

The hyperplane observer's helicity states as discussed[8] are written symbolically as $|(ms)\epsilon\mathbf{p}\lambda\rangle_H$ and satisfy

$$\begin{aligned} P^{i'} |(ms)\epsilon\mathbf{p}\ \rangle_H &= p^{i'} |(ms)\epsilon\mathbf{p}\lambda\rangle_H, \\ P^{0'} |(ms)\epsilon\mathbf{p}\ \rangle_H &= \epsilon[p^{i'}p^{i'} + m^2]^{\frac{1}{2}} |(ms)\epsilon\mathbf{p}\ \rangle_H, \\ \hat{\epsilon}(\eta) |(ms)\epsilon\mathbf{p}\lambda\rangle_H &= \epsilon |(ms)\epsilon\mathbf{p}\lambda\rangle_H, \\ \pi(\eta) |(ms)\epsilon\mathbf{p}\lambda\rangle_H &= \lambda |(ms)\epsilon\mathbf{p}\lambda\rangle_H, \end{aligned} \tag{2.9}$$

where $\epsilon(\eta)$ and $\pi(\eta)$ are the hyperplane observer's sign of the energy and helicity operators, respectively. The usual Lorentz-invariant normalization is taken for these states, so that

$$\begin{aligned} &\langle \eta; (ms')\epsilon'\mathbf{p}'\lambda' \mid \eta; (ms)\epsilon\mathbf{p}\lambda\rangle \\ &\qquad = 2\omega(p)\delta_{s's}\delta_{\epsilon'\epsilon}\delta_{\lambda'\lambda}\delta^3(\mathbf{p}' - \mathbf{p}), \end{aligned} \tag{2.10}$$

with $\omega(p) = (p^{i'}p^{i'} + m^2)^{\frac{1}{2}}$. The helicity states of Eq. (2.9) are furthermore seen in I[8] to be related to the helicity states used by the superobserver by the transformation which takes the superobserver tetrad $\{e^{\mu}\}$ into the tetrad $\{\xi^{\mu'}\}$ used by the hyperplane observer.

Like the Jacob and Wick[4] helicity states, the canonical states of Foldy[3] also form a basis for the $m^2 > 0$ irreducible representation carrier spaces of the Poincaré group, but differ from the helicity states in that the spin label λ is interpreted as the component of angular momentum in the e^3 direction, instead of the $\mathbf{p}$ direction. The superobserver's general canonical state $|(ms)\epsilon\mathbf{p}\lambda\rangle_c$ is related to the rest state by

$$|(ms)\epsilon\mathbf{p}\lambda\rangle_c = L_0^{\epsilon}(\mathbf{p}) |(ms)\epsilon\tilde{\mathbf{p}}\lambda\rangle, \tag{2.11}$$

where $\tilde{\mathbf{p}} = (\epsilon m, \mathbf{0})$ and

$$L_0^{\epsilon}(\mathbf{p}) = H^{\epsilon}(p) R(\mathbf{p}). \tag{2.12}$$

Here, the transformation $l_0^{\epsilon}(\mathbf{p})$ is the Lorentz transformation which takes the standard momentum $\tilde{\mathbf{p}} = (\epsilon m, \mathbf{0})$ into $p = (\epsilon[m^2 + p^2]^{\frac{1}{2}}, \mathbf{p})$.

Analogous to Eq. (2.9) for the helicity states, the canonical states are seen to satisfy

$$\begin{aligned} P_{\mu}P^{\mu} |(ms)\epsilon\mathbf{p}\lambda\rangle &= m^2 |(ms)\epsilon\mathbf{p}\lambda\rangle, \\ W_{\mu}W^{\mu} |(ms)\epsilon\mathbf{p}\lambda\rangle &= -m^2 s(s+1) |(ms)\epsilon\mathbf{p}\lambda\rangle, \\ P^{\mu} |(ms)\epsilon\mathbf{p}\lambda\rangle &= p^{\mu} |(ms)\epsilon\mathbf{p}\lambda\rangle, \\ \hat{\epsilon} |(ms)\epsilon\mathbf{p}\lambda\rangle &= \epsilon |(ms)\epsilon\mathbf{p}\lambda\rangle, \\ J^3 |(ms)\epsilon\mathbf{p}\lambda\rangle &= \lambda |(ms)\epsilon\mathbf{p}\lambda\rangle, \end{aligned} \tag{2.13}$$

where $p^0 \equiv \epsilon[m^2 + \mathbf{p} \cdot \mathbf{p}]^{\frac{1}{2}}$. Following a procedure similar to that used for the helicity states,[8] the canonical states are seen to transform under a general Lorentz transformation $l(\omega)$ with operator $L(\omega)$ as

$$L(\omega) |p\lambda\rangle_c = \sum_{\mu} D^{s}_{\mu\lambda}[\tilde{l}^{\epsilon}(l(\omega))] |(\mathbf{lp})\mu\rangle, \tag{2.14}$$

where

$$\tilde{l}^{\epsilon}[l(\omega)] = l_0^{\epsilon}(lp)^{-1} l(\omega) l_0^{\epsilon}(p) \tag{2.15}$$

is an element of the little group $R(3)$.

One could now proceed in a manner parallel to the preceding treatment of helicity states in I[8] to generalize Foldy's[3] canonical representation to a hyperplane. However, Macfarlane[10] has shown that the canonical representation is related to the helicity representation by

$$\begin{aligned} |\mathbf{p}\alpha\rangle_c &= R_0^{s}(\mathbf{p}) |\mathbf{p}\alpha\rangle_H \\ &= \sum_{\lambda} D^{s}_{\lambda\alpha}[r_0(\mathbf{p})] |\mathbf{p}\lambda\rangle_H, \end{aligned} \tag{2.16}$$

where $R_0^{s}[r_0(\mathbf{p})]$ is the rotation operator corresponding to $r_0(\mathbf{p})$ which acts only on the spin variables. Thus, the hyperplane generalization of canonical states may be made directly from the results of the helicity representation generalization.

In terms of the hyperplane helicity states of I[8], the hyperplane canonical states are defined by

$$|\eta; \mathbf{p}\alpha\rangle_c = R_0^s[\eta; \mathbf{p}]\,|\eta; \mathbf{p}\alpha\rangle_H = \sum_\lambda D^s_{\lambda\alpha}[r_0(\eta; \mathbf{p})]\,|\eta; \mathbf{p}\lambda\rangle_H, \quad (2.17)$$

with $r_0(\eta; \mathbf{p})$ being the hyperplane rotation which takes $p = [p^{1'}, p^{2'}, p^{3'}]$ into the $\xi^{3'}$ direction. From Eq. (2.1) and the unitarity of the rotation matrices, the normalization is seen to be the same as for helicity states:

$$\langle \eta; (ms')\epsilon'\mathbf{p}'\alpha' \,|\, \eta; (ms)\epsilon\mathbf{p}\alpha\rangle_c = 2\omega(p)\delta_{s's}\delta_{\epsilon'\epsilon}\delta_{\lambda'\lambda}\delta^3(\mathbf{p}'^{i'} - \mathbf{p}^{i'}), \quad (2.18)$$

with $\omega(p) = [m^2 + p^{i'}p^{i'}]^{\frac{1}{2}}$.

Corresponding to Eq. (53) of I,[8] the hyperplane canonical states may be written in terms of the superobserver momenta $k^\mu = l^\mu{}_i(\eta)p^{i'}$ by

$$|\tilde{\eta}; k(\eta)\alpha(\eta)\rangle_c \equiv |\eta; \mathbf{p}(\eta)\lambda(\eta)\rangle_c = \sum_\lambda D^s_{\lambda\alpha}[r_0(\eta; \mathbf{p})]\,|\tilde{\eta}; k(\eta)\lambda(\eta)\rangle_H. \quad (2.19)$$

Parallel to the treatment given in I,[8] the hyperplane canonical states are seen to be related to the superobserver's canonical states by

$$|\tilde{\eta}; k(\eta)\alpha(\eta)\rangle_c = L_0^\epsilon(\eta; p')L(\eta)L_0^\epsilon(p)^{-1}\,|\mathbf{p}\lambda\rangle, \quad (2.20)$$

where $k^\mu = l^\mu{}_i(\eta)p^{i'} = p^\mu - \eta^\mu(\eta p)$ and the L_0^ϵ are the corresponding hyperplane and superobserver operators of Eq. (2.12). Also, $L(\eta)$ is that operator which corresponds to the Lorentz transformation $l(\eta)$ which takes the superobserver tetrad $\{e^\mu\}$ into the hyperplane tetrad $\{\xi^{\mu'}\}$. It then follows by the same argument as given in I[8] that

$$|\tilde{\eta}; k(\eta)\alpha(\eta)\rangle_c = L(\eta)\,|(l^{-1}p)\alpha\rangle_c = \sum_\mu D^s_{\mu\alpha}[\tilde{t}(l(\eta))]\,|\mathbf{p}\mu c\rangle, \quad (2.21)$$

where $k^\mu = p^\mu - \eta^\mu(\eta p)$ and $\tilde{t}$ is the little group element, given by Eq. (2.15), which corresponds to the Lorentz transformation $l(\eta)$. It is then easily verified that the hyperplane canonical states of Eq. (2.20) satisfy

$$K^\mu\,|\tilde{\eta}; k\alpha\rangle_c = k^\mu\,|\tilde{\eta}; k\alpha\rangle_c,$$
$$J^{3'}\,|\tilde{\eta}; k\alpha\rangle_c = \alpha\,|\tilde{\eta}; k\alpha\rangle_c. \quad (2.22)$$

3. THE NEWTON–WIGNER LOCALIZED STATES

In constructing their localized states, Newton and Wigner[1] have postulated that in order to form a basis for a position representation of elementary systems, such state must transform within the $m^2 > 0$ irreducible representations of the Poincaré group. This implies that such states may be partially labeled by the mass, spin, and sign of the energy. However, since the treatment of localized states by Newton and Wigner[1] is done in terms of the Bargmann–Wigner[11] representation, such a simple labeling scheme is not available, owing to the fact that the solutions of the wave equations

$$\gamma^k_\alpha P_k \Psi = \mu\Psi \quad (3.1)$$

are not in fully reduced form. As a consequence, the resulting localized states and corresponding position operator take relatively complicated forms.

In this development, the Jacob and Wick[4] helicity formalism is used, since it provides a more convenient labeling scheme for states which transform within a given irreducible representation of the Poincaré group. Indeed, it is seen that in terms of the canonical representation of Foldy,[3] which is related to the helicity representation by a momentum-dependent rotation of the type given in Eq. (2.16), the resulting localized states and position operator are of especially simple form. Of course, since the Newton–Wigner[1] states are unique, and there exists a transformation, given by Pursey,[12] which relates the Bargmann–Wigner[11] representation with the helicity representation, one may, in principle, derive the results found herein from the Newton–Wigner[1] results. Fortunately, though, the convenience of the helicity labeling scheme allows one to retrace the Newton–Wigner[1] development with much less difficulty.

In parallel with the Newton–Wigner[1] treatment, the set $S_0(ms\epsilon)$ of states localized at the origin of the tetrad $\{e^\mu : \mu = 0, 1, 2, 3\}$ is first written as a superposition of $(2s + 1)$ helicity states

$$|(ms)\epsilon\mathbf{x} = \mathbf{0}\alpha; x^0 = 0\rangle = \sum_\lambda \int d\tilde{p}\,\Phi^{(ms)\epsilon}_{\lambda\alpha}(\mathbf{p})\,|(ms)\epsilon\mathbf{p}\lambda\rangle_H. \quad (3.2)$$

Here, $d\tilde{p}$ is the Lorentz-invariant volume element $[2\omega(p)]^{-1}\,d^3p$ and the functions $\Phi^{(ms)\epsilon}_{\lambda\alpha}(\mathbf{p})$ remain to be determined, as do the allowable values of the parameter α. The sets of states $S_x(ms\epsilon)$ over each point $(x^0, \mathbf{x})$ may now be defined in terms of the set $S_0(s)$ of Eq. (3.2) via the translation $t(x^0, \mathbf{x}): (0, \mathbf{0}) \to (x^0, \mathbf{x})$ with operator $T(t)$:

$$|(ms)\epsilon\mathbf{x}\alpha; x^0\rangle = T(t(x^0, \mathbf{x}))\,|(ms)\epsilon\mathbf{0}\alpha; 0\rangle. \quad (3.3)$$

Since helicity states transform under translations as

$$T[t(x, \mathbf{x})]\,|(ms)\epsilon\mathbf{p}\lambda\rangle_H = e^{-ix_\mu p^\mu}\,|(ms)\epsilon\mathbf{p}\lambda\rangle_H, \quad (3.4)$$

the states defined by Eq. (3.4) are given by

$$|(ms)\epsilon\mathbf{x}\alpha; x^0\rangle = \sum_\lambda \int d\tilde{p}\,\Phi^{(ms)\epsilon}_{\lambda\alpha}(\mathbf{p})e^{-ix_\mu p^\mu}\,|(ms)\epsilon p\lambda\rangle_H, \quad (3.5)$$

where $p^0 = \epsilon(m^2 + \mathbf{p}\cdot\mathbf{p})^{\frac{1}{2}}$. Where there is no ambiguity, the irreducible representation labels m, s, and ϵ will be suppressed.

In order uniquely to define localized states, Newton and Wigner[1] have postulated that the set $S_0(m, s, \epsilon)$ of states defined by Eq. (3.2) satisfies the following conditions:

(a) The set S_0 is a linear set.

(b) The set $S_0(s)$ is invariant under rotations about the origin and under spatial and temporal inversions.

(c) A state defined over an arbitrary point ($x^0 = 0$, $\mathbf{x}$) by Eq. (3.5) is orthogonal to each state of S_0.

(d) The states of S_0 satisfy certain continuity conditions, so that the infinitesimal generators of the Poincaré group are applicable.

Postulate (c) may be called the "locality" condition, in that it provides a group-theoretical criterion for the spatial localization of relativistic states. One sees that postulate (c) and Eqs. (3.2) and (3.4), together with the orthogonality Eq. (2.10) for helicity states, lead to

$$\langle \mathbf{0}\alpha'; 0 \mid \mathbf{x}\alpha; 0\rangle = \delta_{\alpha'\alpha}\delta^3(\mathbf{x}) = \sum_\lambda \int d\tilde{p}\Phi^*_{\lambda\alpha'}(\mathbf{p})\Phi_{\lambda\alpha}(\mathbf{p})e^{ip\cdot x}, \quad (3.6)$$

so that the $\Phi_{\lambda\alpha}(\mathbf{p})$ must satisfy

$$\sum_\lambda \Phi^*_{\lambda\alpha'}(\mathbf{p})\Phi_{\lambda\alpha}(\mathbf{p}) = (2\pi)^{-3}2\omega(\mathbf{p})\delta_{\alpha'\alpha}. \quad (3.7)$$

Moreover, the rotational invariance of S_0 imposed by postulate (b) implies that, for an arbitrary rotation $r(\alpha\beta\gamma)$ with operator $R(r)$, the states of Eq. (3.2) must satisfy

$$R(r)\,|\mathbf{0}, \alpha; 0\rangle = \sum_{\alpha'} C_{\alpha'\alpha}(r)\,|\mathbf{0}\alpha'; 0\rangle_H, \quad (3.8)$$

where the coefficients $C_{\alpha'\alpha}$ may depend only on the rotation r and the spin s, with the sum being over the entire range of the parameter α'. It is assumed at this point that the range of α is, in fact, that of the helicities, from $-s$ to s by integral steps, and that at most the coefficients $C(r)$ are related to the rotation matrix $D^s(r)$ by a fixed unitary transformation U:

$$C^s(r) = U^{-1}D^s(r)U. \quad (3.9)$$

These assumptions are justified when it is shown that, with this choice, the states defined by Eq. (3.2) form a basis for S_0. In particular, there is no loss in generality in taking $U_{\alpha\beta} = \delta_{\alpha\beta}$, so that

$$R(r)\,|\mathbf{0}\alpha; 0\rangle = \sum_{\alpha'} D^s_{\alpha'\alpha}(r)\,|\mathbf{0}\alpha'; 0\rangle. \quad (3.10)$$

By using the Lorentz transformation property of the helicity states together with Eqs. (3.2) and (3.10), it follows that

$$R(r)\,|\mathbf{0}\alpha; 0\rangle = \sum_{\lambda'\lambda}\int d\tilde{p}D^s_{\lambda'\lambda}(r)\Phi_{\lambda\alpha}(r^{-1}\mathbf{p})\,|\mathbf{p}\lambda'\rangle_H = \sum_{\alpha'\lambda}\int d\tilde{p}\Phi_{\lambda\alpha'}(\mathbf{p})D^s_{\alpha'\alpha}(r)\,|\mathbf{p}\lambda\rangle_H, \quad (3.11)$$

so that the $\Phi_{\lambda\alpha}$ must satisfy

$$\sum_{\mu'\mu} D^s_{\alpha\mu'}(r^{-1})\Phi_{\mu'\mu}(\mathbf{p})D^s_{\mu\beta}(r) = \Phi_{\alpha\beta}(r^{-1}\mathbf{p}). \quad (3.12)$$

Thus, by noting the similarity of Eq. (3.12) with the transformation property of the rotation matrix $D[r_0(\mathbf{p})]$,

$$D[r_0(r^{-1}\mathbf{p})] = D(r^{-1})D[r_0(\mathbf{p})]D(r), \quad (3.13)$$

where $r_0(\mathbf{p})$ is the rotation which takes the $\mathbf{p}$ direction into the e^3 direction, one is led to choose

$$\Phi^{(ms)\epsilon}_{\lambda\alpha}(\mathbf{p}) = (2\pi)^{-\frac{3}{2}}[2\omega(\mathbf{p})]^{\frac{1}{2}}D^s_{\lambda\alpha}[r_0(\mathbf{p})]. \quad (3.14)$$

It is now a simple matter to verify that, with this choice of $\Phi_{\lambda\alpha}$, both the localization condition of Eq. (3.7) and the rotational symmetry condition of Eq. (3.12) are satisfied.

The localized states defined by Eq. (3.2) may now be written as

$$|\mathbf{0}\alpha; 0\rangle = (2\pi)^{-\frac{3}{2}}\sum_\lambda \int d^3p[2\omega(p)]^{-\frac{1}{2}}D^s_{\lambda\alpha}[r_0(\mathbf{p})]\,|\mathbf{p}\lambda\rangle_H. \quad (3.15)$$

However, by recalling the relationship between the helicity representation and the canonical representation given by Eq. (2.16), the localized states of Eq. (3.2) may be written in an even simpler form as

$$|\mathbf{0}\alpha; 0\rangle = (2\pi)^{-\frac{3}{2}}\int d^3p[2\omega(p)]^{-\frac{1}{2}}\,|\mathbf{p}\alpha\rangle_c. \quad (3.16)$$

Hence, as with the canonical representation, the parameter α is to be interpreted as the component of spin in the e^3 direction.

Foldy[3] has shown that, under space inversion P and Wigner time reversal,[13] the canonical states transform according to

$$P\,|\epsilon\mathbf{p}\lambda\rangle_c = e^{i\theta_p}(\epsilon)^{2s}\,|\epsilon - \mathbf{p}\lambda\rangle_c,$$
$$T\,|\epsilon\mathbf{p}\lambda\rangle_c = e^{i\theta_T(\epsilon)}(-1)^{s-\lambda}\,|\epsilon - \mathbf{p} - \lambda\rangle_c, \quad (3.17)$$

where θ_p and $\theta_\tau(\epsilon)$ are fixed phases. Thus the states given by Eq. (3.16) are seen to transform as

$$P\,|\epsilon\mathbf{0}\alpha; 0\rangle = e^{i\theta_p}(\epsilon)^{2s}\,|\epsilon\mathbf{0}\alpha; 0\rangle,$$
$$T\,|\epsilon\mathbf{0}\alpha; 0\rangle = e^{i\theta_T(\epsilon)}(-1)^{s-\alpha}\,|\epsilon\mathbf{0} - \alpha; 0\rangle, \quad (3.18)$$

so that, with the choice Eq. (3.14) for $\Phi_{\lambda\alpha}(\mathbf{p})$, postulate (b) is completely satisfied, as are postulates (a) and (d).

In order to show that the $(2s+1)$ states of Eq. (3.16) form a basis for $S_0(ms\epsilon)$, it is sufficient to show that there is no additional state $|\Psi_a^{(ms)\epsilon}\rangle \neq 0$, which is in $S_0(ms\epsilon)$ such that

$$\langle \mathbf{0}\alpha; 0 \,|\, \Psi_a^{(ms)\epsilon}\rangle = 0, \tag{3.19}$$

for each $-s \leq \alpha \leq s$. Since the canonical states form a basis for the given irreducible representation of the Poincaré group, any such state $|\Psi_a\rangle$ may be written as

$$|\Psi_a\rangle = \sum_\alpha \int d\tilde{p}\Psi_{\alpha a}(\mathbf{p})\,|\mathbf{p}\alpha\rangle_c, \tag{3.20}$$

where

$$\Psi_{\alpha a}(\mathbf{p}) = {}_c\langle \mathbf{p}\alpha \,|\, \Psi_a\rangle. \tag{3.21}$$

Then, from Eqs. (2.10) and (3.16), one sees that

$$\langle \mathbf{0}\alpha; 0 \,|\, \Psi_a\rangle = (2\pi)^{-\frac{3}{2}}\int d^3p\Psi_{\alpha a}(\mathbf{p})[2\omega(p)]^{-\frac{1}{2}} = 0, \tag{3.22}$$

for all $-s \leq \alpha \leq s$. However, since $|\Psi_a\rangle$ is assumed to be in S_0, then it must be orthogonal to each element of S_0 translated by an arbitrary distance $\mathbf{x}$. In particular,

$$\langle \mathbf{x}\alpha; 0 \,|\, \Psi_a\rangle = (2\pi)^{-\frac{3}{2}}\int d^3p\Psi_{\alpha a}(\mathbf{p})[2\omega(p)]^{-\frac{1}{2}}e^{-ip\cdot x} = 0, \tag{3.23}$$

for $\mathbf{x} \neq 0$. Then, by combining Eq. (3.22) with Eq. (3.23),

$$\int d^3p\Psi_{\alpha a}(\mathbf{p})[2\omega(p)]^{-\frac{1}{2}}e^{-ip\cdot x} = 0, \tag{3.24}$$

for all $\mathbf{x}$. But $[\omega(p)]^{-\frac{1}{2}}$ is well defined, so that any such state $\Psi_{\alpha a}(\mathbf{p})$ must identically vanish. Hence, the set of $(2s+1)$ states given by Eq. (3.16) forms a basis for $S_0(ms\epsilon)$. By the same argument, the set of states defined by Eq. (3.5) over each $x = (x^0, \mathbf{x})$ forms a basis for the set $S_x(ms\epsilon)$ over x. It must be stressed, however, that from postulate (c), the sets S_x are localized only with respect to other sets $S_{x'}$ over the same instantaneous hyperplane. Owing to the extra factor $e^{-ip^0x^0}$ which would appear in $\langle \mathbf{x}\alpha; x^0 \,|\, \mathbf{0}\alpha; 0\rangle$, it is clear that sets $S_x(x^0)$ and S_0 $(x^0 = 0)$ over two different instantaneous hyperplanes are not orthogonal. Thus, the "time" parameter x^0 is seen as a label specifying the particular hyperplane over which the sets S_x are defined, instead of as an eigenvaluelike parameter as are the x^i.

The localized states over an arbitrary point $\mathbf{x}$ on the $x^0 = 0$ instantaneous hyperplane are, from Eqs. (3.5) and (3.16),

$$|\mathbf{x}\alpha; 0\rangle = (2\pi)^{-\frac{3}{2}}\int d^3p[2\omega(p)]^{-\frac{1}{2}}e^{ip\cdot x}\,|p\alpha\rangle_c. \tag{3.25}$$

Equation (3.25) may be inverted in the usual way to yield

$$|p\alpha\rangle_c = (2\pi)^{-\frac{3}{2}}[2\omega(p)]^{\frac{1}{2}}\int d^3x e^{-ip\cdot x}\,|\mathbf{x}\alpha; 0\rangle \tag{3.26}$$

and

$${}_c\langle \mathbf{p}\alpha' \,|\, \mathbf{x}\alpha; 0\rangle = \langle \mathbf{x}\alpha; 0 \,|\, \mathbf{p}\alpha'\rangle_c^* = \delta_{\alpha'\alpha}(2\pi)^{-\frac{3}{2}}[2\omega(p)]^{\frac{1}{2}}e^{ip\cdot x}. \tag{3.27}$$

The Hermitian position operator, of which the localized states of Eq. (3.16) are eigenstates, may now be constructed in a manner similar to that used by Newton and Wigner.[1] On the $x^0 = 0$ instantaneous hyperplane, the operators X^i $(x^0 = 0)$ are defined by

$$X^i(0)\,|\mathbf{x}\alpha; 0\rangle = x^i\,|\mathbf{x}\alpha; 0\rangle. \tag{3.28}$$

From Eq. (3.25), this may be written as

$$X^i(0)\,|\mathbf{x}\alpha 0\rangle = (2\pi)^{-\frac{3}{2}}\int d^3p[2\omega(p)]^{-\frac{1}{2}}x^ie^{ip\cdot x}\,|\mathbf{p}\alpha\rangle_c. \tag{3.29}$$

One now needs only to note that

$$[2\omega(p)]^{\frac{1}{2}}x^ie^{ip\cdot x} = [2\omega(p)]^{\frac{1}{2}}\left(\frac{1}{i}\frac{\partial}{\partial p^i}e^{ip\cdot x}\right) = i\left(\frac{\partial}{\partial p_i} + [2\omega^2(p)]^{-1}p^i\right)[2\omega(p)]^{\frac{1}{2}}e^{ip\cdot x} \tag{3.30}$$

to see that the momentum-space representation of the $X^i(0)$ is given by

$$X^i(0) = i\left(\frac{\partial}{\partial p_i} + [2\omega^2(p)]^{-1}p^i\right). \tag{3.31}$$

This is exactly the same result as that found by Newton and Wigner,[1] but which in their case is applicable only to the spin-zero irreducible representations. Owing to the choice of the canonical representation in this development, however, the position operator of Eq. (3.31) may be applied within the irreducible representations for any spin and for either sign of the energy, as has been noted by Berg.[14]

The commutation rules satisfied by the $X^i(0)$ are readily found from Eq. (3.31) to be

$$\begin{aligned}[X^i, X^j] &= 0,\\ [P^i, X^j] &= -i\delta^{ij},\\ [P^0, X^i] &= -iP^i/P^0,\end{aligned} \tag{3.32}$$

as previously shown by Newton and Wigner.[1] Here P^0 is given by $\pm(\mathbf{P}\cdot\mathbf{P} + m^2)^{\frac{1}{2}}$. Moreover, since from Eqs. (3.10) and (3.28) the X^i are seen to transform under rotations like the P^i,

$$R^{-1}(r)X^iR(r) = r^i_jX^j, \tag{3.33}$$

it then follows that

$$[J^i, X^j] = i\epsilon^{ijk}X^k. \tag{3.34}$$

The Newton–Wigner[1] position operator is given at time x^0 by

$$X^i(x^0) = T^{-1}[t(x^0, \mathbf{0})]X^i(0)T[t(x^0, \mathbf{0})], \tag{3.35}$$

where $T[t(x^0, \mathbf{0})]$ is the purely timelike translation operator $e^{-ix^0P^0}$. It then follows from the identity

$$e^A\mathcal{O}e^{-A} = \mathcal{O} + [A, \mathcal{O}] + (1/2!)[A, [A, \mathcal{O}]] + \{\text{higher order commutators}\}, \tag{3.36}$$

with $A = e^{-ix^0P^0}$ and $\mathcal{O} = X^i(0)$, together with Eq. (3.32), that the $X^i(t)$ take the well-known form[2]

$$X^i(x^0) = X^i(0) + x^0P^i/P^0. \tag{3.37}$$

4. THE HYPERPLANE LOCALIZED STATES

In order to generalize the results of Newton and Wigner[1] to an arbitrary hyperplane, it is first necessary to generalize the defining postulates (b) and (c) to an arbitrary spacelike hyperplane with tetrad specified by Eq. (2.5). Of course, conditions (a) and (d) are hyperplane independent and so remain unaltered. As was the case with the development of the hyperplane helicity states,[8] the hyperplane observer will define all indexed quantities with respect to the special tetrad defined by Eq. (2.5), and all states and functions are labeled by η in order to specify the tetrad to which the parameters refer. The notation for primed indices and brackets instead of parentheses to distinguish vector components will also be retained.

In parallel with Eq. (3.2) the hyperplane localized states will first be defined over the origin as a superposition of the hyperplane helicity states of I [8]:

$$|\eta, 0; (ms)\epsilon\mathbf{0}\alpha\rangle \equiv \sum_\lambda \int d\tilde{p}\Phi^{(ms)\epsilon}_{\lambda\alpha}[\eta; \mathbf{p}]\,|\eta; (ms)\epsilon\mathbf{p}\lambda\rangle_H, \tag{4.1}$$

where $d\tilde{p} \equiv [2\omega(p)]^{-1}\,dp^{1'}\,dp^{2'}\,dp^{3'}$ and $\mathbf{p} = [p^{1'}, p^{2'}, p^{3'}]$ is the 3-momentum as seen by the hyperplane observer. The states over an arbitrary point $[x^{0'} = \tau, \mathbf{x}]$ are defined by

$$|\eta, \tau; (ms)\epsilon\mathbf{x}\alpha\rangle = T[\eta; t[\tau, \mathbf{x}]]\,|\eta, 0; (ms)\epsilon\mathbf{0}\alpha\rangle, \tag{4.2}$$

where

$$t: [0, \mathbf{0}] \to [\tau, \mathbf{x}], \tag{4.3}$$

the parameter τ being the hyperplane observer's "time" parameter. The translation operator $T[\eta, t]$ is given by Eq. (2.8) in terms of the hyperplane parameters and Poincaré generators. In parallel with Eq. (3.5), the general state is then given by

$$|\eta, \tau; (ms)\epsilon\mathbf{x}\alpha\rangle = \sum_\lambda \int d\tilde{p}\Phi^{(ms)\epsilon}_{\lambda\alpha}[\eta; \mathbf{p}]e^{-ix_{\mu'}p^{\mu'}}\,|\eta; (ms)\epsilon\mathbf{p}\lambda\rangle_H, \tag{4.4}$$

where $p^{0'} = \epsilon\omega(p) = \epsilon(m^2 + p^{i'}p^{i'})^{\frac{1}{2}}$. As was the case in the development for the instantaneous hyperplane, the labels m, s, and $\epsilon = \epsilon(\eta)$ will be suppressed when there is no ambiguity.

The postulates of Newton and Wigner[1] may now be generalized to an arbitrary hyperplane by the following:

(a′) The set $S_0(\eta; ms\epsilon)$ is a linear set.

(b′) The hyperplane set $S_0(\eta; ms\epsilon)$ localized at the origin is invariant under hyperplane rotations and hyperplane spatial and temporal inversions.

(c′) A state defined over an arbitrary point $[0, \mathbf{x}]$ by Eq. (4.4) is orthogonal to each state of $S_0(\eta; ms\epsilon)$.

(d′) The states of $S_0(\eta; ms\epsilon)$ satisfy certain continuity conditions, so that the infinitesimal generators of the Poincaré group are applicable.

Conditions (a′) and (b′) are identical with conditions (a) and (d). The argument leading to the choice of satisfactory hyperplane functions $\Phi_{\lambda\alpha}[\eta; \mathbf{p}]$ follows in exact parallel with that presented in the preceding section, the only difference being the replacement of the instantaneous hyperplane parameters ϵ, λ, and p^μ by the corresponding hyperplane parameters $\epsilon(\eta)$, $\lambda(\eta)$, and $p^{\mu'}$. The hyperplane result corresponding to Eq. (3.14) is, in terms of the hyperplane parameters,

$$\Phi^{(ms)\epsilon}_{\lambda(\eta)\alpha(\eta)}[\eta; \mathbf{p}] = (2\pi)^{-\frac{3}{2}}[2\omega(p)]^{\frac{1}{2}}D^s_{\lambda(\eta)\alpha(\eta)}[\eta; r_0(\eta, \mathbf{p})], \tag{4.5}$$

where $r_0(\eta, \mathbf{p})$ is the hyperplane rotation which takes $\mathbf{p} = [p^{1'}, p^{2'}, p^{3'}]$ into the $\xi^{3'}$ direction and

$$\omega(p) = [m^2 + p^{i'}p^{i'}]^{\frac{1}{2}}.$$

In terms of the hyperplane helicity states, the localized states of Eq. (4.1) take the form

$$|\eta, 0; \mathbf{0}\alpha\rangle = (2\pi)^{-\frac{3}{2}}\sum_\lambda \int d^3p[2\omega(p)]^{-\frac{1}{2}}D^s_{\lambda\alpha}[\eta; r_0(\eta, \mathbf{p})]\,|\eta; \mathbf{p}\lambda\rangle_H \tag{4.6}$$

or, in terms of the hyperplane observer's canonical states,

$$|\eta, 0; \mathbf{0}\alpha\rangle = (2\pi)^{-\frac{3}{2}}\int d^3p[2\omega(p)]^{-\frac{1}{2}}\,|\eta; \mathbf{p}\alpha\rangle_c. \tag{4.7}$$

The hyperplane states over arbitrary $x^{i'}$ are correspondingly given by

$$|\eta, 0; \mathbf{x}\alpha\rangle = (2\pi)^{-\frac{3}{2}} \int d^3p [2\omega(p)]^{-\frac{1}{2}} e^{ip^{i'}x^{i'}} |\eta; \mathbf{p}\alpha\rangle_c, \tag{4.8}$$

with the hyperplane observer's position operator given as

$$X^{i'} = X^{i'}(0) + \tau P^{i'}/P^{0'}, \tag{4.9}$$

with

$$X^{i'}(0) = i\left(\frac{\partial}{\partial p_{i'}} + [2\omega^2(p)]^{-1} p^{i'}\right). \tag{4.10}$$

In order to now write the hyperplane states of Eq. (4.6) in terms of the superobserver's momenta and space–time parameters, it is necessary to recall the following relations from Sec. 2:

$$|\tilde{\eta}; (ms)\epsilon(\eta)k^{\mu}(\eta)\lambda(\eta)\rangle_c \equiv |\eta; (ms)\epsilon(\eta)p^{i'}(\eta)\lambda(\eta)\rangle_c, \tag{2.19'}$$

and the invariant volume element

$$d\tilde{p} = \frac{dp^{1'}\, dp^{2'}\, dp^{3'}}{2\omega(\mathbf{p})} = \frac{d^4k\delta(\eta k)}{2(m^2 - k_{\mu}k^{\mu})^{\frac{1}{2}}} \equiv d\tilde{k}, \tag{4.11}$$

where

$$k^{\mu} = (\xi_{i'})^{\mu} p^{i'} = l^{\mu}{}_{i}(\eta) p^{i'} \tag{4.12}$$

is the hyperplane 3-momentum as described by the superobserver. With the definition

$$|\tilde{\eta}, 0; (ms)\epsilon(\eta)0\alpha\rangle \equiv |\eta, 0; (ms)\epsilon(\eta)\mathbf{0}\alpha\rangle, \tag{4.13}$$

Eq. (4.7) may thus be written in terms of the superobserver momenta parameters as

$$|\tilde{\eta}, 0; 0\alpha\rangle = (2\pi)^{-\frac{3}{2}} \int d^4k\delta(\eta k)[2\omega(k)]^{-\frac{1}{2}} |\tilde{\eta}; k\alpha\rangle_c, \tag{4.14}$$

where $\omega(k) \equiv [m^2 - k_{\mu}k^{\mu}]^{\frac{1}{2}}$.

The general states over the point $x = [x^{0'}, x^{i'}] = (x^0, x^i)$ are then found by applying to Eq. (4.7) the hyperplane translation $t(\eta; x^{\mu'})$, with operator

$$T[\eta; t(x^{\mu'})] = e^{-i(y_{\mu}K^{\mu} + \tau H)}, \tag{2.8'}$$

where

$$y^{\mu} = l^{\mu}{}_{i}(\eta)x^{i'}, \quad \tau = x^{0'},$$
$$x^{\mu} = y^{\mu} + \eta^{\mu}\tau. \tag{4.15}$$

The hyperplane generalization of Eq. (3.5) is then seen to be

$$|\tilde{\eta}, \tau; y(\eta)\alpha(\eta)\rangle = (2\pi)^{-\frac{3}{2}} \int d^4k\delta(\eta k)[2\omega(k)]^{-\frac{1}{2}} \times e^{-i[y_{\mu}k^{\mu} + \epsilon\omega(k)\tau]} |\tilde{\eta}; k(\eta)\alpha(\eta)\rangle_c. \tag{4.16}$$

The hyperplane localization condition corresponding to that given by Eq. (3.6) is, in terms of the hyperplane states,

$$\langle \tilde{\eta}, 0; y'\alpha' \mid \tilde{\eta}, 0; y\alpha\rangle = \delta_{\alpha'\alpha}\delta^3[\eta; y' - y], \tag{4.17}$$

where, with Eq. (4.15), the hyperplane Dirac distribution $\delta^3[\eta; y' - y]$ is given by

$$\delta^3[\eta; y' - y] \equiv (2\pi)^{-3} \int dp^{1'}\, dp^{2'}\, dp^{3'} e^{ip^{i'}(x^{i'} - x^{i'})}$$
$$= (2\pi)^{-3} \int d^4k\delta(\eta k) e^{-ik_{\mu}(y^{\mu} - y'^{\mu})}. \tag{4.18}$$

As is the case with localized states on an instantaneous hyperplane, the localization criterion of Eq. (4.17) applies only to states defined over a given hyperplane, with normal η and hyperplane "time" τ. Neither do the hyperplane localized states transform covariantly under hyperplane Lorentz transformations $l(\eta; \omega)$, as is seen from the noncovariance of the localization condition Eq. (4.17). However, the sets $S_x(\eta)$, for all η and all x, do furnish a manifestly covariant realization of localized states in the sense that the superobserver has, by Eq. (4.16), a basis set for the Newton–Wigner[1] "position" representation for each inertial frame.

The hyperplane position operator used by the superobserver may now be defined by

$$Y^{\mu}(\eta, \tau) = (\xi^{i'})^{\mu} X^{i'}(\tau) = l^{\mu}{}_{i}(\eta) X^{i'}(\tau), \tag{4.19}$$

where the hyperplane operator $X^{i'}(\tau)$ has been given by Eqs. (4.9) and (4.10). From Eq. (2.7), $K^{\mu}(\eta) = l^{\mu}{}_{i}(\eta)P^{i'}$ and $H(\eta) = P^{0'}$, so that Eq. (4.19) becomes

$$Y^{\mu}(\eta, \tau) = Y^{\mu}(\eta, 0) + \tau \frac{K^{\mu}(\eta)}{H(\eta)}, \tag{4.20}$$

the local hyperplane operator first written in this form by Fleming.[6] Here,

$$Y^{\mu}(\eta, 0) = il^{\mu}{}_{i}(\eta)\left(\frac{\partial}{\partial p_{i'}} + [2\omega^2(p)]^{-1} p^{i'}\right). \tag{4.21}$$

In terms of the hyperplane momenta $k^{\mu} = l^{\mu}{}_{i}(\eta)p^{i'}$, it follows immediately that

$$Y^{\mu}(\eta, 0) = i\left(\frac{\partial}{\partial k_{\mu}} + [2\omega^2(k)]^{-1} k^{\mu}\right). \tag{4.22}$$

The connection between the hyperplane states localized over the origin and the instantaneous hyperplane localized states may now be demonstrated. From Eq. (4.14), with the connection between the respective hyperplane and superobserver canonical

states given by Eq. (2.19), it follows that

$$|\tilde{\eta}, 0; 0\alpha(\eta)\rangle = (2\pi)^{-\frac{3}{2}} \int d\tilde{k}[2\omega(k)]^{\frac{1}{2}} |\tilde{\eta}; k(\eta)\alpha(\eta)\rangle_c$$
$$= (2\pi)^{-\frac{3}{2}} \int d\tilde{p}'[2\omega(p')]^{\frac{1}{2}} |\eta; p'(\eta)\alpha(\eta)\rangle_c, \tag{4.23}$$

where $k^\mu = l^\mu{}_i(\eta)p'^i$. By taking $p' = l^{-1}(\eta)p$ and inserting the result of Eq. (2.21), Eq. (4.23) becomes

$$|\tilde{\eta}, 0; 0\alpha(\eta)\rangle = (2\pi)^{-\frac{3}{2}} \int d(l^{-1}\mathbf{p})[2\omega(l^{-1}p)]^{\frac{1}{2}} L(\eta) |(l^{-1}p)\alpha\rangle_c$$
$$= (2\pi)^{-\frac{3}{2}} \int d\tilde{p}[2\omega(l^{-1}\mathbf{p})]^{\frac{1}{2}} L(\eta) |(l^{-1}p)\alpha\rangle_c, \tag{4.24}$$

where $(l^{-1}\mathbf{p})$ is the spatial part of the 4-momentum $l^{-1}(\eta)p$. By noting that

$$\omega(l^{-1}\mathbf{p}) = [m^2 - l_\mu{}^i l^\nu{}_i p^\mu p_\nu]^{\frac{1}{2}} = \eta_\mu p^\mu \tag{4.25}$$

and using Eq. (2.21), Eq. (4.24) becomes

$$|\tilde{\eta}, 0; 0\alpha(\eta)\rangle = (2\pi)^{-\frac{3}{2}} \sum_\beta \int d\tilde{p}[2(\eta p)]^{\frac{1}{2}} D^s_{\beta\alpha}[\tilde{l}^\epsilon(l(\eta))] |p\beta\rangle_c. \tag{4.26}$$

Here $\tilde{l}^\epsilon(l(\eta))$ is the rotation given by Eq. (2.15), which is momentum dependent. By applying the inversion Eq. (3.26) to Eq. (4.26), the connection is thus seen to be

$$|\tilde{\eta}, 0; 0\alpha(\eta)\rangle = \sum_\beta \int d^3x \Psi^{s\epsilon}_{\beta\alpha}(\eta; \mathbf{x}) |\mathbf{x}\beta; 0\rangle, \tag{4.27}$$

where

$$\Psi^{s\epsilon}_{\beta\alpha}(\eta; \mathbf{x}) = (2\pi)^{-3} \int d^3p[(\eta p)/\omega(p)]^{\frac{1}{2}} D^s_{\beta\alpha}[\tilde{l}^\epsilon(l(\eta))] e^{-ip\cdot x}. \tag{4.28}$$

For the special case of the instantaneous hyperplane, $l^\nu_\mu(\eta) = \delta^\nu_\mu$, so that

$$\tilde{T}(l(\eta)) = \mathbb{1},$$
$$D^s_{\beta\alpha}[\tilde{l}(l(\tilde{\eta}))] = \delta_{\beta\alpha}, \tag{4.29}$$

and

$$(\tilde{\eta}p) = \omega(p) = (m^2 + \mathbf{p}\cdot\mathbf{p})^{\frac{1}{2}}.$$

Then, Eq. (4.28) reduces to

$$\Psi^s_{\beta\alpha}(\tilde{\eta}; \mathbf{x}) = \delta_{\beta\alpha}(2\pi)^{-3} \int d^3p e^{-ip\cdot x}$$
$$= \delta_{\beta\alpha}\delta^3(\mathbf{x}), \tag{4.30}$$

so that Eq. (4.27) becomes the identity

$$|\tilde{\eta}, 0; 0\alpha(\tilde{\eta})\rangle = |\mathbf{x} = 0\alpha; 0\rangle. \tag{4.31}$$

ACKNOWLEDGMENTS

The author would like to thank Professor D. L. Pursey for many interesting discussions, and the National Aeronautics and Space Administration for its financial assistance during the period in which this work was performed.

* Work was performed in the Ames Laboratory of the U.S. Atomic Energy Commission. Contribution No. 2579.

† Present address: Department of Applied Mathematics, University of Liverpool, Liverpool 3, United Kingdom.

[1] T. D. Newton and E. P. Wigner, Rev. Mod. Phys. **21**, 400 (1949).
[2] M. H. L. Pryce, Proc. Roy. Soc. (London) **195A**, 62 (1948).
[3] L. L. Foldy, Phys. Rev. **102**, 568 (1956).
[4] M. Jacob and G.-C. Wick, Ann. Phys. (N.Y.) **7**, 404 (1959).
[5] C. Møller, Commun. Dublin Inst. Advanced Studies A, No. 5 (1949) (unpublished).
[6] G. N. Fleming, Phys. Rev. **137**, B188 (1965).
[7] G. N. Fleming, J. Math. Phys. **7**, 1959 (1966).
[8] S. C. McDonald, J. Math. Phys. **10**, 1234 (1969).
[9] C. L. Hammer, S. C. McDonald, and D. L. Pursey, Phys. Rev. **171**, 1349 (1968).
[10] A. J. Macfarlane, J. Math. Phys. **4**, 490 (1963).
[11] V. Bargmann and E. P. Wigner, Proc. Natl. Acad. Sci. U.S. **34**, 211 (1948).
[12] D. L. Pursey, Nucl. Phys. **53**, 174 (1964); Ann. Phys. (N.Y.) **32**, 157 (1965).
[13] E. P. Wigner, Nachr. Akad. Wiss. Göttingen, Math.-Physik Kl., IIa, **31**, 546 (1932).
[14] R. A. Berg, J. Math. Phys. **6**, 34 (1965).

Branching of Solutions to Some Nonlinear Eigenvalue Problems

EDWARD T. DEAN* AND PAUL L. CHAMBRÉ†
Department of Nuclear Engineering, University of California
Berkeley, California 94720

(Received 4 September 1969)

The multiplicity of solutions to nonlinear eigenvalue problems is analyzed by a method originally proposed by Hammerstein, and the results are given in tables which display the relation between the number of solutions for eigenvalues in the neighborhood of a critical eigenvalue and the properties of the nonlinear function and its derivatives. Several general types of nonlinear functions are considered, and a simple method of estimating the critical eigenvalue for each type is presented. Since these functions describe physical phenomena, a stability analysis is given for the cases of multiple solutions in order to determine which solution represents the observed physical state. The results are applied to the following nonlinear eigenvalue problems: (i) nonlinear heat generation and the temperature distribution in conducting solids; (ii) temperature distribution in a heat-conducting gas undergoing chemical reactions, leading to a thermal explosion; (iii) nonlinear effects of temperature-dependent viscosity on the temperature distribution of a fluid flowing in a pipe; (iv) neutron flux distribution in a reactor for temperature-dependent cross sections.

1. INTRODUCTION

The current interest in nonlinear eigenvalue problems is devoted to the study of the branching of solutions for certain values of the eigenparameter, known as "bifurcation points."[1] The investigation of these particular characteristics of such problems is important since many physical situations can be described by boundary-value problems of this kind.[1–6] Because these physical problems usually involve the equilibrium distribution of a certain quantity, it is necessary to determine if the solution is multivalued for particular circumstances.

Millman and Keller treat the problem of determining the steady-state temperature distribution due to a nonlinear heat source or sink. They use perturbation theory to obtain a solution and, in the process, make some comments concerning the location of bifurcation points and the nature of the solution branching. As we show later, however, perturbation theory can be used only in a limited way to determine the nature of the branching solutions. We use for this purpose a different method, first proposed by Hammerstein,[7] and independent of the inherent limitations of perturbation theory.

We show that, for some functions $f(x, u)$ in Eq. (2.1) having characteristics of a practically important nature, there are two, one, or no real solutions to the eigenvalue problem for certain values of the eigenparameter. If there is a bifurcation point producing multiple solutions for some eigenvalues, its numerical value can be estimated for certain types of nonlinear terms, which include some of practical importance. If the solutions to the so-called "steady-state" problem (2.1) and (2.2) are perturbed an infinitesimal amount and the related time-dependent problem (4.1) is studied, then it is possible to determine the relative stability associated with the steady-state solutions. If there happen to be two equilibrium solutions of (2.1) and (2.2), one can obtain both of the solutions by the method proposed by Shampine.[8] This method is one of successive approximations by which the successive iterate converges monotonely and uniformly upwards to the lesser solution or downwards to the greater solution, depending on the initial trial function.

We consider four examples of practical importance: (i) nonlinear heat generation and the temperature distribution in conducting solids; (ii) temperature distribution in a heat-conducting gas undergoing chemical reaction, leading to a thermal explosion; (iii) nonlinear effects of temperature-dependent viscosity on the temperature distribution of a fluid flowing in a pipe; (iv) neutron flux distribution in a reactor for temperature-dependent cross sections.

2. BRANCHING OF SOLUTIONS

We study the branching of solutions of

$$Lu = \lambda f(x, u), \quad x \in D, \tag{2.1}$$

$$Bu = 0, \quad x \in \partial D, \tag{2.2}$$

where $x = (x_1, x_2, \cdots, x_m)$, L is the uniformly elliptic self-adjoint operator

$$Lu = -\sum_{i,j=1}^{m} \frac{\partial}{\partial x_i}\left(a_{ij}(x)\frac{\partial u}{\partial x_j}\right) + a_0(x)u,$$

and the boundary conditions are of the form

$$Bu \equiv \alpha(x)u + \beta(x)\frac{\partial u}{\partial \nu}.$$

The coefficients $a_{ij}(x) = a_{ji}(x)$ are continuously differentiable, $a_0(x) \geq 0$ is continuous, and, for all unit vectors $\xi = (\xi_1, \xi_2, \cdots, \xi_m)$,

$$\sum_{i,j=1}^{m} a_{ij}(x)\xi_i\xi_j \geq a > 0, \quad x \in D.$$

The functions $\alpha(x)$ and $\beta(x)$ are piecewise continuous with $\alpha(x) \equiv 1$ and $\beta(x) \equiv 0$ on ∂D_1, where $\partial D_1 + \partial D_2 = \partial D$ and ∂D_1 is of positive measure. On ∂D_2, $\alpha(x) \geq 0$, $\beta(x) \geq 0$, and $\partial/\partial\nu$ is the conormal derivative

$$\frac{\partial u}{\partial \nu} \equiv \sum_{i,j=1}^{m} n_i(x)a_{ij}(x)\frac{\partial u}{\partial x_j},$$

with $n(x) = (n_1(x), n_2(x), \cdots, n_m(x))$ being the outer unit normal to ∂D at x.

With these assumptions, and using the Green's function for the operator $-L$, we transform the nonlinear eigenvalue problem (2.1) and (2.2) into the equivalent nonlinear integral equation

$$u(x; \lambda) = \lambda \int_D G(x; t)f(t, u(t; \lambda))\, dt. \tag{2.3}$$

This is the equation considered by Hammerstein,[7] and later by Vainberg and Trenogin.[9]

Let $u_0(x)$ be a solution of (2.3) for $\lambda = \lambda_b$; that is,

$$u_0(x) = \lambda_b \int_D G(x; t)f(t, u_0(t))\, dt, \tag{2.4}$$

where λ_b is the simple principal eigenvalue and $\phi(x)$ is the corresponding normalized eigenfunction, if it exists, of the equation

$$L\phi = \lambda_b f_u(x, u_0(x))\phi, \quad x \in D, \tag{2.5}$$

$$B\phi = 0, \quad x \in \partial D. \tag{2.6}$$

$\phi(x)$ is assumed to be real and strictly positive in D. We note that $\phi(x)$ is also the principal eigenfunction of the equivalent integral equation

$$\phi(x) = \lambda_b \int_D G(x; t)f_u(t, u_0(t))\phi(t)\, dt. \tag{2.7}$$

Here $f_{u^{(n)}}(x, u(x))$ denotes the nth derivative of f with respect to u.

If v is to describe the solution of (2.3) in the neighborhood of u_0 for λ in the neighborhood of λ_b, then we set $v = u - u_0$ and $\mu = \lambda - \lambda_b$. Assuming that $f(x, u)$ is an analytic function of u so that it can be expanded in terms of a Taylor series about $u = u_0$, Hammerstein obtains, from Eq. (2.3),

$$\begin{aligned} v(x; \mu) - \lambda_b \int_D G(x; t)f_u(t, u_0(t))v(t; \mu)\, dt \\ = \mu \int_D G(x; t)[f(t, u_0(t)) + v(t; \mu)f_u(t, u_0(t))]\, dt \\ + (\lambda_b + \mu)\int_D G(x; t) \\ \times [v^2(t; \mu)f_{u^{(2)}}(t, u_0(t))/2! + \cdots]\, dt. \end{aligned} \tag{2.8}$$

In order that Eq. (2.8) be solvable, it is necessary that the right-hand side of (2.8) be orthogonal to $\phi^+(x)$, the normalized eigenfunction of the transposed kernel $G(t; x)f_u(x, u_0(x))$ corresponding to the characteristic value λ_b. As can be verified by multiplying (2.7) by $f_u(x, u_0(x))$, $\phi^+(x)$ is given by

$$\phi^+(x) = \sigma f_u(x, u_0(x))\phi(x),$$

where $\sigma^2 = (\int_D (f_u\phi)^2\, dx)^{-1}$. Since $\phi^+(x)$ is the solution to a homogeneous equation, one can choose the appropriate sign for σ when $f_u > 0$ or $f_u < 0$, to insure that $\phi^+ > 0$, just as $\phi > 0$.

We let

$$\begin{aligned} \epsilon(\mu) &\equiv \int_D \phi(t)v(t; \mu)\, dt \\ &\equiv \int_D \phi(t)[u(t; \lambda_b + \mu) - u_0(t)]\, dt; \end{aligned}$$

this orthogonality condition leads, after some effort, to the "branch equation"

$$\sum_{m=2}^{\infty} L_{m0}\epsilon^m + \sum_{m=0}^{\infty}\epsilon^m \sum_{l=1}^{\infty}\mu^l L_{ml} = 0. \tag{2.9}$$

Here the first few "branch coefficients" are

$$L_{01} = \frac{1}{\lambda_b}\int_D dx \int_D dt G(x; t)f(t, u_0(t))\phi^+(x), \tag{2.10}$$

$$L_{20} = \tfrac{1}{6}\lambda_b \int_D dx \int_D dt G(x; t)f_{uu}(t, u_0(t))\phi^2(t)\phi^+(x), \tag{2.11}$$

and if it happens that $f_{uu}(x, u_0(x)) \equiv 0$ on D, L_{11} and L_{30} reduce to

$$L_{11} = \tfrac{1}{2}\int_D dx \int_D dt G(x; t)f_u(t, u_0(t))\phi(t)\phi^+(x), \tag{2.12}$$

$$L_{30} = \tfrac{1}{24}\lambda_b \int_D dx \int_D dt G(x; t)f_{u^{(3)}}(t, u_0(t))\phi^3(t)\phi^+(x). \tag{2.13}$$

In general, if

$$\begin{aligned} f_{uu}(x, u_0(x)) \equiv 0,\ f_{u^{(3)}}(x, u_0(x)) \\ \equiv 0, \cdots, f_{u^{(n-1)}}(x, u_0(x)) \equiv 0 \end{aligned}$$

on D, then $L_{20} = L_{30} = \cdots = L_{(n-1),0} = 0$ and L_{n0} reduces to

$$L_{n0} = \frac{\lambda_b}{(n+1)!} \int_D dx \int_D dt G(x, t) f_{u^{(n)}}(t, u_0(t)) \times \phi^n(t)\phi^+(x). \quad (2.14)$$

Because of the nature of the eigenfunctions $\phi(x)$ and $\phi^+(x)$, it is clear that the properties of $f(x, u)$ and its partial derivatives with respect to u determine the branch coefficients, and these in turn prescribe the nature of the bifurcation. A classification is made below.

Solving the implicit Eq. (2.9) for ϵ in terms of μ, for small values of μ, will detail the nature of the branching of solutions in the neighborhood of the bifurcation point $\lambda = \lambda_b$. Such an expression for ϵ can be obtained by "Newton's diagram" or "polygon" method,[10] for specific functions $f(x, u)$.

In addition, Eq. (2.8) can be solved[11] for small values of the dependent variable $v(x; \mu)$, thereby giving the solution $u(x; \lambda)$ in the neighborhood of the known solution. For our purposes, it is sufficient to note that the solution $v(x; \mu) \equiv u(x; \lambda_b + \mu) - u_0(x)$ can be expressed as a double series in powers of ϵ and μ, ϵ being determined by (2.9). This series is of the form

$$v(x; \mu) = \sum_{m=1}^{\infty} \epsilon^m v_{m0}(x) + \sum_{m=0}^{\infty} \epsilon^m \sum_{n=1}^{\infty} \mu^n v_{mn}(x). \quad (2.15)$$

Since ϵ and μ are restricted to small values, the positivity or negativity of $v(x; \mu)$ is determined by the largest term. Suppose, for instance, that $\epsilon(\mu)$ is given by $\pm(K_1\mu)^{\frac{1}{2}}[1 + W_1(\mu)]$, where $W_1(\mu) \to 0$ as $\mu \to 0$ and where K_1 is such that $(K_1\mu) > 0$ for small positive or negative values of μ. Then the largest term of (2.15) is $\pm(K_1\mu)^{\frac{1}{2}}v_{10}(x)$. It can be shown[12] that $v_{10}(x)$ is just equal to $\phi(x)$, the positive eigenfunction, so that the sign of $v(x; \mu)$ is determined by the sign preceding this largest term. Hence, in this case, $u(x; \lambda) \equiv u_0(x) + v(x; \lambda - \lambda_b)$ is greater than or less than $u_0(x)$, depending on the sign of ϵ.

Suppose, in another instance, that $\epsilon(\mu)$ is given by $K_2\mu[1 + W_2(\mu)]$, where $W_2(\mu) \to 0$ as $\mu \to 0$. Then in this case, the largest term of (2.15) is given by $\mu[K_2v_{10}(x) + v_{01}(x)]$. If it happens that $f(t, u_0(t)) \equiv 0$, one can show[12] that $v_{01}(x) = 0$. Hence, this largest term reduces to $\mu K_2\phi(x)$, so that the sign of $v(x; \mu)$ is again determined by the sign of ϵ.

We now consider functions $f(x, u)$ with special properties, which are associated with problems of practical importance.

3. SPECIAL CASES

A. $f(x, u_0(x)) \equiv 0$, $f_u(x, u_0(x)) \not\equiv 0$

In this and subsequent analyses, we assume that there exists a solution to (2.5) and (2.6), a necessary but not sufficient condition for branching, where $u_0(x)$ is taken to be positive in D.

In this case, $L_{01} = 0$ and L_{11}, L_{20}, and L_{30} are the same as in Sec. 2.

If $f_{uu}(x, u_0(x)) \not\equiv 0$, the Newton diagram method gives $\epsilon(\mu) = (-L_{11}\mu/L_{20})[1 + \omega_1(\mu)]$, where $\omega_1(\mu) \to 0$ as $\mu \to 0$. This indicates that $u(x; \lambda)$ is single valued for λ in the neighborhood of λ_b. Moreover, from the closing statements of Sec. 2, $u(x, \lambda)$ is greater than or less than $u_0(x)$, depending on the sign of $L_{11}\mu/L_{20}$.

If $f_{uu}(x, u_0(x)) \equiv 0$, by chance, and $f_{u^{(3)}}(x, u_0(x)) \not\equiv 0$, then the Newton diagram method gives $\epsilon(\mu) = \pm(-L_{11}\mu/L_{30})^{\frac{1}{2}}[1 + \omega_2(\mu)]$, where $\omega_2(\mu) \to 0$ as $\mu \to 0$. This implies that there are two real solutions for $\lambda > \lambda_b$, when $\operatorname{sgn} L_{11} = -\operatorname{sgn} L_{30}$. The situation is reversed when $\operatorname{sgn} L_{11} = \operatorname{sgn} L_{30}$.

Suppose, for example, that $f(x, u_0(x)) \equiv 0$,

$$f_u(x, u_0(x)) < 0, \quad f_{uu}(x, u_0(x)) \equiv 0,$$

and $f_{u^{(3)}}(x, u_0(x)) > 0$. Then the Newton diagram method yields $\epsilon(\mu) = \pm(-L_{11}\mu/L_{30})^{\frac{1}{2}}[1 + \omega_2(\mu)]$, as indicated in the previous paragraph. Since $\phi(x)$ is real and strictly positive and $f_u(x, u_0(x)) < 0$, we conclude that $L_{11} < 0$. Likewise, since $f_{u^{(3)}}(x, u_0(x)) > 0$, it follows that $L_{30} > 0$. Thus, the quotient $-L_{11}/L_{30}$ is positive, and $\epsilon(\mu)$ is real for $\mu > 0$ only. We conclude that there are two real values of ϵ, that is, two solutions $u(x; \lambda)$, for $\lambda > \lambda_b$ and no real values of ϵ for $\lambda < \lambda_b$. Furthermore, according to the arguments given at the end of Sec. 2, the two solutions $u_1(x; \lambda)$ and $u_2(x; \lambda)$ are such that $u_1(x; \lambda) \leq u_0(x) \leq u_2(x; \lambda)$. If $u_0(x) > 0$, then the two solutions are also positive.

Similar considerations lead to the results summarized in Tables I and II. The plus and minus signs indicate that $u > u_0$ and $u < u_0$, respectively.

In general, if $f(x, u_0(x)) \equiv 0$, $f_u(x, u_0(x)) < 0$,

$$f_{uu}(x, u_0(x)) \equiv f_{u^{(n-1)}}(x, u_0(x)) \equiv 0,$$

and $f_{u^{(n)}}(x, u_0(x)) \not\equiv 0$, then similar arguments lead to the results summarized in Table III. If $f_u(x, u_0(x)) > 0$, the situation in Table III is reversed, as in Table II.

Qualitative curves based on the considerations leading to Tables I and II and depicting $u_m \equiv \max u$ vs λ, are given in Fig. 1 for some cases of branching when there are double solutions for some values of λ.

We note in Fig. 1(a) that the branch corresponding to $u_m > \max u_0$ has positive slope. That is, $d\lambda/du_m > 0$.

TABLE I. The number of real solutions $u(x;\lambda)$ for λ in the neighborhood of λ_b when $f(x, u_0(x)) \equiv 0$ and $f_u(x, u_0(x)) < 0$ for all $x \in D$, where the (+) and (−) signs indicate $u(x;\lambda) > u_0(x)$ and $u(x;\lambda) < u_0(x)$, respectively.

	$f_{uu}(x, u_0(x)) \not\equiv 0$		$f_{uu}(x, u_0(x)) \equiv 0$	
	$f_{uu}(x, u_0(x)) > 0$	$f_{uu}(x, u_0(x)) < 0$	$f_{u^{(3)}}(x, u_0(x)) > 0$	$f_{u^{(3)}}(x, u_0(x)) < 0$
$\lambda > \lambda_b$	1:(+)	1:(−)	2:(+), (−)	0
$\lambda < \lambda_b$	1:(−)	1:(+)	0	2:(+), (−)

TABLE II. The number of real solutions $u(x;\lambda)$ for λ in the neighborhood of λ_b when $f(x, u_0(x)) \equiv 0$ and $f_u(x, u_0(x)) > 0$ for all $x \in D$, where the (+) and (−) signs indicate $u(x;\lambda) > u_0(x)$ and $u(x;\lambda) < u_0(x)$, respectively.

	$f_{uu}(x, u_0(x)) \not\equiv 0$		$f_{uu}(x, u_0(x)) \equiv 0$	
	$f_{uu}(x, u_0(x)) > 0$	$f_{uu}(x, u_0(x)) < 0$	$f_{u^{(3)}}(x, u_0(x)) > 0$	$f_{u^{(3)}}(x, u_0(x)) < 0$
$\lambda > \lambda_b$	1:(−)	1:(+)	0	2:(+), (−)
$\lambda < \lambda_b$	1:(+)	1:(−)	2:(+), (−)	0

TABLE III. The number of real solutions $u(x;\lambda)$ for λ in the neighborhood of λ_b when $f(x, u_0(x)) \equiv 0$ and $f_u(x, u_0(x)) < 0$ for all $x \in D$, and $f_{uu}(x, u_0(x)) \equiv f_{u^{(3)}}(x, u_0(x)) \equiv \cdots \equiv f_{u^{(n-1)}}(x, u_0(x)) \equiv 0$; the (+) and (−) signs indicate $u(x;\lambda) > u_0(x)$ and $u(x;\lambda) < u_0(x)$, respectively.

	n even		n odd	
	$f_{u^{(3)}}(x, u_0(x)) > 0$	$f_{u^{(3)}}(x, u_0(x)) < 0$	$f_{u^{(n)}}(x, u_0(x)) > 0$	$f_{u^{(n)}}(x, u_0(x)) < 0$
$\lambda > \lambda_b$	1:(+)	1:(−)	2:(+), (−)	0
$\lambda < \lambda_b$	1:(−)	1:(+)	0	2:(+), (−)

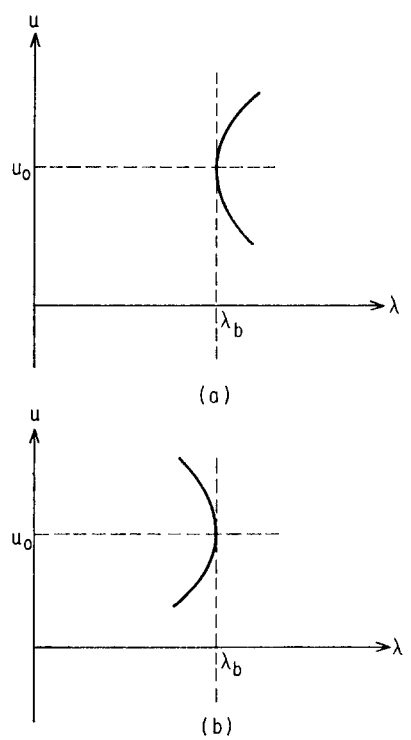


FIG. 1. Branching of solutions for the case when $f(x, u_0(x)) \equiv 0 \equiv f_{uu}(x, u_0(x))$ and (a) $f_u(x, u_0(x))/f_{u^{(3)}}(x, u_0(x)) < 0$, (b) $f_u(x, u_0(x))/f_{u^{(3)}}(x, u_0(x)) > 0$.

The other branch, corresponding to $u_m \equiv \max u < \max u_0$ has negative slope $d\lambda/du_m < 0$. The opposite is true for Fig. 1(b). We show in Sec. 5 that $\lambda = \lambda_b$ is a point of neutral stability.

B. $f(x, u_0(x)) \not\equiv 0$, $f_u(x, u_0(x)) \not\equiv 0$

In this case, the relevant branch coefficients L_{01} and L_{20} are the same as given in Sec. 2. Then the Newton diagram method gives $\epsilon(\mu) = \pm(-L_{01}\mu/L_{20})^{\frac{1}{2}} \times [1 + \omega_3(\mu)]$, where $\omega_3(\mu) \to 0$ as $\mu \to 0$. Thus, there are two solutions $u(x;\lambda)$ for $\lambda < \lambda_b$, and there is no solution for $\lambda > \lambda_b$, when $\operatorname{sgn} L_{01} = \operatorname{sgn} L_{20}$. The reversed situation holds when $\operatorname{sgn} L_{01} = -\operatorname{sgn} L_{20}$. Furthermore, the two solutions $u_1(x;\lambda)$ and $u_2(x;\lambda)$ are such that $\max u_1 < \max u_0 < \max u_2$.

For the particular case of

$$f(x, u_0(x)) > 0, \quad f_u(x, u_0(x)) > 0,$$

we summarize the results in Table IV.

We give the qualitative curves in Fig. 2 for the particular case described in Table IV. Again, we take note that in Fig. 2(a) the branch corresponding to $u_m = \max u > \max u_0$ has positive slope $d\lambda/du_m > 0$; the other branch, corresponding to $u_m < \max u_0$ has negative slope, implying that $d\lambda/du_m < 0$. In this case also, $\lambda = \lambda_b$ is a point of neutral stability, as is shown.

Once it has been established that branching occurs, an estimate of the numerical value of the bifurcation

TABLE IV. The number of real solutions $u(x;\lambda)$ for λ in the neighborhood of λ_b when $f(x, u_0(x)) > 0$ and $f_u(x, u_0(x)) > 0$, where the (+) and (−) signs indicate $u(x;\lambda) > u_0(x)$ and $u(x;\lambda) < u_0(x)$, respectively.

	$f_{uu}(x, u_0(x)) > 0$	$f_{uu}(x, u_0(x)) < 0$
$\lambda > \lambda_b$	0	2:(+), (−)
$\lambda < \lambda_b$	2:(+), (−)	0

point is essential. In problems of physical interest this point represents a critical value, beyond which, or below which, an equilibrium state cannot exist.

4. BOUNDS ON λ_b

The difficulty in calculating the critical value results from the fact that this value of λ is determined by the simultaneous equations (2.4) and (2.7). Joseph[10] gives a technique of obtaining an upper bound on the bifurcation point when the nonlinear term $f(x, u)$ is of the form $f(x, u) = u + G(u)$, where $G(u) \geq G(0) = 1$ for $u \geq 0$, which is assumed on physical grounds, and the situation of Fig. 2(b) is assumed to exist.

However, it is possible to obtain a fairly accurate value for λ_b for more general cases.

In the special case where $u_0(x) \equiv 0$, that is, $f(x, 0) \equiv 0$ but $f_u(x, 0) \not\equiv 0$ on D, and $Bu \equiv u = 0$ on ∂D, then the difficulty mentioned above does not exist. The bifurcation point can be obtained directly from Eq. (2.7), since Eq. (2.4) is satisfied for any λ_b.

If $f(x, u)$ can be written in the form $f(x, u) = h(x)g(u)$, then λ_b is bounded above for the situation of Fig. 2(b) by ub (λ_b), and below for the situation of Fig. 2(a) by lb (λ_b), where

$$\text{ub}\,(\lambda_b) = \Lambda_0 \max_{\psi>0} \psi/g(\psi), \tag{4.1}$$

$$\text{lb}\,(\lambda_b) = \Lambda_0 \min_{\psi>0} \psi/g(\psi). \tag{4.2}$$

Here Λ_0 is the principal eigenvalue of the associated Helmholtz equation

$$L\zeta = \Lambda h(x)\zeta, \quad x \in D, \tag{4.3}$$

$$B\zeta = 0, \quad x \in \partial D. \tag{4.4}$$

To show this, we write

$$0 = \int_D (\zeta_0 Lu - uL\zeta_0)\, dx$$

$$= \lambda \int_D \zeta_0 hg(u)\, dx - \Lambda_0 \int_D \zeta_0 hu\, dx,$$

where $\zeta_0(x)$ is the eigenfunction associated with Λ_0 and use is made of the self-adjointness of the operator L and its associated boundary conditions.

Then

$$\lambda/\Lambda_0 = \int_D \zeta_0 h(u/g(u))g(u)\, dx \Big/ \int_D \zeta_0 hg(u)\, dx,$$

where it is assumed that the integral $\int_D \zeta_0 hg(u)\, dx$ does not vanish. Thus,

$$\min_{\psi\geq 0} \frac{\psi}{g(\psi)} \leq \frac{\lambda}{\Lambda_0} \leq \max_{\psi\geq 0} \frac{\psi}{g(\psi)}, \tag{4.5}$$

which gives the bounds on λ_b stated above for the situations of Fig. 2.

Note that the inequality (4.5) is quite general, but acquires significant meaning for the specific functions $f(x, u)$ whose properties result in the situations described by Fig. 2.

We now turn to the interesting question of the physical meaning of the existence of two solutions for some values of λ. Clearly, one of the solutions must, in some sense, be associated with a more stable state when the physical system is subjected to a small perturbation. Hence, we attend to the question of stability in order to determine which solution represents the stable state observed in nature.

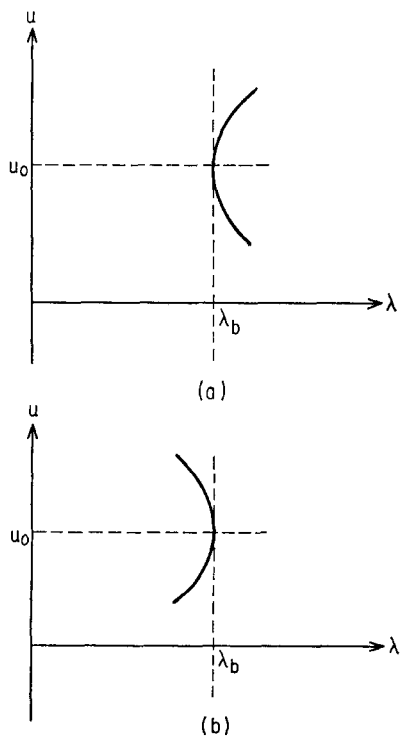


FIG. 2. Branching of solutions for the case when $f(x, u_0(x)) > 0$, $f_u(x, u_0(x)) > 0$, and (a) $f_{uu}(x, u_0(x)) < 0$, (b) $f_{uu}(x, u_0(x)) > 0$.

5. STABILITY ANALYSIS

For this purpose we follow the example of Joseph[4] and treat the related time-dependent problem

$$\frac{\partial u^*}{\partial t} + Lu^* = \lambda f(x, u^*), \quad x \in D, \tag{5.1}$$

$$Bu^* = 0, \quad x \in \partial D, \tag{5.2}$$

where λ and $f(x, u^*)$ are such that there are two solutions to (2.1) and (2.2). We seek solutions of the form

$$u^*(x, t; \lambda) = u(x, \lambda) + u_\beta(x, \lambda)e^{-\beta t},$$

where $u(x, \lambda)$ is the steady-state solution and $u_\beta(x, \lambda)$ represents a small initial perturbation.

The function $f(x, u^*)$, assumed to be analytic in u^*, can be expanded in a Taylor series about $u^* = u$, and, since u_β is taken to be an infinitesimal disturbance, the second- and higher-order terms can be neglected. Then the equations for u_β are given by

$$Lu_\beta = [\beta + \lambda f_u(x, u(x; \lambda))]u_\beta, \quad x \in D, \tag{5.3}$$

$$Bu_\beta = 0, \quad x \in \partial D, \tag{5.4}$$

from which we can determine the value of β associated with each solution. If $\beta > 0$, then the solution is stable; conversely, $\beta < 0$ means that the solution is unstable. $\beta = 0$ implies neutral stability.

Taking the derivative of both sides of equations (2.1) and (2.2) with respect to $u_m = \max u$ yields

$$L\dot{u} = \lambda f_u(x, u)\dot{u} + \dot{\lambda} f(x, u), \quad x \in D, \tag{5.5}$$

$$B\dot{u} = 0, \quad x \in \partial D, \tag{5.6}$$

where $\dot{u} = \partial u/\partial u_m$ and $\dot{\lambda} = d\lambda/du_m$.

A. Point of Neutral Stability $f(x, u_0(x)) \not\equiv 0$

If $f(x, u_0(x)) \not\equiv 0$ for $x \in D$, then, when $\dot{\lambda} = 0$ and $\beta = 0$, a solution of (5.3) and (5.4) exists and is given by $u_\beta(x) = Au(x)$, where A is an arbitrary constant. In the limit, as λ tends to λ_b and $u(x; \lambda)$ becomes $u_0(x)$, Eqs. (5.5) and (5.6) are satisfied by $\dot{u}(x) = \phi(x)$ and $\dot{\lambda} = 0$. [See Eqs. (2.5) and (2.6)]. Thus, $\lambda = \lambda_b$ is a point of neutral stability and the branches in Fig. 2 have slope infinity there $du_m/d\lambda = \infty$.

B. Stable and Unstable Branches $f(x, u_0(x)) \not\equiv 0$

We write

$$0 = \int_D (\dot{u}Lu_\beta - u_\beta L\dot{u})\, dx = \int_D (\beta\dot{u}u_\beta - \dot{\lambda}u_\beta f(x, u))\, dx.$$

In the neighborhood of λ_b where $\dot{\lambda} \neq 0$, we have

$$\beta/\dot{\lambda} = \int_D u_\beta f(x, u)\, dx \Big/ \int_D \dot{u}u_\beta\, dx,$$

and in the limit as $\dot{\lambda}$ goes to zero, i.e., λ tends to λ_b,

$$d\beta/d\dot{\lambda}|_{\dot{\lambda}=0} = \int_D \phi(x)f(x, u_0(x))\, dx \Big/ \int_D \phi^2(x)\, dx$$

$$= \int_D \phi(x)f(x, u_0(x))\, dx.$$

We conclude that in the neighborhood of λ_b the constant β and the slope of the branch, $du_m/d\lambda = 1/\dot{\lambda}$, have the same sign if $f(x, u_0(x)) > 0$. If $f(x, u_0(x)) < 0$, they have opposite sign. Thus, in Fig. 2(a), the stable branch is the one corresponding to $u_m = \max u > \max u_0$, and the unstable branch is the one corresponding to $u_m < \max u_0$. In Fig. 2(b), the situation is reversed.

C. Point of Neutral Stability $f(x, u_0(x)) \equiv 0$

Taking the derivative of both sides of Eqs. (4.4) and (4.5) with respect to $u_m = \max u$ yields

$$L\ddot{u} = \lambda f_u(x, u)\ddot{u} + f_{uu}(x, u)\dot{u}^2 + 2\lambda f_u(x, u)\dot{u} + \ddot{\lambda} f(x, u), \quad x \in D,$$

$$B\ddot{u} = 0, \quad x \in \partial D,$$

where $\ddot{\lambda} = \partial^2\lambda/\partial u_m^2$ and $\ddot{u} = \partial^2 u/\partial u_m^2$.

If $f_u(x, u_0(x)) \not\equiv 0$ on D and $f(x, u_0(x)) \equiv 0 \equiv f_{uu}(x, u_0(x))$, then, when $\dot{\lambda} = 0$ and $\beta = 0$ in the limit as λ tends to λ_b and $u(x; \lambda)$ becomes $u_0(x)$, there is a solution of (5.3) and (5.4), given by $u_\beta(x) = A\ddot{u}(x) \equiv A\phi(x)$. Thus, $\lambda = \lambda_b$ is a point of neutral stability and both branches in Figs. 1(a) and 1(b) have slope infinity there.

D. Stable and Unstable Branches $f(x, u_0(x)) \equiv 0$

Employing the same arguments of Sec. B above, we find for the case $f(x, u_0(x)) \equiv 0$ that in the limit

$$\left.\frac{d\beta}{d\dot{\lambda}}\right|_{\dot{\lambda}=0} = 2\int_D \phi^2(x)f_u(x, u_0(x))\, dx.$$

We conclude, then, that in the neighborhood of λ_b, the constant β and the slope of the branch $du_m/d\lambda = 1/\dot{\lambda}$, have the same sign if $f_u(x, u_0(x)) > 0$, and the opposite sign if $f_u(x, u_0(x)) < 0$.

Thus, the stable branches in Fig. 1 are those corresponding to $u_m = \max u > \max u_0$, if $f_{u^{(3)}}(x, u_0(x)) < 0$, or those corresponding to $u_m < \max u_0$, if

$$f_{u^{(3)}}(x, u_0(x)) > 0.$$

The other branches are unstable to a small disturbance. We now consider four examples of physical significance.

6. EXAMPLES

A. Nonlinear Heat Generation and the Temperature Distribution in Conducting Solids

Millman and Keller consider the problem

$$\Delta T - \lambda S(T) = 0, \quad x \in D, \tag{6.1}$$

$$\frac{\partial T}{\partial n} = \alpha(T - T_0), \quad x \in \partial D, \tag{6.2}$$

where T_0 is the constant temperature outside D and $S(T_0) = 0$. They employ perturbation theory to solve (5.1) and (5.2) for T in the neighborhood of T_0, and to make comments on the bifurcation of the solution.

Since T_0 is a constant and $S(T_0) = 0$ on D, $T(x) = T_0$ is a solution to (6.1) and (6.2) for all values of λ and, if branching occurs, the bifurcation point is, by (2.5) and (2.6), simply the principal eigenvalue of the Sturm–Liouville system

$$\Delta\phi - \Lambda S'(T_0)\phi = 0, \quad x \in D,$$

$$\frac{\partial \phi}{\partial n} = \alpha(\phi - T_0), \quad x \in \partial D.$$

Suppose $S''(T_0) = 0$ and $S'(T_0) < 0$. Then, by what has gone before (see Table I), we conclude that there are two solutions for $\lambda > \lambda_b$ and no solution for $\lambda < \lambda_b$ when $S'''(T_0) > 0$, and that there are two solutions for $\lambda < \lambda_b$ and no solution for $\lambda > \lambda_b$ when $S'''(T_0) < 0$. Furthermore, the stability analysis in Sec. 5D indicates that the smaller solution is the stable solution, and the larger solution is unstable, if $S'''(T_0) > 0$; the reverse is true if $S'''(T_0) < 0$. This agrees with the results of Millman and Keller.

Suppose $S''(T_0) \neq 0$ and $S'(T_0) < 0$. Then, by the results for this special case (shown in Table I), there is only one solution for λ in the neighborhood of λ_b.

However, Millman and Keller conclude, on the basis of their perturbation analysis, that there are again two solutions. It is believed that this conclusion results from extending the perturbation treatment beyond the limits of its validity. The analysis of the previous sections has no such intrinsic limits.

B. Temperature Distribution in an Exothermic Chemically Reacting System

One of the co-authors[2] treats the problem of thermal explosions in a chemically reacting system. The governing equation for the steady-state temperature distribution is given by

$$\frac{d^2\theta}{dx} + \frac{k}{x}\frac{d\theta}{dx} + \delta e^{\theta} = 0, \quad x \in [0, 1], \tag{6.3}$$

$$\theta'(0) = \theta(1) = 0, \tag{6.4}$$

where $k = 0$ corresponds to an infinite plane-parallel vessel and $k = 1$ corresponds to a cylindrical vessel of infinite length.

Since $f(x, \theta) = e^{\theta}$ and all its derivatives are positive, the nature of the bifurcation of solutions to this problem is described by Table IV and Fig. 2(b). There are two solutions for $\lambda < \lambda_b$, and the stability analysis of Sec. 5B predicts that the stable solution is the smaller one, while the larger solution is unstable. There is no solution for $\lambda > \lambda_b$, indicating that $\lambda > \lambda_b$ is a threshold value for a thermal explosion. Making use of (4.1), we see that this critical value is bounded above by

$$\lambda_b \leq \Lambda_0 \max_{\psi \geq 0} \frac{\psi}{\exp \psi} = \frac{\Lambda_0}{e}, \quad \text{at} \quad \psi = 1,$$

where Λ_0 is the principal eigenvalue for the associated Helmholtz equation,

$$\frac{d^2\zeta}{dx^2} + \frac{k}{x}\frac{d\zeta}{dx} + \Lambda\zeta = 0, \quad x \in D,$$

$$\zeta'(0) = \zeta(1) = 0, \quad x \in \partial D.$$

For $k = 0$, we have $\Lambda_0 = \frac{1}{4}\pi^2$; for $k = 1$, $\Lambda_0 = 5.784$, i.e., the first positive root of $J_0(\Lambda^{\frac{1}{2}}) = 0$. The upper bound on λ_b is then 0.908 for $k = 0$, and 2.127 for $k = 1$. This compares well with the known values for λ_b of 0.893 and 2.000, respectively.

C. Nonlinear Temperature-Dependent Viscosity Effects on Temperature Distribution in a Flowing Fluid

Joseph[4] considers the problem of fluid flowing in a channel, where the viscosity is a nonlinear function of the temperature. The appropriate equations for the temperature distribution for Hagen–Poiseuille flow in a cylindrical vessel of infinite length are

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{d\theta}{dr}\right) + \lambda r^2(1 + \theta + a\theta^2) = 0, \quad r[0, 1], \tag{6.5}$$

$$\theta'(0) = \theta(1) = 0, \tag{6.6}$$

where a is a positive constant.

For positive θ, $f(r, \theta) = r^2(1 + \theta + a\theta^2)$ is strictly positive. In addition, $f_\theta(r, \theta) > 0$ and $f_{\theta\theta}(r, \theta) > 0$. We then conclude from Table IV and Fig. 2(b) that there are two solutions for $\lambda < \lambda_b$ and none for $\lambda > \lambda_b$. The stability analysis shows that the smaller solution is stable, and the larger solution is associated with instabilities. Physically, these criteria indicate that stable flow is impossible beyond a critical value of the parameter λ, which includes such physical constants as the channel diameter, reference viscosity, and axial pressure gradient. This critical value is bounded

above by

$$\lambda_b \leq \Lambda_0 \max_{\psi \geq 0} \frac{\psi}{1 + \psi + a\psi^2} = \frac{\Lambda_0}{(2\sqrt{a}) + 1},$$

$$\text{for} \quad \psi = \frac{1}{\sqrt{a}},$$

where Λ_0 is the principal eigenvalue of the associated Helmholtz equation,

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{d\zeta}{dr}\right) + \Lambda r^2 \zeta = 0, \quad r \in [0, 1],$$

$$\zeta'(0) = \zeta(1) = 0.$$

In this case $\Lambda_0 = 23.14$, the first positive root of $J_0(\frac{1}{2}\Lambda^{\frac{1}{2}}) = 0$.

D. Neutron Flux Distribution in a Reactor for Temperature-Dependent Cross Sections

The neutron flux distribution in a nuclear reactor can be approximated by the 1-velocity diffusion equation

$$\Delta\phi + \lambda L^{-2}\phi = 0, \quad x \in D, \tag{6.7}$$

$$\phi = 0, \quad x \in \partial D, \tag{6.8}$$

where L is the "diffusion length."

Moskalev[14] incorporates the temperature dependence of the cross sections by writing (6.7) and (6.8) as

$$\Delta\phi + \lambda a(T)\phi = 0, \quad x \in D, \tag{6.9}$$

$$\phi = 0, \quad x \in \partial D, \tag{6.10}$$

where $a(T) = L_0^{-2}(T_0/T)^{\frac{1}{2}}e^{[-2\beta(T-T_0)]}$, T_0 is the boundary temperature, and β is a small positive constant.

We assume that the temperature depends on the neutron flux according to

$$T - T_c = c\phi,$$

where T_c is the reactor coolant temperature and c is another constant.

Letting $A \equiv (T_0/T_c)^{\frac{1}{2}}e^{-2\beta(T_c/T_0)}/L_0^2$, $\alpha \equiv 2\beta T_c$, and $u \equiv c\phi/T_c$, we can write (6.9) and (6.10) as

$$\Delta u + \lambda A u e^{-\alpha u}/(1 + u)^{\frac{1}{2}} = 0, \quad x \in D, \tag{6.11}$$

$$u = 0, \quad x \in \partial D. \tag{6.12}$$

Since $u(x) \equiv 0$ satisfies (6.11) and (6.12) for all λ, if branching occurs, the bifurcation point is simply, by (2.5) and (2.6), the principal eigenvalue of the Sturm–Liouville system,

$$\Delta\phi + \Lambda A\phi = 0, \quad x \in D,$$

$$\phi = 0, \quad x \in \partial D.$$

Computing the derivatives of the nonlinear term and noting that, for known reactors, $\max u < 1$ since $T < 2T_c$, we find that the first derivative is positive, but the second derivative is negative. Then, from Table II, we conclude that there is only one solution for $\lambda > \lambda_b$ which is positive, since $u_0(x) \equiv 0$, and one negative solution for $\lambda < \lambda_b$, which is disregarded as physically meaningless.

Thus, λ_b again represents a "critical" value, even though no branching occurs. This example also illustrates that the satisfaction by some λ_b of Eqs. (2.4) and (2.7) simultaneously is a necessary but not sufficient condition for branching to occur at λ_b.

* Work done in partial fulfillment of the requirements for the Ph.D. degree, University of California, Berkeley.
† Part of this work was supported under NONR 222(95).

[1] M. H. Millman and J. B. Keller, J. Math. Phys. **10**, 342 (1969).
[2] P. L. Chambré, J. Chem. Phys. **20**, 1795 (1952).
[3] D. D. Joseph, Intern. J. Heat. Mass Transfer **8**, 281 (1965).
[4] D. D. Joseph, Phys. Fluids **7**, 1761 (1964).
[5] H. B. Keller and D. S. Cohen, J. Math. Mech. **16**, 1361 (1967).
[6] S. A. Kaganov, Prikl. Mat. Mekh. **31**, 1081 (1967).
[7] A. Hammerstein, Acta Math. **54**, 117 (1930).
[8] L. F. Shampine, J. Math. Mech. **17**, 1065 (1968).
[9] M. M. Vainberg and V. A. Trenogin, Russ. Math. Surv. **7**, 1 (1962).
[10] See, for example, Ref. 9, pp. 34–44, or E. Hille, *Analytic Function Theory* (Ginn and Co., Boston, 1959), pp. 105–109.
[11] See, for example, Ref. 9, pp. 13–14 and pp. 46–49.
[12] See Ref. 9, pp. 30–32.
[13] D. D. Joseph, Quart. Appl. Math. **23**, 343 (1966).
[14] O. B. Moskalev, Atomic Energy (USSR) **16**, 304 (1964).

Exact Solution of a Family of Integral Equations of Anisotropic Scattering*

H. KAGIWADA† AND R. KALABA
Department of Electrical Engineering, University of Southern California
Los Angeles, California 90007

(Received 20 October 1969)

It is shown that the solution of a certain Cauchy system provides the solution of a family of integral equations occurring in the theory of anisotropic scattering in a finite slab. Numerical experiments show that the Cauchy system is readily solved numerically, even in the case of very strong forward scattering.

I. INTRODUCTION

A key role in the theory of anisotropic scattering in a finite slab is played by the integral equation

$$\begin{aligned} J(t, v, x, u) &= \tfrac{1}{2}c(v, -u)e^{-(x-t)/u} \\ &\quad + \int_t^x \int_0^1 c(v, -v')e^{-(y-t)/v'}J(y, -v', x, u)\frac{dv'}{v'}\,dy \\ &\quad + \int_0^t \int_0^1 c(v, +v')e^{-(t-y)/v'}J(y, v', x, u)\frac{dv'}{v'}\,dy, \\ &\quad 0 \le t \le x,\ 0 \le u \le 1,\ -1 \le v \le +1. \quad (1) \end{aligned}$$

It is assumed that the interval length x is sufficiently small that Eq. (1) possesses a unique solution. For the physical background and analytical and computational approaches, see Refs. 1–3. In this paper it is shown that the family of integral equations (1) is equivalent to a Cauchy system which can readily be handled computationally. First, the theory is given,[4] and then results of numerical experiments are described.

II. STATEMENT OF CAUCHY SYSTEM

We consider the Cauchy system for the auxiliary function S,

$$S(v, u, 0) = 0, \quad (2)$$

where

$$\begin{aligned} S_x(v, u, x) &= -(u^{-1} + v^{-1})S(v, u, x) + 2c(v, -u) \\ &\quad + \int_0^1 c(v, v')S(v', u, x)\frac{dv'}{v'} \\ &\quad + 2\int_0^1 S(v, v', x)\frac{dv'}{v'}\left(\tfrac{1}{2}c(-v', -u)\right. \\ &\quad \left. + \tfrac{1}{4}\int_0^1 \frac{dv''}{v''}\,c(-v', v'')S(v'', u, x)\right), \\ &\quad x \ge 0,\ 0 \le v, u \le 1. \quad (3) \end{aligned}$$

We also consider the Cauchy system for the function J,

$$\begin{aligned} J_x(t, v, x, u) &= -u^{-1}J(t, v, x, u) + \int_0^1 \left(c(-v', -u)\right. \\ &\quad \left. + \tfrac{1}{2}\int_0^1 \frac{dv''}{v''}\,c(-v', v'')S(v'', u, x)\right)J(t, v, x, v')\frac{dv'}{v'}, \\ &\quad x \ge t,\ 0 \le u \le 1,\ -1 \le v \le +1. \quad (4) \end{aligned}$$

The initial condition on the function J at $x = t$ is

$$\begin{aligned} J(t, v, t, u) &= \tfrac{1}{2}c(v, -u) + \tfrac{1}{4}\int_0^1 \frac{dv'}{v'}\,c(v, v')S(v', u, t), \\ &\quad 0 \le u \le 1,\ -1 \le v \le +1. \quad (5) \end{aligned}$$

It is assumed that this Cauchy system possesses a unique solution, at least for x sufficiently small.

The aim is to show that the Cauchy system above and the integral equation (1) are equivalent.[4] These equations have been derived previously on physical grounds.[3]

III. DERIVATION OF CAUCHY SYSTEM

We begin by differentiating both sides of Eq. (1) with respect to x, which yields

$$\begin{aligned} J_x(t, v, x, u) &= -\frac{1}{2u}\,c(v, -u)e^{-(x-t)/u} \\ &\quad + \int_0^1 \frac{dv'}{v'}\,c(v, -v')e^{-(x-t)/v'}J(x, -v', x, u) \\ &\quad + \int_t^x \int_0^1 \frac{dv'}{v'}\,c(v, -v')e^{-(y-t)/v'}J_x(y, -v', x, u)\,dy \\ &\quad + \int_0^t \int_0^1 \frac{dv'}{v'}\,c(v, +v')e^{-(t-y)/v'}J_x(y, v', x, u)\,dy. \quad (6) \end{aligned}$$

By regarding Eq. (6) as an integral equation for the function J_x, keeping Eq. (1) in mind, and using the superposition principle for linear systems, we see

that the solution of Eq. (6) for the function J_x is

$$J_x(t, v, x, u) = -\frac{1}{u} J(t, v, x, u) + 2\int_0^1 J(x, -v', x, u)J(t, v, x, v')\frac{dv'}{v'}. \tag{7}$$

For $J(x, v, x, u)$ we may write, according to Eq. (1),

$$J(x, v, x, u) = \tfrac{1}{2}c(v, -u) + \int_0^x \int_0^1 \frac{dv'}{v'} c(v, v')e^{-(x-y)/v'}J(y, v', x, u)\, dy. \tag{8}$$

We introduce the function S by means of the definition

$$S(v, u, x) = 4\int_0^x e^{-(x-y)/v}J(y, v, x, u)\, dy, \quad 0 \le v, u \le 1, x \ge 0. \tag{9}$$

Equation (8) may now be rewritten

$$J(x, v, x, u) = \tfrac{1}{2}c(v, -u) + \tfrac{1}{4}\int_0^1 \frac{dv'}{v'} c(v, v')S(v', u, x). \tag{10}$$

Next, we differentiate both sides of Eq. (9) with respect to x. This provides the equation

$$S_x(v, u, x) = -\frac{1}{v} S(v, u, x) + 4J(x, v, x, u) + 4\int_0^1 e^{-(x-y)/v}\left(-\frac{1}{u} J(y, v, x, u) + 2\int_0^1 J(x, -v', x, u)J(y, v, x, v')\frac{dv'}{v'}\right) dy, \tag{11}$$

where use has been made of Eq. (7) for the function J_x.

Equation (2) follows directly from the definition in Eq. (9), where the function S is defined in terms of the function J. The differential equation for the function S, Eq. (3), comes directly from Eq. (11) by using Eqs. (8)–(10). The initial condition for the function J at $x = t$ in Eq. (5) follows directly from Eq. (8) with the variable x set equal to t. The differential equation (4) for the function J, valid for $x \ge t$, follows immediately from Eqs. (7) and (10).

IV. VALIDATION OF CAUCHY SYSTEM

Now, we shall show that the solution of the Cauchy system (2)–(5) for the functions S and J provides a solution of the integral equation (1). Our first task is to demonstrate that

$$S(v, u, x) = 4\int_0^x e^{-(x-y)/v}J(y, v, x, u)\, dy, \quad x \ge 0, 0 \le v, u < 1. \tag{12}$$

Let the function Q be defined by the equation

$$Q(v, u, x) = 4\int_0^x e^{-(x-y)/v}J(y, v, x, u)\, dy, \quad x \ge 0, 0 \le v, u \le 1. \tag{13}$$

It clearly satisfies the initial condition

$$Q(v, u, 0) = 0. \tag{14}$$

Furthermore, differentiation of both sides of Eq. (13) with respect to x shows that

$$Q_x(v, u, x) = -\frac{1}{v} Q(v, u, x) + 4J(x, v, x, u) + 4\int_0^x e^{-(x-y)/v}\left[-\frac{1}{u} J(y, v, x, u) + \int_0^1 \left(c(-v', -u) + \tfrac{1}{2}\int_0^1 \frac{dv''}{v''} c(-v', v'')S(v'', u, x)\right) \times J(y, v, x, v')\frac{dv'}{v'}\right] dy. \tag{15}$$

The last equation may be rewritten as

$$Q_x(v, u, x) = -\left(\frac{1}{v} + \frac{1}{u}\right)Q(v, u, x) + 4J(x, v, x, u) + 4\int_0^1 c(-v', -u)\frac{dv'}{v'}\int_0^x e^{-(x-y)/v}J(y, v, x, v')\, dy + 2\int_0^1 \int_0^1 \frac{dv''}{v''}\frac{dv'}{v'} c(-v', v'')S(v'', u, x) \times \int_0^1 e^{-(x-y)/v}J(y, v, x, v')\, dy \tag{16}$$

or

$$Q_x(v, u, x) = -\left(\frac{1}{u} + \frac{1}{u}\right)Q(v, u, x) + 2c(v, -u) + \int_0^1 \frac{dv'}{v'} c(v, v')S(v', u, x) + \int_0^1 c(-v', -u)\frac{dv'}{v'} Q(v, v', x) + \tfrac{1}{2}\int_0^1 \int_0^1 \frac{dv''}{v''}\frac{dv'}{v'} c(-v', v'')S(v'', u, x)Q(v, v', x). \tag{17}$$

Assuming that the linear Cauchy system for the function Q in Eqs. (17) and (14) has a unique solution and keeping in mind the Cauchy system for the function S,

we see that

$$Q = S \tag{18}$$

or

$$S(v, u, x) = 4\int_0^x e^{-(x-y)/v} J(y, v, x, u)\, dy, \qquad x \geq 0,\ 0 \leq v, u \leq 1. \tag{19}$$

Next, we introduce the function M by the equation

$$\begin{aligned} M(t, v, x, u) &= \tfrac{1}{2}c(v, -u)e^{-(x-t)/u} \\ &\quad + \int_t^x \int_0^1 \frac{dv'}{v'}\, c(v, -v')e^{-(y-t)/v'} J(y, -v', x, u)\, dy \\ &\quad + \int_0^t \int_0^1 \frac{dv'}{v'}\, c(v, v')e^{-(t-y)/v'} J(y, +v', x, u)\, dy, \\ &\quad 0 \leq t \leq x,\ 0 \leq u \leq 1,\ -1 \leq v \leq +1. \end{aligned} \tag{20}$$

At $x = t$ we find that

$$\begin{aligned} M(t, v, t, u) &= \tfrac{1}{2}c(v, -u) + \int_0^t \int_0^1 \frac{dv'}{v'}\, c(v, v')e^{-(t-y)/v'} J(y, v', t, u)\, dy \\ &= \tfrac{1}{2}c(v, -u) + \int_0^1 \frac{dv'}{v'}\, c(v, v') \int_0^t e^{-(t-y)/v'} J(y, v', t, u)\, dy \\ &= \tfrac{1}{2}c(v, -u) + \tfrac{1}{4}\int_0^1 \frac{dv'}{v'}\, c(v, v')Q(v', u, t) \\ &= \tfrac{1}{2}c(v, -u) + \tfrac{1}{4}\int_0^1 \frac{dv'}{v'}\, c(v, v')S(v', u, t) \\ &= J(t, v, t, u). \end{aligned} \tag{21}$$

In addition, differentiation of both sides of Eq. (20) with respect to x shows that

$$\begin{aligned} M_x(t, v, x, u) &= -\frac{1}{2u}\, c(v, -u)e^{-(x-t)/u} \\ &\quad + \int_0^1 \frac{dv'}{v'}\, c(v, -v')e^{-(x-t)/v'} J(x, -v', x, u)\, dy \\ &\quad + \int_t^x \int_0^1 \frac{dv'}{v'}\, c(v, -v')e^{-(y-t)/v'}\, dy \\ &\quad \times \left(-\frac{1}{u} J(y, -v', x, u) \right. \\ &\quad \left. + 2\int_0^1 J(x, -v'', x, u)J(y, -v', x, v'')\, \frac{dv''}{v''}\right) \\ &\quad + \int_0^t \int_0^1 \frac{dv'}{v'}\, c(v, v')e^{-(t-y)/v'}\, dy \\ &\quad \times \left(-\frac{1}{u} J(y, v', x, u) \right. \\ &\quad \left. + 2\int_0^1 J(x, -v'', x, u)J(y, v', x, v'')\, \frac{dv''}{v''}\right). \end{aligned} \tag{22}$$

By collecting terms, this equation becomes

$$\begin{aligned} M_x(t, v, x, u) &= -\frac{1}{u} M(t, v, x, u) \\ &\quad + 2\int_0^1 \frac{dv'}{v'} J(x, -v', x, u)M(t, v, x, v'), \quad x \geq t. \end{aligned} \tag{23}$$

From our uniqueness assumption, it follows that

$$M(t, v, x, u) \equiv J(t, v, x, u), \quad x \geq t, \tag{24}$$

which is precisely the family of integral equations (1).

V. NUMERICAL RESULTS

The general scheme for solving the Cauchy system for the functions S and J numerically is to approximate the integrals occurring by the use of Gaussian quadrature formulas, which transforms the differential–integral equations into a system of ordinary differential equations.[5,6]

In the event that

$$c(v, u) = \text{const} = \tfrac{1}{2}\lambda, \tag{25}$$

the integral equation (1) becomes

$$J(t, x, u) = \tfrac{1}{4}\lambda e^{-(x-t)/u} + \tfrac{1}{2}\lambda \int_0^x E_1(|t - y|)J(y, x, u)\, dy, \qquad 0 \leq t \leq x,\ 0 \leq u \leq 1, \tag{26}$$

where, as usual,

$$E_1(|t - y|) = \int_0^1 e^{-|t-y|/z} \frac{dz}{z}. \tag{27}$$

In addition, the Cauchy system for the function S becomes

$$\begin{aligned} S_x(v, u, x) &= -\left(\frac{1}{u} + \frac{1}{v}\right) S(v, u, x) \\ &\quad + \tfrac{1}{2}\lambda \int_0^1 S(v', u, x)\, \frac{dv'}{v'} \\ &\quad + \tfrac{1}{2}\lambda \int_0^1 S(v, v', x)\, \frac{dv'}{v'} \\ &\quad + \tfrac{1}{4}\lambda \int_0^1 \int_0^1 S(v, v', x)S(v'', u, x)\, \frac{dv''}{v''}\, \frac{dv'}{v'}. \end{aligned} \tag{28}$$

This is a well-known result for the case of isotropic scattering. Numerical results for S and J, based on the numerical solution of the Cauchy system, using the method of lines, are available in Refs. 5 and 7.

To test the effects of stronger and stronger forward scattering, we selected a phase function of the form[3]

$$p(\cos\theta) = k(b - \cos\theta)^{-1}, \quad b > 1, \tag{29}$$

which corresponds to

$$c(v, u) = (\lambda k)2[(b - uv)^2 - (1 - u^2)(1 - v^2)]^{-\frac{1}{2}}. \tag{30}$$

The constant k fulfills the normalization condition

$$k = 2[\log(b + 1)/(b - 1)]^{-1}. \tag{31}$$

The closer b is to unity, the stronger the forward scattering is.

We found that, with $b = 1.1$ and $\lambda = 1.0$, use of a Gaussian quadrature formula of order 7 and an integration step size of 0.01 in an Adams–Moulton fourth-order integration routine resulted in about five accurate digits in the evaluation of the function S for $0 \leq x \leq 1$, $0 \leq v, u \leq 1$. An IBM 7044 computer was used. When $b = 1.01$, about three accurate figures were obtained. When $b = 1.001$, we used a Gaussian quadrature formula of order 15 and a step size of 0.001 and obtained about one accurate figure. The front-to-back ratio passes from about 20 to 2000 in these cases. This shows the complete feasibility of the method from the numerical viewpoint, even for strongly peaked phase functions.

VI. DISCUSSION

The method presented, when augmented by quasilinearization,[8] can be used to solve inverse problems. In this class of problems, we measure diffusely reflected radiation and wish to infer local scattering properties of the medium.

Extensions to cases involving internal sources of radiation and reflecting surfaces are readily made.

* Supported by the National Science Foundation under Grant No. GF-294, and the National Institutes of Health under Grants No. GM-16197-01 and GM-16437-01.

† Consultant, Rand Corp., Santa Monica, Calif. 90406.

[1] V. V. Sobolev, Dokl. Akad. Nauk SSSR **179**, 41 (1968) [Sov. Phys.—Dokl. **13**, 180 (1968)].
[2] A. Leonard and T. Mullikin, J. Math. Phys. **5**, 399 (1964).
[3] H. Kagiwada and R. Kalaba, The Rand Corp., Santa Monica, Calif., RM-5245-PR, 1967.
[4] H. Kagiwada and R. Kalaba, J. Math. Anal. Appl. **23**, 540 (1968).
[5] R. Bellman, H. Kagiwada, R. Kalaba, and S. Ueno, J. Math. Anal. Appl. **12**, 541 (1965).
[6] H. Kagiwada, R. Kalaba, and A. Schumitzky, Calcolo **5**, 63 (1968).
[7] R. Bellman, H. Kagiwada, and R. Kalaba, J. Quant. Spectry. Radiative Transfer **6**, 291 (1966).
[8] R. Bellman and R. Kalaba, *Quasilinearization and Nonlinear Boundary Value Problems* (American Elsevier, New York, 1965).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 5 MAY 1970

Groups of Curvature Collineations in Riemannian Space-Times Which Admit Fields of Parallel Vectors

GERALD H. KATZIN, JACK LEVINE,* AND WILLIAM R. DAVIS
Departments of Physics and Mathematics, North Carolina State University*
Raleigh, North Carolina

(Received 22 July 1969; Revised Manuscript Received 8 December 1969)

By definition, a Riemannian space V_n admits a symmetry called a curvature collineation (CC) if the Lie derivative with respect to some vector ξ^i of the Riemann curvature tensor vanishes. It is shown that if a V_n admits a parallel vector field, then it will admit groups of CC's. It follows that every space–time with an expansion-free, shear-free, rotation-free, geodesic congruence admits groups of CC's, and hence gravitational pp waves admit such groups of symmetries.

1. INTRODUCTION

A Riemannian space V_n with curvature tensor[1] R^i_{jkm} is said to admit a symmetry called a curvature collineation[2] (CC) if $\pounds_\xi R^i_{jkm} = 0$ for some vector field[3] ξ^i. In this paper, we show that any V_n which admits a field of parallel vectors also admits groups of CC's defined by ξ^i related to the parallel field vectors in a simple manner. It follows that every space–time which admits a shear-free, expansion-free, rotation-free, geodesic congruence admits groups of CC's. As a consequence, all plane-fronted pure gravitational waves ($R_{ij} = 0$) with parallel ray vectors (pp waves) admit groups of CC's.

It can be shown (see Ref. 2) that a necessary and

To test the effects of stronger and stronger forward scattering, we selected a phase function of the form[3]

$$p(\cos\theta) = k(b - \cos\theta)^{-1}, \quad b > 1, \tag{29}$$

which corresponds to

$$c(v, u) = (\lambda k)2[(b - uv)^2 - (1 - u^2)(1 - v^2)]^{-\frac{1}{2}}. \tag{30}$$

The constant k fulfills the normalization condition

$$k = 2[\log(b + 1)/(b - 1)]^{-1}. \tag{31}$$

The closer b is to unity, the stronger the forward scattering is.

We found that, with $b = 1.1$ and $\lambda = 1.0$, use of a Gaussian quadrature formula of order 7 and an integration step size of 0.01 in an Adams–Moulton fourth-order integration routine resulted in about five accurate digits in the evaluation of the function S for $0 \leq x \leq 1$, $0 \leq v, u \leq 1$. An IBM 7044 computer was used. When $b = 1.01$, about three accurate figures were obtained. When $b = 1.001$, we used a Gaussian quadrature formula of order 15 and a step size of 0.001 and obtained about one accurate figure. The front-to-back ratio passes from about 20 to 2000 in these cases. This shows the complete feasibility of the method from the numerical viewpoint, even for strongly peaked phase functions.

VI. DISCUSSION

The method presented, when augmented by quasilinearization,[8] can be used to solve inverse problems. In this class of problems, we measure diffusely reflected radiation and wish to infer local scattering properties of the medium.

Extensions to cases involving internal sources of radiation and reflecting surfaces are readily made.

* Supported by the National Science Foundation under Grant No. GF-294, and the National Institutes of Health under Grants No. GM-16197-01 and GM-16437-01.

† Consultant, Rand Corp., Santa Monica, Calif. 90406.

[1] V. V. Sobolev, Dokl. Akad. Nauk SSSR **179**, 41 (1968) [Sov. Phys.—Dokl. **13**, 180 (1968)].
[2] A. Leonard and T. Mullikin, J. Math. Phys. **5**, 399 (1964).
[3] H. Kagiwada and R. Kalaba, The Rand Corp., Santa Monica, Calif., RM-5245-PR, 1967.
[4] H. Kagiwada and R. Kalaba, J. Math. Anal. Appl. **23**, 540 (1968).
[5] R. Bellman, H. Kagiwada, R. Kalaba, and S. Ueno, J. Math. Anal. Appl. **12**, 541 (1965).
[6] H. Kagiwada, R. Kalaba, and A. Schumitzky, Calcolo **5**, 63 (1968).
[7] R. Bellman, H. Kagiwada, and R. Kalaba, J. Quant. Spectry. Radiative Transfer **6**, 291 (1966).
[8] R. Bellman and R. Kalaba, *Quasilinearization and Nonlinear Boundary Value Problems* (American Elsevier, New York, 1965).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 5 MAY 1970

Groups of Curvature Collineations in Riemannian Space-Times Which Admit Fields of Parallel Vectors

GERALD H. KATZIN, JACK LEVINE,* AND WILLIAM R. DAVIS
Departments of Physics and Mathematics, North Carolina State University*
Raleigh, North Carolina

(Received 22 July 1969; Revised Manuscript Received 8 December 1969)

By definition, a Riemannian space V_n admits a symmetry called a curvature collineation (CC) if the Lie derivative with respect to some vector ξ^i of the Riemann curvature tensor vanishes. It is shown that if a V_n admits a parallel vector field, then it will admit groups of CC's. It follows that every space–time with an expansion-free, shear-free, rotation-free, geodesic congruence admits groups of CC's, and hence gravitational pp waves admit such groups of symmetries.

1. INTRODUCTION

A Riemannian space V_n with curvature tensor[1] R^i_{jkm} is said to admit a symmetry called a curvature collineation[2] (CC) if $\pounds_\xi R^i_{jkm} = 0$ for some vector field[3] ξ^i. In this paper, we show that any V_n which admits a field of parallel vectors also admits groups of CC's defined by ξ^i related to the parallel field vectors in a simple manner. It follows that every space–time which admits a shear-free, expansion-free, rotation-free, geodesic congruence admits groups of CC's. As a consequence, all plane-fronted pure gravitational waves ($R_{ij} = 0$) with parallel ray vectors (pp waves) admit groups of CC's.

It can be shown (see Ref. 2) that a necessary and

sufficient condition for a V_n to admit a CC may be expressed in the form[4]

$$(h_{im;j} + h_{mj;i} - h_{ij;m})_{;k} - (h_{km;j} + h_{mj;k} - h_{kj;m})_{;i} = 0, \quad (1.1)$$

where

$$h_{ij} \equiv \pounds_\xi g_{ij} = \xi_{i;j} + \xi_{j;i}, \quad (1.1')$$

with g_{ij} the components of the metric tensor of the V_n. Subcases of interest for this paper are (see Ref. 2) (a) special curvature collineations (SCC's), (b) affine collineations (AC's), and (c) motions (M's), defined by

$$\text{SCC:} \quad h_{ij;km} = 0, \quad (1.2a)$$

$$\text{AC:} \quad h_{ij;k} = 0, \quad (1.2b)$$

$$\text{M:} \quad h_{ij} = 0, \quad (1.2c)$$

respectively.

For purposes of this paper, we formulate the following definitions: A CC is a "proper" CC (Prop CC) if it is not an SCC. An SCC is a proper SCC (Prop SCC) if it is not an AC. An AC is a proper AC (Prop AC) if it is not an M. The terminology "improper CC" refers to any SCC (proper or not proper).

2. GROUPS OF CC's IN A V_n

Assume that a V_n admits a parallel vector field λ_i ($\lambda_{i;j} = 0$), $\lambda_i \equiv S_{,i}$, where S is a nonconstant scalar.[5] Let $F(S)$ be an arbitrary function of S which is four-times differentiable.[6] If[7]

$$\xi_i \equiv F_{,i} = F'\lambda_i, \quad (2.1)$$

then we find from the definition (1.1) of h_{ij} that

$$h_{ij} = 2F''\lambda_i\lambda_j, \quad (2.2a)$$

$$h_{ij;k} = 2F'''\lambda_i\lambda_j\lambda_k, \quad (2.2b)$$

$$h_{ij;km} = 2F''''\lambda_i\lambda_j\lambda_k\lambda_m. \quad (2.2c)$$

It immediately follows from (1.1) and (2.2c) that ξ_i of (2.1) in general defines a CC and that ξ_i defines a normal congruence of curves.

It follows from (2.1) and (2.2) that ξ_i defines an improper CC, which is a Prop SCC, Prop AC, or M, if $F(S)$ is a cubic, quadratic, or linear polynomial in S, respectively, and defines a Prop CC if $F'''' \neq 0$.

We now consider group properties of improper CC's. If we define the generators[8] $X_{(\alpha)} \equiv \xi^{(\alpha)i}\partial/\partial x^i$, where $\xi^{(\alpha)i} \equiv S^\alpha\lambda^i$, $\alpha = 0, 1, 2$,[9] we then find that $[X_0, X_1] = \Lambda X_0$, $[X_0, X_2] = 2\Lambda X_1$, $[X_1, X_2] = \Lambda X_2$, where $\Lambda \equiv \lambda_i\lambda^i = \text{const.}$

In the case of Prop CC's, $F'''' \neq 0$, we consider a set of r linearly independent functions $F_{(\alpha)}(S)$, $\alpha = 1, 2, \cdots, r$, and the corresponding generators

$$Y_\alpha \equiv \frac{\xi^{(\alpha)i}\partial}{\partial x^i} = \frac{F'_{(\alpha)}\lambda^i\partial}{\partial x^i}.$$

It is found that $[Y_\alpha, Y_\beta] = 0$ if $\Lambda = 0$. We summarize the above results in the following theorem.

Theorem 2.1: A V_n which admits a parallel vector field $\lambda_i \equiv S_{,i}$ admits (a) a Prop SCC: $\xi^{(2)i} \equiv S^2\lambda^i$, (b) a Prop AC: $\xi^{(1)i} \equiv S\lambda^i$, and (c) an M: $\xi^{(0)i} \equiv \lambda^i$. The generators $X_\alpha \equiv \xi^{(\alpha)i}\partial/\partial x^i$ define a 3-parameter group of Prop SCC's, which is Abelian if λ^i is a null vector, $\Lambda = 0$. The V_n also admits a Prop CC defined by the vector $\xi^i \equiv F'(S)\lambda_i$, $F'''' \neq 0$. In case $\Lambda = 0$, the V_n admits an r-parameter Abelian group of Prop CC's defined by r vectors $\xi^{(\alpha)i} \equiv F_{(\alpha)}\lambda^i$, where $F_{(\alpha)}(S)$, $F''''_{(\alpha)}(S) \neq 0$, $\alpha = 1, 2, \cdots, r$, is any given set of r linearly independent but otherwise arbitrary scalar functions for any $r = 1, 2, \cdots$.

3. GROUPS OF SYMMETRIES CONCOMITANT WITH PROP SCC

Without the *a priori* assumption of the existence of a parallel vector field λ^i, let us assume that a V_n admits a Prop SCC defined by a vector η^i. It follows (see Ref. 2, Corollary 5.1) that the V_n then admits, in general,[10] a nontrivial field of parallel vectors $N_k \equiv M_{,k}$, where $M \equiv g^{ij}(\eta_{i;j} + \eta_{j;i}) = 2\eta^i_{;i}$. The existence of this parallel field N_k thus implies by use of Theorem 2.1 the existence of additional CC's (proper or improper) in the V_n. In particular, if we choose for one of these additional CC's a Prop SCC defined by vector $\xi^{(2)i} \equiv M^2N^i$, it can be shown that, in general, $\xi^{(2)i}$ and η^i are linearly independent.

Theorem 3.1: A V_n which admits a Prop SCC does, in general, admit a nontrivial parallel field and, hence, additional CC's as determined by Theorem 2.1.

On the other hand, if we assume that a V_n admits a Prop SCC based upon a parallel field λ^i as described in Theorem 2.1, then the parallel field N^i (as described above) can easily be shown to be linearly dependent upon λ^i; hence, no additional CC's can be generated by reuse of Theorem 2.1 with reference to vector N^i.

4. APPLICATIONS IN SPACE-TIME

We now consider a space–time which admits a geodesic congruence which we take to be defined by the vector field u^i, $u^i_{;j}u^j = 0$. It can be shown that we may write

$$u^{(\alpha)}_{i;j} = \sigma^{(\alpha)}_{ij} + \omega^{(\alpha)}_{ij} + \theta^{(\alpha)}P^{(\alpha)}_{ij},$$

$$u^{(\alpha)}_i u^{(\alpha)i} = \delta^\alpha_1, \quad \alpha = 0, 1,$$

where $\sigma^{(\alpha)}_{ij}$, $\omega^{(\alpha)}_{ij}$, and $\theta^{(\alpha)}$ are the shear, rotation, and expansion, respectively, and where $P^{(\alpha)}_{ij}$ is proportional

to the appropriate projection operator corresponding to whether u^i is null or nonnull.[11] From the above equations, it immediately follows that if a space–time admits a shear-free, rotation-free, expansion-free, geodesic congruence (either null or nonnull), then it admits a parallel vector field. Hence, we have the next theorem.

Theorem 4.1: Every space–time which admits an expansion-free, shear-free, rotation-free, geodesic congruence of curves defined with respect to the vector field u^i admits a group of CC's as described by Theorem 2.1, wherein the parallel vector field λ^i is taken to be u^i.[12]

Since plane-fronted pure gravitational waves,[13] $R_{ij} = 0$, with parallel (null) ray vector k_i, $k_i k^i = 0$, $k_{i;j} = 0$, are characterized by an expansion-free, rotation-free, shear-free, null geodesic congruence (which is essentially defined by the ray vector k_i), we may state our last theorem.

Corollary 4.1: Every plane-fronted pure gravitational ($R_{ij} = 0$) wave with a null parallel ray vector k_i (pp wave) admits a group of CC's as described by Theorem 2.1, wherein the parallel vector field λ^i is taken to be k^i.[14]

It is of interest to note that a subclass of the Lichnerowicz radiation space–times[15] also admits groups of CC's as a consequence of Theorem 2.1. These space–times are considered in detail by Levine and Zund.[16]

ACKNOWLEDGMENT

The authors wish to thank Dr. J. W. York, Jr., for his helpful comments.

[1] Indices take on values $1, \cdots, n$. The Einstein summation convention is employed.

[2] G. H. Katzin, J. Levine, and W. R. Davis, J. Math. Phys. **10**, 617 (1969).

[3] $\pounds_\xi$ denotes the Lie derivative with respect to the vector ξ^i. For a discussion of the theory of Lie derivatives, refer to K. Yano, *The Theory of Lie Derivatives and its Applications* (North-Holland, Amsterdam, 1957).

[4] Covariant differentiation is indicated by a semicolon (;) and partial differentiation by a comma (,).

[5] A canonical form for the metrics of those Riemannian spaces which admit fields of parallel vectors is well known. Refer to L. P. Eisenhart, Ann. Math. (2) **9**, 762 (1938).

[6] This notation is used throughout the paper with this understanding.

[7] The prime (′) indicates differentiation with respect to S, i.e., d/dS.

[8] L. P. Eisenhart, *Continuous Groups of Transformations* (Dover, New York, 1961).

[9] The notation S^α indicates S to the αth power.

[10] Obviously, the parallel field N_k is trivial if $\eta^i_{[;ik]} = 0$.

[11] For a signature of -2, the projection operator for the case $u_i u^i = 1$ is given by ${}_{(3)}h_{ij} = g_{ij} - u_i u_j = 3P^{(1)}_{ij}$ [see, for example, J. L. Synge, *Relativity: The General Theory* (North-Holland, Amsterdam, 1960), p. 172]. In the null case, $u_i u^i = 0$, the appropriate projection operator takes the form ${}_{(3)}h_{ij} = g_{ij} + u_i u_j/(u_k v^k)^2 - (u_i u_j + u_j v_i)/u_k v^k = 2P^{(0)}_{ij}$, where v_j is any timelike vector field satisfying $v^k_{;j} u^j = 0$ [see, for example, P. Jordan, J. Ehlers, and R. Sachs, Akad. Wiss. Lit. (Mainz) Abhandl. Math. Nat. Kl. **1**, 12 (1961) or R. Sachs, in *Recent Developments in General Relativity* (Macmillan, New York, 1962), pp. 395–407].

[12] The Einstein static cosmological space–time with line element

$$ds^2 = -(1 + \tfrac{1}{4}K_0 r^2)^{-2}[(dx^1)^2 + (dx^2)^2 + (dx^3)^2] + (dx^4)^2,$$

where $r^2 \equiv (x^1)^2 + (x^2)^2 + (x^3)^2$, is an example of a space–time with $\theta^{(1)} = 0$, $\sigma^{(1)}_{ij} = 0$, $\omega^{(1)}_{ij} = 0$, which is known to admit CC's. See G. H. Katzin, J. Levine, and W. R. Davis, Tensor, New Ser., **21**, 52 (1970).

[13] J. Ehlers and W. Kundt, in *Gravitation: An Introduction to Current Research*, L. Witten, Ed. (Wiley, New York, 1962). Also see F. A. E. Pirani, in *Lectures on General Relativity* (*Brandeis, Summer Institute, 1964*) (Prentice-Hall, Englewood Cliffs, N.J., 1965), p. 354.

[14] It can be shown that the electromagnetic-field tensor F_{ij} of a plane electromagnetic wave in Minkowski space–time is invariant in the sense that $\pounds_\xi F_{ij} = 0$, where $\xi^i \equiv F'(S)k^i$ with k^i the propagation vector of the wave, so that $k^i_{,j} = 0$ and $k_i \equiv S_{,i}$. Hence, the groups of CC's admitted by a pp gravitational wave which are based upon the existence of a parallel vector field in a nonflat Riemannian space–time have this type analog in a pp electromagnetic wave in Minkowski space–time.

[15] A. Lichnerowicz, Ann. Mat. Pura Appl. **50**, 1 (1960).

[16] J. Levine and J. D. Zund, Ann. Mat. Pura Appl. (IV) **80**, 373 (1968).

$\overline{E2}$-Parametrization of $SL(2, C)$*

G. J. IVERSON AND G. MACK
International Centre for Theoretical Physics, Trieste, Italy, and Center for Theoretical Studies, University of Miami, Coral Gables, Florida

(Received 10 July 1969)

We consider the $\overline{E2}$-parametrization of unimodular 2×2 matrices $A \in SL(2, C)$, which is of the form

$$A = E_1 e^{-\frac{1}{2}a\sigma_3} V E_2, \quad \text{with} \quad V = e^{\frac{1}{2}i\pi\sigma_2} \quad \text{and} \quad E = \begin{pmatrix} e^{-\frac{1}{2}i\phi} & 0 \\ z & e^{\frac{1}{2}i\phi} \end{pmatrix} \in \overline{E2}.$$

$\overline{E2}$ is a covering of the group of Euclidean motions in the plane. We compute the correspondingly factorized matrix elements of the unitary representations of $SL(2, C)$ in an $\overline{E2}$-basis; the result is given in Eq. (6). As a fringe benefit we obtain an integral transform which amounts to expansion in terms of Meijer G-functions and which generalizes the familiar Hankel transform. The results of this paper are useful, e.g., for computing vertex functions in the theory of massless particles with continuous spin.

1. INTRODUCTION AND RESULTS

Matrix elements of unitary irreducible representations of $SL(2, C)$ in an $\overline{E2}$-basis have recently been studied by Chang and O'Raifeartaigh.[1] They used the canonical parametrization of unimodular 2×2 matrices $A \in SL(2, C)$:

$$A = U_1 e^{-\frac{1}{2}a\sigma_3} U_2, \quad U_{1,2} \in SU(2).$$

It is the purpose of the present paper to give the appropriately factorized $SL(2, C)$ matrix elements in the $\overline{E2}$-basis for an alternative parametrization of A. Every matrix $A \in SL(2, C)$ with $A_{12} \neq 0$ may be written as

$$A = E_1 e^{-\frac{1}{2}a\sigma_3} V E_2, \quad \text{where} \quad V = e^{\frac{1}{2}i\pi\sigma_2} = i\sigma_2, \quad -\infty < a < +\infty, \quad (1)$$

and

$$E = \begin{pmatrix} e^{-\frac{1}{2}i\varphi} & 0 \\ 0 & e^{\frac{1}{2}i\varphi} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \xi & 1 \end{pmatrix} \begin{pmatrix} e^{-\frac{1}{2}i\psi} & 0 \\ 0 & e^{\frac{1}{2}i\psi} \end{pmatrix} = \begin{pmatrix} e^{-\frac{1}{2}i(\varphi+\psi)} & 0 \\ \xi e^{\frac{1}{2}i(\varphi-\psi)} & e^{\frac{1}{2}i(\varphi+\psi)} \end{pmatrix},$$

$$0 \leq \varphi + \psi < 4\pi, \quad 0 \leq \varphi - \psi < 4\pi, \quad 0 < \xi < +\infty. \quad (2)$$

The matrices E form a covering $\overline{E2}$ of the group $E2$ of Euclidean motions in the plane (i.e., rotations and translations). For E_2 we may, without loss of generality, take $\varphi_2 = 0$, so that A is specified by the six real parameters φ_1, ξ_1, ψ_1, a, ξ_2, and ψ_2. A parametrization related to (1) and (2) has been used in partial-wave analysis.[2] It is also of importance in the theory of massless particles with continuous spin.[3]

There are two series of unitary irreducible representations of $SL(2, C)$ labeled by a pair $\chi = [l_0, c]$:

principal series: l_0 integer or half-integer, c pure imaginary; (3a)

supplementary series: $l_0 = 0$, $0 < c < 1$. (3b)

Their decomposition in terms of (a direct integral of) unitary irreducible representations of $\overline{E2}$ is known[1,4]: Each faithful unitary $\overline{E2}$ representation occurs with multiplicity one, and no others occur (with nonzero measure). The faithful unitary representations of $\overline{E2}$ are labeled by a real number $\rho > 0$. A corresponding (improper) basis of the $SL(2, C)$ representation space may be labeled by $|\rho, \lambda\rangle$:

$$(M_3 - \lambda)\,|\rho, \lambda\rangle = 0, \quad (\pi_1^2 + \pi_2^2 - \rho^2)\,|\rho, \lambda\rangle = 0, \quad (4)$$

where

$$\lambda = l_0, l_0 \pm 1, \cdots,$$

i.e., λ assumes either all integer or all half-odd integer values. $\pi_1 = M_1 - N_2$, $\pi_2 = M_2 + N_1$, and M_3 are the generators of $\overline{E2}$; M_k, and N_k are the generators of rotations and Lorentz boosts, respectively. The $\overline{E2}$ Casimir operator is $\pi_1^2 + \pi_2^2$.

We adopt the normalization

$$\langle \rho', \lambda' \,|\, \rho, \lambda \rangle = (\rho)^{-1}\delta(\rho' - \rho)\delta_{\lambda\lambda'}. \quad (5)$$

The matrix elements of the unitary irreducible representation $T_\chi(A)$ may then be written as

$$\langle \rho\lambda |\, T_\chi(A)\,|\rho'\lambda'\rangle = \sum_\kappa d^\rho_{\lambda\kappa}(E_1) r^\chi_{\rho\rho'\kappa}(a) d^{\rho'}_{-\kappa\lambda'}(E_2), \quad \kappa, \lambda, \lambda' = l_0, l_0 \pm 1, \cdots, \quad (6a)$$

with

$$d^\rho_{\lambda\kappa}(E) \equiv \langle \rho\lambda |\, T_\chi(E)\,|\rho\kappa\rangle = e^{-i\lambda\varphi} J_{\lambda-\kappa}(\rho\xi) e^{-i\kappa\psi} \quad (6b)$$

and

$$r^{\chi}_{\rho\rho',\kappa}(a) = 2\Delta^{-1}(\rho/\rho')^{-i\,\mathrm{Im}\,c} \times G^{20}_{04}((\rho\rho'/\Delta)^2\,|\tfrac{1}{2}(l_0-\kappa-c), \tfrac{1}{2}(-l_0-\kappa+c), \tfrac{1}{2}(-l_0+\kappa-c), \tfrac{1}{2}(l_0+\kappa+c)), \quad \Delta \equiv 4e^{-a}. \tag{6c}$$

Here G^{20}_{04} is the familiar Meijer G-function.[5] Expressions (6) are valid both for the principal and the supplementary series. Equation (6c) simplifies[5] for the special case of the Majorana representations to

$$r^{\chi}_{\rho\rho'\kappa}(a) = 2(\rho\rho'\Delta)^{-\frac{1}{2}}J_{-2\kappa}[4(\rho\rho'/\Delta)^{\frac{1}{2}}] \tag{6d}$$

for

$$\chi = [\tfrac{1}{2}, 0], \quad \kappa = \pm\tfrac{1}{2}, \pm\tfrac{3}{2}, \cdots,$$

or

$$\chi = [0, \tfrac{1}{2}], \quad \kappa = 0, \pm 1, \pm 2, \cdots.$$

The explicit expression (6) could, e.g., be used in harmonic analysis on the Lorentz group and is also useful in the theory of massless particles with continuous spin.[3] Besides, we obtain, from the group representation law, inversion formulas for an integral transform which provides expansions in terms of Meijer G-functions (Sec. 3).

2. PROOFS

A. $\overline{E2}$-Parametrization

Every unimodular 2×2 matrix $A' \in SL(2, C)$ with $A'_{22} \neq 0$ may be decomposed as follows[6]:

$$A' = kz \quad \text{with} \quad k = \begin{pmatrix} \mu & \lambda \\ 0 & \mu^{-1} \end{pmatrix}, \quad z = \begin{pmatrix} 1 & 0 \\ w & 1 \end{pmatrix}, \quad w, \mu, \lambda \in \mathbb{C}.$$

The matrix $z \in \overline{E2}$ already, and every k may be written in the form

$$k = V^{-1}Ee^{-\frac{1}{2}a\sigma_3}V, \quad \text{with} \quad E = \begin{pmatrix} e^{-\frac{1}{2}i\phi} & 0 \\ w' & e^{\frac{1}{2}i\phi} \end{pmatrix} \in \overline{E2}.$$

This is achieved by taking the parameters w', a, and ϕ as

$$\phi = 2\arg\mu, \quad a = 2\ln|\mu|, \quad \text{and} \quad w' = -e^{\frac{1}{2}a}\lambda.$$

Application of this to $A' = V^{-1}A$ gives the desired result (2).

B. Matrix Elements

We use Gel'fand's description[7] of (unitary) irreducible representations of $SL(2, C)$ in terms of homogeneous functions of a complex 2-spinor-variable $\zeta = (\zeta_1, \zeta_2) \neq (0, 0)$. Homogeneity reads

$$f(\mu\zeta) = \mu^{l_0+c-1}\bar{\mu}^{-l_0+c-1}f(\zeta), \quad \text{for} \quad \mu \in \mathbb{C}, \tag{7}$$

and the transformation law is

$$[T_{\chi}(A)f](\zeta) = f(\zeta A). \tag{8}$$

The corresponding generators are

$$M_k = \tfrac{1}{2}\left(\zeta\sigma_k\frac{\partial}{\partial\zeta} - \frac{\partial}{\partial\zeta^+}\sigma_k\zeta^+\right),$$
$$N_k = \tfrac{1}{2}i\left(\zeta\sigma_k\frac{\partial}{\partial\zeta} + \frac{\partial}{\partial\zeta^+}\sigma_k\zeta^+\right), \tag{9}$$

where the σ_k are Pauli matrices.

The representations of the principal and supplementary series are unitary by virtue of the following invariant scalar products: for the principal series,

$$(f, g) = \int d\mu(\zeta)\bar{f}(\zeta)g(\zeta) \tag{10a}$$

and, for the supplementary series,

$$(f, g) = \int d\mu(\zeta)\int d\mu(\omega)\bar{f}(\omega)K_{\chi}(\omega, \zeta)g(\zeta). \tag{10b}$$

K_{χ} is Gel'fand's intertwining operator[7] and is given by

$$K_{\chi}(\zeta, \omega) = N^{-1}(\omega\epsilon\zeta)^{-l_0-c-1}\overline{(\omega\epsilon\zeta)}^{l_0-c-1}. \tag{11}$$

The normalization factor N will be fixed later. ϵ is the 2-dimensional antisymmetric tensor

$$\omega\epsilon\zeta = \omega_1\zeta_2 - \omega_2\zeta_1.$$

Note that $\omega\epsilon\zeta$ is an $SL(2, C)$ invariant on account of $\epsilon A^{-1}\epsilon^{-1} = A^{\mathrm{T}}$, with T standing for the transpose. Finally, the measure is given by

$$d\mu(\zeta) = (4\pi^2)^{-1}(\tfrac{1}{2}i)^2\, d^2\zeta\, d^2\bar{\zeta}\delta(\tfrac{1}{2}\zeta\underset{\sim}{q}\zeta^+ - 1), \tag{12}$$

where

$$\underset{\sim}{q} = q^{\mu}\sigma_{\mu}, \quad q \text{ positive lightlike.}$$

σ_0 is the 2×2 unit matrix. The integrals (10) are well known to be independent of the positive lightlike vector q in the definition of the measure. This is a consequence of the homogeneity condition (7) (see Appendix) and ensures the Lorentz invariance of the scalar products (10). The Hilbert spaces on which the unitary representations act consist of homogeneous functions with finite norm (f, f) as given by Eq. (10) (cf. Ref. 7).

The explicit form of the basis vectors $|\rho, \lambda\rangle$ in this realization of the representation space are also known[4]:

$$|\rho, \lambda\rangle = f^{\rho\chi}_{\lambda}(\zeta), \quad \lambda = l_0, l_0 \pm 1, \cdots,$$
$$f^{\rho\chi}_{\lambda}(\zeta) = \rho^{-\mathrm{Re}\,c}\,|\zeta_2|^{2c-2}\,e^{i\lambda\phi}J_{l_0+\lambda}(\rho\,|\zeta_1/\zeta_2|)e^{il_0\alpha}, \tag{13}$$

where

$$\alpha + \phi = 2\arg\zeta_1, \quad \alpha - \phi = 2\arg\zeta_2 - 2\pi,$$

arg standing for the phase of a complex number as usual. These basis vectors satisfy Eqs. (4) and (5) and

transform correctly under the subgroup $\overline{E2}$:

$$(T_\chi(E)f^{\rho\chi}_\lambda)(\zeta) = f^{\rho\chi}_\lambda(\zeta E) = \sum_{\lambda'} d^\rho_{\lambda'\lambda}(E) f^{\rho\chi}_{\lambda'}(\zeta), \quad (14)$$

with d^ρ given by Eq. (6b). Furthermore, one finds the relation for suitable normalization N, in Eq. (11),

$$\int d\mu(\omega) K_\chi(\zeta, \omega) f^{\rho\chi}_\lambda(\omega) = \rho^{2i\,\mathrm{Im}\,c} f^{\rho-\chi}_\lambda(\zeta), \quad (15)$$

where $-\chi = [-l_0, -c]$. This is essentially a consequence of the Lorentz invariance of the intertwining operator K_χ, which implies that both sides of Eq. (15) obey the same differential equations (4). Equation (15) may also be checked by explicit computation; in this way one finds

$$N = 2^{-2c+2}\Gamma(|l_0| - c)/\Gamma(|l_0| + c + 1).$$

We are now ready to compute the matrix elements of $T^\chi(A)$. We obtain from Eqs. (8), (10), and (13), for the principal series,

$$\langle \rho\lambda |\, T_\chi(A)\, | \rho'\lambda' \rangle = \int d\mu(\zeta) \bar{f}^{\rho\chi}_\lambda(\zeta) f^{\rho'\chi}_{\lambda'}(\zeta A),$$

and, for the supplementary series,

$$\begin{aligned}\langle \rho\lambda |\, T_\chi(A)\, | \rho'\lambda' \rangle &= \int d\mu(\zeta) \int d\mu(\omega)\, \bar{K}_\chi(\zeta, \omega) \bar{f}^{\rho\chi}_\lambda(\omega) f^{\rho'\chi}_{\lambda'}(\zeta A) \\ &= \int d\mu(\zeta) \bar{f}^{\rho-\chi}_\lambda(\zeta) f^{\rho'\chi}_{\lambda'}(\zeta A).\end{aligned}$$

Here we used Eq. (15) to obtain the last equation. Introducing the parametrization (1) and using (14) and the Lorentz invariance of the scalar product, we obtain

$$\begin{aligned}\langle \rho\lambda |\, T_\chi(A)\, | \rho'\lambda' \rangle &= \sum_{\kappa\kappa'} \langle \rho\lambda |\, T_\chi(E_1)\, | \rho\kappa \rangle \langle \rho\kappa |\, T_\chi(e^{-\frac{1}{2}a\sigma_3} V)\, | \rho'\kappa' \rangle \\ &\quad \times \langle \rho'\kappa' |\, T_\chi(E_2)\, | \rho'\lambda' \rangle \\ &= \sum_{\kappa\kappa'} d^\rho_{\lambda\kappa}(E_1) \check{r}^\chi_{\rho\kappa,\rho'\kappa'}(a)\, d^{\rho'}_{\kappa'\lambda'}(E_2),\end{aligned}$$

with

$$\check{r}^\chi_{\rho\kappa,\rho'\kappa'}(a) = \int d\mu(\zeta) \bar{f}^{\rho,\pm\chi}_\kappa(\zeta) \bar{f}^{\rho'\chi}_{\kappa'}(\zeta e^{-\frac{1}{2}a\sigma_3} V),$$

and $f(\zeta)$ as given by Eq. (13). $\pm\chi$ is $+\chi$ for the principal and $-\chi$ for the supplementary series. We introduce new parameters of integration

$$-\zeta_1/\zeta_2 = xe^{i\phi}, \quad \arg \zeta_2 = \beta,$$

so that

$$d\mu(\zeta) = (4\pi^2)^{-1} x\, dx\, d\phi\, d\beta|_{|\zeta_2|=1}.$$

The integrations over β, ϕ may be carried out trivially, resulting in a Kronecker $\delta_{-\kappa,\kappa'}$. The remaining integration may be performed with the help of the formula for the regularized integral[3,8]:

$$\begin{aligned}&\int_0^\infty dx x^{\sigma-1} J_m(ax) J_n\left(\frac{b}{x}\right) \\ &\quad = \tfrac{1}{2} a^{-\sigma/2} b^{\sigma/2} G^{20}_{04}((\tfrac{1}{4}ab)^2 \,|\, \tfrac{1}{4}(2n - \sigma), \tfrac{1}{4}(2m + \sigma), \\ &\qquad\qquad \tfrac{1}{4}(-2n - \sigma), \tfrac{1}{4}(-2m + \sigma)) \\ &\qquad m, n \text{ integer}, \quad \sigma \in \mathbb{C}.\end{aligned}$$

The result is as indicated in Eq. (6c).

3. MEIJER G-TRANSFORM

The group representation law implies

$$\begin{aligned}\sum_\kappa \int \sigma\, d\sigma\, \langle \rho\lambda |\, T_\chi(A)\, | \sigma\kappa \rangle \langle \sigma\kappa |\, T_\chi(A^{-1})\, | \rho'\lambda' \rangle \\ = \delta_{\lambda\lambda'} \rho^{-1} \delta(\rho' - \rho).\end{aligned}$$

Specializing to $A = V$ and inserting the expression (6) for the matrix elements of T^χ, we find the completeness relation

$$\begin{aligned}&(-1)^{2l_0} \tfrac{1}{4} \int_0^\infty \sigma\, d\sigma \\ &\quad \times G^{20}_{04}((\tfrac{1}{4}\rho\sigma)^2 \,|\, \tfrac{1}{2}(l_0 - \kappa - c), \tfrac{1}{2}(-l_0 - \kappa + c), \\ &\qquad\qquad \tfrac{1}{2}(-l_0 + \kappa - c), \tfrac{1}{2}(l_0 + \kappa + c)) \\ &\quad \times G^{20}_{04}((\tfrac{1}{4}\sigma\rho')^2 \,|\, \tfrac{1}{2}(l_0 + \kappa - c), \tfrac{1}{2}(-l_0 + \kappa + c), \\ &\qquad\qquad \tfrac{1}{2}(-l_0 - \kappa - c), \tfrac{1}{2}(l_0 - \kappa + c)) \\ &= \rho^{-1}\delta(\rho' - \rho). \qquad (16)\end{aligned}$$

This is true for any $\chi = [l_0, c]$ in either the principal or supplementary series and arbitrary κ such that $\kappa + l_0$ is integer.

Consider now a function $f(x)$ such that $\int d\rho f(\rho) |\rho\lambda\rangle$ is normalizable, i.e.,

$$\int_0^\infty x^{-1}\, dx\, |f(x)|^2 < \infty. \qquad (17)$$

Define

$$\begin{aligned}g(y) &= (-)^{2l_0} \tfrac{1}{2} \int_0^\infty x\, dx\, f(x) \\ &\quad \times G^{20}_{04}((\tfrac{1}{4}xy)^2 \,|\, \tfrac{1}{2}(l_0 - \kappa - c), \tfrac{1}{2}(-l_0 - \kappa + c), \\ &\qquad \tfrac{1}{2}(-l_0 + \kappa - c), \tfrac{1}{2}(l_0 + \kappa + c)). \qquad (18a)\end{aligned}$$

Then the following inversion formula holds:

$$\begin{aligned}f(x) &= \tfrac{1}{2} \int_0^\infty y\, dy\, g(y) \\ &\quad \times G^{20}_{04}((\tfrac{1}{4}xy)^2 \,|\, \tfrac{1}{2}(l_0 + \kappa - c), \tfrac{1}{2}(-l_0 + \kappa + c), \\ &\qquad \tfrac{1}{2}(-l_0 - \kappa - c), \tfrac{1}{2}(l_0 - \kappa + c)). \qquad (18b)\end{aligned}$$

Note that (17a) and (17b) are of the same form except for a change of sign of κ and the over-all factor $(-)^{2l_0}$.

Equations (18) provide a whole class of integral transforms labeled by (χ, κ), where $\chi \equiv [l_0, c]$ satisfies either (3a) or (3b) and κ is arbitrary integer, for l_0 integer, and arbitrary half-odd-integer otherwise. We propose the name "Meijer G-transform of order (χ, κ)." For the special cases $\chi = [\frac{1}{2}, 0]$ and $\chi = [0, \frac{1}{2}]$ it goes over into the familiar Hankel transform of order 2κ for the function $\check{f}(x) = xf(\frac{1}{2}x^2)$.

ACKNOWLEDGMENTS

The authors are grateful to Professor A. Salam and Professor P. Budini, and to the International Atomic Energy Agency for the hospitality kindly extended to them at the International Centre for Theoretical Physics, Trieste. We also wish to thank Professor B. Kurşunoğlu for his kind hospitality at the Center for Theoretical Studies, Coral Gables, where part of the work was done.

APPENDIX

For the convenience of the reader we give here a useful reformulation of results of Gel'fand on invariant bilinear forms.[7,9] Let $\zeta = (\zeta_1, \zeta_2) \neq (0, 0)$ be a complex 2-spinor-variable and assign to every 4-vector q^μ a matrix $q = q^\mu \sigma_\mu$. Here σ_κ are Pauli matrices and σ_0 is the 2×2 unit matrix. We then have the following:

Lemma: Suppose $h(\mu\zeta) = |\mu|^{-4}h(\zeta)$ for $\mu \in \mathbb{C}$. Then

$$\int_{\mathbb{C}^2} d^2\zeta \, d^2\bar{\zeta}\delta(\tfrac{1}{2}\zeta q \zeta^+ - 1)h(\zeta)$$

is independent of the positive lightlike vector q^μ.

* Work supported in part by the United States Air Force under grant No. AF AFOSR 1268-67.

[1] S.-J. Chang and L. O'Raifeartaigh, J. Math. Phys. **10**, 21 (1969).
[2] J. F. Boyce, R. Delbourgo, A. Salam, and J. Strathdee, "Partial-Wave Analysis" (Part I), ICTP, Trieste, Report IC/67/9 (unpublished).
[3] G. J. Iverson and G. Mack, "Quantum Fields and Interactions of Massless Particles: The Continuous Spin Case," Center for Theoretical Studies, Coral Gables, preprint CTS-T-Phys-69-2 [submitted to Ann. Phys. (N.Y.)].
[4] S. Ström, Arkiv Fysik **34**, 215 (1967).
[5] H. Bateman, *Higher Transcendental Functions* (McGraw-Hill, New York, 1953), Vol. I.
[6] M. A. Naimark, *Linear Representations of the Lorentz Group* (Pergamon, London, 1964).
[7] I. M. Gel'fand, M. I. Graev, and N. Ya. Valenkin, *Generalized Functions* (Academic, New York, 1966), Vol. 5.
[8] H. Bateman, *Tables of Integral Transforms* (McGraw-Hill, New York, 1954), Vol. 2, Formula 19.3.(45).
[9] V. Gorgé and H. Leutwyler, Helv. Phys. Acta **41**, 171 (1968).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 5 MAY 1970

Problem of the Disordered Chain

TA-FENG LIN
The Rockefeller University, New York, New York 10021

(Received 19 December 1969)

The equivalence of the methods of Dyson and Kac is proven in the 1-dimensional problem of the disordered chain, i.e., the determination of the spectrum of the frequencies of normal modes of a system of coupled oscillators, where the masses and/or spring constants are random variables.

INTRODUCTION

The problem of the disordered chain is to determine the spectrum of the frequencies of normal modes of a system of coupled oscillators when the masses and (or) spring constants are random variables. There exists an extensive literature, beginning with a pioneering paper of Dyson,[1] dealing with the 1-dimensional case governed by the Hamiltonian

$$H = \sum \frac{p_k^2}{2m_k} + \tfrac{1}{2}\sum \kappa_j(x_{j+1} - x_j)^2$$

(with or without periodic boundary conditions).

Recently, Kac[2] proposed yet another approach to the problem which was based on a number of plausible but unproved conjectures. In this paper, we prove that the results of Kac and Dyson are equivalent.

1. THE METHOD OF DYSON: RANDOM MASSES

Let us first consider the case in which all the springs are of equal strength and choose the units so that the elastic moduli $\kappa_j \equiv 1$. Then the frequencies of

Equations (18) provide a whole class of integral transforms labeled by (χ, κ), where $\chi \equiv [l_0, c]$ satisfies either (3a) or (3b) and κ is arbitrary integer, for l_0 integer, and arbitrary half-odd-integer otherwise. We propose the name "Meijer G-transform of order (χ, κ)." For the special cases $\chi = [\frac{1}{2}, 0]$ and $\chi = [0, \frac{1}{2}]$ it goes over into the familiar Hankel transform of order 2κ for the function $\check{f}(x) = xf(\frac{1}{2}x^2)$.

ACKNOWLEDGMENTS

The authors are grateful to Professor A. Salam and Professor P. Budini, and to the International Atomic Energy Agency for the hospitality kindly extended to them at the International Centre for Theoretical Physics, Trieste. We also wish to thank Professor B. Kurşunoğlu for his kind hospitality at the Center for Theoretical Studies, Coral Gables, where part of the work was done.

APPENDIX

For the convenience of the reader we give here a useful reformulation of results of Gel'fand on invariant bilinear forms.[7,9] Let $\zeta = (\zeta_1, \zeta_2) \neq (0, 0)$ be a complex 2-spinor-variable and assign to every 4-vector q^μ a matrix $q = q^\mu \sigma_\mu$. Here σ_κ are Pauli matrices and σ_0 is the 2×2 unit matrix. We then have the following:

Lemma: Suppose $h(\mu\zeta) = |\mu|^{-4}h(\zeta)$ for $\mu \in \mathbb{C}$. Then

$$\int_{\mathbb{C}^2} d^2\zeta \, d^2\bar{\zeta}\delta(\tfrac{1}{2}\zeta q \zeta^+ - 1)h(\zeta)$$

is independent of the positive lightlike vector q^μ.

* Work supported in part by the United States Air Force under grant No. AF AFOSR 1268-67.

[1] S.-J. Chang and L. O'Raifeartaigh, J. Math. Phys. **10**, 21 (1969).
[2] J. F. Boyce, R. Delbourgo, A. Salam, and J. Strathdee, "Partial-Wave Analysis" (Part I), ICTP, Trieste, Report IC/67/9 (unpublished).
[3] G. J. Iverson and G. Mack, "Quantum Fields and Interactions of Massless Particles: The Continuous Spin Case," Center for Theoretical Studies, Coral Gables, preprint CTS-T-Phys-69-2 [submitted to Ann. Phys. (N.Y.)].
[4] S. Ström, Arkiv Fysik **34**, 215 (1967).
[5] H. Bateman, *Higher Transcendental Functions* (McGraw-Hill, New York, 1953), Vol. I.
[6] M. A. Naimark, *Linear Representations of the Lorentz Group* (Pergamon, London, 1964).
[7] I. M. Gel'fand, M. I. Graev, and N. Ya. Valenkin, *Generalized Functions* (Academic, New York, 1966), Vol. 5.
[8] H. Bateman, *Tables of Integral Transforms* (McGraw-Hill, New York, 1954), Vol. 2, Formula 19.3.(45).
[9] V. Gorgé and H. Leutwyler, Helv. Phys. Acta **41**, 171 (1968).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 5 MAY 1970

Problem of the Disordered Chain

TA-FENG LIN
The Rockefeller University, New York, New York 10021

(Received 19 December 1969)

The equivalence of the methods of Dyson and Kac is proven in the 1-dimensional problem of the disordered chain, i.e., the determination of the spectrum of the frequencies of normal modes of a system of coupled oscillators, where the masses and/or spring constants are random variables.

INTRODUCTION

The problem of the disordered chain is to determine the spectrum of the frequencies of normal modes of a system of coupled oscillators when the masses and (or) spring constants are random variables. There exists an extensive literature, beginning with a pioneering paper of Dyson,[1] dealing with the 1-dimensional case governed by the Hamiltonian

$$H = \sum \frac{p_k^2}{2m_k} + \tfrac{1}{2}\sum \kappa_j(x_{j+1} - x_j)^2$$

(with or without periodic boundary conditions).

Recently, Kac[2] proposed yet another approach to the problem which was based on a number of plausible but unproved conjectures. In this paper, we prove that the results of Kac and Dyson are equivalent.

1. THE METHOD OF DYSON: RANDOM MASSES

Let us first consider the case in which all the springs are of equal strength and choose the units so that the elastic moduli $\kappa_j \equiv 1$. Then the frequencies of

the normal modes $0 \leq \omega_1 \leq \omega_2 \leq \cdots \leq \omega_N$ are the roots of the equation

$$\Delta_N = \begin{vmatrix} 2 - m_1\omega^2 & -1 & 0 & \cdots & & 0 \\ -1 & 2 - m_2\omega^2 & -1 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot \\ 0 & \cdots & & 0 & -1 & 2 - m_N\omega^2 \end{vmatrix} = 0. \tag{1.1}$$

We now assume that $m_1, m_2, \cdots, m_N$ are independent, identically distributed random masses with distribution function $R(m)$. The problem is to find the limiting frequency distribution $H(\omega)$, defined as

$$H(\omega) = \lim_{N\to\infty} \frac{1}{N} \langle H_N(\omega) \rangle = \lim_{N\to\infty} \frac{1}{N} \left\langle \sum_{\omega_n < \omega} 1 \right\rangle. \tag{1.2}$$

Dyson[1] derived the formula

$$\int_0^\infty \log\left(1 + \frac{\omega^2}{\xi}\right) dH(\omega) = \int_0^\infty dF(x) \int_0^\infty dR(m) \log\left(1 + \frac{1}{x} + \frac{1}{\xi m}\right) = \left\langle \log\left(1 + \frac{1}{X} + \frac{1}{\xi m}\right) \right\rangle, \quad 0 < \xi, \tag{1.3}$$

where $F(x)$ is the distribution of the random variable X with property (a); i.e., if X_0 is defined by the formula

$$X_0 = \xi m + X(1 + X)^{-1}, \tag{1.4}$$

where X is independent of m, then X_0 has the same distribution function as X, i.e., $F(x)$. Dyson's derivation was somewhat complicated, and we derive (1.3) by a simpler method suggested by Bellman.[3]

Let $D_N = D_N(\xi; m_1, \cdots, m_N)$ be defined for $\xi > 0$ as follows:

$$D_N = \begin{vmatrix} 2 + m_1\xi & -1 & 0 & \cdots & 0 & 0 \\ -1 & 2 + m\xi & -1 & \cdots & 0 & 0 \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ 0 & 0 & 0 & \cdots & -1 & 2 + m_N\xi \end{vmatrix}. \tag{1.5}$$

It is easily checked that

$$D_N = m_1 \cdots m_N \prod_{n=1}^{N} (\xi + \omega_n^2)$$

and, hence,

$$\lim_{N\to\infty} \frac{1}{N} \langle \log D_N \rangle = \langle \log m \rangle + \int_0^\infty \log(\xi + \omega^2)\, dH(\omega). \tag{1.6}$$

From (1.5), we have the familiar recursion formula

$$\frac{D_N}{D_{N-1}} = 2 + m_1\xi - \frac{D_{N-2}}{D_{N-1}}, \tag{1.7}$$

where clearly m_1 and D_{N-1}/D_{N-2} are independent. Let G_N be the distribution function of D_N/D_{N-1}; then, as $N \to \infty$, the limiting distribution function G should have property (b); i.e., if Y has distribution function $G(y)$ and Y is defined by the formula

$$Y_0 = 2 + m\xi - Y^{-1}, \tag{1.8}$$

with m and Y independent, then Y_0 has the same distribution function as Y, i.e., $G(y)$. Using the identity

$$D_N = \frac{D_N}{D_{N-1}} \cdot \frac{D_{N-1}}{D_{N-2}} \cdots \frac{D_1}{D_0} \cdot D_0, \quad D_0 \equiv 1,$$

one has

$$\lim_{N\to\infty} \frac{1}{N} \langle \log D_N \rangle = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left\langle \log \frac{D_n}{D_{n-1}} \right\rangle = \langle \log Y \rangle. \tag{1.9}$$

From (1.6), (1.9), and property (b) of Y, we obtain

$$\int_0^\infty \log\left(1 + \frac{\omega^2}{\xi}\right) dH(\omega) = \langle \log Y \rangle - \langle \log \xi m \rangle = \left\langle \log\left(1 + \frac{1}{\xi m} + \frac{Y - 1}{\xi m Y}\right) \right\rangle, \tag{1.10}$$

where m and Y are independent. In Sec. 3, we prove that (1.3) and (1.10) are, in fact, equivalent.

2. THE METHOD OF KAC: RANDOM MASSES

In a lecture given at the Trondheim Theoretical Physics Seminar in 1968, Kac[2] proposed a completely different approach. Though not simpler than the above approach, it is of some independent interest and it does yield what is probably the most explicit formula for the transform in (1.3) or (1.9).

First of all, Kac introduces the periodic boundary condition so that D_N is now given by the formula

$$D_N = \begin{vmatrix} 2+m_1\xi & -1 & 0 & \cdots & 0 & -1 \\ -1 & 2+m_2\xi & -1 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ -1 & 0 & 0 & & -1 & 2+m_N\xi \end{vmatrix}. \tag{2.1}$$

He then shows that

$$\langle D_N^{\frac{1}{2}k}\rangle = C_{N,k}\lambda^N(k), \quad k = 1, 2, \cdots, \tag{2.2}$$

where $C_{N,k}$ tends to 1 as $N \to \infty$ and where $\lambda(k)$ is the maximum eigenvalue of the Hilbert–Schmidt kernel $K_{(k)}(\mathbf{x}, \mathbf{y})$ which is given by

$$K_{(k)}(\mathbf{x}, \mathbf{y}) = (\pi^{-\frac{1}{2}})^k [G(\xi\,\|\mathbf{x}\|^2)]^{\frac{1}{2}} e^{-\|\mathbf{x}-\mathbf{y}\|^2} [G(\xi\,\|\mathbf{y}\|^2)]^{\frac{1}{2}}. \tag{2.3}$$

Here $\|\mathbf{x}\|$ denotes the k-dimensional Euclidean length of the vector $\mathbf{x} = (x_1, \cdots, x_k)$ and

$$G(\xi\,\|\mathbf{x}\|^2) = \int_0^\infty e^{-\xi m\|\mathbf{x}\|^2}\, dR(m). \tag{2.4}$$

In trying to extend formula (2.2) to k's which are not necessarily positive integers, Kac[2] noted that, since the principal eigenfunction of (2.3) is an eigenfunction of spherical symmetry, $\lambda(k)$ is also the maximum eigenvalue of the kernel:

$$L_k(x, y) = (2\pi^{-\frac{1}{2}})(xy)^{\frac{1}{2}(k-1)} e^{-(x^2+y^2)} [G(\xi x^2) G(\xi y^2)]^{\frac{1}{2}} \times \sum_{n=0}^{\infty} \frac{\Gamma(n+\frac{1}{2})}{(2n)!\,\Gamma(n+\frac{1}{2}k)} (2xy)^{2n}, \quad 0 \le x, y. \tag{2.5}$$

Since $L_k(x, y)$ is defined for all positive k's, it is natural to conjecture that

$$\langle D_N^{-\frac{1}{2}k}\rangle \sim \lambda^N(k), \tag{2.6}$$

for all $k > 0$.

If one goes one step further and conjectures that (2.6) is valid even if

$$k = \epsilon/N,$$

and that, for small k,

$$\lambda(k) \sim 1 - \alpha k + \cdots, \tag{2.7}$$

then one is led to the conjecture that

$$\lim_{N\to\infty} \frac{1}{N} \langle \log D_N\rangle = 2 \lim_{k\to 0} \frac{1-\lambda(k)}{k} = 2\alpha. \tag{2.8}$$

(It is easy to prove that $\lambda(k) \to 1$ as $k \to 0$.)

By a method suggested by Wilson (see the Appendix of Ref. 2), it is then easy to obtain

$$2\alpha = \langle \log(2 + \xi m)\rangle - 2\sum_{n=1}^{\infty} \frac{\lambda_n}{1-\lambda_n} f_n^2, \tag{2.9}$$

where

$$f_n = \int_0^\infty x^{-\frac{1}{2}} G^{\frac{1}{2}}(\xi x^2) e^{-x^2} U_n(x)\, dx \tag{2.10}$$

and the λ_n and the $U_n(x)$ are the eigenvalues and normalized eigenfunction of the kernel.

$$M_0(x, y) = 2(xy)^{-\frac{1}{2}} e^{-(x^2+y^2)} [G(\xi x^2) G(\xi y^2)]^{\frac{1}{2}} \times \sum_{n=1}^{\infty} \frac{(xy)^{2n}}{n!\,(n-1)!}, \quad 0 \le x, y. \tag{2.11}$$

We justify the conjecture (2.8) by showing that it leads to Dyson's result.

3. EQUIVALENCE OF THE DYSON AND KAC METHODS

We seek an eigenfunction of (2.5) in the form

$$\phi(x) = x^{\frac{1}{2}(k-1)} G^{\frac{1}{2}}(\xi x^2) \int_0^\infty e^{-ux^2}\, dQ_k(u). \tag{3.1}$$

Substituting this into the integral equation

$$\int_0^\infty L_k(x, y)\phi(y)\, dy = \lambda(k)\phi(x), \tag{3.2}$$

we obtain after an elementary calculation

$$\lambda(k) \int_0^\infty e^{-ux^2}\, dQ_k(u) = \int_0^\infty \int_0^\infty dR(m) \frac{e^{-(u+\xi m)/(1+u+\xi m)x^2}}{(1+u+\xi m)^{\frac{1}{2}k}}\, dQ_k(u) = \int_0^1 (1-v)^{\frac{1}{2}k} e^{-vx^2}\, dP_k(v), \tag{3.3}$$

where $P_k(v)$ is defined for $0 \le v \le 1$ by the formula

$$P_k(v) = \int_0^{v/(1-v)} Q_k\left(\frac{v}{1-v} - m\right) dR\left(\frac{m}{\xi}\right). \tag{3.4}$$

It is now clear that, in order for (3.3) to be fulfilled, it is enough that we have, for $0 \leq u \leq 1$,

$$Q_k(u) = \int_0^u (1-v)^{\frac{1}{2}k}\, dP_k(v) \Big/ \int_0^1 (1-v)^{\frac{1}{2}k}\, dP_k(v). \tag{3.5}$$

In view of (3.4), this constitutes an integral equation for $Q_k(u)$; if a nondecreasing solution can be found, the corresponding $\phi(x)$ as given by (3.1) is nonnegative and, hence, it is the principal eigenfunction. It also follows that

$$\lambda(k) = \int_0^1 (1-v)^{\frac{1}{2}k}\, dP_k(v). \tag{3.6}$$

In Appendix A, we prove that a nondecreasing solution of (3.5) indeed exists.

If one extends the definition of $Q_k(u)$ and $P_k(u)$ by setting

$$Q_k(u) = 0, \quad u \leq 0,$$
$$Q_k(u) = 1, \quad 1 \leq u,$$

we see that $Q_k(u)$ and $P_k(u)$ are distribution functions. Therefore, it is possible to select a subsequence $k_n \downarrow 0$ such that

$$Q_{k_n}(u) \to Q_0(u), \tag{3.7}$$

where $Q_0(u)$ is again a distribution function. It follows almost at once that Q_0 satisfies the integral equation

$$Q_0(u) = \int_0^u dP_0(v), \tag{3.8}$$

with P_0 given by the formula

$$P_0(v) = \int_0^{v/(1-v)} Q_0\left(\frac{v}{1-v} - m\right) dR\left(\frac{m}{\xi}\right). \tag{3.9}$$

Note also that Q_0 can also be characterized as follows [property (c)]: If Z has the distribution function Q_0 and if Z_0 is defined by the formula

$$Z_0 = 1 - 1/(1 + \xi m + Z), \tag{3.10}$$

with m and Z independent, then Z_0 also has the distribution function Q_0.

Now, it follows from the fact that all the P_k are distribution functions and, from (1.6) and (2.8), that

$$\begin{aligned}\lim_{N\to\infty} \frac{1}{N} \langle \log D_N \rangle &= \int_0^\infty \log(1+v)\, dP_0(v) \\ &= \langle \log(1 + \xi m + Z) \rangle \\ &= \langle \log m \rangle + \int_0^\infty \log(\xi + \omega^2)\, dH(\omega).\end{aligned} \tag{3.11}$$

We note that there is a slight difference between (1.5) and (2.1) if we use D_N in (1.5) instead of D_N in (2.1), and a modification of $C_{N,k}$ in (2.2) should be made. This is clearly a minor matter and the result is not affected.

We now prove that (1.3), (1.10), and (3.11) are equivalent. To see this, we start from (1.8) and rewrite it in the form

$$(Y_0 - 1) = m\xi + (Y-1)/[1 + (Y-1)], \tag{3.12}$$

with m and Y independent. Comparing (3.12) with (1.4), we see clearly that $Y - 1$ and X have the same distribution function F. Thus, (1.4) implies that

$$\begin{aligned}&\langle \log [1 + X^{-1} + (\xi m)^{-1}] \rangle \\ &\quad = \langle \log (X + \xi m + \xi m X) \rangle - \langle \log (\xi m X) \rangle \\ &\quad = \langle \log X_0(1 + X) \rangle - \langle \log (\xi m X) \rangle \\ &\quad = \langle \log (1 + X) \rangle - \langle \log (\xi m) \rangle \\ &\quad = \langle \log Y \rangle - \langle \log \xi m \rangle,\end{aligned}$$

which shows that (1.3) and (1.10) are equivalent. To show that they are also equivalent to (3.11), let us start from (3.10), which is rewritten in the form

$$\frac{1}{1 - Z_0} = 1 + \xi m + Z = 2 + \xi m - \frac{1}{(1-Z)^{-1}}, \tag{3.13}$$

with m and Z independent. Comparing (3.13) and (1.8), we see that Y and $(1-Z)^{-1}$ have the same distribution function or that Z and $(Y-1)/Y$ have the same distribution function P_0. From (1.10), we have immediately that

$$\left\langle \log \left(1 + \frac{1}{\xi m} + \frac{Y-1}{\xi m Y}\right) \right\rangle = \langle \log(\xi m + 1 + Z) \rangle - \langle \log \xi m \rangle,$$

which means that (3.11) is equivalent to (1.10).

4. THE PROBLEM OF RANDOM VARIABLES $\{\kappa_j/m_j, \kappa_j/m_{j+1}\}_j$

Let us discuss another type of disordered chain introduced by Dyson.[1] We assume now that

$$\{\kappa_j/m_j, \kappa_j/m_{j+1}\}_j$$

are independent random variables with common distribution function $G(\mu)$ concentrated on the nonnegative real axis. Dyson's result in this case is

$$\begin{aligned}\int_0^\infty \log\left(1 + \frac{\omega^2}{\xi}\right) dH(\omega) &= 2\int_0^\infty \log(1+u)\, dF(u) \\ &= 2\langle \log(1+U) \rangle,\end{aligned} \tag{4.1}$$

where $F(u)$ is the distribution function of U with the

property (d) that if

$$U_0 = 1/(1 + U), \tag{4.2}$$

where μ is independent of U and has the distribution function G, then the distribution function of U_0 is the same as that of U, i.e., F.

In the special case

$$G(\mu) = \int_0^\mu e^{-t}\, dt, \quad 0 \le \mu, \tag{4.3}$$

$F(u)$ can be found explicitly, with the result

$$F(u) = \int_0^u f(s)\, ds = \int_0^u \frac{1}{K}\frac{e^{-\xi s}}{(1+s)}\, ds, \tag{4.4}$$

where

$$K = \int_0^\infty \frac{1}{1+s}\, e^{-\xi s}\, ds. \tag{4.4'}$$

We first give the Kac approach to this problem. Let

$$D_N = \begin{vmatrix} \kappa_N + \kappa_1 + \xi m_1 & \kappa_1 & 0 & \cdots & 0 & -\kappa_N \\ -\kappa_1 & \kappa_1 + \kappa_2 + \xi m_2 & -\kappa_2 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ -\kappa_{N-1} & 0 & 0 & \cdots & -\kappa_{N-2} & \kappa_{n-2} + \kappa_{N-1} + \xi m_N \end{vmatrix}$$

$$= m_1 \cdots m_N \begin{vmatrix} \mu_1 + \mu_2 + \xi & -\mu_2 & 0 & \cdots & 0 & -\mu_1 \\ -\mu_3 & \mu_3 + \mu_4 + \xi & -\mu_4 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ -\mu_{2N} & 0 & 0 & \cdots & -\mu_{2N-1} & \mu_{2N-1} + \mu_{2N} + \xi \end{vmatrix}$$

$$= m_1 \cdots m_N \cdot \tilde{D}_N, \tag{4.5}$$

where, by assumption, $\mu_1, \mu_2, \cdots, \mu_{2N}$ are independent with common distribution function $G(\mu)$. Then, as before,

$$\lim_{N\to\infty} \frac{1}{N} \langle \log \tilde{D}_N \rangle = \int_0^\infty \log(\xi + \omega^2)\, dH(\omega), \tag{4.6}$$

where $\tilde{D}_N$ is defined by (4.5).

Now the formula

$$\frac{1}{D_N^{\frac{1}{2}}} = (\pi)^{-\frac{1}{2}N} \int_{-\infty}^\infty dx_1 \cdots \int_{-\infty}^\infty dx_N \times \exp\left(-\xi \cdot \sum_1^N m_n x_n^2 - \sum_1^N \kappa_n (x_{n+1} - x_n)^2\right), \quad x_{N+1} = x_1,$$

$$= (m_1 \cdots m_N)^{-\frac{1}{2}} \pi^{-\frac{1}{2}N} \int_{-\infty}^\infty dx_1 \cdots \int_{-\infty}^\infty dx_N \times \exp\left\{-\xi \sum_1^N x_n^2 - \left[\left(\frac{\kappa_n}{m_{n+1}}\right)^{\frac{1}{2}} x_{n+1} - \left(\frac{\kappa_n}{m_n}\right)^{\frac{1}{2}} x_n\right]^2\right\}$$

makes it clear that

$$\langle (\tilde{D}_N)^{-\frac{1}{2}} \rangle = c_N \lambda^N(1), \tag{4.7}$$

where c_N is a constant which tends to 1 as $N \to \infty$ and $\lambda(1)$ is the maximum eigenvalue of the integral kernel $K_{(1)}(x, y)$.

$$K_{(1)}(x, y) = \pi^{-\frac{1}{2}} e^{-\frac{1}{2}\xi x^2} \langle \exp\{-[(\mu x)^{\frac{1}{2}} - (\nu y)^{\frac{1}{2}}]^2\} \rangle e^{-\frac{1}{2}\xi y^2},$$

where μ and ν are independent random variables with common distribution function G. In general, we have, for $k = 1, 2, \cdots$,

$$\langle (\tilde{D}_N)^{-\frac{1}{2}k} \rangle = C_{N,k} \lambda^N(k), \tag{4.8}$$

where $C_{N,k}$ tends to 1 as $N \to \infty$ and $\lambda(k)$ is now the maximum eigenvalue of the integral kernel $K_{(k)}(\mathbf{x}, \mathbf{y})$.

$$K_{(k)}(\mathbf{x}, \mathbf{y}) = \pi^{-\frac{1}{2}k} e^{-\frac{1}{2}\xi \|\mathbf{x}\|^2} \times \langle \exp[-\mu \|\mathbf{x}\|^2 + 2(\mu\nu)^{\frac{1}{2}}(\mathbf{x} \cdot \mathbf{y}) - \nu \|\mathbf{y}\|^2] \rangle e^{-\frac{1}{2}\xi \|\mathbf{y}\|^2}, \tag{4.9}$$

where $\mathbf{x}$ is the vector $(x_1, x_2, \cdots, x_k)$.

Using the fact that, for k an integer, the principal eigenfunction is an eigenfunction of spherical symmetry, we can easily check that $\lambda(k)$ is also the maximum eigenvalue of the Hilbert–Schmidt kernel,

$$L_k(x, y) = 2\pi^{-\frac{1}{2}} (xy)^{\frac{1}{2}(k-1)} e^{-\frac{1}{2}\xi(x^2+y^2)} \times \sum_{n=0}^\infty \frac{\Gamma(n + \frac{1}{2})}{(2n)!\, \Gamma(n + \frac{1}{2}k)} \langle \mu^n e^{-\mu x^2} \rangle \langle \mu^n e^{-\mu y^2} \rangle (2xy)^{2n}, \quad 0 < x, y, \tag{4.10}$$

where μ has the distribution function G.

Kac now makes the same conjecture as before, i.e.,

$$\langle (\tilde{D}_N)^{-\frac{1}{2}k} \rangle \sim \lambda^N(k), \tag{4.11}$$

for all $k > 0$; then, by analogous reasoning, formulas

similar to (2.6) and (2.8) are obtained leading finally to the conjecture that

$$\lim_{N\to\infty} \frac{1}{N} \langle \log \tilde{D}_N \rangle = 2\alpha = -2\lambda'(0). \quad (4.12)$$

We now give a justification of the conjectures using essentially the same method as above. Again we seek the principal eigenfunction in the form

$$\phi_k(x) = x^{\frac{1}{2}(k-1)} e^{-\frac{1}{2}\xi x^2} \int_0^\infty e^{-vx^2}\, dQ_k(v). \quad (4.13)$$

Substituting this into the integral equation with the kernel $L_k(x, y)$, we obtain

$$\begin{aligned}
&\lambda(k) \int_0^\infty e^{-vx^2}\, dQ_k(v) \\
&= \frac{2}{\pi^{\frac{1}{2}}} \sum_{n=0}^{\infty} \frac{\Gamma(n+\frac{1}{2})}{(2n)!\,\Gamma(n+\frac{1}{2}k)} \\
&\quad \times \left\langle (\mu\nu)^n e^{-\nu x^2} \int_0^\infty y^{2n+k-1} e^{-1\xi+\mu+V_k)}\, dy \right\rangle (2x)^{2n} \\
&= \sum_{n=0}^{\infty} \frac{x^{2n}}{n!} \langle (\xi+\mu+V_k)^{-\frac{1}{2}k} \cdot (\mu\nu)^n e^{-\nu x^2} \rangle \\
&= \langle (\xi+\mu+V_k)^{-\frac{1}{2}k} \exp[-(\xi+V_k)\nu x^2/(\xi+\mu+V_k)] \rangle, \quad (4.14)
\end{aligned}$$

where V_k has distribution function $Q_k(v)$.

As in Sec. 3, we prove that the $Q_k(v)$ exists and is unique for $k > 0$ and that $\lambda(k) > 0$ is given by the formula

$$\lambda(k) = \langle (\xi+\mu+V_k)^{-\frac{1}{2}k} \rangle. \quad (4.15)$$

Again, we prove that

$$2\alpha = -2\lambda'(0) = \langle \log(\xi+\mu+V) \rangle, \quad (4.16)$$

where μ is independent of V and V is a random variable whose distribution function Q_0 is such that if

$$V_0 = \mu(\xi+V)/(\xi+\nu+V), \quad (4.17)$$

where μ, ν, and V are independent and where μ and ν have the distribution function G, then V_0 has the same distribution function as V, i.e., Q_0 [property (e)].

To show the equivalence of (4.1) and (4.6), it suffices to prove that (4.12) and (4.16) imply

$$\langle \log(\xi+\mu+V) \rangle = \log\xi + 2\langle \log(1+U) \rangle. \quad (4.18)$$

Rewriting (4.17) in the form

$$\frac{V_0}{\xi} = \frac{\mu}{\xi(1+\nu/\xi(1+V/\xi))}$$

and comparing it with (4.2), we see clearly that V/ξ has property (d). Thus, from (4.2), we have

$$\xi(1+U_0)(1+U) = \xi+\mu+\xi U,$$

and, therefore,

$$\begin{aligned}
\log\xi + 2\langle \log(1+U) \rangle &= \langle \log[\xi(1+U_0)(1+U)] \rangle \\
&= \langle \log(\xi+\mu+\xi U) \rangle \\
&= \langle \log(\xi+\mu+V) \rangle,
\end{aligned}$$

which is (4.18).

We note that Bellman's method can also be applied and that it leads to an equivalent result.

In the special case, when $G(\mu)$ is given by (4.3) [and $F(u)$ by (4.4)], the kernel is

$$\begin{aligned}
L_k(x, y) &= \frac{2}{\pi^{\frac{1}{2}}} (xy)^{\frac{1}{2}(k-1)} e^{-\frac{1}{2}\xi(x^2+y^2)} \\
&\quad \times \sum_{n=0}^{\infty} \frac{\Gamma(n+\frac{1}{2})(n!)^2}{(2n)!\,\Gamma(n+\frac{1}{2}k)} \frac{(2xy)^{2n}}{(1+x^2)^{n+1}(1+y^2)^{n+1}}, \\
&\qquad 0 < x, y,
\end{aligned}$$

and the corresponding maximum eigenvalue is

$$\lambda(k) = \langle (\xi+\mu+\nu U_k)^{-\frac{1}{2}k} \rangle, \quad 0 < k.$$

Here, as before, μ, ν, and U_k are independent, μ and ν have the same distribution function G, and

$$\begin{aligned}
&\lambda(k) \left\langle \frac{1}{1+U_k x} \right\rangle \\
&\quad = \left\langle (\xi+\mu+\nu U_k)^{-\frac{1}{2}k} \left(1 + \frac{\xi+\nu U_k}{\xi+\mu+\nu U_k}\right)^{-1} \right\rangle.
\end{aligned}$$

The derivative of $\lambda(k)$, at $k = 0$, now satisfies the equation

$$2\alpha = -2\lambda'(0) = \langle \log(\xi+\mu+\nu W) \rangle,$$

where W has the property (f), i.e.,

$$W_0 = (\xi+\nu W)/(\xi+\mu+\nu W)$$

(μ, ν, and W are independent, and μ and ν have the distribution function G) has the same distribution function as W.

It is not hard to check that W and $(1+U)^{-1}$ have the same distribution function and, since the density function of U is known explicitly [see (4.4)], we can also determine the density function of W.

ACKNOWLEDGMENT

I wish to express my gratitude to Professor M. Kac for his suggestions regarding the work and his discussion on this problem and, especially, for his help in the writing of the manuscript.

APPENDIX A

Let B be the space of all probability measures concentrated on the closed interval [0, 1] and consider

a mapping L^* from B into B defined for $Q \in B$:

$$(L^*Q)(u) = \int_0^u (1-v)^{\frac{1}{2}k}\, dP(v) \Big/ \int_0^1 (1-v)^{\frac{1}{2}k}\, dP(v), \qquad 0 \le u \le 1,$$

where $P(v)$ is defined by the formula

$$P(v) = \int_0^{v/(1-v)} Q\left(\frac{v}{1-v} - m\right) dR\left(\frac{m}{\xi}\right), \quad 0 \le v \le 1.$$

The mapping L^* is easily seen to be 1-to-1 and (weakly) continuous. We also note that B is a convex and (weakly) compact set, so that the fixed-point theory applies. Hence, there exists $Q_k \in B$ such that

$$Q_k(u) = (L^*Q_k)(u) = \int_0^u (1-v)^{\frac{1}{2}k}\, dP_k(v) \Big/ \int_0^u (1-v)^{\frac{1}{2}k}\, dP_k(v);$$

i.e., Q_k satisfies (3.5).

APPENDIX B

Professor Kac called the author's attention to the following interesting and highly plausible conjecture.

As mentioned in the text (see also Ref. 2), we have for k a positive integer

$$\langle D_N^{-\frac{1}{2}k} \rangle = \sum_j \lambda_j^N(k), \tag{B1}$$

where the λ_j are the eigenvalues of the kernel

$$(\pi^{-\frac{1}{2}})^k [G(\xi \|\mathbf{x}\|^2)]^{\frac{1}{2}} e^{-\|\mathbf{x}-\mathbf{y}\|} [G(\xi \|\mathbf{y}\|^2)]^{\frac{1}{2}}. \tag{B2}$$

Because the kernel is rotation invariant, its eigenfunctions are of the form

$$f(\|\mathbf{x}\|) \cdot S(\mathbf{x}/\|\mathbf{x}\|), \tag{B3}$$

where S is a k-dimensional spherical harmonic.

As is well known, the eigenvalues $\lambda_j(k)$ can be obtained as follows.

Let

$$L_k^{(n)}(r, \rho) = (r\rho)^{\frac{1}{2}(k-1)} e^{-(r^2+\rho^2)} [G(\xi r^2)]^{\frac{1}{2}} [G(\xi \rho^2)]^{\frac{1}{2}} O_k^{(n)}(r, \rho), \tag{B4}$$

where

$$O_k^{(n)}(r, \rho) = \frac{2}{\Gamma(\frac{1}{2}k)} \times \int_0^\pi e^{2r\rho \cos\psi} \tilde{C}_n^{\frac{1}{2}k-1}(\cos\psi) \sin^{k-2}\psi\, d\psi \Big/ \int_0^\pi \sin^{k-2}\psi\, d\psi, \tag{B5}$$

and the $\tilde{C}_n^\nu$ are the normalized Gegenbauer polynomials; i.e.,

$$\tilde{C}_n^\nu(x) = C_n^\nu(x)/C_n^\nu(1), \tag{B6}$$

and the $C_n(x)$ are the standard Gegenbauer polynomials defined by the equation

$$(1 - 2xt + t^2)^{-\nu} = \sum_{n=0}^\infty C_n^\nu(x) t^n. \tag{B7}$$

Let the $\lambda_i(n, k)$ be the eigenvalues of $L_k^{(n)}(r, \rho)$. Then the eigenvalues of (B2) are the $\lambda_i(n, k)$, each counted with multiplicity

$$(2n + k - 2)\Gamma(n + k - 2)/[\Gamma(k-1)\Gamma(n+1)].$$

In particular,

$$\langle D_N^{-\frac{1}{2}k} \rangle = \sum_{n=0}^\infty (2n + k - 2) \times [\Gamma(n + k - 2)/\Gamma(k-1)\Gamma(n+1)] \sum_{i=1}^\infty \lambda_i^N(n, k). \tag{B8}$$

Now, it is easily ascertainable that the right-hand side of (B8) makes sense for all real k. Therefore, it is tempting to conjecture that (B8) holds for all real k.

[1] F. J. Dyson, in *Mathematical Physics in One Dimension*, E. Lieb and D. C. Mattis, Eds. (Academic, New York, 1967).
[2] M. Kac, Nordisk Institute for Teoretisk Atomfysik, Blegdamsvej 17, Copenhagen, Denmark, Preprint No. DK-2100, 1968.
[3] R. Bellman, Phys. Rev. **101**, 19 (1956).

Irreducible Cartesian Tensors. III. Clebsch–Gordan Reduction*

J. A. R. COOPE
Department of Chemistry, University of British Columbia
Vancouver, British Columbia, Canada

(Received 28 October 1966)

The reduction of products of irreducible Cartesian tensors is formulated generally by means of 3-j tensors. These are special cases of the invariant mappings discussed in Part II [J. A. R. Coope and R. F. Snider, J. Math. Phys. **11**, 993 (1970)]. The 3-j formalism is first developed for a general group. Then, the 3-j tensors and spinors for the rotation group are discussed in detail, general formulas in terms of elementary invariant tensors being given. The 6-j and higher n-j symbols coincide with the familiar ones. Interrelations between Cartesian and spherical tensor methods are emphasized throughout.

INTRODUCTION

This paper treats the Clebsch–Gordan reduction for Cartesian tensors within the context of the general theory of Paper II.[1] The tensors to be reduced are now tensor products of two irreducible tensors. The appropriate linear mappings to reduced forms are specified by invariant tensors which are the analogs of the 3-j symbols of spherical tensor theory[2,3] and are, therefore, termed 3-j tensors. These 3-j tensors are conveniently described as linear combinations of tensor products of the elementary invariant tensors for the group, and this representation gives a concrete meaning to the 3-j formalism as a specification of Cartesian tensor operations. On the other hand, they can be represented as "contractions" between the primitive projections for the three symmetries involved and can be explicitly related to spherical tensor 3-j symbols by a resolution of the three projections.

A special case of the Cartesian Clebsch–Gordan reduction for the rotation group has been discussed recently by Barsella and Fabri.[4,5] The full 3-j tensors are not required for this case which, in our language, is the coupling of natural forms to natural forms, nor for the re-embedding of the coupled natural forms, but they are required for tensor-averaging applications of the type of the Wigner–Eckart theorem.[6,7]

Section I summarizes the general theory for reduction of Cartesian tensors, which is the basis for the development of the Clebsch–Gordan formalism. Some details concerning unitary representations are included in Appendix A2. Section II develops the 3-j formalism for a general group, and Sec. III discusses the 1-j tensors and their relation to the spherical 1-j symbols. The remaining sections develop the 3-j formalism in detail for $SO(3)$ and $SU(2)$, i.e., for tensors and spinors under the rotation group.

I. CARTESIAN FORMALISM

We have a group G of linear transformations on a real or complex finite vector space $\mathfrak{X} = \mathfrak{X}^1$, the space of contravariant vectors, a dual space $\overline{\mathfrak{X}} = \mathfrak{X}_1$ of covariant vectors defined by an invariant scalar product, i.e., defined to transform contragrediently according to the inverse transpose representation of G, and we consider induced transformations on $\mathfrak{X}^n_m$, the space of mixed tensors of order (n, m), with n contravariant and m covariant indices. Dual bases $\mathbf{e}_\sigma(n \mid \bar{m})$ and $\mathbf{e}^\sigma(m \mid \bar{n})$ satisfy $\mathbf{e}_\sigma : \mathbf{e}^\tau = \delta^\tau_\sigma$.[8] If $\mathbf{e}_\sigma$ is irreducible of symmetry j, then the dual tensor $\mathbf{e}^\sigma$ is irreducible of symmetry $\bar{j}$, where $\bar{j}$ denotes the contragredient representation.[9,10] For unitary representations, $\bar{j}$ coincides with the complex conjugate representation j^*.

The reduction of the space $\mathfrak{X}^n_m$ into subspaces irreducible under the transformations induced by G is achieved by construction of appropriate projections $\mathbf{\Pi}^j_\alpha$, where j is the symmetry index and α a multiplicity index. $\mathbf{\Pi}^j_\alpha(n\bar{m} \mid m\bar{n})$ is an invariant tensor in $\mathfrak{X}^{m+n}_{m+n}$, and besides being idempotent it must satisfy the condition[11] that the range of the map $\mathbf{\Pi}^j_\alpha : \mathfrak{X}^n_m \to \mathfrak{X}^n_m$ (contraction with right-hand indices) is dual to the range of the dual map $\mathbf{\Pi}^j_\alpha : \mathfrak{X}^m_n \to \mathfrak{X}^m_n$ (contraction with left-hand indices), so that it has a spectral resolution $\mathbf{\Pi}^j_\alpha = \sum_\sigma \mathbf{e}_\sigma \mathbf{e}^\sigma$. $\mathbf{\Pi}^j_\alpha$ projects out an irreducible subspace of symmetry j from $\mathfrak{X}^n_m$ and the dual subspace of symmetry $\bar{j}$ from $\mathfrak{X}^m_n$.

The $\mathbf{\Pi}^j_\alpha$, being invariant tensors, can be constructed from tensor products of the few *elementary* invariant tensors appropriate to the group. These can be interpreted as defining elementary tensor operations—contractions, embeddings, and the substitution of indices. They are usually directly related to the definition of the group, and, in principle, characterize the possible projections $\mathbf{\Pi}^j_\alpha$ and, hence, the

possible irreducible representations completely (Appendix A1).

As discussed in Paper II, any invariant tensor $\mathsf{T}^{(0)}(m \mid \bar{n})$ maps an irreducible subspace $H^j_{n\alpha} \subset \mathfrak{X}^n$ either into the null space or onto an irreducible subspace $H^j_{m\beta} \subset \mathfrak{X}^m$ of the same symmetry. The mapping $\mathsf{T}^{(0)}: H^j_{n\alpha} \to H^j_{m\beta}$ is one-to-one and can be inverted. Thus an irreducible subspace H^j of a space of convenient order can be selected for each symmetry j as a standard space for symmetry j, and every irreducible tensor $\mathsf{F}^{(j)}$ can be represented as a tensor $\mathsf{f}^{(j)}$ in the standard space invariantly embedded:

$$\mathsf{F}^{(j)}(n) = \mathsf{T}^{(0)}(n \mid \bar{m}_j) \cdot \mathsf{f}^{(j)}(m_j). \tag{1}$$

This observation reduces the problem of decomposition of contravariant tensors essentially to a knowledge of a set of projections, denoted by $\mathsf{E}^j(m_j \mid \bar{m}_j)$, into the standard spaces, one for each symmetry j. Where an invariant mapping between spaces of different order is possible, i.e., where covariant elementary invariant tensors exist, the standard space is naturally chosen in the space of least tensorial order consistent with symmetry, and the tensor $\mathsf{f}^{(j)}$ in Eq. (1) may then be called the *reduced form* of $\mathsf{F}^{(j)}$. If the symmetry j occurs in the space of minimal order with multiplicity one, then tensors in H^j have a particularly simple form, being eigentensors of all invariant automorphisms $\mathsf{T}^{(0)}$: $\mathfrak{X}^{m_j} \to \mathfrak{X}^{m_j}$ and, in particular, totally symmetric or totally antisymmetric to permutations of indices—such tensors are appropriately said to be in *natural form*,[12] $\mathsf{f}^{(j)}$ being the natural form of $\mathsf{F}^{(j)}$. Where the multiplicity is not one, it may still be possible to find a relative invariant mapping to a space of relative tensors[13] of lower order and multiplicity one. The inverse map requires a relative tensor of opposite weight so that the embedded relative natural tensor is again an absolute tensor.[14]

Given the standard projections E^j, the reduction of nth-order tensors proceeds by contraction to the minimal rank, projection with E^j, and inversion of the mapping. Precisely, we construct a maximum number (equal to the multiplicity N^j_n of j in $\mathfrak{X}^n$) of linearly independent invariant tensors $\mathsf{T}_p(n \mid \bar{m}_j)$, satisfying $\mathsf{T}_p = \mathsf{T}_p \cdot \mathsf{E}^j$ (so that[1] the image of T_p: $\mathfrak{X}^{m_j} \to \mathfrak{X}^n$ is an irreducible subspace of $\mathfrak{X}^n$ of symmetry j), together with the dual tensors[15] $\mathsf{T}^p(m_j \mid \bar{n})$, satisfying $\mathsf{T}^p = \mathsf{E}^j \cdot \mathsf{T}^p$, and

$$\mathsf{T}^p(m_j \mid \bar{n}) \odot^n \mathsf{T}_q(n \mid \bar{m}_j) = \delta^p_q \mathsf{E}^j(m_j \mid \bar{m}_j). \tag{2}$$

Then the projection into the total invariant but, in general, reducible subspace $H^j_n = \sum_\alpha H^j_{n\alpha}$ of symmetry j (usually the only significant decomposition) is

$$\mathbf{\Pi}^j(n \mid \bar{n}) = \sum_p \mathsf{T}_p(n \mid \bar{m}_j) \odot^{m_j} \mathsf{T}^p(m_j \mid \bar{n}). \tag{3}$$

The T^p extract standard or natural tensors $\mathsf{f}^{(j)p}$ from a given tensor F, and the T_p re-embed them.[16] The T_p can be constructed by applying E^j to arbitrary invariant tensors $\mathsf{S}^{(0)}_p(n \mid \bar{m}_j)$ of appropriate order, and a basis for the T^p can be similarly obtained. In the case of compact and finite groups, G can be taken as a set of unitary transformations on $\mathfrak{X}$; the dual space is essentially the adjoint space and, for any basis T_p, a Hermitian metric g_{pq} is defined, which can be used to raise and lower multiplicity indices (Appendix A2). The Clebsch–Gordan reduction is a specialization of this procedure.

II. CLEBSCH–GORDAN REDUCTION

Consider the decomposition of the tensor product of two irreducible subspaces H^{j_1} and H^{j_2}, the standard spaces for symmetries j_1 and j_2, into a sum of subspaces $(H^{j_1} \otimes H^{j_2})^{j_3}$ of symmetry j_3 (in general, reducible) and of order $n = m_1 + m_2$:

$$H^{j_1} \otimes H^{j_2} = \sum_{j_3} (H^{j_1} \otimes H^{j_2})^{j_3}. \tag{4}$$

This decomposition can be accomplished by the general procedure for tensors of order n (Sec. I), i.e., by constructing a basis $\mathsf{T}_p(n \mid \bar{m}_3)$. Since, however, it is only necessary to span that part of $H^{j_3}_n$ in $H^{j_1} \otimes H^{j_2}$, it is possible to take a smaller basis, that of invariant tensors $\mathsf{T}^{j_1 j_2}_{j_3 p}(n \mid \bar{m}_3)$, satisfying

$$\mathsf{T}^{j_1 j_2}_{j_3 p} = \mathsf{E}^{j_1}\mathsf{E}^{j_2} : \mathsf{T}^{j_1 j_2}_{j_3 p} \cdot \mathsf{E}^{j_3}. \tag{5}$$

Only a number $M(j_1 j_2 \bar{j}_3)$ of such tensors can be linearly independent, equal to the multiplicity of j_3 in $j_1 \otimes j_2$. A set can be constructed by replacing $\mathsf{T}^{j_1 j_2}_{j_3 p}$ on the right of Eq. (5) by arbitrary invariant tensors $\mathsf{S}^{(0)}_p(n \mid \bar{m}_3)$ of order $(m_1 + m_2, m_3)$. With a dual basis[15] $\mathsf{T}^{j_3 p}_{j_1 j_2}$, defined to satisfy

$$\mathsf{T}^{j_3 p}_{j_1 j_2} = \mathsf{E}^{j_3} \cdot \mathsf{T}^{j_3 p}_{j_1 j_2} : \mathsf{E}^{j_1}\mathsf{E}^{j_2} \tag{6}$$

and

$$\mathsf{T}^{j_3 p}_{j_1 j_2} : \mathsf{T}^{j_1 j_2}_{j_3 q} = \delta^p_q \mathsf{E}^{j_3}, \tag{7}$$

the resolution of the projection for $(H^{j_1} \otimes H^{j_2})^{j_3}$ is

$$\mathbf{\Pi}^{j_3}(m_1 + m_2 \mid \bar{m}_1 + \bar{m}_2) = \sum_p \mathsf{T}^{j_1 j_2}_{j_3 p}(m_1 + m_2 \mid \bar{m}_3) \odot^{m_3} \mathsf{T}^{j_3 p}_{j_1 j_2}(m_3 \mid \bar{m}_1 + \bar{m}_2). \tag{8}$$

Summation over j_3 gives the identity for $H^{j_1} \otimes H^{j_2}$, i.e., the completeness relation

$$\sum_{j_3} \mathbf{\Pi}^{j_3}(m_1 + m_2 \mid \bar{m}_1 + \bar{m}_2) = \mathsf{E}^{j_1}(m_1 \mid \bar{m}_1)\mathsf{E}^{j_2}(m_2 \mid \bar{m}_2). \tag{9}$$

Thus, if $A^{(j_1)}(m_1)$ and $B^{(j_2)}(m_2)$ are standard irreducible tensors, the $T^{j_3p}_{j_1j_2}$ couple them to standard tensors $C^{(j_3p)}(m_3)$,

$$C^{(j_3p)} = A^{(j_1)}B^{(j_2)} : T^{j_3p}_{j_1j_2}, \tag{10}$$

and these are re-embedded by the $T^{j_1j_2}_{j_3p}$ in $\mathfrak{X}^{m_1+m_2}$,

$$C^{(j_3)}(m_1+m_2) = \sum_p C^{(j_3)p}(m_3)\cdot T^{j_1j_2}_{j_3p}, \tag{11}$$

in such a way that

$$\sum_{j_3} C^{(j_3)} = A^{(j_1)}B^{(j_2)}. \tag{12}$$

Here $C^{(j_3)}$ is of symmetry j_3 but is, in general, reducible.

The invariant tensors $T^{j_3p}_{j_1j_2}$ and $T^{j_1j_2}_{j_3p}$ are Cartesian "Clebsch–Gordan tensors" of a type analogous to the (unitary) Wigner coefficients for the rotation group. The Cartesian "3-j tensors," which we denote by $V_p(j_1,j_2,\bar{j}_3)$ and $V^p(j_3,\bar{j}_1,\bar{j}_2)$, are defined by Eqs. (5) and (6) in essentially the same way, but are normalized in a way appropriate to bases for tensors of symmetry $j=0$ in $H^{j_1}\otimes H^{j_2}\otimes H^{\bar{j}_3}\subset \mathfrak{X}^{m_1+m_2}_{m_3}$,

$$V^p\begin{pmatrix} j_3 & \bar{j}_1 & \bar{j}_2 \\ t & \bar{r} & \bar{s}\end{pmatrix} : V_q\begin{pmatrix} j_1 & j_2 & \bar{j}_3 \\ r & s & \bar{t}\end{pmatrix} = \delta^p_q, \tag{13}$$

rather than by Eq. (7).[15] Here, and in much of the following, we do not indicate tensorial orders but use generic symbols r, s, $\bar{r}$, r', etc., to distinguish whole sets of Cartesian indices $r=\{r_1r_2\cdots r_m\}$, etc. The orthogonality and completeness relations with respect to j_3 are now

$$[j_3]V^p\begin{pmatrix} j_3' & \bar{j}_1 & \bar{j}_2 \\ t' & \bar{r} & \bar{s}\end{pmatrix} : V_q\begin{pmatrix} j_1 & j_2 & \bar{j}_3 \\ r & s & \bar{t}\end{pmatrix} = \delta_{j_3j_3'}\delta^p_q E^{j_3}(t'\mid \bar{t}) \tag{14}$$

and

$$\sum_{j_3p}[j_3]V_p\begin{pmatrix} j_1 & j_2 & \bar{j}_3 \\ r' & s' & \bar{t}\end{pmatrix}\cdot V^p\begin{pmatrix} j_3 & \bar{j}_1 & \bar{j}_2 \\ t & \bar{r} & \bar{s}\end{pmatrix} = E^{j_1}(r'\mid \bar{r})E^{j_2}(s'\mid \bar{s}). \tag{15}$$

There is considerable symmetry within a given triad $(j_1j_2\bar{j}_3)$ or dual triad $(\bar{j}_1\bar{j}_2j_3)$, and the index p can be associated with the triad as a whole. The orthogonality of Eq. (13) implies three relations

$$[j_i]V^p(j_3,\bar{j}_1,\bar{j}_2) : V_q(j_1,j_2,\bar{j}_3) = \delta^p_q E^{j_i}, \tag{16}$$

where j_i is any one of j_1, j_2, j_3 and the contraction is over the indices not associated with j_i.[17] Moreover, the uniqueness of the coupling[18] $j\otimes\bar{j}\to 0$ implies that the dimensions of the spaces $(H^{j_1}\otimes H^{j_2})^{j_3}$, $(H^{j_1}\otimes H^{\bar{j}_3})^{\bar{j}_2}$, and $(H^{j_2}\otimes H^{\bar{j}_3})^{\bar{j}_1}$ coincide with the dimension of $(H^{j_1}\otimes H^{j_2}\otimes H^{\bar{j}_3})^0$. Thus, the same set of 3-j tensors $V^p(j_3,\bar{j}_1,\bar{j}_2)$ and $V_p(j_1,j_2,\bar{j}_3)$ describes the coupling $j_1\otimes j_2\to j_3$ of two contravariant tensors, the couplings $j_1\otimes\bar{j}_3\to\bar{j}_2$ and $j_2\otimes\bar{j}_3\to\bar{j}_1$ of a contravariant and a covariant tensor to a covariant tensor, for which the projection is of the type

$$\Pi^{\bar{j}_1}(s\bar{t}\mid \bar{s}'t') = \sum_p V_p\begin{pmatrix} j_1 & j_2 & \bar{j}_3 \\ r & s & \bar{t}\end{pmatrix}\bigcirc^r V^p\begin{pmatrix} j_3 & \bar{j}_1 & \bar{j}_2 \\ t' & \bar{r} & \bar{s}'\end{pmatrix}, \tag{17}$$

the coupling $j_1\otimes j_2\otimes\bar{j}_3\to 0$ of one covariant and two contravariant tensors to an invariant, for which the projection is

$$\Pi^{(0)}(rs\bar{t}\mid \bar{r}'\bar{s}'t') = \sum_p V_p\begin{pmatrix} j_1 & j_2 & \bar{j}_3 \\ r & s & \bar{t}\end{pmatrix}V^p\begin{pmatrix} j_3 & \bar{j}_1 & \bar{j}_2 \\ t' & \bar{r}' & \bar{s}'\end{pmatrix}, \tag{18}$$

and the four dual couplings. Equation (18) implies that the invariant part of a tensor triple product is

$$(A^{(j_1)}(r)B^{(j_2)}(s)C^{(\bar{j}_3)}(\bar{t}))^{(0)} = \sum_p \lambda^p V_p\begin{pmatrix} j_1 & j_2 & \bar{j}_3 \\ r & s & \bar{t}\end{pmatrix}, \tag{19}$$

where λ^p is the pth triple scalar product

$$\lambda_p = A^{(j_1)}(r)B^{(j_2)}(s)C^{(\bar{j}_3)}(\bar{t}) : V^p\begin{pmatrix} j_3 & \bar{j}_1 & \bar{j}_2 \\ t & \bar{r} & \bar{s}\end{pmatrix}. \tag{20}$$

This is the tensor property underlying the Wigner–Eckart theorem for a general group. Appendix A3 illustrates some of these properties for $GL(3)$.

For $j_1, j_2, \bar{j}_3$ distinct, we are free to define symbols such as $V_p(j_1,j_2,\bar{j}_3)$, $V_p(j_2,j_1,\bar{j}_3)$, and $V_p(j_1,\bar{j}_3,j_2)$ independently, and to impose simple "symmetry conditions." However, if $j_1=j_2$, then transposition of indices r and s ($r_i\leftrightarrow s_i$) is a symmetry operation, and the set of V_p, $p=1,2,\cdots,M(j_1j_1\bar{j}_3)$, is a basis for a representation of S_2, the symmetric group of order two. That is, from the completeness of the V_p

$$V_q\begin{pmatrix} j_1 & j_1 & \bar{j}_3 \\ s & r & \bar{t}\end{pmatrix} = \sum_p S^p_q V_p\begin{pmatrix} j_1 & j_1 & \bar{j}_3 \\ r & s & \bar{t}\end{pmatrix}, \tag{21}$$

where $\mathbf{S}^2=\mathbf{1}$. By a change of basis, the representation can be decomposed into irreducible representations for which

$$V_\kappa\begin{pmatrix} j_1 & j_1 & \bar{j}_3 \\ s & r & \bar{t}\end{pmatrix} = \pm V_\kappa\begin{pmatrix} j_1 & j_1 & \bar{j}_3 \\ r & s & \bar{t}\end{pmatrix}, \tag{22}$$

according as $H^{j_3}_\kappa(2m_1)$ lies in the symmetric or antisymmetric part of $H^{j_1}\otimes H^{j_1}$. The number of even and odd symbols V_κ is a characteristic of the triad.

In the unitary case (Appendix A2), we define the adjoint of a 3-j tensor,

$$V^\dagger_p\begin{pmatrix} j_3 & \bar{j}_1 & \bar{j}_2 \\ t & \bar{r} & \bar{s}\end{pmatrix} = V_p\begin{pmatrix} j_1 & j_2 & \bar{j}_3 \\ r & s & \bar{t}\end{pmatrix}^*, \tag{23}$$

and a positive-definite Hermitian metric for the multiplicity indices,

$$g_{pq} = \mathsf{V}_p^\dagger \vdots \mathsf{V}_q. \tag{24}$$

Then the inverse tensors are given by

$$\mathsf{V}^p = \sum_q g^{pq}\mathsf{V}_q^\dagger. \tag{25}$$

It is possible, though not necessarily convenient, to orthogonalize the V_p so that $g_{pq} = \delta_{pq}$ and $\mathsf{V}^p = \mathsf{V}_p^\dagger$, in which case the individual terms

$$\mathbf{\Pi}_p^{j_3} = [j_3]\mathsf{V}_p \cdot \mathsf{V}^p \tag{26}$$

are orthogonal projections giving a *complete* decomposition of $H^{j_1} \otimes H^{j_2}$. Further, by a unitary transformation with respect to multiplicity indices, the tensors $\mathsf{C}^{(j_3)p}$ for a particular product $\mathsf{A}^{(j_1)}\mathsf{B}^{(j_2)}$ can be made orthogonal, each being embedded in a different orthogonal subspace.

When j_3 and $\bar{j}_3$ are both contained in $\mathfrak{X}^n$ and, in particular, whenever $j_3 = \bar{j}_3$, then completely covariant and contravariant 3-j tensors can be defined. It is appropriate in this case to start with new symmetrical[19] definitions

$$\mathsf{V}_p\begin{pmatrix} j_1 & j_2 & j_3 \\ r & s & t \end{pmatrix} = \mathsf{E}^{j_1}(r \mid \bar{r}')\mathsf{E}^{j_2}(s \mid \bar{s}')\mathsf{E}^{j_3}(t \mid \bar{t}') \vdots \mathsf{V}_p\begin{pmatrix} j_1 & j_2 & j_3 \\ r' & s' & t' \end{pmatrix}, \tag{27}$$

$$\mathsf{V}^p\begin{pmatrix} \bar{j}_1 & \bar{j}_2 & \bar{j}_3 \\ \bar{r} & \bar{s} & \bar{t} \end{pmatrix} = \mathsf{V}^p\begin{pmatrix} \bar{j}_1 & \bar{j}_2 & \bar{j}_3 \\ \bar{r}' & \bar{s}' & \bar{t}' \end{pmatrix} \vdots \mathsf{E}^{j_1}(r' \mid \bar{r})\mathsf{E}^{j_2}(s' \mid \bar{s})\mathsf{E}^{j_3}(t' \mid \bar{t}), \tag{28}$$

and

$$\mathsf{V}_p \vdots \mathsf{V}^q = \delta_p^q.$$

In the unitary case, we define $\mathsf{V}_p^\dagger \vdots \mathsf{V}_q = g_{pq}$, so that

$$\mathsf{V}^p\begin{pmatrix} \bar{j}_1 & \bar{j}_2 & \bar{j}_3 \\ \bar{r} & \bar{s} & \bar{t} \end{pmatrix} = \sum_q g^{pq}\mathsf{V}_q\begin{pmatrix} j_1 & j_2 & j_3 \\ r & s & t \end{pmatrix}^*. \tag{29}$$

For two j's being equal, the only restriction on "phases" is again that the V_p are a basis for a representation of S_2. However, for $j_1 = j_2 = j_3$, the set of V_p are a basis for a representation of S_3, the symmetric group of order 3, and the irreducible representations characteristic of the triad may include the doubly degenerate representation. With the simplest choice of symmetries, i.e., with the $\mathsf{V}(j,j,j)$ bases for irreducible representations of S_3, some V_p will be symmetric to all exchanges and some antisymmetric, but, in general, there will also be pairs $\mathsf{V}_{\kappa 1}$, $\mathsf{V}_{\kappa 2}$ transforming into one another under the permutations.[20] For a multiplicity-free triad, the representation of S_3 is, of course, 1-dimensional, and $\mathsf{V}(j,j,j)$ is either symmetric or antisymmetric to single exchanges. These symmetries appear in an interesting way in Cartesian theory—for $SO(3)$ and $SU(2)$, for example, they appear directly through the antisymmetry of the elementary tensors ϵ_{rst} and ϵ_{rs}.

The Cartesian 3-j tensors can be related to spherical-tensor 3-j symbols by a spectral resolution of the projections in the defining equations according to

$$\mathsf{E}^j(r \mid \bar{r}') = \sum_m \mathbf{e}_m^j(r)\mathbf{e}^{\bar{j}m}(\bar{r}'). \tag{30}$$

From Eqs. (27) and (28), we have, assuming the phase choices consistent,

$$\mathsf{V}_p\begin{pmatrix} j_1 & j_2 & j_3 \\ r & s & t \end{pmatrix} = \sum_{m_1,m_2,m_3}\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}_p \mathbf{e}_{m_1}^{j_1}(r)\mathbf{e}_{m_2}^{j_2}(s)\mathbf{e}_{m_3}^{j_3}(t), \tag{31}$$

where

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}_p = \mathsf{V}_p\begin{pmatrix} j_1 & j_2 & j_3 \\ r & s & t \end{pmatrix} \vdots \mathbf{e}^{\bar{j}_1 m_1}(\bar{r})\mathbf{e}^{\bar{j}_2 m_2}(\bar{s})\mathbf{e}^{\bar{j}_3 m_3}(\bar{t}), \tag{32}$$

and

$$\mathsf{V}^p\begin{pmatrix} \bar{j}_1 & \bar{j}_2 & \bar{j}_3 \\ \bar{r} & \bar{s} & \bar{t} \end{pmatrix} = \sum_{m_1,m_2,m_3}\begin{pmatrix} \bar{j}_1 & \bar{j}_2 & \bar{j}_3 \\ \bar{m}_1 & \bar{m}_2 & \bar{m}_3 \end{pmatrix}^p \mathbf{e}^{\bar{j}_1 m_1}(\bar{r})\mathbf{e}^{\bar{j}_2 m_2}(\bar{s})\mathbf{e}^{\bar{j}_3 m_3}(\bar{t}), \tag{33}$$

where

$$\begin{pmatrix} \bar{j}_1 & \bar{j}_2 & \bar{j}_3 \\ \bar{m}_1 & \bar{m}_2 & \bar{m}_3 \end{pmatrix}^p = \mathsf{V}^p\begin{pmatrix} \bar{j}_1 & \bar{j}_2 & \bar{j}_3 \\ \bar{r} & \bar{s} & \bar{t} \end{pmatrix} \vdots \mathbf{e}_{m_1}^{j_1}(r)\mathbf{e}_{m_2}^{j_2}(s)\mathbf{e}_{m_3}^{j_3}(t). \tag{34}$$

Equations (32) and (34) imply that the spherical-tensor symbols so defined satisfy

$$[j_3]\sum_{m_1,m_2}\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}_p\begin{pmatrix} \bar{j}_1 & \bar{j}_2 & \bar{j}_3' \\ \bar{m}_1 & \bar{m}_2 & \bar{m}_3' \end{pmatrix}^q = \delta_{j_3 j_3'}\delta_p^q\delta_{m_3}^{m_3'} \tag{35}$$

and

$$\sum_{j_3,p,m_3}[j_3]\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}_p\begin{pmatrix} \bar{j}_1 & \bar{j}_2 & \bar{j}_3 \\ \bar{m}_1' & \bar{m}_2' & \bar{m}_3 \end{pmatrix}^p = \delta_{m_1}^{m_1'}\delta_{m_2}^{m_2'}. \tag{36}$$

In the unitary case,

$$\begin{pmatrix} \bar{j}_1 & \bar{j}_2 & \bar{j}_3 \\ \bar{m}_1 & \bar{m}_2 & \bar{m}_3 \end{pmatrix}^p = \sum_q g^{pq}\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}_q^*, \tag{37}$$

and Eqs. (35)–(36) go over into the unitarity conditions of Derome and Sharp,[3,21] provided the basis is chosen as orthonormal ($g_{pq} = \delta_{pq}$). It may, however, be possible to choose the 3-j tensors to be of simpler form if they are not required to be orthogonal.

III. 1-j TENSORS

The coupling of two irreducible tensors to an invariant is a special case of the Clebsch–Gordan reduction, one j being zero, but it is distinguished because the invariant coupling tensors, here called 1-j tensors in analogy to the so-called 1-j symbols of spherical-tensor theory, define the metric properties of the irreducible subspaces and orthogonality properties of the irreducible tensors.

The mixed 1-j tensor, which is the analog of the spherical symbol $\delta_m^{m'}$, is naturally taken precisely as the projection $\mathsf{E}^j(r \mid \bar{s})$, numerical factors being completely determined by idempotency. It is thus unnecessary to introduce also a symbol $\mathsf{E}^{\bar{j}}(\bar{s} \mid r)$ for projections on the dual space, because idempotency determines the phases, and necessarily

$$\mathsf{E}^{\bar{j}}(\bar{s} \mid r) \equiv \mathsf{E}^j(r \mid \bar{s}). \tag{38}$$

The uniqueness of the coupling $j \otimes \bar{j} \to 0$, that is, the fact that $j_1 \otimes j_2$ contains $j = 0$ if and only if $j_2 = \bar{j}_1$ and then with multiplicity one,[18] follows, in fact, directly from the properties of the projections. The invariant part of the tensor product of two standard irreducible tensors $\mathsf{A}^{(j)}(r)$, $\mathsf{B}^{(\bar{j}')}(\bar{s})$, one contravariant and the other covariant, say $(\mathsf{A}^{(j)}(r)\mathsf{B}^{(\bar{j}')}(\bar{s}))^{(0)} = \mathbf{\Lambda}^{(0)}(r \mid \bar{s})$, satisfies $\mathbf{\Lambda}^{(0)} = \mathsf{E}^j \cdot \mathbf{\Lambda}^{(0)} \cdot \mathsf{E}^{j'}$ and therefore, by Theorems 2 and 3 of Paper II (the latter a form of Schur's lemma), must vanish for $j \neq j'$ and for $j = j'$ must be a multiple of E^j. (This is actually a generalization of Theorem 4, Paper II.) Thus

$$(\mathsf{A}^{(j)}(r)\mathsf{B}^{(\bar{j}')}(\bar{s}))^{(0)} = \lambda\delta_{jj'}\mathsf{E}^j(r \mid \bar{s}), \tag{39}$$

where E^j also, trivially, gives the coupling to the scalar

$$\lambda = [j]^{-1}\mathsf{B}^{(\bar{j})}(\bar{s}) \cdot \mathsf{E}^j(s \mid \bar{r}) \cdot \mathsf{A}^{(j)}(r) \tag{40}$$

$$= [j]^{-1}\mathsf{B}^{(\bar{j})} \cdot \mathsf{A}^{(j)}. \tag{41}$$

In contrast to spherical tensor theory[2,3] which treats all the group representations in an equivalent way, completely covariant and contravariant 1-j tensors do not, in general, exist in Cartesian theory because j and $\bar{j}$ are not, in general, both contained in $\mathfrak{X}^n$, but one in $\mathfrak{X}^n$ and the other in the dual space. If, however, they do both occur in $\mathfrak{X}^n$ and, hence, both also in $\mathfrak{X}_n$, then covariant and contravariant 1-j tensors, which we denote by $\mathbf{g}_j(\bar{r}, \bar{s})$ and $\mathbf{g}^j(r, s)$ and which can be chosen to coincide, respectively, with $[j]^{\frac{1}{2}}\mathsf{V}(^{j0\bar{j}}_{r0\bar{s}})$ and $[j]^{\frac{1}{2}}\mathsf{V}(^{\bar{j}0j}_{\bar{r}0s})$, do exist and are closely related to the spherical 1-j symbols, the spherical tensor metric.[2]

The uniqueness of the coupling $j \otimes \bar{j} \to 0$ implies that the definitions

$$\mathbf{g}_j(\bar{r}, \bar{s}) = \mathbf{g}_j(\bar{r}', \bar{s}') : \mathsf{E}^{\bar{j}}(r' \mid \bar{r})\mathsf{E}^j(s' \mid \bar{s}) \tag{42}$$

and

$$\mathbf{g}^j(r, s) = \mathsf{E}^{\bar{j}}(r \mid \bar{r}')\mathsf{E}^j(s \mid \bar{s}') : \mathbf{g}^j(r', s') \tag{43}$$

are unique to within numerical factors. These are partly determined by the requirement that $\mathbf{g}^j$ be the inverse of $\mathbf{g}_j$, in the sense

$$\mathbf{g}^j(r, s) \cdot \mathbf{g}_j(\bar{r}, \bar{s}') = \mathsf{E}^j(s \mid \bar{s}'), \tag{44}$$

which implies also

$$\mathbf{g}^j(r, s) \cdot \mathbf{g}_j(\bar{r}', \bar{s}) = \mathsf{E}^j(r \mid \bar{r}'), \tag{45}$$

[cf. Eqs. (16)]. In the unitary case, the numerical factors are determined up to a phase factor by

$$\mathbf{g}^j = \mathbf{g}_j^\dagger \tag{46}$$

[that is, $\mathbf{g}^j(r, s) = \mathbf{g}_j(\bar{r}, \bar{s})^*$] which is consistent with Eqs. (42)–(43) because E^j is self-adjoint.

The uniqueness of the coupling also implies that the 1-j tensors for $\bar{j}$ are, to within numerical factors, the transpose of those for j, i.e.,

$$\mathbf{g}_j(\bar{r}, \bar{s}) = \lambda\mathbf{g}_{\bar{j}}(\bar{s}, \bar{r}), \tag{47}$$

$$\mathbf{g}^j(r, s) = \lambda^{-1}\mathbf{g}^{\bar{j}}(s, r), \tag{48}$$

where, in the unitary case, λ is at most a phase factor. For $j \neq \bar{j}$, we can define $\mathbf{g}_j$ and $\mathbf{g}_{\bar{j}}$ independently. However, for $j = \bar{j}$, in which case $\mathbf{g}_j(\bar{r}, \bar{s})$ is the metric for H^j, Eq. (47) implies $\mathbf{g}_j(\bar{r}, \bar{s}) = \lambda^2\mathbf{g}_j(\bar{r}, \bar{s})$, whence necessarily

$$\lambda = \pm 1, \tag{49}$$

the metric being symmetric or antisymmetric according to whether $j = 0$ is contained in the symmetric or antisymmetric part of $H^j \otimes H^j$, i.e., according to whether the representations are "integral" or "half-integral."[3,22] For spinors, for example,[23]

$$\mathbf{g}_j(\bar{r}, \bar{s}) = (\epsilon_{r,s})^{2j}, \tag{50}$$

with the symmetry $(-1)^{2j}$.

The relation to the spherical symbols is given by specialization of Eqs. (31)–(34), which here take the form

$$\mathbf{g}_j(\bar{r}, \bar{s}) = \sum_{m,m'} (j)_{mm'}\mathbf{e}^{jm}(\bar{r})\mathbf{e}^{\bar{j}m'}(\bar{s}), \tag{51}$$

where

$$(j)_{mm'} = \mathbf{g}_j(\bar{r}, \bar{s}) : \mathbf{e}_m^{\bar{j}}(r)\mathbf{e}_{m'}^j(s), \tag{52}$$

and similarly,

$$\mathbf{g}^j(r, s) = \sum_{m,m'} (j)^{mm'}\mathbf{e}_m^{\bar{j}}(r)\mathbf{e}_{m'}^j(s), \tag{53}$$

where

$$(j)^{mm'} = \mathbf{g}^j(r, s) : \mathbf{e}^{jm}(\bar{r})\mathbf{e}^{\bar{j}m'}(\bar{s}). \tag{54}$$

The symbols $(j)_{mm'}$ and $(j)^{mm'}$ have the usual properties of the spherical 1-j symbols.[3] Equation (51) implies that they couple spherical tensors to invariants in accord with the equivalence

$$\mathbf{g}_j(\bar{r}, \bar{s}) : \mathsf{A}^{(\bar{j})}(r)\mathsf{B}^{(j)}(s) = \sum_{m,m'} (j)_{mm'} A^{\bar{j}m} B^{jm'}, \tag{55}$$

where $A^{\bar{j}m} = \mathbf{e}^{jm}(r) \cdot \mathsf{A}^{(j)}(r)$ is a spherical tensor component of $\mathsf{A}^{\bar{j}}$. Similarly, Eqs. (44)–(45) imply the relations

$$\begin{aligned} \sum_m (j)^{m'm}(j)_{m''m} &= \delta^{m'}_{m''}, \\ \sum_m (j)^{mm'}(j)_{mm''} &= \delta^{m'}_{m''}, \end{aligned} \tag{56}$$

and Eq. (49) the usual symmetry for $j = \bar{j}$,

$$(j)_{mm'} = \pm(j)_{m'm}, \tag{57}$$

for the integral and half-integral representations. In the unitary case, from Eqs. (46), (52), and (54),

$$(j)^{mm'} = (j)^*_{mm'}. \tag{58}$$

The Cartesian 1-j tensors raise and lower indices in standard tensors in the manner

$$\begin{aligned} \mathbf{g}_j(\bar{r}, \bar{s}) \cdot \mathsf{A}^j(s) &= \mathsf{A}^{\bar{j}}(\bar{r}), \\ \mathbf{g}^j(r, s) \cdot \mathsf{A}^{\bar{j}}(\bar{r}) &= \mathsf{A}^j(s), \end{aligned} \tag{59}$$

so that, for $j = \bar{j}$,

$$\mathsf{A}^j(r) \cdot \mathsf{B}^j(\bar{r}) = (-1)^{2j}\mathsf{A}^j(\bar{r}) \cdot \mathsf{B}^j(r). \tag{60}$$

But note, however, for $j \neq \bar{j}$ that

$$\mathbf{g}_j(\bar{r}, \bar{s}) \cdot \mathsf{A}^j(r) = 0, \tag{61}$$

[cf. Eq. (47)]. In particular, we can lower indices for a spherical-tensor basis set in two ways,

$$\begin{aligned} \mathbf{e}^j_m(\bar{r}) &= \mathbf{g}_j(\bar{r}, \bar{s}) \cdot \mathbf{e}^j_m(s), \\ &= \sum_{m'} (j)_{m'm}\mathbf{e}^{\bar{j}m'}(\bar{r}), \end{aligned} \tag{62}$$

and using these $\mathbf{e}^j_m(\bar{r})$ tensors, we have the following representation of the spherical metric:

$$(j)_{mm'} = \mathbf{e}^{\bar{j}}_m(r) \cdot \mathbf{e}^j_{m'}(\bar{r}). \tag{63}$$

For ambivalent groups[22] and, in particular, for simply reducible[22] groups, $j = \bar{j}$ for all representations and, therefore, a covariant 1-j tensor exists for all symmetries. This is true, in particular, for the vector representations (basis in $\mathfrak{X}^1$), and it implies the existence of an invariant "metric" in the tensor space for all ambivalent groups, i.e., a covariant elementary invariant tensor of *second order* defining the invariant product of two contravariant vectors.[24]

For orthogonal groups with $\delta_{r_1r_2}$ invariant, covariant and contravariant indices are not distinguished, and $\mathbf{g}_j(\bar{r}, \bar{s})$ and the transpose of $\mathbf{g}^j(s, r)$ can be chosen to coincide with $\mathsf{E}^j(r \mid \bar{s})$. Then Eq. (38) or Eqs. (46)–(48) express the symmetry $\mathsf{E}^{\bar{j}}(s \mid r) = \mathsf{E}^j(r \mid s) = [\mathsf{E}^j(s \mid r)]^*$ for the complex representations $(j \neq \bar{j})$—e.g., for $SO(2)$ with $j = m$ and $\bar{j} = -m$,[1,23] we have

$$\begin{aligned} \mathsf{E}^m(r \mid s) &= 2^{-m}(\delta_{rs} + i\epsilon_{rs})^m \\ &= 2^{-m}(\delta_{sr} - i\epsilon_{sr})^m = \mathsf{E}^{-m}(s \mid r). \end{aligned} \tag{64}$$

For real representations $(j = \bar{j})$, Eq. (38) implies that E^j is symmetric,

$$\mathsf{E}^j(r \mid s) = \mathsf{E}^j(s \mid r), \tag{65}$$

and, since E^j is self-adjoint, also real, and, by Eqs. (47)–(49), that real representations are integral. Conversely, a real symmetric E^j implies $j = \bar{j}$.

IV. 3-j TENSORS FOR THE ROTATION GROUP

For the rotation group $SO(3)$, it is unnecessary to distinguish covariant and contravariant indices. The elementary invariant tensors are δ_{rs} and ϵ_{rst}, related, respectively, to the dot and cross products. The natural form for a tensor of symmetry j is a traceless symmetric tensor of order j, and the projections have the representation[1,23]

$$\mathsf{E}^j(r \mid s) = \sum_p c_p \delta^{j-2p}_{rs}\delta^p_{rr}\delta^p_{ss}, \tag{66}$$

where the coefficients

$$c_p(j) = (-1)^p \frac{(j!)^2(2j - 2p)!}{(2j)!\, p!\, (j - p)!\, (j - 2p)!} \tag{67}$$

are, apart from the normalization $c_0 = 1$, the coefficients of Legendre polynomials. The fact that the E^j are symmetric to the interchange $r \leftrightarrow s$ implies that the group is ambivalent and the representations integral.

The 3-j tensors are obtained by applying the projections E^{j_1}, E^{j_2}, E^{j_3} to arbitrary invariant tensors $\mathbf{S}^{(0)}_p(r, s, t)$ of order $J = (j_1 + j_2 + j_3)$. Since the E^j symmetrize indices and annihilate trace terms of the type $\delta_{r_1r_2}$, only elementary tensors of the types δ_{rs}, δ_{st}, δ_{rt}, and ϵ_{rst} need be considered, and $\mathbf{S}^{(0)}$ may, without loss of generality, be restricted to the form

$$\begin{aligned} \mathbf{S}^{(0)} &= \delta^\alpha_{st}\delta^\beta_{tr}\delta^\gamma_{rs}, \qquad J \text{ even}, \\ \mathbf{S}^{(0)} &= \epsilon_{rst}\delta^\alpha_{st}\delta^\beta_{tr}\delta^\gamma_{rs}, \quad J \text{ odd}. \end{aligned} \tag{68}$$

The exponents must be positive or zero and such that the total numbers of r, s, and t indices are, respectively, j_1, j_2, and j_3. Thus, for J even,

$$\begin{aligned} \beta + \gamma &= j_1, \\ \gamma + \alpha &= j_2, \\ \alpha + \beta &= j_3, \end{aligned} \tag{69}$$

with the unique solutions $\alpha = \frac{1}{2}J - j_1$, etc., which are nonnegative if and only if the triangle conditions

$$|j_i - j_j| \leq j_k \leq j_i + j_j, \quad \text{cyclic}, \tag{70}$$

are satisfied. For J odd, the exponents $\alpha = \frac{1}{2}(J-1) - j_1$, etc., are again unique, and the triangle conditions must again be satisfied. This completes the proof within Cartesian theory (i.e., based on the existence of just the two elementary invariants δ_{rs} and ϵ_{rst}) that $SO(3)$ is simply reducible with respect to tensor representations according to the Clebsch–Gordan series $j_3 = j_1 + j_2, j_1 + j_2 - 1, \cdots, |j_1 - j_2|$.

The 3-j tensors themselves are conveniently taken in an unnormalized form, denoted by $\mathsf{T}(j_1, j_2, j_3)$, such that the leading term [one or other of the terms in Eq. (68)] has coefficient unity. For J even,

$$\mathsf{T}\begin{pmatrix} j_1 & j_2 & j_3 \\ r & s & t \end{pmatrix} = \mathsf{E}^{j_1}\mathsf{E}^{j_2}\mathsf{E}^{j_3} \vdots (\delta_{st}^{\frac{1}{2}J-j_1}\delta_{tr}^{\frac{1}{2}J-j_2}\delta_{rs}^{\frac{1}{2}J-j_3}), \tag{71}$$

and for J odd,

$$\mathsf{T}\begin{pmatrix} j_1 & j_2 & j_3 \\ r & s & t \end{pmatrix} = \mathsf{E}^{j_1}\mathsf{E}^{j_2}\mathsf{E}^{j_3} \vdots (\epsilon_{rst}\delta_{st}^{\frac{1}{2}(J-1)-j_1}\delta_{tr}^{\frac{1}{2}(J-1)-j_2}\delta_{rs}^{\frac{1}{2}(J-1)-j_3}). \tag{72}$$

The normalized 3-j tensors are then

$$\mathsf{V}(j_1, j_2, j_3) = \Omega^{-\frac{1}{2}}\mathsf{T}(j_1, j_2, j_3), \tag{73}$$

where

$$\Omega_{j_1 j_2 j_3} = \mathsf{T}(j_1, j_2, j_3) \vdots \mathsf{T}(j_1, j_2, j_3). \tag{74}$$

These 3-j tensors have the usual maximum symmetries. For J even, $\mathsf{T}(j_1, j_2, j_3)$ is invariant to all permutations of j_1, j_2, j_3, but for J odd, it has the symmetry of ϵ_{rst}—changing sign for odd permutations and also, if we insist that ϵ_{rst} be an absolute tensor,[13] under inversion of the coordinate system.

The terms δ_{rs} and ϵ_{rst} in Eqs. (71)–(72) take contractions between the three projections. Thus the 3-j tensors also have the following representation: For J even,

$$\begin{aligned}\mathsf{T}\begin{pmatrix} j_1 & j_2 & j_3 \\ r & s & t \end{pmatrix} = \sum_{\text{all } x,y,z} & \mathsf{E}^{j_1}(r \mid y_1 \cdots y_\beta z_1 \cdots z_\gamma) \\ & \times \mathsf{E}^{j_2}(s \mid z_1 \cdots z_\gamma x_1 \cdots x_\alpha) \\ & \times \mathsf{E}^{j_3}(t \mid x_1 \cdots x_\alpha y_1 \cdots y_\beta),\end{aligned} \tag{75}$$

where $\alpha = \frac{1}{2}J - j_1$, etc., and for J odd,

$$\begin{aligned}\mathsf{T}\begin{pmatrix} j_1 & j_2 & j_3 \\ r & s & t \end{pmatrix} = \sum_{\text{all } x,y,z} & \epsilon_{x_0 y_0 z_0}\mathsf{E}^{j_1}(r \mid x_0 y_1 \cdots y_\beta z_1 \cdots z_\gamma) \\ & \times \mathsf{E}^{j_2}(s \mid y_0 z_1 \cdots z_\gamma x_1 \cdots x_\alpha) \\ & \times \mathsf{E}^{j_3}(t \mid z_0 x_1 \cdots x_\alpha y_1 \cdots y_\beta),\end{aligned} \tag{76}$$

where $\alpha = \frac{1}{2}(J-1) - j_1$, etc. There are two especially simple cases. For the coupling to maximum j,

$$\mathsf{T}(j_1, j_2, j_1 + j_2) = \mathsf{E}^{j_3}(r_1 \cdots r_{j_1} s_1 \cdots s_{j_2} \mid t_1 \cdots t_{j_3}), \tag{77}$$

and, in particular, for $j_2 = 0$,

$$\mathsf{T}(j, 0, j) = \mathsf{E}^{j}(r \mid t). \tag{78}$$

For the coupling to $j_3 = j_1 + j_2 - 1$,

$$\mathsf{T}(j_1, j_2, j_1 + j_2 - 1) = \sum_x \epsilon_{rsx}\mathsf{E}^{j_3}(x r_1 \cdots r_{j_1-1} s_1 \cdots s_{j_2-1} \mid t_1 \cdots t_{j_3}), \tag{79}$$

and, in particular, for $j_2 = 1$,

$$\mathsf{T}(j, 1, j) = \sum_x \epsilon_{rsx}\mathsf{E}^{j}(r_1 \cdots r_{j-1} x \mid t_1 \cdots t_j). \tag{80}$$

Equation (79) is proved by observing that the right side is, in fact, traceless and symmetric in r, s, and t, and has the correct leading term.

Expansion of the projections E^j in Eqs. (71)–(72), in terms of elementary tensors, gives the expressions

$$\begin{aligned}\mathsf{T}\begin{pmatrix} j_1 & j_2 & j_3 \\ r & s & t \end{pmatrix} = \sum_{a,b,c} x_{a,b,c}[\delta_{rr}^{a}\delta_{ss}^{b}\delta_{tt}^{c}\delta_{st}^{\frac{1}{2}J-j_1+a-b-c} \\ \times \delta_{tr}^{\frac{1}{2}J-j_2-a+b-c}\delta_{rs}^{\frac{1}{2}J-j_3-a-b+c}]\end{aligned} \tag{81}$$

for J even, and

$$\begin{aligned}\mathsf{T}\begin{pmatrix} j_1 & j_2 & j_3 \\ r & s & t \end{pmatrix} = \sum_{a,b,c} x_{a,b,c}[\epsilon_{rst}\delta_{rr}^{a}\delta_{ss}^{b}\delta_{tt}^{c}\delta_{st}^{\frac{1}{2}(J-1)-j_1+a-b-c} \\ \times \delta_{tr}^{\frac{1}{2}(J-1)-j_2-a+b-c}\delta_{rs}^{\frac{1}{2}(J-1)-j_3-a-b+c}]\end{aligned} \tag{82}$$

for J odd. Here a, b, and c are the numbers of trace terms. The $x_{a,b,c}$ may be called 3-j *coefficients*[25] (with respect to a nonorthonormal basis of tensor products of elementary invariant tensors) and are such that $\mathsf{T}(j_1, j_2, j_3)$ is traceless and symmetric in the three sets of indices. Table I gives the 3-j tensors for $J \leq 11$ explicitly in this form, together with the norms Ω. General formulas for the $x_{a,b,c}$ and Ω are obtained in Secs. VI and VIII.

For the unnormalized 3-j tensors in Table I, the coupling formulas take the form

$$\mathsf{C}^{(j_3)}(t) = \left(\frac{[j_3]}{\Omega}\right)^{\frac{1}{2}}\mathsf{T}\begin{pmatrix} j_1 & j_2 & j_3 \\ r & s & t \end{pmatrix} : \mathsf{A}^{(j_1)}(r)\mathsf{B}^{(j_2)}(s) \tag{83}$$

and

$$\mathsf{C}^{(j_3)}(r, s) = \left(\frac{[j_3]}{\Omega}\right)^{\frac{1}{2}}\mathsf{T}\begin{pmatrix} j_1 & j_2 & j_3 \\ r & s & t \end{pmatrix} \cdot \mathsf{C}^{(j_3)}(t). \tag{84}$$

Only the coefficients of type $x_{0,0,c}$ are required for the coupling to natural form since $\mathsf{A}^{(j_1)}(r)$ and $\mathsf{B}^{(j_2)}(s)$

TABLE I. Cartesian 3-j tensors for the rotation group for $J = j_1 + j_2 + j_3 \leq 11$. $\mathsf{V}(j_1, j_2, j_3) = \Omega^{-\frac{1}{2}}\mathsf{T}(j_1, j_2, j_3)$. Repeated indices are symmetrized (see Ref. 19).

J	3-j tensor	Expression	$\Omega = \mathsf{T} \vdots \mathsf{T}$
2	T(1, 1, 0)	δ_{rs}	3
3	T(1, 1, 1)	ϵ_{rst}	6
4	T(2, 2, 0)	$\delta^2_{rs} - \frac{1}{3}\delta_{rr}\delta_{ss}$	5
	T(2, 1, 1)	$\delta_{rs}\delta_{rt} - \frac{1}{3}\delta_{rr}\delta_{st}$	5
5	T(2, 2, 1)	$\epsilon_{rst}\delta_{rs}$	$\frac{15}{2}$
6	T(3, 3, 0)	$\delta^3_{rs} - \frac{3}{5}\delta_{rr}\delta_{ss}\delta_{rs}$	7
	T(3, 2, 1)	$\delta^2_{rs}\delta_{rt} - \frac{2}{5}\delta_{rr}\delta_{rs}\delta_{st} - \frac{1}{5}\delta_{rr}\delta_{ss}\delta_{rt}$	7
	T(2, 2, 2)	$\delta_{rs}\delta_{rt}\delta_{st} - \frac{1}{3}(\delta_{rr}\delta^2_{st} + \delta_{ss}\delta^2_{rt} + \delta_{tt}\delta^2_{rs}) + \frac{2}{9}\delta_{rr}\delta_{ss}\delta_{tt}$	$\frac{35}{12}$
7	T(3, 3, 1)	$\epsilon_{rst}(\delta^2_{rs} - \frac{1}{5}\delta_{rr}\delta_{ss})$	$\frac{28}{3}$
	T(3, 2, 2)	$\epsilon_{rst}(\delta_{rs}\delta_{rt} - \frac{1}{5}\delta_{rr}\delta_{st})$	7
8	T(4, 4, 0)	$\delta^4_{rs} - \frac{6}{7}\delta_{rr}\delta_{ss}\delta^2_{rs} + \frac{3}{35}\delta^2_{rr}\delta^2_{ss}$	9
	T(4, 3, 1)	$\delta^3_{rs}\delta_{rt} - \frac{3}{7}(\delta_{rr}\delta^2_{rs}\delta_{st} + \delta_{rr}\delta_{ss}\delta_{rs}\delta_{rt}) + \frac{3}{35}\delta^2_{rr}\delta_{ss}\delta_{st}$	9
	T(4, 2, 2)	$\delta^2_{rs}\delta^2_{rt} - \frac{4}{7}\delta_{rr}\delta_{rs}\delta_{rt}\delta_{st} - \frac{1}{7}\delta_{rr}(\delta_{ss}\delta^2_{rt} + \delta_{tt}\delta^2_{rs}) + \frac{3}{35}\delta^2_{rr}\delta^2_{st} + \frac{1}{35}\delta^2_{rr}\delta_{ss}\delta_{tt}$	9
	T(3, 3, 2)	$\delta^2_{rs}\delta_{rt}\delta_{st} - \frac{2}{5}(\delta_{rr}\delta_{rs}\delta^2_{st} + \delta_{ss}\delta_{rs}\delta^2_{rt}) - \frac{1}{3}\delta_{tt}\delta^3_{rs} - \frac{1}{25}\delta_{rr}\delta_{ss}\delta_{rt}\delta_{st} + \frac{2}{35}\delta_{rr}\delta_{ss}\delta_{tt}\delta_{rs}$	$\frac{14}{5}$
9	T(4, 4, 1)	$\epsilon_{rst}(\delta^3_{rs} - \frac{3}{7}\delta_{rr}\delta_{ss}\delta_{rs})$	$\frac{45}{4}$
	T(4, 3, 2)	$\epsilon_{rst}(\delta^2_{rs}\delta_{rt} - \frac{1}{7}\delta_{rr}\delta_{ss}\delta_{rt} - \frac{2}{7}\delta_{rr}\delta_{rs}\delta_{st})$	$\frac{15}{2}$
	T(3, 3, 3)	$\epsilon_{rst}(\delta_{rs}\delta_{rt}\delta_{st} - \frac{1}{5}(\delta_{rr}\delta^2_{st} + \delta_{ss}\delta^2_{rt} + \delta_{tt}\delta^2_{rs}) + \frac{2}{25}\delta_{rr}\delta_{ss}\delta_{tt})$	$\frac{21}{5}$
10	T(5, 5, 0)	$\delta^5_{rs} - \frac{10}{9}\delta_{rr}\delta_{ss}\delta^3_{rs} + \frac{5}{21}\delta^2_{rr}\delta^2_{ss}\delta_{rs}$	11
	T(5, 4, 1)	$\delta^4_{rs}\delta_{rt} - \frac{4}{9}\delta_{rr}\delta^3_{rs}\delta_{st} - \frac{2}{3}\delta_{rr}\delta_{ss}\delta^2_{rs}\delta_{rt} + \frac{4}{21}\delta^2_{rr}\delta_{ss}\delta_{rs}\delta_{st} + \frac{1}{21}\delta^2_{rr}\delta^2_{ss}\delta_{rt}$	11
	T(5, 3, 2)	$\delta^3_{rs}\delta^2_{rt} - \frac{2}{3}\delta_{rr}\delta^2_{rs}\delta_{rt}\delta_{st} - \frac{1}{3}\delta_{rr}\delta_{ss}\delta_{rs}\delta^2_{rt} - \frac{1}{9}\delta_{rr}\delta_{tt}\delta^3_{rs} + \frac{2}{21}\delta^2_{rr}\delta_{rs}\delta^2_{st} + \frac{2}{21}\delta^2_{rr}\delta_{ss}\delta_{rt}\delta_{st} + \frac{1}{21}\delta^2_{rr}\delta_{ss}\delta_{tt}\delta_{rs}$	11
	T(4, 4, 2)	$\delta^3_{rs}\delta_{rt}\delta_{st} - \frac{3}{7}(\delta_{rr}\delta^2_{rs}\delta^2_{st} + \delta_{ss}\delta^2_{rs}\delta^2_{rt}) - \frac{1}{3}\delta_{tt}\delta^4_{rs} - \frac{3}{49}\delta_{rr}\delta_{ss}\delta_{rs}\delta_{rt}\delta_{st} + \frac{3}{49}(\delta^2_{rr}\delta_{ss}\delta^2_{st} + \delta_{rr}\delta^2_{ss}\delta^2_{rt}) + \frac{17}{49}\delta_{rr}\delta_{ss}\delta_{tt}\delta^2_{rs} - \frac{2}{49}\delta^2_{rr}\delta^2_{ss}\delta_{tt}$	$\frac{165}{56}$
	T(4, 3, 3)	$\delta^2_{rs}\delta^2_{rt}\delta_{st} - \frac{4}{7}\delta_{rr}\delta_{rs}\delta_{rt}\delta^2_{st} - \frac{2}{5}(\delta_{ss}\delta_{rs}\delta^3_{rt} + \delta_{tt}\delta^3_{rs}\delta_{rt}) + \frac{2}{35}\delta^2_{rr}\delta^3_{st} + \frac{1}{35}(\delta_{rr}\delta_{ss}\delta^2_{rt}\delta_{st} + \delta_{rr}\delta_{tt}\delta^2_{rs}\delta_{st}) + \frac{12}{35}\delta_{rr}\delta_{ss}\delta_{tt}\delta_{rs}\delta_{rt} - \frac{1}{25}\delta^2_{rr}\delta_{ss}\delta_{tt}\delta_{st}$	$\frac{11}{5}$
11	T(5, 5, 1)	$\epsilon_{rst}(\delta^4_{rs} - \frac{2}{3}\delta_{rr}\delta_{ss}\delta^2_{rs} + \frac{1}{21}\delta^2_{rr}\delta^2_{ss})$	$\frac{66}{5}$
	T(5, 4, 2)	$\epsilon_{rst}(\delta^3_{rs}\delta_{rt} - \frac{1}{3}\delta_{rr}\delta^2_{rs}\delta_{st} - \frac{1}{3}\delta_{rr}\delta_{ss}\delta_{rs}\delta_{rt} + \frac{1}{21}\delta^2_{rr}\delta_{ss}\delta_{st})$	$\frac{33}{4}$
	T(5, 3, 3)	$\epsilon_{rst}(\delta^2_{rs}\delta^2_{rt} - \frac{4}{9}\delta_{rr}\delta_{rs}\delta_{rt}\delta_{st} - \frac{1}{9}(\delta_{rr}\delta_{ss}\delta^2_{rt} + \delta_{rr}\delta_{tt}\delta^2_{rs}) + \frac{2}{63}\delta^2_{rr}\delta^2_{st} + \frac{1}{63}\delta^2_{rr}\delta_{ss}\delta_{tt})$	$\frac{22}{3}$
	T(4, 4, 3)	$\epsilon_{rst}(\delta^2_{rs}\delta_{rt}\delta_{st} - \frac{2}{7}(\delta_{rr}\delta_{rs}\delta^2_{st} + \delta_{ss}\delta_{rs}\delta^2_{rt}) - \frac{1}{5}\delta_{tt}\delta^3_{rs} - \frac{3}{49}\delta_{rr}\delta_{ss}\delta_{rt}\delta_{st} + \frac{31}{245}\delta_{rr}\delta_{ss}\delta_{tt}\delta_{rs})$	$\frac{99}{28}$

are already traceless. Indeed, all that is required is to contract the product $\mathsf{A}^{(j_1)}\mathsf{B}^{(j_2)}$ to a tensor of order j_3 by $(\frac{1}{2}J - j_3)$ dot products or, for J odd, $[\frac{1}{2}(J-1) - j_3]$ dot products and one cross product, and to make the result traceless and symmetric. The latter could be accomplished by projecting with E^{j_3}, but many terms obtained would vanish. For J even, the nonzero terms are just

$$\mathsf{C}^{(j_3)} = \left(\frac{[j_3]}{\Omega}\right)^{\frac{1}{2}} \sum_c x_{0,0,c}\{(\mathsf{A}^{(j_1)} \bigcirc^{\frac{1}{2}J - j_3 - c} \mathsf{B}^{(j_2)})\mathsf{U}^c\}, \quad (85)$$

and for J odd, the terms in the parentheses are

$$\{(\mathsf{A}^{(j_1)} \times \bigcirc^{\frac{1}{2}(J-1) - j_3 - c} \mathsf{B}^{(j_2)})\mathsf{U}\},$$

where U is the unit tensor $(\delta_{tt'})$ and the braces indicate symmetrization of indices. Similarly, according to Eq. (84), only coefficients of type $x_{a,b,0}$ are required for re-embedding since $\mathsf{C}^{(j_3)}(t)$ is already traceless in t. Table II illustrates the reduction for $j_1, j_2 = 1, 2$. The first part of the table, which gives the natural forms, agrees with the results of Barsella and Fabri.[4]

The full 3-j tensors (all the coefficients $x_{a,b,c}$) are required for tensor-averaging applications according to Eqs. (19) and (20). For J even, the (single) scalar is often easily determined from the component with all indices equal to z. Writing for the unnormalized 3-j tensors

$$(\mathsf{A}^{(j_1)}\mathsf{B}^{(j_2)}\mathsf{C}^{(j_3)})^{(0)} = \mu\mathsf{T}(j_1, j_2, j_3), \quad (86)$$

where $\mu = \Omega^{-1}\mathsf{T} \vdots \mathsf{A}^{(j_1)}\mathsf{B}^{(j_2)}\mathsf{C}^{(j_3)}$, we have, by Eq. (84),

$$\mu = \left[\sum_{a,b,c} x_{a,b,c}\right]^{-1} \times (A^{(j_1)}(z \cdots z)B^{(j_2)}(z \cdots z)C^{(j_3)}(z \cdots z))^{(0)}. \quad (87)$$

For J odd, from Eq. (85),

$$\mu = j_1 j_3\left[\sum_{a,b,c} x_{a,b,c}\right]^{-1} \times (A^{(j_1)}(yz \cdots z)B^{(j_2)}(zz \cdots z)C^{(j_3)}(xz \cdots z))^{(0)}. \quad (88)$$

Section VIII gives formulas for the required coefficient sums. Appendix (A4) gives a simple example of the application to matrix elements (Cartesian Wigner–Eckart theorem).

TABLE II. Coupling of irreducible natural tensors. $j_1, j_2 \leq 2$. A, B are second-order tensors, **a** and **b** are vectors. Natural coupled forms are written in invariant notations (see Ref. 5): U is the unit tensor; braces imply symmetrized indices. Repeated indices are symmetrized, not summed (unless indicated).

j_1	j_2	j_3	Natural forms $C^{(j_3)}$	Embedded forms $C^{(j_3)}(j_1+j_2)$	Expanded embedded forms[a]
1	1	0	$3^{-\frac{1}{2}}(\mathbf{a}\cdot\mathbf{b})$	$3^{-\frac{1}{2}}C^{(0)}\delta_{rs}$	$\frac{1}{3}(\mathbf{a}\cdot\mathbf{b})\delta_{rs}$
		1	$2^{-\frac{1}{2}}(\mathbf{a}\times\mathbf{b})$	$2^{-\frac{1}{2}}\Sigma_t C^{(1)}_t\epsilon_{rst}$	$\frac{1}{2}(a_rb_s - a_sb_r)$
		2	$\{\mathbf{ab}\} - \frac{1}{3}(\mathbf{a}\cdot\mathbf{b})\mathsf{U}$	$C^{(2)}_{rs}$	$\frac{1}{2}(a_rb_s + a_sb_r) - \frac{1}{3}(\mathbf{a}\cdot\mathbf{b})\delta_{rs}$
2	1	1	$(\frac{3}{5})^{\frac{1}{2}}(\mathsf{A}\cdot\mathbf{b})$	$(\frac{3}{5})^{\frac{1}{2}}(C^{(1)}_r\delta_{rs} - \frac{1}{3}C^{(1)}_s\delta_{rr})$	$\frac{3}{5}(\mathsf{A}\cdot\mathbf{b})_r\delta_{rs} - \frac{1}{5}(\mathsf{A}\cdot\mathbf{b})_s\delta_{rr}$
		2	$(\frac{3}{5})^{\frac{1}{2}}\{\mathsf{A}\times\mathbf{b}\}$	$(\frac{3}{5})^{\frac{1}{2}}\Sigma_t C^{(2)}_{rt}\epsilon_{rst}$	$\frac{2}{3}A_{rr}b_s - \frac{2}{3}A_{rs}b_r + \frac{1}{3}(\mathsf{A}\cdot\mathbf{b})_s\delta_{rr} - \frac{1}{3}(\mathsf{A}\cdot\mathbf{b})_r\delta_{rs}$
		3	$\{\mathsf{Ab} - \frac{2}{5}(\mathsf{A}\cdot\mathbf{b})\mathsf{U}\}$	$C^{(3)}_{rrs}$	$\frac{1}{3}A_{rr}b_s + \frac{2}{3}A_{rs}b_r - \frac{2}{15}(\mathsf{A}\cdot\mathbf{b})_s\delta_{rr}$ $- \frac{4}{15}(\mathsf{A}\cdot\mathbf{b})_r\delta_{rs}$
2	2	0	$5^{-\frac{1}{2}}(\mathsf{A}:\mathsf{B})$	$5^{-\frac{1}{2}}C^{(0)}(\delta^2_{rs} - \frac{1}{3}\delta_{rr}\delta_{ss})$	$\frac{1}{5}(\mathsf{A}:\mathsf{B})(\delta^2_{rs} - \frac{1}{3}\delta_{rr}\delta_{ss})$
		1	$(\frac{3}{5})^{\frac{1}{2}}(\mathsf{A}\cdot\times\mathsf{B})$	$(\frac{3}{5})^{\frac{1}{2}}\Sigma_t C^{(1)}_t\epsilon_{rst}\delta_{rs}$	$\frac{2}{5}[(\mathsf{A}\cdot\mathsf{B})_{rs} - (\mathsf{A}\cdot\mathsf{B})_{sr}]\delta_{rs}$
		2	$(\frac{1\,2}{7})^{\frac{1}{2}}\{\mathsf{A}\cdot\mathsf{B} - \frac{1}{3}(\mathsf{A}:\mathsf{B})\mathsf{U}\}$	$(\frac{1\,2}{7})^{\frac{1}{2}}[C^{(2)}_{rs}\delta_{rs} - \frac{1}{3}(C^{(2)}_{rr}\delta_{ss} + C^{(2)}_{ss}\delta_{rr})]$	$\frac{6}{7}[(\mathsf{A}\cdot\mathsf{B})_{rs} + (\mathsf{A}\cdot\mathsf{B})_{sr}]\delta_{rs}$ $- \frac{4}{7}[(\mathsf{A}\cdot\mathsf{B})_{rr}\delta_{ss} + (\mathsf{A}\cdot\mathsf{B})_{ss}\delta_{rr}]$ $+ (\mathsf{A}:\mathsf{B})(-\frac{4}{7}\delta^2_{rs} + \frac{8}{21}\delta_{rr}\delta_{ss})$
		3	$\{\mathsf{A}\times\mathsf{B}\} - \frac{1}{5}\{(\mathsf{A}\cdot\times\mathsf{B})\mathsf{U}\}$	$\Sigma_t C^{(3)}_{rst}\epsilon_{rst}$	$\frac{1}{2}(A_{rr}B_{ss} - A_{ss}B_{rr})$ $- \frac{2}{5}[(\mathsf{A}\cdot\mathsf{B})_{rs} - (\mathsf{A}\cdot\mathsf{B})_{sr}]\delta_{rs}$
		4	$\{\mathsf{AB} - \frac{4}{7}(\mathsf{A}\cdot\mathsf{B})\mathsf{U} + \frac{2}{35}(\mathsf{A}:\mathsf{B})\mathsf{U}^2\}$	$C^{(4)}_{rrss}$	$\frac{1}{6}(A_{rr}B_{ss} + A_{ss}B_{rr}) + \frac{2}{3}A_{rs}B_{rs}$ $- \frac{4}{21}[(\mathsf{A}\cdot\mathsf{B})_{rs} + (\mathsf{A}\cdot\mathsf{B})_{sr}]\delta_{rs}$ $- \frac{2}{21}[(\mathsf{A}\cdot\mathsf{B})_{rr}\delta_{ss} + (\mathsf{A}\cdot\mathsf{B})_{ss}\delta_{rr}]$ $+ \frac{2}{105}(\mathsf{A}:\mathsf{B})(2\delta^2_{rs} + \delta_{rr}\delta_{ss})$

[a] It may be far from evident, in general, that the sum of the embedded coupled tensors is, in fact, the tensor product of the original tensors, as the completeness relation requires. Using the identity $[A_{rr}B_{ss} + A_{ss}B_{rr} - 2A_{rs}B_{rs} + 2[(\mathsf{A}\cdot\mathsf{B})_{rr}\delta_{ss} + (\mathsf{A}\cdot\mathsf{B})_{ss}\delta_{rr} - (\mathsf{A}\cdot\mathsf{B})_{rs}\delta_{rs} - (\mathsf{A}\cdot\mathsf{B})_{sr}\delta_{rs}] + (\mathsf{A}:\mathsf{B})(\delta_{rs}{}^2 - \delta_{rr}\delta_{ss})] \equiv 0$, $C^{(4)}(4) = (A^2 \otimes B^2)^{(2)}$ can be written in the alternative form $C_{rrss}{}^{(4)} = \frac{1}{2}(A_{rr}B_{ss} + A_{ss}B_{rr}) - \frac{6}{7}[(\mathsf{A}\cdot\mathsf{B})_{rs} + (\mathsf{A}\cdot\mathsf{B})_{sr}]\delta_{rs} + \frac{4}{7}[(\mathsf{A}\cdot\mathsf{B})_{rr}\delta_{ss} + (\mathsf{A}\cdot\mathsf{B})_{ss}\delta_{rr}] + \frac{1}{35}(\mathsf{A}:\mathsf{B})(13\delta_{rs}{}^2 - 11\delta_{rr}\delta_{ss})$. It is then evident for $j_1 = j_2 = 2$ that the sum of the embedded tensors for j_3 even is $\frac{1}{2}(A_{rr}B_{ss} + A_{ss}B_{rr})$, and for j_3 odd, $\frac{1}{2}(A_{rr}B_{ss} - A_{ss}B_{rr})$. The identity is obtained directly by contracting $A_{a_1a_2}B_{b_1b_2}$ with the antisymmetrized (in r_1, s_1, a_1, b_1) tensor product $\Sigma \pm \mathfrak{P}(\delta_{a_1}{}^{a}\delta_{b_1}{}^{b}\delta_{r_1}{}^{r_2}\delta_{s_1}{}^{s_2})$ which vanishes because it is antisymmetric in four indices (impossible for 3-space).

V. 3-*j* SPINORS

For spinors under the rotation group, or more precisely for $SU(2)$, the elementary invariant tensors are δ^{ν}_{μ} and the antisymmetric $\epsilon_{\mu\nu}$ and $\epsilon^{\mu\nu}$.[26] The natural spinors of symmetry j are simply symmetric spinors of order $2j$, and the projections $\mathsf{E}^j(2j \mid \overline{2j}) = (\delta^{\mu'}_{\mu})^{2j}$ are simply symmetrization operators.

The 3-*j* spinors are, therefore, invariant spinors of order $2j_1 + 2j_2 + 2j_3 = 2J$, separately symmetric in the three sets of indices. The completely covariant 3-*j* spinors are, therefore, to within normalization and phase of the form $(\epsilon_{\nu\sigma})^{\alpha}(\epsilon_{\sigma\mu})^{\beta}(\epsilon_{\mu\nu})^{\gamma}$. The exponents satisfy the condition that the total numbers of μ, ν, and σ indices are, respectively, $2j_1$, $2j_2$, and $2j_3$, with the unique solutions $\alpha = J - 2j_1$, etc., nonnegative if and only if the triangle conditions are satisfied. For maximum symmetry, we define

$$\mathsf{T}^{sp}\begin{pmatrix} j_1 & j_2 & j_3 \\ \bar{\mu} & \bar{\nu} & \bar{\sigma} \end{pmatrix} = (\epsilon_{\nu\sigma})^{J-2j_1}(\epsilon_{\sigma\mu})^{J-2j_2}(\epsilon_{\mu\nu})^{J-2j_3} \quad (89)$$

as the unnormalized completely covariant 3-*j* spinors. By successively raising indices with the 1-*j* tensors $g^{j_3}(\sigma, \sigma') = (\epsilon^{\sigma\sigma'})^{2j_3}$, etc., we obtain the maximum-symmetry mixed symbols

$$\mathsf{T}^{sp}\begin{pmatrix} j_1 & j_2 & j_3 \\ \bar{\mu} & \bar{\nu} & \sigma \end{pmatrix} = (-1)^{J-2j_1}(\delta^{\sigma}_{\nu})^{J-2j_1}(\delta^{\sigma}_{\mu})^{J-2j_2}\epsilon^{J-2j_3}_{\mu\nu} \quad (90)$$

and

$$\mathsf{T}^{sp}\begin{pmatrix} j_1 & j_2 & j_3 \\ \bar{\mu} & \nu & \sigma \end{pmatrix} = (-1)^{J-2j_3}(\epsilon^{\nu\sigma})^{J-2j_1}(\delta^{\sigma}_{\mu})^{J-2j_2}(\delta^{\nu}_{\mu})^{J-2j_3}, \quad (91)$$

and the completely contravariant

$$\mathsf{T}^{sp}\begin{pmatrix} j_1 & j_2 & j_3 \\ \mu & \nu & \sigma \end{pmatrix} = (\epsilon^{\nu\sigma})^{J-2j_1}(\epsilon^{\sigma\mu})^{J-2j_2}(\epsilon^{\mu\nu})^{J-2j_3}, \quad (92)$$

the latter numerically identical to the covariant spinor, being just the adjoint of a real quantity. Particularly simple cases are

$$\mathsf{T}^{sp}\begin{pmatrix} j_1 & j_2 & j_1+j_2 \\ \bar{\mu} & \bar{\nu} & \bar{\sigma} \end{pmatrix} = (\epsilon_{\nu\sigma})^{2j_2}(\epsilon_{\sigma\mu})^{2j_1} \quad (93)$$

and

$$\mathsf{T}^{sp}\begin{pmatrix} j & 0 & j \\ \bar{\mu} & 0 & \bar{\sigma} \end{pmatrix} = (\epsilon_{\sigma\mu})^{2j}. \quad (94)$$

The norm of the 3-*j* spinors is easily obtained by direct contraction:

$$\Omega^{sp} = \mathsf{T}^{sp\dagger} \vdots \mathsf{T}^{sp} = \frac{(J+1)!\,(J-2j_1)!\,(J-2j_2)!\,(J-2j_3)!}{(2j_1)!\,(2j_2)!\,(2j_3)!}. \quad (95)$$

A property for $SU(2)$, useful in connection with the tensors under $SO(3)$, is the result of cross contraction (of, say, a μ index with a ν index). There are just

$(J+1)(J-2j_3)$ ways of contracting a μ with a ν in Eq. (92) to a nonnull result, out of a total of $4j_1j_2$ ways. Hence

$$\sum_{\mu_1\nu_1}\epsilon^{\mu_1\nu_1}\mathsf{T}^{sp}\begin{pmatrix}j_1 & j_2 & j_3\\ \bar{\mu} & \bar{\nu} & \bar{\sigma}\end{pmatrix} = \frac{(J+1)(J-2j_3)}{4j_1j_2}\,\mathsf{T}^{sp}\begin{pmatrix}j_1-\frac{1}{2} & j_2-\frac{1}{2} & j_3\\ \bar{\mu} & \bar{\nu} & \bar{\sigma}\end{pmatrix}. \quad (96)$$

VI. INTERRELATIONS

The 3-j spinors are extremely simple, and this simplicity underlies the elegant proofs of spherical tensor properties by the "spinor method." Such derivations have generally been based on the properties of spinor invariants of the type $u_1v_2 - v_1u_2$ rather than on the definition of (Cartesian) 3-j spinors as such. Our normalized 3-j spinors $\mathsf{V}^{sp} = \Omega^{-\frac{1}{2}}\mathsf{T}^{sp}$ are related to the spherical tensor 3-j coefficients with the usual phases by Eqs. (31)–(34), with

$$\mathbf{e}^j_m(\mu) = \mathbf{e}^{jm}(\bar{\mu}) = \begin{pmatrix}2j\\ j-m\end{pmatrix}^{-\frac{1}{2}}(\delta^\mu_1)^{j+m}(\delta^\mu_2)^{j-m}. \quad (97)$$

Equation (31) gives

$$\Omega^{-\frac{1}{2}}\mathsf{T}^{sp}\begin{pmatrix}j_1 & j_2 & j_3\\ \mu & \nu & \sigma\end{pmatrix} = \sum_{m_1m_2m_3}\begin{pmatrix}j_1 & j_2 & j_3\\ m_1 & m_2 & m_3\end{pmatrix} \times \begin{pmatrix}2j_1\\ j_1-m_1\end{pmatrix}^{-\frac{1}{2}}\begin{pmatrix}2j_2\\ j_2-m_2\end{pmatrix}^{-\frac{1}{2}}\begin{pmatrix}2j_3\\ j_3-m_3\end{pmatrix}^{-\frac{1}{2}} \times (\delta^\mu_1)^{j_1+m_1}(\delta^\mu_2)^{j_1-m_1}(\delta^\nu_1)^{j_2+m_2} \times (\delta^\nu_2)^{j_2-m_2}(\delta^\sigma_1)^{j_3+m_3}(\delta^\sigma_2)^{j_3-m_3}, \quad (98)$$

and Eq. (32) gives the inversion. The familiar equation[27] for the 3-j coefficients,

$$(v_1w_2 - w_1v_2)^{J-2j_1}(w_1u_2 - u_1w_2)^{J-2j_2}(u_1v_2 - v_1u_2)^{J-2j_3} = [(J+1)!\,(J-2j_1)!\,(J-2j_2)!\,(J-2j_3)!]^{\frac{1}{2}} \times \sum_{m_1,m_2,m_3}\begin{pmatrix}j_1 & j_2 & j_3\\ m_1 & m_2 & m_3\end{pmatrix}\frac{u_1^{j_1+m_1}u_2^{j_1-m_1}v_1^{j_2+m_2}v_2^{j_2-m_2}w_1^{j_3+m_3}w_2^{j_3-m_3}}{[(j_1+m_1)!\,(j_1-m_1)!\,(j_2+m_2)!\,(j_2-m_2)!\,(j_3+m_3)!\,(j_3-m_3)!]^{\frac{1}{2}}}, \quad (99)$$

is obtained from Eq. (98) by contracting $2j_1$-, $2j_2$-, and $2j_3$-fold, respectively, with first-order spinors u_μ, v_ν, and w_σ, and noting that for each m_i there are $\binom{2j_i}{j_i-m_i}$ identical terms.

For tensors, the analog of Eq. (98) is complicated, because the $\mathbf{e}^j_m(r)$ are complicated. However, the analog of Eq. (99) is quite simple: For $\hat{\mathbf{u}} = \hat{\mathbf{u}}(\theta, \phi)$, a unit vector $\mathbf{e}^{jm}\cdot(\hat{\mathbf{u}})^j = \bar{Y}^{jm}(\theta,\phi)$ is a spherical harmonic normalized according to $\sum_m|\bar{Y}^{jm}|^2 = 2^j(j!)^2/(2j)!$.[28] By contracting Eq. (31) with $(\hat{\mathbf{u}})^{j_1}(\hat{\mathbf{v}})^{j_2}(\hat{\mathbf{w}})^{j_3}$ and noting that trace terms such as δ_{rr} give $\hat{\mathbf{u}}\cdot\hat{\mathbf{u}} = 1$, we obtain for J even

$$\Omega^{-\frac{1}{2}}\sum_{a,b,c}x_{a,b,c}\cos\theta_{23}^{\frac{1}{2}J-j_1+a-b-c} \times \cos\theta_{31}^{\frac{1}{2}J-j_2-a+b-c}\cos\theta_{12}^{\frac{1}{2}J-j_3-a-b+c} = \sum_{m_1m_2m_3}\begin{pmatrix}j_1 & j_2 & j_3\\ m_1 & m_2 & m_3\end{pmatrix} \times \bar{Y}^{j_1m_1}(\theta_1\phi_1)\,\bar{Y}^{j_2m_2}(\theta_2\phi_2)\,\bar{Y}^{j_3m_3}(\theta_3\phi_3), \quad (100)$$

which can be regarded as a generalization of the addition theorem for spherical harmonics. It reduces to the addition theorem for $j_2 = 0$, $\hat{\mathbf{u}} = \hat{\mathbf{w}}$. For J odd, the analogous equation differs by the change $\frac{1}{2}J \to \frac{1}{2}(J-1)$ and a volume factor $[\hat{\mathbf{u}}\times\hat{\mathbf{v}}\cdot\hat{\mathbf{w}}]$ on the left. For $\hat{\mathbf{u}} = \hat{\mathbf{v}} = \hat{\mathbf{w}}$, Eq. (100) gives the useful formula

$$\sum_{a,b,c}x_{a,b,c} = (-1)^{\frac{1}{2}J}\Omega^{\frac{1}{2}}(N_{j_1}N_{j_2}N_{j_3})^{-\frac{1}{2}}\begin{pmatrix}j_1 & j_2 & j_3\\ 0 & 0 & 0\end{pmatrix}, \quad (101)$$

where $N_j = (2j)!/[2^j(j!)^2]$.

The 3-j tensors can be related to the 3-j spinors by substituting spinor expressions for the elementary invariant tensors U and ϵ[28]:

$$\delta_{rs} \sim (\epsilon_{\mu,\nu})^2,$$
$$\delta_{rr'} \sim (\epsilon_{\mu,\mu'})^2,$$
$$\epsilon_{rst} \sim (\sqrt{2})\,\epsilon_{\mu\nu}\epsilon_{\nu\sigma}\epsilon_{\sigma\mu} = (\sqrt{2})\,\mathsf{T}^{sp}\begin{pmatrix}1 & 1 & 1\\ \mu & \nu & \sigma\end{pmatrix}. \quad (102)$$

In $(\epsilon_{\mu\mu'})^2$, the two μ indices are symmetrized as are the two μ', but not the μ and μ'. Substitution in Eqs. (82)–(83) must give expressions proportional to the corresponding (integral j) 3-j spinors, which are therefore invariant to further symmetrization of the spinor indices. Symmetrization, however, annihilates all the terms originally involving δ_{rr}, δ_{ss}, or δ_{tt}, while making the leading term into the form of Eq. (89). Thus we

have the isomorphism

$$\mathsf{T}\begin{pmatrix} j_1 & j_2 & j_3 \\ r & s & t \end{pmatrix} \sim \mathsf{T}^{sp}\begin{pmatrix} j_1 & j_2 & t_3 \\ \mu & \nu & \sigma \end{pmatrix}, \qquad J \text{ even},$$
$$\mathsf{T}\begin{pmatrix} j_1 & j_2 & j_3 \\ r & s & t \end{pmatrix} \sim (\sqrt{2})\,\mathsf{T}^{sp}\begin{pmatrix} j_1 & j_2 & j_3 \\ \mu & \nu & \sigma \end{pmatrix}, \quad J \text{ odd}. \tag{103}$$

The factor $\sqrt{2}$ for J odd arises from Eq. (102).

These relations give the normalizing factors for the 3-j tensors:

$$\Omega\,(\text{tensor}) = \Omega\,(\text{spinor}), \quad \text{for } J \text{ even},$$
$$\Omega\,(\text{tensor}) = 2\Omega\,(\text{spinor}), \quad \text{for } J \text{ odd}, \tag{104}$$

in conjunction with Eq. (95). They also give the formula for cross contraction of 3-j tensors. By double application of Eq. (96), we have for the tensors that

$$\sum_{r_1 s_1} \delta_{r_1 s_1} \mathsf{T}\begin{pmatrix} j_1 & j_2 & j_3 \\ r & s & t \end{pmatrix} = \frac{J(J+1)(J-2j_3)(J-2j_3-1)}{4j_1 j_2(2j_1-1)(2j_2-1)} \times \mathsf{T}\begin{pmatrix} j_1-1 & j_2-1 & j_3 \\ r & s & t \end{pmatrix}, \tag{105}$$

which is valid for J even or odd and useful for deriving recursion formulas. Equation (105), and hence Ω, may also be obtained directly, though less easily, from Eqs. (81) and (82) and Table VI.

VII. HIGHER n-j SYMBOLS

The higher n-j symbols are scalar quantities and therefore essentially the same in Cartesian tensor, spinor, and spherical tensor theory. With the maximum symmetry, 3-j tensors and spinors of Eqs. (71)–(72) and (89)–(92), normalized to unity, and with definitions of higher n-j symbols analogous to the spherical tensor definitions, the higher n-j symbols coincide with the usual ones. Thus, for tensors or spinors,

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ j_4 & j_5 & j_6 \end{Bmatrix} = \sum_{r,s,t,u,v,w} \mathsf{V}\begin{pmatrix} j_1 & j_2 & j_3 \\ r & s & t \end{pmatrix}\mathsf{V}\begin{pmatrix} j_1 & j_5 & j_6 \\ \bar{r} & \bar{v} & w \end{pmatrix} \times \mathsf{V}\begin{pmatrix} j_4 & j_2 & j_6 \\ u & \bar{s} & \bar{w} \end{pmatrix}\mathsf{V}\begin{pmatrix} j_4 & j_5 & j_3 \\ \bar{u} & v & \bar{t} \end{pmatrix}, \tag{106}$$

with $\mathsf{V} = \Omega^{-\frac{1}{2}}\mathsf{T}$.

It is straightforward to verify Eq. (106) for some 3-j tensors in Table I by carrying out the contractions with the help of the relation $\sum_r \epsilon_{rst}\epsilon_{ruv} = \delta_{su}\delta_{tv} - \delta_{sv}\delta_{tu}$. For example,

$$\tfrac{1}{3\cdot 6}\sum_{r,s,t,u,v,w} \epsilon_{rst}\epsilon_{rvw}\epsilon_{usw}\epsilon_{uvt} = \tfrac{1}{6} = \begin{Bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{Bmatrix}, \tag{107}$$

in agreement with $\mathsf{V}(1,1,1) = \Omega^{-\frac{1}{2}}\mathsf{T}(1,1,1) = 6^{-\frac{1}{2}}\epsilon_{rst}$. Indeed, the relative ease of carrying out such contractions, compared to analogous summations over *numerically specified* spherical 3-j symbols, illustrates the point that formal recoupling theory is not always a necessity in low-order Cartesian-tensor calculations, although valuable, even for $j = 2$.

The n-j symbols can also be represented in an interesting way as complete contractions between n projections E^j. Thus, if $(j_a + j_b + j_c)$ is even for all four triads, Eqs. (75) and (106) give

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ j_4 & j_5 & j_6 \end{Bmatrix} = [\Omega_{123}\Omega_{156}\Omega_{426}\Omega_{453}]^{-\frac{1}{2}} \times \sum_{\text{all indices}} \mathsf{E}^{j_1}(r_2^{13}r_3^{12} \mid r_5^{16}r_6^{15}) \times \mathsf{E}^{j_2}(r_1^{23}r_3^{21} \mid r_4^{26}r_6^{24}) \times \mathsf{E}^{j_3}(r_1^{32}r_2^{31} \mid r_4^{35}r_5^{34})\mathsf{E}^{j_4}(r_2^{46}r_6^{42} \mid r_5^{43}r_3^{45}) \times \mathsf{E}^{j_5}(r_1^{56}r_6^{51} \mid r_4^{53}r_3^{54})\mathsf{E}^{j_6}(r_2^{64}r_4^{62} \mid r_1^{65}r_5^{61}), \tag{108}$$

where $r_a^{bc} \equiv r_a^{cb}$ denotes a set of $\frac{1}{2}(j_b + j_c - j_a)$ dummy indices. For triads with $(j_a + j_b + j_c)$ odd, there are cross products as given by Eq. (76).

The n-j symbols have the usual recoupling significance. For example, if $\mathsf{A}^{(j_\mu)}$ and $\mathsf{B}^{(k_\nu)}$ are irreducible tensors which can be separately averaged (e.g., tensor operators in different Hilbert spaces), and $[\mathsf{A}^{(j_\mu)} \otimes \mathsf{B}^{(j_\nu)}]^{(l_\sigma)}$ their coupling to symmetry l_σ according to Eq. (83), then averaging over both the A's and B's, we get

$$([\mathsf{A}^{(j_1)} \otimes \mathsf{B}^{(k_1)}]^{(l_1)}[\mathsf{A}^{(j_2)} \otimes \mathsf{B}^{(k_2)}]^{(l_2)}[\mathsf{A}^{(j_3)} \otimes \mathsf{B}^{(k_3)}]^{(l_3)})^{(0)} = \lambda\Omega^{-\frac{1}{2}}_{l_1 l_2 l_3}\mathsf{T}(l_1, l_2, l_3), \tag{109}$$

where

$$\lambda = [(2l_1+1)(2l_2+1)(2l_3+1)]^{\frac{1}{2}}\begin{Bmatrix} j_1 & j_2 & j_3 \\ k_1 & k_2 & k_3 \\ l_1 & l_2 & l_3 \end{Bmatrix}\lambda_A\lambda_B, \tag{110}$$

where

$$\lambda_A = \Omega^{-\frac{1}{2}}_{j_1 j_2 j_3}\mathsf{T}(j_1, j_2, j_3) \vdots\, \mathsf{A}^{(j_1)}\mathsf{A}^{(j_2)}\mathsf{A}^{(j_3)}, \tag{111}$$

with a corresponding expression for λ_B, analogous to the relation between reduced matrix elements of commuting operators.

VIII. GENERAL FORMULAS FOR 3-j TENSORS

A. $J = j_1 + j_2 + j_3$ Even

The 3-j tensors are determined to within normalization by the fact that they are invariant tensors separately traceless and symmetric in the three sets of indices. By requiring single contractions of Eq.

TABLE III. General formulas for the 3-j coefficients, $J = j_1 + j_2 + j_3$ even.

$$c_p(j) = x\begin{pmatrix} j & 0 & j \\ p & 0 & p \end{pmatrix} = (-1)^p \frac{(j!)^2(2j-2p)!}{(2j)!\,p!\,(j-p)!\,(j-2p)!},$$

$$x_{a,0,0} = (-1)^a \frac{2^a j_1!\,(2j_1-2a)!\,(\tfrac{1}{2}J-j_2)!\,(\tfrac{1}{2}J-j_3)!}{a!\,(2j_1)!\,(j_1-a)!\,(\tfrac{1}{2}J-j_2-a)!\,(\tfrac{1}{2}J-j_3-a)!},$$

$$x_{a,b,0} = (-1)^b \frac{2^{b-a} j_1!\,j_2!\,(2j_2-2b)!\,(\tfrac{1}{2}J-j_2)!\,(\tfrac{1}{2}J-j_3)!}{(2j_1)!\,(2j_2)!\,(j_2-b)!\,(\tfrac{1}{2}J-j_1+a-b)!\,(\tfrac{1}{2}J-j_3-a-b)!} \times \sum_\lambda (-1)^\lambda \frac{2^{2\lambda}(2j_1-2\lambda)!\,(\tfrac{1}{2}J-j_1+\lambda)!}{\lambda!\,(a-\lambda)!\,(b-a+\lambda)!\,(j_1-\lambda)!\,(\tfrac{1}{2}J-j_2-\lambda)!},$$

$$x_{a,b,c} = (-1)^c \frac{2^{c-a-b} j_1!\,j_2!\,j_3!\,(2j_3-2c)!\,(\tfrac{1}{2}J-j_2)!\,(\tfrac{1}{2}J-j_3)!}{(2j_1)!\,(2j_2)!\,(2j_3)!\,(j_3-c)!\,(\tfrac{1}{2}J-j_1+a-b-c)!\,(\tfrac{1}{2}J-j_2-a+b-c)!} \times \sum_{\lambda,\mu}\left(\frac{(-1)^\mu 2^{2\mu}(2j_2-2\mu)!\,(\tfrac{1}{2}J-j_2-\lambda+\mu)!}{(j_2-\mu)!\,(a-\lambda)!\,(b-\mu)!\,(c-a-b+\lambda+\mu)!\,(\tfrac{1}{2}J-j_3-\lambda-\mu)!} \times \sum_\nu (-1)^\nu 2^{2\nu} \frac{(2j_1-2\nu)!\,(\tfrac{1}{2}J-j_1+\nu)!}{\nu!\,(\lambda-\nu)!\,(\mu-\lambda+\nu)!\,(j_1-\nu)!\,(\tfrac{1}{2}J-j_2-\nu)!}\right).$$

Sum: $$\sum_{a,b,c} x_{a,b,c} = 2^{\frac{1}{2}J}(\tfrac{1}{2}J)! \frac{j_1!\,j_2!\,j_3!\,(J-2j_1)!\,(J-2j_2)!\,(J-2j_3)!}{(2j_1)!\,(2j_2)!\,(2j_3)!\,(\tfrac{1}{2}J-j_1)!\,(\tfrac{1}{2}J-j_2)!\,(\tfrac{1}{2}J-j_3)!}.$$

Normalization: $$\Omega = \mathsf{T} : \mathsf{T} = \frac{(J+1)!\,(J-2j_1)!\,(J-2j_2)!\,(J-2j_3)!}{(2j_1)!\,(2j_2)!\,(2j_3)!}.$$

(81) with respect to r, s, and t to vanish, we obtain the following set of recursion formulas for the coefficients $x_{a,b,c}$[29]:

$$\begin{aligned}
&(a+1)(2j_1-2a-1)x_{a+1,b,c} + \beta\gamma x_{a,b,c} \\
&\quad + \tfrac{1}{2}\beta(\beta+1)x_{a,b,c-1} + \tfrac{1}{2}\gamma(\gamma+1)x_{a,b-1,c} = 0, \\
&(b+1)(2j_2-2b-1)x_{a,b+1,c} + \alpha\gamma x_{a,b,c} \\
&\quad + \tfrac{1}{2}\alpha(\alpha+1)x_{a,b,c-1} + \tfrac{1}{2}\gamma(\gamma+1)x_{a-1,b,c} = 0, \\
&(c+1)(2j_3-2c-1)x_{a,b,c+1} + \alpha\beta x_{a,b,c} \\
&\quad + \tfrac{1}{2}\alpha(\alpha+1)x_{a,b-1,c} + \tfrac{1}{2}\beta(\beta+1)x_{a-1,b,c} = 0,
\end{aligned} \tag{112}$$

where $\alpha = \frac{1}{2}J - j_1 + a - b - c$, $\beta = \frac{1}{2}J - j_2 - a + b - c$, $\gamma = \frac{1}{2}J - j_3 - a - b + c$. The direct solution of these would be difficult in general, but for two special cases it is simple. For $j_1 = j_3 = j$ and $j_2 = 0$, the "Legendre" coefficients $c_a(j)$ of the projections E^j are obtained, and for $a = b = 0$ we have

$$(c+1)(2j_3-2c-1)x_{0,0,c+1} = -\alpha\beta x_{0,0,c}, \tag{113}$$

with the solution, which agrees with the results of Barsella and Fabri,[4] given in Table III.

To obtain $x_{a,b,c}$ in the general case, we start from the explicit expansion of $\mathsf{T}(j_1, j_2, j_1+j_2)$ as given by Eq. (77). We write[23]

$$\mathsf{E}^{j_3}(t_1 \cdots t_{j_3} \mid r_1 \cdots r_\alpha s_1 \cdots s_\beta) = \sum_{p,x,y} c_p(j_3) N_{x,y}(\alpha\beta \mid p) \delta^\kappa_{st} \delta^\sigma_{tr} \delta^\tau_{rs} \delta^p_{tt} \delta^x_{rr} \delta^y_{ss}, \tag{114}$$

where $\kappa = \beta - p + x - y$, $\sigma = \alpha - p - x + y$, $\tau = p - x - y$, and the $c_p(j_3)$ are the coefficients for the projection E^{j_3} [Eq. (67)]. Here $N_{x,y}(\alpha\beta \mid p)$ is equal to the number of ways of selecting x pairs of type A–A, y of type B–B, and $p - x - y$ of type A–B from totals of α indistinguishable objects A and β indistinguishable objects B, normalized such that

$$\sum_{x,y} N_{x,y}(\alpha\beta \mid p) = 1, \tag{115}$$

i.e., divided by the number of ways of selecting p pairs when A and B are not distinguishable:

$$N_{x,y}(\alpha\beta \mid p) = \frac{\alpha!\,\beta!\,p!\,(\alpha+\beta-2p)!}{(\alpha+\beta)!}\left(\frac{2^{p-x-y}}{(\alpha-p-x+y)!\,(\beta-p+x-y)!\,(p-x-y)!\,x!\,y!}\right). \tag{116}$$

In particular, $N_{0,0}(\alpha\beta \mid p)$ gives the relation

$$x_{0,0,p} = c_p(j_3) N_{0,0}(\tfrac{1}{2}J - j_2, \tfrac{1}{2}J - j_1 \mid p) \quad \text{for } J \text{ even}, \tag{117}$$

between the coefficients c_p for the projections and the coefficients $x_{0,0,p}$ for the coupling of irreducible tensors to natural form.[30]

Now according to Eq. (75), we can write, in general, that

$$\mathsf{T}(j_1 j_2 j_3) = \sum_{\lambda,\mu} x_{\lambda,\mu,0} \mathsf{E}^{j_3}(t_1 \cdots t_{j_3} \mid r_1 \cdots r_\alpha s_1 \cdots s_\beta) \times \delta^{\frac{1}{2}J - j_3 - \lambda - \mu}_{rs} \delta^\lambda_{rr} \delta^\mu_{ss}, \tag{118}$$

where $\alpha = \frac{1}{2}J - j_2 - \lambda + \mu$, $\beta = \frac{1}{2}J - j_1 + \lambda - \mu$.

TABLE IV. General formulas for the 3-j coefficients, $J = j_1 + j_2 + j_3$ odd.[a]

$$x_{a,b,c} = (-1)^c \frac{2^{c-a-b} j_1!\, j_2!\, j_3!\, (2j_3 - 2c)!\, (\frac{1}{2}[J-1] - j_2)!\, (\frac{1}{2}[J-1] - j_3)!}{(2j_1)!\, (2j_2)!\, (2j_3)!\, (j_3 - c)!\, (\frac{1}{2}[J-1] - j_1 + a - b - c)!\, (\frac{1}{2}[J-1] - j_2 - a + b - c)!}$$
$$\times \sum_{\lambda,\mu} \left((-1)^\mu \frac{2^{2\mu} (2j_2 - 2\mu)!\, (\frac{1}{2}[J-1] - j_2 - \lambda + \mu)!}{(j_2 - \mu)!\, (a - \lambda)!\, (b - \mu)!\, (c - a - b + \lambda + \mu)!\, (\frac{1}{2}[J-1] - j_3 - \lambda - \mu)!} \right.$$
$$\left. \times \sum_\nu (-1)^\nu \frac{2^{2\nu} (2j_1 - 2\nu)!\, (\frac{1}{2}[J-1] - j_1 + \nu)!}{\nu!\, (\lambda - \nu)!\, (\mu - \lambda + \nu)!\, (j_1 - \nu)!\, (\frac{1}{2}[J-1] - j_2 - \nu)!} \right).$$

Sum: $$\sum_{a,b,c} x_{a,b,c} = 2^{\frac{1}{2}(J+1)} (\tfrac{1}{2}[J+1])! \frac{j_1!\, j_2!\, j_3!\, (J-2j_1)!\, (J-2j_2)!\, (J-2j_3)!}{(2j_1)!\, (2j_2)!\, (2j_3)!\, (\frac{1}{2}[J-1] - j_1)!\, (\frac{1}{2}[J-1] - j_2)!\, (\frac{1}{2}[J-1] - j_3)!}.$$

Normalization: $$\Omega = \mathsf{T} \vdots \mathsf{T} = 2 \frac{(J+1)!\, (J-2j_1)!\, (J-2j_2)!\, (J-2j_3)!}{(2j_1)!\, (2j_2)!\, (2j_3)!}.$$

[a] The formula for $x_{a,b,c}$ is related to that for J even by the substitution $J \to J - 1$, but the formulas for Ω and $\Sigma\, x_{a,b,c}$ are not similarly related.

The expansion of E^{j_3} in Eq. (118) according to Eq. (114) and the comparison of the result to Eq. (81) give an equation

$$x_{a,b,c} = c_c(j_3) \sum_{\lambda,\mu} x_{\lambda,\mu,0} \times N_{a-\lambda,b-\mu}(\tfrac{1}{2}J - j_2 - \lambda + \mu, \tfrac{1}{2}J - j_1 + \lambda - \mu \mid c), \quad (119)$$

relating the general coefficients to coefficients with one subscript zero. By taking b, for example, equal to zero and noting that $N_{a-\lambda,-\mu}$ is zero unless μ, necessarily nonnegative, is zero, we relate the coefficients with one subscript zero to those with two subscripts zero:

$$x_{a,b,0} = \sum_\lambda x_{\lambda,0,0} c_b(j_2) \times N_{a-\lambda,0}(\tfrac{1}{2}J - j_3 - \lambda, \tfrac{1}{2}J - j_1 + \lambda \mid b). \quad (120)$$

The combination of Eqs. (119) and (120) gives the general formula

$$x_{a,b,c} = c_c(j_3) \sum_{\lambda,\mu,\nu} x_{\lambda,0,0} c_\nu(j_2) \times N_{a-\mu,b-\nu}(\tfrac{1}{2}J - j_2 - \mu + \nu, \tfrac{1}{2}J - j_1 + \mu - \nu \mid c) \times N_{\mu-\lambda,0}(\tfrac{1}{2}J - j_3 - \lambda, \tfrac{1}{2}J - j_1 + \lambda \mid \nu). \quad (121)$$

Substitution of the expressions for c_p, $x_{\lambda,0,0}$, and $N_{x,y}(\alpha\beta \mid p)$ from Eq. (67), Table III, and Eq. (116) gives the explicit formulas in Table III. These are unsymmetrical and appear simplest to use with $j_1 < j_2 < j_3$. Some particular cases are given in Tables V and VI.

The coefficient sum $\sum_{a,b,c} x_{a,b,c}$, useful in connection with Eq. (87) and for numerical checks, and given by Eq. (101) or directly from the recursion formulas of Sec. VIIIC, is written explicitly in Table III. From Eqs. (115) and (119) we have, further, that

$$\sum_{a,b} x_{a,b,0} = \left(\sum_c c_c \right)^{-1} \sum_{a,b,c} x_{a,b,c} = N_{j_3} \sum_{a,b,c} x_{a,b,c}, \quad (122)$$

where $N_j = (2j)!/[2^j (j!)^2]$, and

$$\sum_{a,b} x_{a,b,c} = c_c \sum_{a,b} x_{a,b,0}. \quad (123)$$

B. $J = j_1 + j_2 + j_3$ Odd

The condition that the right side of Eq. (82) is traceless in r, s, and t leads to recursion formulas for the coefficients for J odd, identical to Eqs. (112) except for the change $\frac{1}{2}J \to \frac{1}{2}(J-1)$ in α, β, and γ. The formulas for the $x_{a,b,c}$ are, therefore, closely similar to those for J even and, indeed, are obtained from the latter *as written in Table III*, simply by the substitution $J \to J - 1$.

The proof is similar to that for J even. We consider the expansion of $\mathsf{T}(j_1, j_2, j_1 + j_2 - 1)$ in the form

$$\sum_k \epsilon_{rsk} \mathsf{E}^{j_3}(k r_1 \cdots r_{j_1-1} s_1 \cdots s_{j_1-1} \mid t_1 \cdots t_{j_3}) = \sum_{p,x,y} c_p(j_3) M_{x,y}(j_1 - 1, j_2 - 1 \mid p) \epsilon_{rst} \delta^\sigma_{rt} \delta^\kappa_{st} \delta^\tau_{rs} \delta^p_{tt} \delta^x_{rr} \delta^y_{ss}, \quad (124)$$

where $\sigma = j_1 - 1 - p - x + y$, $\kappa = j_2 - 1 - p + x - y$, $\tau = p - x - y$, and where $M_{x,y}$ is related to the previous $N_{x,y}$ by[31]

$$M_{x,y}(\alpha\beta \mid p) = [(\alpha + \beta - 2p + 1)/(\alpha + \beta + 1)] \times N_{x,y}(\alpha\beta \mid p). \quad (125)$$

The expansion of

$$\mathsf{T}(j_1, j_2, j_3) = \sum_{\lambda,\mu,k} x_{\lambda,\mu,0} \epsilon_{rsk} \mathsf{E}^{j_3}(k r_1 \cdots r_\alpha s_1 \cdots s_\beta \mid t) \times \delta^{\frac{1}{2}(J-1)-j_3-\lambda-\mu}_{rs} \delta^\lambda_{rr} \delta^\mu_{ss}, \quad (126)$$

TABLE V. Special cases: $j_3 = j_1 + j_2 - p$ for $p = 0, 1, 2, 3$.

Case	$x_{a,b,c}$	$\Omega \equiv \mathsf{T} \vdots \mathsf{T}$	$\sum x_{a,b,c}$ [a]
$j_3 = j_1 + j_2$	$(-1)^c \dfrac{2^{c-a-b} j_1!\, j_2!\, j_3!\, (2j_3 - 2c)!}{(j_1 - a + b - c)!\, (j_2 + a - b - c)!\, (j_3 - c)!\, (2j_3)!\, a!\, b!\, (c - a - b)!}$	$2j_3 + 1$	$N_{j_3}^{-1}$
$j_3 = j_1 + j_2 - 1$	$(-1)^c \dfrac{2^{c-a-b}(j_1 - 1)!\, (j_2 - 1)!\, j_3!\, (2j_3 - 2c)!}{(j_1 - a + b - c - 1)!\, (j_2 + a - b - c - 1)!\, (j_3 - c)!\, (2j_3)!\, a!\, b!\, (c - a - b)!}$	$\dfrac{(2j_3 + 1)(2j_3 + 2)}{2j_1 j_2}$	$\frac{1}{2}(j_3 + 1) N_{j_3}^{-1}$
$j_3 = j_1 + j_2 - 2$	$\left\{(-1)^c \dfrac{2^{c-a-b}(j_1 - 1)!\, (j_2 - 1)!\, j_3!\, (2j_3 - 2c)!}{(j_1 - a + b - c - 1)!\, (j_2 + a - b - c - 1)!\, (j_3 - c)!\, (2j_3)!\, a!\, b!\, (c - a - b + 1)!} \times \left[(c - a - b + 1) - \dfrac{2aj_2}{2j_1 - 1} - \dfrac{2bj_1}{2j_2 - 1}\right]\right\}$	$\dfrac{(2j_3 + 1)(2j_3 + 2)(2j_3 + 3)}{2j_1 j_2 (2j_1 - 1)(2j_2 - 1)}$	$\dfrac{(j_3 + 1)}{(2j_1 - 1)(2j_2 - 1)} N_{j_3}^{-1}$
$j_3 = j_1 + j_2 - 3$	$\left\{(-1)^c \dfrac{2^{c-a-b}(j_1 - 2)!\, (j_2 - 2)!\, j_3!\, (2j_3 - 2c)!}{(j_1 - a + b - c - 2)!\, (j_2 + a - b - c - 2)!\, (j_3 - c)!\, 2j_3!\, a!\, b!\, (c - a - b + 1)!} \times \left[(c - a - b + 1) - \dfrac{2a(j_2 - 1)}{2j_1 - 1} - \dfrac{2b(j_1 - 1)}{2j_2 - 1}\right]\right\}$	$\dfrac{3(2j_3 + 1)(2j_3 + 3)(j_3 + 1)(j_3 + 2)}{j_1(j_1 - 1)(2j_1 - 1)j_2(j_2 - 1)(2j_2 - 1)}$	$\dfrac{3(j_3 + 1)(j_3 + 2)}{2(2j_1 - 1)(2j_2 - 1)} N_{j_3}^{-1}$

[a] $N_j = (2j)!/2^j(j!)^2$ is the leading coefficient of the Legendre polynomials.

where $\alpha = \frac{1}{2}(J-1) - j_2 - \lambda + \mu$, $\beta = \frac{1}{2}(J-1) - j_1 + \lambda - \mu$, leads to

$$x_{a,b,c} = \sum_{\lambda,\mu} c_c x_{\lambda,\mu,0} \times M_{a-\lambda,b-\mu}(\tfrac{1}{2}[J-1] - j_2 - \lambda + \mu, \tfrac{1}{2}[J-1] - j_1 + \lambda - \mu \mid c) \quad (127)$$

and

$$x_{a,b,0} = \sum_{\lambda} x_{\lambda,0,0} c_b \times M_{a-\lambda,0}(\tfrac{1}{2}[J-1] - j_3 - \lambda, \tfrac{1}{2}[J-1] - j_1 + \lambda \mid b), \quad (128)$$

the analogs of Eqs. (119) and (120). The final formula is given in Table IV.

Coefficient sums can be determined from the tensor recursion formulas (Sec. VIIIC), or indirectly from the correspondence to spherical tensors [Eq. (34)],[28] which yields

$$\sum x_{a,b,c} = (-1)^{\frac{1}{2}(J+1)} \left[\frac{j_1(j_1+1) j_3(j_3+1)}{N_{j_1} N_{j_2} N_{j_3}} \Omega\right]^{\frac{1}{2}} \times \begin{pmatrix} j_1 & j_2 & j_3 \\ -1 & 0 & 1 \end{pmatrix}, \quad J \text{ odd}, \quad (129)$$

where $N_j = (2j)!/[2^j (j!)^2]$. Table IV records $\sum x_{a,b,c}$. From Eq. (127), we have also[31]

$$\sum_{a,b} x_{a,b,0} = \frac{2}{j_3 + 1} N_{j_3} \sum x_{abc} \quad (130)$$

and

$$\sum_{a,b} x_{a,b,c} = \frac{j_3 - 2c}{j_3} c_c(j_3) \sum_{a,b} x_{a,b,0}. \quad (131)$$

C. Recursion Formulas: J Even or Odd

Application of the projections E^{j_1+1} and E^{j_2+1} to the tensor product $\mathsf{T}(j_1, j_2; j_3) \otimes \mathsf{U}$ where $\mathsf{U} = (\delta_{rs})$, together with Eq. (105), gives an equation for $\mathsf{T}(j_1+1, j_2+1, j_3)$:

$$\begin{aligned} &\mathsf{T}\begin{pmatrix} j_1+1 & j_2+1 & j_3 \\ r & s & t \end{pmatrix} \\ &= \delta_{rs} \mathsf{T}\begin{pmatrix} j_1 & j_2 & j_3 \\ r & s & t \end{pmatrix} - \frac{j_1}{(2j_1+1)} \delta_{rr} \mathsf{T}\begin{pmatrix} j_1 & j_2 & j_3 \\ sr\cdots r & s & t \end{pmatrix} \\ &\quad - \frac{j_2}{(2j_2+1)} \delta_{ss} \mathsf{T}\begin{pmatrix} j_1 & j_2 & j_3 \\ r & rs\cdots s & t \end{pmatrix} \\ &\quad + \frac{J(J+1)(J-2j_3)(J-2j_3-1)}{4(2j_1+1)(2j_1-1)(2j_2+1)(2j_2-1)} \\ &\quad \times \delta_{rr}\delta_{ss} \mathsf{T}\begin{pmatrix} j_1-1 & j_2-1 & j_3 \\ r & s & t \end{pmatrix}. \end{aligned} \quad (132)$$

The symmetrization convention is used, and the notation $sr\cdots r$ means that one r index is replaced

TABLE VI. Particular 3-j coefficients, $(a+b+c) \leq 3$. The parameter x equals $\frac{1}{2}J$ for $J = (j_1+j_2+j_3)$ even, $\frac{1}{2}(J-1)$ for J odd.

Coefficient	Value
$x_{0,0,0}$	1
$x_{1,0,0}$	$-\dfrac{(x-j_2)(x-j_3)}{(2j_1-1)}$
$x_{2,0,0}$	$\dfrac{(x-j_2)(x-j_2-1)(x-j_3)(x-j_3-1)}{2(2j_1-1)(2j_1-3)}$
$x_{1,1,0}$	$-\dfrac{(x-j_3)(x-j_3-1)}{2(2j_1-1)(2j_2-1)}[(2j_1+2j_2-2x-1)-2(x-j_1)(x-j_2)]$
$x_{3,0,0}$	$-\dfrac{(x-j_2)(x-j_2-1)(x-j_2-2)(x-j_3)(x-j_3-1)(x-j_3-2)}{6(2j_1-1)(2j_1-3)(2j_1-5)}$
$x_{2,1,0}$	$\dfrac{(x-j_2)(x-j_3)(x-j_3-1)(x-j_3-2)}{2(2j_1-1)(2j_1-3)(2j_2-1)}[(2j_2-1)-(x-j_1)(x-j_2+1)]$
$x_{1,1,1}$	$\dfrac{(x-j_1)(x-j_2)(x-j_3)}{2(2j_1-1)(2j_2-1)(2j_3-1)}[(x-j_1)(j_2+j_3-1)+(x-j_2)(j_1+j_3-1)+(x-j_3)(j_1+j_2-1)$ $-2(x-j_1)(x-j_2)(x-j_3)+2x-1]$

by an s index. Equation (132) can be used as a recursion formula, but the substitutions $r \to s$ and $s \to r$ are inconvenient. Another relation of similar type,

$$\mathsf{T}\begin{pmatrix} j_1-1 & j_2+1 & j_3 \\ r & s & t \end{pmatrix} = \frac{2j_1(2j_1-1)}{(J-2j_2)(J-2j_2-1)}\mathsf{T}\begin{pmatrix} j_1 & j_2 & j_3 \\ sr\cdots r & s & t \end{pmatrix} - \frac{J(J+1)(J-2j_3)(J-2j_3-1)}{2(2j_2+1)(2j_2-1)(J-2j_2)(J-2j_2-1)} \times \delta_{ss}\mathsf{T}\begin{pmatrix} j_1-1 & j_2-1 & j_3 \\ r & s & t \end{pmatrix}, \tag{133}$$

is obtained by choosing the coefficients to make the right side traceless in s and to have the same leading term as the left side. The coefficients in (132) and (133) are the same for J even and odd.[32] The combination of Eqs. (132) and (133) gives a recursion formula not involving the substitutions $r \leftrightarrow s$:

$$\begin{aligned}\mathsf{T}(j_1+1, j_2+1, j_3) &= \mathsf{T}(j_1, j_2, j_3)\delta_{rs} \\ &- \frac{(J-2j_2)(J-2j_2-1)}{2(2j_1+1)(2j_1-1)}\delta_{rr}\mathsf{T}(j_1-1, j_2+1, j_3) \\ &- \frac{(J-2j_1)(J-2j_1-1)}{2(2j_2+1)(2j_2-1)}\delta_{ss}\mathsf{T}(j_1+1, j_2-1, j_3) \\ &- \frac{J(J+1)(J-2j_3)(J-2j_3-1)}{4(2j_1+1)(2j_1-1)(2j_2+1)(2j_2-1)} \\ &\times \delta_{rr}\delta_{ss}\mathsf{T}(j_1-1, j_2-1, j_3),\end{aligned} \tag{134}$$

again valid for J even or odd. In the case $j_1 = j_2 = j$ and $j_3 = 0$, this reduces to the recursion formula of Paper II for the projections E^j, that is, essentially to the recursion formula for Legendre polynomials.

Note added in proof: In the recursion formulas, $\mathsf{T}(j_1, j_2, j_3)$ should be taken as zero if $(j_1 j_2 j_3)$ is not a proper triad satisfying the triangle conditions. The meaning of the substitutions of indices, $r \to s$, in Eqs. (131) and (132) is clarified by an example. For $j_1 = 2$, $j_2 = j_3 = 1$, Eq. (132) gives, correctly,

$$\begin{aligned}\mathsf{T}(1,2,1) &= 6\{\tfrac{1}{2}(\delta_{rs}\delta_{st}+\delta_{ss}\delta_{rt})-\tfrac{1}{3}\delta_{rs}\delta_{st}\}-\tfrac{10}{3}\delta_{ss}\delta_{rt} \\ &= \delta_{rs}\delta_{st}-\tfrac{1}{3}\delta_{ss}\delta_{rt}.\end{aligned}$$

Useful recursion formulas of a different type involve contractions with the elementary tensor ϵ. It can be shown that

$$\sum_{x,y}\epsilon_{rxy}\mathsf{T}\begin{pmatrix} j_1 & j_2 & j_3 \\ xr\cdots & s & yt\cdots \end{pmatrix} = -\beta\frac{(J+1)(J-2j_1)(J-2j_2)}{4j_1j_3(2j_3-1)}\mathsf{T}\begin{pmatrix} j_1 & j_2 & j_3-1 \\ r & s & t \end{pmatrix}, \tag{X1}$$

where $\beta = 1$ for J even, $\beta = 2$ for J odd. The left side is obviously traceless symmetric, and the numerical factor follows from the tensor formulas or, more easily, from the spinor correspondence. It is then straightforward to verify the following recursion relation:

$$\mathsf{T}\begin{pmatrix} j_1 & j_2 & j_3+1 \\ r & s & t \end{pmatrix} = \beta^{-1}\frac{2j_1}{J-2j_3}\sum_x \epsilon_{rxt}\mathsf{T}\begin{pmatrix} j_1 & j_2 & j_3 \\ xr\cdots & s & t \end{pmatrix} + \frac{(J+1)(J-2j_1)(J-2j_2)}{2(2j_3+1)(2j_3-1)(J-2j_3)}\delta_{tt}\mathsf{T}\begin{pmatrix} j_1 & j_2 & j_3-1 \\ r & s & t \end{pmatrix} \tag{X2}$$

since the right side is traceless in all sets of indices and has the correct leading term. This relation is particularly useful because it allows all 3-j tensors for the coupling $j_1 \otimes j_2 \to j_3$, to be obtained from the particularly simple tensors for $j_3 = j_1 + j_2$ or $j_3 = |j_1 - j_2|$, by stepping j_3 up or down. Related to these formulas is the cross substitution relation

$$\begin{aligned}\mathsf{T}\begin{pmatrix} j_1 & j_2 & j_3 \\ sr\cdots & rs\cdots & t\end{pmatrix} &= \tfrac{1}{2}\{[j_3(j_3+1) - j_1(j_1+1) - j_2(j_2+1)]/j_1 j_2\} \\ &\times \mathsf{T}\begin{pmatrix} j_1 & j_2 & j_3 \\ r & s & t\end{pmatrix} \\ &+ \frac{J(J+1)(J-2j_3)(J-2j_3-1)}{4j_1j_2(2j_1-1)(2j_2-1)} \\ &\times \delta_{rs}\mathsf{T}\begin{pmatrix} j_1-1 & j_2-1 & j_3 \\ r & s & t\end{pmatrix}, \end{aligned} \tag{X3}$$

which expresses the effect of the Casimir operator $\mathcal{J}^2(j_1 + j_2) = \mathcal{J}^2(j_1) + \mathcal{J}^2(j_2) + 2\mathcal{J}_1 \cdot \mathcal{J}_2$, for tensors in $\mathfrak{X}^{j_1+j_2}$, on $\mathsf{T}(j_1, j_2, j_3)$. The tensor on the left is given by the average of all $j_1 j_2$ single transpositions, $r \leftrightarrow s$.

All these relations can be expressed as relations on the coefficients x_{abc}, by the substitution of the explicit expressions of Eqs. (81) and (82). By contraction with $(\mathbf{u})^{j_1}(\mathbf{v})^{j_2}(\mathbf{w})^{j_3}$ they can also be regarded as recursion formulas for angular functions $P_{j_1 j_2 j_3}(\theta_{12}, \theta_{23}, \theta_{13})$ of the type given in Eq. (100). The substitutions of indices can be related to differential operators of the type $\mathbf{v} \cdot (\partial/\partial \mathbf{u})$. Insight into the relations is obtained by expressing them as recoupling transformations, which also relate the tensor norms Ω to the 6-j symbols. Thus,

$$\Omega_{j_1 j_2 j_3} = [j_1][j_2][j_3]\begin{Bmatrix} j_1 & j_2 & j_3 \\ \frac{1}{2}J - j_1 & \frac{1}{2}J - j_2 & \frac{1}{2}J - j_3\end{Bmatrix}^2 \tag{X4}$$

for J even.

IX. TENSORS NOT IN NATURAL FORM

The preceding formalism applies in principle to the coupling of irreducible tensors not in natural form, provided the natural projections are replaced [e.g., in Eqs. (27)–(28)] by the appropriate nth-order projections $\mathbf{\Pi}^j_\alpha(n \mid \bar{n})$.[33] However, equivalently, the coupling can be described in terms of the embedded reduced forms. If $\mathsf{A}^{(j_a)}(n_a)$, $\mathsf{B}^{(j_b)}(n_b)$, and $\mathsf{C}^{(j_c)}(n_c)$ are irreducible tensors and T_a, T_b, T_c describe their embeddings so that, in the notation of Eq. (1), e.g.,

$$\mathsf{A}^{(j_a)}(n_a) = \mathsf{T}(n_a \mid \bar{m}_a) \cdot \mathbf{a}^{(j_a)}(m_a),$$

where $\mathbf{a}^{(j_a)}$ is the reduced form, then the invariant part of their product is

$$(\mathsf{A}^{(j_a)}\mathsf{B}^{(j_b)}\mathsf{C}^{(j_c)})^{(0)} = \lambda \mathsf{T}\begin{pmatrix} j_a & j_b & j_c \\ n_a & n_b & n_c\end{pmatrix}, \tag{135}$$

where the more general 3-j tensor is given in terms of the natural form 3-j tensor by

$$\begin{aligned}&\mathsf{T}\begin{pmatrix} j_a & j_b & j_c \\ n_a & n_b & n_c\end{pmatrix} \\ &= \mathsf{T}_a(n_a \mid \bar{m}_a)\mathsf{T}_b(n_b \mid \bar{m}_b)\mathsf{T}_c(n_c \mid \bar{m}_c) \vdots \mathsf{T}\begin{pmatrix} j_a & j_b & j_c \\ m_a & m_b & m_c\end{pmatrix}.\end{aligned} \tag{136}$$

The reduced tensor λ is independent of the embedding, provided the embeddings are properly normalized, i.e.,

$$\lambda = \Omega^{-1}\mathsf{T}\begin{pmatrix} j_a & j_b & j_c \\ \bar{n}_a & \bar{n}_b & \bar{n}_c\end{pmatrix} \vdots \mathsf{A}^{(j_a)}\mathsf{B}^{(j_b)}\mathsf{C}^{(j_c)}, \tag{137}$$

$$\lambda = \Omega^{-1}\mathsf{T}\begin{pmatrix} j_a & j_b & j_c \\ \bar{m}_a & \bar{m}_b & \bar{m}_c\end{pmatrix} \vdots \mathbf{a}^{(j_a)}\mathbf{b}^{(j_b)}\mathbf{c}^{(j_c)}, \tag{138}$$

provided

$$\mathsf{T}^a(m_a \mid \bar{n}_a) \cdot \mathsf{T}_a(n_a \mid \bar{m}_a) = \mathsf{E}^{(j_a)}(m_a \mid \bar{m}_a), \tag{139}$$

etc. Similarly, in contraction together of different 3-j tensors of the form (136) according to Eq. (106), the T_a, T_b, and T_c are contracted together to natural projections E^j and produce no change in the result. Thus the 6-j and higher n-j symbols are independent of embedding.

X. HYBRID 3-j SYMBOLS

Before Racah's spherical tensor operator methods,[34] it was common to consider matrix elements of Cartesian tensor operators between spherical tensor states.[35] What is involved is essentially a "hybrid" 3-j tensor, with two spherical-tensor indices and one set of Cartesian indices, which can be defined by

$$\begin{aligned}\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & s & m_3\end{pmatrix} &= \sum_{m_2}\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3\end{pmatrix}\mathbf{e}^{j_2}_{m_2}(s) \\ &= \mathsf{V}\begin{pmatrix} j_1 & j_2 & j_3 \\ r & s & t\end{pmatrix} : \mathbf{e}^{j_1 m_1}(\bar{r})\mathbf{e}^{j_3 m_3}(\bar{t}),\end{aligned} \tag{140}$$

with the second expression for integral j_1, j_3 only. For these, the Wigner–Eckart theorem takes the form[36]

$$\langle \psi^{(j_1)}_{m_1} | A^{(j_2)}_{s_1 s_2 \cdots s_{j_2}} | \psi^{(j_3)}_{m_3} \rangle = \lambda\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & \bar{s} & \bar{m}_3\end{pmatrix}, \tag{141}$$

where λ, the reduced matrix element, coincides with that for the spherical tensor operator $A^{(j_2)}_{m_2} = \mathsf{A}^{(j_2)}(\bar{s}) \cdot \mathbf{e}^{j_2}_{m_2}(s)$ in the usual spherical tensor theory.[37]

The disadvantage is simply that a general formula for the hybrid 3-j tensors is complicated. The formal properties, on the other hand, are analogous to those of the purely spherical and purely Cartesian 3-j tensors—e.g., the same 6-j and 9-j symbols—so that there is no particular disadvantage if the final results are to be scalars, indices being formally summed. A simple result, for example, is

$$\sum_{m_1,m_3} |\langle \psi_{m_1}^{(j_1)}| \mathsf{A}^{(j_2)}(\bar{s}) |\psi_{m_3}^{(j_3)}\rangle|^2 = |\lambda|^2 \mathsf{E}^{j_2}(s \mid \bar{s}). \quad (142)$$

XI. DISCUSSION

Tensor averages arise through integrations over coordinates or, for tensor operators, from the trace, and also in connection with quantum matrix elements. However, they are usually calculated efficiently not by integration, but by projection onto the appropriate subspace H_n^0 of invariant tensors—this, for example, is the point of the Wigner–Eckart theorem for matrix elements. The projection is straightforward in principle if a basis for H_n^0 is known,[38] and the problem may be greatly simplified if symmetry conditions on the tensor restrict the possibilities to a smaller subspace $H_n^{0'} \subset H_n^0$ of invariant tensors of the given symmetry. Thus, an important case in which $H_n^{0'}$ is one- (or zero-) dimensional is that of a product of two irreducible tensors [Eq. (39)] and another for $SO(3)$ is that of any totally symmetric tensor. Here the invariant part must be proportional to $\delta_{rr}^{\frac{1}{2}n}$,[23] and normalization gives

$$\begin{aligned} (\mathsf{A}(r))^{(0)} &= a(n+1)^{-1}\delta_{rr}^{\frac{1}{2}n}, \quad n \text{ even}, \\ &= 0, \qquad\qquad\quad n \text{ odd}, \end{aligned} \quad (143)$$

where $a = A_{ppqqrr\cdots}$ (summation convention).[39]

The tensorial content of the Wigner–Eckart theorem[6,7] for a multiplicity-free group is, likewise, that the invariant part of the product of three irreducible tensors lies in a one-dimensional $H_n^{0'}$ (whose basis is the 3-j tensor). For a nonmultiplicity-free group, the analog of this property is given by Eqs. (19) and (20). However, the distinguishing property of the *triple* tensor product has been lost. The product of three irreducible tensors is no different in principle from that of any number of irreducible tensors.

For the invariant part of an N-fold product, we have analogously that

$$(\mathsf{A}^{(j_1)}\mathsf{B}^{(j_2)}\cdots\mathsf{F}^{(j_N)})^{(0)} = \sum_{p=1}^{M} \lambda^p \mathsf{V}_p(j_1, j_2, \cdots, j_N), \quad (144)$$

where λ^p, the pth N-fold scalar product ("reduced invariant tensor") is

$$\lambda^p = (\mathsf{A}^{(j_1)}\mathsf{B}^{(j_2)}\cdots\mathsf{F}^{(j_N)}) \odot^N \mathsf{V}^p(j_1, j_2, \cdots, j_N). \quad (145)$$

The V_p and V^p could logically be called N-j tensors (not to be confused with the scalar n-j symbols) and are defined in the Cartesian case by equations of the type

$$\begin{aligned} &\mathsf{V}_p(j_1, j_2, \cdots, j_N) \\ &\quad = (\mathsf{E}^{j_1}\mathsf{E}^{j_2}\cdots\mathsf{E}^{j_N}) \odot^N \mathsf{V}_p(j_1, j_2, \cdots, j_N), \end{aligned} \quad (146)$$

[cf. Eqs. (27)–(28)] with appropriate modification if some indices are covariant, some contravariant, and with $\mathsf{V}^p \odot^N \mathsf{V}_q = \delta_q^p$. The multiplicity $M(j_1, j_2, \cdots, j_N)$ is given by the possible parentage schemes for coupling N tensors to symmetry $j = 0$, and parentage also gives one possible set of (normalized) N-j tensors, analogous to the generalized Clebsch–Gordan coefficients of Yutsis, Levinson, and Vanagas,[40]

$$\begin{aligned} \mathsf{V}_p = \mathsf{V}_{p_1}(j_1, j_2, \bar{k}_2) &\odot^{k_2} \mathsf{V}_{p_2}(k_2, j_3, \bar{k}_3) \\ &\times \odot^{k_3} \cdots \odot^{k_{N-2}} \mathsf{V}_{p_{N-2}}(k_{N-2}, j_{N-1}, j_N), \end{aligned} \quad (147)$$

though not usually the simplest set. For Cartesian tensors under the rotation group, the N-j tensors are simply a set of linearly independent invariant tensors separately traceless and symmetric in N sets of indices.

ACKNOWLEDGMENT

It is a pleasure to acknowledge many interesting discussions with R. F. Snider.

APPENDIX A

1. Elementary Invariant Tensors

These are usually directly related to the definition of the group and, in principle, completely determine the possible projections into invariant and, in particular, irreducible subspaces and, hence, the possible irreducible representations. From this standpoint, the inequivalent irreducible representations of the general linear group $GL(v)$ remain irreducible and inequivalent under the real linear group $GL'(v)$ and unitary group $U(v)$, because all three are characterized by the existence of the same single elementary invariant tensor, the unit tensor δ_r^s. The unimodular groups $SL(v)$ and $SU(v)$ are defined by the existence of the additional elementary invariant tensors $\epsilon_{r_1r_2\cdots r_v}$ and $\epsilon^{s_1s_2\cdots s_v}$ (the permutation symbols). The fact[10] that irreducible representations of $GL(v)$ and $U(v)$ remain irreducible under $SL(v)$ and $SU(v)$ corresponds to that fact that the *mixed* tensor $\mathbf{\Pi}_\alpha^j$ can only contain the *covariant* $\epsilon_{r_1r_2\cdots r_v}$ and *contravariant* $\epsilon^{s_1s_2\cdots s_v}$ together, so that in view of the relation

$$\epsilon_{r_1r_2\cdots r_v}\epsilon^{s_1s_2\cdots s_v} = \sum_{\mathfrak{P}} (-1)^{\mathfrak{P}} \mathfrak{P}(\delta_{r_1}^{s_1}\delta_{r_2}^{s_2}\cdots\delta_{r_v}^{s_v}), \quad (A1)$$

where the sum is over all permutations of the r (or s) indices, there are no new possibilities for the projections. Since, however, there are invariant mappings between spaces of different order, representations inequivalent under $GL(v)$ and $U(v)$ may be equivalent under $SL(v)$ and $SU(v)$. Similarly, the orthogonal and symplectic groups are defined by the invariance of the metric tensor, i.e., by existence of a covariant second-order elementary invariant tensor, symmetric, δ_{rs}, or skew-symmetric, together with the corresponding contravariant tensor. In contrast to the change from $U(v)$ to $SU(v)$ the irreducible representations of $SO(v)$ are, for v even, essentially different from $O(v)$, because $\epsilon_{r_1\cdots s_1\cdots s_{\frac{1}{2}v}}$, with covariant and contravariant indices now equivalent, can appear *singly* in the projections.

2. Unitary Representations

Representations of compact and finite groups can always be made unitary,[10] and in this case we assume G itself to be a group of unitary transformations on $\mathfrak{X}$. The Hermitian inner product of two tensors of the same order,

$$\langle \mathsf{A}, \mathsf{B}\rangle = \sum_{r_1\cdots r_n} (A^{r_1\cdots r_n})^* B^{r_1\cdots r_n}, \qquad \text{(A2)}$$

is then invariant under G. The orthogonality relations[10] for group representations imply that basis tensors $\mathbf{e}^j_{\alpha\sigma}(n)$ for unitary irreducible representations (α a multiplicity index) satisfy

$$\langle \mathbf{e}^j_{\alpha\sigma}, \mathbf{e}^{j'}_{\beta\tau}\rangle = \lambda^j_{\alpha\beta}\delta_{\sigma\tau}\delta_{jj'}; \qquad \text{(A3)}$$

where $\lambda^j_{\alpha\beta}$ is independent of σ and τ. Consequently, we may assume that tensor bases are taken orthonormal with respect to the inner product, and we define the dual space by

$$\mathbf{e}^\sigma(\bar{r}) = \mathbf{e}^\dagger_\sigma(\bar{r}) = \mathbf{e}_\sigma(r)^* \qquad \text{(A4)}$$

as the adjoint space.

The adjoint $\mathsf{T}^\dagger$ of a tensor mapping $\mathsf{T}(n \mid \bar{n})$: $\mathfrak{X}^n \to \mathfrak{X}^n$ is obtained by the interchange of covariant and contravariant indices and complex conjugation, and it is convenient to use this terminology also for mappings T: $\mathfrak{X}^n \to \mathfrak{X}^m$ between spaces of different order. Thus

$$\langle \mathbf{e}_\sigma(m), \mathsf{T}\cdot\mathbf{e}_\tau(n)\rangle = \langle \mathsf{T}^\dagger\cdot\mathbf{e}_\sigma(m), e_\tau(n)\rangle, \qquad \text{(A5)}$$

where

$$\mathsf{T}^\dagger(r \mid \bar{s}) = \mathsf{T}(s \mid \bar{r})^*. \qquad \text{(A6)}$$

The permutation symbol $\epsilon_{r_1r_2\cdots r_n}$ is the adjoint of $\epsilon^{r_1r_2\cdots r_n}$, and the projections E^j are self-adjoint.

The invariant mapping $\mathsf{T}^{(0)}$: $H^j_{n\alpha} \to H^j_{m\beta}$ of an irreducible subspace is conformal with respect to the inner product. For if $\mathsf{T}^{(0)}(m \mid \bar{n})$ is any invariant tensor, then $\mathsf{T}^{(0)\dagger} \odot^m \mathsf{T}^{(0)}$ maps $\mathfrak{X}^n$ into itself, and by Schur's lemma,[10]

$$\begin{aligned}\langle \mathbf{e}^j_{\beta\sigma}, \mathbf{e}^j_{\beta\tau}\rangle &= \langle \mathsf{T}^{(0)}\cdot\mathbf{e}^j_{\alpha\sigma}, \mathsf{T}^{(0)}\cdot\mathbf{e}^j_{\alpha\tau}\rangle \\ &= \langle \mathbf{e}^j_{\alpha\sigma}, (\mathsf{T}^{(0)\dagger}\cdot\mathsf{T}^{(0)})\cdot\mathbf{e}^j_{\alpha\tau}\rangle = \lambda\langle \mathbf{e}^j_{\alpha\sigma}, \mathbf{e}^j_{\beta\tau}\rangle, \qquad \text{(A7)}\end{aligned}$$

where λ is independent of σ and τ, a result which also follows from Eq. (A3) and the fact that invariant maps preserve symmetry.

Given an arbitrary basis $\mathsf{T}_p(n \mid \bar{m}_j)$ for the reduction of nth-order tensors, we can always in the unitary case define a metric by which the multiplicity indices p can be raised or lowered. The adjoint $\mathsf{T}^\dagger_p$ of $\mathsf{T}_p = \mathsf{T}_p\cdot\mathsf{E}^j$ satisfies $\mathsf{T}^\dagger_p = \mathsf{E}^j\cdot\mathsf{T}^\dagger_p$, since E^j is self-adjoint, and consequently the dot product $\mathsf{T}^\dagger_p \odot^n \mathsf{T}_q$ is proportional to E^j (Paper II, Theorem 3). We write

$$\mathsf{T}^\dagger_p(m_j \mid \bar{n}) \odot^n \mathsf{T}_q(n \mid \bar{m}_j) = g_{pq}\mathsf{E}^j(m_j \mid \bar{m}_j). \qquad \text{(A8)}$$

Regarding the $\mathsf{T}_p(n \mid \bar{m}_j)$ as a linearly dependent basis for the $N^j_n[j]$-dimensional subspace H^j_n, the quantity $\mathsf{T}^\dagger_p\cdot\mathsf{T}_q$ can be regarded as a metric for the basis, which factors according to (A8) into $\mathsf{E}^j(m_j \mid \bar{m}_j)$ for the tensor indices and into a positive-definite Hermitian metric g_{pq} for the multiplicity indices. This is a generalization of the results of Paper II. Note, however, that the g_{pq} for spinors in Paper II was not defined in this way and was not positive definite. The definition of g_{pq} here differs by a factor $[j]$ from that of Eq. (24) for the 3-j tensors, because the T_p and $\mathsf{V}_p(j_1, j_2, j_3)$ are normalized differently.

The dual basis for H^j_n can now be obtained by raising indices with the inverse tensor,

$$\mathsf{T}^p(m_j \mid \bar{n}) = \sum_q g^{pq}\mathsf{T}^\dagger_q(m_j \mid \bar{n}), \qquad \text{(A9)}$$

Eq. (2) being clearly satisfied. If the basis is orthogonal ($g_{pq} = \delta_{pq}$), the individual (orthogonal) projections $\mathsf{T}_p \odot^{m_j} \mathsf{T}^\dagger_p = \mathbf{\Pi}^j_p$ give a complete decomposition of H^j_n. By a unitary transformation of the basis, a unique set of $\mathbf{\Pi}^j_p$ can be obtained for any given tensor $\mathsf{F}(n)$ such that the reduced tensors $\mathsf{f}^{(j)p}$ are also orthogonal [a generalization of the result given in Paper II for $SO(3)$].

Equations (39) and (41) contain the orthogonality relations for irreducible Cartesian tensors in natural form—with analogous relations in terms of the $\mathbf{g}_j$ and $\mathbf{g}^j$ where these exist. In the unitary case, according to (A4), the complex conjugate of a tensor "transforms covariantly," [41] and from Eq. (39),

$$K^{-1}\int \mathsf{A}^{(j)}(r)^*\mathsf{B}^{(j')}(s)\,d\mathfrak{X} = \delta_{jj'}[j]^{-1}\langle \mathsf{A}^{(j)}, \mathsf{B}^{(j)}\rangle E^j(s \mid \bar{r}), \qquad \text{(A10)}$$

where the integration is over the "unit sphere" in $\mathfrak{X}$ and K is the (finite) integration volume. For tensor operators in a finite Hilbert space of dimension K, e.g., a space of angular-momentum states, the left side is replaced by $K^{-1}\,\mathrm{Tr}_Q\,(\mathsf{A}^{(j)\dagger}\mathsf{B}^{(j')})$, the trace Tr_Q being over the Hilbert space.

3. Example

We give for illustration some 3-j tensors for $GL(3)$.[42] The only elementary invariant tensor is the mixed tensor δ^s_r. Table VII gives a set of projections for natural tensors of orders 1, 2, and 3 in terms of δ^s_r. The third-order, 8-dimensional representation, denoted by $3E$, occurs with multiplicity two, and we have arbitrarily chosen the standard tensors as the eigentensors of the cyclic permutation (123) with eigenvalue $\omega = e^{\frac{2}{3}\pi i}$ and projection $\mathsf{E}^{3E} = \mathbf{\Pi}^{3E}_{+}$. The total projection for (reducible) tensors of symmetry $3E$ is $\mathbf{\Pi}^{3E} = \mathbf{\Pi}^{3E}_{+} + \mathbf{\Pi}^{3E}_{-}$, where $\mathbf{\Pi}^{3E}_{-} = (\mathbf{\Pi}^{3E}_{+})^*$.[43] The unnormalized 3-j tensors $\mathsf{T}(j, 0, \bar{j})$ coincide with the projections (1-j tensors), and their norms Ω are therefore the same as the traces of the projections, i.e., equal to the dimensions $[j]$ of the representations. The unnormalized 3-j tensors are defined by

$$\mathsf{T}\begin{pmatrix} j_1 & j_2 & j_3 \\ r & s & \bar{t} \end{pmatrix} = \sum_{\text{all } x,y} \mathsf{E}^{j_1}(r \mid \bar{x})\mathsf{E}^{j_2}(s \mid \bar{y})\mathsf{E}^{j_3}(x, y \mid \bar{t}),$$

$$\mathsf{T}\begin{pmatrix} j_3 & \bar{j}_1 & \bar{j}_2 \\ t & \bar{r} & \bar{s} \end{pmatrix} = \sum_{\text{all } x,y} \mathsf{E}^{j_3}(t \mid \bar{x}, \bar{y})\mathsf{E}^{j_1}(x \mid \bar{r})\mathsf{E}^{j_2}(y \mid \bar{s}), \tag{A11}$$

[cf. Eq. (75)].

The 3-j representation of the coupling of two contravariant vectors to a symmetric tensor is thus (using the summation convention)

$$\begin{aligned} C^{t_1t_2} &= \mathsf{T}\begin{pmatrix} 2S & \bar{1} & \bar{1} \\ t & \bar{r} & \bar{s} \end{pmatrix} a^r b^s \\ &= \tfrac{1}{2}(a^{t_1}b^{t_2} + a^{t_2}b^{t_1}). \end{aligned} \tag{A12}$$

The same 3-j tensor couples a contravariant vector and a symmetric covariant tensor to a covariant vector:

$$\begin{aligned} c_s &= A_{t_1t_2}b^r\mathsf{T}\begin{pmatrix} 2S & \bar{1} & \bar{1} \\ t & \bar{r} & \bar{s} \end{pmatrix} \\ &= A_{sr}b^r. \end{aligned} \tag{A13}$$

Finally, the invariant part of the tensor product of a covariant symmetric tensor with two contravariant vectors is[44]

$$\begin{aligned} (A_{t_1t_2}b^rc^s)^{(0)} &= \lambda\mathsf{T}\begin{pmatrix} 1 & 1 & \bar{2}S \\ r & s & \bar{t} \end{pmatrix} \\ &= \tfrac{1}{2}\lambda(\delta^r_{t_1}\delta^s_{t_2} + \delta^s_{t_1}\delta^r_{t_2}), \end{aligned} \tag{A14}$$

where

$$\begin{aligned} \lambda &= \Omega^{-1}A_{t_1t_2}b^rc^s\mathsf{T}\begin{pmatrix} 2S & \bar{1} & \bar{1} \\ t & \bar{r} & \bar{s} \end{pmatrix} \\ &= \tfrac{1}{6}A_{rs}b^rc^s. \end{aligned} \tag{A15}$$

4. Spin-1 System

The fine structure of a spin-1 system is very conveniently considered in a Cartesian basis, especially in "weak field" situations,[45] and provides a very simple but interesting example of the Cartesian Wigner–Eckart theorem. The basis states $\boldsymbol{\phi}_r$ (i.e., an ordinary vector basis $\mathbf{i}, \mathbf{j}, \mathbf{k}$) satisfy

$$S_r\boldsymbol{\phi}_s = i\sum_t \epsilon_{rst}\boldsymbol{\phi}_t, \tag{A16}$$

in terms of the usual rotation generators (spin operators). Any operator for a spin-1 system can be expanded in terms of the natural form irreducible tensor operators

$$\begin{aligned} \bar{Y}^{(0)} &= 1, \\ \bar{Y}^{(1)}_s &= S_s, \\ \bar{Y}^{(2)}_{s_1s_2} &= \tfrac{1}{2}(S_{s_1}S_{s_2} + S_{s_2}S_{s_1}) - \tfrac{1}{3}\mathbf{S}^2\delta_{s_1s_2}, \end{aligned} \tag{A17}$$

whose matrix elements are proportional, respectively, to the 3-j tensors

$$\begin{aligned} \mathsf{T}(1, 0, 1) &= \delta_{rt}, \\ \mathsf{T}(1, 1, 1) &= \epsilon_{rst}, \\ \mathsf{T}(1, 2, 1) &= \tfrac{1}{2}(\delta_{rs_1}\delta_{ts_2} + \delta_{rs_2}\delta_{ts_1}) - \tfrac{1}{3}\delta_{rt}\delta_{s_1s_2} \end{aligned} \tag{A18}$$

(Table I). The proportionality constant for $\bar{Y}^{(1)}$ is conveniently found from $\langle\boldsymbol{\phi}_y|\,S_z\,|\boldsymbol{\phi}_x\rangle = i$ [Eq. (A16)], and that for $\bar{Y}^{(2)}$ from $\langle\boldsymbol{\phi}_z|\,\bar{Y}^{(2)}_{zz}\,|\boldsymbol{\phi}_z\rangle = -\tfrac{1}{3}\mathbf{S}^2 = -\tfrac{2}{3}$;

$$\begin{aligned} \langle\boldsymbol{\phi}_r|\,1\,|\boldsymbol{\phi}_t\rangle &= \delta_{rt}, \\ \langle\boldsymbol{\phi}_r|\,S_s\,|\boldsymbol{\phi}_t\rangle &= i\epsilon_{rst}, \\ \langle\boldsymbol{\phi}_r|\,\bar{Y}^{(2)}_{s_1s_2}\,|\boldsymbol{\phi}_t\rangle &= -[\tfrac{1}{2}(\delta_{rs_1}\delta_{ts_2} + \delta_{rs_2}\delta_{ts_1}) - \tfrac{1}{3}\delta_{rt}\delta_{s_1s_2}]. \end{aligned} \tag{A19}$$

Therefore, for a spin Hamiltonian of the form

$$\mathcal{H} = \mathbf{H}\cdot\mathbf{S} + \mathsf{D} : \bar{Y}^{(2)}, \tag{A20}$$

where D is, without loss of generality, traceless and symmetric, the Hamiltonian matrix is

$$\mathcal{H}_{rt} = i\sum_s H_s\epsilon_{rst} - D_{rt}. \tag{A21}$$

In particular, in the principle axis coordinate system of D,

$$\mathcal{H} = \begin{bmatrix} -D_{xx} & -iH_z & iH_y \\ iH_z & -D_{yy} & -iH_x \\ -iH_y & iH_x & -D_{zz} \end{bmatrix}. \tag{A22}$$

TABLE VII. 3-j tensors for $GL(3)$. $\mathsf{V} = \Omega^{-\frac{1}{2}}\mathsf{T}$. The T tensors have relative normalization of the Clebsch-Gordan type (cf. $\mathsf{T}^{j}_{j_1 j_2}$ and $\mathsf{T}^{j_3}_{j_1 j_2}$ of Sec. II). Symmetries j are labeled by the tensorial order and the permutation symmetry, S totally symmetric, A totally antisymmetric, and E the doubly degenerate representation of S_3, i.e., the 8-dimensional multiplicity-two representation of $GL(3)$. $\omega = e^{(2\pi i/3)}$, and $\mathsf{T}^\dagger$ is the transpose conjugate.

Projection	3-j tensor	Expression	$\Omega = \mathrm{Tr}\,\mathsf{E}^j = [j]$
$\mathsf{E}^1(r \mid t)$	$\mathsf{T}\begin{pmatrix}1 & 0 & \bar{1}\\ r & s & \bar{t}\end{pmatrix}$	δ^r_t	3
$\mathsf{E}^{2S}(r \mid t)$	$\mathsf{T}\begin{pmatrix}2S & 0 & \bar{2}S\\ r & s & \bar{t}\end{pmatrix}$	$(\delta^r_t)^2 = \frac{1}{2}(\delta^r_{t_1}\delta^r_{t_2} + \delta^r_{t_1}\delta^r_{t_2})$	6
$\mathsf{E}^{2A}(r \mid t)$	$\mathsf{T}\begin{pmatrix}2A & 0 & \bar{2}A\\ r & s & \bar{t}\end{pmatrix}$	$\frac{1}{2}(\delta^r_{t_1}\delta^r_{t_2} - \delta^r_{t_1}\delta^r_{t_2})$	3
$\mathsf{E}^{3S}(r \mid t)$	$\mathsf{T}\begin{pmatrix}3S & 0 & \bar{3}S\\ r & s & \bar{t}\end{pmatrix}$	$(\delta^r_t)^3$	10
$\mathsf{E}^{3E}(r \mid t)$	$\mathsf{T}\begin{pmatrix}3E & 0 & \bar{3}E\\ r & s & \bar{t}\end{pmatrix}$	$\mathbf{\Pi}^{3E}_{+} = \frac{1}{3}(\delta^r_{t_1}\delta^r_{t_2}\delta^r_{t_3} + \omega\delta^r_{t_1}\delta^r_{t_2}\delta^r_{t_3} + \omega^*\delta^r_{t_1}\delta^r_{t_2}\delta^r_{t_3})$	8
$\mathsf{E}^{3A}(r \mid t)$	$\mathsf{T}\begin{pmatrix}3A & 0 & \bar{3}A\\ r & s & \bar{t}\end{pmatrix}$	$\frac{1}{6}(\delta^r_{t_1}\delta^r_{t_2}\delta^r_{t_3} - \delta^r_{t_1}\delta^r_{t_2}\delta^r_{t_3} - \delta^r_{t_1}\delta^r_{t_2}\delta^r_{t_3} - \delta^r_{t_1}\delta^r_{t_2}\delta^r_{t_3} + \delta^r_{t_1}\delta^r_{t_2}\delta^r_{t_3} + \delta^r_{t_1}\delta^r_{t_2}\delta^r_{t_3})$	1

3-j tensors	Expression	$\Omega = \mathsf{T}^\dagger \vdots \mathsf{T}$
$\mathsf{T}\begin{pmatrix}1 & 1 & \bar{2}S\\ r & s & \bar{t}\end{pmatrix} = \mathsf{T}^\dagger\begin{pmatrix}2S & \bar{1} & \bar{1}\\ t & \bar{r} & \bar{s}\end{pmatrix}$	$\mathsf{E}^{2S}(rs \mid t_1t_2) = \frac{1}{2}(\delta^r_{t_1}\delta^s_{t_2} + \delta^s_{t_1}\delta^r_{t_2})$	6
$\mathsf{T}\begin{pmatrix}1 & 1 & 2A\\ r & s & \bar{t}\end{pmatrix} = \mathsf{T}^\dagger\begin{pmatrix}2A & \bar{1} & \bar{1}\\ t & \bar{r} & \bar{s}\end{pmatrix}$	$\mathsf{E}^{2A}(rs \mid t_1t_2) = \frac{1}{2}(\delta^r_{t_1}\delta^s_{t_2} - \delta^s_{t_1}\delta^r_{t_2})$	3
$\mathsf{T}\begin{pmatrix}1 & 2S & \bar{3}S\\ r & s & \bar{t}\end{pmatrix} = \mathsf{T}^\dagger\begin{pmatrix}3S & \bar{1} & \bar{2}S\\ t & \bar{r} & \bar{s}\end{pmatrix}$	$\mathsf{E}^{3S}(rs_1s_2 \mid t_1t_2t_3)$	10
$\mathsf{T}\begin{pmatrix}1 & 2A & \bar{3}A\\ r & s & \bar{t}\end{pmatrix} = \mathsf{T}^\dagger\begin{pmatrix}3A & \bar{1} & \bar{2}A\\ t & \bar{r} & \bar{s}\end{pmatrix}$	$\mathsf{E}^{3A}(rs_1s_2 \mid t_1t_2t_3)$	1
$\mathsf{T}\begin{pmatrix}1 & 2S & \bar{3}E\\ r & s & \bar{t}\end{pmatrix} = \mathsf{T}^\dagger\begin{pmatrix}3E & \bar{1} & \bar{2}S\\ t & \bar{r} & \bar{s}\end{pmatrix}$ $\mathsf{T}\begin{pmatrix}1 & 2A & \bar{3}E\\ r & s & \bar{t}\end{pmatrix} = \mathsf{T}^\dagger\begin{pmatrix}3E & \bar{1} & \bar{2}A\\ t & \bar{r} & \bar{s}\end{pmatrix}$	$\frac{1}{6}[(\delta^r_{t_1}\delta^s_{t_2}\delta^s_{t_3} + \omega\delta^s_{t_1}\delta^s_{t_2}\delta^r_{t_3} + \omega^*\delta^s_{t_1}\delta^r_{t_2}\delta^s_{t_3})$ $\pm (\delta^r_{t_1}\delta^s_{t_2}\delta^s_{t_3} + \omega\delta^s_{t_1}\delta^s_{t_2}\delta^r_{t_3} + \omega^*\delta^s_{t_1}\delta^r_{t_2}\delta^s_{t_3})]$	4

The eigenstates $\psi_\mu = \sum_r \phi_r c_{r\mu}$ are specified by vectors $\mathbf{c}_\mu$, orthonormal in the sense $\langle\psi_\mu \mid \psi_\nu\rangle = \mathbf{c}^*_\mu \cdot \mathbf{c}_\nu = \delta_{\mu\nu}$, and hence, from (A19), the transition moment is proportional to a cross product:

$$\begin{aligned}\langle\psi_\mu|\,\mathbf{S}\,|\psi_\nu\rangle &= i\mathbf{c}^*_\mu \cdot \mathsf{T}(1,1,1)\cdot\mathbf{c}_\nu \\ &= i\mathbf{c}_\nu \times \mathbf{c}^*_\mu. \end{aligned} \tag{A23}$$

These considerations are particularly simple for spin 1, but they can be generalized to any integral spin. For spin j the states are natural tensors of order j, with linearly dependent basis states $\phi_{r_1r_2\cdots r_j} = \phi_r$ normalized according to[36]

$$\langle\phi_r \mid \phi_t\rangle = \mathsf{E}^j(r \mid \bar{t}), \tag{A24}$$

and with components[46]

$$\phi_r(q) = \mathsf{E}^j(q \mid \bar{r}).$$

The total spin $\mathbf{S}$ in $\mathfrak{X}^j$ is defined in the usual way, in terms of the generators in $\mathfrak{X}^1$, as a sum of operators of the type $1(1)\otimes 1(2)\cdots\otimes\mathbf{S}(i)\cdots\otimes 1(j)$. The matrix elements of irreducible tensor operators $\bar{\mathsf{Y}}^k$ are proportional to the 3-j tensors $\mathsf{T}(j,k,j)$ and, in

particular, for those in (A17),[47]

$$\langle \boldsymbol{\phi}_r | S_s | \boldsymbol{\phi}_t \rangle = (i)^j \mathsf{T}\begin{pmatrix} j & 1 & j \\ r & s & t \end{pmatrix}, \tag{A25}$$

$$\langle \boldsymbol{\phi}_r | \bar{Y}_s^{(2)} | \boldsymbol{\phi}_t \rangle = -j(2j-1)\mathsf{T}\begin{pmatrix} j & 2 & j \\ r & s & t \end{pmatrix}. \tag{A26}$$

Eigenstates $\psi_\mu = \sum_r \boldsymbol{\phi}_r \mathsf{C}_\mu(r)$ of a spin Hamiltonian $\mathcal{H}$ are specified by traceless symmetric tensors C_μ of order j, satisfying

$$[\mathcal{H}(r \mid t) - \lambda \mathsf{E}^j(r \mid t)] \cdot \mathsf{C}_\mu(t) = 0, \tag{A27}$$

and orthonormal in the sense $\langle \psi_\mu \mid \psi_\nu \rangle = \mathsf{C}_\mu^* \cdot \mathsf{C}_\nu = \delta_{\mu\nu}$. From Eqs. (A25) and (80), the transition moment is proportional to

$$\langle \psi_\mu | \mathbf{S} | \psi_\nu \rangle = (i)^j (\mathsf{C}_\nu \bigcirc^{j-1} \times \mathsf{C}_\mu^*), \tag{A28}$$

which is the generalization of Eq. (A23).

* Supported by the National Research Council of Canada and the President's Research Fund of the University of British Columbia.

[1] J. A. R. Coope and R. F. Snider, J. Math. Phys. **11**, 993 (1970), referred to throughout as Paper II.

[2] E. P. Wigner, "On the Matrices Which Reduce the Kronecker Products of Representations of S.R. Groups," in *Quantum Theory of Angular Momentum*, L. C. Biedenharn and H. van Dam, Eds. (Academic, New York, 1965).

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

[4] B. Barsella and E. Fabri, Nuovo Cimento Suppl., Ser. **1**, **2**, 293 (1964).

[5] J. A. R. Coope, R. F. Snider, and F. R. McCourt, J. Chem. Phys. **43**, 2269 (1965), Paper I. The coupling of tensors with vectors, given here, is a special case.

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

[7] C. Eckart, Rev. Mod. Phys. **2**, 305 (1930).

[8] We use essentially the same notation as in Paper II. The dimension of an irreducible representation is here denoted by $[j]$ rather than D^j to conform with common usage for the CG problem. In $\mathsf{e}_\sigma(n \mid \bar{m})$, the n and $\bar{m}$ indicate numbers of contravariant and covariant indices. In $\mathsf{e}_\sigma(n \mid \bar{m}) : \mathsf{e}^\sigma(m \mid \bar{n})$, each dot indicates scalar contraction over whole sets of indices according to a suitable convention, e.g., the first of a set with the first, etc. Where ambiguity could occur, we write, e.g., $\mathsf{e}_\alpha(n \mid \bar{m}) \bigcirc^m \mathsf{e}^\sigma(m \mid \bar{n})$. In Sec. IIff. we frequently omit the designation of tensorial order and substitute a generic symbol for the particular Cartesian indices, $\mathsf{e}(r \mid \bar{s}) = e^{r_1 r_2 \cdots r_n}_{s_1 s_2 \cdots s_m}$.

[9] H. Weyl, *The Theory of Groups and Quantum Mechanics* (Dover, New York, 1949). Hamermesh (Ref. 10) calls $\bar{j}$ the "adjoint representation."

[10] M. Hamermesh, *Group Theory* (Addison-Wesley, Reading, Mass., 1962).

[11] For unitary representations this requirement reduces to the self-adjointness condition on an orthogonal projection.

[12] While suitable for $SO(3)$ and $SU(2)$, the term is perhaps optimistic in general, where the selection of a standard H^j may be arbitrary.

[13] For example, G. E. Hay, *Vector and Tensor Analysis* (Dover, New York, 1953).

[14] See, e.g., the discussion of the improper rotation groups $O(3)$ and $O(2)$ in Paper II.

[15] $\mathsf{T}_p(n \mid \bar{m}_j)$ and $\mathsf{T}^p(m_j \mid \bar{n})$ are, for fixed components m_j, dual tensors of symmetry j and $\bar{j}$ in $\mathfrak{T}^n$ and $\mathfrak{T}_n$. Dual bases for the invariant part of $\mathfrak{T}^n_{m_j}$ and $\mathfrak{T}^{m_j}_n$ would be (relatively) normalized: $\mathsf{T}^p : \mathsf{T}_q = \delta^p_q$.

[16] For an arbitrary basis the individual terms $\mathsf{T}_p \cdot \mathsf{T}^p$ are not necessarily projections of the type $\mathbf{\Pi}^j_\alpha$—in the unitary case $\mathsf{T}_p \cdot \mathsf{T}^p$ is not self-adjoint unless the basis is orthogonal.

[17] The right sides must be proportional to a projection, and further contraction must give Eq. (13).

[18] In the unitary case, this follows (e.g., Ref. 10) from the group orthogonality relations. More generally, if A and B are irreducible tensors whose linearly independent components A_r and B_s transform according to the irreducible representations $U_{r'r}(\mathcal{R})$ and $V_{s's}(\mathcal{R})$, then $\sum_{r,s} A_r C_{rs} B_s$ is invariant for all A and B of the given symmetries if and only if $\mathbf{U}\mathbf{C}\tilde{\mathbf{V}} = \mathbf{C}$ or $\mathbf{U}\mathbf{C} = \mathbf{C}\tilde{\mathbf{V}}^{-1}$ for all $\mathcal{R}$. By Schur's lemmas [H. Boerner, *Representations of Groups* (North-Holland, Amsterdam, 1963)], $\mathbf{U}$ and $\tilde{\mathbf{V}}^{-1}$ are equivalent and, if the coordinate systems are chosen so that $\mathbf{U} \equiv \mathbf{V}$, then $\mathbf{C}$ is a multiple of the unit matrix and the invariant is $\sum_r A_r B_r$. See also Sec. III.

[19] For $j_3 = \bar{j}_3$ half-integral, there is a phase change from the previous definitions of mixed symbols, in the sense that if new mixed symbols are defined by raising and lowering the t indices in the completely covariant and contravariant tensors $\mathsf{V}^p(\bar{j}_1\bar{j}_2 j_3)$ and $\mathsf{V}_p(j_1, j_2, \bar{j}_3)$, using the 1-$j$ tensors defined in Sec. III with λ in Eq. (47) equal where possible to one, then the new mixed symbols are related to the previous ones of Eqs. (5), (6), and (13) by $\mathsf{V}^p(\bar{j}_1\bar{j}_2 j_3) = \pm\mathsf{V}^p(j_3, \bar{j}_1, \bar{j}_2)$, plus sign for $j_3 \neq \bar{j}_3$ or $j_3 = \bar{j}_3$ integral, minus for $j_3 = \bar{j}_3$ half-integral.

[20] The point that the 3-j symbols for nonmultiplicity-free groups can not necessarily be chosen such that their absolute values are invariant to permutations of the j's and corresponding indices, when $j_1 = j_2 = j_3$, is due to J.-R. Derome, J. Math. Phys. **7**, 612 (1966). We believe the situation is most clearly understood by recognizing the set of symbols for a triad with three j's equal as a basis for a representation of S_3, as is done in the text. The essential point is conveniently illustrated by considering a generalization—symbols for the coupling of N rather than three irreducible tensors to an invariant (Sec. XI). For all j's equal, the set of "N-j tensors" are a basis for S_N, and for $N > 3$ degenerate representations are expected even for the rotation group. The simplest example is the coupling of four vectors, for which the multiplicity $M(1, 1, 1, 1)$ is three. A convenient nonorthogonal set of Cartesian 4-j tensors is $\mathsf{V}_1 = \delta_{rs}\delta_{tu}$, $\mathsf{V}_2 = \delta_{rt}\delta_{su}$, $\mathsf{V}_3 = \delta_{ru}\delta_{st}$, and clearly the doubly degenerate representation of S_4, as well as the totally symmetric representation, occurs.

[21] The correspondence in notation is

$$(j_1 j_2 j_3)_{pm_1 m_2 m_3} \longleftrightarrow \begin{pmatrix} \bar{j}_1 & \bar{j}_2 & \bar{j}_3 \\ \bar{m}_1 & \bar{m}_2 & \bar{m}_3 \end{pmatrix}^p.$$

[22] E. P. Wigner, Am. J. Math. **63**, 57 (1941).

[23] As in Paper II we use the symmetrization convention that repeated indices without suffix, as in $\epsilon^{2j}_{r,s} = \epsilon_{rs}\epsilon_{rs} \cdots \epsilon_{rs}$, are completely symmetrized, e.g., $\epsilon^2_{rs} = \frac{1}{2}(\epsilon_{r_1 s_1}\epsilon_{r_2 s_2} + \epsilon_{r_1 s_2}\epsilon_{r_2 s_1})$.

[24] Thus $SU(3)$, with ϵ_{rst} as the only covariant elementary invariant tensor, is not ambivalent, whereas $SU(2)$, which is ambivalent, has the second order ϵ_{rs}. The converse is not true. $SO(2)$ with two elementary covariant tensors δ_{rs} and ϵ_{rs} is not ambivalent.

[25] They could well be written $x\begin{pmatrix} j_1 & j_2 & j_3 \\ a & b & c \end{pmatrix}$ but are, however, completely symmetric to permutations of columns. The antisymmetry of $\mathsf{T}(j_1, j_2, j_3)$ for J odd is contained in the basis tensor ϵ_{rst}, not in the $x_{a,b,c}$ coefficients.

[26] $\epsilon_{12} = \epsilon^{12} = -\epsilon_{21} = -\epsilon^{21} = 1$.

[27] Equation (99) coincides with that of Regge [T. Regge, Nuovo Cimento **10**, 544 (1958)] on making the change in notation $u_1 \leftrightarrow u_2$, $v_1 \leftrightarrow v_2$, $w_1 \leftrightarrow w_2$.

[28] The e^j_m, with subscript m, are taken in the present article to be in natural phase relation $(\mathcal{J}_+\mathsf{e}^j_m = +[j(j+1) - m(m+1)]^{\frac{1}{2}}\mathsf{e}^j_{m+1})$. Equations (31)–(34) hold between the spherical 3-j symbols, and the 3-j tensors with the phases we have assigned [$x_{000} = +1$ in Eqs. (81) and (82)], provided the absolute phase of the tensors $\mathsf{e}^j_m(r)$ is fixed by the assignment $\mathsf{e}^j_0 = (+i)^j N_j^{\frac{1}{2}} \mathsf{E}^j \cdot (\mathbf{k})^j$. Thus $\mathsf{e}^1_1 = -i2^{-\frac{1}{2}} \times (\mathbf{i} + i\mathbf{j})$. Note that the $\bar{Y}^{jm}$ in Eq. (100) are imaginary for j odd. Equation (99) shows that the 3-j spinor phases defined by Eq. (89), together with the phases of the $\mathsf{e}^j_m(v)$ of Eq. (97), are likewise consistent with the spherical symbols. The isometric mapping of contravariant symmetric spinor components to tensor components is given by the complex conjugate of the matrix $U^r_{v_1 v_2}$ of Paper II, Eq. (77). In particular the isometric mapping between 3-j spinors and 3-j tensors has the convenient phases given in Eqs. (102) and (103).

[29] The linear independence of the terms for different a, b, c follows from the uniqueness of the $\mathsf{T}(j_1, j_2, j_3)$ already proved. However, it is also implied by the possibility (Ref. 9) of an invariant decomposition of symmetric tensors $\mathsf{A}(r)$, $\mathsf{B}(s)$, $\mathsf{C}(t)$, and hence of their tensor product, into multiple traceless parts (cf. the discussion of E^j in Paper II).

[30] See the discussion of Eq. (85). $N_{0,0}$ gives the fractional number of nonvanishing traces in $\mathsf{A}^{(j_1)} \odot \mathsf{B}^{(j_2)}$.

[31] $\sum_{x,y} M_{x,y}(\alpha\beta \mid p) = (\alpha + \beta - 2p + 1)/(\alpha + \beta + 1)$ is not unity.

[32] That is, do *not* substitute $J \to J - 1$.

[33] nth-order tensors of symmetry j are not necessarily irreducible. If, however, $e^j_{\alpha\sigma}(n)$ is irreducible, then it can be shown to determine completely the irreducible subspace which contains it and, therefore, the appropriate projection $\mathbf{\Pi}^j_2$.

[34] G. Racah, Phys. Rev. **62**, 438 (1942).

[35] In particular, the matrix elements of a vector; see, e.g., E. U. Condon and G. H. Shortley, *Theory of Atomic Spectra* (Cambridge U.P., London, 1935).

[36] The Wigner–Eckart theorem is naturally expressed in terms of covariant components because the indices shown for the states are essentially basis labels.

[37] It can be calculated conveniently from the $zz\cdots z$ component of $\mathsf{A}^{(j_2)}$; e.g., for J even,

$$(-1)^{j_1}\begin{pmatrix} j_1 & j_2 & j_3 \\ 0 & 0 & 0 \end{pmatrix}\lambda = (+i)^{j_2}[(2j_2)!/2^{j_2}(j_2!)^2]^{\frac{1}{2}} \times \langle \psi_0^{(j_1)} | \mathbf{A}^{(j_2)}zz(\cdots z) \mid \psi^{(j_3)}\rangle.$$

[38] The basis is a particular case $\mathsf{T}_p(n \mid 0)$ of the $\mathsf{T}_p(n \mid \bar{m}_j)$ of Sec. I. Such basis tensors for crystallographic groups are discussed by G. F. Smith, J. Math. Phys. **5**, 1612 (1964).

[39] This result is given in a less condensed notation by R. F. Snider and C. F. Curtiss, Phys. Fluids **1**, 122 (1958).

[40] A. P. Yutsis, I. B. Levinson, and V. V. Vanagas, *Matematicheskii Apparat Teorii Momenta Kolichestva Dvizheniya* (Vilnius, 1960) [English translation: A. Sen and R. N. Sen, *Mathematical Apparatus of the Theory of Angular Momentum* (Israel Program for Scientific Translations, 1962)].

[41] N.B., if $y = \mathcal{R}x$, the covariant transformation is $y^* = (\mathcal{R}x)^*$ not $y^* = \mathcal{R}x^*$.

[42] It is intended to illustrate the definitions of Sec. II, not the general reduction formalism of Sec. I and Paper II. The well-known reduction of tensors under $GL(v)$ is not simplified by the formalism of Sec. I, there being no covariant invariant tensors and therefore no invariant maps to spaces of lower order.

[43] If A^{3E} is an irreducible tensor satisfying $\mathbf{\Pi}^{3E} \cdot \mathsf{A}^{3E} = \mathsf{A}^{3E}$, then it is mapped to the standard form by the permutation (23), i.e., by the invariant tensor $\mathsf{T}^{(0)}(r' \mid \bar{r}) = \delta_{r_1 r_1'}\delta_{r_2 r_3'}\delta_{r_3 r_2'}$.

[44] This invariant part is, of course, not given simply by integration over angles, and integrals over the whole space $\mathfrak{X}$ do not, in general, converge.

[45] M. S. de Groot and J. H. van der Waals, Mol. Phys. **2**, 333 (1959); **3**, 190 (1960); P. Kottis and R. Lefebvre, J. Chem. Phys. **39**, 396 (1963).

[46] These basis states are precisely the $\mathsf{T}_p(n \mid \bar{m}_j)$ of Sec. I with $n = m_j = j$ and p unnecessary.

[47] While the proportionality constants (reduced matrix elements) are found straightforwardly from Eqs. (87), (88) and Tables III, IV, both here and for the $\bar{\mathsf{Y}}^k$ in general, that for (A25) is more simply obtained by observing first that

$$\langle \boldsymbol{\Phi}_{yz\ldots z} | S_z | \boldsymbol{\Phi}_{xz\ldots z}\rangle = i\langle \boldsymbol{\Phi}_{yz\ldots z} \mid \boldsymbol{\Phi}_{yz\ldots z}\rangle = i\mathsf{E}^j(yz\cdots z \mid yz\cdots z),$$

and second, from Eq. (80), that

$$\mathsf{T}\begin{pmatrix} j & 1 & j \\ yz\cdots z & z & xz\cdots z \end{pmatrix} = j^{-1}\mathsf{E}^j(xz\cdots z \mid xz\cdots z) = j^{-1}\mathsf{E}^j(yz\cdots z \mid yz\cdots z).$$

Inequalities for Appell Functions

B. M. MINTON*
Physics Department, University of Calgary, Calgary, Alberta, Canada

(Received 2 December 1969)

The asymptotic expansion of one of Appell's generalizations of the Jacobi function is given for one parameter becoming large while the other is kept fixed. Inequalities are given which may be useful when both parameters become large.

1. INTRODUCTION

In a recent paper,[1] a double Sommerfeld–Watson transformation of a 2-variable expansion of the scattering amplitude was conjectured. The expansion was given as a double sum of polynomials of two variables, where the polynomials are Appell's generalization of the 1-variable Jacobi polynomials.

In contrast to the situation with Jacobi functions, the asymptotic parametric behavior of the Appell functions is unknown and so, in the Sommerfeld–Watson transformation, it was not possible to show that contributions from semicircles at infinity could be neglected.

We examine here the behavior of the Appell functions and give inequalities which indicate that the previous conjectures were justified.

2. APPELL FUNCTIONS

The functions with which we are concerned are Appell functions of the second and third kind. In the usual notation,[2] these are given by

$$E_{\mu\nu}(x, y) \equiv F^2[\mu + \nu + 2; -\mu, -\nu; 1, 1; x, y] \quad (1)$$

and

$$\bar{G}_{\mu\nu}(x^1, y^1) \equiv F^3[\mu + 1, \nu + 1; \mu + 1, \nu + 1; 2\mu + 2\nu + 3; 1/x^1, 1/y^1]. \quad (2)$$

Here μ and ν are complex with Re (μ), Re $(\nu) > -\frac{1}{2}$; we consider x, y, x^1, y^1 complex although we give an inequality for $\bar{G}_{\mu\nu}(x^1, y^1)$ only when x^1 and y^1 are real. $E_{\mu\nu}(x, y)$ and $\bar{G}_{\mu\nu}(x^1, y^1)$ are defined from Eqs. (1) and (2) for $|x| + |y| < 1$ and $|x^1|, |y^1| > 1$, respectively. We are most interested in what happens when $|\mu|$ and $|\nu|$ become large, either separately or together.

3. ASYMPTOTIC LIMIT: $|\mu| \to \infty$

When one parameter becomes large, say $|\mu| \to \infty$ at fixed ν, we may obtain the required asymptotic behavior using the results of Watson[3] for the one-variable Jacobi functions. Thus, by definition, from Eq. (1) we have

$$E_{\mu\nu}(x, y) = \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \frac{(\mu + \nu + 2)_{p+q}(-\mu)_p(-\nu)_q}{p!\, p!\, q!\, q!} x^p y^q$$
$$= \sum_{p=0}^{\infty} \frac{(\mu + \nu + 2)_p(-\mu)_p}{p!\, p!} x^p \times {}_2F_1[\mu + \nu + 2 + p, -\nu; 1; y], \quad (3)$$

where ${}_2F_1[a, b; c; x]$ is the ordinary Gauss hypergeometric function and

$$(a)_p = \Gamma(a + p)/\Gamma(a).$$

We may then use the relations between hypergeometric series[4] to rewrite the hypergeometric function in Eq. (3) as

$$\Gamma(-\nu)\Gamma(\nu + 1)\,{}_2F_1[\mu + \nu + 2 + p, -\nu; 1; y]$$
$$= e^{i\pi\nu} \frac{\Gamma(-\mu - \nu - 1 - p)\Gamma(-\nu)}{\Gamma(-\mu - 2\nu - 1 - p)} (-y)^{\nu} \times {}_2F_1\left[-\nu, -\nu; -\mu - 2\nu - p - 1; \frac{1}{y}\right]$$
$$+ e^{i\pi(\nu+1)} \frac{\Gamma(-\mu - \nu - 1 - p)\Gamma(\nu + 1)}{\Gamma(-\mu - p)} \times y^{-\nu-1}(1 - y)^{-\mu-1-p} \times {}_2F_1\left[\nu + 1, \nu + 1; -\mu - p; 1 - \frac{1}{y}\right] \quad (4)$$

under the usual conditions. Now for $|\mu| \to \infty$ with $|\arg \mu| < \pi$ and $|\nu|$ finite, we may determine the asymptotic limit of the ${}_2F_1$ functions on the right-hand side of Eq. (4),[5] i.e.,

$${}_2F_1[-\nu, -\nu; -\mu - 2\nu - p - 1; 1/y] \to 1,$$
$${}_2F_1[\nu + 1, \nu + 1; -\mu - p; 1 - 1/y] \to 1 \quad (5)$$

under these conditions. Since we are concerned with μ and ν in the right half complex plane, Eq. (5) holds for all p since p is positive; in fact, the larger p is, the better this limit is. After some manipulation we obtain,

from Eqs. (3)–(5),

$$E_{\mu\nu}(x, y) \sim e^{i\pi\nu}(-y)^{\nu}\frac{\Gamma(-\mu-\nu-1)}{\Gamma(-\mu-2\nu-1)\Gamma(\nu+1)}P_{\mu}^{(0,2\nu+1)}(1-2x) + e^{i\pi(\nu+1)}(1-y)^{-\mu-1}y^{-\nu-1}P_{\mu}\left(1+\frac{2x}{1-y}\right). \quad (6)$$

This expression is then taken to mean that, in the limit as $|\mu| \to \infty$ in the right half-plane with $|\nu|$ finite, $E_{\mu\nu}(x, y)$ behaves in the same way as the functions on the right-hand side of Eq. (6) in this same limit. The right-hand side of Eq. (6) contains Jacobi[6] and Γ functions, whose behavior in this limit is known, and so the problem is essentially solved. Equation (6) again illustrates the connection between the Appell and Jacobi functions and it is interesting to note that the first Jacobi function in Eq. (6) is that considered by Balachandran and Nuyts[7] in an alternative 2-variable expansion of the amplitude.

4. ASYMPTOTIC LIMIT: $|\mu|, |\nu| \to \infty$

For both $|\mu|$ and $|\nu| \to \infty$, it is clear that the approach of the preceding section does not work [Eq. (5) does not hold] and the derivation of rigorous asymptotic forms for $E_{\mu\nu}(x, y)$ and $\bar{G}_{\mu\nu}(x^1, y^1)$ has proved too difficult. We are able to show that

$$|E_{\mu\nu}(x, y)| < f(\mu, \nu: x, y)$$

and

$$|\bar{G}_{\mu\nu}(x^1, y^1)| < g(\mu, \nu: x^1, y^1), \quad (7)$$

where $f(\mu, \nu: x, y)$ and $g(\mu, \nu: x^1, y^1)$ have known properties. The results obtained hold for all μ, ν, x, y, x^1, and y^1 satisfying certain conditions and do not, in fact, require that $|\mu|$ and $|\nu|$ be large. It is hoped, however, that the inequalities are best in this limit.

We again consider $E_{\mu\nu}(x, y)$ to illustrate the method which is based on the use of a Mellin–Barnes double-integral representation[8] of a suitable transformation of $E_{\mu\nu}(x, y)$. We have then[2]

$$E_{\mu\nu} = (1-x-y)^{-\mu-\nu-2} \times F^2[\mu+\nu+2; \mu+1, \nu+1; 1, 1; p, q] = (1-x-y)^{-\mu-\nu-2}F^2[p, q], \quad (8)$$

where

$$p = x/(x+y-1), \quad q = y/(x+y-1),$$

and x and y are such that both sides of Eq. (8) are defined. For $F^2[p, q]$ defined by Eq. (8), we have

$$F^2[p, q] = -\frac{1}{4\pi^2}\frac{J}{\Gamma(\mu+\nu+2)\Gamma(\mu+1)\Gamma(\nu+1)},$$

with

$$J = \int_{c_1}\int_{c_2}\frac{\Gamma(\mu+\nu+2+s+t)\Gamma(\mu+1+s)\Gamma(\nu+1+t)\Gamma(-s)\Gamma(-t)}{\Gamma(1+s)\Gamma(1+t)}(-p)^s(-q)^t\,ds\,dt, \quad (9)$$

where c_1 and c_2 are contours in the s and t planes, respectively, chosen to separate the poles of $\Gamma(-s)$ and $\Gamma(-t)$ from those of $\Gamma(\mu+1+s)$, $\Gamma(\nu+1+t)$, and $\Gamma[\mu+\nu+2+s+t]$. In the $s(t)$ plane, $c_{1(2)}$ is taken to be the line $s(t) = -\frac{1}{2} + ip$, $-\infty < p < \infty$, together with a semicircle at infinity in the right half-plane. We must have Re (μ), Re $(\nu) > -\frac{1}{2}$ for this to be possible. Contributions from semicircles at infinity may be shown to be zero for $|p|, |q| < 1$ and $|\arg(-p)|, |\arg(-q)| < \frac{1}{2}\pi$. If we had used the original F^2 function in Eq. (1) instead of the transformed function $F^2[p, q]$, we would have needed to indent the straight-line contours to exclude poles of $\Gamma(-\mu+s)$ and $\Gamma(-\nu+t)$ at $s = \mu - m$ and $t = \nu - n$, $m, n = 0, 1, 2, \cdots$, and it would have been impossible to satisfy conditions like $|\arg(-x)| < \frac{1}{2}\pi$ for x real. Thus, we have

$$J = \int_{-\frac{1}{2}-\infty i}^{-\frac{1}{2}+\infty i} ds \int_{-\frac{1}{2}-\infty i}^{-\frac{1}{2}+\infty i} dt\,\frac{\Gamma(\mu+\nu+2+s+t)\Gamma(\mu+1+s)\Gamma(\nu+1+t)\Gamma(-s)\Gamma(-t)}{\Gamma(1+s)\Gamma(1+t)}(-p)^s(-q)^t. \quad (10)$$

This integral is convergent for $|\arg(-p)|, |\arg(-q)| < \frac{1}{2}\pi$ without the restrictions on $|p|$ and $|q|$ which apply to Eq. (9). Substituting in

$$s = -\tfrac{1}{2} + i[k - \text{Im}\,(\mu)],$$
$$t = -\tfrac{1}{2} + i[l - \text{Im}\,(\nu)]$$

and taking the modulus of both sides using

$$|\Gamma(x+iy)| = |\Gamma(x-iy)|,$$

we find

$$|J| \le (pq)^{-\frac{1}{2}}\exp[\text{Im}\,(\mu)\arg(-p) + \text{Im}\,(\nu)\arg(-q)]I, \quad (11)$$

with

$$I = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|\Gamma[\text{Re}\,(\mu)+\text{Re}\,(\nu)+ik+il+1] \times \Gamma[\text{Re}\,(\mu)+\tfrac{1}{2}+ik]\Gamma[\text{Re}\,(\nu)+\tfrac{1}{2}+il]| \times \exp[-k\arg(-p) - l\arg(-q)]\,dk\,dl. \quad (12)$$

The dependence on Im (μ) and Im (ν) is thus simply extracted from the integral. To proceed further, we need to approximate the Γ functions in the integrand in Eq. (12). We do this using Eq. (A4) of the Appendix. First, however, we split I of Eq. (12) into four parts for the four separate ranges of integration $-\infty < k \leq 0$, $0 \leq k < \infty$, $-\infty < l \leq 0$, and $0 \leq l < \infty$. If

$$I = I_1 + I_2,$$

where

$$I_1 = \left(\int_0^\infty \int_0^\infty + \int_{-\infty}^0 \int_{-\infty}^0\right) dk\, dl B(l, k),$$

$$I_2 = \left(\int_{-\infty}^0 \int_0^\infty + \int_0^\infty \int_{-\infty}^0\right) dk\, dl B(l, k), \quad (13)$$

and where $B(l, k)$ is the integrand in Eq. (12), then

$$I_1 < 2\int_0^\infty \int_0^\infty |\Gamma(\mathrm{Re}\,(\mu) + \mathrm{Re}\,(\nu) + 1 + ik + il) \times \Gamma(\mathrm{Re}\,(\mu) + \tfrac{1}{2}ik)\Gamma(\mathrm{Re}\,(\nu) + \tfrac{1}{2} + il)| \times \exp\,[k\,|\arg(-p)| + l\,|\arg(-q)|]\, dk\, dl. \quad (14)$$

For large y, $|\Gamma(x + iy)| \sim e^{-\pi/2|y|}$, and so the integral in Eq. (14) converges for $|\arg(-p)|, |\arg(-q)| < \pi$. Substituting in the Γ-function approximations from Eq. (A4) of the Appendix, we obtain

$$I_1 < B(4\gamma/\pi)^{-2\,\mathrm{Re}\,(\mu)-2\,\mathrm{Re}\,(\nu)}[\Gamma(2\,\mathrm{Re}\,(\mu) + 2)\Gamma(2\,\mathrm{Re}\,(\nu) + 2)\Gamma(2\,\mathrm{Re}\,(\mu) + 2\,\mathrm{Re}\,(\nu) + 3)]^{\frac{1}{2}} \times \int_0^\infty \int_0^\infty \exp\,[k(|\arg(-p)| + 2\gamma - \pi) + l(|\arg(-q)| + 2\gamma - \pi)] \times \frac{\{[\mathrm{Re}\,(\mu) + k + 1][\mathrm{Re}\,(\nu) + l + 1][\mathrm{Re}\,(\mu) + \mathrm{Re}\,(\nu) + \frac{3}{2} + k + l]\}^{\frac{1}{2}}}{|\mathrm{Re}\,(\mu) + \frac{1}{2} + ik|\,|\mathrm{Re}\,(\nu) + \frac{1}{2} + il|\,|\mathrm{Re}\,(\mu) + \mathrm{Re}\,(\nu) + 1 + ik + il|}\, dk\, dl. \quad (15)$$

The integral in Eq. (15) now converges for

$$|\arg(-p)| < \pi - 2\gamma, \quad |\arg(-q)| < \pi - 2\gamma,$$

where γ is such that $\gamma = e^{-\gamma} \sim 0.6$ and B is a numerical constant independent of μ and ν.

Since the most important contribution to the integral in Eq. (15) comes from l and k small, we may approximate the last term in the integrand by

$$[\mathrm{Re}\,(\mu)\,\mathrm{Re}\,(\nu)\,\mathrm{Re}\,(\mu + \nu)]^{-\frac{1}{2}}$$

to give

$$I_1 < B^1(4\gamma/\pi)^{-2\,\mathrm{Re}\,(\mu)-2\,\mathrm{Re}\,(\nu)}\,[\Gamma(2\,\mathrm{Re}\,(\mu) + 2)\Gamma(2\,\mathrm{Re}\,(\nu) + 2)]^{\frac{1}{2}} \times [\Gamma(2\,\mathrm{Re}\,(\mu) + 2\,\mathrm{Re}\,(\nu) + 3)]^{\frac{1}{2}}\,[\mathrm{Re}\,(\mu)\,\mathrm{Re}\,(\nu)\,\mathrm{Re}\,(\mu + \nu)]^{-\frac{1}{2}}. \quad (16)$$

Constants of integration such as $|\arg(-p)| + 2\gamma - \pi$ have been absorbed into the constant B^1. A similar result is obtained for I_2 with the convergence conditions

$$|\arg(-p)| < \tfrac{1}{2}\pi - \gamma, \quad |\arg(-q) < \tfrac{1}{2}\pi - \gamma.$$

Finally, collecting together all relevant terms and simplifying, we obtain

$$|E_{\mu\nu}(x, y)| < B\left(\frac{16\gamma^2}{\pi^2}\,|1 - x - y|\right)^{-\mathrm{Re}\,(\mu)-\mathrm{Re}\,(\nu)} [\mathrm{Re}\,(\mu)\,\mathrm{Re}\,(\nu)\,\mathrm{Re}\,(\mu + \nu)]^{-\frac{1}{2}} \times \exp\,[\mathrm{Im}\,(\mu)\arg(x) + \mathrm{Im}\,(\nu)\arg(y)] \times \frac{[\Gamma(2\,\mathrm{Re}\,(\mu) + 2)\Gamma(2\,\mathrm{Re}\,(\nu) + 2)\Gamma(2\,\mathrm{Re}\,(\mu) + 2\,\mathrm{Re}\,(\nu) + 3)]^{\frac{1}{2}}}{|\Gamma(\mu + 1)\Gamma(\nu + 1)\Gamma(\mu + \nu + 2)|}, \quad (17)$$

for $\mathrm{Re}\,(\mu)$, $\mathrm{Re}\,(\nu) > -\frac{1}{2}$, and

$$|\arg(x) - \arg(1 - x - y)| < \tfrac{1}{2}\pi - \gamma,$$
$$|\arg(y) - \arg(1 - x - y)| < \tfrac{1}{2}\pi - \gamma. \quad (18)$$

Although Eqs. (8) and (9) were only valid with restrictions on $|x|$ and $|y|$, Eq. (10) converges for all values of x and y for which $|p|$ and $|q|$ remain finite, subject only to the restrictions on $|\arg(-p)|$ and $|\arg(-q)|$. Equation (10) may then be used to define the continuation of the F^2 function for values of $|p|$ and $|q|$, for which Eqs. (8) and (9) are not defined. Of course, we may only close the contours in the right half-planes to obtain the original F function if we have the restriction $|p|, |q| < 1$. Since we have only used Eq.

(10) in obtaining the result (17), it follows that Eq. (17) is valid for all x and y for which conditions (18) are satisfied and for which $|p|$ and $|q|$ are finite.

Equation (17) now gives the required results as the large μ and ν behavior of all the factors on the right-hand side is known. This result may be simplified by use of the Legendre duplication formula or Eq. (A5) to give a form reminiscent of Watson's asymptotic formulas for the Jacobi functions, i.e.,

$$\begin{aligned}|E_{\mu\nu}(x, y)| < {} & B^1((4\gamma^2/\pi^2)\,|1 - x - y|)^{-\mathrm{Re}\,(\mu)-\mathrm{Re}\,(\nu)} \\ & \times \exp\left[\mathrm{Im}\,(\mu)\arg(x) + \mathrm{Im}\,(\nu)\arg(y)\right. \\ & \left. + \pi\,|\mathrm{Im}\,(\mu)| + \pi\,|\mathrm{Im}\,(\nu)|\right] \\ & \times O(\mu^{-\frac{1}{2}}\nu^{-\frac{1}{2}}(\mu + \nu)^{-1}).\end{aligned}$$

We may use a similar though more complicated method to find the behavior of $\bar{G}_{\mu\nu}(x^1, y^1)$. We then find

$$\begin{aligned}&|\bar{G}_{\mu\nu}(x^1, y^1)| \\ &\quad < B\left|\frac{\Gamma(2\mu + 2\nu + 3)}{\Gamma(\nu + 1)\Gamma(\nu + 1)\Gamma(\mu + 1)\Gamma(\mu + 1)}\right| \\ &\qquad \times (2\gamma/\pi)^{-\mathrm{Re}\,(\mu)-2\,\mathrm{Re}\,(\nu)}\,|\mu^{-\frac{1}{2}}\nu^{-\frac{1}{2}}(\mu + \nu)^{-\frac{1}{2}}| \\ &\qquad \times \frac{\Gamma(2\,\mathrm{Re}\,(\mu) + 2)\Gamma(2\,\mathrm{Re}\,(\nu) + 2)}{[\Gamma(4\,\mathrm{Re}\,(\mu) + 4\,\mathrm{Re}\,(\nu) + 5)]^{\frac{1}{2}}}, \qquad (19)\end{aligned}$$

again for $\mathrm{Re}\,(\mu)$, $\mathrm{Re}\,(\nu) > -\frac{1}{2}$ and where we have assumed for simplicity that x^1 and y^1 are real and greater than unity. The results (18) and (19) are now in simple enough form to allow them to be used in connection with the Sommerfeld–Watson transform referred to previously.

ACKNOWLEDGMENT

I should like to thank Dr. J. M. Charap for his help and encouragement with this work.

APPENDIX: THREE Γ-FUNCTION INEQUALITIES

Erber[8] quotes an integral of Cauchy and derives a Γ-function inequality, namely,

$$|\Gamma(x + iy)| < 2(\tfrac{1}{4}\pi y)^{x+\frac{1}{2}}\frac{[\Gamma(2x + 2)]^{\frac{1}{2}}}{|x + iy|}\,e^{1-\pi/2|y|}, \qquad (A1)$$

for $x > 0$ and y real, $|y| \geq 2/\pi$.

We use the same integral to obtain two other bounds which are more suitable for our purpose. Cauchy's integral is

$$\begin{aligned}&|\Gamma(x + iy)|^2 \\ &\quad = \pi 2^{-2x}\frac{\Gamma(2x + 1)}{|x + iy|^2}\,e^{-\pi\nu}\left(\int_0^\pi d\psi(\sin\psi)^{2x}e^{-2\psi|\nu|}\right)^{-1}. \qquad (A2)\end{aligned}$$

Now

$$\int_0^\pi d\psi(\sin\psi)^{2x}e^{-2\psi|\nu|} > \int_0^\gamma d\psi(\sin\psi)^{2x}e^{-2\psi|\nu|}, \quad \gamma < \pi.$$

For $\gamma < \frac{1}{2}\pi$ we have $\sin\psi \geq 2\psi/\pi$, and, if we take γ such that $\gamma = e^{-\gamma}$, we also have for the range of integration $e^{-\psi} \geq \psi$. Thus,

$$\begin{aligned}&\int_0^\pi d\psi(\sin\psi)^{2x}e^{-2\psi|\nu|} \\ &\qquad > \left(\frac{2}{\pi}\right)^{2x}\int_0^\gamma \psi^{2x}e^{-2\psi|\nu|}\,d\psi \\ &\qquad > \left(\frac{2}{\pi}\right)^{2x}\int_0^\gamma \psi^{2x+2|\nu|}\,d\psi \\ &\qquad = \left(\frac{2}{\pi}\right)^{2x}(2x + 2\,|y| + 1)^{-1}\gamma^{2x+1}e^{-2\gamma|\nu|}. \qquad (A3)\end{aligned}$$

Thus, substituting from Eq. (A3) into Eq. (A2), we get

$$\begin{aligned}&|\Gamma(x + iy)| \\ &< (\pi\gamma)^{-\frac{1}{2}}\left(\frac{4\gamma}{\pi}\right)^{-x}e^{|\nu|(\gamma-\frac{1}{2}\pi)}\,\frac{[\Gamma(2x + 1)]^{\frac{1}{2}}(x + |y| + \frac{1}{2})^{\frac{1}{2}}}{|x + iy|}, \qquad (A4)\end{aligned}$$

where $\gamma = e^{-\gamma} \simeq 0.6$.

Similarly, we may show that

$$\begin{aligned}&|\Gamma(x + iy)| \\ &> \pi^{-\frac{1}{2}}2^{-x+\frac{1}{2}}(|y|^{\frac{1}{2}}/|x + iy|)e^{-\frac{1}{2}\pi|\nu|}[\Gamma(2x + 1)]^{\frac{1}{2}}. \qquad (A5)\end{aligned}$$

We use Eq. (A4) in the $E_{\mu\nu}(x, y)$ case and Eqs. (A1), (A4), and (A5) in the $\bar{G}_{\mu\nu}(x^1, y^1)$ case.

* Supported in part by the National Research Council of Canada.

[1] J. M. Charap and B. M. Minton, J. Math. Phys. **10**, 1823 (1969).
[2] P. Appell and J. Kampe de Feriet, *Fonctions hypergeometriques et hyperspheriques* (Gauthier-Villars, Paris, 1926).
[3] G. N. Watson, Trans. Cambridge Phil. Soc. **22**, 277 (1918).
[4] *Higher Transcendental Functions*, A. Erdélyi, Ed. (McGraw-Hill, New York, 1953), Vol. 1, p. 106, Eq. (27).
[5] Reference 4, p. 76, Eq. (10).
[6] Reference 4, Vol. 2, p. 169.
[7] A. P. Balachandran and J. Nuyts, Phys. Rev. **172**, 1821 (1968).
[8] T. Erber, Arch. Ratl. Mech. Anal. **4**, 342 (1959–60). This paper gives a bound for another F^2 function based on the Mellin–Barnes representation, but the choice of parameters leads to simplifications and the bound is not suitable for our purposes.

Diffraction by a Circular Cavity

DONALD LUDWIG
Courant Institute of Mathematical Sciences, Division of Electromagnetic Research
251 *Mercer St., New York, N.Y.* 10012

(Received 11 July 1969)

The reduced wave equation $\Delta u + k^2n^2u = 0$ is considered, where $n = 1$ in the exterior of a circular cylinder of radius a and $n = N < 1$ in the interior. The solution and its first derivatives are required to be continuous for $r = a$. The behavior of the solution for large ka is described, especially in the vicinity of a ray incident upon the cylinder at the critical angle. The principal novelty of the present results is a description of the solution in certain regions by a combination of geometrical optics (ray contributions) and whispering gallery modes (a kind of diffracted modes). These results are in qualitative agreement with those of Chen, but certain details in Chen's treatment require elaboration. The present results do not agree with those of Nussenzveig for $N < 1$, especially in the geometrical interpretation. This revision of Nussenzveig's work brings it into agreement with Keller's geometrical theory of diffraction. In fact, part of the present work can be combined with Chen's results to give a fairly complete geometrical theory for diffraction by a convex, transparent object. However, the present treatment does not give a uniform asymptotic expansion of the solution, since the expansion is not given in a transition region in the immediate vicinity of the critically reflected ray.

1. INTRODUCTION

The reduced wave equation $\Delta u + k^2n^2u = 0$ is considered, where $n = 1$ in the exterior of a circular cylinder of radius a and $n = N < 1$ in the interior. The solution and its first derivatives are required to be continuous for $r = a$. The behavior of the solution for large ka is described, especially in the vicinity of a ray incident upon the cylinder at the critical angle. The principal novelty of the present results is a description of the solution in certain regions by a combination of geometrical optics (ray contributions) and whispering gallery modes (a kind of diffracted modes). These results are in qualitative agreement with those of Chen,[1,2] but certain details in Chen's treatment require elaboration. The present results do not agree with those of Nussenzveig[3] for $N < 1$, especially in the geometrical interpretation. This revision of Nussenzveig's work brings it into agreement with Keller's geometrical theory of diffraction. In fact, Sec. 4 and Appendix C of the present work can be combined with Chen's results to give a fairly complete geometrical theory for diffraction by a convex, transparent object. However, the present treatment does not give a uniform asymptotic expansion of the solution, since the expansion is not given in a transition region in the immediate vicinity of the critically reflected ray.

Section 2 contains a description of the exact solution and certain asymptotic approximations, which are obtained by replacing Bessel and Hankel functions by their asymptotic expansions. In Sec. 3, a formal analysis of the solution is given in terms of various saddle-point and residue contributions, and these terms are interpreted geometrically. Section 4 contains a description of the integration contours appropriate to the various regions in which the solution simplifies and thus summarizes the asymptotic behavior of the solution. The relationship of the present work to previous work is described in Sec. 5.

2. THE EXACT SOLUTION

In this section the exact solution for a plane wave incident upon a circular cylinder is given, and it is transformed into a form suitable for asymptotic evaluation. This involves replacing the Bessel and Hankel functions by their Debye expansions or by their uniform asymptotic expansions in terms of Airy functions. In certain regions, it is also desirable to represent the solution in terms of successive surface interactions (internal reflections, etc.). These interactions correspond to successive terms in a geometrical series (the Debye series).

In terms of polar coordinates r and θ, the solution $u(r, \theta, k)$ is required to satisfy the following:

$$\Delta u + k^2u = 0 \quad \text{for} \quad r > a; \tag{2.1}$$

$$\Delta u + k^2N^2u = 0 \quad \text{for} \quad r < a, \quad N < 1; \tag{2.2}$$

$$u \text{ and } \partial u/\partial r \text{ are continuous for } r = a; \tag{2.3}$$

$$u = e^{ikr\cos\theta} + u_s(r, \theta, k) \quad \text{for} \quad r \geq a, \tag{2.4}$$

where u_s satisfies the radiation condition

$$\lim_{r\to\infty} r^{\frac{1}{2}}\left(\frac{\partial u_s}{\partial r} - iku_s\right) = 0. \tag{2.5}$$

By use of the Poisson summation formula, the incident field may be represented in the form

$$e^{ikr\cos\theta} = \sum_{m=-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i\lambda(\frac{1}{2}\pi-\theta+2\pi m)} J_\lambda(kr)\, d\lambda. \tag{2.6}$$

Separation of variables leads to a solution in the form

$$u(r, \theta, k) = u^{(1)}(r, \theta, k) + u^{(2)}(r, \theta, k) \quad \text{for} \quad r \geq a, \tag{2.7}$$

where

$$u^{(1)}(r, \theta, k) = -\tfrac{1}{2} \sum_{m=-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i\lambda(\frac{1}{2}\pi - \theta + 2\pi m)} A H_\lambda^{(1)}(kr) \frac{H_\lambda^{(2)}(ka)}{H_\lambda^{(1)}(ka)}\, d\lambda, \tag{2.8}$$

$$u^{(2)}(r, \theta, k) = \tfrac{1}{2} \sum_{m=-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i\lambda(\frac{1}{2}\pi - \theta + 2\pi m)} H_\lambda^{(2)}(kr)\, d\lambda. \tag{2.9}$$

Here we have set

$$A = \left(\frac{H_\lambda^{(2)'}(ka)}{H_\lambda^{(2)}(ka)} - N \frac{J_\lambda'(kNa)}{J_\lambda(kNa)}\right) \Big/ \left(\frac{H_\lambda^{(1)'}(ka)}{H_\lambda^{(1)}(ka)} - N \frac{J_\lambda'(kNa)}{J_\lambda(kNa)}\right). \tag{2.10}$$

For $r \leq a$, the solution is given by

$$u(r, \theta, k) = \tfrac{1}{2} \sum_{m=-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i\lambda(\frac{1}{2}\pi - \theta + 2\pi m)} B J_\lambda(kNr) \frac{H_\lambda^{(2)}(ka)}{J_\lambda(kNa)}\, d\lambda, \tag{2.11}$$

where

$$B = \left(\frac{H_\lambda^{(1)'}(ka)}{H_\lambda^{(1)}(ka)} - \frac{H_\lambda^{(2)'}(ka)}{H_\lambda^{(2)}(ka)}\right) \Big/ \left(\frac{H_\lambda^{(1)'}(ka)}{H_\lambda^{(1)}(ka)} - N \frac{J_\lambda'(kNa)}{J_\lambda(kNa)}\right). \tag{2.12}$$

Asymptotic Expansions of the Cylinder Functions

Since the vicinity of a critically incident ray is of primary concern, the region of integration near $\lambda = kNa$ is most significant. By assumption, kNa is large, and hence the Debye expansions of the Hankel functions are valid (see Sommerfeld[4]): If $\alpha^2 < r^2$, then

$$H_{k\alpha}^{(1),(2)}(kr) \sim [2/\pi k(r^2 - \alpha^2)^{\frac{1}{2}}]^{\frac{1}{2}} \times \exp\left[\pm ik((r^2 - \alpha^2)^{\frac{1}{2}} - \alpha \cos^{-1}(\alpha/r)) \mp \tfrac{1}{4} i\pi\right]. \tag{2.13}$$

If α is near Na, then $J_{k\alpha}(kNa)$ can be uniformly approximated in terms of the Airy function. We prefer to use $V(\omega)$, defined by

$$V(\omega) = \text{Ai}\,(-\omega), \tag{2.14}$$

where Ai is the Airy function (see Erdélyi[5]). Then, if $\alpha < Nr$,

$$J_{k\alpha}(kNr) \sim \frac{2^{\frac{1}{2}}}{k^{\frac{1}{3}}} \left(\frac{\rho(r, \alpha)}{N^2 r^2 - \alpha^2}\right)^{\frac{1}{4}} V(k^{\frac{2}{3}} \rho(r, \alpha)), \tag{2.15}$$

where

$$\tfrac{2}{3}[\rho(r, \alpha)]^{\frac{3}{2}} = (N^2 r^2 - \alpha^2)^{\frac{1}{2}} - \alpha \cos^{-1}(\alpha/Nr). \tag{2.16}$$

If $\alpha > Nr$, the same formula holds if $\rho(r, \alpha)$ is defined by

$$\tfrac{2}{3}[-\rho(r, \alpha)]^{\frac{3}{2}} = \alpha \cosh^{-1}(\alpha/Nr) - (\alpha^2 - N^2 r^2)^{\frac{1}{2}}. \tag{2.17}$$

The function $\rho(r, \alpha)$ is regular where $\alpha = Nr$, and it has a simple zero there.

If the approximations (2.13) and (2.15) are inserted into (2.8), the result is

$$u^{(1)}(r, \theta, k) \sim \sqrt{k} \sum_{m=-\infty}^{\infty} \int_{-a}^{a} e^{ik\phi(r, \theta, \alpha, m)} \times Z(r, \alpha) \frac{V(\omega) + \zeta V'(\omega)}{V(\omega) - \zeta V'(\omega)}\, d\alpha + v^{(1)}(r, \theta, k), \tag{2.18}$$

where

$$\phi(r, \theta, \alpha, m) = (r^2 - \alpha^2)^{\frac{1}{2}} - 2(a^2 - \alpha^2)^{\frac{1}{2}} + \alpha[-\cos^{-1}(\alpha/r) + 2\cos^{-1}(\alpha/a) + \tfrac{1}{2}\pi - \theta + 2\pi m], \tag{2.19}$$

$$Z(r, \alpha) = [2\pi(r^2 - \alpha^2)^{\frac{1}{2}}]^{-\frac{1}{2}} e^{\frac{1}{4} i\pi}, \tag{2.20}$$

$$\zeta = -ik^{-\frac{1}{3}} \rho_r(a, \alpha)/\phi_r(a, \alpha), \tag{2.21}$$

$$\omega = k^{\frac{2}{3}} \sigma(\alpha) = k^{\frac{2}{3}} \rho(a, \alpha), \tag{2.22}$$

$$v^{(1)}(r, \theta, k) = -\tfrac{1}{2} \sum_{m=-\infty}^{\infty} \int_{|\lambda| \geq ka} e^{i\lambda(\frac{1}{2}\pi - \theta + 2\pi m)} A H_\lambda^{(2)}(ka) \frac{H_\lambda^{(1)}(kr)}{H_\lambda^{(1)}(ka)}\, d\lambda. \tag{2.23}$$

Similarly, if (2.13) is inserted into (2.9), the result is

$$u^{(2)}(r, \theta, k) \sim k \sum_{m=-\infty}^{\infty} \int_{-a}^{a} e^{ik\phi_i(r, \theta, \alpha, m)} Z(r, \alpha)\, d\alpha + v^{(2)}(r, \theta, k), \tag{2.24}$$

where

$$\phi_i(r, \theta, \alpha, m) = -(r^2 - \alpha^2)^{\frac{1}{2}} + \alpha[\cos^{-1}(\alpha/r) + \tfrac{1}{2}\pi - \theta + 2\pi m], \tag{2.25}$$

$$v^{(2)}(r, \theta, k) = \tfrac{1}{2} \sum_{m=-\infty}^{\infty} \int_{|\lambda| \geq ka} e^{i\lambda(\frac{1}{2}\pi - \theta + 2\pi m)} H_\lambda^{(2)}(kr)\, d\lambda. \tag{2.26}$$

For $r \leq a$, we obtain

$$u(r, \theta, k) \sim \sqrt{k} \sum_{m=-\infty}^{\infty} \int_{-a}^{a} e^{ik\psi(\theta,\alpha,m)} \times \frac{V(k^{\frac{2}{3}}\rho(r, \alpha))g(r, \alpha)2}{V(\omega) - \zeta V'(\omega)} d\alpha + v(r, \theta, k). \quad (2.27)$$

Here,

$$\psi(\theta, \alpha, m) = -(a^2 - \alpha^2)^{\frac{1}{2}} + \alpha[\cos^{-1}(\alpha/a) + \tfrac{1}{2}\pi - \theta + 2\pi m], \quad (2.28)$$

$$g(r, \alpha) = [2\pi(a^2 - \alpha^2)^{\frac{1}{2}}]^{-\frac{1}{2}} \left(\frac{\rho(r, \alpha)(N^2a^2 - \alpha^2)}{(N^2r^2 - \alpha^2)\rho(a, \alpha)}\right)^{\frac{1}{4}} e^{\frac{1}{4}i\pi}, \quad (2.29)$$

$$v(r, \theta, k) = \tfrac{1}{2} \sum_{m=-\infty}^{\infty} \int_{|\lambda| \geq a} e^{i\lambda(\frac{1}{2}\pi - \theta + 2\pi m)} BH_{\lambda}^{(2)}(ka) \frac{J_{\lambda}(kNr)}{J_{\lambda}(kNa)} d\lambda. \quad (2.30)$$

The Debye Series

Integrals such as (2.18) and (2.27) are difficult to deal with because of the factor $V(\omega) - \zeta V'(\omega)$ in the denominator. Even if the asymptotic expansion of the Airy function is introduced, the saddle-point method is not applicable, because the asymptotic expansion of V corresponds to waves moving in two different directions. In certain circumstances, this difficulty can be overcome by separating the denominator into two pieces and expanding the quotient in a geometric series. Each term can be interpreted in terms of interactions at the interface and reflections at the origin.

We first define the $V_{\pm}(\omega)$, which correspond to outgoing and incoming waves:

$$V_{\pm}(\omega) = V(e^{\pm\frac{2}{3}\pi i}\omega). \quad (2.31)$$

The connection between V, V_+, and V_- is given by

$$V(\omega) = e^{\frac{1}{3}i\pi}V_-(\omega) + e^{-\frac{1}{3}i\pi}V_+(\omega). \quad (2.32)$$

For $\omega \gg 1$, the first term in the asymptotic expansion of V_+ is given by

$$V_{\pm}(\omega) \sim \frac{e^{\pm\frac{1}{12}i\pi}}{2(\sqrt{\pi})\omega^{\frac{1}{4}}} e^{\pm\frac{2}{3}i\omega^{\frac{3}{2}}}. \quad (2.33)$$

Thus (2.32) and (2.33) show that the denominator in (2.18) and (2.27) contains two different phases where $\omega \gg 1$, i.e., where $\rho(a, \alpha) > 0$. We write

$$V(\omega) - \zeta V'(\omega) = e^{\frac{1}{3}i\pi}[V_-(\omega) - \zeta V'_-(\omega)](1 - R), \quad (2.34)$$

where

$$R = e^{\frac{1}{3}i\pi}[V_+(\omega) - \zeta V'_+(\omega)]/[V_-(\omega) - \zeta V'_-(\omega)]. \quad (2.35)$$

After insertion of (2.34), (2.27) becomes (formally)

$$u(r, \theta, k) \sim \sqrt{k} \sum_{m=-\infty}^{\infty} \sum_{l=0}^{\infty} \int_{-a}^{a} e^{ik\psi(\theta,\alpha,m)} V(k^{\frac{2}{3}}\rho(r, \alpha)) \times 2\frac{e^{-\frac{1}{3}i\pi}g(r, \alpha)R^l d\alpha}{V_-(\omega) - \zeta V'_-(\omega)} + v(r, \theta, k). \quad (2.36)$$

An alternate form of (2.36) is obtained by inserting (2.32), namely,

$$u(r, \theta, k) \sim \sum_{m=-\infty}^{\infty} \sum_{l=0}^{\infty} [u_l^-(r, \theta, k) + u_l^+(r, \theta, k)] + v(r, \theta, k), \quad (2.37)$$

where

$$u_l^{\pm}(r, \theta, k) = [e^{-\frac{2}{3}\pi i}]\sqrt{k} \int_{-a}^{a} e^{ik\psi(\theta,\alpha,m)} \times \frac{V_{\pm}(k^{\frac{2}{3}}\rho(r, \alpha))2g(r, \alpha)R^l d\alpha}{V_-(\omega) - \zeta V'_-(\omega)}. \quad (2.38)$$

Here the factor in brackets appears only in the definition of u_l^+.

We note that u_0^- is identical to the integral in (2.27), if V is replaced by V_- throughout. This amounts to considering solutions of the original boundary-value problem with radial dependence $H_{\lambda}^{(2)}(kNr)$ instead of $J_{\lambda}(kNr)$ for $r \leq a$. Such solutions would be obtained by imposing a "radiation condition" at the interface, instead of a regularity condition at $r = 0$. The radiation condition amounts to the requirement that the solution for $r < a$ have a phase which increases with distance from the boundary (see Keller and Lewis[6]).

Solutions with radial dependence $H_{\lambda}^{(2)}(kNr)$ [or $V_-(k^{\frac{2}{3}}\rho(r, \alpha))$] are unsatisfactory because of the singularity at $r = 0$. This singularity is cancelled by addition of a term with radial dependence $H_{\lambda}^{(1)}(kNr)$ [or $V_+(k^{\frac{2}{3}}\rho(r, \alpha))$]. This explains the term u_0^+ in (2.37). Geometrically, u_0^+ represents a wave incident upon the interface from inside. The interaction produces a reflected wave inside (u_1^-) and a transmitted wave outside. Thus, each term in (2.37) has an interpretation in terms of successive interactions at the interface and at the origin.

An analogous decomposition is possible for the scattered field (2.18). Replacing V by V_- throughout, we define $u_r(r, \theta, k)$ by

$$u_r(r, \theta, k) = \sqrt{k} \sum_{m=-\infty}^{\infty} \int_{-a}^{a} e^{ik\phi(r,\theta,\alpha,m)} \times Z(r, \alpha) \frac{V_-(\omega) + \zeta V'_-(\omega)\, d\alpha}{V_-(\omega) - \zeta V'_-(\omega)}. \quad (2.39)$$

We write

$$u^{(1)}(r, \theta, k) = u_r(r, \theta, k) + u_r'(r, \theta, k) + v^{(1)}(r, \theta, k), \tag{2.40}$$

where

$$u_r'(r, \theta, k) = \sqrt{k} \sum_{m=-\infty}^{\infty} \int_{-a}^{a} \frac{e^{ik\phi(r,\theta,\alpha,m)} Z(r, \alpha) e^{\frac{1}{6}i\pi} \zeta \, d\alpha}{\pi(V - \zeta V')(V_- - \zeta V_-')}. \tag{2.41}$$

Here we have used the Wronskian relation

$$V'(\omega)V_-(\omega) - V(\omega)V_-'(\omega) = e^{\frac{1}{6}i\pi}/2\pi. \tag{2.42}$$

After substitution of (2.34), (2.41) becomes (formally)

$$u_r'(r, \theta, k) = \sqrt{k} \sum_{m=-\infty}^{\infty} \sum_{l=0}^{\infty} \int_{-a}^{a} e^{ik\phi(r,\theta,\alpha,m)} \times Z(r, \alpha) \frac{e^{-\frac{1}{6}i\pi} \zeta R^l \, d\alpha}{\pi(V_- - \zeta V_-')^2}. \tag{2.43}$$

The interpretation of (2.43) is similar to the interpretation of (2.37).

3. FORMAL ANALYSIS AND GEOMETRICAL INTERPRETATION OF THE SOLUTION

This section contains a formal treatment of the solution, to show in detail how various parts of the solution are given by geometrical optics. First, the geometry of transmission and reflection at an interface is described. These simple notions lead at once to complicated systems of rays for diffraction by a transparent cylinder. When the Airy functions which appear in the integrands of the previous section are replaced by their asymptotic expansions, these integrands can be interpreted in terms of reflection and transmission coefficients. The saddle-point method then shows that (2.24) yields the incident field, (2.36) represents the transmitted and multiply reflected field inside the cylinder, (2.39) represents the reflected field outside the cylinder, and (2.43) represents the transmitted, multiply reflected, and retransmitted field outside the cylinder. Certain poles and residues of (2.18) and (2.27) are also examined and interpreted in terms of whispering gallery modes, i.e., waves which travel just inside the interface and shed energy outside at the critical angle.

Geometrical Optics at the Interface

We look for approximate solutions of (2.1)–(2.5) for large k in the form

$$u_s(r, \theta) = e^{ik\phi_s(r,\theta)} Z_s(r, \theta), \quad \text{for} \quad r \geq a,$$
$$u_1(r, \theta) = e^{ik\phi_1(r,\theta)} Z_1(r, \theta), \quad \text{for} \quad r \leq a. \tag{3.1}$$

It follows from substitution of (3.1) into (2.1) and (2.2) that

$$(\nabla\phi_1)^2 = N^2, \quad (\nabla\phi_s)^2 = 1. \tag{3.2}$$

The ray direction is the direction of $\nabla\phi$. Further details are given by Keller and Lewis.[6]

The boundary condition (2.3) implies that

$$a \cos\theta = \phi_s(a, \theta) = \phi_1(a, \theta), \tag{3.3}$$

$$1 + Z_s = Z_1, \quad \text{for} \quad r = a, \tag{3.4}$$

$$\cos\theta + \left(\frac{\partial\phi_s}{\partial r}\right) Z_s = \left(\frac{\partial\phi_1}{\partial r}\right) Z_1, \quad \text{for} \quad r = a. \tag{3.5}$$

In view of (3.3), the tangential components of $\nabla\phi_s$, $\nabla\phi_1$, and $\nabla(r\cos\theta)$ are equal for $r = a$; it follows from (3.2) that

$$\sin\gamma_i = N \sin\gamma_t, \quad \text{for} \quad r = a, \tag{3.6}$$

where γ_i and γ_t are the angles of incidence and refraction, respectively. Now Z_s and Z_1 can be determined from (3.4) and (3.5): We define

$$\xi = (N^2 - \sin^2\gamma_i)^{\frac{1}{2}}/\cos\gamma_i; \tag{3.7}$$

then

$$Z_s = R_{22} = (1 - \xi)/(1 + \xi), \quad \text{for} \quad r = a, \tag{3.8}$$

$$Z_1 = T_{21} = 2/(1 + \xi), \quad \text{for} \quad r = a. \tag{3.9}$$

These coefficients are called the reflection and transmission coefficients for incidence from outside the cylinder. In a similar way, we can prescribe an incident field in region 1, and outgoing conditions on the remainder of the solution. The result is

$$R_{11} = (\xi - 1)/(\xi + 1), \tag{3.10}$$

$$T_{12} = 2\xi/(1 + \xi). \tag{3.11}$$

The preceding theory breaks down if $\sin\gamma_i = N$, since then $\gamma_t = \frac{1}{2}\pi$, and the transmitted ray is tangent to the circle. However, if $\sin\gamma_i > N$, then we can find a complex ϕ_1, and we require that $\operatorname{Im}\phi_1$ increase as r decreases. Such a solution is called an evanescent wave, and the field is said to be totally reflected in the exterior. If (3.7) is replaced by

$$\xi = i(\sin^2\gamma_i - N^2)^{\frac{1}{2}}/\cos\gamma_i, \tag{3.12}$$

then (3.8) and (3.9) remain unchanged.

The Geometry of the Rays

The critically incident rays correspond to $\sin\gamma_i = N$, and they strike the cylinder at heights $\pm Na$. These rays are reflected at the same angle, and the region between these reflected rays is the region of ordinary reflection (denoted by O.R.; see Fig. 1). This region is covered by rays which are reflected from the cylinder. between A and B. Initial values on these rays are given by

(3.3), (3.8), and (3.9). The phase and amplitude on each ray can then be computed by solving the eikonal and transport equations. The region between the critically reflected rays and the shadow boundaries (which begin at C and D) is the region of total reflection (denoted by T.R.). It is covered by rays which are reflected between A and C and between B and D. Initial values are given by (3.3), (3.12), (3.8), and (3.9). The shadow region can be treated by creeping wave theory. (See Rulf,[7] and Lewis, Bleistein, and Ludwig[8] for details.)

The ray geometry inside the cylinder is shown in Fig. 2. A ray incident at E between A and B gives rise to a transmitted field corresponding to the refracted ray EF. The ray strikes the cylinder again at F, where it gives rise to a reflected ray FG and a transmitted ray. The amplitude of the transmitted ray at F is determined by means of the reflection coefficient R_{11}, and the amplitude of the transmitted ray is determined by means of the transmission coefficient T_{12}. A similar procedure yields appropriate reflected and transmitted rays and amplitudes at G, H, etc. It will be seen later that there are caustics along each ray FG, GH, etc., and hence a caustic correction must be included in the computation of amplitudes and phases.

The field, as described above, presents the following complications: The region of total reflection and the shadow region are covered by rays of type $T_{21}R_{11}^{l}T_{12}$, for each value of $l \geq 0$. Thus, in T.R., at each point there are infinitely many geometrical rays. In addition, for sufficiently large values of l, a chain of reflected rays which encircles the interior of the cylinder an arbitrary number of times can be found. However, as E approaches A, the direction of EF is nearly tangential, and so geometrical optics must fail. These difficulties, due to infinitely many geometrical rays and rays which are nearly tangential, are the principal novelties of the present problem.

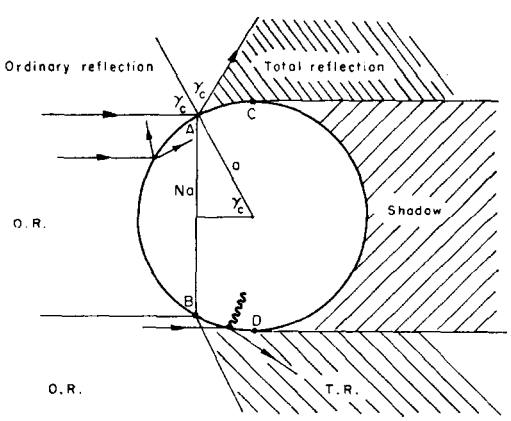


FIG. 1. The geometry of the rays outside the cylinder.

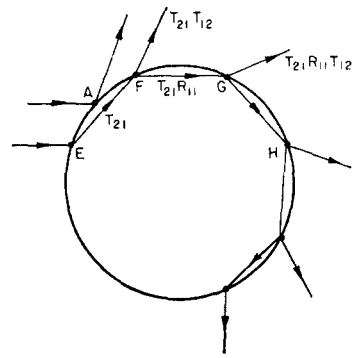


FIG. 2. The geometry of the rays inside the cylinder.

Asymptotic Expansions of the Airy Functions

The integrals (2.38), (2.39), and (2.43) are not suitable for application of the saddle-point method, since the integrands contain Airy functions. However, in certain regions the asymptotic expansions of the Airy functions are valid, and the saddle-point method can be applied. Here we proceed formally, in preparation for the more careful treatment of Sec. 4.

We shall assume that $\rho(r, \alpha) > 0$ and $\sigma(\alpha) = \rho(a, \alpha) > 0$. In that case, (2.33) is valid, and substitution of (2.33) into (2.39) yields

$$u_r(r, \theta, k) \sim \sqrt{k} \sum_{m=-\infty}^{\infty} \int_{\sigma>0} e^{ik\phi(r,\theta,\alpha,m)} \times Z(r, \alpha) \frac{1 - \xi(\alpha)}{1 + \xi(\alpha)} d\alpha, \quad (3.13)$$

where

$$\xi(\alpha) = \frac{(\sqrt{\sigma})\rho_r(a, \alpha)}{\phi_r(a, \theta, \alpha, m)} = \frac{(N^2a^2 - \alpha^2)^{\frac{1}{2}}}{(a^2 - \alpha^2)^{\frac{1}{2}}}. \quad (3.14)$$

Similarly, substitution of (2.33) into (2.43) yields

$$u_r'(r, \theta, k) \sim \sqrt{k} \sum_{m=-\infty}^{\infty} \sum_{l=0}^{\infty} \int_{\sigma>0} e^{ik\phi(r,\theta,\alpha,m,l)} Z(r, \alpha)(-i)^{l+1} \times T_{21}(\alpha)(R_{11}(\alpha))^l T_{12}(\alpha)\, d\alpha, \quad (3.15)$$

where T_{21}, R_{11}, and T_{12} are given by (3.8)–(3.11) with ξ given by (3.14), and

$$\phi'(r, \theta, \alpha, m, l) = \phi(r, \theta, \alpha, m) + \tfrac{4}{3}(l + 1)\sigma^{\frac{3}{2}}. \quad (3.16)$$

Substitution of (2.33) into (2.38) yields

$$u_l^{\pm}(r, \theta, k, m) \sim \int_{\substack{\sigma>0 \\ \rho>0}} e^{ik\phi^{\pm}(r,\theta,\alpha,m,l)} \times Z_{\pm}(r, \alpha)(-i)^l T_{21}(\alpha)(R_{11}(\alpha))^l\, d\alpha, \quad (3.17)$$

where

$$\phi^{\pm}(r, \theta, \alpha, m, l) = \phi(a, \theta, \alpha, m) \pm \tfrac{2}{3}\rho^{\frac{3}{2}} + \tfrac{2}{3}(2l + 1)\sigma^{\frac{3}{2}}, \quad (3.18)$$

$$Z_{-}(r, \alpha) = [(N^2a^2 - \alpha^2)/(N^2r^2 - \alpha^2)]^{\frac{1}{4}} Z(a, \alpha), \quad (3.19)$$

$$Z_{+}(r, \alpha) = -iZ_{-}(r, \alpha). \quad (3.20)$$

It is easy to check that the integrands of (3.13), (3.15), and (3.17) are given by geometrical optics, if the incident field is the integrand of (2.24). Note that $\phi_i(r, \theta, \alpha, m)$, $\phi(r, \theta, \alpha, m)$, and $\phi^-(r, \theta, \alpha, m, 0)$ are all equal where $r = a$, and the corresponding amplitudes are related by $R_{22}(\alpha)$ and $T_{21}(\alpha)$. In order to make the solution regular for $r < a$, the term u_0^+ must be added. A discussion of this point and its relationship to geometrical optics is given in Appendix A. Since u_0^+ corresponds to a wave incident from the inside, transmitted and reflected terms are required to satisfy the boundary conditions. These terms correspond to setting $l = 0$ in (3.15) and to u_1^- in (2.37) and (3.17). Similar interpretations hold for the higher values of l.

The Incident Field

We apply the saddle-point method to (2.24). The stationary point $\hat{\alpha}(r, \theta)$ is obtained by setting the derivative of ϕ_i equal to zero:

$$\frac{\partial}{\partial \alpha} \phi_i(r, \theta, \hat{\alpha}, m) = \cos^{-1} \frac{\hat{\alpha}}{r} + \tfrac{1}{2}\pi - \theta + 2\pi m = 0. \tag{3.21}$$

This implies that

$$\hat{\alpha}(r, \theta) = r \cos (\theta - \tfrac{1}{2}\pi - 2\pi m), \tag{3.22}$$

$$\phi_i(r, \theta, \hat{\alpha}, m) = -r \cos \theta. \tag{3.23}$$

Thus $\hat{\alpha}$ is constant along lines parallel to the x axis (the incident rays) and $\phi_i(r, \theta, \hat{\alpha}, m)$ is the incident phase. A short calculation shows that the amplitude given by the saddle-point method is also correct.

It is apparent from (3.21) that the different values of m contribute saddle points on different sheets of the covering of the (x, y) plane by the (r, θ) plane. Since we will confine our attention to a small portion of the (x, y) plane, only the value $m = 0$ will make a significant contribution to the field.

The Field Inside the Cylinder

We apply the saddle-point method to (3.17) to verify geometrical optics. We set $m = 0$ throughout, in order to concentrate on phenomena in the vicinity of the critically incident ray.

The stationary point for $u_0^-(r, \theta, k, 0)$ [denoted by $\alpha_0^-(r, \theta)$] is obtained by setting the derivative of ϕ^- equal to zero:

$$\frac{\partial}{\partial \alpha} \phi_0^-(r, \theta, \alpha_0^-, 0)$$

$$= \cos^{-1} \frac{\alpha_0^-}{a} + \tfrac{1}{2}\pi - \theta + \cos^{-1} \frac{\alpha_0^-}{Nr} - \cos^{-1} \frac{\alpha_0^-}{Na} = 0. \tag{3.24}$$

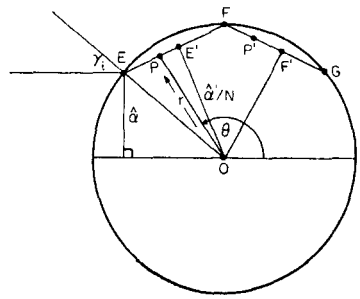


FIG. 3. Interpretation of α_0^-.

We set

$$\Phi_0^-(r, \theta) = \phi_0^-(r, \theta, \alpha_0^-, 0) = -[a^2 - (\alpha_0^-)^2]^{\frac{1}{2}} + N\{[a^2 - (\alpha_0^-)^2/N^2]^{\frac{1}{2}} - [r^2 - (\alpha_0^-)^2/N^2]^{\frac{1}{2}}\}. \tag{3.25}$$

Since $(\nabla \phi_0^-)^2 = N^2$ identically in α, it follows that

$$\nabla \phi_0^- \cdot \nabla \frac{\partial}{\partial \alpha} \phi_0^- = 0, \tag{3.26}$$

i.e., $(\partial/\partial\alpha)\phi_0^-$ is constant along rays for ϕ_0^- and, hence, $\alpha_0^-(r, \theta)$ is constant along the rays for Φ_0^-. Where $r = a$, (3.25) agrees with (3.23) and, hence, $\alpha_0^-(a, \theta) = \hat{\alpha}(a, \theta)$. Now (3.25) can be interpreted geometrically as in Fig. 3. On EE', we have $\alpha_0^- = \hat{\alpha} = a \sin \gamma_i$, and $OE' = \hat{\alpha}/N = a \sin \gamma_t$, according to Snell's law. Then

$$EP = EE' - PE' = [a^2 - (\hat{\alpha}^2/N^2)]^{\frac{1}{2}} - [r^2 - (\hat{\alpha}^2/N^2)]^{\frac{1}{2}}. \tag{3.27}$$

Comparison with (3.25) shows that $\Phi_0^-(r, \theta)$ is the sum of the incident phase at E and $N\,\overline{EP}$ (which represents the time required to travel from E to P in medium 1). The amplitude at P can be verified in a similar fashion (see Appendix A).

The saddle-point evaluation of u_0^- breaks down at E', since $\rho(r, \alpha_0^-) = 0$ there. This difficulty is due to the complicated behavior of the Bessel function where order and argument are nearly equal. This difficulty is treated in Appendix A. As is explained there, between E' and F, u_0^- must be replaced by u_0^+. The saddle-point evaluation of u_0^+ shows that the phase of u_0^+ at P is the sum of the phase at E' and $N\,E'P$, when P is between E' and F. Similar discussions can be given for $u_l^\pm$ when $l > 0$.

The Scattered Field

Here we apply the saddle-point method to (3.13) and (3.15) in order to verify geometrical optics in the exterior of the cylinder. As before, we set $m = 0$. The saddle point $\tilde{\alpha}(x)$ of (3.13) satisfies

$$\frac{\partial}{\partial \alpha} \phi(r, \theta, \tilde{\alpha}, 0) = 2 \cos^{-1} \frac{\tilde{\alpha}}{a} + \tfrac{1}{2}\pi - \theta - \cos^{-1} \frac{\tilde{\alpha}}{r} = 0. \tag{3.28}$$

We set

$$\Phi(r, \theta) = \phi(r, \theta, \tilde{\alpha}) = -(a^2 - \tilde{\alpha}^2)^{\frac{1}{2}} + (r^2 - \tilde{\alpha}^2)^{\frac{1}{2}} - (a^2 - \tilde{\alpha}^2)^{\frac{1}{2}}. \tag{3.29}$$

By the same argument as for α_0^-, we see that $\tilde{\alpha}(r, \theta)$ is constant along rays of $\Phi(r, \theta)$. Where $r = a$, $\tilde{\alpha} = \hat{\alpha} = a \sin \gamma_i$. The interpretation of Φ is given in Fig. 4. The segments OR and OQ are perpendicular to the extensions of the incident and reflected rays. Since the angle of incidence equals the angle of reflection, $OQ = OR = \tilde{\alpha}$. Hence Φ at P is equal to the sum of the incident phase at E and EP. From the saddle-point method, we obtain

$$u_r(r, \theta, k) \sim e^{ik\Phi(r,\theta)} R_{22}(\tilde{\alpha}) \times [2(r^2 - \tilde{\alpha}^2)^{\frac{1}{2}}/(a^2 - \tilde{\alpha}^2)^{\frac{1}{2}} - 1]^{-\frac{1}{2}}. \tag{3.30}$$

The saddle point of α_l of (3.21) is obtained from

$$\frac{\partial}{\partial \alpha} \phi'(r, \theta, \alpha_l, 0, l) = -\cos^{-1} \frac{\alpha_l}{r} + 2 \cos^{-1} \frac{\alpha_l}{a} + \tfrac{1}{2}\pi - \theta - 2(l + 1) \cos^{-1} \frac{\alpha_l}{Na} = 0. \tag{3.31}$$

We define

$$\Phi_l'(r, \theta) = \phi'(r, \theta, \alpha_l, 0, l) = -(a^2 - \alpha_l^2)^{\frac{1}{2}} + (r^2 - \alpha_l^2)^{\frac{1}{2}} - (a^2 - \alpha_l^2)^{\frac{1}{2}} + N2(l + 1)\left(a^2 - \frac{\alpha_l^2}{N^2}\right)^{\frac{1}{2}}. \tag{3.32}$$

If $l = 0$, as is shown in Fig. 5, then Φ_l' is the sum of the incident phase at E, plus $N\overline{EF}$, plus $\overline{FP}$. Similar interpretations hold for higher values of l. The saddle-point evaluation of (3.15) yields

$$u_r'(r, \theta, k) \sim \sum_{l=0}^{\infty} e^{ik\Phi_l'(r,\theta)}(-i)^l T_{21}(\alpha_l)(R_{11}(\alpha_l))^l T_{12}(\alpha_l) \times \left[1 - \frac{2(r^2 - \alpha_l^2)^{\frac{1}{2}}}{(a^2 - \alpha_l^2)^{\frac{1}{2}}} + \frac{2(l + 1)(r^2 - \alpha_l^2)^{\frac{1}{2}}}{(N^2a^2 - \alpha_l^2)^{\frac{1}{2}}}\right]^{-\frac{1}{2}}. \tag{3.33}$$

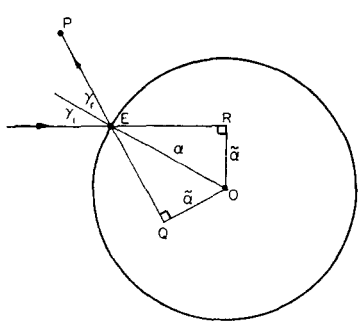


FIG. 4. Interpretation of $\tilde{\alpha}$ and Φ.

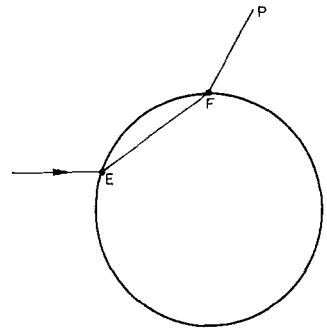


FIG. 5. Interpretation of Φ_0'.

The factor $(-i)^l$ can be interpreted as a phase shift, since each internally reflected ray passes through a caustic (see Appendix A). The factor in brackets can be interpreted as a ray density.

The totally reflected field corresponds to a real saddle point $\tilde{\alpha}$ of (2.18) with $Na < \tilde{\alpha} < a$. In this case, $\sigma < 0$ and (see Erdélyi[5])

$$V(k^{\frac{2}{3}}\sigma) \sim [(-\sigma)^{-\frac{1}{4}}/2(\sqrt{\pi})k^{\frac{1}{6}}]e^{-\frac{2}{3}k(-\sigma)^{\frac{3}{2}}}. \tag{3.34}$$

Then (2.18) becomes (formally)

$$u^{(1)}(r, \theta, k) \sim \sqrt{k} \sum_{m=-\infty}^{\infty} \int_{\sigma<0} e^{ik\phi(r,\theta,\alpha,m)} \times Z(r, \alpha) \frac{1 - i\xi_t(\alpha)}{1 + i\xi_t(\alpha)} d\alpha, \tag{3.35}$$

where

$$\xi_t(\alpha) = \frac{(\sqrt{-\sigma})\rho_r(a, \alpha)}{\phi_r(a, \alpha)} = \frac{(\alpha^2 - N^2a^2)^{\frac{1}{2}}}{(a^2 - \alpha^2)^{\frac{1}{2}}}. \tag{3.36}$$

Now (3.35) has the same form as (3.13). A similar procedure yields (3.30), with

$$R_{22}(\tilde{\alpha}) = [1 - i\xi_t(\tilde{\alpha})]/[1 + i\xi_t(\tilde{\alpha})]. \tag{3.37}$$

Residues near $\alpha = Na$

In addition to the saddle points which have been described above, there are poles in the integrals (2.18) and (2.27) just above the real axis. Here it will be shown that these poles correspond to whispering gallery modes (excited by the critically incident ray) which propagate just inside the cylinder and continually shed energy to the outside. The transmitted wave leaves the cylinder at the critical angle, and hence it is analogous to a lateral wave.

The Debye series was introduced in Sec. 2 in order to apply the saddle-point method, and the resulting approximations are predicted by geometrical optics. These approximations break down near the critically incident ray, which is refracted tangentially inside the cylinder. Therefore, we might expect some sort of

diffracted waves which are excited by the critically incident ray, i.e., which correspond to values of α near Na. Since the Debye series converges very poorly near $\alpha = Na$, we use the more basic formulas (2.18) and (2.27).

The poles of (2.18) and (2.27) correspond to values of ω for which

$$V(\omega_p) - \zeta V'(\omega_p) = 0, \quad p = 1, 2, \cdots. \tag{3.38}$$

If q_p is a root of V, then to first order

$$\omega_p \sim q_p + \zeta. \tag{3.39}$$

The corresponding values of α can be approximated from (2.22) and (2.16):

$$\alpha_p \sim Na - k^{-\frac{2}{3}}(\tfrac{1}{2}Na)^{\frac{1}{3}}(q_p + \zeta). \tag{3.40}$$

Since $\operatorname{Im} \zeta < 0$, we have $\operatorname{Im} \alpha_p > 0$; and since $\zeta = O(k^{-\frac{1}{3}})$, the α_p are close to the real axis.

After formally taking a sum of residues, (2.27) becomes

$$u(r, \theta, k) \sim \sum_{p=1}^{\infty} (2\pi i) k^{-\frac{1}{6}} \frac{2}{V'(\omega_p)} \left[\frac{e^{ik\psi} V(k^{\frac{2}{3}}\zeta) g}{\partial \rho / \partial \alpha}\right]_{\alpha=\alpha_p}. \tag{3.41}$$

Since $\alpha_p - Na = O(k^{-\frac{2}{3}})$, we have $\rho(r, \alpha_p) < 0$ unless $r - a = O(k^{-\frac{2}{3}})$. In view of (3.34), the residues are exponentially small inside the cylinder except in a strip of width $O(k^{-\frac{2}{3}})$ near the boundary. On the boundary, from (2.28) we have

$$\psi(a, \theta, \alpha_p) = -(a^2 - \alpha_p^2)^{\frac{1}{2}} + \alpha_p \left(\cos^{-1} \frac{\alpha_p}{a} + \frac{\pi}{2} - \theta + 2\pi m\right). \tag{3.42}$$

By neglecting the imaginary part of α_p, this expression can be interpreted as in Fig. 6. We choose $CA = \alpha_p$, and hence the ray at C is incident near the critical angle. Then $\sphericalangle COP$ is given by $\cos^{-1} \alpha_p/a + \pi/2 - \theta$. To the lowest order, $\alpha_p = Na$, and hence $\psi \sim \phi_i(C) + N\,CP$, where CP denotes the arc length

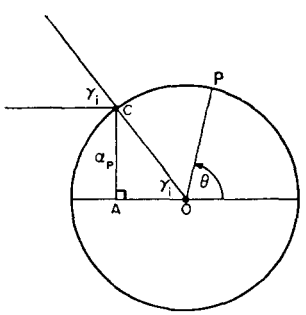


FIG. 6. Interpretation of the phase on the boundary.

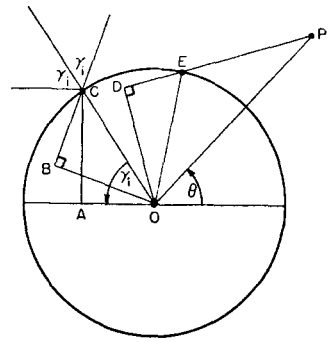


FIG. 7. Interpretation of the phase away from the boundary.

from C to P. The term of order $k^{-\frac{2}{3}}$ in α_p shifts the phase of each mode slightly, and the last term (involving ζ) has order k^{-1} with positive imaginary part. This term may be interpreted as an exponential decay in amplitude (independent of frequency) as CP increases, due to energy being shed to the exterior.

After formally taking a sum of residues, (2.18) becomes

$$u^{(1)}(r, \theta, k) \sim \sum_{p=1}^{\infty} (2\pi i) k^{-\frac{1}{6}} 2\zeta \left[\frac{e^{ik\phi} Z}{\partial \rho / \partial \alpha}\right]_{\alpha=\alpha_p}. \tag{3.43}$$

Again, neglecting the imaginary part of α_p, we may interpret $\phi(r, \theta, \alpha_p)$ as in Fig. 7. We have $CA = OB = OD = \alpha_p$, $DE = (a^2 - \alpha_p^2)^{\frac{1}{2}}$, and $DP = (r^2 - \alpha_p^2)^{\frac{1}{2}}$. Furthermore,

$$\cos^{-1}(\alpha_p/a) = \tfrac{1}{2}\pi - \sphericalangle COA = \sphericalangle EOD. \tag{3.44}$$

Hence we obtain

$$\phi(r, \theta, \alpha_l) = \phi_i(c) + \overline{EP} + \alpha_p\,(\sphericalangle\, EOC). \tag{3.45}$$

To lowest order, $\alpha_p = Na$ and $\phi(r, \theta, \alpha_p)$ consists of the sum of the incident phase at C, $N\,CE$, and EP. The higher-order terms in α_p can be interpreted just as in (3.42). We conclude that each term in (3.43) represents a wave which travels just inside the cylinder, as described above, and then is shed to the exterior at an angle near the critical angle. It is interesting to note that the residues which have been obtained represent a limiting case of waves which are incident near the critical angle, are transmitted nearly tangentially, internally reflected many times, and emerge at an angle near the critical angle.

4. ASYMPTOTIC EVALUATION OF THE SOLUTION

This section contains an asymptotic evaluation of the solution, based upon the concepts and approximations of the preceding sections. The original integral is decomposed and the contour of integration is deformed in different ways in the various regions

of space. In the region of ordinary reflection (see Fig. 1), for $\alpha < Na$, the integrand is split into terms corresponding to u_r and u_r' as defined by (2.39) and (2.41). Only u_r contributes a saddle point, and the rest of the scattered field is exponentially small. This agrees with the prediction of geometrical optics that only the incident and reflected fields are present in O.R.

In the region of total reflection, a more complicated decomposition is required. For a certain $\alpha^* < Na$, the integrand is split into terms corresponding to u_r and u_r' for $\alpha < \alpha^*$, and the form (2.43) is used for u_r'. Each term in the latter sum has a saddle point. For $\alpha > \alpha^*$, the original form (2.18) is used, but the contour is deformed into the upper half-plane, picking up some residues near $\alpha = Na$. There is an additional saddle point for $u^{(1)}$ between Na and a. The saddle points for u_r' correspond to the transmitted, multiply reflected and transmitted rays, the residues correspond to the whispering gallery modes, and the saddle point for $u^{(1)}$ yields the totally reflected field. In the shadow regions, the saddle point for $u^{(1)}$ is replaced by a sum of residues, which correspond to creeping waves. The other terms are as before.

Inside the cylinder, the ray geometry is complicated by the presence of caustics of the internally reflected field. However, near the boundary $r = a$, a close analogy holds with the treatment for the exterior. The main difference is the evanescent field (corresponding to a complex saddle point) which contributes to the field near the region of total reflection.

The Region of Ordinary Reflection

As is explained above, we use a decomposition analogous to (2.39) and (2.41) for $\alpha \leq \alpha^*$. The point α^* is chosen midway between $\tilde{\alpha}(r, \theta)$ and Na [see (3.28)]. Thus (2.18) becomes

$$
\begin{aligned}
&u^{(1)}(r, \theta, k) \\
&\sim \sqrt{k}\int_{-a}^{\alpha^*} e^{ik\phi(r,\theta,\alpha,0)} Z(r, \alpha)\frac{V_-(\omega) + \zeta V_-'(\omega)}{V_-(\omega) - \zeta V_-'(\omega)}\,d\alpha \\
&+ \sqrt{k}\int_{-\alpha}^{\alpha^*} e^{ik\phi(r,\theta,\alpha,0)}\frac{Z(r, \alpha)e^{\frac{1}{6}i\pi}\zeta\,d\alpha}{\pi(V(\omega)-\zeta V'(\omega))(V_-(\omega)-\zeta V_-'(\omega))} \\
&+ \sqrt{k}\int_{\alpha^*}^{a} e^{ik\phi(r,\theta,\alpha,0)} Z(r, \alpha)\frac{V(\omega) + \zeta V'(\omega)}{V(\omega) - \zeta V'(\omega)}\,d\alpha \\
&+ \sum_{m\neq 0}\sqrt{k}\int_{-a}^{a} e^{ik\phi(r,\theta,\alpha,m)} \\
&\times Z(r, \alpha)\frac{V(\omega) + \zeta V'(\omega)}{V(\omega) - \zeta V'(\omega)}\,d\alpha + v^{(1)}(r, \theta, k). \qquad (4.1)
\end{aligned}
$$

Now the contours can be deformed as shown in Fig. 8. Because α^* is not near Na, the asymptotic expansions of V and V_- are valid in the first two integrals of (4.1).

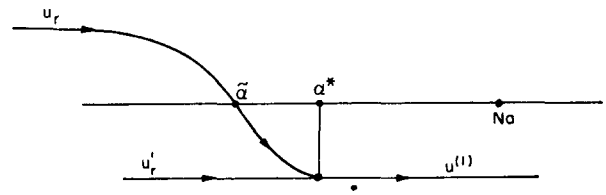


FIG. 8. The contour in the region of O.R.

Moreover, it is apparent from (3.28) and the geometrical interpretation of α that $\partial\phi(r, \theta, \alpha, m)/\partial\alpha$ is positive or negative if m is positive or negative. Hence, each term in the summation with $m \neq 0$ can be made exponentially small by deforming the contour into the upper or lower half-plane, respectively. For a similar reason, the integrals from α^* to a and $v^{(1)}$ are negligible. Thus, we obtain

$$
\begin{aligned}
u^{(1)}(r, \theta, k) &\sim \sqrt{k}\int_{a}^{\alpha^*} e^{ik\phi(r,\theta,\alpha,0)} Z(r, \alpha)\frac{1 - \xi(\alpha)}{1 + \xi(\alpha)}\,d\alpha \\
&+ \sqrt{k}\int_{a}^{\alpha^*} e^{ik\phi(r,\theta,\alpha,0)}\frac{e^{\frac{1}{3}k\sigma^{\frac{1}{3}}}4\zeta(\sqrt{\omega})Z(r, \alpha)\,d\alpha}{(1 + i\zeta\sqrt{\omega})^2}. \qquad (4.2)
\end{aligned}
$$

The integrand of the second term of (4.2) is the term in (3.15) with $m = l = 0$. It follows from the interpretation of (3.31) that $\partial\phi'(r, \theta, \alpha, 0, 0)/\partial\alpha < 0$ in O.R. and, hence, the second term in (4.2) is negligible after the contour is deformed into the lower half-plane. Hence, $u^{(1)}$ is approximated by the saddle-point contribution described in (3.28)–(3.30).

The Region of Total Reflection, and the Shadow

Here the integrand for $u^{(1)}$ is decomposed by use of a number $\alpha^* < Na$. Poles having real parts greater than α^* and saddle points less than α^* will appear in the evaluation of the solution. There is some arbitrariness in the representation of the solution; this corresponds to the fact that the residues and saddle-point contributions have very similar geometrical interpretations. We note that the saddle points cluster at Na, while the poles cluster at $-\infty$. Hence, there is a position for α^* which yields approximately equal numbers of saddle points and residue contributions. It would also be desirable for the last residue retained to have the same order of magnitude as the last saddle point retained. Both of these conditions are satisfied if $Na - \alpha^* = O(k^{-\frac{1}{2}})$. Then the number of saddle points and residues retained is $O(k^{\frac{1}{2}})$ and the last terms retained are $O(k^{-\frac{1}{2}})$. There is a remainder integral, which also is $O(k^{-\frac{1}{2}})$ except in transition regions. These transition regions are present since the positions of the various saddle points depend upon the observation point. In other words, each transmitted wave $T_{21}R_{11}^{l}T_{12}$ has its own transition region

from O.R. to T.R. If l and k are large and fixed, the corresponding saddle-point contributions may or may not be included, depending upon the position of the observation point.

We decompose $u^{(1)}$ as follows: Let $L = k^{\frac{1}{3}}\mu$, where μ and μ^{-1} are bounded. We define Φ^* by

$$\Phi^*(r, \theta, \alpha) = (r^2 - \alpha^2)^{\frac{1}{2}} - 2(a^2 - \alpha^2)^{\frac{1}{2}} + \alpha[-\cos^{-1}(\alpha/r) + 2\cos^{-1}(\alpha/a) + \tfrac{1}{2}\pi - \theta] + (2L + 3) \times [(N^2a^2 - \alpha^2)^{\frac{1}{2}} - \alpha\cos^{-1}(\alpha/Na)], \tag{4.3}$$

and we determine $\alpha^*(r, \theta)$ as the stationary value of Φ^*:

$$\left.\frac{\partial\Phi^*}{\partial\alpha}\right|_{\alpha=\alpha^*} = -\cos^{-1}\frac{\alpha^*}{r} + 2\cos^{-1}\frac{\alpha^*}{a} + \tfrac{1}{2}\pi - \theta - (2L + 3)\cos^{-1}\frac{\alpha^*}{Na} = 0. \tag{4.4}$$

Note that α^* is not a saddle point of ϕ_L or ϕ_{L+1} but lies between their saddle points. We write

$$u^{(1)}(r, \theta, k) = I_1 + I_2 + I_3 + I_4 + v^{(1)}(r, \theta, k) + \sum_{m\neq 0}\sqrt{k}\int_{-a}^{a} e^{il\phi}Z\frac{V + \zeta V'}{V - \zeta V'}\,d\alpha, \tag{4.5}$$

where

$$I_1 = \sqrt{k}\int_{-a}^{\alpha^*} e^{ik\phi_Z}\frac{V_-(\omega) + \zeta V'_-(\omega)}{V_-(\omega) - \zeta V'_-(\omega)}\,d\alpha, \tag{4.6}$$

$$I_2 = \sum_{l=0}^{L}\sqrt{k}\int_{-\infty}^{\alpha^*}\frac{e^{ik\phi}Ze^{-\frac{1}{6}i\pi}\zeta R^l}{\pi(V_-(\omega) - \zeta V'_-(\omega))^2}\,d\alpha, \tag{4.7}$$

$$I_3 = \sqrt{k}\int_{-\infty}^{\alpha^*}\frac{e^{ik\phi}Z\zeta e^{\frac{1}{6}i\pi}R^{L+1}\,d\alpha}{\pi(V(\omega) - \zeta V'(\omega))(V_-(\omega) - \zeta V'_-(\omega))}, \tag{4.8}$$

$$I_4 = \sqrt{k}\int_{\alpha^*}^{a} e^{ik\phi}Z\frac{V + \zeta V'}{V - \zeta V'}\,d\alpha. \tag{4.9}$$

The last two terms in (4.5) can be estimated as in O.R.

We deform the contours as follows (see Fig. 9): The end points are moved from α^* to β^*, where

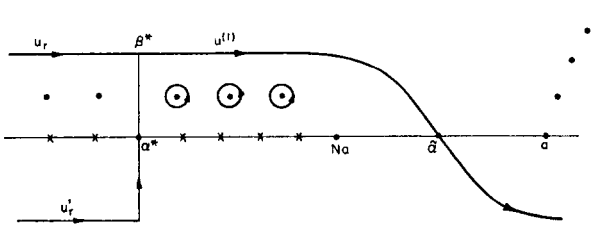


FIG. 9. The contour in the region of T.R.

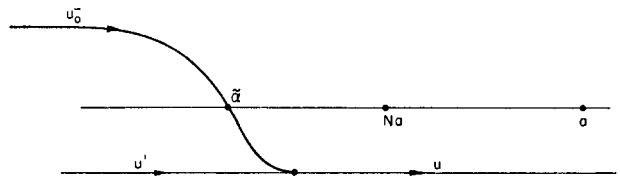


FIG. 10. The contour adjacent to O.R.

$\operatorname{Re}\alpha^* = \operatorname{Re}\beta^*$, $\operatorname{Im}\beta^* = O(k^{-\frac{5}{8}})$. The contour for I_1 is moved into the upper half-plane. No residues are contributed, since I_1 does not have poles near the real axis. The contours for I_2 and I_3 are deformed into the lower half-plane passing through α^*. The contour for I_4 is deformed into the upper half-plane near β^*, but it crosses the real axis at the saddle point $\tilde{\alpha}$ between Na and a. Residues are contributed from the poles just above the real axis and to the right of α^*. On the deformed contour, I_1 is exponentially small, since $\partial\phi(r, \omega, \alpha, 0)/\partial\alpha > 0$ if $\alpha < Na$ and the observation point is in T.R. Each term in I_2 can be evaluated by the saddle-point method, since all of the saddle points for $l \leq L$ are to the left of α^*. These contributions are described in (3.31)–(3.33). Similarly, I_4 is asymptotically given by a sum of residues [see (3.43)–(3.45)] plus a saddle-point contribution [see (3.35)–(3.37)]. The integral I_3 is estimated in Appendix B. Provided that the segment (α^*, β^*) does not pass near a pole, it is shown that I_3 has order $k^{-\frac{1}{2}}$.

We omit the details of the treatment of the shadow region, since it is quite analogous to the above. The residues from poles near $\alpha = a$ correspond to creeping waves (see Rulf,[7] and Lewis, Bleistein, and Ludwig[8]).

The Interior of the Cylinder

The field inside the cylinder is extremely complicated, due to the presence of ray families for each value of l (the number of internal reflections). For $l > 0$, each such family has a caustic curve. Hence, the number of saddle-point contributions to be expected varies considerably from point to point inside the cylinder, and the integration contours will change accordingly. Here we confine our attention to the vicinity of the boundary $r \sim a$, and near the critically incident ray $\theta \sim \theta_c$ (defined below). There the deformation of the contour is completely analogous to the previous ones.

We define

$$\theta_c = \tfrac{1}{2}\pi + \cos^{-1}(Na/a). \tag{4.10}$$

If $\theta > \theta_c$ and $r \sim a$, the observation point is adjacent to the region of ordinary reflection. The solution (2.37) may be rewritten as $u = u_0^- + u'$. The contour of integration is shown in Fig. 10, where $\tilde{\alpha}$ denotes

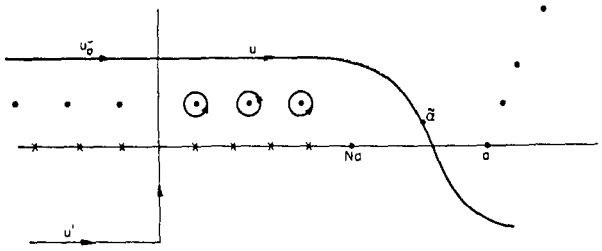


FIG. 11. The contour adjacent to T.R.

the saddle point of u_0^- (the transmitted ray). Further away from the boundary, saddle points for $l > 0$ may be encountered in pairs. The corresponding contours have a loop above the real axis, to the left of $\alpha = Na$.

If $\theta < \theta_c$ and $r \sim a$, the observation point is adjacent to T.R. The appropriate contour is shown in Fig. 11. The integrand for u has a saddle point in the upper half-plane, above the interval (Na, a). This contribution yields the evanescent field, due to total reflection of the incident ray. There are also residue contributions, which correspond to whispering gallery modes. All of these contributions decay exponentially as r decreases. There are also saddle-point contributions from the various terms which comprise u', which correspond to internally reflected rays.

5. COMPARISON WITH PREVIOUS RESULTS

The most relevant previous work is due to Chen[1,2] and Nussenzveig.[3] Chen extends Keller's geometrical theory to cover transparent bodies, and the present results are in agreement with his. Nussenzveig (correctly) criticizes Chen's choice of contour, but these difficulties are removed by our more elaborate treatment. Chen's results, together with Sec. 4 and Appendix C of the present work, give a fairly complete description of diffraction by a general transparent, convex object. A more rigorous and complete treatment could probably be given by imitating the method of Ludwig.[9]

The bulk of Nussenzveig's work is concerned with the case $N > 1$, which is quite different from the case $N < 1$. For the latter case, Nussenzveig obtains residue contributions from poles in the lower half-plane (below $\alpha = Na$). He interprets these residues as modes which propagate from the region of total reflection into the region of ordinary reflection. This interpretation is not correct, since the minus sign in front of ζ_{10}^+ in his equation (4.82) is not taken into account. Therefore, the direction of propagation of his residues should be from the region of ordinary reflection to the region of total reflection, as with the residues described in Appendix C. However, since the poles are in the lower half-plane, the amplitude of each term increases with increasing phase. Therefore, these modes transport energy from a source in the interior of the cylinder to the exterior; indeed, the amplitude of each mode increases exponentially as r decreases, for $r < a$. We conclude that these modes are unphysical and are not excited by a wave incident upon the cylinder. The treatment of Streifer and Kodis[10] utilizes similar modes and suffers from the same defect.

Buldyrev and Lanin treat the Green's function for diffraction by a transparent cylinder in a series of papers.[11–13] Their results agree with the present ones, except in the region of total reflection (see Ref. 13). In this region, Buldyrev and Lanin[13] represent the solution as a sum of M saddle-point contributions, plus a "head wave" S_M^+, where M has order $p^{\frac{1}{3}}$ and $p = kNa$ in our notation. It is shown below that this value for M is too large, since the saddle-point method is not valid for so many terms. Such a large value for M is not plausible, since it involves use of the asymptotic expansion of the Airy function where its argument is bounded.

If one attempts to overcome the above difficulty by a choice of $M \ll p^{\frac{1}{3}}$, then the saddle point of ϕ_{0M} moves away from $\zeta = 1$ (see Fig. 6 on p. 141 of Ref. 13), and the end points $T_\pm$ [see (2.5)] must be $\gg 1$ in order that $\exp[ip\phi_{0M}(\zeta)]$ be negligible outside of the interval (ζ_-, ζ_+) (see p. 142 of Ref. 13). Therefore, the integral for $G_M(\gamma, p)$ cannot be dealt with as simply as in Ref. 13. However, if their contour C_3 is deformed into a descent contour, then the portion near $\zeta = 1$ is pushed into the upper half-plane. The result is described in Sec. 4 above.

Buldyrev and Lanin justify the choice $M = O(p^{\frac{1}{3}})$ on the basis that "two waves exist separately if the optical path difference is greater than the wavelength" (see p. 140 of Ref. 13). The validity of the method can be checked mathematically as follows: An integral of the form

$$I(p) = \int f(\zeta) e^{ip\phi(\zeta)}\, d\zeta \tag{5.1}$$

yields a saddle-point contribution at ζ_M if a number δ can be found such that, for $|\zeta - \zeta_M| \leq \delta$,

$$p\phi''(\zeta_M)\frac{\delta^2}{2} \gg p\phi'''(\zeta)\frac{\delta^3}{3!}, \tag{5.2}$$

$$p\phi''(\zeta_M)\frac{\delta^2}{2} \gg 1, \tag{5.3}$$

$$f(\zeta_M) \gg f'(\zeta)\delta. \tag{5.4}$$

Conditions (5.2) and (5.4) justify expansion of ϕ and f about $\zeta = \zeta_M$, while condition (5.3) is necessary in order to evaluate the simplified integral. If $M = O(p^{\frac{1}{3}})$ as in Ref. 13, then $\phi''(\zeta_M) = O(p^{\frac{2}{3}})$ and $\phi''' = O(p^{\frac{1}{3}})$. Then (5.2) and (5.4) require that

$$p^{-\frac{2}{3}} \gg \delta \gg p^{-\frac{1}{6}}. \tag{5.5}$$

The function f is defined by (1.4) of Ref. 13; it involves the quotient of Airy functions of argument $[p\phi(1, \zeta)]^{\frac{2}{3}}$, raised to the Mth power. Hence, (5.5) implies that

$$f'(\zeta_M)\delta \gg p^{\frac{1}{6}} f(\zeta_M), \tag{5.6}$$

which contradicts (5.4).

ACKNOWLEDGMENTS

The author is indebted to J. B. Keller, H. M. Nussenzveig, and the late R. M. Lewis for a number of helpful discussions of this problem.

This research was supported by the Office of Naval Research under Contract NONR 285(48).

APPENDIX A: THE INTERNALLY REFLECTED FIELD

The discussion of the field inside the cylinder which is given in Sec. 3 suffers from two defects: the method is only valid where $\rho(r, \alpha) > 0$, and the substitution of $u_l^+(r, \theta, k)$ for $u_l^-(r, \theta, k)$ beyond the midpoint of each internally reflected ray requires justification. Here we shall remove these defects and also discuss the caustics inside the cylinder by applying the method of stationary phase for double integrals.

Assuming as before that $\sigma(\alpha) > 0$, we substitute (2.33) into (2.36) to obtain

$$\begin{aligned} u(r, \theta, k) \sim k^{\frac{2}{3}} \sum_{m=-\infty}^{\infty} \sum_{l=0}^{\infty} & \\ \times \int_{\sigma>0} & \exp\left[ik(\psi(r, \theta, \alpha, m) + \tfrac{2}{3}(2l+1)\sigma^{\frac{3}{2}})\right] \\ \times\ & V(k^{\frac{2}{3}}\rho(r, \alpha))(2)^{\frac{1}{2}}(-i)^l T_{21}(\alpha)[R_{11}(\alpha)]^l \\ \times & \left[\frac{\rho(r, \alpha)(N^2a^2 - \alpha^2)}{(a^2 - \alpha^2)(N^2r^2 - \alpha^2)}\right]^{\frac{1}{4}} d\alpha. \end{aligned} \tag{A1}$$

Now we use the integral representation for V in order to obtain a double integral: From Erdélyi[5] we have

$$V(k^{\frac{2}{3}}\rho) = \frac{k^{\frac{1}{3}}}{2\pi} \int_{-\infty}^{\infty} e^{ik[\rho\beta - \frac{1}{3}\beta^3]}\, d\beta, \tag{A2}$$

and, hence,

$$\begin{aligned} u(r, \theta, k) \sim \frac{k}{2\pi} \sum_{m=-\infty}^{\infty} \sum_{l=0}^{\infty} & \iint e^{ik\Psi}(2)^{\frac{1}{2}}(-i)^l T_{21}(\alpha)[R_{11}(\alpha)^l] \\ & \times \left[\frac{\rho(r, \alpha)(N^2a^2 - \alpha^2)}{(a^2 - \alpha^2)(N^2r^2 - \alpha^2)}\right]^{\frac{1}{4}} d\alpha\, d\beta, \end{aligned} \tag{A3}$$

where

$$\Psi = \psi(r, \theta, \alpha, m) + \tfrac{2}{3}(2l+1)\sigma^{\frac{3}{2}} + \rho(r, \alpha)\beta - \tfrac{1}{3}\beta^3. \tag{A4}$$

The method of stationary phase is applicable to (A3), provided that the Hessian of Ψ is different from zero (see below) without any further restrictions on ρ. The conditions for Ψ to be stationary are

$$\psi_\alpha + (2l+1)(\sqrt{\sigma})\sigma_\alpha + \beta\rho_\alpha = 0, \quad \rho - \beta^2 = 0. \tag{A5}$$

We note that (A5) implies that $\phi^\pm$ [given by (3.18)] is stationary. The cases $\beta \gtrless 0$ correspond to stationary points for $u_l^\pm$.

We define Δ (the Hessian of Ψ) by means of

$$\begin{aligned} \Delta = \begin{vmatrix} \Psi_{\alpha\alpha} & \Psi_{\alpha\beta} \\ \Psi_{\alpha\beta} & \Psi_{\beta\beta} \end{vmatrix} &= -2 \left| \frac{\beta}{(N^2r^2 - \alpha^2)^{\frac{1}{2}}} \right| \\ &- 2\beta\left[\frac{1}{(N^2a^2 - \alpha^2)^{\frac{1}{2}}} - \frac{1}{(a^2 - \alpha^2)^{\frac{1}{2}}} + \frac{2l}{(N^2a^2 - \alpha^2)^{\frac{1}{2}}}\right]. \end{aligned} \tag{A6}$$

The caustics of the field correspond to zeros of Δ, and the ordinary method of stationary phase is not applicable at caustics. Since $N < 1$, the term in brackets in (A6) is positive for all l, and hence $\Delta < 0$ if $\beta > 0$. If $l = 0$, then $\Delta < 0$ even if $\beta < 0$; but if $l > 0$, then Δ will vanish if $\beta < 0$, for some value of r between α/N and α. Geometrically, these facts imply that in Fig. 3 there is no caustic between E and F, but there is a casutic between F and F', etc. The solution has no anomalies at E', F', etc.; the difficulties there are due to the deficiencies of the representation (2.37). The method of stationary phase in two dimensions can be applied to (A3) to verify the results of geometrical optics (with caustic corrections). We omit the details.

APPENDIX B: THE ERROR INTEGRAL IN T.R.

Here we estimate the error integral I_3, given by (4.8). It follows from the definition of L and from (4.4) that

$$Na - \alpha^* = O(k^{-\frac{1}{2}}). \tag{B1}$$

Hence, the Airy functions in the denominator of (4.8) may be replaced by their asymptotic expansions [see (2.32) and (2.33)]:

$$I_3 \sim \sqrt{k} \int_{-\infty}^{\beta^*} \frac{e^{ik\Phi^*} \xi T_{21} R_{11}^{L+1}\, d\alpha}{(2\pi)^{\frac{1}{2}}(a^2 - \alpha^2)^{\frac{1}{4}}(\cos W - i \sin W)}. \tag{B2}$$

Here Φ^* is given by (4.3) and

$$W = \tfrac{2}{3}\omega^{\frac{3}{2}} - \tfrac{1}{4}\pi = k(N^2a^2 - \alpha^2)^{\frac{1}{2}} - k\alpha \cos^{-1}(\alpha/Na) - \tfrac{1}{4}\pi. \quad \text{(B3)}$$

From (B1), it follows that

$$\xi = O(k^{-\frac{1}{4}}) \quad \text{(B4)}$$

and

$$\frac{\partial W}{\partial \alpha} = O(k^{\frac{1}{4}}). \quad \text{(B5)}$$

After introducing W as a new variable, we obtain

$$I_3 \sim \sqrt{\pi}\int \frac{e^{ik\Phi^*}\xi T_{21}R_{11}^{L+1}\,dW}{(2\pi)^{\frac{1}{2}}(a^2-\alpha^2)^{\frac{1}{2}}(\cos W - i\xi \sin W)W_\alpha}. \quad \text{(B6)}$$

When (B4) and (B5) are applied to (B6), the result is that the portion of the integral where $\alpha - \alpha^* = O(k^{-\frac{5}{8}})$ has order $k^{-\frac{1}{2}}$ if $\cos W - i\xi \sin W$ is bounded away from zero. The remainder of the integral (in the lower half-plane) is exponentially small.

APPENDIX C: WHISPERING GALLERY MODES

The following discussion is motivated by the residue form (3.41). However, it is independent of the rest of the paper, and the results are valid for a general convex boundary which is concave towards the faster medium. The basic idea is to find solutions of the differential equations and jump conditions which have a caustic inside but close to the boundary. The resulting equations are very similar to those of Ludwig[14] and Lewis, Bleistein, and Ludwig.[8]

We denote the solution in the exterior of S by $U(x, k)$, and the solution in the interior of S by $u(x, k)$. S is assumed to be smooth and convex. The differential equations are

$$\Delta U + k^2U = 0, \quad \Delta u + k^2N^2u = 0, \quad N < 1, \quad \text{(C1)}$$

and the boundary conditions are

$$u = U, \qquad \frac{\partial u}{\partial n} = \frac{\partial U}{\partial n}, \quad \text{for } x \text{ on } S. \quad \text{(C2)}$$

Motivated by Refs. 8 and 14, we make the ansatz

$$u(x, k) = \exp\{ik[\theta_0(x) + k^{-\frac{2}{3}}\theta_1(x)]\} \times [V(k^{\frac{2}{3}}\rho_0(x) + \rho_1(x))g_0(x) + ik^{-\frac{1}{3}}V'(k^{\frac{2}{3}}\rho_0(x) + \rho_1(x))h_0 + \cdots]. \quad \text{(C3)}$$

Here V is given by (2.14): $V''(\omega) + \omega V(\omega) = 0$. The corresponding ansatz for U is

$$U(x, k) = \exp[ik(\phi_0(x) + k^{-\frac{2}{3}}\phi_1(x))] \times [Z_1(x)/k^{\frac{1}{3}} + \cdots]. \quad \text{(C4)}$$

We require that $\rho_0(x) = 0$ on S; this implies that the caustic (where $\rho_0 + k^{-\frac{2}{3}}\rho_1 = 0$) is close to the boundary.

The boundary conditions (C2) cannot be satisfied unless the phases match; thus we must have

$$\theta_0(x) = \phi_0(x), \quad \theta_1(x) = \phi_1(x), \quad \text{for } x \text{ on } S. \quad \text{(C5)}$$

The usual procedure of geometrical optics (see Keller and Lewis[6]) implies that

$$[\nabla(\phi_0 + k^{-\frac{2}{3}}\phi_1)]^2 = 1 \quad \text{(C6)}$$

to first two orders, and hence

$$(\nabla\phi_0)^2 = 1, \quad \nabla\phi_0 \cdot \nabla\phi_1 = 0. \quad \text{(C7)}$$

The second equation of (C7) implies that ϕ_1 is constant along rays. Similarly, equations (1.21)–(1.22) of Ref. 14 imply that

$$(\nabla\theta_0)^2 + \rho_0(\nabla\rho_0)^2 = N^2, \quad \text{(C8)}$$

$$\nabla\theta_0 \cdot \nabla\rho_0 = 0, \quad \text{(C9)}$$

$$2\nabla\theta_0 \cdot \nabla\theta_1 + 2\rho_0\nabla\rho_0 \cdot \nabla\rho_1 + \rho_1(\nabla\rho_0)^2 = 0, \quad \text{(C10)}$$

$$\nabla\theta_1 \cdot \nabla\rho_0 + \nabla\theta_0 \cdot \nabla\rho_1 = 0. \quad \text{(C11)}$$

Note that since $\rho_0 = 0$ on S, (C8) and (C9) imply that $\partial\theta_0/\partial n = 0$ and $(\nabla\theta_0)^2 = N^2$ on S. Then (C5) and (C7) imply that

$$\frac{\partial\phi_0}{\partial n} = (1 - N^2)^{\frac{1}{2}}, \quad \text{on } S, \quad \text{(C12)}$$

and hence the exterior rays leave the boundary at the critical angle.

In order to apply the boundary conditions, we compute

$$\frac{\partial U}{\partial n} = \left[ik\left(\frac{\partial\phi_0}{\partial n}\right)\frac{Z_1}{k^{\frac{1}{3}}} + \cdots\right]\exp[ik(\phi_0 + k^{-\frac{2}{3}}\phi_1)], \quad \text{(C13)}$$

$$\frac{\partial u}{\partial n} = \left[k^{\frac{2}{3}}\left(\frac{\partial\rho_0}{\partial n}\right)V'(\rho_1)g_0 + \cdots\right]\exp[ik(\theta_0 + k^{-\frac{2}{3}}\theta_1)]. \quad \text{(C14)}$$

Hence, (C2) implies

$$\left(\frac{\partial\rho_0}{\partial n}\right)V'(\rho_1)g_0 = i\left(\frac{\partial\phi_0}{\partial n}\right)Z_1, \quad \text{(C15)}$$

$$V(\rho_1) = 0, \quad \text{(C16)}$$

$$V'(\rho_1)h_0/i = Z_1. \quad \text{(C17)}$$

The present equation (C16) differs slightly from the residue equation (3.38), since the roots of (3.38) are not real, but near the real axis. The present procedure is equivalent to the earlier one, but more convenient for calculations.

Now (C15) and (C17) imply that

$$h_0 = \frac{\partial\rho_0/\partial n}{\partial\phi_0/\partial n}\, g_0 . \tag{C18}$$

From Eq. (1.23) of Ref. 14 (at $\rho_0 = 0$), we have

$$2\nabla\theta_0 \cdot \nabla g_0 + \Delta\theta_0 g_0 + (\nabla\rho_0)^2 h_0 = 0, \tag{C19}$$

and thus (C18) implies

$$2\nabla\theta_0 \cdot \nabla g_0 + \Delta\theta_0 g_0 + (\nabla\rho_0)^2 \frac{\partial\rho_0/\partial n}{\partial\phi_0/\partial n}\, g_0 = 0. \tag{C20}$$

If we define p by means of

$$2N\frac{dp}{ds} = (\nabla\rho_0)^2 \frac{\partial\rho_0/\partial n}{\partial\phi_0/\partial n} = \frac{2}{a(1-N^2)^{\frac{1}{2}}}, \tag{C21}$$

where s is the arclength along a surface ray and a is the radius of curvature of the surface ray, and if we set

$$G = e^{p} g_0, \tag{C22}$$

then

$$2N\frac{dG}{ds} + \Delta\theta_0 G = 0. \tag{C23}$$

Equation (C23) together with (C16) would result from imposing the condition $u = 0$ at the boundary. The interpretation and analysis of these equations is essentially contained in Lewis, Bleistein, and Ludwig.[8] The exponential decay factor e^{-p} in g_0 is due to the energy continually being shed to the exterior. Note that $u = O(k^{-\frac{1}{3}})$ at the boundary; hence the loss is relatively small. It is interesting that, in two dimensions, from (C21),

$$\frac{dp}{d\omega} = \frac{1}{N(1-N^2)^{\frac{1}{2}}}, \tag{C24}$$

where ω is the angle formed between the tangent to the boundary and the x axis.

The wave in the exterior is analogous to a lateral wave, since the rays for ϕ_0 meet the boundary at the critical angle. The wave in the interior is analogous to a whispering gallery mode for the interior problems with boundary condition $u = 0$. These modes are excited in a boundary layer near a critically incident ray. The diffraction coefficients are determined from the residue calculation (see Chen[2]).

[1] Y. M. Chen, J. Math. Phys. **5**, 820 (1964).
[2] Y. M. Chen, J. Math. Phys. **6**, 1332 (1965).
[3] H. M. Nussenzveig, J. Math. Phys. **10**, 82 (1969).
[4] A. Sommerfeld, *Partial Differential Equations* (Academic, New York, 1949).
[5] A. Erdélyi, *Asymptotic Expansions* (Dover, New York, 1956).
[6] J. B. Keller and R. M. Lewis, *Asymptotic Theory of Wave Propagation and Diffraction* (Wiley, New York, to appear).
[7] B. Rulf, Commun. Pure Appl. Math. **21**, 67 (1968).
[8] R. M. Lewis, N. Bleistein, and D. Ludwig, Commun. Pure Appl. Math. **20**, 295 (1967).
[9] D. Ludwig, Commun. Pure Appl. Math. **20**, 103 (1967).
[10] W. Streifer and R. Kodis, Quart. Appl. Math. **22**, 193 (1964).
[11] V. S. Buldyrev, Z. Vycisl. Mat. Mat. Fiz. **4**, No. 4, Suppl., 275 (1964).
[12] V. S. Buldyrev and A. I. Lanin, Trud. Mat. Inst. Steklov **92**, 147 (1966).
[13] V. S. Buldyrev and A. I. Lanin, Z. Vycisl. Mat. Mat. Fiz. **6**, 90 (1966) [USSR Comp. Math. Math. Phys. **6**, 128 (1966)].
[14] D. Ludwig, Commun. Pure Appl. Math. **19**, 215 (1966).

Noninvariance Groups in the Second-Quantization Picture and Their Applications*

M. MOSHINSKY AND C. QUESNE†
Instituto de Física, Universidad de México, México 20, D.F., Mexico

(Received 10 November 1969)

We investigate the existence of noninvariance groups in the second-quantization picture for fermions distributed in a finite number of states. The case of identical fermions in a single shell of angular momentum j is treated in detail. We show that the largest noninvariance group is a unitary group $U(2^{2j+1})$. The explicit form of its generators is given both in the m scheme and in the seniority–angular-momentum basis. The full set of 0-, 1-, 2-, $\cdots$, $(2j+1)$-particle states in the j shell is shown to generate a basis for the single irreducible representation [1] of $U(2^{2j+1})$. The notion of complementary subgroups within a given irreducible representation of a larger group is defined, and the complementary groups of all the groups commonly used in classifying the states in the j shell are derived within the irreducible representation [1] of $U(2^{2j+1})$. These concepts are applied to the treatment of many-body forces, the state-labeling problem, and the quasiparticle picture. Finally, the generalization to more complex configurations is briefly discussed.

1. INTRODUCTION

The concept of the noninvariance group[1] is a very attractive one since, for a given dynamical problem, all the states belong to a single irreducible representation (IR) of the group. This leads to the establishment of many relations for the matrix elements of operators between these states as well as to more systematic classification schemes for the states themselves.

Unfortunately, these noninvariance groups have been discussed only for very simple[1] 1-particle problems, and even in those cases they lead to noncompact groups. For many-body problems, even with simple types of interactions, their determination seems at present to be out of the question.

On the other hand, if we are working in the second-quantized picture (SQP) with a finite number of states $\mathcal{N}$, the most general interactions can be formulated in terms of polynomial functions of the generators[2] of a unitary group $U(\mathcal{N})$. This suggests the possibility that we may find, in the SQP, noninvariance groups of a simple type, which may also be independent of the characteristics of the interactions we would like to use.

Actually, noninvariance groups of this type have been known for some time. To indicate their characteristics explicitly, we restrict ourselves to the SQP for identical fermions in a single shell of angular momentum j. The anticommuting creation and annihilation operators are then denoted respectively by

$$b_m^+, \quad b^{m'}, \quad m, m' = j, \cdots, -j, \tag{1.1}$$

with

$$\{b_m^+, b^{m'}\}_+ = \delta_m^{m'}, \quad \{b_m^+, b_{m'}^+\}_+ = \{b^m, b^{m'}\}_+ = 0. \tag{1.2}$$

In the work of Helmers[3] and others, and more explicitly in that of Judd,[4] the orthogonal group of $4j+3$ dimensions $O^+(4j+3)$ is introduced, the generators of whose Lie algebra are

$$b_{m_1}^+ b_{m_2}^+, \quad b^{m_1} b^{m_2}, \quad b_{m_1}^+ b^{m_2} - b^{m_2} b_{m_1}^+, \quad b_{m_1}^+, \quad b^{m_2}. \tag{1.3}$$

It is then shown[4] that all states

$$|m_1 \cdots m_n\rangle \equiv b_{m_1}^+ b_{m_2}^+ \cdots b_{m_n}^+ |0\rangle, \quad m_1 > m_2 \cdots > m_n, \quad n = 0, 1, 2, \cdots, 2j+1, \tag{1.4}$$

where $|0\rangle$ (corresponding to $n = 0$) is the vacuum state, belong to the single IR

$$[\tfrac{1}{2}\ \tfrac{1}{2} \cdots \tfrac{1}{2}] \equiv [1^{2j+1}] \tag{1.5}$$

of $O^+(4j+3)$. Thus the group $O^+(4j+3)$ is a noninvariance group in the SQP for a single shell of angular momentum j.

In this paper, we introduce a noninvariance group, of which $O^+(4j+3)$ is actually a subgroup, which proves to be much more useful in the discussion of the matrix elements of our dynamical problem. To derive this group, we note first that the number of states (1.4), because of the anticommuting nature of the b^+'s, is given by

$$\sum_n \binom{2j+1}{n} = 2^{2j+1}. \tag{1.6}$$

Thus, the number of operators of the type

$$b_{m_1}^+ \cdots b_{m_n}^+ b^{m'_{n'}} \cdots b^{m_{1'}}, \quad m_1 > \cdots > m_n, \quad m'_1 > \cdots > m'_{n'}, \quad n, n' = 0, 1, \cdots, 2j+1, \tag{1.7}$$

is $(2^{2j+1})^2$, if we include among them the operator 1. The commutator of two operators of the type (1.7)

clearly must give us back again a linear combination of operators of the type (1.7), and so these operators form a Lie algebra. In the next section, we show that this Lie algebra corresponds to a unitary group of 2^{2j+1} dimensions $U(2^{2j+1})$. We clearly see from (1.7) and (1.3) that $U(2^{2j+1})$ contains $O^+(4j+3)$ as a subgroup, so then $U(2^{2j+1})$ is also a noninvariance group in the SQP. In fact, we prove that all states (1.4) belong to the single IR [1] of $U(2^{2j+1})$.

We now proceed to determine linear combinations of the operators (1.7) that satisfy the usual commutation relations for the generators of the group $U(2^{2j+1})$, thus proving in a constructive fashion that we are dealing with a group of this type.

2. THE GENERATORS OF THE GROUP $U(2^{2j+1})$

In our analysis we are only interested in the effect of the operators (1.7) on the complete set of states (1.4), i.e., in matrix elements of the type

$$\langle \bar{m}_1 \cdots \bar{m}_{\bar{n}} | \, b^+_{m_1} \cdots b^+_{m_n} b^{m'_{n'}} \cdots b^{m_1} | \bar{m}'_1 \cdots \bar{m}'_{\bar{n}'} \rangle. \tag{2.1}$$

We can then prove that the operators (1.7) define the Lie algebra of $U(2^{2j+1})$ if we can find $(2^{2j+1})^2$ linearly independent combinations of the operators (1.7) such that their matrices, with respect to the full set of states (1.4), have zeros everywhere, except at a single position where they take the value 1.

The matrix whose elements are (2.1) clearly has zero values for all $\bar{n} < n$ and $\bar{n}' < n'$. For $\bar{n} = n$ and $\bar{n}' = n'$, the matrix element (2.1) is

$$\delta^{\bar{m}_1}_{m_1} \cdots \delta^{\bar{m}_n}_{m_n} \delta^{m_1}_{\bar{m}'_1} \cdots \delta^{m'_{n'}}_{\bar{m}'_{n'}} \tag{2.2}$$

and, thus, the submatrix with n particles in the bra and n' in the ket has a 1 at some position and zero everywhere else. From (2.1), we see that all other matrix elements will be zero, unless

$$\bar{n} - n = \bar{n}' - n', \tag{2.3}$$

and, thus, we have only zeros when $\bar{n} = n$, $\bar{n}' > n'$ and $\bar{n} > n$, $\bar{n}' = n'$. By enumerating the states in an order that follows the increasing number of particles, the full matrix then looks as follows:

$$\begin{array}{c} \\ \bar{n} < n \\ \bar{n} = n \\ \bar{n} > n \end{array} \begin{array}{c} \begin{array}{ccc} \bar{n}' < n' & \bar{n}' = n' & \bar{n}' > n' \end{array} \\ \left[\begin{array}{c|c|c} 0 & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline 0 & 0 & \end{array}\right] \end{array}. \tag{2.4}$$

The blocks marked 0 contain only zeros, while the block marked 1 contains 1 at some position and zeros everywhere else. The unmarked block for $\bar{n} > n$ and $\bar{n}' > n'$ has not as yet been determined.

The unmarked block in (2.4) is quite a complicated expression with, in general, values different from zero at those positions in which (2.3) is satisfied. On the other hand, the matrix representation of a generator of $U(2^{2j+1})$ would be given by (2.4) if the unmarked block had also the value zero.

This suggests, then, that to get the generators of $U(2^{2j+1})$ we multiply the operator (1.7) by some type of projection operator which does not change the value of those blocks in the matrix representation (2.4) that have already been determined, but which makes the block for $\bar{n} > n$ and $\bar{n}' > n'$ equal to zero. This is easily achieved with the help of the number operator

$$N \equiv \sum_{m=-j}^{j} b^+_m b^m. \tag{2.5}$$

The operator

$$\prod_{r=n'+1}^{2j+1} \left(\frac{N-r}{n'-r}\right), \tag{2.6}$$

when acting on a state with a number of particles $\bar{n}' > n'$, clearly gives zero, while for $\bar{n}' \leq n'$ it gives the same state multiplied by a constant which for $\bar{n}' = n'$ is just 1. It is clear, therefore, that the operators

$$b^+_{m_1} \cdots b^+_{m_n} b^{m'_{n'}} \cdots b^{m_1} \prod_{r=n'+1}^{2j+1} \left(\frac{N-r}{n'-r}\right) \tag{2.7}$$

are precisely the generators of the group $U(2^{2j+1})$.

While the explicit expression of the generators is very simple, they are not yet in normal form, i.e., they are not expressed as linear combinations of products of creation and annihilation operators in which all the creation are to the left of the annihilation operators. Before proceeding to put them in normal form, we change our states, and thus our generators, from the basis (1.4) to the one in which they are characterized by seniority and total angular momentum. This new basis is much more relevant to the physical applications we are contemplating.

3. THE GENERATORS OF $U(2^{2j+1})$ IN THE SENIORITY-ANGULAR-MOMENTUM BASIS

If we are working in a single shell of angular momentum j, the physically relevant classification of our states[5] includes the number of particles n, the seniority v, the total angular momentum J, and projection M. These states are linear combinations of the states (1.4) given by

$$|nv\beta JM\rangle \equiv P^+_{nv\beta JM} |0\rangle, \tag{3.1}$$

with

$$P^+_{nv\beta JM} \equiv \sum_{m_1 \cdots m_n} A^{m_1 \cdots m_n}_{nv\beta JM} b^+_{m_1} \cdots b^+_{m_n}, \tag{3.2}$$

where the A's are appropriate coefficients.[2] The index

β, not yet defined, distinguishes between the possibly repeated states of given J corresponding to a definite seniority v.

The states (3.1) are characterized by the following chain of groups, below each one of which we give the IR to which they correspond:

$$\begin{array}{ccccccccc} U(2^{2j+1}) & \supset & O^+(4j+3) & \supset & O^+(4j+2) & \supset & U(2j+1) \supset Sp(2j+1) & \supset & O^+(3) \\ [1] & & [\tfrac{1}{2}^{2j+1}] & & [\tfrac{1}{2}\,\tfrac{1}{2}\cdots\tfrac{1}{2}(-1)^n\tfrac{1}{2}] & & [1^n] \quad\quad [1^v] & & J \end{array} . \tag{3.3}$$

The fact that the state (3.1) belongs to the IR J, $[1^v]$, $[1^n]$ of the ordinary proper rotation group $O^+(3)$, the symplectic group $Sp(2j+1)$, and the unitary group $U(2j+1)$ is well known.[5] Judd has shown that it also belongs to the IR $[\frac{1}{2}\,\frac{1}{2}\cdots\frac{1}{2}(-1)^n\frac{1}{2}]$ of the $O^+(4j+2)$ whose generators are given by (1.3) when we exclude $b^+_{m_1}$ and b^{m_2}. We have already mentioned that all states (1.4) belong to the IR $[\frac{1}{2}^{2j+1}]$ of $O^+(4j+3)$. Finally, we see from the representation of the generators (2.7) of $U(2^{2j+1})$, given by (2.4) when we put 0 in the unmarked block, that the states (1.4) belong to the IR [1] of $U(2^{2j+1})$. This would, then, also apply to the states (3.1) which are linear combinations of (1.4).

Let us now designate, by vectors **x** and **y**, respectively, the full sets of 2^{2j+1} states given by (1.4) and (3.1). From (3.2), these two vectors are related by a unitary matrix **A** which has only elements different from zero on blocks along the diagonal that correspond to the same number of particles n in (1.4) and (3.1). Thus, we have

$$\mathbf{y} = \mathbf{A}\mathbf{x}. \tag{3.4}$$

Then we denote by **B** the matrix of operators whose elements

$$B^{m_1'\cdots m_n'}_{m_1\cdots m_n} \tag{3.5}$$

are given by (2.7). Clearly, the matrix of generators of $U(2^{2j+1})$ in the basis (3.1) is obtained from **B** through the transformation

$$\mathbf{ABA}^+. \tag{3.6}$$

Before applying this transformation, we note that the operator N is unaffected by it. This can be seen from the fact that the representation of N with respect to the basis (1.4) or (3.1) is given also by a matrix with blocks along the diagonal, but these blocks are unit matrices multiplied by the number of particles. Thus, the matrix of N and **A** commute, and we can write

$$\begin{aligned} &\mathbf{ABA}^+ \\ &= \Big\| \sum_{m_1\cdots m_n} \sum_{m_1'\cdots m_n'} A^{m_1\cdots m_n}_{nv\beta JM} b^+_{m_1}\cdots b^+_{m_n} b^{m'n'}\cdots b^{m_1'} \\ &\quad \times A^{m_1'\cdots m_n'*}_{n'v'\beta'J'M'} \prod_{r=n'+1}^{2j+1}\left(\frac{N-r}{n'-r}\right)\Big\| \\ &= \Big\| P^+_{nv\beta JM} P^{n'v'\beta'J'M'} \prod_{r=n'+1}^{2j+1}\left(\frac{N-r}{n'-r}\right)\Big\|, \end{aligned} \tag{3.7}$$

where

$$P^{n'v'\beta'J'M'} = (P^+_{n'v'\beta'J'M'})^+. \tag{3.8}$$

Denoting now the generators of $U(2^{2j+1})$ in the basis (3.1) by $\mathbf{C}^{\gamma}_{\gamma'}$ with $\gamma \equiv nv\beta JM$, we see that they can be written in the form

$$\mathbf{C}^{n'v'\beta'J'M'}_{nv\beta JM} = P^+_{nv\beta JM} \prod_{r=1}^{2j+1-n'}\left(\frac{N-r}{-r}\right) P^{n'v'\beta'J'M'}, \tag{3.9}$$

where we put the projection operator between the creation and the annihilation polynomials, making use of the relation

$$b^m N = (N+1)b^m. \tag{3.10}$$

We now return to the question raised in the previous section, concerning the possibility of expressing the generators (3.9) of $U(2^{2j+1})$ in normal form, i.e., as linear combinations of the $(2^{2j+1})^2$ independent normal operators

$$\mathbf{D}^{n'v'\beta'J'M'}_{nv\beta JM} \equiv P^+_{nv\beta JM} P^{n'v'\beta'J'M'}. \tag{3.11}$$

For this purpose, we first express the projection operator in (3.9) in normal form. Since the operator N is a Casimir operator of the group $U(2j+1)$, whose generators are

$$\mathbf{C}^{m'}_{m} \equiv b^+_m b^{m'}, \tag{3.12}$$

the projection operator in (3.9) can be expanded only in terms of normal operators (3.11) with $n = n'$ that are contracted with respect to the quantum numbers v, β, J, and M that characterize the n-particle state. In fact, we prove by induction that

$$\prod_{r=1}^{2j+1-n'}\left(\frac{N-r}{-r}\right) = \sum_{s=0}^{2j+1-n'}(-1)^s \sum_{v\beta JM} P^+_{sv\beta JM} P^{sv\beta JM}. \tag{3.13}$$

For $n' = 2j$ the equality obviously holds, taking into account the definition (2.5) of N. Assuming that it holds for a given n', we proceed now to prove that it is also valid for $n' - 1$. The operator in the latter case is

$$\begin{aligned} &\left[\prod_{r=1}^{2j+1-n'}\left(\frac{N-r}{-r}\right)\right]\left(\frac{N+n'-2j-2}{n'-2j-2}\right) \\ &= \sum_{s=0}^{2j+1-n'} \frac{(-1)^s(s+n'-2j-2)}{(n'-2j-2)} \sum_{v\beta JM} P^+_{sv\beta JM} P^{sv\beta JM} \\ &\quad + \sum_{s=0}^{2j+1-n'} \frac{(-1)^s}{(n'-2j-2)} \sum_{\substack{v\beta JM \\ m}} P^+_{sv\beta JM} b^+_m b^m P^{sv\beta JM}. \end{aligned} \tag{3.14}$$

We can now use the relation

$$P^+_{s\nu\beta J M} b^+_m = \sum_{\nu'\beta' J' M'} (s+1)^{\frac{1}{2}} \langle JjMm \mid J'M' \rangle \times \langle j^s \nu\beta J; j \,|\} j^{s+1} \nu'\beta' J' \rangle P^+_{s+1\nu'\beta' J' M'}, \tag{3.15}$$

where $\langle \mid \rangle$, $\langle \mid\} \rangle$ are, respectively, a Clebsch–Gordan and a fractional parentage coefficient (fpc) and the factor $(s+1)^{\frac{1}{2}}$ comes from the fact that[6] we are dealing with anticommuting creation operators and not antisymmetric states. Substituting (3.15) in the last part of (3.14) together with a similar relation for the Hermitian conjugate operator $b^m P^{s\nu\beta JM}$ and using the orthonormality properties[7] of fpc's and Clebsch–Gordan coefficients, we can immediately prove that the projection operator (3.14) has again the form (3.13), but with n' replaced by $n'-1$.

Now, replacing the projection operator in (3.9) by (3.13) and using a generalization of the development (3.15), i.e.,

$$P^+_{n\nu\beta JM} P^+_{s\bar\nu\bar\beta\bar J\bar M} = \sum_{\hat\nu\hat\beta\hat J\hat M} \left(\frac{(n+s)!}{n!\,s!}\right)^{\frac{1}{2}} \langle J\bar J M\bar M \mid \hat J\hat M\rangle \times \langle j^n \nu\beta J; j^s \bar\nu\bar\beta\bar J \,|\} j^{n+s} \hat\nu\hat\beta\hat J\rangle P^+_{n+s\hat\nu\hat\beta\hat J\hat M}, \tag{3.16}$$

where $\langle \mid\} \rangle$ is now an s-particle fpc and the term $[(n+s)!/n!\,s!]^{\frac{1}{2}}$ appears for the same reasons[6] as $(s+1)^{\frac{1}{2}}$ in (3.15), we can express the generators (3.9) of $U(2^{2j+1})$ in the following normal form:

$$\mathbf{C}^{n'\nu'\beta' J' M'}_{n\,\nu\,\beta\,J\,M} = \sum_s \sum_{\substack{\hat\nu\hat\beta\hat J\hat M \\ \hat\nu'\hat\beta'\hat J'\hat M'}} \left[(-1)^s \frac{1}{s!}\left(\frac{(n+s)!\,(n'+s)!}{n!\,n'!}\right)^{\frac{1}{2}} \times \sum_{\bar\nu\bar\beta\bar J\bar M} [\langle J\bar J M\bar M \mid \hat J\hat M\rangle\langle J'\bar J M'\bar M \mid \hat J'\hat M'\rangle \times \langle j^n \nu\beta J; j^s \bar\nu\bar\beta\bar J \,|\} j^{n+s}\hat\nu\hat\beta\hat J\rangle \times \langle j^{n'}\nu'\beta' J'; j^s \bar\nu\bar\beta\bar J \,|\} j^{n'+s}\hat\nu'\hat\beta'\hat J'\rangle] \mathbf{D}^{n'+s\,\hat\nu'\hat\beta'\hat J'\hat M'}_{n+s\,\hat\nu\,\hat\beta\,\hat J\,\hat M}\right]. \tag{3.17}$$

The s-particle fpc can be expressed in terms of the single-particle fpc by well-known procedures,[7] and so the expansion of the generators of $U(2^{2j+1})$ in terms of normal operators can be given explicitly, if we know the single-particle fpc in the j shell.

Using properties of the fpc and Clebsch–Gordan coefficients, we could invert the relation (3.17), but it is actually much simpler to note that an alternative form of the generator (3.9) is given by

$$\mathbf{C}^{n'\nu'\beta' J' M'}_{n\,\nu\,\beta\,J\,M} = |n\nu\beta JM\rangle\langle n'\nu'\beta' J' M'| = P^+_{n\nu\beta JM} |0\rangle\langle 0| P^{n'\nu'\beta' J' M'}. \tag{3.18}$$

On the other hand, the normal operator (3.11) can be written, when introducing a unit operator between P^+ and P, as

$$\mathbf{D}^{n'\nu'\beta' J' M'}_{n\,\nu\,\beta\,J\,M} = P^+_{n\nu\beta JM}\left(\sum_{s\bar\nu\bar\beta\bar J\bar M} |s\bar\nu\bar\beta\bar J\bar M\rangle\langle s\bar\nu\bar\beta\bar J\bar M|\right) P^{n'\nu'\beta' J' M'} = \sum_{s\bar\nu\bar\beta\bar J\bar M} P^+_{n\nu\beta JM} P^+_{s\bar\nu\bar\beta\bar J\bar M} |0\rangle\langle 0| P^{s\bar\nu\bar\beta\bar J\bar M} P^{n'\nu'\beta' J' M'}. \tag{3.19}$$

Using (3.16) again, we obtain

$$\mathbf{D}^{n'\nu'\beta' J' M'}_{n\,\nu\,\beta\,J\,M} = \sum_s \sum_{\substack{\hat\nu\hat\beta\hat J\hat M \\ \hat\nu'\hat\beta'\hat J'\hat M'}} \left[\frac{1}{s!}\left(\frac{(n+s)!\,(n'+s)!}{n!\,n'!}\right)^{\frac{1}{2}} \times \sum_{\bar\nu\bar\beta\bar J\bar M} (\langle J\bar J M\bar M \mid \hat J\hat M\rangle\langle J'\bar J M'\bar M \mid \hat J'\hat M'\rangle \times \langle j^n \nu\beta J; j^s \bar\nu\bar\beta\bar J \,|\} j^{n+s}\hat\nu\hat\beta\hat J\rangle \times \langle j^{n'}\nu'\beta' J'; j^s \bar\nu\bar\beta\bar J \,|\} j^{n'+s}\hat\nu'\hat\beta'\hat J'\rangle) \mathbf{C}^{n'+s\,\hat\nu'\hat\beta'\hat J'\hat M'}_{n+s\,\hat\nu\,\hat\beta\,\hat J\,\hat M}\right]. \tag{3.20}$$

The relations (3.17) and (3.20) between the generators of $U(2^{2j+1})$ and the normal operators (3.11) are the basic ones for the applications we want to present in this paper. Before proceeding with their discussion, we introduce the concept of complementary group and indicate those of its properties that will be useful in the following discussion.

4. THE CONCEPT OF COMPLEMENTARY GROUP

The concept we want to introduce appears implicitly in the relations between the noninvariance group $O^+(4j+2)$ and its subgroups $Sp(2j+1)$ and $SU^Q(2)$, where the latter is the quasispin group.

As mentioned above, the generators of $O^+(4j+2)$ are given by

$$b^+_{m_1} b^+_{m_2}, \quad \tfrac{1}{2}(b^+_{m_1} b^{m_2} - b^{m_2} b^+_{m_1}) = \mathbb{C}^{m_2}_{m_1} - \tfrac{1}{2}\delta^{m_2}_{m_1}, \quad b^{m_1} b^{m_2}. \tag{4.1}$$

The symplectic group $Sp(2j+1)$ is a subgroup of $O^+(4j+2)$ and its generators are[2,3]

$$\Lambda^{m_2}_{m_1} = \mathbb{C}^{m_2}_{m_1} + (-1)^{m_1+m_2}\mathbb{C}^{-m_1}_{-m_2}. \tag{4.2}$$

If we now consider all the invariants with respect to the symplectic group $Sp(2j+1)$ that we can form from the generators (4.1) of $O^+(4j+2)$, we arrive at the

generators of the quasispin group $SU^Q(2)$, i.e.,

$$Q_+ = \tfrac{1}{2}\sum_m (-1)^{j-m} b_m^+ b_{-m}^+,$$
$$Q_0 = \tfrac{1}{2}\Big[\sum_m b_m^+ b^m - (j+\tfrac{1}{2})\Big], \qquad (4.3)$$
$$Q_- = -\tfrac{1}{2}\sum_m (-1)^{j-m} b^m b^{-m},$$

which satisfy the usual commutation rules of the spherical components of angular momentum.

The generators of $Sp(2j+1)$ and $SU^Q(2)$, of course, commute. Their direct product

$$Sp(2j+1) \times SU^Q(2) \qquad (4.4)$$

is a subgroup of $O^+(4j+2)$, and we could characterize alternatively the states (3.1) by the IR of this subgroup rather than follow the chain

$$O^+(4j+2) \supset U(2j+1) \supset Sp(2j+1)$$

of (3.3). The important point to note is that the IR of $SU^Q(2)$ and $Sp(2j+1)$ contained in the IR $[\frac{1}{2}\,\frac{1}{2}\cdots\frac{1}{2}(-1)^n\frac{1}{2}]$ of $O^+(4j+2)$ are not independent but have the 1-to-1 correspondence[4]

$$Q = \tfrac{1}{2}[(j+\tfrac{1}{2}) - v], \qquad (4.5)$$

where v is the seniority and Q the value of the quasispin, as can be seen easily from the Casimir operators of both groups. This relation suggests the following definition of complementarity for groups.[8]

Let us consider a direct product of two groups H_1 and H_2, which are subgroups of a larger group H. We refer to the groups H_1 and H_2 as "complementary" within a definite IR of H, if there is a 1-to-1 correspondence between all the IR of H_1 and of H_2 contained in this IR of H. Thus, for example, $SU^Q(2)$ and $Sp(2j+1)$ are complementary within the IR $[\frac{1}{2}\,\frac{1}{2}\cdots\frac{1}{2}(-1)^n\frac{1}{2}]$ of $O^+(4j+2)$.

We proceed to discuss the complementarity relation when the larger group H is $U(2^{2j+1})$. All states (3.1) belong to the single IR [1] of $U(2^{2j+1})$, and we could ask ourselves which are the complementary groups H_2 when H_1 is taken as any of the subgroups of $U(2^{2j+1})$ in the chain (3.3). We saw above that the complementary group to $Sp(2j+1)$ was obtained by considering all symplectic invariants formed from the generators of $O^+(4j+2)$. This suggests, then, that we consider all possible invariants, with respect to the group H_1 that we can form from the generators of $U(2^{2j+1})$, and identify the resulting group, trying to prove at the same time that it satisfies the complementarity definition with respect to H_1.

We proceed to implement this program for all the groups of the chain (3.3) starting with its last member $O^+(3)$.

If we look at the generators (3.9) of $U(2^{2j+1})$, it is immediately clear that invariants with respect to $O^+(3)$ can be formed when we put $J = J'$, $M = M'$, and sum over M, i.e., the operators we get can be denoted by

$$C^{n'v'\beta'J}_{n\,v\,\beta\,J} \equiv \sum_M \mathbf{C}^{n'v'\beta'JM}_{n\,v\,\beta\,JM}. \qquad (4.6)$$

Now from the matrix expression of the operators $\mathbf{C}^{\gamma'}_{\gamma}$, $\gamma \equiv nv\beta JM$, we conclude that they satisfy the commutation relations of the generators of the unitary group, i.e.,

$$[\mathbf{C}^{\gamma'}_{\gamma}, \mathbf{C}^{\gamma''}_{\gamma'''}] = \mathbf{C}^{\gamma''}_{\gamma}\delta^{\gamma'}_{\gamma'''} - \mathbf{C}^{\gamma'}_{\gamma'''}\delta^{\gamma''}_{\gamma}, \qquad (4.7)$$

where the δ's are symbolic expressions for the product of Kronecker δ's in all the quantum numbers. Then, looking at the definition (4.6), we conclude from (4.7) that two operators $C^{n'v'\beta'J}_{n\,v\,\beta\,J}$ commute when they belong to different J's and that, when they have the same J, they satisfy a commutation relation of the type (4.7) in the remaining indices. It is clear, therefore, that the operators (4.6) are the generators of a direct sum of unitary groups associated with each possible value of J, whose dimension $d(J)$ is the same as the number of states in the shell of angular momentum j that have total angular momentum J, divided by the degeneracy $2J+1$ of these states. The group of invariants with respect to $O^+(3)$ that we can form from the generators of $U(2^{2j+1})$ is then

$$\sum_J \oplus\, U(d(J)). \qquad (4.8)$$

The generators of the group (4.8) commute, of course, with the components of angular momentum[2]

$$L_q = [j(j+1)]^{\frac{1}{2}} \sum_{m,m'} \langle j1m'q \,|\, jm\rangle \mathbf{C}^{m'}_{m}$$
$$= \tfrac{1}{2}[j(j+1)]^{\frac{1}{2}} \sum_{m,m'} \langle j1m'q \,|\, jm\rangle \Lambda^{m'}_{m}, \qquad q = 1, 0, -1, \quad (4.9)$$

which are the generators of $O^+(3)$. It remains to see whether there is a 1-to-1 correspondence between the IR of the group (4.8) and those of $O^+(3)$, when the states we are analyzing all belong to the single IR [1] of $U(2^{2j+1})$.

The answer to this question is immediate. The IR of the group (4.8) can be characterized by the set of all the IR of the groups in the direct sum. Now, if we are dealing with states belonging to a definite IR J of $O^+(3)$, we see that they belong to the IR [1] of $U(d(J))$ for the same J and to the IR [0] of all the other groups in the direct sum. Thus, there is a 1-to-1 correspondence between the IR of $O^+(3)$ and that of the group (4.8) and so, if we identify $O^+(3)$ with H_1, we could denote (4.8) by H_2.

The complementary groups for the other groups in the chain (3.3) follow now by a similar reasoning. For $Sp(2j+1)$, the invariants we can form from the

generators of $U(2^{2j+1})$ are clearly

$$C^{n'v}_{n\,v} \equiv \sum_{\beta J M} \mathbf{C}^{n'v\beta JM}_{n\,v\beta JM}. \tag{4.10}$$

These generators commute with those of the group $Sp(2j+1)$ given by (4.2). They are the generators of the direct sum of unitary groups

$$\sum_{v} \oplus\, U(d(v)), \tag{4.11}$$

whose dimension $d(v)$ is the same as the number of states in the shell of angular momentum j that have seniority v divided by the degeneracy of the states with this seniority. By the same arguments as in the case of $O^+(3)$ we conclude that (4.11) is the complementary group to $Sp(2j+1)$. Since the quasispin group is the complementary group to $Sp(2j+1)$ within the IR $[\frac{1}{2}\,\frac{1}{2}\cdots\frac{1}{2}\,(-1)^n\frac{1}{2}]$ of $O^+(4j+2)$, which is a subgroup of $U(2^{2j+1})$, it is clear that the quasispin group is a subgroup of (4.11). Moreover, the generators of (4.11) can be written as linear combinations of powers of the quasispin operators of the type $Q^n_+ Q^{n_0}_0 Q^{n_-}_-$.

For the group $U(2j+1)$ the invariants we can form are clearly

$$C^n_n \equiv \sum_{v\beta J M} \mathbf{C}^{nv\beta JM}_{nv\beta JM}, \tag{4.12a}$$

and so the complementary group is

$$\sum_{n} \oplus\, U(1). \tag{4.13}$$

For the group $O^+(4j+2)$, the complementary group is

$$U(1) \oplus U(1), \tag{4.14}$$

where the generators for these two unitary groups are

$$C^+_+ \equiv \sum_{n(\text{even})} \sum_{v\beta JM} \mathbf{C}^{nv\beta JM}_{nv\beta JM}, \quad C^-_- \equiv \sum_{n(\text{odd})} \sum_{v\beta JM} \mathbf{C}^{nv\beta JM}_{nv\beta JM}, \tag{4.15a}$$

respectively. Finally for $O^+(4j+3)$ the complementary group is just $U(1)$ with generator

$$C \equiv \sum_{n=0}^{2j+1} \sum_{v\beta JM} \mathbf{C}^{nv\beta JM}_{nv\beta JM}. \tag{4.16a}$$

Note that alternative forms for (4.12a), (4.15a), and (4.16a) are respectively

$$C^n_n = \prod_{\substack{r=0 \\ r\neq n}}^{2j+1} \left(\frac{N-r}{n-r}\right), \tag{4.12b}$$

$$C^+_+ = \prod_{\substack{r=1 \\ r\ \text{odd}}}^{2j} \left(\frac{N-r}{n-r}\right), \quad C^-_- = \prod_{\substack{r=0 \\ r\ \text{even}}}^{2j+1} \left(\frac{N-r}{n-r}\right), \tag{4.15b}$$

$$C = 1, \tag{4.16b}$$

where 1 is the unity operator.

We have thus obtained the complementary groups to all the subgroups in the chain (3.3) within the single IR [1] of $U(2^{2j+1})$. The IR of the complementary group H_2, together with those of a canonical[9] set of its subgroups, would then completely characterize the states associated with a given IR of the original group H_1.

We make use of the concepts developed in this section in the following applications.

5. APPLICATION TO MANY-BODY FORCES

In the SQP the operator associated with the Hamiltonian is in general a sum of operators associated with 0-, 1-, 2-, 3-, $\cdots$, $(2j+1)$-body forces, i.e.,

$$H = \sum_{p=0}^{2j+1} H^{(p)}, \tag{5.1}$$

whose explicit expressions are given by

$$H^{(p)} = \sum_{v\beta v'\beta' J} \langle pv\beta J|\, H^{(p)}\, |pv'\beta' J\rangle\, D^{pv'\beta' J}_{pv\,\beta\,J}, \tag{5.2}$$

where

$$D^{pv'\beta' J}_{pv\,\beta\,J} = \sum_{M} P^{+}_{pv\beta JM} P^{pv'\beta' JM}. \tag{5.3}$$

We have implicitly assumed that all $H^{(p)}$ are invariant under rotations, so their matrix elements are diagonal in J and M and independent of M. Thus, the operators (5.2), instead of being linear combinations[2] of the normal operators (3.11) with $n = n'$, are given in terms of the contracted operators (5.3).

The operator associated with the Hamiltonian H in the SQP can also be written as

$$H = \sum_{nv\beta JM} \sum_{n'v'\beta' J' M'} |nv\beta JM\rangle\, \langle nv\beta JM|\, H\, |n'v'\beta' J' M'\rangle \times \langle n'v'\beta' J' M'|. \tag{5.4}$$

Taking into account that the matrix element of H only connects states with the same number of particles, i.e., $n = n'$, and that H is invariant under rotations, we can, using (3.18) and (4.6), write

$$H = \sum_{n} \sum_{v\beta v'\beta' J} \langle nv\beta J|\, H\, |nv'\beta' J\rangle\, C^{nv'\beta' J}_{nv\,\beta\,J}. \tag{5.5}$$

Now, from the relation (3.17) between the generators (3.9) of $U(2^{2j+1})$ and the normal operators (3.11), we can derive a similar relation between the generators (4.6) of the complementary group to $O^+(3)$ and the normal operators (5.3). Using this relation, we write H in terms of $D^{pv'\beta' J}_{pv\,\beta\,J}$ as

$$H = \sum_{nv\beta v'\beta' J} \Bigg(\langle nv\beta J|\, H\, |nv'\beta' J\rangle \sum_{p\hat{v}\hat{\beta}\hat{v}'\hat{\beta}'\hat{J}} (-1)^{p-n} \times \frac{p!}{n!\,(p-n)!} \sum_{\bar{v}\bar{\beta}\bar{J}} \langle j^n v\beta J;\, j^{p-n}\bar{v}\bar{\beta}\bar{J}\,|\}\, j^p\hat{v}\hat{\beta}\hat{J}\rangle \times \langle j^n v'\beta' J;\, j^{p-n}\bar{v}\bar{\beta}\bar{J}\,|\}\, j^p\hat{v}'\hat{\beta}'\hat{J}\rangle D^{p\hat{v}'\hat{\beta}'\hat{J}}_{p\hat{v}\,\hat{\beta}\,\hat{J}} \Bigg). \tag{5.6}$$

Comparing (5.6) with (5.1) and (5.2), we arrive at the following relation between the matrix elements of the p-body Hamiltonian and the matrix elements of the full Hamiltonian H:

$$\langle pv\beta J|\, H^{(p)}\, |pv'\beta' J\rangle = \sum_{n=0}^{p}\Bigg[(-1)^{p-n}\frac{p!}{n!\,(p-n)!}\sum_{\hat{v}\hat{\beta}\hat{v}'\hat{\beta}'\hat{J}} \times \Bigg(\sum_{\bar{v}\bar{\beta}\bar{J}}\langle j^n\hat{v}\hat{\beta}\hat{J}; j^{p-n}\bar{v}\bar{\beta}\bar{J}\,|\}\, j^p v\beta J\rangle \times \langle j^n\hat{v}'\hat{\beta}'\hat{J}; j^{p-n}\bar{v}\bar{\beta}\bar{J}\,|\}\, j^p v'\beta' J\rangle \times \langle n\hat{v}\hat{\beta}\hat{J}|\, H\, |n\hat{v}'\hat{\beta}'\hat{J}\rangle\Bigg)\Bigg]. \quad (5.7)$$

The matrix elements of H appearing in (5.5) could be taken, in many cases, directly from experience, if we assume that the single-shell description is appropriate. They could be given, for example, by the binding energy of nuclei with closed shells of protons and a single open shell of neutrons, as well as by the energy levels of excited states of these same nuclei. If a full set of matrix elements of H can be obtained in this way we could, from (5.7), determine the contributions of effective p-body forces in the shell under consideration. An analysis for nuclei with either protons or neutrons in an open $f_{\frac{7}{2}}$ shell has been presented elsewhere.[10]

6. APPLICATION TO THE STATE-LABELING PROBLEM

When classifying states by the IR of a noncanonical[9] chain of groups such as (3.3), the labels provided by these IR may not be sufficient. For example, we generally need, for a given seniority v [IR of $Sp(2j+1)$], extra quantum numbers β to distinguish between repeated states of definite J [IR of $O^+(3)$].

We need, therefore, extra operators whose eigenvalues provide new quantum numbers of the β type. The procedure for obtaining these operators has been discussed elsewhere[11] and, for the particular case just mentioned, they can be derived as follows: Consider the generators $\Lambda^{m_1}_{m_2}$ defined by (4.2) and form linear combinations of powers of them that are scalars with respect to the rotation group $O^+(3)$. As examples, we have

$$\sum_q\Bigg[(-1)^q\sum_{m_1m_2}\Big(\langle jkm_1q\,|\,jm_2\rangle\Lambda^{m_1}_{m_2}\Big) \times\sum_{m_1'm_2'}\Big(\langle jkm_1'-q\,|\,jm_2'\rangle\Lambda^{m_1'}_{m_2'}\Big)\Bigg], \quad (6.1a)$$

$$\sum_{m_1m_2m_3}\Lambda^{m_2}_{m_1}\Lambda^{m_3}_{m_2}\Lambda^{m_1}_{m_3}, \quad (6.1b)$$

etc. The operators of the type (6.1), when applied to an arbitrary state (3.1), obviously cannot change the quantum numbers n, v, J, M of the state. Clearly, then, their matrix with respect to the full set of states (3.1) breaks into blocks characterized by v and J and of dimension equal to the number of different values of β we have. Diagonalization of these submatrices then characterizes the states by the eigenvalues of the operators of the type (6.1).

We may require one or more[12] operators of the type (6.1) to completely characterize the states. It is interesting, therefore, to analyze the properties of this family of operators. We note first that the set of linearly independent operators of the type (6.1) is finite, since in the SQP we have no more than $(2^{2j+1})^2$ operators of all types. Furthermore, the commutator of two operators of type (6.1) is again of type (6.1), as it is a polynomial function of the $\Lambda^{m_2}_{m_1}$ and is invariant under $O^+(3)$. Thus, the set of all linearly independent operators of this type form a Lie algebra.

To find what the type of Lie algebra is, we recall that the operators (6.1) do not change the quantum numbers n, v, J, M of the states, and also that their matrix elements are independent of M [because the operators are scalars with respect to $O^+(3)$] and of n [because the matrix element of $\Lambda^{m_2}_{m_1}$ with respect to the states (3.1) is independent of n and this also applies to any function of the generators of the symplectic group]. Thus, all operators of the type (6.1) can be expressed as linear combinations of

$$C^{v\beta' J}_{v\beta\ J} \equiv \sum_{n=0}^{2j+1}\sum_{M=-J}^{J} \mathbf{C}^{nv\beta' JM}_{nv\beta\ JM}. \quad (6.2)$$

By a similar analysis to the one presented in Sec. 4, we see that the operators (6.2) are generators of the direct sum of unitary groups

$$\sum_{(v,J)}\oplus\, U(d(v,J)), \quad (6.3)$$

where the dimension $d(v,J)$ stands for the number of independent states of definite J that we have for a given seniority v. In this enumeration we regard as identical all states that differ only by M or n.

We have identified the group to which belongs the full set of linearly independent state-labeling operators. Had we decided to use one or more of these operators in the actual characterization of the states, it would have been quite convenient to expand them in terms of the operators (6.2), in which we use some arbitrary labeling procedure for the β's. This would facilitate the determination of the eigenvalues of these operators, as the matrix elements of the generators (6.2) with respect to the states (3.1) can be obtained immediately.

7. APPLICATION TO THE QUASIPARTICLE PICTURE

In recent years, the concept of quasiparticle[13] has played an important role in many branches of physics. While the quasispin group (4.3) is related to it, the concept itself has not yet been discussed within a general group-theoretical framework. With the help of the noninvariance group $U(2^{2j+1})$ and its subgroups in the chain (3.3), we shall initiate this discussion.

We note first that our creation operators b_m^+ and annihilation operators b^m are given respectively in covariant and contravariant components. We can raise or lower indices,[14] using the metric tensor $(-1)^{j-m}\delta_{m,-m'}$, and thus obtain

$$b^{+m} = (-1)^{j+m}b_{-m}^+, \quad b_m = (-1)^{j-m}b^{-m}. \tag{7.1}$$

The Bogoliubov–Valatin transformation (BVT) gives us the new creation and annihilation operators $\mathfrak{b}_m^+$ and $\mathfrak{b}^m$, related to the previous ones by

$$\mathfrak{b}_m^+ = ub_m^+ + w(-1)^{j-m}b^{-m}, \tag{7.2a}$$

$$\mathfrak{b}^m = w^*(-1)^{j-m}b_{-m}^+ + u^*b^m, \tag{7.2b}$$

where u and w are arbitrary complex numbers satisfying the condition

$$uu^* + ww^* = 1. \tag{7.3}$$

We could, if we so wished, write u and w in the form

$$u = e^{-\frac{1}{2}i\varphi}\cos\tfrac{1}{2}\theta\, e^{-\frac{1}{2}i\chi}, \quad w = e^{\frac{1}{2}i\varphi}\sin\tfrac{1}{2}\theta\, e^{-\frac{1}{2}i\chi}, \tag{7.4}$$

where φ, θ, and χ are Euler angles in some abstract space.

The operators $\mathfrak{b}_m^+$ and $\mathfrak{b}^{m'}$ satisfy the anticommutation rules (1.2) and so, in terms of them, we can construct the generators of a new $U(2^{2j+1})$ group given by (3.9) when $\mathfrak{b}_m^+$ and $\mathfrak{b}^{m'}$ replace b_m^+ and $b^{m'}$. It is obvious, though, that these new generators will be linear combinations of the old ones, and so the new $U(2^{2j+1})$ is identical to the one whose generators are (3.9). From (1.3), it is clear that the new and the original $O^+(4j+3)$ groups are also identical and that this continues to hold for $O^+(4j+2)$.

Furthermore, one can easily see from the BVT (7.2) that the generators of the symplectic group remain invariant under it, i.e.,

$$\mathfrak{b}_{m_1}^+\mathfrak{b}^{m_2} + (-1)^{m_1+m_2}\mathfrak{b}_{-m_2}^+\mathfrak{b}^{-m_1} = b_{m_1}^+b^{m_2} + (-1)^{m_1+m_2}b_{-m_2}^+b^{-m_1}. \tag{7.5}$$

From the expression (4.9) for the generators of $O^+(3)$, we conclude that they remain invariant under the BVT, as can also be seen from the fact that $O^+(3)$ is a subgroup of $Sp(2j+1)$.

Thus, in the chain (3.3), all the groups transform into themselves under the BVT, except for $U(2j+1)$. As the IR of this group $[1^n]$ is characterized by the number of particles n, it is clear that the states associated with the new creation and annihilation operators (7.2) will be linear combinations of the states (3.1) with the same v, β, J, and M, but different n. In fact, this combination is very easily determined when we notice that the states (3.1) can also be written[4] in terms of the quasispin and its projection, i.e.,

$$|nv\beta JM\rangle = |QM_Q\beta JM\rangle, \tag{7.6}$$

where Q and v are related by (4.5) and

$$M_Q = \tfrac{1}{2}[n - (j+\tfrac{1}{2})]. \tag{7.7}$$

Clearly, then, the BVT carries the original states (3.1) into new ones of the form

$$\sum_{M'_Q} |QM'_Q\beta JM\rangle\, \mathfrak{D}^Q_{M'_QM_Q}(\varphi, \theta, \chi), \tag{7.8}$$

where the $\mathfrak{D}$'s are the familiar[15] Wigner functions associated with the angles introduced in (7.4). This result becomes obvious when we note that the BVT gives rise to a rotation of the generators (4.3) of the quasispin group.

The advantage of the quasiparticle picture is that now the Hamiltonian (5.5) can be expressed in terms of generators of the complementary group to $O^+(3)$ associated with $\mathfrak{b}_m^+$ and $\mathfrak{b}^{m'}$ with coefficients that depend on the matrix elements of H, as well as on the parameters u and w. We could manipulate these parameters to achieve simplifications in the Hamiltonian that allow us to determine the eigenstates in terms of (7.8). Then, we have a solution to the problem in terms of states in which the number of particles is not conserved. Standard projection techniques[13] permit us to derive from this state a solution with a definite number of particles.

8. GENERALIZATION TO MORE COMPLEX CONFIGURATIONS

So far, our SQP has been restricted to a shell where all the 1-particle states have total angular momentum j. It is clear, though, that generalizations to more complex configurations are obvious. Thus, if we have two types of particles in the shell, e.g., protons and neutrons, we just introduce the isotopic spin projection τ and have the creation and annihilation operators

$$b_{m\tau}^+, \quad b^{m'\tau'}, \quad m, m' = j, \cdots, -j, \quad \tau, \tau' = \pm\tfrac{1}{2}. \tag{8.1}$$

In the case of several shells, we can introduce for these operators the index ρ that takes the values

$$\rho = nlm_lm_s\tau, \quad \rho = nljm\tau, \tag{8.2}$$

etc., depending on whether we use l–s coupling, for which m_l, m_s are the projection of orbital angular momentum and spin, or j–j coupling, or other ways of characterizing the states.

All the steps in the previous sections can be extended to these generalized operators, though expansions such as (3.17) would require the knowledge of more complex fpc's.

A generalization of particular interest concerns all the states in a single level of the harmonic oscillator for which

$$\rho = lm_l m_s \tau, \quad l = \nu, \nu - 2, \cdots, 1 \text{ or } 0, \quad (8.3)$$

where ν is the total number of quanta in the given level. The $U(2j+1)$ group now becomes[2]

$$U(2(\nu+1)(\nu+2))$$

which, among its subgroups, has the symmetry group $U(3)$ of the harmonic oscillator. Applications of the noninvariance group $U(2^{2(\nu+1)(\nu+2)})$ of this problem are planned in a future publication.

* Work supported by Comisión Nacional de Energía Nuclear (México).

† Chercheur agréé, Institut Interuniversitaire des Sciences Nucléaires; on leave of absence from Université Libre de Bruxelles, Belgium.

[1] N. Mukunda, L. O'Raifeartaigh, and E. C. G. Sudarshan, Phys. Rev. Letters **15**, 1041 (1965); Phys. Rev. **139**, B839 (1965).
[2] M. Moshinsky, *Group Theory and the Many Body Problem* (Gordon and Breach, New York, 1968).
[3] K. Helmers, Nucl. Phys. **23**, 594 (1961).
[4] B. R. Judd, in *Group Theory and Applications*, E. M. Loebl, Ed. (Academic, New York, 1968).
[5] M. Hamermesh, *Group Theory and Its Application to Physical Problems* (Addison-Wesley, Reading, Mass., 1964), Chap. 11.
[6] C. Quesne, Rev. Mex. Fís. **19** (1970) (to be published).
[7] A. de-Shalit and I. Talmi, *Nuclear Shell Theory* (Academic, New York, 1963), Chaps. 26 and 28.
[8] M. Moshinsky and C. Quesne, Phys. Letters. **29B**, 482 (1969).
[9] M. Moshinsky, J. Math. Phys. **4**, 1128 (1963).
[10] C. Quesne, Phys. Letters **31B**, 7 (1970).
[11] V. Bargmann and M. Moshinsky, Nucl. Phys. **23**, 177 (1961).
[12] M. Moshinsky and J. G. Nagel, Phys. Letters **5**, 173 (1963).
[13] M. H. Macfarlane, *Lectures in Theoretical Physics* (University of Colorado Press, Boulder, Colo., 1966), p. 621.
[14] E. P. Wigner, *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra* (Academic, New York, 1959), p. 293.
[15] See Ref. 14, pp. 166–170.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 5 MAY 1970

Geometrical Derivation of the Conservation Laws

JERRY L. ALDERMAN* AND OTTO BERGMANN
Department of Physics, George Washington University, Washington, D.C. 20037

(Received 17 December 1969)

Conservation laws which are customarily obtained by invoking the invariance of a Lagrangian density under the transformations of some intrinsic symmetry group may be given a completely geometrical treatment within the context of the theory by Yang and Mills. The formalism is extended from unitary transformations to general linear transformations and the concepts of parallel transfer, covariant differentiation, and intrinsic curvature tensor are discussed. Conservation laws follow from the assumed invariance of the Lagrangian under parallel transfer defined with a general affine connection which is a direct sum of intrinsic and space–time affinities. Conservation of generalized charge is a consequence of the arbitrariness of that part of the affinities which operates in the intrinsic space, and conservation of energy and momentum is related to the arbitrariness of the part of the affinities which operates on the space–time indices. Some comments on Palatini's derivation of Einstein's equation of general relativity are made.

I. GEOMETRY IN THE INTRINSIC SPACE

Unitary transformations in three dimensions can be defined by the invariance of a Hermitian form

$$\sum_{\alpha=1}^{3} \phi_\alpha^* \psi_\alpha \quad (1)$$

under linear transformations. In this paper, we use geometrical concepts which can be cast in a familiar analytical form, even for complex-valued vectors, if van der Waerden's[1] distinction between tensor indices is used. The notation ϕ_α instead of ϕ_α^* is obviously more general, since higher tensors with any number of dotted and undotted indices can be considered. The difference lies, of course, in the transformation laws

$$\psi'_\alpha = S_\alpha^\beta \psi_\beta, \quad \phi'_\alpha = S_\alpha^\beta \phi_\beta = (S_\alpha^\beta)^* \phi_\beta. \quad (2)$$

The most general method for constructing invariants makes use of a dual space which can be formulated in terms of contravariant vectors with the transformation laws

$$\psi^{\alpha'} = (S^{-1})_\beta^\alpha \psi^\beta, \quad \phi^{\dot{\alpha}'} = (S^{-1})_{\dot{\beta}}^{\dot{\alpha}} \phi^{\dot{\beta}} = (S^{-1})_\beta^{\alpha *} \phi^{\dot{\beta}}. \quad (3)$$

The familiar rules for contraction of indices holds

etc., depending on whether we use l–s coupling, for which m_l, m_s are the projection of orbital angular momentum and spin, or j–j coupling, or other ways of characterizing the states.

All the steps in the previous sections can be extended to these generalized operators, though expansions such as (3.17) would require the knowledge of more complex fpc's.

A generalization of particular interest concerns all the states in a single level of the harmonic oscillator for which

$$\rho = lm_l m_s \tau, \quad l = \nu, \nu - 2, \cdots, 1 \text{ or } 0, \quad (8.3)$$

where ν is the total number of quanta in the given level. The $U(2j+1)$ group now becomes[2]

$$U(2(\nu+1)(\nu+2))$$

which, among its subgroups, has the symmetry group $U(3)$ of the harmonic oscillator. Applications of the noninvariance group $U(2^{2(\nu+1)(\nu+2)})$ of this problem are planned in a future publication.

* Work supported by Comisión Nacional de Energía Nuclear (México).

† Chercheur agréé, Institut Interuniversitaire des Sciences Nucléaires; on leave of absence from Université Libre de Bruxelles, Belgium.

[1] N. Mukunda, L. O'Raifeartaigh, and E. C. G. Sudarshan, Phys. Rev. Letters **15**, 1041 (1965); Phys. Rev. **139**, B839 (1965).
[2] M. Moshinsky, *Group Theory and the Many Body Problem* (Gordon and Breach, New York, 1968).
[3] K. Helmers, Nucl. Phys. **23**, 594 (1961).
[4] B. R. Judd, in *Group Theory and Applications*, E. M. Loebl, Ed. (Academic, New York, 1968).
[5] M. Hamermesh, *Group Theory and Its Application to Physical Problems* (Addison-Wesley, Reading, Mass., 1964), Chap. 11.
[6] C. Quesne, Rev. Mex. Fís. **19** (1970) (to be published).
[7] A. de-Shalit and I. Talmi, *Nuclear Shell Theory* (Academic, New York, 1963), Chaps. 26 and 28.
[8] M. Moshinsky and C. Quesne, Phys. Letters. **29B**, 482 (1969).
[9] M. Moshinsky, J. Math. Phys. **4**, 1128 (1963).
[10] C. Quesne, Phys. Letters **31B**, 7 (1970).
[11] V. Bargmann and M. Moshinsky, Nucl. Phys. **23**, 177 (1961).
[12] M. Moshinsky and J. G. Nagel, Phys. Letters **5**, 173 (1963).
[13] M. H. Macfarlane, *Lectures in Theoretical Physics* (University of Colorado Press, Boulder, Colo., 1966), p. 621.
[14] E. P. Wigner, *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra* (Academic, New York, 1959), p. 293.
[15] See Ref. 14, pp. 166–170.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 5 MAY 1970

Geometrical Derivation of the Conservation Laws

JERRY L. ALDERMAN* AND OTTO BERGMANN
Department of Physics, George Washington University, Washington, D.C. 20037

(Received 17 December 1969)

Conservation laws which are customarily obtained by invoking the invariance of a Lagrangian density under the transformations of some intrinsic symmetry group may be given a completely geometrical treatment within the context of the theory by Yang and Mills. The formalism is extended from unitary transformations to general linear transformations and the concepts of parallel transfer, covariant differentiation, and intrinsic curvature tensor are discussed. Conservation laws follow from the assumed invariance of the Lagrangian under parallel transfer defined with a general affine connection which is a direct sum of intrinsic and space–time affinities. Conservation of generalized charge is a consequence of the arbitrariness of that part of the affinities which operates in the intrinsic space, and conservation of energy and momentum is related to the arbitrariness of the part of the affinities which operates on the space–time indices. Some comments on Palatini's derivation of Einstein's equation of general relativity are made.

I. GEOMETRY IN THE INTRINSIC SPACE

Unitary transformations in three dimensions can be defined by the invariance of a Hermitian form

$$\sum_{\alpha=1}^{3} \phi_\alpha^* \psi_\alpha \quad (1)$$

under linear transformations. In this paper, we use geometrical concepts which can be cast in a familiar analytical form, even for complex-valued vectors, if van der Waerden's[1] distinction between tensor indices is used. The notation ϕ_α instead of ϕ_α^* is obviously more general, since higher tensors with any number of dotted and undotted indices can be considered. The difference lies, of course, in the transformation laws

$$\psi'_\alpha = S_\alpha^\beta \psi_\beta, \quad \phi'_\alpha = S_\alpha^\beta \phi_\beta = (S_\alpha^\beta)^* \phi_\beta. \quad (2)$$

The most general method for constructing invariants makes use of a dual space which can be formulated in terms of contravariant vectors with the transformation laws

$$\psi^{\alpha'} = (S^{-1})_\beta^\alpha \psi^\beta, \quad \phi^{\dot{\alpha}'} = (S^{-1})_{\dot{\beta}}^{\dot{\alpha}} \phi^{\dot{\beta}} = (S^{-1})_\beta^{\alpha *} \phi^{\dot{\beta}}. \quad (3)$$

The familiar rules for contraction of indices holds

here too, provided that the indices are either both dotted or both undotted. A fully covariant form of (1) is obtained if we introduce an intrinsic metric tensor $\rho^{\alpha\beta}$, which is equal to the contravariant "unity" and which is to be distinguished from the mixed Kronecker deltas δ_{α}^{β} and δ_{α}^{β}. We write instead of (1) with the usual summation convention

$$\rho^{\alpha\beta}\phi_{\alpha}\psi_{\beta}. \tag{4}$$

The Hermitian property of the metric tensor in the intrinsic space,

$$\rho^{\alpha\beta} = (\rho^{\alpha\beta})^* = \rho^{\beta\alpha}, \tag{5}$$

is an invariant property and, since $\rho^{\alpha\beta}$ is nonsingular, we may introduce an intrinsic covariant metric tensor by

$$\rho^{\alpha\mu}\rho_{\alpha\gamma} = \delta_{\gamma}^{\mu}, \quad \rho^{\alpha\mu}\rho_{\beta\mu} = \delta_{\beta}^{\alpha}. \tag{6}$$

We use these Hermitian tensors to raise and lower indices. Unitary transformations can now alternatively be defined as those transformations which preserve the special form of the metric tensors as co- and contravariant "unities." Clearly, the notation can be simplified conceptually if only unitary transformations are considered, but the greater generality is needed in later sections and is, in any case, a useful investment for future research.

We base our physical consideration on a Lagrangian as a function of classical fields and their derivatives $\mathfrak{L}(\psi_{\alpha}; \psi_{\alpha,n})$, where the Greek indices $\alpha = 1, 2, 3$ and Latin indices $n = 0, 1, 2, 3$ are the space–time indices.[2] The invariance under (2) and (3) will hold only if the coefficients in the intrinsic transformations are independent of the space–time coordinates. According to an analysis of the concept of a local field theory by Yang and Mills,[3] one should demand that the theory remain invariant even with space–time dependent transformations S_{α}^{β}, and this requires a generalization of the derivatives of the field functions. This generalized derivative is called a covariant derivative, since no confusion with the familiar concept in Riemannian geometry and in general relativity can arise; in fact, the general affine connection is just a direct sum[3] of the affinities of space–time, the intrinsic space, and, possibly, the spinor space. We introduce the concept of the covariant derivative via the geometrical concept of parallel transfer.[4] We call $\bar{\psi}_{\alpha}(x + dx)$ parallel to $\psi_{\alpha}(x)$ if

$$\bar{\psi}_{\alpha}(x + dx) = \psi_{\alpha}(x) + \Gamma_{\alpha n}^{\lambda}\psi_{\lambda}(x), \tag{7}$$

where $\bar{\psi}_{\alpha}(x + dx)$ is an intrinsic vector at $x + dx$. If scalars are to remain invariant under parallel transfer, i.e.,

$$\bar{\psi}^{\alpha}(x + dx)\bar{\psi}_{\alpha}(x + dx) = \psi^{\alpha}(x)\psi_{\alpha}(x), \tag{8}$$

we must have

$$\bar{\psi}^{\alpha}(x + dx) = \psi^{\alpha}(x) - \Gamma_{\kappa n}^{\alpha}\psi^{\kappa}(x). \tag{9}$$

There are obvious analogs to (7)–(9) for vectors with dotted indices. The covariant derivatives are defined by

$$\begin{aligned}\psi_{\alpha|n} &= [\psi_{\alpha}(x + dx) - \bar{\psi}_{\alpha}(x + dx)]/dx^n \\ &= \psi_{\alpha,n} - \Gamma_{\alpha n}^{\lambda}\psi_{\lambda},\end{aligned} \tag{10}$$

with the further assumption that the product rule for differentiation holds here too, which enables us to define covariant derivatives of tensors of higher rank. For the space–time-dependent transformations (2) and (3), the affinities transform like

$$\Gamma_{\alpha n}^{\lambda'} = (S^{-1})_{\nu}^{\lambda}S_{\alpha}^{\mu}\Gamma_{\mu n}^{\nu} + (S^{-1})_{\nu}^{\lambda}S_{\alpha,n}^{\nu} \tag{11}$$

and can, therefore, be transformed to zero at one point, but unitary transformations do not suffice for this purpose. The affinities can be transformed away everywhere if the integrability conditions for $S_{\alpha,n}^{\nu}$ in (11) are satisfied, i.e., if the intrinsic curvature tensor

$$R_{\nu mn}^{\mu} = -\Gamma_{\nu m,n}^{\mu} + \Gamma_{\nu n,m}^{\mu} + \Gamma_{\kappa m}^{\mu}\Gamma_{\nu n}^{\kappa} - \Gamma_{\kappa n}^{\mu}\Gamma_{\nu m}^{\kappa} \tag{12}$$

vanishes. The tensor character of (12) follows from

$$\psi_{\alpha|mn} - \psi_{\alpha|nm} = R_{\alpha mn}^{\beta}\psi_{\beta}. \tag{13}$$

The last equation holds only in flat space–time, since a covariant derivative with respect to a space–time index occurs. We do not record the additional terms needed for a curved space–time, since the complete expression would require a specification of the behavior of ψ_{α} under general transformations of space–time.

It can easily be shown that the intrinsic metric tensor $\rho_{\alpha\beta}$ can always be transformed to the standard form *everywhere*, if we remain within the framework of classical field theory. This can be shown by Taylor expansions of $\rho_{\alpha\beta}$ and a nonunitary S_{α}^{β} around an arbitrary point of space–time or, alternatively, by a diagonalization of $\rho_{\alpha\beta}$ with the help of a unitary S_{α}^{β} and a subsequent reduction of the space–time dependent diagonal elements to unity by appropriate scale transformations. We mention the second alternative for later reference, and now assume that the metric tensor is equal to the unit matrix everywhere. It follows that the covariant derivative of the metric tensor is proportional to the Hermitian part of the affinity, distinguished by a bar under the indices

$$\rho_{\alpha\beta|n} = -\rho_{\alpha\mu}\Gamma_{\beta n}^{\mu} - \rho_{\dot{\nu}\beta}\Gamma_{\dot{\alpha}n}^{\dot{\nu}} = -2\Gamma_{\underline{\alpha\beta}n}. \tag{14}$$

Thus, the latter must be an intrinsic tensor under unitary transformations. This can be verified directly

from (11), using the invariance and coordinate independence of $\rho_{\dot\alpha\beta}$ and the expression

$$(S^{-1})^{\beta}_{\alpha} = \rho_{\alpha\dot\mu}\rho^{\beta\dot\nu}S^{\dot\mu}_{\dot\nu}, \tag{15}$$

which is valid for unitary transformations. Another fundamental intrinsic tensor or, rather, tensor density is needed later. The alternating symbol $\epsilon^{\alpha\beta\gamma}$ with vanishing covariant derivatives is numerically invariant and we may assume that its values are 0 or ± 1. Covariant derivatives of densities are defined as usual.

We return now to the question of transforming $\rho_{\dot\alpha\beta}$ to standard form. With the help of unitary transformations, we can only diagonalize $\rho_{\dot\alpha\beta}$ with space–time-dependent diagonal elements. To complete the reduction, we can use nonunitary scale transformations, or we can make a suitable space–time coordinate transformation, using general covariance of the theory. In the second case, we remain within the group of unitary transformations, but, in either case, the Hermitian part of the affinities will be modified. We conclude with a comment on general covariance. It is generally agreed[5] that this principle does not have the character of a fundamental law of nature, since almost any theory can be expressed in a generally covariant form and whether or not the generality is preferred is only a matter of convenience or practical necessity. But the analysis by Yang and Mills[6] presents an important argument in favor of general covariance, since the assumption of a *special* frame of reference for *all* space–time points is difficult to reconcile with the local character of a theory.

II. PARALLEL TRANSLATION AND CONSERVATION LAWS[7]

According to (8), any intrinsic scalar is left invariant if all its parts are parallel-translated an infinitesimal distance dx^n. This invariance holds for an arbitary set of affinities. If the intrinsic symmetry group expresses a fundamental symmetry of physics, any proposed Lagrangian must be an intrinsic scalar and thus must satisfy

$$\bar{\mathfrak{L}}(x^n + dx^n) - \mathfrak{L}(x^n) = 0. \tag{16}$$

We show that this invariance gives rise to the same conservation laws as would be obtained in the usual way by an application of Noether's law. Since parallel transfer has only been defined for tensors and not for their ordinary derivatives, it is imperative that the Lagrangian be expressed in a manifestly covariant form. In particular, the anti-Hermitian part of the affinities and their derivatives may appear in the Lagrangian only in the form of the covariant derivatives and in the curvature tensor. However, it will be convenient to consider the whole affinities, rather than their Hermitian and anti-Hermitian parts, as the fundamental elements of the theory and to replace the Hermitian part, wherever it occurs separately, by the covariant derivative of the metric tensor. Writing then[8]

$$\mathfrak{L} = \mathfrak{L}(\psi_\alpha;\ \psi_{\alpha|n};\ \rho^{\dot\alpha\beta};\ \rho^{\dot\alpha\beta}_{|n};\ R^{\alpha}_{\beta mn}), \tag{17}$$

the procedure is to express the invariance of $\mathfrak{L}$ in the form

$$\bar\delta\mathfrak{L} = C^{\beta}_{\alpha}\Gamma^{\alpha}_{\beta n}\,dx^n + C^{\dot\beta}_{\dot\alpha}\Gamma^{\dot\alpha}_{\dot\beta n}\,dx^n \tag{18}$$

and to conclude on the basis of the arbitrariness of the affinities and of the displacements dx^n that

$$C^{\beta}_{\alpha} = 0, \quad C^{\dot\beta}_{\dot\alpha} = 0, \tag{19}$$

since the affinities in (18) are not related to those which are contained in the covariant derivatives and in the curvature tensor in (17). We get now

$$\frac{\partial\mathfrak{L}}{\partial\psi_\beta}\psi_\alpha + \frac{\partial\mathfrak{L}}{\partial\psi_{\beta|m}}\psi_{\alpha|m} - \frac{\partial\mathfrak{L}}{\partial\rho^{\dot\kappa\alpha}}\rho^{\dot\kappa\beta} - \frac{\partial\mathfrak{L}}{\partial\rho^{\dot\kappa\alpha}_{|m}}\rho^{\dot\kappa\beta}_{|m} + \frac{\partial\mathfrak{L}}{\partial R^{\kappa}_{\beta mn}}R^{\kappa}_{\alpha mn} - \frac{\partial\mathfrak{L}}{\partial R^{\alpha}_{\lambda mn}}R^{\beta}_{\lambda mn} = 0, \tag{20}$$

but we observe that $\partial\mathfrak{L}/\partial\psi_\beta$ stands here for the partial derivative with $\psi_{\beta|m}$ kept constant. The derivative $\partial\mathfrak{L}/\partial\psi_\beta$ with $\psi_{\beta,m}$ kept fixed is given by

$$\frac{\partial\mathfrak{L}}{\partial\psi_\beta} - \frac{\partial\mathfrak{L}}{\partial\psi_{\mu|n}}\Gamma^{\beta}_{\mu n}. \tag{21}$$

Adding and subtracting $(\partial\mathfrak{L}/\partial\psi_{\beta|n})_{,n}$, we obtain, for the derivative $\partial\mathfrak{L}/\partial\psi_\beta$ in (20),

$$\frac{\delta\mathfrak{L}}{\delta\psi_\beta} + \left(\frac{\partial\mathfrak{L}}{\partial\psi_{\beta|n}}\right)_{|n}. \tag{22}$$

This shows, incidently, that the Euler–Lagrange derivative is a covariant expression. We perform the same transformation for the derivative with respect to $\rho^{\dot\kappa\beta}$ and get, instead of (20),

$$\frac{\delta\mathfrak{L}}{\delta\psi_\beta}\psi_\alpha - \frac{\delta\mathfrak{L}}{\delta\rho^{\dot\kappa\alpha}}\rho^{\dot\kappa\beta} + \frac{\partial\mathfrak{L}}{\partial R^{\kappa}_{\beta mn}}R^{\kappa}_{\alpha mn} - \frac{\partial\mathfrak{L}}{\partial R^{\alpha}_{\lambda mn}}R^{\beta}_{\lambda mn} + \left(\frac{\partial\mathfrak{L}}{\partial\psi_{\beta|m}}\psi_\alpha - \frac{\partial\mathfrak{L}}{\partial\rho^{\dot\kappa\alpha}_{|m}}\rho_{\dot\kappa\beta}\right)_{|m} = 0. \tag{23}$$

The term containing the curvature tensor can be transformed, by substitution of (12),

$$\frac{\partial\mathfrak{L}}{\partial R^{\kappa}_{\beta mn}}R^{\kappa}_{\alpha mn} - \frac{\partial\mathfrak{L}}{\partial R^{\alpha}_{\lambda mn}}R^{\beta}_{\lambda mn} = 2\left(\frac{\partial\mathfrak{L}}{\partial R^{\alpha}_{\beta mn}}\right)_{|mn}. \tag{24}$$

If the Euler–Lagrange derivatives with respect to

ψ_β and $\rho^{\dot\kappa\beta}$ vanish, we obtain a covariant conservation law

$$J^{\beta\ m}_{\ \alpha\ |m} = 0, \tag{25}$$

with

$$J^{\beta\ m}_{\ \alpha} = \frac{\partial \mathfrak{L}}{\partial \psi_{\beta|m}} \psi_\alpha - \frac{\partial \mathfrak{L}}{\partial \rho^{\dot\kappa\alpha}_{\ |m}} \rho^{\dot\kappa\beta} + 2\left(\frac{\partial \mathfrak{L}}{\partial R^{\alpha}_{\beta nm/|n}}\right). \tag{26}$$

The current vector, which is an intrinsic tensor of the second rank, will not be real; but if the Lagrangian is real, the complex conjugate of (26) is also conserved and we can construct a real current vector, as well as a purely imaginary one. Before discussing this point further, we derive another identity and a second form of the conservation law. We have already mentioned that, in a manifestly covariant theory, the affinities can appear in the Lagrangian only implicitly in the form of the covariant derivatives of tensors and in the curvature tensor. If equations of no higher order than the second are desired, the derivatives of the affinities appear only in the curvature tensor. Thus,

$$\frac{\partial \mathfrak{L}}{\partial \Gamma^{\alpha}_{\beta m}} = -\frac{\partial \mathfrak{L}}{\partial \psi_{\beta|m}} \psi_\alpha + \frac{\partial \mathfrak{L}}{\partial \rho^{\dot\kappa\alpha}_{\ |m}} \rho^{\dot\kappa\beta} + 2\left(\frac{\partial \mathfrak{L}}{\partial R^{\alpha}_{\lambda mn}} \Gamma^{\beta}_{\lambda n} - \frac{\partial \mathfrak{L}}{\partial R^{\kappa\ n}_{\beta mn}} \Gamma^{\kappa}_{\alpha n}\right) \tag{27}$$

and

$$\frac{\partial \mathfrak{L}}{\partial \Gamma^{\alpha}_{\beta m,n}} = -2\frac{\partial \mathfrak{L}}{\partial R^{\alpha}_{\beta mn}}. \tag{28}$$

The Euler–Lagrange derivatives with respect to the affine connections are

$$\frac{\delta \mathfrak{L}}{\delta \Gamma^{\alpha}_{\beta m}} = -J^{\beta}_{\alpha m}. \tag{29}$$

We find from (28) that

$$\frac{\partial \mathfrak{L}}{\partial \Gamma^{\alpha}_{\beta m,n}} = -\frac{\partial \mathfrak{L}}{\partial \Gamma^{\alpha}_{\beta n,m}} \tag{30}$$

and we derive from (23), (24), and (29) an identity between the various Euler–Lagrange derivatives,

$$\frac{\delta \mathfrak{L}}{\delta \psi_\beta} \psi_\alpha - \frac{\delta \mathfrak{L}}{\delta \rho^{\dot\kappa\alpha}} \rho^{\dot\kappa\beta} - \left(\frac{\delta \mathfrak{L}}{\delta \Gamma^{\alpha}_{\beta m/|m}}\right) = 0, \tag{31}$$

which is a consequence of the general covariance of the theory. Of course, (31) does not go beyond (25); since it is customary to consider the affinities as field variables [thus setting (29) equal to zero], the conservation law (25) is not very meaningful and (31) says only that the metric *must* also be a field variable, i.e., its Euler–Lagrange derivative must vanish.

Omitting now the Euler–Lagrange derivatives, a true conservation law follows from (26), (28), and (30):

$$j^{\beta\ m}_{\ \alpha\ ,m} = 0, \tag{32}$$

$$j^{\beta\ m}_{\ \alpha} = 2\left(\frac{\partial \mathfrak{L}}{\partial R^{\alpha}_{\beta mn/,n}}\right) = \left(\frac{\partial \mathfrak{L}}{\partial \Gamma^{\alpha}_{\beta n,m/,n}}\right), \tag{33}$$

where the currents are defined by

$$j^{\beta\ m}_{\ \alpha} = \frac{\partial \mathfrak{L}}{\partial \psi_{\beta|m}} \psi_\alpha - \frac{\partial \mathfrak{L}}{\partial \rho^{\dot\kappa\alpha}_{\ |m}} \rho^{\dot\kappa\beta} + 2\left(\frac{\partial \mathfrak{L}}{\partial R^{\mu}_{\beta mn}} \Gamma^{\mu}_{\alpha n} - \frac{\partial \mathfrak{L}}{\partial R^{\alpha}_{\nu mn}} \Gamma^{\beta}_{\nu n}\right). \tag{34}$$

Real and purely imaginary combinations can be formed as mentioned earlier. The expression (34) is not covariant in the intrinsic space, i.e., not gauge invariant, and this is true for the real as well as the imaginary combination. However, the integral of $j^{\beta 0}_{\alpha}$ over a spacelike hypersurface is gauge invariant, as the first expression (33) shows.[9] This is quite sufficient for the interpretation of the generalized charges, but it is interesting to observe that (34) is modified under a general transformation (2) by the addition of a term

$$\frac{\partial \mathfrak{L}}{\partial R^{\mu}_{\nu mn}} [(S^{-1})^{\beta}_{\nu} S^{\mu}_{\alpha}]_{,n}, \tag{35}$$

and that this term does not occur if $\partial \mathfrak{L}/\partial R^{\mu}_{\nu mn}$ is proportional to δ^{ν}_{μ}, which can be true only for the real part of the said derivative. Tentatively assuming then that

$$\frac{\partial \mathfrak{L}}{\partial R^{\mu}_{\nu mn}} + \frac{\partial \mathfrak{L}}{\partial R_{\nu\ mn}^{\ \mu}} \sim \delta^{\nu}_{\mu}, \tag{36}$$

the Hermitian part of the current vector will be gauge invariant under unitary transformations, i.e., it will be an intrinsic tensor. The anti-Hermitian part of the current vector will have the same transformation property as in the theory of Yang and Mills. Equation (36) is, of course, not necessary for the interpretation of the conservation laws.

III. THE AFFINE COMPLEX

The geometrical concepts used for the intrinsic space are familiar from the application of differential geometry to space–time, as used in general relativity. We will discuss here briefly the extension of the geometry of the intrinsic space, when space–time is curved and the space–time affinities are needed. In particular, we show that the derivation of the conservation laws given in the previous section can be extended to give conservation of energy and momentum. Such a derivation should be contrasted with the usual appeal to the invariance of the Lagrangian under certain symmetry operations, one of which invariably is space and

time translations and rotations in the general sense in space–time, whereas any other symmetry operations are coincidental. In the following derivation, conservation laws follow from invariance under parallel translation in space–time and the multitude of conserved quantities has its origin in the fact that the affine complex is a direct sum of mutually independent affinities, operating in different spaces, like space–time on the one hand and the unitary intrinsic space on the other hand.

The Lagrangian (17) is now assumed to be a space–time scalar density, rather than a scalar, and we will assume that $\mathfrak{L}$ now also depends on the metric g^{mn} and the space–time curvature tensor through the contracted form. We omit a dependence on $g^{mn}{}_{|p}$, since these covariant derivatives vanish in Riemannian geometry. However, the ordinary derivatives of g^{mn} enter in the space–time affinities, which cannot be considered as independent variables, since parallel translation cannot be defined for them. Therefore, we begin our considerations by assuming independence of the space–time affinities and the metric, as in the affine theories and in the approach by Palatini.[10] The link with the metric theory will be established later.

Because $\mathfrak{L}$ is a scalar density, parallel translations lead to

$$\bar{\mathfrak{L}}(x + dx) - \mathfrak{L}(x) = \Gamma^n_{nm}\mathfrak{L}\, dx^m \tag{37}$$

instead of (16). We now assume again that (37) holds for parallel translations with completely arbitrary affine connections and conclude that

$$-\frac{\partial\mathfrak{L}}{\partial g^{ms}}g^{ns} - \frac{\partial\mathfrak{L}}{\partial g^{sm}}g^{sn} + \frac{\partial\mathfrak{L}}{\partial R_{ns}}R_{ms} + \frac{\partial\mathfrak{L}}{\partial R_{sn}}R_{sm} + \frac{\partial\mathfrak{L}}{\partial\psi_{\alpha|n}}\psi_{\alpha|m} + \frac{\partial\mathfrak{L}}{\partial\rho^{\alpha\beta}_{|n}}\rho^{\alpha\beta}_{|m} + 2\frac{\partial\mathfrak{L}}{\partial R^{\alpha}_{\beta ns}}R^{\alpha}_{\beta ms} - \mathfrak{L}\delta^n_m = 0. \tag{38}$$

Besides (38), there is still another identity associated with the invariance under parallel translation. Clearly, $\mathfrak{L}$ cannot depend explicitly on the coordinates x^n. This can be derived formally from (37) by assuming such a dependence and conclude from the fact that dx^n is also completely arbitrary that, in fact,

$$\left(\frac{\partial\mathfrak{L}}{\partial x^m}\right)_{\text{explicit}} = 0. \tag{39}$$

We rewrite all ordinary derivatives in terms of covariant derivatives. Here we need (13) and

$$\rho^{\alpha\beta}_{|mn} - \rho^{\alpha\beta}_{|nm} = -R^{\alpha}_{\mu mn}\rho^{\mu\beta} - R^{\beta}_{\nu mn}\rho^{\alpha\nu} \tag{40}$$

and the Bianchi identity for the intrinsic curvature tensor, which is obtained just as in Riemannian geometry by covariant differentiations of (13):

$$R^{\alpha}_{\beta mn|p} + R^{\alpha}_{\beta np|m} + R^{\alpha}_{\beta pm|n} = 0. \tag{41}$$

With the covariant energy–momentum tensor

$$T^n_m = \left(\mathfrak{L} - \frac{\partial\mathfrak{L}}{\partial R_{i\kappa}}R_{i\kappa}\right)\delta^n_m - 2\frac{\partial\mathfrak{L}}{\partial R^{\alpha}_{\beta nq}}R^{\alpha}_{\beta mq} - \frac{\partial\mathfrak{L}}{\partial\psi_{\alpha|n}}\psi_{\alpha|m} - \frac{\partial\mathfrak{L}}{\partial\rho^{\alpha\beta}_{|n}}\rho^{\alpha\beta}_{|m}, \tag{42}$$

we can write the identity (39) as

$$T^n_{m|n} - R^{\alpha}_{\beta nm}\frac{\delta\mathfrak{L}}{\delta\Gamma^{\alpha}_{\beta n}} + \left(\frac{\partial\mathfrak{L}}{\partial R_{i\kappa}}\right)_{|m}R_{i\kappa} - \Gamma^{\beta}_{\alpha m}\left(\frac{\delta\mathfrak{L}}{\delta\psi_{\alpha}}\psi_{\beta} - \frac{\delta\mathfrak{L}}{\delta\rho^{\mu\beta}}\rho^{\mu\alpha}\right) = 0, \tag{43}$$

where use has been made of (38). On physical grounds, it is obvious that we have to demand that the Euler–Lagrange derivatives with respect to the intrinsic affinities vanish. Since the latter are not uniquely determined by the intrinsic metric, they would be arbitrary and the physical system under consideration would not be closed. The field equations for the space–time geometry are more difficult because of the dependence of the affinities on the metric. We need the total Euler–Lagrange derivative with respect to g^{mn}. A somewhat lengthy calculation gives[11]

$$\frac{\delta\mathfrak{L}}{\delta g^{mn}} \equiv \frac{\partial\mathfrak{L}}{\partial g^{mn}} + \frac{\partial\mathfrak{L}}{\partial g^{nm}} + g_{m\kappa}\left(\frac{\delta\mathfrak{L}}{\delta\Gamma^{n}_{\kappa l}}\right)_{|l} + g_{n\kappa}\left(\frac{\delta\mathfrak{L}}{\delta\Gamma^{m}_{\kappa l}}\right)_{|l} - g_{\kappa n}g_{lm}g^{ij}\left(\frac{\delta\mathfrak{L}}{\delta\Gamma^{i}_{\kappa l}}\right)_{|j} = 0. \tag{44}$$

This is, in essence, the fundamental equation of Palatini. If the Lagrangian depends only on scalar fields, apart from the space–time metric, one finds that

$$\frac{\delta\mathfrak{L}}{\delta\Gamma^{p}_{mn}} = \left(\frac{\partial\mathfrak{L}}{\partial R_{mn}}_{|p}\right) - \left(\frac{\partial\mathfrak{L}}{\partial R_{mq}}_{|q}\right)\delta^n_p, \tag{45}$$

and this would vanish for the usual Lagrangian of general relativity. More generally, the last three terms in (44) cannot contribute, as can be seen by recalling the derivation of (44) from an action principle. Multiplying (44) with δg^{mn}, whose covariant derivatives vanish, and integrating over a space–time domain, assuming that δg^{mn} vanish at the boundary, each of these terms becomes a divergence of a contravariant vector density. Using Gauss' law, they are seen to cancel. The identity (38) can now be written with (42)

and (44) as

$$-\frac{\delta\mathfrak{L}}{\delta g^{mn}} - T_{nm} - \frac{\partial\mathfrak{L}}{\partial R_{i\kappa}} R_{i\kappa} g_{mn} + \frac{\partial\mathfrak{L}}{\partial R_{pq}} (R_{mq} g_{pn} + R_{nq} g_{pm}) = 0, \quad (46)$$

which is equivalent to Einstein's equation if the appropriate Lagrangian is inserted. It should be noted that many of the equations in this section are not correct if the Lagrangian contains space–time spinor or tensor fields. It should also be pointed out that the relation

$$\frac{\partial\mathfrak{L}}{\partial\Gamma^{p}_{mn,q}} = -\frac{\partial\mathfrak{L}}{\partial R_{mn}} \delta^{q}_{p} + \frac{\partial\mathfrak{L}}{\partial R_{mq}} \delta^{n}_{p} \quad (47)$$

can be used to derive an ordinary divergence conservation law of energy and momentum.

We conclude this section with a few general remarks. Looking at the formalism from the point of view of unitary symmetry, the essential extension consists in the addition of a Hermitian tensor and a space–time vector to the affine connection. If such a field exists, it could carry a mass field without destroying gauge invariance. The mass terms in the Lagrangian would have the form

$$\tfrac{1}{2} M^2 \rho_{\{n}^{\alpha\beta} \rho_{\{m}^{\mu\nu} g^{mn} \rho_{\alpha\nu} \rho_{\beta\mu}, \quad (48)$$

which reduces in a standard reference frame in the intrinsic space to a mass term for the Hermitian part of the affinities $\Gamma^{\alpha}_{\beta n}$. The characteristic feature of such fields is their coupling with other intrinsic tensor fields, viz., bilinear in the fields and additive to the ordinary derivative of the tensor field. It is suggestive to call this coupling the general form of a *minimal coupling*. Finally, we would like to say that the assumption of a direct sum of the affinities may be too narrow and other assumptions, like an affine complex which is not separable in space–time spin and intrinsic spin, should not be disregarded with our present knowledge of elementary particles. We hope to return to this possibility in a forthcoming paper.

ACKNOWLEDGMENT

Both authors thank the Graduate School of the George Washington University for financial support through an NSF institutional grant. Part of the work was carried out while one author (O. B.) enjoyed the hospitality of the Aspen Center of Physics in Aspen, Colorado.

* Present address: Vitro Laboratories, Silver Spring, Maryland.

[1] B. L. van der Waerden, Nachr. Akad. Wiss. Göttingen, Math.-Physik. Kl. 100 (1929); see also W. L. Bade and H. Jehle, Rev. Mod. Phys. **25**, 714 (1953).

[2] To save space, we consider here only one field ψ, which is a space–time scalar and an intrinsic vector, although the most important family of particles would require a space–time spinor field with a pair of intrinsic indices $\dot{\alpha}\beta$. Except for Sec. III, the space–time will be assumed to be flat.

[3] C. N. Yang and R. L. Mills, Phys. Rev. **96**, 191 (1954); later papers on related formalism are S. L. Glashow and M. Gell-Mann, Ann Phys. (N.Y.) **15**, 437 (1961); H. G. Loos, J. Math. Phys. **8**, 2114 (1967); R. P. Treat, Nuovo Cimento **50**, 871 (1967); and C. A. Uzes, J. Math. Phys. **10**, 1885 (1969).

[4] See for instance S. Helgason, *Differential Geometry and Symmetric Spaces* (Academic Press Inc., New York, 1962).

[5] See, for instance P. Havas, *Delaware Seminar in the Foundations of Physics* (Springer-Verlag, Berlin, 1967), pp. 128 and 9.

[6] See Ref. 3.

[7] A preliminary report was given in Bull. Am. Phys. Soc. (11) **13**, 1379 (1968). Further details and a possible interpretation can be found in the Ph.D. thesis of J. L. Alderman (George Washington University, 1968).

[8] We omit here the Hermitian conjugate fields $\psi_{\dot{\alpha}}$ and $\rho^{\beta\dot{\alpha}}$ which must appear in a real-valued Lagrangian.

[9] See Ref. 3.

[10] References and a critical summary can be found in W. Pauli, *Theory of Relativity* (Pergamon Press Ltd., London, 1958), particularly in Footnote 8 of the Appendix, pp. 214 and 5. See also E. Schrödinger, *Space-Time Structure* (Cambridge University Press, Cambridge, England, 1950).

[11] The analogous equation for the space–time spinor case will be given in a forthcoming paper.

Lorentz-Invariant Gravitational Perturbations and the Evaluation of Generalized Green's Functions*

HUBERT GOENNER†
Department of Physics, Temple University, Philadelphia, Pa. 19122

(Received 4 September 1969)

The Lorentz-invariant perturbation theory of classical nonlinear field theories, notably the fast-motion approximation for a system of gravitationally interacting particles, is used to illustrate the appearance of generalized Green's functions (GGF's) in physics. The (retarded) GGF's are defined as the convolution of the usual (retarded) Green's function with certain other linear functionals. A manifestly Lorentz-invariant technique for the evaluation of the convolution integral is described and applied to two important classes of GGF's. Time-symmetric GGF's are also discussed briefly. The technique exhibited provides the tools for the calculation of the higher-order approximations in a Lorentz-invariant perturbational approach to the classical n-body problem.

1. INTRODUCTION

In order to introduce the concept of a generalized Green's function, by way of example we start with the wave equation in flat space–time[1] R_4,

$$\Box\phi(x) = f(x). \tag{1.1}$$

The retarded solution of Eq. (1.1) is obtained as the convolution of the source term $f(x)$ with the retarded Green's function $D^{\mathrm{R}}(x)$:

$$\phi(x) = D^{\mathrm{R}}(x) * f(x) = \frac{1}{4\pi}\int d^4y\, D^{\mathrm{R}}(x-y)f(y). \tag{1.2}$$

$f(x)$ is usually taken to be a locally summable function. However, in many applications one allows $f(x)$ to be a generalized function itself.[2] For example, if $f(x)$ is to describe a line source generated by a moving particle with world line $a_\mu = a_\mu(s)$,

$$f(x) = \int_{-\infty}^{+\infty} ds w(s)\delta^4(x - a(s)), \tag{1.3}$$

then Eq. (1.2) is rewritten to give

$$\phi(x) = \int_{-\infty}^{+\infty} ds w(s) D^{\mathrm{R}}(x) * \delta^4(x - a(s)). \tag{1.4}$$

By virtue of the properties of the Dirac δ, $D^{\mathrm{R}}(x) * \delta^4(x-a) = D^{\mathrm{R}}(x-a)$, Eq. (1.4) reduces to the familiar retarded solution of Eqs. (1.1) and (1.3) (Lienard–Wiechert potential).

There are physical situations where the generalized function of Eq. (1.3) does not have pointlike support.[3] For example, $\delta^4(x-a)$ might be replaced by $D^{\mathrm{R}}(x-a)$ itself. By using the relation[4]

$$D^{\mathrm{R}}(x) * D^{\mathrm{R}}(x-a) = \tfrac{1}{2}\theta(x-a), \tag{1.5}$$

which will be derived in Sec. 4, Eq. (1.4) then goes over into

$$\phi(x) = \tfrac{1}{2}\int_{-\infty}^{+\infty} ds w(s)\theta(x - a(s)) = \tfrac{1}{2}\int_{-\infty}^{s_{\mathrm{R}}} ds w(s). \tag{1.6}$$

Indeed, nontrivial examples are provided by a Lorentz-invariant perturbation approach to the description of the motion of interacting particles in classical nonlinear field theories. Schieve[5] has discussed a method of successive approximation for the solution of the field equations for the coupled electromagnetic and charge-symmetric scalar meson fields in which the GGF's appear implicitly.[6] Similarly, Einstein's equations of gravitation are reduced to wave equations by a Lorentz-invariant perturbation theory, with only the source term $f(x)$ changing in successive orders of approximation.[7] In the linear approximation, $f(x)$ takes the form (1.3). In second approximation, however, in addition to such terms the inhomogeneity of the wave equation contains expressions of the structure

$$\sum_{kj}^{N}\int_{-\infty}^{+\infty} ds_k w(s_k)\int_{-\infty}^{+\infty} ds_j \bar{w}(s_j) O\left(\frac{\partial}{\partial a_k^\alpha}, \frac{\partial}{\partial a_j^\beta}\right) \times D^{\mathrm{R}}(x - a(s_k)) D^{\mathrm{R}}(x - a(s_j)), \tag{1.7}$$

where $O(\partial/\partial a_k^\alpha, \partial/\partial a_j^\beta)$ stands for a linear differential operator and the sum runs over the various particles generating the gravitational field. As a formal solution of the wave equation with source (1.7), one obtains

$$\phi(x) = \sum_{k,j}\int_{-\infty}^{+\infty} ds_k w(s_k)\int_{-\infty}^{+\infty} ds_j \bar{w}(s_j) O\left(\frac{\partial}{\partial a_k^\alpha}, \frac{\partial}{\partial a_j^\beta}\right) \times D^{\mathrm{R}}(x) * D^{\mathrm{R}}(x - a(s_k)) D^{\mathrm{R}}(x - a(s_j)). \tag{1.8}$$

With the convolution $D^{\mathrm{R}}(x) * D^{\mathrm{R}}(x-a)\cdot D^{\mathrm{R}}(x-b)$, one of the objects appears to which the name *generalized Green's function* (GGF) is assigned.[7]

In Sec. 2 a definition of such GGF's is given as well as a description of their mathematical nature. Section 3 develops a Lorentz-invariant technique for evaluation of the convolution integral of the GGF's which, in Sec. 4, is applied to two special classes of GGF's. In a final Sec. 5, time-symmetric GGF's are discussed briefly.

2. THE GENERALIZED GREEN'S FUNCTION

A. Definition

With retarded GGF $G^{\mathrm{R}}(x)$, we define the convolution of the retarded Green's function with another *generalized function* $F(x) \in \mathfrak{F}$:

$$G^{\mathrm{R}}(x) = D^{\mathrm{R}}(x) * F(x). \tag{2.1}$$

As indicated, $F(x)$ will be selected from a special set $\mathfrak{F}$ of generalized functions,the construction of which (by the help of Dirac's δ and the retarded Green's function) will be described below. (Fx) has to be well behaved so that the rhs of Eq. (2.1) exists (Sec. 2B).

The definitions of the advanced and symmetric GGF's $G^{\mathrm{A}}(x)$ and $G^{\mathrm{S}}(x)$ are obtained from Eq. (2.1) by replacing $D^{\mathrm{R}}(x)$ by $D^{\mathrm{A}}(x)$ and $D^{\mathrm{S}}(x)$, respectively.[8] In general, the GGF's are no longer elementary solutions of the wave equation, but satisfy

$$\Box G^{\mathrm{R}}(x) = F(x), \quad F(x) \in \mathfrak{F} \tag{2.2}$$

B. Construction of the set $\mathfrak{F}$

First we consider the infinite-dimensional vector space $\mathfrak{V}$ over the ring $\mathfrak{R}$ of C^∞-functions, with bounded support in R_4, whose basis elements are the functionals

$$d_0 = \delta^4(x), \quad d_n^{\mathrm{R}}(x, a_1, \cdots, a_n) = \prod_{k=1}^{n} D^{\mathrm{R}}(x - a_k(s_k)), \quad n = 1, 2, 3, \cdots, \tag{2.3}$$

with $x_\alpha - a_{k\alpha}$, $k = 1, 2, \cdots, n$, in R_4. The general element of $\mathfrak{V}$ is

$$B_N(x, f_1, \cdots, f_N) = \sum_{j=0}^{N} f_j(x, a_1, \cdots, a_j) d_j^{\mathrm{R}}(x, a_1, \cdots, a_j)$$

with $f_j \in \mathfrak{R}$. The elements d_n^{R} of (2.3) have a two-sided nature, as well as the test functions $\phi(x, s_1, \cdots, s_n)$ by which the corresponding linear functionals are defined.

(1) For fixed $x \in R_4$, d_n^{R} is a generalized function on the space $\mathfrak{D}_n^*(s_1, \cdots, s_n)$ of all C^∞-test functions $\phi(x, s_1, \cdots, s_n)$ with bounded support in the product space $R_n = R \times R \times \cdots \times R$ (n factors, R is the real line). d_n^{R} itself has pointlike support in R_n.

(2) For fixed variables $s_1, \cdots, s_n$ (or fixed $a_{k\alpha}$, $k = 1, \cdots, n$), d_n^{R} is a Lorentz-invariant generalized function in R_4, its support being the intersection of n half light cones. It is defined on the space $\overline{\mathfrak{D}}(x)$ of all C^∞-test functions with compact support in R_4

In general, then, the linear functional d_n^{R} is defined on the product space $\mathfrak{D}_n = \mathfrak{D}_n^* \times \overline{\mathfrak{D}}$. The multiplication with a C^∞-function $f(x, s_1, \cdots, s_n)$ on $R_4 \times R_n$ is well defined. The same holds for the direct product (with respect to the variables s_k) $d_n^{\mathrm{R}}(x, a_1, \cdots, a_n) \times d_m^{\mathrm{R}}(x, b_1, \cdots, b_m)$ on

$$\mathfrak{D}_{n,m} = (\mathfrak{D}_n^* \times \mathfrak{D}_m^*) \times \overline{\mathfrak{D}} = \mathfrak{D}_{n+m}^* \times \overline{\mathfrak{D}}.$$

Next, one introduces the convolution product in space–time R_4 as a multiplication for the elements B_N of $\mathfrak{V}$: $B_N(x, f_1, \cdots, f_N) * B_M(x, g_1, \cdots, g_M)$. One notes that $B_0(x, f_0 = 1)$ serves as (left and right) unit element with regard to this multiplication. In order that the convolution

$$d_n^{\mathrm{R}}(x, a_1, \cdots, a_n) * d_m^{\mathrm{R}}(x, b_1, \cdots, b_m)$$

is defined on R_4, the supports of d_n^{R} and d_m^{R} have to satisfy one of the following sufficient requirements[9]: (a) Either d_n^{R} or d_m^{R} has bounded support, (b) the supports of both are on the same half of the null cone through x. The second condition is satisfied by definition of the d_n^{R}.

Obviously, the introduction of the convolution product for the elements of $\mathfrak{V}$ does *not* construct an algebra over $\mathfrak{V}$. In general,

$$B_N(x, f_1, \cdots, f_N) * B_M(x, g_1, \cdots, g_M) \neq B_L(x, b_1, \cdots, b_L).$$

For generalized functions built up from elements

$$d_n^{\mathrm{S}} = \prod_{k=1}^{n} D^{\mathrm{S}}(x - a_k(j_k)),$$

where $D^{\mathrm{S}}(x)$ is the *symmetric* Green's function (Sec. 5), the condition (b) fails to hold. Moreover, the support of d_n^{S}, being the intersection of n *full* light cones, might be unbounded. The situation is illuminated by the following criteria: $d_2^{\mathrm{S}}(x, a_1, a_2)$ has bounded support if $a_{1\alpha} - a_{2\alpha}$ is timelike; $d_3^{\mathrm{S}}(x, a_1, a_2, a_3)$ has bounded support if the 2-flat spanned by $a_{1\alpha} - a_{2\alpha}$ and $a_{1\alpha} - a_{3\alpha}$ is timelike or null; d_n^{S}, $n \geq 4$, has

bounded support if *none* of the Grammian determinants

$$\Delta(p, ijk) = \begin{vmatrix} (a_p - a_i)^2 & (a_p - a_i, a_p - a_j) & (a_p - a_i, a_p - a_k) \\ (a_p - a_i, a_p - a_j) & (a_p - a_j)^2 & (a_p - a_j, a_p - a_k) \\ (a_p - a_i, a_p - a_k) & (a_p - a_j, a_p - a_k) & (a_p - a_k)^2 \end{vmatrix} \tag{2.4}$$

has rank < 3 ($a_{p\alpha}, a_{i\alpha}, a_{j\alpha}, a_{k\alpha}$ are any four different vectors out of the set $a_{1\alpha}, \cdots, a_{n\alpha}$). The convolution $d_n^S * d_m^S$ may not be defined if the listed conditions are not met.

C. Two Special Classes of GGF's

The first of the special classes of GGF's contains the so-called *generating* GGF's of order n, $g_n(x, a_1, \cdots, a_n)$ of Bertotti and Plebański.[7] They may be obtained by the following recursive construction:

$$g_1(x, a_1) = d_1^R(x, a_1) = D^R(x - a_1), \quad \text{for} \quad n = 1,$$

$$g_{n\kappa}(x, a_1, \cdots, a_n) = D^R(x) * [g_{jp_j}(x, a_1, \cdots, a_j) \times g_{n-jq_j}(x, a_{j+1}, \cdots, a_n)], \quad \text{for} \quad n \geq 2, \tag{2.5}$$

with

$$j = 1, 2, \cdots, J_n,$$

$$J_n = \tfrac{1}{2}n, \quad \text{for } n \text{ even},$$

$$= \tfrac{1}{2}(n-1), \quad \text{for } n \text{ odd},$$

$$P_j = 1, 2, \cdots, m_j; q_j = 1, 2, \cdots, m_{n-j},$$

$$\kappa = 1, 2, \cdots, m_n,$$

and

$$m_n = \sum_{j=1}^{J_{n-q(n)}} m_j m_{n-j} - \tfrac{1}{2}q(n) J m_{J_n}(J m_{J_n} - 1),$$

where

$$q(n) = 1, \quad \text{for } n \text{ even},$$

$$= 0, \quad \text{for } n \text{ odd}.$$

m_j is the number of generating GGF's of order j ($m_1 = 1$). The generating GGF of order j contains $j - 1$ convolutions. Since $m_j = 1$ for $j \leq 3$, we put $g_{j1} = g_j$ for $j \leq 3$.

The second class of GGF's is defined by taking one of the basis vectors of $\mathfrak{B}$ as $F(x)$ of Eq. (2.1):

$$\begin{aligned} D^R(x, a_1, \cdots, a_n \,|\, f) &= D^R(x) * d_n^R(x, a_1, \cdots, a_n) \\ &\quad \times f(x, a_1, \cdots, a_m, C_1, \cdots, C_e) \\ &= \frac{1}{4\pi} \int d^4y D^R(x - y) \\ &\quad \times \prod_{k=1}^{n} D^R(y - a_k) f(y, a_1, \cdots, a_m, C_1, \cdots, C_e), \end{aligned} \tag{2.6}$$

with f of class C^∞. We know that f could carry tensor indices. However, we prefer to work with Lorentz scalars; these must be constructed from $(x - a_i, x - a_j), i \neq j$, $(x - a_j, C_p)$, (C_p, C_q), $|x - a_j|^2$, with $i, j = 1, 2, \cdots, m$ and $p, q = 1, 2, \cdots, e$. The $C_{1\mu}, \cdots, C_{e\mu}$ are constant vectors [for example, $da^\alpha(s)/ds$] which do not depend on the variable of integration in Eq. (2.6).

Both types of generalized functions (2.5) and (2.6) will appear in the solutions of nonlinear classical field equations in the various orders of the Lorentz-invariant perturbation theory mentioned in Sec. 1. The field of the interacting particles is expressed as a functional of the motion of the particles. This functional formally is written as an integral over the particles' world lines, the kernel of which is given by GGF's. Consequently, one would like to evaluate the convolution product [Eqs. (2.1), (2.5), and (2.6)] by actually carrying out the integration over R_4. It is of special interest to know whether the appearing GGF's are *regular* or *singular* generalized functions on the space $\mathfrak{D}_n^*$ of test functions $\phi(x, s_1, \cdots, s_n)$. In the latter case, expressions like Eq. (1.8) will have only *symbolic* meaning. One cannot use them to calculate, for example, the equations of motion of the system of particles without first providing for a suitable regularization.[10] Examples of both kinds will appear among the two classes of GGF's treated in Sec. 4.

3. MANIFESTLY LORENTZ-INVARIANT EVALUATION OF GGF's

Bertotti and Plebański[7] have devised a graphical method in order to easily write down the GGF's appearing in Eq. (2.5). By carrying out the integration in a special coordinate system and subsequently recasting the result into Lorentz-covariant form, they evaluate $D^R(x, a_1, \cdots, a_n \,|\, 1)$ for $n = 2, 3$. This approach seems to be practical only in the evaluation of GGF's of low orders.

A. Tetrad Technique

In order to evaluate the GGF by manifestly Lorentz-covariant operations, we span flat space-time by an in general not orthogonal tetrad $n_\alpha^{(\sigma)}$:

$$y_\alpha - x_\alpha = \kappa n_\alpha^{(1)} + \lambda n_\alpha^{(2)} + \mu n_\alpha^{(3)} + \nu n_\alpha^{(4)}. \tag{3.1}$$

The integration variables introduced instead of y are κ, λ, μ, and ν and the element of 4-volume equals

$$d^4y = [-G]^{\frac{1}{2}} \, d\kappa \, d\lambda \, d\mu \, d\nu, \tag{3.2}$$

with G being the Grammian determinant of the tetrad

$$G = \det(n_\sigma^{(\alpha)} n^{(\beta)\sigma}).$$

In the following, the tetrad legs are picked from among the vectors $R_{k\alpha} = x_\alpha - a_\alpha$ and $C_{k\alpha}$ occurring in Eqs. (2.5) and (2.6). Depending on how many linearly independent *timelike* vectors are available, we shall work with tetrads of 3, 2, or 1 timelike legs. Discussing these cases separately we start with the following.

1. (3, 1) Tetrad

$$n_\alpha^{(A)} = R_{A\alpha}, \quad n_\alpha^{(4)} = \epsilon_\alpha, \quad A = 1, 2, 3,$$
$$n_\alpha^{(A)} \epsilon^\alpha = 0, \quad \epsilon^2 = -1. \tag{3.3}$$

$R_{A\alpha}$ are any three linearly independent timelike vectors from among the $R_{k\alpha}$, $k = 1, 2, \cdots, n$. ϵ^α is spacelike and normed and need not be specified any further.[11]

Upon introduction of new integration variables u and v_A by

$$u = |x - y|^2, \quad v_A = |y - a_A|^2, \tag{3.4}$$

one obtains from Eq. (3.2)

$$d^4y = \tfrac{1}{8}[-G]^{-\frac{1}{2}} v^{-1}(u, v_A)\, du\, dv_1\, dv_2\, dv_3, \tag{3.5}$$

with[12]

$$v^2 = \tfrac{1}{4}(x_A + r_A) H_{AB}^{-1} (x_B + r_B) - u. \tag{3.6}$$

Here the 3-vectors

$$x_A = \begin{pmatrix} v_1 - u \\ v_2 - u \\ v_3 - u \end{pmatrix}, \quad r_A = -\begin{pmatrix} R_1^2 \\ R_2^2 \\ R_3^2 \end{pmatrix} \tag{3.7a}$$

and the 3×3 matrix

$$H_{AB} = ((R_A^\sigma R_{B\sigma})), \quad H = \det H_{AB} = -G, \tag{3.7b}$$

have been introduced. We note that

$$r_A H_{AB}^{-1} r_B = \tfrac{1}{4} H^{-1} \Omega, \tag{3.8}$$

where Ω agrees with the quantity defined in the footnote on p. 197 of Ref. 7 and is displayed in Appendix A.

The quantities appearing in the weight function f of Eq. (2.6) have to be expressed as functions of the variables u and v_A:

$$\begin{aligned}(y - a_i, y - a_j) &= u + (R_i R_j) + \tfrac{1}{2}(x_A + r_A) H_{AB}^{-1} (g_B^{(i)} + g_B^{(j)}) \\ &\quad - \nu(u, v_A)[(g_A^{(i)} H_{AB}^{-1} g_B^{(i)} - R_i^2)^{\frac{1}{2}} \\ &\quad + (g_A^{(j)} H_{AB}^{-1} g_B^{(j)} - R_j^2)^{\frac{1}{2}}],\end{aligned} \tag{3.9}$$

with

$$g_A^{(i)} = \begin{pmatrix} (R_1 R_i) \\ (R_2 R_i) \\ (R_3 R_i) \end{pmatrix}. \tag{3.7c}$$

Equation (3.9) is valid for $1 \leq i, j \leq m$. It is advantageous to use the relations

$$H_{AB}^{-1} g_B^{(i)} = \delta_A^{(i)}, \quad \text{if } 1 \leq i \leq 3,$$

and

$$g_A^{(i)} H_{AB}^{-1} g_B^{(j)} = (R_C R_i)\delta_j^C, \quad \text{if } 1 \leq i, j \leq 3. \tag{3.10}$$

In the case of $1 \leq i, j \leq 3$, Eq. (3.9), for example, reduces to

$$(y - a_A, y - a_B) = \tfrac{1}{2}(v_A + v_B) - \tfrac{1}{2} R_{AB}^2, \tag{3.9'}$$

where

$$R_{AB\alpha} = R_{A\alpha} - R_{B\alpha} = a_{B\alpha} - a_{A\alpha}.$$

Finally, we list

$$\begin{aligned}(y - a_i, C_j) &= (R_i C_j) + \tfrac{1}{2}(x_A + r_A) H_{AB}^{-1} s_B^{(j)} \\ &\quad - \nu(u, v_A)(s_A^{(j)} H_{AB}^{-1} s_B^{(j)} - C_j^2)^{\frac{1}{2}},\end{aligned} \tag{3.11}$$

with

$$s_A^{(j)} = \begin{pmatrix} (R_1 C_j) \\ (R_2 C_j) \\ (R_3 C_j) \end{pmatrix}. \tag{3.7d}$$

We cannot achieve greater generality since the most convenient choice of tetrad legs depends much on the special GGF to be evaluated. In addition, the transformation $\kappa, \lambda, \mu, \nu \to u, v_A$ might have to be replaced by a more suitable one.

2. (2, 2) Tetrad

$$n_\alpha^{(j)} = R_{j\alpha}, \quad n_\alpha^{(3)} = \epsilon_\alpha^{(1)}, \quad n_\alpha^{(4)} = \epsilon_\alpha^{(2)}, \quad i, j = 1, 2,$$
$$n_\alpha^{(j)} \epsilon^{(i)\alpha} = 0, \quad \epsilon_\alpha^{(1)} \epsilon^{(2)\alpha} = 0, \quad \epsilon_\alpha^{(i)} \epsilon^{(i)\alpha} = -1. \tag{3.12a}$$

It is convenient to introduce plane polar coordinates ρ, ϕ in the spacelike 2-flat spanned by $\epsilon_\alpha^{(1)}$ and $\epsilon_\alpha^{(2)}$:

$$\mu = \rho \sin\phi, \quad \nu = \rho \cos\phi, \quad \rho \geq 0, \quad 0 \leq \phi < 2\pi. \tag{3.13a}$$

If no timelike $R_{j\alpha}$ are available, we take two linearly independent spacelike vectors $R_{j\alpha}$ and choose the following legs:

$$n_\alpha^{(j)} = R_{j\alpha}, \quad n_\alpha^{(3)} = \epsilon_\alpha^{(1)}, \quad n_\alpha^{(4)} = \epsilon_\alpha^{(2)}, \quad i, j = 1, 2,$$
$$n_\alpha^{(j)} \epsilon^{(i)\alpha} = 0, \quad \epsilon_\alpha^{(1)} \epsilon^{(2)\alpha} = 0, \quad \epsilon_\alpha^{(1)} \epsilon^{(1)\alpha} = -1,$$
$$\epsilon_\alpha^{(2)} \epsilon^{(2)\alpha} = +1. \tag{3.12b}$$

The vectors $\epsilon_\alpha^{(1)}$ and $\epsilon_\alpha^{(2)}$ span a timelike 2-flat in which we are going to use pseudopolar coordinates[13]

$$\mu = \rho \sinh\phi, \quad \nu = \rho \cosh \phi,$$
$$\rho \geq 0, \quad -\infty < \phi < +\infty. \tag{3.13b}$$

Next, for both cases, the integration variables u and v_A of Eq. (3.4) are introduced, but only for $A = 1, 2$:

$$\kappa, \lambda, \mu, \nu \to \kappa, \lambda, \rho, \phi \to u, v_1, v_2, \phi. \tag{3.14}$$

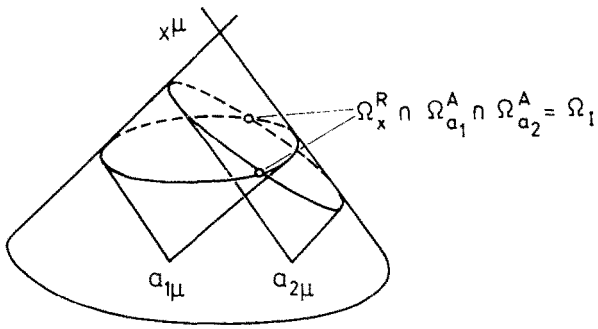


FIG. 1. Support of GGF $D^R(x, a_1, a_2 | f)$.

By use of Eqs. (3.12), (3.13), and (3.14) the four-volume element Eq. (3.2) becomes

$$d^4y = \tfrac{1}{8}[\epsilon\sigma_{12}]^{-\frac{1}{2}}\, du\, dv_1\, dv_2\, d\phi, \tag{3.15}$$

with

$$\epsilon = +1, \quad \text{for tetrad Eq. (3.12b)},$$
$$= -1, \quad \text{for tetrad Eq. (3.12a)},$$

and[14]

$$\sigma_{12} = R_1^2R_2^2 - (R_1R_2)^2. \tag{3.16}$$

An alternative to the (2, 2) tetrad (3.12) for which $R_{1\alpha}$ and $C_{1\alpha}$ have been selected as timelike legs is treated in Appendix B2.

Corresponding formulas for the case of a (1, 3) tetrad may be worked out without difficulty. However, this case is not treated here since we will have no use for it in the applications of Sec. 4.

B. Range of Integration

In order comfortably to evaluate the GGF's (2.6), the defining integral is rewritten in terms of the time-symmetric Green's function $\delta(|x-y|^2)$. The supports of the retarded Green's functions $D^R(x-y)$, $D^R(a_1-y), \cdots, D^R(a_n-y)$ are called Ω_x^R, $\Omega_{a_1}^R$, $\cdots$, $\Omega_{a_n}^R$, respectively. Then, $D^R(x, a_1, \cdots, a_n|f)$ will be nonvanishing only if the intersection

$$\Omega_I = \Omega_x^R \cap \Omega_{a_1}^A \cap \cdots \cap \Omega_{a_n}^A \neq \phi \tag{3.17}$$

is not empty.

A discussion of all possible locations for the points a_k, $k = 1, 2, \cdots, n$, with respect to the field point x shows that $\Omega_I \neq \phi$ only if (Fig. 1)

(1) $R_{k\alpha}$ is *not* spacelike, $k = 1, 2, \cdots, n$, and pointing into the future,
(2) $R_{ij\alpha}$ is *not* timelike $(i \neq j)$.

Both conditions take care of the label "retarded" for the Green's function and may be expressed by the product of step functions[15]

$$P \overset{\text{def}}{=} \prod_{i=1}^{n} \theta(R_i) \prod_{\substack{i,j \\ i\neq j}}^{1\leq i,j\leq n} \theta[-R_{ij}^2]. \tag{3.18}$$

Using $D^R(x-y) = 2\theta(x-y)\delta(|x-y|^2)$ and Eqs. (3.17) and (3.18), we may reformulate Eq. (2.6) as

$$D^R(x, a_1, \cdots, a_n/f) = 2^{n+1}\prod_{i=1}^{n}\theta(R_i)\prod_{\substack{i,j \\ i\neq j}}^{1\leq i,j\leq n}\theta[-R_{ij}^2]\bar{D}(x, a_1, \cdots, a_n/f), \tag{3.19}$$

where

$$\bar{D}(x, a_1, \cdots, a_n/f) = \frac{1}{4\pi}\int_{U(\Omega_I)} d^4y\,\delta(|x-y|^2)f(y, a_j, C_k)\prod_{i=1}^{n}\delta(|y-a_i|^2). \tag{3.20}$$

$U(\Omega_I)$ labels the range of integration: it is such a neighborhood of Ω_I that $U(\Omega_I) \cap \Omega_x^A = \{x\}$. This condition eliminates contributions from the future half of the lightcone.

In general, $-\infty < u < \infty$, $-\infty < v_A < \infty$ for the new integration variables. However, these ranges have to be modified depending on the nature of Ω_I (compare Appendix C).

4. APPLICATION OF TECHNIQUE OF INTEGRATION TO SPECIAL GGF's

In applying the manifestly Lorentz-invariant technique of integration presented in Sec. 3, we keep in mind that the results to be expected depend on the assumption that neither the Grammian determinant of the tetrad nor the Jacobians of the coordinate transformations made [Eqs. (3.4) and (3.14)] vanish. This requirement imposes restrictions on the location of the field point[16] x. We first deal with the GGF's of Eq. (2.6).

(1) *Case* $n = 1$: We shall choose two special classes of functions f.

(a) $f = f(|x-b|^2)$. Applying formulas (3.12a), (3.13a), and (3.15) for the (2, 2) tetrad and Eq. (3.19), one arrives at

$$D^R(x) * D^R(x-a)f(|x-b|^2) = \tfrac{1}{4}\theta(R_1)(-\sigma_{12})^{-\frac{1}{2}}\int_{w_1}^{w_2} dw f(w), \tag{4.1}$$

with

$$w_1 = (R_2R_{21}) - (-\sigma_{12})^{\frac{1}{2}}, \quad w_2 = (R_2R_{21}) + (-\sigma_{12})^{\frac{1}{2}}$$

[compare Eq. (C4)]. With $f = 1$ one obtains the result quoted in Eq. (1.5).

(b) $F = f[(x-a, \dot{a})]$. Although this case could be treated as in Sec. 4A, it is convenient to introduce timelike tetrads $R_{1\alpha}$, $\dot{a}_\alpha$ in place of $R_{1\alpha}$, $R_{2\alpha}$ (Appendix B). Then

$$D^R(x) * D^R(x-a)f[(x-a, \dot{a})] = \tfrac{1}{2}\theta(R_1)\rho^{-1}\int_{k_1}^{k_2} dk f(k), \tag{4.2}$$

with $k_1 = \frac{1}{2}[(R_1\dot{a}) - \rho]$ and $k_2 = \frac{1}{2}[(R_1\dot{a}) + \rho]$ where $\rho = [(R_1\dot{a})^2 - \dot{a}^2R_1^2]^{\frac{1}{2}}$. Especially, if $f[(x-a,\dot{a})] = (x-a,\dot{a})^{-l}$, one obtains

$$D^R(x) * D^R(x-a)(x-a,\dot{a})^{-l} = \tfrac{1}{2}\theta(x-a)\rho^{-1}\begin{cases}\rho, & l=0\\ \log\left|\dfrac{(R_1\dot{a})+\rho}{(R_1\dot{a})-\rho}\right|, & l=1,\\ 2^{2l-2}(1-l)^{-1}(\dot{a}^2R_1^2)^{1-l} & \\ \quad\times\{[(R_1\dot{a})+\rho]^{l-1} & \\ \quad -[(R_1\dot{a})-\rho]^{l-1}\}, & l>1.\end{cases} \tag{4.3}$$

For $l \geq 1$, Eq. (4.3) is valid only if $R_1^2 \neq 0$. The GGF's $D^R(x) * D^R(x-a)(x-a,\dot{a})^{-l}$ are singular linear functionals. As a consequence, it is not possible to apply the wave operator to the rhs of Eq. (4.3) in order to verify Eq. (2.2). The resulting expression does not exist. The integrals (4.3) have been evaluated for $l=1$ and $l=3$ by Stephani[17] by a different method.

(2) *In the case of* $n=2$ we consider

$$f = f[(x-a_1,\dot{a}_1),(x-a_2,\dot{a}_2)].$$

Again one applies the formulas (3.12a), (3.13a), and (3.15) for the (2, 2) tetrad and Eq. (3.19). However, now $(y-a_1,\dot{a}_1)$ and $(y-a_2,\dot{a}_2)$ have also to be expressed by the integration variables u, v_1, v_2, and ϕ. The calculation leads to

$$\begin{aligned}(y-a_1,\dot{a}_1) &= c + d\cos\phi + e\sin\phi,\\ (y-a_2,\dot{a}_2) &= a + b\cos\phi,\end{aligned} \tag{4.4}$$

with constants $a, b, \cdots, e$ listed in Appendix E. Therefore,

$$\begin{aligned}&D^R(x) * D^R(x-a_1)D^R(x-a_2)\\ &\quad\times f[(x-a_1,\dot{a}_1),(x-a_2,\dot{a}_2)]\\ &= \theta(R_1)\theta(R_2)\theta(-R_{12}^2)(-\sigma_{12})^{-\frac{1}{2}}\\ &\quad\times\frac{1}{4\pi}\int_0^{2\pi}d\phi f[c+d\cos\phi+e\sin\phi, a+b\cos\phi].\end{aligned} \tag{4.5}$$

The choice $f=1$ leads to the generating GGF of order[18] 2:

$$\begin{aligned}g_2(x,a_1,a_2) &= D^R(x) * D^R(x-a_1)D^R(x-a_2)\\ &= \tfrac{1}{2}\theta(R_1)\theta(R_2)\theta(-R_{12}^2)(-\sigma_{12})^{-\frac{1}{2}},\\ &\qquad \sigma_{12}\neq 0.\end{aligned} \tag{4.6}$$

In Appendix E [Eqs. (E3) and (E4)] the further special choices $f = \ln(x-a_1,\dot{a}_1)$, $f = \ln(x-a_2,\dot{a}_2)$, $f = (x-a_1,\dot{a}_1)^{-l}(x-a_2,\dot{a}_2)^{-m}$, $l, m \geq 1$ are discussed.

(3) *In the case* $n \geq 3$ we restrict the application of the technique of integration to $f=1$. This is done only for reasons of avoiding lengthy general formulas; Eqs. (3.3), (3.5), (3.6), (3.9), and (3.19) lead to[19]

$$\begin{aligned}&D^R(x) * \prod_{k=1}^{n} D^R(x-a_k)\\ &= \frac{1}{\pi}2^{n-2}\Omega^{-\frac{1}{2}}\prod_{k=1}^{n}\theta(R_k)\prod_{\substack{k,j\\ k\neq j}}^{1\leq k,j\leq n}\theta(-R_{kj}^2)\\ &\quad\times\prod_{j=4}^{n}\delta[R_j^2 + r_A H_{AB}^{-1}g_B^{(j)} - \tfrac{1}{2}\Omega^{\frac{1}{2}}(g_A^{(j)}H_{AB}^{-1}g_B^{(j)} - r_j^2)^{\frac{1}{2}}].\end{aligned} \tag{4.7}$$

For $n=3$ the result $Y_3(x,a,b,c)$ of Ref. 7 is regained (except for a factor of 2).

(4) Next, we try to evaluate the *generating* Green's function of third order:

$$\begin{aligned}&g_3(x,a_1,a_2,a_3)\\ &= \frac{1}{4\pi}\int d^4y D^R(x-y)D^R(y-a_3)g_2(y,a_1,a_2).\end{aligned} \tag{4.8}$$

Applying (4.6), we use the fact that $\sigma_{12}(x)$ of Eq. (3.16) can be rewritten as

$$-\sigma_{12} = (R_1R_{12})^2 - R_1^2R_{12}^2.$$

The vectors $R_{ij\alpha}$ do not depend on the integration variable y. Hence, Eq. (3.11) with $C_{j\alpha} = R_{12\alpha}$ can be used to derive

$$\begin{aligned}-\sigma_{12} &= (y-a_1, y-a_2)^2 - |y-a_1|^2\cdot|y-a_2|^2\\ &= \tfrac{1}{4}[v_1+v_2-R_{12}^2]^2 - v_1v_2.\end{aligned} \tag{4.9}$$

Furthermore, the step functions in Eq. (4.6) may be expressed by help of the timelike vector[20] $R_{3\alpha}$:

$$\begin{aligned}&\theta(y-a_1)\theta(y-a_2)\\ &\qquad = \theta[(y-a_1,R_3)]\cdot\theta[(y-a_2,R_3)].\end{aligned} \tag{4.10}$$

Finally, it turns out to be convenient to transform v_1 and v_2 into two new integration variables x and y by

$$\begin{aligned}v_1 &= x(-\sigma_{13})^{\frac{1}{2}} + (R_1R_{13}),\\ v_2 &= y(-\sigma_{23})^{\frac{1}{2}} + (R_2R_{23}),\end{aligned} \tag{4.11}$$

in order to simplify the limits of integrations (compare Appendix C). The (3, 1) tetrad formulas, together with Eqs. (4.9)–(4.11), then lead to

$$\begin{aligned}&g_3(x,a_1,a_2,a_3)\\ &= (8\pi)^{-1}\theta(R_3)\theta(-R_{12}^2)\theta[(R_1R_3)-\tfrac{1}{2}R_3^2]\theta[(R_2R_3)-\tfrac{1}{2}R_3^2]\\ &\quad\times\int_{-1}^{+1}dx\int_{-1}^{+1}dy(x^2+y^2+2\bar{p}_2xy+p_0)^{-\frac{1}{2}}\\ &\quad\times(q_2x^2+q_2'y^2+2\bar{q}_2xy+2q_1x+2q_1'y+q_0)^{-\frac{1}{2}},\end{aligned} \tag{4.12}$$

with

$$\bar{p}_2 = -t_{12}(\sigma_{13}\sigma_{23})^{-\frac{1}{2}}, \quad p_0 = -R_3^2(\sigma_{13}\sigma_{23})^{-1}H,$$
$$q_2 = \sigma_{13}, \quad q_2' = \sigma_{23}, \quad \bar{q}_2 = (\sigma_{13}\sigma_{23})^{\frac{1}{2}},$$
$$q_1 = -(-\sigma_{13})^{\frac{1}{2}}R_{12\sigma}(2R_2^\sigma - R_3^\sigma),$$
$$q_1' = (-\sigma_{23})^{\frac{1}{2}}R_{12\sigma}(2R_1^\sigma - R_3^\sigma),$$
$$q_0 = 4R_{12}^2(R_1R_{13}) - [R_{12\sigma}(2R_1^\sigma - R_3^\sigma)]^2.$$

As discussed further in Appendix D, the remaining integrations in Eq. (4.12) cannot be expressed by elementary functions but lead to an integral of a Jacobian elliptical function. However, in the case of $R_{3\alpha}$ being a null vector, Eq. (4.12) shows that $\bar{p}_2 = -1$, $p_0 = 0$, and

$$g_2(x, a_1, a_2, a_3) \sim \int_{-1}^{+1} dx \int_{-1}^{+1} dy (x-y)^{-1} \times (q_2x^2 + q_2'y^2 + 2\bar{q}_2xy + 2q_1x + 2q_1'y + q_0)^{-\frac{1}{2}},$$

so that at least one integration can be carried through explicitly. In general, already in the third step of the Lorentz-invariant perturbation expansion, one will have to deal with either an integral representation or a series expansion for the GGF's.

5. TIME-SYMMETRIC GGF's

From the construction of the GGF's (Sec. 2B), especially from the form of the elements (2.3), it is clear that the time-symmetric GGF is not just a linear combination of the retarded and advanced GGF's. This is most easily exemplified by the time-symmetric GGF corresponding to Eq. (2.6):

$$D^{\mathrm{S}}(x, a_1, \cdots, a_n \,|\, f) = \tfrac{1}{4}\pi \int d^4y \prod_{i=1}^{n} D^{\mathrm{S}}(y - a_n) D^{\mathrm{S}}(x - y) f(y_\sigma, c_\sigma). \quad (5.1)$$

With $D^{\mathrm{S}}(x) = \frac{1}{2}[D^{\mathrm{R}}(x) + D^{\mathrm{A}}(x)]$, one may rewrite Eq. (5.1) to be

$$D^{\mathrm{S}}(x, a_1, \cdots, a_n \,|\, f) = D^{\mathrm{R}}(x, a_1, \cdots, a_n \,|\, f) + D^{\mathrm{A}}(x, a_1, \cdots, a_n \,|\, f) + D^{\mathrm{mix}}(x, a_1, \cdots, a_n \,|\, f), \quad (5.2)$$

where[21]

$$D^{\mathrm{mix}}(x, a_1, \cdots, a_n \,|\, f) = (4\pi)^{-1} \sum_{j=0}^{n-1} \int d^4y \mathfrak{P}_j \Big(D^{\mathrm{R}}(x - y) \prod_{i=1}^{j} D^{\mathrm{R}}(y - a_i) \times \prod_{m=j+1}^{n} D^{\mathrm{A}}(y - a_m) \Big) f(y_\sigma, C_\sigma). \quad (5.3)$$

Eq. (5.2) reflects the nonlinearity of the underlying field equations which has been used to construct the set $\mathfrak{F}$ of Sec. 2B.

As in the integral (2.6), we have to deal with the intersection of a number of half light cones in D^{mix}. However, the two conditions listed in Sec. 2B are no longer valid: one discovers that there may be a non-empty intersection of the half light cones even if some of the vectors $R_{i\alpha}$ are *spacelike*.

Also, as already mentioned in Sec. 2B, it is to be expected that the support of the integrand in Eq. (5.1) becomes unbounded if the variables $a_{k\alpha}$ are free to assume all possible space–time positions with respect to the field point x_α. This may be illustrated by example of the GGF $g_2(x, a_1, a_2)$ of Eq. (4.6). A detailed examination[22] shows that Eq. (5.1) can be rewritten as

$$D^{\mathrm{s}}(x, a_1, a_2, \,|\, 1) = \{[\theta(R_1)\theta(R_2) + \theta(-R_1)\theta(-R_2)] + \theta(-R_1^2) \times [\theta(R_2)\theta(R_{21}) + \theta(-R_2)\theta(R_{12})] + \theta(-R_2^2)[\theta(R_1)\theta(R_{12}) + \theta(-R_1)\theta(R_{21})]\} \times \bar{D}(x, a_1, a_2 \,|\, 1) + 2\theta(-R_1^2)\theta(-R_2^2)\theta(-R_{12}^2)\bar{D}(x, a_1, a_2 \,|\, 1), \quad (5.4)$$

with $\bar{D}(x, a_1, a_2 \,|\, 1)$ from Eq. (3.20). The terms resulting from the first and second lines are finite. In this case, $\bar{D}$ is calculated by use of a (2, 2) tetrad with timelike legs $R_{2\alpha}$, $R_{1\alpha}$; $R_{2\alpha}$, $R_{21\alpha}$; $R_{1\alpha}$, $R_{12\alpha}$, respectively, and displayed in Eq. (4.6) (step functions disregarded). However, as already indicated by the step functions, $\bar{D}(x, a_1, a_2 \,|\, 1)$ in the last line of Eq. (5.4) has to be calculated by help of Eqs. (3.12b), (3.13b), and (3.15). One arrives at an integral $\int_{-\infty}^{+\infty} d\phi$ instead of $\int_0^{2\pi} d\phi$ for the previous cases, that is, one obtains an *infinite* contribution to Eq. (5.4).

In order to avoid the difficulties encountered, one might propose a change in the definition (5.1) and (5.2) to

$$D'(x, a_1, \cdots, a_n \,|\, f) = D^{\mathrm{R}}(x, a_1, \cdots, a_n \,|\, f) + D^{\mathrm{A}}(x, a_1, \cdots, a_n \,|\, f). \quad (5.5)$$

However, the use of such expressions in the Lorentz-invariant perturbation theory of Einstein's theory of gravitation (Sec. 1) amounts to a further truncation of the nonlinearity of Einstein's equations. On the other

hand, adoption of Eq. (5.5) frees one from the difficulty that necessary conditions for the existence of symmetric GGF's defined by Eq. (5.1) are not known.

6. CONCLUSIONS

The weak-field (fast-motion) approximation for a system of gravitationally interacting point particles has been used to illustrate the appearance of generalized Green's functions in physics. The character of the interaction of the particles in an n-body system, as it appears in various stages of approximation, may be read off from the GGF's. The objectives of the ensuing discussion have been threefold:

(1) to describe the mathematical nature of GGF's;
(2) to exhibit a manifestly Lorentz-invariant technique for the evaluation of the convolution representation (2.1) of the GGF's;
(3) to discuss time-symmetric GGF's.

The technique exhibited provides the tools for the calculation of higher-order approximations in a Lorentz-invariant perturbational approach to the classical n-body problem. Actual calculations going beyond the second approximation (i.e., beyond the first approximative step taking into account the nonlinearity of the theory) will have to cope with the rapidly increasing complexity of the GGF's. Furthermore, the example (4.3) shows that the GGF's can be *singular* generalized functions (with singularity at the retarded point) so that integrals over the particle's world lines of the type of (1.8) may not exist. Consequently, in order to write down the field of the interacting particles (or equations of motion, etc.) one will first have to provide for a suitable regularization. This situation is found already in the second order of the fast-motion approximation for gravitationally interacting particles.

ACKNOWLEDGMENTS

The author is indebted to Professor P. Havas for his kind hospitality at Temple University. He is very grateful to both Professor Havas and Professor K. Westpfahl, Freiburg, for discussions and constructive criticism.

APPENDIX A: SUPPLEMENTARY FORMULAS FOR THE TETRAD TECHNIQUE

(1) The explicit components of the matrix

$$H_{AB}^{-1} = H^{-1}\begin{pmatrix} \sigma_{23} & t_{12} & t_{13} \\ t_{12} & \sigma_{13} & t_{23} \\ t_{13} & t_{23} & \sigma_{12} \end{pmatrix} \tag{A1}$$

are

$$\sigma_{AB} = R_A^2 R_B^2 - (R_A R_B)^2,$$
$$t_{AB} = (R_A R_C)(R_B R_C) - (R_A R_B) R_C^2,$$
$$A, B, C \text{ cyclic permutations of } 1, 2, 3.$$

(2) By use of Eq. (3.1), one concludes from Eq. (3.4) that

$$u = p_A H_{AB} p_B - v^2,$$
$$v_A = p_B H_{BC} p_C + 2H_{AB} p_B + R_A^2 - v^2 \tag{A2}$$

and by subtraction

$$x_A = 2H_{AB} p_B - r_A, \quad A = 1, 2, 3, \tag{A3}$$

with x_A, r_A, H_{AB} defined by Eq. (3.7) and

$$p_A = \begin{pmatrix} \kappa \\ \lambda \\ \mu \end{pmatrix}.$$

If (A3) is solved for p_A and the result entered into the first equation (A2), then Eq. (3.6) follows. Explicitly,

$$\begin{aligned} &H(x_A + r_A) H_{AB}^{-1} (x_B + r_B) \\ &= \sigma_{23}(v_1 - u)^2 + \sigma_{13}(v_2 - u)^2 + \sigma_{12}(v_3 - u)^2 \\ &\quad + 2t_{12}(v_1 - u)(v_2 - u) + 2t_{13}(v_1 - u)(v_3 - u) \\ &\quad + 2t_{23}(v_2 - u)(v_3 - u) \\ &\quad - 2(R_1^2\sigma_{23} + R_2^2 t_{12} + R_3^2 t_{13})(v_1 - u) \\ &\quad - 2(R_1^2 t_{12} + R_2^2\sigma_{13} + R_3^2 t_{23})(v_2 - u) \\ &\quad - 2(R_1^2 t_{13} + R_2^2 t_{23} + R_3^2\sigma_{12})(v_3 - u) + \tfrac{1}{4}\Omega, \end{aligned} \tag{A4}$$

with

$$\begin{aligned} \tfrac{1}{4}\Omega &= (R_1^2)^2\sigma_{23} + 2R_1^2 R_2^2 t_{12} + 2R_1^2 R_3^2 t_{13} \\ &\quad + (R_2^2)^2\sigma_{13} + 2R_2^2 R_3^2 t_{23} + (R_3^2)^2\sigma_{12}. \end{aligned} \tag{A5}$$

In Eq. (4.7), $v(0, 0, 0, 0)$ has been used, whereas in Eq. (4.12), $v(0, v_1, v_2, 0)$ has been used.

(3) For the case of the (2, 2) tetrad, we list another useful formula which expresses the coordinate ρ of Eq. (3.13) as function of u, v_A, $A = 1, 2$:

$$\begin{aligned} &\rho^2(u, v_A) \\ &= \epsilon u + \tfrac{1}{4}(-\epsilon\sigma_{12})^{-1}[R_1^2(v_2 - u)^2 + R_2^2(v_1 - u)^2 \\ &\quad - 2(R_1 R_2)(v_1 - u)(v_2 - u) - 2R_2^2(R_1 R_{12})(v_1 - u) \\ &\quad + 2R_1^2(R_2 R_{12})(v_2 - u) + R_1^2 R_2^2 R_{12}^2]. \end{aligned} \tag{A6}$$

[ϵ is defined after Eq. (3.15), Sec. 3A.]

APPENDIX B: ALTERNATIVE (2, 2) TETRAD FOR CALCULATION OF (4.2)

In (3.12a) the leg $n_\alpha^{(2)} = R_{2\alpha}$ is replaced by

$$n_\alpha^{(2)} = \dot{a}_\alpha, \tag{B1}$$

with all other relations of (3.12a) being unchanged.

The transformation (3.14) is changed to

$$\kappa, \lambda, \mu, \nu \to \kappa, \lambda, \rho, \phi \to \kappa, u, v_1, \phi. \quad \text{(B2)}$$

Therefore, Eq. (3.15) has to be replaced by

$$d^4y = -\tfrac{1}{4}(R_1\dot{a}_1)^{-1}\rho_1\, d\kappa\, du\, dv_1\, d\phi, \quad \text{(B3)}$$

with $\rho_1 = [(R_1\dot{a}_1)^2 - \dot{a}_1^2R_1^2]^{\frac{1}{2}}$. The range of the variable κ on $u = v_1 = 0$ is

$$\tfrac{1}{2}[1 - \rho_1^{-1}(R_1\dot{a}_1)] \leq \kappa \leq \tfrac{1}{2}[1 + \rho_1^{-1}(R_1\dot{a}_1)] \quad \text{(B4)}$$

and

$$(y - a_1, \dot{a}_1) = \tfrac{1}{2}(R_1\dot{a}_1)^{-1}[v_1 - u + R_1^2 + 2\rho_1^2(1 - \kappa)]. \quad \text{(B5)}$$

APPENDIX C: RANGE OF INTEGRATION

The range of integration $-\infty < u < +\infty$, $-\infty < v_A < \infty$, $A = 1, 2, 3$, has to be modified according to the various possible intersections of half light cones.

Case of (3, 1) Tetrad, Eq. (3.3)

(1) On $u = 0$:

$$-\infty < v_A < \infty. \quad \text{(C1)}$$

(2) On $u = v_3 = 0$:

$$(R_1R_{13}) - (-\sigma_{13})^{\frac{1}{2}} \leq v_1 \leq (R_1R_{13}) + (-\sigma_{13})^{\frac{1}{2}},$$
$$(R_2R_{23}) - (-\sigma_{23})^{\frac{1}{2}} \leq v_2 \leq (R_2R_{23}) + (-\sigma_{23})^{\frac{1}{2}}, \quad \text{(C2)}$$

and similar formulas for exchange of variables.

(3) On $u = v_1 = v_2 = 0$:

$$(-\sigma_{12})^{-1}[R_1^2t_{13} + R_2^2t_{23} + R_3^2\sigma_{12} - (-HR_1^2R_2^2R_{12}^2)^{\frac{1}{2}}] \leq v_3 \leq (-\sigma_{12})^{-1}[R_1^2t_{13} + R_2^2t_{23} + R_3^2\sigma_{12} + (-HR_1^2R_2^2R_{12}^2)^{\frac{1}{2}}]. \quad \text{(C3)}$$

Case of (2, 2) Tetrad, Eq. (3.12)

On $u = v_2 = 0$:

$$(R_1R_{12}) - (-\sigma_{12})^{\frac{1}{2}} \leq v_1 \leq (R_1R_{12}) + (-\sigma_{12})^{\frac{1}{2}}. \quad \text{(C4)}$$

Equation (C2) has been used for Eq. (4.12).

APPENDIX D: FURTHER DISCUSSION OF $g_3(x, a_1, a_2, a_3)$

The substitution

$$x = \frac{l + m\bar{x} + n\bar{y}}{1 + \bar{x} + \bar{y}}, \quad y = \frac{s + t\bar{x} + u\bar{y}}{1 + \bar{x} + \bar{y}}, \quad \text{(D1)}$$

with free parameters l, m, n, s, t, and u, may be used to transform the integral of Eq. (4.12) to

$$J = \int_{\bar{x}_1}^{\bar{x}_2} d\bar{x} \int_{\bar{y}_1}^{\bar{y}_2} d\bar{y}(1 + \bar{x} + \bar{y})^{-1} \times [(\alpha\bar{x}^2 + \beta\bar{y}^2 + \gamma)(\delta\bar{x}^2 + \epsilon\bar{y}^2 + \xi)]^{-\frac{1}{2}}, \quad \text{(D2)}$$

where the constants $\alpha, \cdots, \xi$ depend on $\tilde{p}_2, \cdots, q_0$ of Eq. (4.13). To achieve this, one has to solve a system of six *nonlinear* (quadratic) equations for the unknowns l, m, n, s, t, and u. In order to perform the integration with respect to $\bar{x}$ in Eq. (D2), one observes that

$$(1 + \bar{x} + \bar{y})^{-1} = -\bar{x}[(1 + \bar{y})^2 - \bar{x}^2]^{-1} + (1 + \bar{y})[(1 + \bar{y})^2 - \bar{x}^2]^{-1}$$

and obtains

$$J = \int_{\bar{y}_1}^{\bar{y}_2} d\bar{y}\left(-\int_{\bar{x}_1}^{\bar{x}_2} \frac{\bar{x}\, d\bar{x}}{[(1 + \bar{y})^2 - \bar{x}^2]\sqrt{R}} + (1 + \bar{y})\int_{\bar{x}_1}^{\bar{x}_2} \frac{d\bar{x}}{[(1 + \bar{y})^2 - \bar{x}^2]\sqrt{R}}\right), \quad \text{(D3)}$$

where

$$R = \alpha\beta[\bar{x}^2 + \mu(\bar{y})][\bar{x}^2 + \nu(\bar{y})],$$

with

$$\mu(y) = \alpha^{-1}(\beta\bar{y}^2 + \gamma),$$
$$\nu(y) = \delta^{-1}(\epsilon\bar{y}^2 + \xi).$$

The $\bar{x}$ integration in the first term can be carried through explicitly. The second integral can be reduced to Jacobian elliptic functions. The transformations needed depend on the signs of μ and ν, respectively.[23] The coefficients $\alpha, \beta, \cdots, \xi$ have not been worked out. They are rather lengthy rational functions of the original $\tilde{p}_2, \cdots, q_0$ of Eq. (4.13).

APPENDIX E: LISTING OF FURTHER SPECIAL GGF's

In this Appendix we are going to further evaluate the GGF's of Eq. (4.5):

$$D^{\mathrm{R}}(x) * D^{\mathrm{R}}(x - a_1)D^{\mathrm{R}}(x - a_2) \times f[(x - a_1, \dot{a}_1), (x - a_2, \dot{a}_2)] = \theta(R_1)\theta(R_2)\theta(-R_{12}^2)(-\sigma_{12})^{-\frac{1}{2}} \times \frac{1}{4\pi}\int_0^{2\pi} d\phi f[c + d\cos\phi + e\sin\phi, a + b\cos\phi], \quad \text{(E1)}$$

with

$$a = \tfrac{1}{2}(-\sigma_{12})^{-1}[R_2^2(lR_{12}) + R_{12}^2(lR_2)],$$
$$b = \tfrac{1}{2}(-\sigma_{12})^{-1}[R_1^2R_2^2R_{12}^2(-\sigma_{12}\dot{a}_2^2 + l^2)]^{\frac{1}{2}},$$
$$c = \tfrac{1}{2}(-\sigma_{12})^{-1}[-R_1^2(kR_{12}) + R_{12}^2(kR_1)],$$
$$d = +\tfrac{1}{2}(-\sigma_{12})^{-1}[+\sigma_{12}(\dot{a}_1\dot{a}_2) + (lk)]\left(\frac{R_1^2R_2^2R_{12}^2}{l^2 - \sigma_{12}\dot{a}_2^2}\right)^{\frac{1}{2}},$$
$$e = -\tfrac{1}{2}(-\sigma_{12})^{-1}\{(l^2 - \dot{a}_2^2\sigma_{12})(k^2 - \dot{a}_1^2\sigma_{12}) - [\sigma_{12}(\dot{a}_1\dot{a}_2) + (lk)]^2\}^{\frac{1}{2}}\left(\frac{R_1^2R_2^2R_{12}^2}{l^2 - \sigma_{12}\dot{a}_2^2}\right)^{\frac{1}{2}}, \quad \text{(E2)}$$

where the notations

$$k_\mu = -R_{1u}(R_{12}\dot{a}_1) + R_{12\mu}(R_1\dot{a}_1),$$
$$l_\mu = R_{2u}(R_{12}\dot{a}_2) - R_{12\mu}(R_2\dot{a}_2)$$

have been used.

For $f = \ln(x - a_1, \dot{a}_1)$ or $f = \ln(x - a_2, \dot{a}_2)$, the integral in Eq. (E1) reduces to the standard integral

$$\int_0^{2\pi} d\phi \ln \begin{Vmatrix} |c + d\cos\phi + e\sin\phi| \\ |a + b\cos\phi| \end{Vmatrix}$$

The same holds for $f = (x - a_1, \dot{a}_1)^{-l}(x - a_2, \dot{a}_2)^{-m}$, $l, m \geq 1$, because

$$\int_0^{2\pi} d\phi (c + d\cos\phi + e\sin\phi)^{-l}(a + b\cos\phi)^{-m}$$
$$= (-1)^{l+m-2} \frac{d^{l+m-2}}{dc^{l-1}\, da^{m-1}}$$
$$\times \int_0^{2\pi} d\phi (c + d\cos\phi + e\sin\phi)^{-1}(a + b\cos\phi)^{-1}.$$

It may suffice to list the following results explicitly:

$$D^{\mathrm{R}}(x) * D^{\mathrm{R}}(x - a_1) D^{\mathrm{R}}(x - a_2) \ln(x - a_2, \dot{a}_2)$$
$$= \tfrac{1}{16}(-\sigma_{12})^{-\frac{1}{2}} \ln |\tfrac{1}{4}(-\sigma_{12})^{-1}$$
$$\times [(-\sigma_{12})^{\frac{1}{2}} P_{a_2} + R_2^2(lR_{12}) + R_{12}^2(lR_{12})]|, \quad \text{(E3)}$$

$$D^{\mathrm{R}}(x) * D^{\mathrm{R}}(x - a_1) D^{\mathrm{R}}(x - a_2)(x - a_1, \dot{a}_1)^{-n}$$
$$\times (x - a_2, \dot{a}_2)^{-m}$$
$$= \theta(R_1)\theta(R_2)\theta(R_{12}^2) \begin{cases} P_{a_2}^{-1}, & n = 0, m = 1, \\ -4\sigma_{12} P_{a_2}^{-3}, & n = 0, m = 2, \\ 2\sigma_{12}(M P_{a_2}^{-1} + N P_{a_1}^{-1}) \\ \quad \times \{M[-R_1^2(kR_{12}) + R_{12}^2(kR_1)] \\ \quad + N[R_2^2(lR_{12}) + R_{12}^2(lR_2)] \\ \quad - (l^2 - \dot{a}_2^2\sigma_{12})(k^2 - \dot{a}_1^2\sigma_{12}) \\ \quad + [\sigma_{12}(\dot{a}_1\dot{a}_2) + (lk)]^2\}^{-1}, \\ \qquad n = 1, m = 1. \end{cases} \quad \text{(E4)}$$

By interchanging of $a_{1\alpha}$ and $a_{2\alpha}$, n and m, respectively, we get further expressions ($\dot{a}_{1\alpha}$ and $\dot{a}_{2\alpha}$ have to be interchanged, too). In (E4) the following abbreviations have been used:

$$P_{a_2} = \{[R_2^2(R_{12}\dot{a}_2) + R_{12}^2(R_2\dot{a}_2)]^2 - \dot{a}_2^2 R_1^2 R_2^2 R_{12}^2\}^{\frac{1}{2}},$$
$$P_{a_1} = \{[R_1^2(R_{21}\dot{a}_1) + R_{12}^2(R_1\dot{a}_1)]^2 - \dot{a}_1^2 R_1^2 R_2^2 R_{12}^2\}^{\frac{1}{2}},$$
$$M = -[+\sigma_{12}(\dot{a}_1\dot{a}_2) + (lk)][R_2^2(lR_{12}) + R_{12}^2(lR_2)]$$
$$+ (l^2 - \dot{a}_2^2\sigma_{12})[-R_1^2(kR_{12}) + R_{12}^2(kR_1)],$$
$$N = -[\sigma_{12}(\dot{a}_1\dot{a}_2) + (lk)][-R_1^2(kR_{12}) + R_{12}^2(kR_1)]$$
$$+ (k^2 - \dot{a}_1^2\sigma_{12})[R_2^2(lR_{12}) + R_{12}^2(lR_2)].$$

APPENDIX F: SUPPLEMENT TO THE DISCUSSION OF $D^{\mathrm{s}}(x, a, b \mid 1)$

The possible relative space–time locations of the points x_α, $a_\alpha(s_a)$, and $b_\alpha(s_b)$ are described by the following list of products of step functions:

$$\begin{aligned}
&(1)\quad \theta(R_1)\theta(R_2)\theta(R_{12}),\\
&(2)\quad \theta(-R_1)\theta(-R_2)\theta(-R_{12}),\\
&(3)\quad \theta(R_1)\theta(-R_2)\theta(R_{12}),\\
&(4)\quad \theta(R_1)\theta(R_2)\theta(-R_{12}^2),\\
&(5)\quad \theta(-R_1)\theta(-R_2)\theta(-R_{12}^2),\\
&(6)\quad \theta(R_2)\theta(-R_{12})\theta(-R_1^2),\\
&(7)\quad \theta(-R_2)\theta(R_{12})\theta(-R_1^2),\\
&(8)\quad \theta(R_1)\theta(-R_2^2)\theta(-R_{12}^2),\\
&(9)\quad \theta(-R_1)\theta(-R_2)\theta(-R_{12}^2),\\
&(10)\quad \theta(R_{12})\theta(-R_1^2)\theta(-R_2^2),\\
&(11)\quad \theta(-R_1^2)\theta(-R_2^2)\theta(-R_{12}^2),
\end{aligned} \quad \text{(F1)}$$

and seven expressions obtained by interchanging indices 1 and 2 in the above. The integral of $D^{\mathrm{s}}(x, a, b \mid 1)$ is—according to Eq. (5.1)—made up of eight different products of retarded and advanced Green's functions, the intersection Ω_I of the supports of which is empty except for just one out of the listed combinations in (F1). Writing the numbers (K) of (F1) before the corresponding term, we obtain as the only nonvanishing contributions

$$\begin{aligned}
&(6)\quad D^{\mathrm{R}}(x - y) D^{\mathrm{A}}(y - a) D^{\mathrm{R}}(y - b),\\
&(7)\quad D^{\mathrm{A}}(x - y) D^{\mathrm{R}}(y - a) D^{\mathrm{A}}(y - b),\\
&(1)\quad D^{\mathrm{R}}(x - y) D^{\mathrm{R}}(y - a) D^{\mathrm{R}}(y - b),\\
&(2)\quad D^{\mathrm{A}}(x - y) D^{\mathrm{A}}(y - a) D^{\mathrm{A}}(y - b),\\
&(11)\quad D^{\mathrm{A}}(x - y) D^{\mathrm{R}}(y - a) D^{\mathrm{R}}(y - b),\\
&(11)\quad D^{\mathrm{R}}(x - y) D^{\mathrm{A}}(y - a) D^{\mathrm{A}}(y - b),
\end{aligned}$$

as well as two further terms following by interchange of a_α and b_α. This then leads to Eq. (5.4).

* Research supported by the Aerospace Research Laboratories of the office of Aerospace Research, U.S. Air Force. The work was begun while holding a NATO Fellowship provided by Deutscher Akademischer Austauschdienst, and while on leave from Institut für Theoretische Physik, Freiburg, Germany.

† Present address: Institut für Theoretische Physik, Bunsenstrasse 9, 34 Göttingen, Germany.

[1] Notation used: Minkowski metric $\eta_{\alpha\beta}$ with signature -2. x or y stand for all four coordinates in R^4. Integrals are to be taken over all of R_4 if not indicated otherwise. Positions of particles are described by $a, b, \cdots$ or by $a_{k_\alpha}(s_a)$, $k = 1, 2, \cdots, n$. The proper time s_a of a particle with world line $a_\alpha = a_\alpha(s_a)$ is defined by $ds_a^2 = \eta_{\alpha\beta}\, da^\alpha da^\beta$. The radius vector pointing from the particle's location to the field point is $R_{k_\alpha} = x_\alpha - a_{k\alpha}$. We further use $R_{ij\alpha} = R_{i\alpha} - R_{j\alpha}$. A dot means differentiation with regard to proper time. $\delta^4(x) = \prod_{i=0}^{3} \delta(x_i)$ is the 4-dimensional Dirac δ. The Green's function satisfies $\Box D(x) = 4\pi\delta^4(x)$ with the wave operator $\Box = \eta^{\alpha\beta}(\partial^2/\partial x^\alpha \partial x^\beta)$. Greek indices run from 0 to 3. For them (as well as for capital Latin indices running from 1 to 2 or 3) the summation convention is used. Small Latin indices usually refer to particle numbers. The inner product in space–time is abbreviated by $(A - B, C)$, (AB), and A^2, respectively.

[2] In its simplest form, such a generalization goes back to P. A. M. Dirac.

[3] The support of a generalized function is the smallest closed set of points outside of which it vanishes. For concepts concerning linear functionals, consult L. Schwartz, *Théorie des distributions* (Hermann, Paris, 1957), I. M. Gel'fand and G. E. Shilov, *Generalized Functions* (Academic, New York, 1964), Vol. I, Chap. I, Sec. 1, and W. Güttinger, Fortschr. Physik **14**, 483 (1966).

[4] Step function

$$\theta(x-a) = 1, \quad \text{for} \quad x_0 - a_0 > |x_k - a_k|,$$
$$= 0, \quad \text{for} \quad x_0 - a_0 < |x_k - a_k|.$$

[5] W. C. Schieve, thesis, Lehigh University, Bethlehem, Pa., 1959.

[6] The generalized Green's functions dealt with in this investigation are defined with the help of Lorentz-invariant Riesz distributions. See M. Riesz, Acta Math. **81**, 1 (1949).

[7] B. Bertotti and J. Plebański, Ann. Phys. (N.Y.) **11**, 169 (1960).

[8] Obviously, GGF's can be defined for other hyperbolic differential equations as well. One just has to substitute for $D^R(x)$ the corresponding solution in Eq. (2.1). We deal here, however, only with the wave equation.

[9] See W. Güttinger, Ref. 3, p. 505, and L. Hörmander, *Linear Partial Differential Equations* (Springer, Berlin, 1964).

[10] Even if the GGF is a regular generalized function, the operation on it by differential operators [as in Eq. (1.8)] might generate a singular linear functional.

[11] In the following, capital Latin indices run always from 1 to 3.

[12] Consult Appendix A for Eq. (3.6) as well as for the explicit components of H_{AB}^{-1}.

[13] The coordinization (3.13b) does cover only that part of the timelike 2-flat which can be reached from the origin by timelike directions. The other points may be obtained by taking

$$\mu = \rho \cosh\phi, \quad \nu = \rho \sinh\phi, \quad \rho \geq 0, \quad -\infty < \phi < \infty.$$

[14] Compare Eq. (A1).

[15] For scalar argument u,

$$\theta(u) = 1, \quad \text{if} \quad u > 0,$$
$$= 0, \quad \text{otherwise.}$$

[16] For example, if the motion of three or more particles is coplanar, the field point x cannot be chosen to lie in this plane, etc.

[17] H. Stephani, Acta Physica Polon. **26**, 1045 (1964).

[18] It seems worthwhile to note that $D(x) * D(x-a_1) \cdot D(x-a_2)$ —with $D(x)$ being the retarded, advanced, symmetric Green's function or the Feynman propagator—may be evaluated by help of the Fourier representation of $D(x)$. Then

$$D(x) * D(x-a_1)D(x-a_2) = -\frac{1}{16\pi^6}\oint_C d^4k\, \frac{e^{ik_\alpha R_1{}^\alpha}}{k^2} \times \oint_{C'} d^4k'\, \frac{e^{ik'_\mu R_2{}^\mu}}{k'^2(k+k')^2}.$$

The loops C, C' around the poles of $(k^2)^{-1}$, $(k'^2)^{-1}$, and $[(k+k')^2]^{-1}$ have to be chosen properly (64 possibilities!). However, the calculations are much longer and cannot be kept Lorentz invariant in each step.

[19] We define $\prod_{j=l}^{l'}(\ )_j = 1$, if $l' < l$.

[20] In the coordinate system in which $R_{3\alpha} = R_{30}\delta_\alpha^0$, one finds

$$\theta(y_0 - a_0) = \theta[(y_0 - a_0)R_{30}]$$

due to $R_{30} > 0$.

[21] The symbol $\mathfrak{P}_j$ means: Form all combinations of $j+1$ elements out of the given set $x_1, a_1, \cdots, a_n$; use them as arguments in the retarded Green's functions, fill in the arguments of the advanced Green's functions with the proper remaining a_j, and add up all such terms to give just one term $\sum_j \cdot (\prod_{l=1}^{0} \equiv 1.)$

[22] See Appendix F.

[23] P. F. Byrd and M. D. Friedman, *Handbook of Elliptic Integrals* (Springer, Berlin, 1954).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 5 MAY 1970

Breaking of Euclidean Symmetry with an Application to the Theory of Crystallization

GÉRARD G. EMCH, HUBERT J. F. KNOPS*
Department of Physics and Astronomy, The University of Rochester, Rochester, New York
AND
EDWARD J. VERBOVEN
Instituut voor theoretische Fysica, Universiteit Nijmegen, The Netherlands

(Received 27 October 1969)

We present a systematic study of the processes by which the original Euclidean invariance of a quantum statistical theory can be broken to produce pure phases with a lower symmetry. Our results provide a rigorous basis for Landau's argument on the nonexistence of critical point in the liquid–solid phase transition. A classification of the possible residual symmetries is obtained, and its connection with spectral and cluster properties is established. Our tools are those of the algebraic approach to statistical mechanics; in particular, we make an extensive use of the KMS condition. None of our proofs involves the separability of the algebra of quasilocal observables.

INTRODUCTION

The problem to which we address ourselves in this paper is that of giving, within the realm of statistical mechanics, a consistent interpretation of the following empirical fact: Whereas all fundamental interactions so far known are at least Euclidean invariant, matter does exist in crystalline equilibrium states which exhibit only a much lower symmetry. A proper understanding of this fact is a prerequisite to any theory of the liquid–solid phase transition.

Therefore, we are faced with the problem of finding a mechanism whereby the original Euclidean symmetry of the theory can be broken in such a manner that the residual symmetry is that of a crystalline

[2] In its simplest form, such a generalization goes back to P. A. M. Dirac.

[3] The support of a generalized function is the smallest closed set of points outside of which it vanishes. For concepts concerning linear functionals, consult L. Schwartz, *Théorie des distributions* (Hermann, Paris, 1957), I. M. Gel'fand and G. E. Shilov, *Generalized Functions* (Academic, New York, 1964), Vol. I, Chap. I, Sec. 1, and W. Güttinger, Fortschr. Physik **14**, 483 (1966).

[4] Step function

$$\theta(x-a) = 1, \quad \text{for} \quad x_0 - a_0 > |x_k - a_k|,$$
$$= 0, \quad \text{for} \quad x_0 - a_0 < |x_k - a_k|.$$

[5] W. C. Schieve, thesis, Lehigh University, Bethlehem, Pa., 1959.

[6] The generalized Green's functions dealt with in this investigation are defined with the help of Lorentz-invariant Riesz distributions. See M. Riesz, Acta Math. **81**, 1 (1949).

[7] B. Bertotti and J. Plebański, Ann. Phys. (N.Y.) **11**, 169 (1960).

[8] Obviously, GGF's can be defined for other hyperbolic differential equations as well. One just has to substitute for $D^R(x)$ the corresponding solution in Eq. (2.1). We deal here, however, only with the wave equation.

[9] See W. Güttinger, Ref. 3, p. 505, and L. Hörmander, *Linear Partial Differential Equations* (Springer, Berlin, 1964).

[10] Even if the GGF is a regular generalized function, the operation on it by differential operators [as in Eq. (1.8)] might generate a singular linear functional.

[11] In the following, capital Latin indices run always from 1 to 3.

[12] Consult Appendix A for Eq. (3.6) as well as for the explicit components of H_{AB}^{-1}.

[13] The coordinization (3.13b) does cover only that part of the timelike 2-flat which can be reached from the origin by timelike directions. The other points may be obtained by taking

$$\mu = \rho \cosh\phi, \quad \nu = \rho \sinh\phi, \quad \rho \geq 0, \quad -\infty < \phi < \infty.$$

[14] Compare Eq. (A1).

[15] For scalar argument u,

$$\theta(u) = 1, \quad \text{if} \quad u > 0,$$
$$= 0, \quad \text{otherwise.}$$

[16] For example, if the motion of three or more particles is coplanar, the field point x cannot be chosen to lie in this plane, etc.

[17] H. Stephani, Acta Physica Polon. **26**, 1045 (1964).

[18] It seems worthwhile to note that $D(x) * D(x-a_1) \cdot D(x-a_2)$ —with $D(x)$ being the retarded, advanced, symmetric Green's function or the Feynman propagator—may be evaluated by help of the Fourier representation of $D(x)$. Then

$$D(x) * D(x-a_1)D(x-a_2) = -\frac{1}{16\pi^6}\oint_C d^4k\, \frac{e^{ik_\alpha R_1{}^\alpha}}{k^2} \times \oint_{C'} d^4k'\, \frac{e^{ik'_\mu R_2{}^\mu}}{k'^2(k+k')^2}.$$

The loops C, C' around the poles of $(k^2)^{-1}$, $(k'^2)^{-1}$, and $[(k+k')^2]^{-1}$ have to be chosen properly (64 possibilities!). However, the calculations are much longer and cannot be kept Lorentz invariant in each step.

[19] We define $\prod_{j=l}^{l'}(\ )_j = 1$, if $l' < l$.

[20] In the coordinate system in which $R_{3\alpha} = R_{30}\delta_\alpha^0$, one finds

$$\theta(y_0 - a_0) = \theta[(y_0 - a_0)R_{30}]$$

due to $R_{30} > 0$.

[21] The symbol $\mathfrak{P}_j$ means: Form all combinations of $j+1$ elements out of the given set $x_1, a_1, \cdots, a_n$; use them as arguments in the retarded Green's functions, fill in the arguments of the advanced Green's functions with the proper remaining a_j, and add up all such terms to give just one term $\sum_j \cdot (\prod_{l=1}^{0} \equiv 1.)$

[22] See Appendix F.

[23] P. F. Byrd and M. D. Friedman, *Handbook of Elliptic Integrals* (Springer, Berlin, 1954).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 5 MAY 1970

Breaking of Euclidean Symmetry with an Application to the Theory of Crystallization

GÉRARD G. EMCH, HUBERT J. F. KNOPS*
Department of Physics and Astronomy, The University of Rochester, Rochester, New York
AND
EDWARD J. VERBOVEN
Instituut voor theoretische Fysica, Universiteit Nijmegen, The Netherlands

(Received 27 October 1969)

We present a systematic study of the processes by which the original Euclidean invariance of a quantum statistical theory can be broken to produce pure phases with a lower symmetry. Our results provide a rigorous basis for Landau's argument on the nonexistence of critical point in the liquid–solid phase transition. A classification of the possible residual symmetries is obtained, and its connection with spectral and cluster properties is established. Our tools are those of the algebraic approach to statistical mechanics; in particular, we make an extensive use of the KMS condition. None of our proofs involves the separability of the algebra of quasilocal observables.

INTRODUCTION

The problem to which we address ourselves in this paper is that of giving, within the realm of statistical mechanics, a consistent interpretation of the following empirical fact: Whereas all fundamental interactions so far known are at least Euclidean invariant, matter does exist in crystalline equilibrium states which exhibit only a much lower symmetry. A proper understanding of this fact is a prerequisite to any theory of the liquid–solid phase transition.

Therefore, we are faced with the problem of finding a mechanism whereby the original Euclidean symmetry of the theory can be broken in such a manner that the residual symmetry is that of a crystalline

group. We propose here such a mechanism and show that our interpretation provides a rigorous basis for Landau's argument according to which there is *no* critical point in the liquid–solid phase transition.

We work in the thermodynamical limit of a quantum system originally enclosed in a finite box. From the knowledge available in the case of "finite statistical mechanics," we infer that the following situation prevails in the thermodynamical limit: (i) the time evolution commutes with all rigid transformations of the 3-dimensional real space; (ii) the usual (grand) canonical ensemble theory provides, via the traditional limiting procedure, a (grand) canonical equilibrium state for the infinite system; (iii) this (grand) canonical equilibrium state is invariant with respect to the time evolution and the entire Euclidean group of transformation of the 3-dimensional real space; and (iv) this state satisfies the Kubo–Martin–Schwinger condition.

We refine the concept of pure phase and show that there exists, under very general conditions, a unique decomposition of the (grand) canonical equilibrium state in its pure phase components (Sec. 1). We study the residual symmetry inherited from Euclidean symmetry when this decomposition is performed (Secs. 2 and 3), and we analyze the related cluster properties in space (Sec. 4). The relevence of our results, for the understanding of the liquid–solid phase transition, is then underlined in Sec. 5.

For general orientation purposes, we now briefly review the antecedents to our Secs. 2 and 3. More detailed references will be provided in the main text. Kastler and Robinson[1] restrict their attention to the translation group $\mathbb{R}^3$ only, and distinguish between three classes of extremal $\mathbb{R}^3$ invariant states: E_{I}, E_{II}, and E_{III}. Their classification is based on the properties of the discrete part of the spectrum of the momentum associated with the state considered. They then notice that their E_{II} states present the particularity that they can be uniquely decomposed into states which are extremal invariant with respect to a discrete (evidently normal) subgroup of $\mathbb{R}^3$. An extension of their results was worked out by Robinson and Ruelle[2] who, however, place themselves in a position where they can only analyze the cases where the residual symmetry is a *normal* subgroup of the Euclidean group, a situation which is unacceptable from the crystallographer's point of view. Moreover, the physicist might object to the fact that the preceding investigations are purely kinematical and, therefore, ignore the dynamical aspect of the theory. The "central decomposition" carried out later by Kastler, Haag, and Michel[3] was at the origin of the considerations of our Sec. 1, where we replace the rather dubious thermodynamical arguments of these authors by a systematic exploitation of the KMS condition. We further show that the physically undesirable separability assumptions essential to their considerations can be bypassed by the assumption that the (grand) canonical thermodynamical equilibrium state is locally normal. Finally, our analysis of the symmetry breaking process goes much further than theirs (see, in particular, Secs. 4 and 5). We also could mention that our earlier contributions[4,5,6] now appear as particular cases of the theory presented here. In this respect, we might remark that, whereas the generalizations considered here are certainly genuine from a mathematical point of view, their actual physical value will ultimately depend on the detailed study of the models. We, nevertheless, want to point out that the paucity of the information available to date on the specifics of the interactions responsible for crystallization makes it worthwhile to give the theory in as general a form as possible, so as to offer maximal freedom to the prospective model builders.

1. KMS STATES

The importance of the Kubo–Martin–Schwinger[7,8] boundary conditions for the rigorous theory of equilibrium statistical mechanics has recently been the object of many speculations.[9–19] In this section, we first recall the definition of this condition and then sum up those of its consequences which we use in the sequel.

Let $\mathcal{A}$ be a C^*-algebra with unit and $'\mathfrak{S}$ be the set of all states on $\mathcal{A}$; an element ϕ of $\mathfrak{S}$ is said to satisfy the KMS *condition* for the natural temperature β $(= 1/kT)$ if

$$\int dt f_{-\beta}(t)\langle\phi; B\alpha_t[A]\rangle = \int dt f_0(t)\langle\phi; \alpha_t[A]B\rangle$$

holds for all A, B in $\mathcal{A}$ and all "test functions" $\hat{f}$ in $\mathfrak{D}$ (i.e., $\hat{f}$ is infinitely differentiable and has compact support). In the above condition, α_t denotes the automorphism of $\mathcal{A}$ describing the time evolution $A \to \alpha_t[A]$, and the function $f_\gamma(t)$ is defined, for each $\hat{f}$ in $\mathfrak{D}$, by

$$f_\gamma(t) \equiv \int d\omega \hat{f}(\omega) e^{i\omega(t+i\gamma)}.$$

We denote by $\mathfrak{S}_\beta$ the set of all states satisfying the KMS condition for the natural temperature β. The fundamental physical assumption of this paper is to interpret $\mathfrak{S}_\beta$ as the set of *all equilibrium states* corresponding to the natural temperature β. This identification rests on the following facts.

First, upon writing $B = I$ in the KMS condition, we conclude from the continuity of $\langle \phi; \alpha_t[A] \rangle$ in t and from the fact that this function is bounded (namely by $\|A\|$) that any state ϕ satisfying the KMS condition is invariant in time.

Second, in the traditional (i.e., finite) quantum statistical mechanics, the KMS condition is satisfied by the canonical equilibrium state $\langle \phi; A \rangle = \operatorname{Tr} e^{-\beta H} A / \operatorname{Tr} e^{-\beta H}$, and this is the only state satisfying the KMS condition. The second part of the latter assertion has to be understood as follows: We suppose that $\mathcal{A}$ is the C^*-algebra of the canonical (anti) commutation relations for a finite region of the space $\mathbb{R}^3$; if ϕ is supposed to be such that only a finite number of particles can simultaneously be in this "box," then [20,21] ϕ can be represented by a density matrix in Fock space. Furthermore, since $\mathcal{A}$ acts irreducibly on Fock space, ρ is uniquely determined by ϕ. If now ϕ is supposed to satisfy the KMS condition, one sees easily that ρ is $e^{-\beta H}/\operatorname{Tr} e^{-\beta H}$ where H is the generator of the time evolution. Hence, in finite statistical mechanics, the KMS condition determines uniquely ϕ as soon as α_t is given. We should notice in this connection that the grand canonical formalism can be obtained in the same way by replacing in the above reasoning H by $H - \mu N$.

Third, in all cases which have been treated explicitly,[11,16,22] the thermodynamical Gibbs state (for a truly infinite system) still satisfies the KMS condition.

Once the use of the KMS condition as a mean to recognize equilibrium states corresponding to the natural temperature β is agreed upon, the identification of *pure phases* is straightforward: These are the equilibrium states (corresponding to the natural temperature β) which cannot be decomposed into equilibrium states (with the same β); hence we interpret the set $\mathfrak{E}_\beta$ of the extreme points of $\mathfrak{S}_\beta$ as the set of all pure phases existing at the natural temperature β. There are some immediate points of contact between this interpretation and some of the properties one would naturally expect to hold for a pure phase. For instance, we saw above that in the case of finite statistical mechanics, $\mathfrak{S}_\beta$ consists of only one state, the latter being then extremal in $\mathfrak{S}_\beta$ and accordingly interpreted as the unique pure phase available in this situation. This fact is related to Yang and Lee's result[23] according to which the thermodynamical limit $(V, N \to \infty V/N = cte)$ is an essential ingredient to the proper description of coexisting pure phases. We shall, furthermore, indicate later in this section that states in $\mathfrak{E}_\beta$ satisfy some cluster properties in space which reinforce the interpretation of these states as pure phases; ergodicity (in time) will also be touched in the course of this section.

In order to make this paper reasonably self-contained, we now want to review the properties of $\mathfrak{S}_\beta$ which we shall use in the sequel and present them in as economical a way as possible. The reader will notice that all results stated below are established without assuming $\mathcal{A}$ to be separable.

We first notice that $\mathfrak{S}_\beta$ is a convex set. Furthermore, upon using the norm continuity of $\alpha_t[A]$ and the fact that $f_\gamma(t)$ is absolutely integrable for every $\hat{f}$ in $\mathfrak{D}$, one can prove straightforwardly that the integrals appearing in the KMS condition are continuous on $\mathfrak{S}$, the latter being equipped with the topology it inherits from the w^* topology[24] of $\mathcal{A}^*$. From this fact it follows that $\mathfrak{S}_\beta$ is closed in this topology; $\mathfrak{S}_\beta$ is bounded in the metric topology and is hence w^* compact. Then, we get from the Krein–Milman theorem:

Lemma 1.1: $\mathfrak{S}_\beta$ is the w^* closed convex hull of $\mathfrak{E}_\beta$.

Physically this means that there exists always enough pure phases (in $\mathfrak{E}_\beta$) so that their ordinary mixtures (i.e., finite convex combinations) can approximate any equilibrium state (in $\mathfrak{S}_\beta$) as close as one wishes. One of the aims of this paper is to establish a canonical decomposition of any ϕ in $\mathfrak{S}_\beta$ into its pure phase components.

The next lemma is a slight generalization of results found in Araki[13] and Araki and Miyata.[14]

Lemma 1.2: Let $\mathcal{A}$ be a concrete C^*-algebra acting on some Hilbert space $\mathcal{H}$, $U(\mathbb{R})$ be a weakly continuous representation of $\mathbb{R}$ on $\mathcal{H}$, and let E_0 denote the projector on the subspace of all vectors of $\mathcal{H}$ which are invariant with respect to $U(\mathbb{R})$. For every (normalized) vector Ψ in $E_0\mathcal{H}$, we form the state ψ on $\mathcal{B}(\mathcal{H})$ defined by $\langle \psi; B \rangle = (\Psi, B\Psi)$ and denote by the same symbol ψ its restriction to $\mathcal{A}''$, $\mathcal{A}$, or $E_0\mathcal{A}''E_0$. If the automorphisms $B \to \alpha_t[B] = U(t)BU(-t)$ map $\mathcal{A}$ onto itself, and if ψ is KMS on $\mathcal{A}$ with respect to α_t, then

(i) ψ is KMS on $\mathcal{A}''$,

(ii) ψ is a vector trace on the von Neumann algebra $E_0\mathcal{A}''E_0$.

If, furthermore, Ψ is cyclic for $\mathcal{A}$ in $\mathcal{H}$, then

(iii) ψ is a faithful trace on $E_0\mathcal{A}''E_0$.

Proof: (i) follows directly from the KMS condition and Kaplanski's density theorem (for details, see Araki,[13] Lemma 2.4; one should notice that his proof indeed does not depend on the cyclicity of Ψ). For (ii) we first notice that $\alpha_t[\mathcal{A}] = \mathcal{A}$ for all t in $\mathbb{R}$ implies that $E_0\mathcal{A}''E_0 = E_0\mathcal{N}E_0$, where $\mathcal{N}$ is the von Neumann

algebra generated by $\mathcal{A}$ and $U(\mathbb{R})$; since $E_0 \in \mathcal{N}$, $E_0\mathcal{A}''E_0$ is then indeed a von Neumann algebra. We next recall from Araki and Miyata[14] (see their Theorems 2.1 and 3.4, with the same remark as above on the irrelevance of the cyclicity of Ψ) that the von Neumann algebra $\mathcal{B} = \mathcal{A}'' \cap U(\mathbb{R})'$ satisfies the following conditions:

(a) $\langle \psi; [A, B] \rangle = 0$ for all A in $\mathcal{A}''$ and B in $\mathcal{B}$

(which can be seen immediately from the KMS condition);

(b) $E_0\mathcal{B} \subseteq E_0\mathcal{A}''E_0$ (which is trivial);

(c) $E_0\mathcal{B} \supseteq E_0\mathcal{A}''E_0$ (which can be obtained from the fact that the time average $\eta: \mathcal{A}'' \to \mathcal{B}$ satisfies $\eta(A)E_0 = E_0AE_0$). We have then from (c) that every X in $E_0\mathcal{A}''E_0$ can be written as $X = YE_0 = E_0Y$ with Y in $\mathcal{B}$, so that $[X^*, X] = E_0[Y^*, Y]E_0$; we have then

$$\langle \psi; [X^*, X] \rangle = (\Psi, E_0[Y^*, Y]E_0\Psi) = \langle \psi; [Y^*, Y] \rangle = 0$$

due to (a). Hence, ψ is indeed a vector trace on $E_0\mathcal{A}''E_0$. If now Ψ is cyclic for $\mathcal{A}$ in $\mathcal{H}$, it is cyclic and, hence,[25] separating for $E_0\mathcal{A}''E_0$. Consequently, $\langle \psi; X^*X \rangle = 0$ for X in $E_0\mathcal{A}''E_0$ implies $\|X\Psi\| = 0$ and, hence, $X = 0$, which is to say that ψ is a faithful trace on $E_0\mathcal{A}''E_0$. QED

Lemma 1.3: Let $\mathcal{A}$ be a C^*-algebra and ϕ, ψ in $\mathfrak{S}_\beta$ with $\psi \leq \lambda\phi$ for some $\lambda \geq 1$; denote by π_ϕ the GNS representation associated to ϕ and by Φ the corresponding cyclic vector. Then

(i) there exists a unique positive Z in the center of $\pi_\phi(\mathcal{A})''$ such that $\langle \psi; A \rangle = (\Phi, Z\pi_\phi(A)\Phi)$ for all A in $\mathcal{A}$;

(ii) ϕ belongs to $\mathfrak{E}_\beta$ if and only if $\pi_\phi(\mathcal{A})''$ is a factor.

Proof: Since ϕ and ψ satisfy the KMS condition, they are invariant with respect to the time evolution entering in this condition; there exists then a unique positive X in $\mathcal{N}'_\phi \equiv \pi_\phi(\mathcal{A})' \cap U_\phi(\mathbb{R})'$ such that $\langle \psi; A \rangle = (\Phi, X^2\pi_\phi(A)\Phi)$ for all A in $\mathcal{A}$. Φ and $\Psi \equiv X\Phi$ satisfy then the assumptions of Lemma 1.2 so that ψ is a vector trace on $E_\phi\pi_\phi(\mathcal{A})''E_\phi$ and is dominated by the faithful vector trace ϕ on this algebra. There is, therefore,[26] a unique positive element Z_0 in the center $\mathfrak{Z}$ of $E_\phi\pi_\phi(\mathcal{A})''E_\phi$ such that $\langle \psi; P \rangle = \langle \phi; Z_0P \rangle$ for all (positive) P in $E_\phi\pi_\phi(\mathcal{A})''E_\phi$. Since $E_\phi\pi_\phi(\mathcal{A})''E_\phi = E_\phi\mathcal{N}_\phi E_\phi$, we have[27] that

$$\mathfrak{Z} = E_\phi(\mathcal{N}_\phi \cap \mathcal{N}'_\phi)E_\phi = (\mathcal{N}_\phi \cap \mathcal{N}'_\phi)E_\phi = (\pi_\phi(\mathcal{A})'' \cap \pi_\phi(\mathcal{A})' \cap U(\mathbb{R})')E_\phi = \mathfrak{Z}_\phi(\mathcal{A})E_\phi.$$

The last of the above equalities is due to the fact[28] that the center $\mathfrak{Z}_\phi(\mathcal{A})$ of $\pi_\phi(\mathcal{A})''$ is left element-wise invariant by $\tilde{\alpha}_t$, the natural extension of α_t to $\pi_\phi(\mathcal{A})''$. Furthermore, from the cyclicity of Φ for $\pi_\phi(\mathcal{A})$ (and, hence, for $\mathfrak{N}_\phi$) it follows that the mapping

$$X[\in \mathfrak{N}'_\phi] \to XE_\phi[\in \mathfrak{N}'_\phi E_\phi]$$

is an isomorphism. There exists then a unique positive element Z in $\mathfrak{Z}_\phi(\mathcal{A})$ such that

$$(\Psi, E_\phi\pi_\phi(A)E_\phi\Psi) = (\Phi, ZE_\phi\pi_\phi(A)E_\phi\Phi) \quad \text{for all } A \text{ in } \mathcal{A}.$$

Since Ψ and Φ are in $E_\phi\mathcal{H}_\phi$ and Z (in $\mathfrak{N}'_\phi$) commutes with E_ϕ (in $\mathfrak{N}_\phi$), the above equality reduces indeed to the first statement of the lemma. The second statement follows then from the first one by the usual *ad absurdo* arguments. QED

The first part of this lemma will be instrumental in the mathematical reasoning to be presented below. The second part has been known[29] (though not proved quite satisfactorily) for some time; it gives a mathematically handy criterion for recognizing pure phases: *An equilibrium state is a pure phase if and only if it is primary.* The latter observation has some immediate physical consequences which we now want to mention, as they confirm further our interpretation of extremal KMS states as pure phases. Let, indeed, G be any amenable symmetry group leaving invariant the pure phase ϕ and acting η asymptotically[30] on $\mathcal{A}$. Then the pure phase ϕ has the cluster property

$$\eta^G\langle \phi; B\alpha_g[A]C \rangle = \langle \phi; A \rangle\langle \phi; BC \rangle$$

for every A, B, C in $\mathcal{A}$. Furthermore, the pure phase ϕ is extremal invariant with respect to G, which is to say that a pure phase cannot be a mixture of states invariant under the same symmetry group G, provided that the latter acts η asymptotically on $\mathcal{A}$.[31] However, in spite of the heuristic argument sketched in Ruelle,[32] we do not see any compelling reason for believing that $\eta^G_\phi[A] = \langle \phi; A \rangle I$ implies, in general, that ϕ is a pure phase. Actually, we will see that whenever crystals do exist, there is at least one extremal $\mathbb{E}^3$-(or $\mathbb{R}^3$) invariant state which is not a pure phase.

We might notice that this remark also bears some incidence on ergodicity. Specifically, if the time evolution were to act η asymptotically on $\mathcal{A}$, then

$$\eta^T\langle \phi; B^*\alpha_t[A]B \rangle = \langle \phi; A \rangle$$

for all A and B in $\mathcal{A}$ with $\langle \phi; B^*B \rangle = 1$; i.e., "local perturbations would die out in time" and ϕ would be ergodic, and, in particular, extremal T invariant. We now want to emphasize that the latter property rests, however essentially, on some weakened version of

η-asymptotic behavior in time. Suppose, indeed, that the equilibrium state ϕ is extremal T invariant (and not only extremal KMS). The fact that ϕ satisfies the KMS condition implies rather trivially that the vector Φ is separating for $\pi_\phi(\mathcal{A})''$; this, together with the fact that ϕ is extremal T invariant, implies in turn that the projector E_ϕ on the T-invariant vectors of $\mathcal{H}_\phi$ is 1-dimensional. Consequently, $\eta_\phi^T[A]\Phi = E_\phi \eta_\phi^T[A]\Phi = \lambda\Phi$ for all A in $\mathcal{A}$; upon using again the fact that Φ is separating for $\pi_\phi(\mathcal{A})''$, we get $\eta_\phi^T[A] = \langle\phi; A\rangle I$. This now implies, indeed, the following weakened version of η asymptotic behavior in time:

$$\eta^T\langle\phi; C^*[\alpha_t[A], B]C\rangle = 0, \quad A, B, C \in \mathcal{A};$$

this condition is referred to by saying that T acts η-asymptotically on $\pi_\phi(\mathcal{A})$. We now anticipate the final result of this section, which is that every equilibrium state can be "decomposed" into pure phases. If it were true that every pure phase were not only extremal KMS but also extremal T invariant, then one would see that the time evolution would have to act η-asymptotically on every ϕ in $\mathfrak{S}_\beta$. This might be an appealing assumption, although there is little evidence for it from exactly soluble models. We shall therefore *not* assume in this paper this otherwise reasonable condition.[33]

To show the existence of a canonical decomposition of an equilibrium state into its pure-phase components we shall need the mathematical concept of "central measure," which we now recall with the generality required by our purpose.

Let $\mathcal{A}$ be a C^*-algebra, π_ϕ be the representation associated to $\phi \in \mathfrak{S}$, and $\mathfrak{Z}_\phi(\mathcal{A})$ be the center of $\pi_\phi(\mathcal{A})''$. We denote by $\tilde{\phi}$ the canonical extension of ϕ to $\pi_\phi(\mathcal{A})''$. There exists a unique[34,35] measure on $\mathfrak{S}$, called the *central measure* of ϕ and denoted μ_ϕ, such that there exists a σ-continuous isomorphism φ from $\mathfrak{Z}_\phi(\mathcal{A})$ onto $\mathfrak{L}^\infty(\mathfrak{S}, \mu_\phi)$ with the property

$$\langle\tilde{\phi}; ZA\rangle = \int \varphi_Z(\psi)\langle\psi; A\rangle \, d\mu_\phi(\psi)$$

for all A in $\mathcal{A}$ and all Z in $\mathfrak{Z}_\phi(\mathcal{A})$. Furthermore,[35] this measure is concentrated in the Baire sense on the set $\mathcal{F}$ of all primary states of $\mathcal{A}$; i.e., $\mu_\phi(\mathcal{C}) = 0$ for every Baire set[36] $\mathcal{C}$ in $\mathfrak{S}$ with $\mathcal{C} \cap \mathcal{F} = \varnothing$. Finally, if we denote by $\pi_\psi : \mathcal{A} \to \mathcal{B}(\mathcal{H}_\psi)$ the GNS representation associated to any ψ in $\mathfrak{S}$, we recall[35] that there exists an isometric mapping U from $\mathcal{H}_\phi$ onto $\mathcal{H}_\mu = \int \mathcal{H}_\psi \, d\mu_\phi(\psi)$ which establishes a unitary equivalence between π_ϕ and the representation π_μ[37] and transforms the center $\mathfrak{Z}_\phi(\mathcal{A})$ of $\pi_\phi(\mathcal{A})''$ into the von Neumann algebra of all diagonal operators[38] on $\mathcal{H}_\mu$. These facts show the connection between the above decomposition of ϕ and the familiar central disintegration of a representation.[39] It should be noted, however, and this is the fundamental contribution of Wils,[35] that the definition of the central measure, its existence and uniqueness, as well as the above-mentioned properties do *not* depend on the separability of $\mathcal{A}$. This is important, as $\mathcal{A}$ turns out to be nonseparable in several physical situations.[40] The price paid for this generality is that μ_ϕ is concentrated on $\mathcal{F}$ in the Baire sense only, rather than in the Borel sense.[41] We have to compensate for this by the condition that the states with which we start our analysis are "locally normal." We shall come to this later. First, we want to establish some results showing the relevance for our purpose of the concept of central measure.

Lemma 1.4: Every state in the support of the central measure of a state in $\mathfrak{S}_\beta$ also belongs to $\mathfrak{S}_\beta$.

Proof: Let F be any μ_ϕ measurable subset of $\mathfrak{S}$ and χ_F in $\mathfrak{L}^\infty(\mathfrak{S}, \mu_\phi)$ be its characteristic function. Define for all A in $\mathcal{A}$

$$\begin{aligned}\langle\phi_F; A\rangle &= \int_{\mathfrak{S}} \chi_F(\psi)\langle\psi; A\rangle \, d\mu_\phi(\psi) \\ &= \langle\tilde{\phi}; \varphi^{-1}(\chi_F)A\rangle \\ &= (\Phi, (\varphi^{-1}(\chi_F))\pi_\phi(A)\Phi).\end{aligned}$$

Since $(\varphi^{-1}(\chi_F))$ belongs to $\mathfrak{Z}_\phi(\mathcal{A})$, it is then invariant[28] under the natural extension $\tilde{\alpha}_t$ of α_t to $\pi_\phi(\mathcal{A})''$; ϕ_F is therefore T invariant. From Lemma 1.2(i), the extension $\tilde{\phi}$ of ϕ to $\pi_\phi(\mathcal{A})''$ again satisfies the KMS condition, and hence so does ϕ_F, whatever F μ_ϕ measurable in $\mathfrak{S}$ is; the conclusion of the lemma follows then straightforwardly from Fubini's theorem and from the fact that $\int f_\gamma(t)\langle\psi; B\alpha_t[A]\rangle \, dt$ is weak* continuous in ψ. QED

In plain language, this lemma means that we do not lose the KMS property in the course of the above "central decomposition." We now want to make use of Choquet's theory.[42] For the sake of completeness, we first recall some of the definitions pertinent to this part of our investigation.

Let $\mathfrak{C}$ (e.g., $\mathfrak{S}$ or $\mathfrak{S}_\beta$; for the latter, see Lemma 1.1) be a compact convex subset of a locally convex linear space $\mathfrak{X}$ (e.g., $\mathcal{A}^*$). A real-valued function f on $\mathfrak{C}$ is said to be *convex* (resp. *affine*) if

$$\begin{aligned}f[\lambda\phi_1 + (1-\lambda)\phi_2] \leq (\text{resp. } =)\ \lambda f[\phi_1] \\ + (1-\lambda)f[\phi_2], \quad \text{for all } \phi_1, \phi_2,\end{aligned}$$

in $\mathfrak{C}$ and $0 \leq \lambda \leq 1$. For instance, the function $f[\phi] = \langle\phi; A\rangle$ defined on $\mathfrak{S}$ by any self-adjoint A in

$\mathcal{A}$ is affine. We denote by $C(\mathfrak{C})$ [resp. $A(\mathfrak{C})$] the set of all convex (resp. affine) continuous functions on $\mathfrak{C}$. For any measure μ on $\mathfrak{C}$, and any f in $C(\mathfrak{C})$, we denote by $\langle \mu; f\rangle$ the integral $\int f(\psi)\, d\mu(\psi)$. Let μ and ν be two measures on $\mathfrak{C}$. We say that μ and ν have the *same resultant*, which we denote $\mu \sim \nu$, whenever $\langle \mu; a\rangle = \langle \nu; a\rangle$ for all a in $A(\mathfrak{C})$. In particular, we say that μ *represents* a point ϕ in $\mathfrak{C}$ whenever $\mu \sim \delta_\phi$ [i.e., $\int a(\psi)\, d\mu(\psi) = a(\phi)$ for all a in $A(\mathfrak{C})$]. ν is said to be *majorized* by μ, a fact which we denote $\nu \prec \mu$, if $\langle \nu; f\rangle \leq \langle \mu; f\rangle$ for all f in $C(\mathfrak{C})$. This introduces a *partial ordering* on the set of all measures on $\mathfrak{C}$. A measure μ is said to be *maximal* if it is maximal with respect to this ordering. Our interest in maximal measures arises heuristically from the feeling that $\nu \prec \mu$ should imply that the support of μ is closer to the extreme points of $\mathfrak{C}$ than is the support of ν, making then maximal measures candidates for the decomposition of an equilibrium state into its pure-phase components.

We say[43] that a nonempty compact convex subset $\mathfrak{C}$ of a locally convex linear space $\mathfrak{X}$ is a *simplex* if for each ϕ in $\mathfrak{C}$ there exists a unique maximal measure representing ϕ.

Theorem 1.1: The set $\mathfrak{S}_\beta$ of all states satisfying the KMS condition for the natural temperature β is a simplex, and, for each ϕ in $\mathfrak{S}_\beta$, the unique maximal measure μ_ϕ representing ϕ on $\mathfrak{S}_\beta$ is the central measure of this state.

Proof: Since $\mathfrak{S}_\beta$ is clearly bounded, we conclude from Lemma 1.1 that $\mathfrak{S}_\beta$ is a compact convex subset of the locally convex linear space $\mathcal{A}^*$. Let ϕ be any state in $\mathfrak{S}_\beta$, μ_ϕ be its central measure, and ν be any measure on $\mathfrak{S}_\beta$ representing ϕ. We want to show that $\nu \prec \mu_\phi$. For every f in $C(\mathfrak{S}_\beta)$ and $\epsilon > 0$, there exists[44] a discrete measure $\nu_{f,\epsilon} = \sum_{i=1}^n \alpha_i \delta_{\phi_i}$ (with $\alpha_i > 0$ $\sum_i \alpha_i = 1$, $\phi_i \in \mathfrak{S}_\beta$) which represents ϕ and satisfies $\langle \nu; f\rangle - \langle \nu_{f,\epsilon}; f\rangle > \epsilon$. Since $\nu_{f,\epsilon}$ represents ϕ, we have, in particular, for all A in $\mathcal{A}$ (considered as affine functions on $\mathfrak{S}_\beta$),

$$\langle \phi; A\rangle = \langle \nu_{f,\epsilon}; A\rangle = \langle \sum_i \alpha_i \phi_i; A\rangle,$$

i.e., $\phi = \sum_i \alpha_i \phi_i$. From Lemma 1.3 we conclude that there exists, for each ϕ_i in this decomposition, a unique positive Z_i in $\mathfrak{Z}_\phi(\mathcal{A})$ such that

$$\langle \phi_i; A\rangle = (\Phi, Z_i \pi_\phi(A)\Phi) \quad \text{for all } A \text{ in } \mathcal{A}.$$

Clearly $\sum_{i=1}^n Z_i = I$. Upon using the decomposition of ϕ with respect to its central measure μ_ϕ, we get $\phi_i = \int_{\mathfrak{S}_\beta} \psi\, d\mu_i(\psi)$ with $d\mu_i = \varphi_{z_i}\, d\mu_\phi$, where φ is the homomorphism from $\mathfrak{Z}_\phi(\mathcal{A})$ onto $\mathfrak{L}^\infty(\mathfrak{S}, \mu_\phi)$ associated to ϕ. For every convex continuous function f on $\mathfrak{S}_\beta$ and every $\epsilon > 0$, we have then

$$\begin{aligned}\langle \nu; f\rangle - \epsilon \leq \langle \nu_{f,\epsilon}; f\rangle &= \sum_{i=1}^n \alpha_i \langle \delta_{\phi_1}; f\rangle = \sum_{i=1}^n \alpha_i f[\phi] \\ &\leq \sum_{i=1}^n \alpha_i \int_{\mathfrak{S}_\beta} f(\psi)\, d\mu_i(\psi) \\ &= \int_{\mathfrak{S}_\beta} \varphi_{(\Sigma_{i=1}^n \alpha_i Z_i)}(\psi) f(\psi)\, d\mu_\phi(\psi) \\ &= \int_{\mathfrak{S}_\beta} f(\psi)\, d\mu_\phi(\psi) = \langle \mu_\phi; f\rangle.\end{aligned}$$

Since $\epsilon > 0$ is arbitrary, we have indeed for every $f \in C(\mathfrak{S}_\beta)$ $\langle \nu; f\rangle \leq \langle \mu_\phi; f\rangle$, where ν is an arbitrary measure on $\mathfrak{S}_\beta$ with $\nu \sim \delta_\phi$. This proves that the central measure μ_ϕ of $\phi \in \mathfrak{S}_\beta$ is the unique maximal measure representing ϕ and defined on $\mathfrak{S}_\beta$. QED

We now proceed further and introduce explicitly the fact that $\mathcal{A}$ is the algebra of all quasilocal observables on the physical system of interest. Specifically, we associate with every cube V in $\mathbb{R}^3$ the Fock space $\mathcal{K}_V^F$, the von Neumann algebra $\mathcal{A}(V)$ of all bounded operators on $\mathcal{K}_V^F$, and the (separable, closed, two-sided) ideal $\mathcal{B}(V)$ of all compact operators on $\mathcal{K}_V^F$. The natural isotony property [namely, $V_1 \subseteq V_2$ implies $\mathcal{A}(V_1) \subseteq \mathcal{A}(V_2)$] makes it possible to define the C^*-inductive limit $\mathcal{A} = \overline{\bigcup \mathcal{A}(V)}$, which is our algebra of *quasilocal observables.*

We may say that a state ϕ on $\mathcal{A}$ is locally normal if, for every cube V, the restriction of ϕ to $\mathcal{A}(V)$ is a normal state. We denote by $\mathfrak{N}$ the set of all locally normal states on $\mathcal{A}$. We recall that $\phi|_{\mathcal{A}(V)}$ normal is equivalent to any of the following conditions: (i) $\phi|_{\mathcal{A}(V)}$ is ultraweakly continuous[45]; (ii) the norm of $\phi|_{\mathcal{B}(V)}$ is one[46]; (iii) there exists a (unique) density matrix ρ_V such that $\langle \phi; A\rangle = \operatorname{Tr} \rho_V A$ for all A in $\mathcal{A}(V)$.[47] As a consequence of property (i) above, and from the fact that the injection $\mathcal{A}(V_1) \subseteq \mathcal{A}(V_2)$ (for $V_1 \subseteq V_2$) is ultraweakly continuous, we see that $\phi|_{\mathcal{A}(V_2)}$ normal implies $\phi|_{\mathcal{A}(V_1)}$ normal. Therefore ϕ belongs to $\mathfrak{N}$ if and only if its restrictions to $\mathcal{A}(V^n)$ are all normal, where V^n denotes the n-fold dilation of a *fixed* cube V and $n = 1, 2, 3, \cdots$.

Lemma 1.5: Let μ be a positive normalized measure on the set $\mathfrak{S}$ of all states on the quasilocal algebra $\mathcal{A}$. If $\mu \sim \delta_\phi$ for some ϕ in $\mathfrak{S}$, then the Borel set $\mathfrak{N}$ has μ measure 1 if and only if ϕ belongs to $\mathfrak{N}$.

Proof: From Lanford and Ruelle,[48,49] we know that the lemma holds when $\mathfrak{N}$ is replaced by

$$\mathfrak{N}_n(V) \equiv \{\phi \in \mathfrak{S} \mid \|\phi|_{\mathcal{B}(V^n)}\| = 1\};$$

the lemma, as stated here, follows then directly from this and from the fact that $\mathfrak{N}$ is the countable intersection of the $\mathfrak{N}_n(V)$. QED

We now come back to our central decomposition. From Theorem 1.1 we know that,for every ϕ in $\mathfrak{S}_\beta$, μ_ϕ is maximal and hence[50] is concentrated, *in the Baire sense*, on the set $\mathfrak{E}_\beta$ of the extreme points of $\mathfrak{S}_\beta$. If $\mathfrak{S}_\beta$ were metrizable (and, in particular, if $\mathcal{A}$ were separable), this would be equivalent to saying that μ_ϕ is concentrated, *in the Borel sense*, on $\mathfrak{E}_\beta$, since then the concepts of Baire and Borel set coincide.[36] The additional restriction $\phi \in \mathfrak{N}$ is physically justified[20,21] by the assumption that the total number of operators for all finite volumes exist; this will now be used to bypass any global assumption of separability on $\mathfrak{S}_\beta$.

Theorem 1.2: (i) The central measure μ_ϕ of any locally normal equilibrium state ϕ (for the natural temperature β) is concentrated, in the Borel sense, on the set $\mathfrak{E}_\beta \cap \mathfrak{N}$ of all locally normal pure phases for this temperature. Furthermore, (ii) μ_ϕ is the only[51] (regular) measure on $\mathfrak{S}_\beta$ which represents ϕ and is concentrated, in the Borel sense, on $\mathfrak{E}_\beta$.

Proof: We adapt to our purpose a similar argument given by Ruelle[49]; our proof differs from his in the fact that we use Theorem 1.1 instead of any assumption of η-asymptotic behavior in time. To this effect we use the remark preceding Lemma 1.5. We first recall[52] that there exists a countable sequence $\{A_i\}$ of self-adjoint elements A_i of $\mathcal{A}$ which separates $\mathfrak{N}$ from $\mathfrak{S}$. We then define the continuous mapping π from $\mathfrak{S}_\beta$ to $\mathbb{R}^\infty$ by

$$\pi[\phi] = \{\langle \phi; A_i \rangle \mid i \in Z^+\}$$

and associate to each ϕ in $\mathfrak{N} \cap \mathfrak{S}_\beta$ the measure ν_ϕ on $\pi[\mathfrak{S}_\beta]$, defined by $\langle \nu_\phi; f \rangle = \langle \mu_\phi; f \circ \pi \rangle$, for all f in $C(\pi[\mathfrak{S}_\beta])$.

Since μ_ϕ is maximal on $\mathfrak{S}_\beta$ (Theorem 1.1), ν_ϕ is maximal on the compact, convex metrizable set $\pi[\mathfrak{S}_\beta]$ and hence is concentrated, in the Borel sense, on the extremal points of $\pi[\mathfrak{S}_\beta]$. μ_ϕ is then concentrated on the inverse image of this set and then, by Lemma 1.5, on $\mathfrak{N} \cap \pi^{-1}[\mathfrak{E}(\pi[\mathfrak{S}_\beta])]$. We now use the fact that $\{A_i\}$ separates $\mathfrak{N}$ from $\mathfrak{S}$ to conclude that $\mathfrak{N} \cap \pi^{-1}[\mathfrak{E}(\pi[\mathfrak{S}_\beta])] \subseteq \mathfrak{N} \cap \mathfrak{S}_\beta$, so that μ_ϕ is concentrated on $\mathfrak{N} \cap \mathfrak{S}_\beta$. This proves the first part of the theorem. To prove the second part, we have only to verify that if $\mu \sim \delta_\phi$ and μ is concentrated, in the Borel sense, on $\mathfrak{E}_\beta$, then $\mu = \mu_\phi$. From the Corollary 9.8 in Phelps,[42] we see that μ is maximal; $\mu = \mu_\phi$ follows then from Theorem 1.1. QED

The last step of this section is to find a condition under which not only μ_ϕ is concentrated on $\mathfrak{E}_\beta \cap \mathfrak{N}$ in the Borel sense but, moreover, the support of μ_ϕ is actually *contained* in $\mathfrak{E}_\beta \cap \mathfrak{N}$. Such a condition turns out to be that ϕ be "transitive" in addition to the previous conditions.

We now suppose that the 3-dimensional Euclidean group $\mathbb{E}^3$ is a symmetry group for the theory, i.e., that there exists for each g in $\mathbb{E}^3$ an automorphism α_g of $\mathcal{A}$ such that $\langle \phi; \alpha_g[A] \rangle$ is continuous in g for each ϕ and each A in $\mathcal{A}$. We assume further that $\alpha_g \alpha_t = \alpha_t \alpha_g$ for all times t and all g in $\mathbb{E}^3$. We denote by α_g^* the action of g in $\mathbb{E}^3$ on $\mathfrak{S}$: $\langle \phi; \alpha_g[A] \rangle \equiv \langle \alpha_g^*[\phi]; A \rangle$. We call *orbit* (with respect to E^3) of a state ψ in $\mathfrak{G}$ the set

$$O_\psi^E = \{\alpha_g^*[\psi] \mid g \in \mathbb{E}^3\}.$$

We now say that a state ϕ in $\mathfrak{S}$ is *transitive* (with respect to $\mathbb{E}^3$) whenever there exists a ψ in $\mathfrak{S}$ such that $\mu_\phi(O_\psi) = 1$. We anticipate from the next section [Theorem (2.1)] that this implies O_ψ compact.

We can now prove the final theorem of this section, namely:

Theorem 1.3: Let ϕ be a transitive, locally normal equilibrium state for the natural temperature β. Then the support of its central measure μ_ϕ is contained in $\mathfrak{E}_\beta \cap \mathfrak{N}$.

Proof: By the definition of transitivity, there exists ψ_0 such that $\mu(O_{\psi_0}) = 1$. ϕ locally normal implies then $O_{\psi_0} \cap \mathfrak{E}_\beta \cap \mathfrak{N} \neq \varnothing$. Suppose, indeed, that this were not the case; since O_{ψ_0} is compact and hence Borel, the preceding theorem would imply that $\mu(O_{\psi_0}) = 0$, which contradicts the assumption. There exists, therefore, at least a ψ in $O_{\psi_0} \cap \mathfrak{E}_\beta \cap \mathfrak{N}$. We have then

$$O_{\psi_0} = O_\psi = \{\alpha_g^*[\psi] g \in \mathbb{E}^3\} \subseteq \mathfrak{E}_\beta \cap \mathfrak{N}.$$

Since $\mu(O_{\psi_0}) = 1$ and O_{ψ_0} is compact and hence closed, we conclude that supp $(\mu_\phi) \subseteq O_{\psi_0}$. QED

Incidentally, we might mention that transitivity is strong enough by itself to allow the dropping of local normality in this theorem.

Theorems1.2(ii) and 1.3 together give the result we want: There exists a unique decomposition of any transitive, locally normal equilibrium state ϕ into its pure-phase components, and this decomposition is provided by the central measure μ_ϕ of ϕ. Furthermore, Theorem 1.3 helps in understanding the physical meaning of the condition of transitivity: All the pure phases ψ which enter into ϕ are obtained from any one

of them, say ψ_0, by a transformation $\psi = \alpha_g[\psi_0]$ with g in $\mathbb{E}^3$, and therefore the symmetry

$$H_\psi^E \equiv \{g \in \mathbb{E}^3 \mid \alpha_g^*(\psi) = \psi\}$$

that ψ inherits from ϕ is a conjugate, with respect to $\mathbb{E}^3$, of the symmetry $H_{\psi_0}^E$ inherited by ψ_0.

2. STRONG TRANSITIVITY AND SYMMETRY BREAKING

In this section we concentrate on the particular class of "strongly transitive" equilibrium states. We first motivate the introduction of this concept in our context and give its precise definition. We then establish a classification of these states, based on the symmetry that their pure-phase components inherit from the Euclidean invariance of the theory.

The analysis of transitive states, started in the preceding section, is now carried further in order to show the necessity of strengthening this notion for a proper understanding of the occurrence of crystalline phases.

We first notice that a Euclidean-invariant, transitive equilibrium (i.e., KMS) state ϕ is necessarily extremal Euclidean invariant. Indeed, ϕ transitive implies that μ_ϕ is $\mathbb{E}^3$ ergodic, i.e., that $\mu_\phi(K)$ is either 0 or 1 for all $\mathbb{E}^3$-invariant, μ_ϕ-measurable subsets K in $\mathfrak{S}_\beta$; the latter condition is, in turn, equivalent[53] to the fact that ϕ is extremal $\mathbb{E}^3$ invariant. This remark suggests the introduction of the *class* P of all states ϕ on $\mathcal{A}$ which are locally normal, extremal invariant with respect to the Euclidean symmetry, and satisfy the KMS condition. We notice incidentally that ϕ in P is either transitive or is such that $\mu_\phi(O_\psi^E) = 0$ for all ψ in $\mathfrak{S}_\beta$. For the time being, we restrict our attention to the class P. This restriction is not of a serious nature since every ($\mathbb{E}^3$-invariant) equilibrium state can be decomposed into extremal $\mathbb{E}^3$-invariant states. We notice that it can be hard to tell, in general, which states "occur" in a decomposition (the support of the measure seems often to be too large); in this case, however, one expects a discrete decomposition (gas–liquid–solid) where this problem does not arise.

The following theorem establishes relations between transitivity and symmetry.

Theorem 2.1: Let ϕ be in P and μ_ϕ denote its central measure. The following three statements are equivalent:

(i) ϕ is transitive;

(ii) $\mathbb{E}^3/H_\psi^E$ is compact for all ψ in supp (μ_ϕ);

(iii) O_ψ^E is compact and homeomorphic to $\mathbb{E}^3/H_\psi^E$ for all ψ in supp (μ_ϕ).

Proof: Clearly (iii) implies (ii). To prove the converse implication, we form the mapping $\bar{\alpha}_\psi$ from $\mathbb{E}^3/H_\psi^E$ onto O_ψ^E defined by $\bar{\alpha}_\psi[\tilde{g}] = \alpha_g^*[\psi]$ where $g\,[\in \mathbb{E}^3] \to \tilde{g}\,[\in \mathbb{E}^3/H_\psi^E]$ is the canonical mapping. (ii) implies (iii) follows then from the fact that $\bar{\alpha}_\psi$ is continuous and bijective. We now prove that (i) implies (ii). To this effect, we notice that there exists a compact subset C of $\mathbb{E}^3$ *and* a countable sequence $\{g_n\}$ of elements in $\mathbb{E}^3$ such that the union of all $C_n = g_n C$ covers $\mathbb{E}^3$. By assumption, there exists at least one ψ_0 such that $\mu_\phi(O_{\psi_0}^E) = 1$. Since the sets $\bar{\alpha}_{\psi_0}[\tilde{C}_n]$ are the images of compact sets under the continuous mapping $\bar{\alpha}_{\psi_0}$, they are compact sets covering $O_{\psi_0}^E$. We recall[54] that ϕ Euclidean invariant implies that the central measure μ_ϕ is invariant under the action of the Euclidean group. Hence $\mu_\phi(O_{\psi_0}^E)$ can differ from zero only when $\mu_\phi(\bar{\alpha}_{\psi_0}[\tilde{C}_n]) = \mu_\phi(\bar{\alpha}_{\psi_0}[\tilde{C}])$ is different from zero. Consider the set $D_1 = C^2 = \{c_1 c_2 \mid c_1, c_2 \in C\}$. If $\mathbb{E}^3/H_{\psi_0}^E$ were not compact, there would exist an element $\tilde{f}_1$ in $\mathbb{E}^3/H_{\psi_0}^E$ which would not be contained in $\tilde{D}_1$. Consider now the compact set $\tilde{D}_2 = \tilde{D} \cup (f_1 D)^\sim$; there would again exist an element $\tilde{f}_2$ in $\mathbb{E}^3/H_{\psi_0}^E$ which would not be contained in $\tilde{D}_2$. We could then construct in this manner an infinite sequence $\{\tilde{f}_0\,(=e), \tilde{f}_1, \tilde{f}_2, \cdots\}$ such that $\tilde{f}_n \notin \tilde{D}_n = \bigcup_{m=0}^{n-1} (f_m D)^\sim$. We now notice[55] that the compact sets $(f_n C_n)^\sim$ would be all disjoint, and so then would the compact sets $\bar{\alpha}_{\psi_0}[f_n C_n)^\sim]$ since $\bar{\alpha}_{\psi_0}$ is bijective. Upon using again the invariance of μ_ϕ under $\mathbb{E}^3$, we conclude that all these sets would have μ_ϕ measure equal to $\mu_\phi(\bar{\alpha}_{\psi_0}[\tilde{C}])$, the latter being different from zero whenever $\mu_\phi(O_{\psi_0}^E) \neq 0$ (as we mentioned above). This would then contradict the boundedness of the measure of $O_{\psi_0}^E$, so that $\mathbb{E}^3/H_{\psi_0}^E$ must be compact. Since the orbit is the image of this compact set under the continuous mapping $\bar{\alpha}_{\psi_0}$, we conclude that this orbit is closed and compact. It coincides, therefore, with the support of μ_ϕ. Every state ψ in supp (μ_ϕ) has for stabilizer a group H_ψ^E which is therefore conjugate to $H_{\psi_0}^E$; consequently, $\mathbb{E}^3/H_\psi^E$ is compact for all ψ in supp (μ_ϕ). This concludes the proof that (i) implies (ii).

We now prove that (iii) implies (i). Since ϕ is Euclidean invariant, we can write[56]

$$\begin{aligned}\langle \phi; A\rangle &= \int_{\mathfrak{S}_\beta} \langle \psi; A\rangle \, d\mu_\phi(\psi) = \eta^E \langle \phi; \alpha_{\hat{g}}[A]\rangle \\ &= \eta^E \int_{\mathfrak{S}_\beta} \langle \psi; \alpha_{\hat{g}}[A]\rangle \, d\mu_\phi(\psi) \\ &= \int_{\mathfrak{S}_\beta} \eta^E \langle \psi; \alpha_{\hat{g}}[A]\rangle \, d\mu_\phi(\psi).\end{aligned}$$

Let, now, F denote an arbitrary Borel subset of $\mathfrak{S}_\beta$ and define ϕ_F by

$$\langle \phi_F; A\rangle = \int_F \eta^E\langle \psi; \alpha_g^\wedge[A]\rangle \, d\mu_\phi(\psi).$$

This positive functional is then invariant with respect to $\mathbb{E}^3$ and is majorized by ϕ. Since the latter is extremal Euclidean invariant, we conclude that $\phi_F = \mu_\phi(F)\phi$; i.e.,

$$\int_F (\eta^E\langle \psi; \alpha_g[A]\rangle - \langle \phi; A\rangle)\, d\mu_\phi(\psi) = 0.$$

We recall that this holds for all Borel sets F in $\mathfrak{S}_\beta$, so that, for each A separately,

$$\eta^E\langle \psi; \alpha_g^\wedge[A]\rangle = \langle \phi; A\rangle, \quad \mu_\phi \text{ almost everywhere.}$$

Since ϕ is locally normal, there exists[52] a countable family $\{A_n\}$ separating ϕ from all other states. Let D_n be the Borel set of those ψ in $\mathfrak{S}_\beta$ for which the above equality holds for $A = A_n$. We have then $\mu_\phi(D_n) = 1$ and, consequently, $\mu_\phi(\bigcap_n D_n) = 1$. Consider now the Borel set $B = \text{supp } (\mu_\phi) \cap \{\bigcap_n D_n\}$; clearly, $\mu_\phi(B) = 1$. Since μ_ϕ is concentrated (see Theorem 1.2) on $\mathfrak{E}_\beta$, there exists at least one extremal KMS state ψ_0 contained in B. Since, in particular, ψ_0 belongs to $\bigcap_n D_n$, we have

$$\langle \phi; A_n\rangle = \eta^E\langle \psi_0; \alpha_g^\wedge[A_n]\rangle, \quad \text{for all } A_n.$$

Since $\{A_n\}$ separates ϕ from all other states, we conclude that $\phi = \eta^E\psi_0$. We now use (iii), specifically that O_ψ^E is compact, to prove that η^E can be written as an integral over $O_{\psi_0}^E$. To this effect, we define the measure ν on $O_{\psi_0}^E$ by

$$\langle \nu; f\rangle = \eta^E\tilde{f},$$

where f runs over the set $\mathfrak{C}(O_{\psi_0}^E)$ of all continuous functions on $O_{\psi_0}^E$ and $\tilde{f}$ in $\mathfrak{C}(\mathbb{E}^3)$ is defined by $\tilde{f}(g) = f(\tilde{\alpha}_{\psi_0}[\tilde{g}])$. The measure ν then satisfies

$$\int_{O^E\psi_0} \langle \psi; A\rangle \, d\nu(\psi) = \eta^E\langle \psi; \alpha_g[A]\rangle = \langle \phi; A\rangle.$$

Hence, ν is a measure which represents ϕ and is concentrated on $\mathfrak{E}_\beta$ ($\supseteq O_{\psi_0}^E$ since $\psi_0 \in \mathfrak{E}_\beta$). From Theorem (1.2), we conclude, therefore, that $\nu = \mu_\phi$, and, in particular, $\mu_\phi(O_{\psi_0}^E) = \nu(O_{\psi_0}^E) = 1$. This concludes the proof of the theorem.[57]

In the sequel we are mainly interested in state ϕ for which the stabilizers H_ψ^E of the pure phases which occur in the decomposition of ϕ contain at least three noncoplanar translations, a necessary condition for H_ψ^E to qualify as a symmetry of crystalline type. This requirement is equivalent to the compactness of $\mathbb{R}^3/H_\psi^R$ ($H_\psi^R \equiv \{\mathbf{a} \in \mathbb{R}^3 \mid g \equiv (I, \mathbf{a}) \in H_\psi^E\}$) which is, however, *not* necessarily implied by the compactness of $\mathbb{E}^3/H_\psi^E$. The next theorem shows in which sense the condition of transitivity has to be strengthened to ensure that $\mathbb{R}^3/H_\psi^R$ is compact. We introduce to this effect the following concept: A state ϕ will be said to be *strongly transitive* if it is transitive and if there exists at least one ψ in supp (μ_ϕ) such that the orbit O_ψ^R of ψ under the action of the group $\mathbb{R}^3$ of all translations is closed.

Theorem 2.2: Let ϕ be in P and μ_ϕ denote its central measure. The following two statements are equivalent:

(i) ϕ is strongly transitive;

(ii) $\mathbb{R}^3/H_\psi^R$ is compact for all ψ in supp (μ_ϕ).

Proof: We first show that (ii) implies (i). Since the compactness of $\mathbb{R}^3/H_\psi^R$ implies that of $\mathbb{E}^3/H_\psi^R$ for all ψ in supp (μ_ϕ), we know from Theorem 2.1 that ϕ is transitive. Furthermore, since O_ψ^R is the image of $\mathbb{R}^3/H_\psi^R$ through a continuous mapping, (ii) implies that O_ψ^R [with ψ in supp (μ_ϕ)] is compact and hence closed. Hence ϕ is indeed strongly transitive. Suppose now that $O_{\psi_0}^R$ is closed and hence compact, for some ψ_0 in supp (μ_ϕ). The first step in our proof of (i) implies (ii) is to construct an invariant measure on $O_{\psi_0}^R$. For each f in $\mathfrak{C}(O_{\psi_0}^R)$, we define

$$\bar{f}(\mathbf{a}) = f(\alpha_{\mathbf{a}}^*[\psi_0]),$$

which is a continuous and bounded function on $\mathbb{R}^3$. Let η^R denote an invariant mean over $\mathbb{R}^3$ and define $\langle \nu; f\rangle = \eta^R\bar{f}$; this yields a finite Radon measure on $O_{\psi_0}^R$ which is invariant under $\mathbb{R}^3$. An argument similar to the one used in the proof of Theorem 2.1 leads to the conclusion that $\mathbb{R}^3/H_{\psi_0}^R$ is compact for all ψ in supp (μ_ϕ). QED

As announced in the beginning of this section, we shall restrict our attention to strongly transitive states. We denote by SP the subclass of P consisting of all strongly transitive states. Since an SP state ϕ satisfies in particular the assumptions of Theorem 1.3, we know that the respective stabilizers of the pure phases occurring in the decomposition of ϕ are conjugate to one another. We shall refer to this conjugate class as the *intrinsic symmetry* of ϕ. We now undertake the classification of SP states according to their possible intrinsic symmetry.

The first possibility is that the SP state ϕ is already primary and, hence, is a pure phase. In this case its intrinsic symmetry is the full Euclidean group $\mathbb{E}^3$, (e.g., gas and liquid). When this occurs, we say that ϕ belongs to the class SP_1.

A second possibility is that the rotation symmetry is broken in the decomposition, but that $H_\psi^R = \mathbb{R}^3$; i.e., the pure phases still possess the original translation symmetry. This occurs in a system exhibiting spontaneous magnetization.[58] We denote this type of states by SP_2.

In the third place, it can happen that H_ψ^R is continuous in one or two dimensions and discrete in the remaining dimension (rotation symmetry is then certainly broken). This type of states is denoted by SP_3.

Finally, there is the possibility that H_ψ^R is discrete (in three dimensions). Since ϕ is an SP state, we recall that H_ψ^R are closed and $\mathbb{R}^3/H_\psi^R$ are compact for all ψ in supp (μ_ϕ). From these facts it follows that each H_ψ^R is generated by three noncoplanar translations (i.e., H_ψ^R is a lattice) and the intrinsic symmetry of ϕ is that of a crystallographic group. This type of states is denoted by SP_4.

Clearly SP_1, SP_2, SP_3, and SP_4 exhaust the class of all SP states. In Sec. 4 we shall show that each of the four types satisfies characteristic cluster properties in space.

3. SPECTRUM PROPERTIES

The aim of this section is to establish a connection between the classification of SP states and certain spectral properties of representations of the translation group $\mathbb{R}^3$ which occur naturally in the study of these states. Aside from its intrinsic interest, this connection will be used in the next section to obtain an alternate characterization of SP_n ($n = 1$ to 4) states in terms of cluster properties.

We first notice that, if ϕ is an Euclidean-invariant state and Ψ is a vector in $\mathcal{H}_\phi$ such that $U_\phi(1, \mathbf{a})\Psi = e^{i\mathbf{a}\cdot\mathbf{p}}\Psi$ for all $\mathbf{a}$ in $\mathbb{R}^3$, we have then, for each Λ in O_3,

$$U_\phi(1, \mathbf{a})U_\phi(\Lambda, 0)\Psi = e^{i\Lambda^{-1}[\mathbf{a}]\cdot\mathbf{p}}U_\phi(\Lambda, 0)\Psi.$$

Whenever $\mathbf{p} \neq 0$, there exists an arbitrary small rotation Λ such that $\Lambda^{-1}[\mathbf{a}] \cdot \mathbf{p} \neq \mathbf{a} \cdot \mathbf{p} \pmod{2\pi}$ and, hence, $(U_\phi(\Lambda, 0)\Psi, \Psi) = 0$; i.e., $\|U_\phi(\Lambda, 0)\Psi - \Psi\| = 2$, which contradicts the continuity of U_ϕ. For reference purposes we take formal notice of this apparently well-known result in the following lemma.

Lemma 3.1: Let ϕ be an Euclidean-invariant state on $\mathcal{A}$ and $U_\phi(\mathbb{R}^3) = \{U_\phi(\mathbf{a}) = e^{i\mathbf{P}\phi\cdot\mathbf{a}} \mid \mathbf{a} \in \mathbb{R}^3\}$ the representation of $\mathbb{R}^3$ obtained via the GNS construction. Then the only discrete point in the spectrum of $\mathbf{P}_\phi$ is $\mathbf{p} = 0$.

Now, let ψ be one of our pure phases [i.e., $\psi \in$ supp (μ_ϕ) where ϕ is an SP state]. In general, ψ is not $\mathbb{R}^3$ invariant. We can, however, restore the translation invariance by considering the state $\chi = \eta^R\psi$ where η^R is an invariant mean over $\mathbb{R}^3$. It is interesting to see that χ still retains a precise memory of the translational symmetry H_ψ^R of the pure phase ψ. This fact is more precisely expressed by the following theorem of which, incidentally, the preceding lemma restricted to SP_1 states is a particular case.

Theorem 3.1: Let ϕ be an SP state, $\psi \in$ supp (μ_ϕ), and $\chi = \eta^R\psi$. Then χ is a KMS state and the discrete spectrum of $\mathbf{P}_\chi$ ($U_\chi(\mathbf{a}) \equiv e^{i\mathbf{P}\chi\cdot\mathbf{a}}$ $\forall$ $\mathbf{a} \in \mathbb{R}^3$) is identical with the reciprocal group ${}^*H_\psi^R$ of H_ψ^R,[59] and is simple.

Proof: Since ψ is invariant with respect to H_ψ^R, χ can be obtained from ψ by averaging the latter over $\mathbb{R}^3/H_\psi^R$. This group is compact since $\phi \in SP$. It possesses then a *unique* invariant measure h, which is called its Haar measure. We can therefore write

$$\chi = \int_{\mathbb{R}^3/H_\psi^R} \psi_{\tilde{\mathbf{a}}}\, dh(\tilde{\mathbf{a}}),$$

where $\mathbf{a} \to \tilde{\mathbf{a}}$ denotes the canonical mapping from $\mathbb{R}^3$ onto $\mathbb{R}^3/H_\psi^R$, and $\psi_{\tilde{\mathbf{a}}} \equiv \alpha_{\mathbf{a}}^*\psi$. By assumption ψ belongs to $\mathcal{E}_\beta$ (Theorem 1.3) and so then do all $\psi_{\tilde{\mathbf{a}}}$; consequently, χ belongs to $\mathfrak{S}_\beta$. This proves the first part of the theorem. The proof of the second part rests on the existence of a particular isometric isomorphism ρ from the center $\mathfrak{Z}_\chi(\mathcal{A})$ of $\pi_\chi(\mathcal{A})''$ onto $\mathfrak{L}^\infty(\mathbb{R}^3/H_\psi^R; h)$. We first construct our ρ. To this effect, we consider the mapping $\bar{\psi}$ from $\mathbb{R}^3/H_\psi^R$ to $\mathfrak{S}_\beta$ defined by $\bar{\psi}(\tilde{\mathbf{a}}) = \psi_{\tilde{\mathbf{a}}}$. Since $\bar{\psi}$ is continuous, the measure $\nu \equiv h \cdot \bar{\psi}^{-1}$ is a regular[60] Borel measure on $\mathfrak{S}_\beta$, which is clearly concentrated on $\mathcal{E}_\beta$. Consequently,[61] ν is maximal. Since ν represents χ and $\mathfrak{S}_\beta$ is a simplex (Theorem 1.1), ν is identical with the central measure μ_χ of χ. There exists, therefore, an isometric isomorphism x from $\mathfrak{Z}_\chi$ onto $\mathfrak{L}^\infty(\mathfrak{S}_\beta, \nu)$ such that

$$\langle\tilde{\chi}; ZA\rangle = \int_{\mathfrak{S}_\beta} x_Z(\psi)\langle\psi; A\rangle\, d\nu(\psi),$$

where $\tilde{\chi}$ denotes the natural extension of χ to $B(\mathcal{H}_\chi)$. Since ν is actually concentrated on $\bar{\psi}(\mathbb{R}^3/H_\psi^R)$, we can rewrite this as

$$\langle\tilde{\chi}; ZA\rangle = \int_{\mathbb{R}^3/H_\psi^R} \rho_Z(\tilde{\mathbf{a}})\langle\psi_{\tilde{\mathbf{a}}}; A\rangle\, dh(\tilde{\mathbf{a}}),$$

with $\rho_Z(\tilde{\mathbf{a}}) \equiv x_Z(\psi_{\tilde{\mathbf{a}}})$ satisfying our requirements. Now, let $\mathbf{p}$ be any eigenvalue of $\mathbf{P}_\chi$; there exists then a vector Ψ in $\mathcal{H}_\chi$ such that $U_\chi(\mathbf{a})\Psi = e^{i\mathbf{p}\cdot\mathbf{a}}\Psi$ for all $\mathbf{a}$ in $\mathbb{R}^3$. We show now that there exists an element Z in $\mathfrak{Z}_\chi(\mathcal{A})$ such that $\Psi = ZX$ where X is, as above, the cyclic vector corresponding to χ via the GNS construction. Since X is cyclic for $\pi_\chi(\mathcal{A})$, there exists a

sequence $\{A_n\}$ of elements of $\mathcal{A}$ such that

$$\lim_{n\to\infty} \|\pi_\chi(A_n)\mathrm{X} - \Psi\| = 0.$$

We form $B_n = \eta_\chi^R(e^{-i\mathbf{p}\cdot\mathbf{a}}\alpha_{\tilde{\mathbf{a}}}[A_n])$ which all belong to $\mathfrak{Z}_\chi(\mathcal{A})$ since, as a consequence of locality, $\mathcal{A}$ is asymptotically Abelian under the action of $\mathbb{R}^3$. One then verifies that $\{B_n\}$ converges weakly, as n tends to infinity, to an element Z in $\mathfrak{Z}_\chi(\mathcal{A})$ such that $Z\mathrm{X} = \Psi$. We now form, for every A in $\mathcal{A}$ and every $\mathbf{a}_0$ in $\mathbb{R}^3$,

$$\begin{aligned} e^{i\mathbf{p}\cdot\mathbf{a}_0}\int_{\mathbb{R}^3/H_\psi^R} \rho_Z(\tilde{\mathbf{a}})\langle\psi_{\tilde{\mathbf{a}}}; A\rangle\, dh(\tilde{\mathbf{a}}) \\ = e^{i\mathbf{p}\cdot\mathbf{a}_0}(\pi_\chi(A^*)\mathrm{X}, \Psi) = \langle\tilde{\chi}; \alpha_{-\tilde{\mathbf{a}}_0}[A]\mathrm{Z}\rangle \\ = \int_{\mathbb{R}^3/H_\psi^R} \rho_Z(\tilde{\mathbf{a}})\langle\psi_{\tilde{\mathbf{a}}}; \alpha_{-\mathbf{a}_0}[A]\rangle\, dh(\tilde{\mathbf{a}}) \\ = \int_{\mathbb{R}^3/H_\psi^R} \rho_Z(\tilde{\mathbf{a}}' + \tilde{\mathbf{a}}_0)\langle\psi_{\tilde{\mathbf{a}}'}; A\rangle\, dh(\tilde{\mathbf{a}}'). \end{aligned}$$

Since A runs over $\mathcal{A}$, $\rho_Z(\tilde{\mathbf{a}} + \tilde{\mathbf{a}}_0) = e^{i\mathbf{p}\cdot\mathbf{a}_0}\rho_Z(\tilde{\mathbf{a}})$, from which we conclude that $\mathbf{p}$ belongs to the reciprocal group ${}^*H_\psi^R$ and that $\rho_Z(\tilde{\mathbf{a}}) = e^{i\mathbf{p}\cdot\mathbf{a}}$. Since ρ is an isomorphism, Z is unique, given $\mathbf{p}$, and hence $\mathbf{p}$ is nondegenerate. To prove that $\mathbf{p}$ is an eigenvalue of $\mathbf{P}_\chi$, whenever $\mathbf{p}$ belongs to ${}^*H_\psi^R$, we construct $Z_\mathbf{p} \equiv \rho^{-1}(f_\mathbf{p})$ [where the function $f_\mathbf{p}$ in $\mathfrak{L}^\infty(\mathbb{R}^3/H_\psi^R; h)$ is defined as $f_\mathbf{p}(\tilde{\mathbf{a}}) = e^{i\mathbf{p}\cdot\mathbf{a}}$] and then simply retrace back the steps of the proof just completed. QED

The next theorem shows that states χ as considered in the preceding theorem do occur naturally in a decomposition of an *SP* state into its $\mathbb{R}^3$-extremal components. We recall[62] that since $\mathcal{A}$ is $\mathbb{R}^3$ Abelian, as a consequence of locality, the set $\mathfrak{S}_R$ of all $\mathbb{R}^3$ invariant states on $\mathcal{A}$ is a simplex. There exists therefore a unique maximal measure ν_ϕ^R representing ϕ in $\mathfrak{S}_R$ and concentrated (in the Baire sense) on the set $\mathfrak{E}_R$ of all extremal $\mathbb{R}^3$-invariant states on $\mathcal{A}$:

$$\langle\phi; A\rangle = \int_{\mathfrak{S}_R} \langle\psi; A\rangle\, d\nu_\phi^R(\psi).$$

Theorem 3.2: Let ϕ be a *SP* state, μ_ϕ its central measure, and ν_ϕ^R the measure decomposing ϕ into its extremal $\mathbb{R}^3$-invariant components. For every χ in supp (ν_ϕ^R) there exists a ψ in supp (μ_ϕ) such that $\chi = \eta^R\psi$.

Proof: We noticed in the proof of the preceding theorem that $\chi \equiv \eta^R\psi$ is $\mathbb{R}^3$ invariant for every ψ in $\mathfrak{S}$. Since $\mathbf{p} = 0$ is a nondegenerate (Theorem 3.1) eigenvalue of $\mathbf{P}_\chi$, χ is now extremal $\mathbb{R}^3$ invariant whenever ψ belongs to supp (μ_ϕ). We now consider the mapping η_0^R obtained as the restriction of η^R to supp (μ_ϕ). Since supp (μ_ϕ) is contained in $\mathfrak{E}_\beta$ (Theorem 1.3), the range of η_0^R is contained in $\mathfrak{E}_R$ by reason of the preceding remark. Since ϕ is by assumption an *SP* state, $\mathbb{R}^3/H_\psi^R$ is compact (Theorem 2.2); from this, one can conclude[63] that η_0^R is continuous. We then form[60] the regular Borel measure $\nu = \mu_\phi \circ (\eta_0^R)^{-1}$ on $\mathfrak{S}_R$. This measure is concentrated, by construction, on $\mathfrak{E}_R$ and is therefore[61] maximal. It clearly represents ϕ, since $\eta^R\phi = \phi$, and is therefore identical to ν_ϕ^R since $\mathfrak{S}_R$ is a simplex. Consequently, every χ in supp (ν_ϕ^R) lies in the support of ν and is of the form $\chi = \eta^R\psi$ with ψ in supp (μ_ϕ). QED

This theorem allows, in particular, the characterization of SP_1 states *among SP states* on the basis of their invariance properties with respect to translations alone.

Theorem 3.3: Let ϕ be an *SP* state. Then the following conditions are equivalent:

(i) $\mathbf{p} = 0$ is a nondegenerate eigenvalue of $\mathbf{P}_\phi$;
(ii) ϕ is extremal $\mathbb{R}^3$ invariant;
(iii) ϕ is a pure phase;
(iv) ϕ is an SP_1 state.

Proof: (i) and (ii) are well known to be equivalent since $\mathcal{A}$ is $\mathbb{R}^3$ Abelian. (iii) and (iv) are equivalent by definition. If ϕ is not extremal $\mathbb{R}^3$ invariant, ν_ϕ^R provides a genuine decomposition of ϕ into KMS states and ϕ is therefore not a pure phase. Hence, (iii) implies (ii). To prove that (ii) implies (iii), we proceed by elimination. Let us suppose that ϕ is an SP_2 state, and ψ be any of its constitutive phases. ψ is translation invariant, but not rotation invariant, so that $\eta^R\psi = \psi$ is not rotation invariant. Consequently, $\eta^R\psi$ is different from ϕ and then (Theorem 3.2) the decomposition of ϕ with respect to η_ϕ^R is not trivial; i.e., ϕ is not extremal $\mathbb{R}^3$ invariant. Suppose now that ϕ is an SP_3 or SP_4 state, $\psi \in$ supp (μ_ϕ), and $\chi = \eta^R\psi$ is $\mathbb{R}^3$ invariant. To assume that $\eta^R\psi$ would also be O_3^+ invariant would then imply that $\eta^R\psi$ is E^3 invariant and, hence, (by Lemma 3.1) $\mathbf{P}_\chi$ would admit only one eigenvalue, namely $\mathbf{p} = 0$. By Theorem 3.1 this would in turn imply that ${}^*H_\psi^R = \{0\}$ and, hence, $H_\psi^R = \mathbb{R}^3$, which contradicts the assumption that ϕ is in SP_3 or SP_4. Hence, $\eta^R\psi$ is not O_3^+ invariant. From there, the proof proceeds as in the SP_2 case. We conclude, therefore, that the only possibility occurring under assumption (ii) is that ϕ is SP_1; i.e., (ii) implies (iii). QED

4. CLUSTER PROPERTIES

In this section we list several cluster properties, and we show that they differentiate between the various classes of intrinsic symmetry of *SP* states.

An $\mathbb{R}^3$-invariant state ϕ is said to be:

(i) *uniform clustering* when

$$\forall A \in \mathcal{A}\ \forall \epsilon > 0, \text{ there exists } R_{\epsilon,A} \text{ such that}$$
$$|\langle \phi; AB\rangle - \langle \phi; A\rangle\langle \phi; B\rangle| < \epsilon \|B\| \ \forall B \in \mathcal{A}(\mathsf{C}(S_R)) \equiv \overline{\bigcup_{V \cap S_R = \phi} \mathcal{A}(V)}$$

(S_R denotes a sphere with radius R);

(ii) *weak mixing* when

$$\eta^R |\langle \phi; \alpha_{\hat{\mathbf{a}}}[A]B\rangle - \langle \phi; A\rangle\langle \phi; B\rangle| = 0;$$

(iii) *partial weak mixing* when one can choose a plane O such that, for every Cartesian coordinate system x, y, z with the x axis not contained in O, one has

$$\eta^x |\eta^{y,z}\langle \phi; \alpha_{\hat{a}}[A]B\rangle - \langle \phi; A\rangle\langle \phi; B\rangle| = 0$$

[η^x denotes mean along the x direction and $\eta^{y,z}$ denotes mean over the (y, z) plane];

(iv) *weak clustering* when

$$\eta^R \langle \phi; \alpha_{\hat{\mathbf{a}}}[A]B\rangle = \langle \phi; A\rangle\langle \phi; B\rangle.$$

Clearly, (i) → (ii) → (iii) → (iv).

In view of Theorem 3.3 it is natural to consider first the SP_1 states separately from the other possible classes. We have the following theorem.

Theorem 4.1: A state $\phi \in P$ is an SP_1 state if and only if ϕ is uniformly clustering.

Proof: Let $\tilde{\phi}$ denote the restriction of ϕ to $\mathcal{A}(V)$, and let E be the projection on the subspace of $\mathcal{H}_\phi$ generated by $\pi_\phi(\mathcal{A}(V))\Phi$; one has then $\pi_{\tilde{\phi}}(\mathcal{A}(V)) \simeq \pi_\phi(\mathcal{A}(V))E$. Since ϕ is locally normal, the representation $\pi_{\tilde{\phi}}$ of $\mathcal{A}(V)$ is[20,21] a direct sum of (irreducible) Fock representations and therefore[64] $\tilde{\phi}$ is primary of type I. The mapping $\pi_\phi(\mathcal{A}(V))'' \to \pi_\phi(\mathcal{A}(V))''E$ is an isomorphism; consequently, $\pi_\phi(\mathcal{A}(V))$ is quasi-equivalent to $\pi_{\tilde{\phi}}(\mathcal{A}(V))$ and therefore[64] $\pi_\phi(\mathcal{A}(V))''$ is also a factor of type I. This structure implies[65] that ϕ is uniform clustering if and only if ϕ is primary. Since ϕ is a KMS state, this means that ϕ is a pure phase and is of type SP_1. QED

We now undertake the characterization of the other types of SP states. As we showed in the course of the proof of Theorem 3.3, they can be nontrivially decomposed into extremal $\mathbb{R}^3$-invariant states. In order to analyze the different cluster properties of states occurring in this decomposition, we need the following lemma.

Lemma 4.2: Let χ be an extremal translation invariant state and let S_D denote the discrete part of the spectrum of $U_\chi(\mathbb{R}^3)$. Then

(i) χ weak mixing is equivalent to $S_D = \{0\}$;

(ii) χ partial weak mixing and not weak mixing is equivalent to S_D contains other points besides $\mathbf{p} = 0$, and all those points lie in a plane;

(iii) χ weak clustering and not partial weak mixing is equivalent to S_D contains points that lie in three noncoplanar directions.

Proof: The equivalence (i) is well known in classical ergodic theory. The proof of the present statement can be obtained[66] by a slight generalization of the proof[67] in the classical case. Incidentally, another possible way to arrive at the same result is to use[68] Godement's theory of means.

We now prove (ii). Suppose that χ is partial weak mixing; i.e., for every direction x not in O, one has

$$\eta^x |\eta^{y,z}(\Psi_1, U_\chi(\hat{\mathbf{a}})\Psi_2) - (\Psi_1, \mathrm{X})(\mathrm{X}, \Psi_2)| = 0, \quad \Psi_1, \Psi_2 \in \mathcal{H}_\chi.$$

Let us denote by D_0 the subspace of $\mathcal{H}_\chi$ of vectors invariant for $U_\chi(0, a_y, a_z)$. When we perform the mean over y, z, we have by the ergodic theorem

$$\eta^x |(\Psi_1, U_\chi(\hat{a}_x)D_0\Psi_2) - (\Psi_1, \mathrm{X})(\mathrm{X}, \Psi_2)| = 0.$$

Since the subspace D_0 is invariant for $U_\chi(a_x)$, we see that an alternative way of expressing the fact that χ is partial weak mixing is

$$\eta^x |(\Psi_1', U_\chi(\hat{a}_x)\Psi_2') - (\Psi_1', \mathrm{X})(\mathrm{X}, \Psi_2')| = 0, \quad \Psi_1', \Psi_2' \in D_0;$$

i.e., $U_\chi(a_x)$ is weakly mixing on D_0.

By (i), this is equivalent with the fact that X is the only eigenvector of $U_\chi(\hat{a}_x)$ in D_0. Obviously, this is true if and only if the equation $U_\chi(\mathbf{a})\Lambda = e^{ip_x a_x}\Lambda$ ($\Lambda \in \mathcal{H}_\chi$) has as the only solution $p_x = 0$ and $\Lambda = \mathrm{X}$; i.e., in the discrete x direction there are no points of S_D. Since this holds for all directions x not in O, we see that all points of S_D lie in a plane. The fact that there actually are points in this plane is by (i) equivalent to χ not weak mixing. This proves (ii). Statement (iii) now follows directly. Since χ is extremal invariant, one has that χ is weakly clustering. The fact that χ is not partial weak mixing is by (ii) equivalent to S_D contains points that do not lie in one plane. QED

This result together with the spectral properties derived in Sec. 3 enables us now to prove the following theorem which complements Theorem 4.1.

Theorem 4.2: Let ϕ be an SP state which is not extremal invariant with respect to translations, and let

ν_ϕ^R denote the measure that decomposes ϕ into extremal translation invariant states. Let $\chi \in \nu_\phi^R$;then (i) ϕ is an SP_2 state if and only if χ is weak mixing; (ii) ϕ is an SP_3 state if and only if χ is partial weak mixing and not weak mixing; (iii) ϕ is an SP_4 state if and only if χ is weakly clustering and not partial weak mixing.

Proof: Owing to Theorem 3.2, there exists a $\psi \in$ supp (μ_ϕ) such that $\chi = \eta^R\psi$. Since ϕ is strongly transitive, the type of ϕ is determined by the symmetry possessed by ψ. The state ϕ is not extremal $\mathbb{R}^3$ invariant; consequently, (Theorem 3.3) ϕ is an SP_2, SP_3, or SP_4 state. In view of Theorem 3.1, we have that the discrete part S_D of the spectrum of $U_\chi(\mathbb{R}^3)$ satisfies $S_D = {}^*H_\psi^R$. Now ϕ is by definition an SP_2 state when $H_\psi^R = \mathbb{R}^3$; i.e., $S_D = {}^*H_\psi = \{0\}$. By Lemma 4.1(i) this is equivalent with χ is weakly mixing. The state ϕ is an SP_3 state when $H_\psi^R \neq \mathbb{R}^3$, but it is continuous in at least one direction. This means that the points of ${}^*H_\psi^R$ lie in a plane orthogonal to this direction. There are other points besides $\mathbf{p} = 0$ in this plane since the closed subgroup H_ψ^R differs from $\mathbb{R}^3$. By Lemma 4.1(ii), this is equivalent with χ is partial weakly mixing but not weakly mixing. Finally, ϕ is an SP_4 state when H_ψ^R is generated by three noncoplanar vectors. In this case the reciprocal group ${}^*H_\psi^R$ is also a 3-dimensional lattice. By Lemma 4.1(iii), this is equivalent to χ weakly clustering but not partial weak mixing. QED

5. CONCLUDING REMARKS

We now want to emphasize in traditional physical terms the contrast between the properties of SP_1 and SP_4 states. The analysis made in this paper provides, in the case of the latter, a mechanism for the symmetry breaking occurring in crystallization. This leads to interpret the (grand) canonical state in the so-called "solid region" of the phase diagram as a mixture of pure phases which individually possess the intrinsic crystalline symmetry H_ϕ^R and which are transformed one into another by (rigid) Euclidean motions. We also proved that the extremal translation invariant state χ, obtained by averaging a pure crystalline phase ψ with respect to all translations in 3-dimensional space, possesses the following characteristic properties. First, the discrete part of the momentum spectrum is nondegenerate and isomorphic to the lattice reciprocal to H_ψ^R. Second, it satisfies only the weakest of all four clustering properties considered here, namely weak clustering. Finally, the (grand) canonical equilibrium state ϕ itself, which is a mixture, does not show any of these cluster properties. In contrast, for SP_1 states, the (grand) canonical equilibrium state ϕ cannot be decomposed, either as a mixture of pure phases or as a mixture of extremal translation invariant states. The discrete momentum spectrum of such a state consists only in the single point $\mathbf{p} = 0$. The strongest of the four cluster properties, namely, uniform clustering, is satisfied by these states, and by them only.

The first distinction between SP_4 and SP_1 states (namely, the fact that the former are decomposable whereas the latter are not) should come as no surprise: It is necessary for the understanding of crystalization as a symmetry-breaking process. The second distinction, in terms of the momentum spectrum, might be of experimental relevance. The third distinction has at least one immediate physical consequence: It provides a firm basis for Landau's argument on the absence of critical point in the liquid–solid phase transition; this argument indeed rests[69] on the assumption that solid and liquid have different kinds of cluster properties in space. Our contribution in this respect is that the analysis carried in this paper (see, in particular, Secs. 3 and 4) justifies this assumption.

ACKNOWLEDGMENTS

In the course of this work we benefited from several discussions with Dr. W. Wils. G. G. Emch wants to thank Dr. C. Aarts for his hospitality at the Faculteit der Wiskunde and Natuurwetenschappen, Universiteit Nijmegen, during the Summer of 1969, where part of this research was carried out.

* Z.W.O. Postdoctoral Fellow.
[1] D. Kastler and D. W. Robinson, Commun. Math. Phys. **3**, 151 (1966).
[2] D. W. Robinson and D. Ruelle, Ann. Inst. Henri Poincaré **4**, 299 (1967).
[3] D. Kastler, R. Haag, and L. Michel, Seminar talk, Marseille, 1967.
[4] G. G. Emch, H. J. F. Knops, and E. J. Verboven, J. Phys. Soc. Japan **26**, Suppl., 301 (1969).
[5] H. J. F. Knops, Ph.D. thesis, Nijmegen, 1969.
[6] H. J. F. Knops and E. J. Verboven, Phys. Letters **29A**, 386 (1969).
[7] R. Kubo, J. Phys. Soc. Japan **12**, 570 (1957).
[8] P. C. Martin and J. Schwinger, Phys. Rev. **115**, 1342 (1959).
[9] R. Haag, N. M. Hugenholtz, and M. Winnink, Commun. Math. Phys. **5**, 215 (1967).
[10] N. M. Hugenholtz, Commun. Math. Phys. **6**, 189 (1967).
[11] D. W. Robinson, Commun. Math. Phys. **7**, 337 (1968).
[12] M. Winnink, thesis, Groningen, 1968.
[13] H. Araki, Preprint RIMS 38, Kyoto University, 1968.
[14] H. Araki and H. Miyata, Preprint RIMS 37, Kyoto University, 1968.
[15] D. Kastler, J. C. T. Pool, and E. T. Poulsen, Commun. Math. Phys. **12**, 175 (1969).
[16] H. Araki, Preprint IHES, 1969.
[17] A. Z. Jadzyk, Commun. Math. Phys. **13**, 142 (1969).
[18] N. M. Hugenholtz and J. D. Wieringa, Commun. Math. Phys. **11**, 183 (1969).
[19] M. Takesaki, Preprint, UCLA and Tohoku University, 1969(?).
[20] G. F. dell'Antonio, S. Doplicher, and D. Ruelle, Commun. Math. Phys. **2**, 223 (1966).
[21] G. F. dell'Antonio and S. Doplicher, J. Math. Phys. **8**, 663 (1967).

[22] Some important facts which we now understand as consequences of the KMS condition were already noticed by H. Araki and E. J. Woods, J. Math. Phys. **4**, 637 (1963).

[23] C. N. Yang and T. D. Lee, Phys. Rev. **87**, 404 (1952).

[24] We recall that this is the topology associated to the point-wise convergence of the linear functionals $\phi(A) = \langle \phi; A \rangle$ on $\mathcal{A}$; see, for instance, Dunford and Schwartz, *Linear Operators* (Interscience, New York, 1957), Part I.

[25] See the remark following Proposition I.6.3 in J. Dixmier, *Les algèbres d'opérateurs dans l'espace Hilbertien* (Gauthier-Villars, Paris, 1967).

[26] J. Dixmier, Ref. 25, Theorem I.6.3.

[27] J. Dixmier, Ref. 25, Corollary to Proposition I.2.2.

[28] H. Araki, Ref. 13, Corollary 2.5.

[29] H. Araki, Ref. 16, Theorem 10.3, for instance.

[30] I.e., with η^G denoting an invariant mean on G, $\eta^G\langle \psi; C^*[\alpha_{\hat{g}}[A], B]C\rangle = 0$ $\forall$ A, B, C in $\mathcal{A}$ and $\forall$ G invariant states ψ on $\mathcal{A}$, we actually only need here the following consequence of this assumption: $\eta^g_\phi[A] \in \mathfrak{Z}_\phi(\mathcal{A})$ for the state ϕ considered.

[31] Throughout this paper we shall assume locality; i.e., $[A, B] = 0$ for all local observables A and B which are spatially disjoint. As a consequence of this assumption, we have $\lim \|[A, \alpha_g[B]]\| = 0$, as $g \to \infty$; this in turn implies all other asymptotic conditions in space which we shall use.

[32] D. Ruelle, *Statistical Mechanics* (Benjamin, New York, 1969).

[33] A previous investigation has been carried out by one of us (Ref. 5) along the lines of the present paper with, however, this condition *assumed* throughout.

[34] S. Sakai, Trans. Am. Math. Soc. **118**, 406 (1965).

[35] W. Wils, Compt. Rend. **267**, 810 (1968).

[36] For the definition of Baire sets and their relation with Borel sets, see for instance P. R. Halmos, *Measure Theory* (Van Nostrand, New York, 1950), Sec. 51.

[37] Defined by $\pi_\mu(A) = \int_{\mathfrak{S}} \pi_\psi(A)\, d\mu_\phi(\psi)$ for all A in $\mathcal{A}$.

[38] We recall that a "diagonal operator" on $\mathcal{H}_\mu$ is a mapping from $\mathcal{H}_\mu$ into itself of the form $\int_{\mathfrak{S}} T_\psi\, d\mu_\phi(\psi)$ where $T_\psi = \lambda(\psi) I_\psi$ with I_ψ denoting the identity on $\mathcal{H}_\psi$ and λ belonging to $\mathfrak{L}^\infty(\mathfrak{S}, \mu_\phi)$.

[39] See, for instance, J. Dixmier, *Les C*-algèbres et leurs représentations* (Gauthier-Villars, Paris, 1964), Sec. 8.

[40] For instance, the C^*-algebra of the canonical commutation relations is not separable; see 3.3.3 in J. Manuceau, Ann. Inst. Henri Poincaré, 139 (1968).

[41] When $\mathcal{A}$ is separable, $\mathfrak{S}$ becomes metrizable and the concepts of Baire and Borel sets coincide so that Wils' theory (Ref. 35) contains Dixmier's (Ref. 39) as a particular case.

[42] For general reference on this theory, see R. R. Phelps, *Lectures on Choquet's theorem*, Van Nostrand Mathematical Series #7 (Van Nostrand, Princeton, N.J. 1966).

[43] There are at least five equivalent characterizations of simplices (Choquet's theorem; see, for instance, Ref. 42, p. 66); the one taken here is the most appropriate to out purpose.

[44] See the proof of Lemma 9.6 of Ref. 42.

[45] Theorem I.4.1 in Dixmier, Ref. 25.

[46] Example I.3.6c in Dixmier, Ref. 25.

[47] C.E. Rickart, *General theory of Banach Algebras* (Van Nostrand, Princeton, 1960), Appendix, Sec. 1.

[48] O. Lanford and D. Ruelle, J. Math. Phys. **8**, 1460 (1967), Proposition 4.2; for a proof see D. Ruelle, Ref. 49.

[49] D. Ruelle, Comm. Math. Phys. **3**, 133 (1966); see, in particular, pp. 146–7.

[50] See, for instance, p. 30 in R. R. Phelps (Ref. 42).

[51] We recall that μ_ϕ is already the unique measure satisfying the axioms of a central measure. The assertion of the present theorem is clearly of a different type. Furthermore, [see Theorem II.3.1 in Knops (Ref. 5)] an even stronger kind of uniqueness occurs, namely, that μ_ϕ is the only (regular) measure on $\mathfrak{S}_\beta$ concentrated in the Baire sense on $\mathfrak{E}_\beta$ whenever η Abelianess in time holds.

[52] Proposition 4.3(ii) in Lanford and Ruelle (Ref. 48).

[53] See Lemma 3 in Kastler (Ref. 3.)

[54] See Corollary to Lemma 2 in Kastler, Ref. 3.

[55] For details, see Knops, Ref. 5, p. 81.

[56] The interchange of the mean and integral can be justified as follows. We first notice that the integrand is a function of positive type in g. The mean η^E can hence be written as $\lim I_n$, as $n \to \infty$, where I_n denotes an integral over a compact subset of $\mathbb{E}^3$. The argument is then completed by a successive application of Fubini's and Lebesgue's theorems.

[57] (iii) implies in particular that $\mathbb{E}^3/H^E_\psi$ and O^E_ψ are Borel isomorphic. In the separable case, this is always the case, irrespective of transitivity; see a remark in the beginning of Sec. 3 in Kastler.

[58] See, in this spirit, G. G. Emch, H. J. F. Knops, and E. J. Verboven, Commun. Math. Phys. **8**, 300 (1968).

[59] We recall that $^*H^R_\psi = \{\mathbf{p} \in \mathbb{R}^3 \mid \mathbf{p} \cdot \mathbf{a} = 0 \pmod{2\pi}\ \forall\ \mathbf{a} \in H^R_\psi\}$.

[60] See Example 53.7 in Halmos, Ref. 36.

[61] Corollary 9.8 in Phelps, Ref. 42.

[62] Corollary 3.2 in Lanford and Ruelle, Ref. 48.

[63] See pp. 104–5 in Knops, Ref. 5.

[64] Prop. 5.4.11 in Dixmier, Ref. 25.

[65] See Lemma 10.2 in Araki, Ref. 16 or Lemma 2.2.2 in Knops, Ref. 5.

[66] See Lemma 2.2.1 in Knops, Ref. 5.

[67] P. R. Halmos, *Lectures on Ergodic Theory* (Chelsea, New York, 1956), p. 40.

[68] See Theorem 3 in S. Doplicher, and D. Kastler, Commun. Math. Phys. **7**, 1 (1968).

[69] G. E. Uhlenbeck, in *Fundamental Problems in Statistical Mechanics II*, E. G. D. Cohen, Ed. (North-Holland, Amsterdam, 1968); see, in particular, p. 17.

Reduction of the Most Degenerate Representations of $SO_0(p, q)$ with Respect to $SO_0(p-1, q-1) \otimes T_{p+q-2}$

L. CASTELL
*International Centre for Theoretical Physics, Trieste, Italy**

(Received 28 September 1967)

The reduction of three principal series of the most degenerate unitary irreducible representations of the identity component of the noncompact rotation groups $SO_0(p, q)$ with respect to the identity component of the inhomogeneous rotation group $SO_0(p-1, q-1) \otimes T_{p+q-2}$ is carried out, using the results of the reduction with respect to the maximal compact subgroup $SO(p) \otimes SO(q)$.

I. INTRODUCTION

In this paper we deal with the principal series of the most degenerate unitary representations of the identity component of the noncompact rotation groups $SO_0(p, q)$. As the carrier space of these representations, we consider the three homogeneous spaces (of rank one under the action of the group)

$$SO_0(p, q)/SO_0(p-1, q), \quad SO_0(p, q)/SO_0(p, q-1),$$

and

$$SO_0(p, q)/SO(p-1, q-1) \otimes T_{p+q-2}.$$

These are homeomorphic to the two hyperboloids H_q^p and H_p^q and to the cone C_q^p in the $(p+q)$-dimensional Minkowski space $M_{p,q}$, $p \geq q \geq 1$,

$$H_q^p: \quad x_1^2 + x_2^2 + \cdots x_p^2 - x_{p+1}^2 - \cdots - x_{p+q}^2 = 1, \tag{1}$$

$$H_p^q: \quad x_1^2 + x_2^2 + \cdots x_p^2 - x_{p+1}^2 - \cdots - x_{p+q}^2 = -1 \tag{2}$$

(for H_p^1 we have the additional restriction $x_{p+1} \geq 1$),

$$C_q^p: \quad x_1^2 + x_2^2 + \cdots x_p^2 - x_{p+1}^2 \cdots - x_{p+q}^2 = 0, \tag{3}$$

$$(x_\mu) \neq (0), \quad (x_\mu) \equiv (-x_\mu)$$

[for C_1^p the equivalence relation $(x_\mu) \equiv (-x_\mu)$ is replaced by condition $x_{p+1} \geq 0$]. In Ref. 1 a biharmonic coordinate system has been used on these manifolds so that the maximal compact subgroup $SO(p) \otimes SO(q)$ factors out. The basis states in one irreducible representation are labelled by $p+q-2$ integer quantum numbers. Since these states belong to the Hilbert space on which the representation has been constructed, the proof of irreducibility, unitarity, etc., of the representation is relatively simple. We shall make use of these results. We construct the quasiregular representation on the three manifolds; however, we label the basis states of one irreducible representation by the quantum numbers of the subgroup $SO_0(p-1, q-1) \otimes T_{p+q-2}$ which is the semidirect product of $SO_0(p-q, q-1)$ with the $(p+q-2)$-dimensional Abelian translation subgroup T_{p+q-2} in $M_{p-1,q-1}$. The quasiregular representations must be unitarily equivalent to those constructed in Ref. 1 and, thus, we can pick out the different irreducible components by direct comparison. In this way, we obtain the reduction of the most degenerate irreducible representations of the principal series of $SO_0(p, q)$ with respect to

$$SO_0(p-1, q-1) \otimes T_{p+q-2}.$$

The third alternative, the reduction of the irreducible representations of $SO_0(p, q)$ with respect to $SO_0(p-1, q)$ or $SO_0(p, q-1)$, has already been treated.[2]

On the hypersurfaces (1, 2, 3), we introduce the coordinate system

$$y_\mu = x_\mu(x_p + x_{p+q})^{-1}, \quad \mu = 1, 2, \cdots, p-1,$$
$$y_{\mu-1} = x_\mu(x_p + x_{p+q})^{-1}, \quad \mu = p+1, \cdots, p+q-1, \tag{4}$$
$$\bar{y} = (x_p + x_{p+q})^{-1}.$$

The restriction on the coordinates is $\bar{y} \neq 0$ for the hyperboloids and $\bar{y} \neq 0$, $(\bar{y}) \equiv (-\bar{y})$ for the cone, besides H_p^1 and C_1^p where we have $\bar{y} > 0$. In this coordinate system all transformations of the group $SO_0(p, q)$ are most conveniently parametrized in the following way:

$$y'_\mu = \lambda L_{\mu\nu} y_\nu + a_\mu, \quad \bar{y}' = \lambda \bar{y}, \tag{5}$$

$$y'_\mu = \frac{y_\mu - b_\mu(y^2 + \epsilon\bar{y}^2)}{\sigma(y, \bar{y})}, \quad \bar{y}' = \frac{\bar{y}}{\sigma(y, \bar{y})}, \tag{6}$$

where

$$y^2 = \sum_1^{p+q-2} g_{\mu\nu} y_\mu y_\nu, \quad \text{etc.},$$

$$g_{11} = g_{22} = \cdots g_{p-1,p-1} = -g_{p,p} = \cdots$$
$$= -g_{p+q-2,p+q-2} = 1, \quad g_{\mu\nu} = 0, \quad \text{for} \quad \mu \neq \nu,$$
$$\sigma(y, \bar{y}) = 1 - 2 \text{ by } + b^2(y^2 + \epsilon\bar{y}^2),$$

$\epsilon = +1, -1, 0$, for H_p^q, H_q^p, and C_q^p, respectively.

TABLE I. Irreducible most degenerate representations of $(SO_0(p-1,q-1)\otimes D)\ \textcircled{S}\ T_{p+q-2}$

$p\geq q>2$	$p>q=2$	$p=q=2$	$p>2, q=1$	$p=2, q=1$
R^+, R^-	R^+_+, R^+_-, R^-	$R^+_+, R^-_{[+]}, R^+_-, R^-_{[-]}$	R^-	R^-_+, R^-_-

The a_μ represent the translations T_{p+q-2} and the b_μ the "special conformal" transformations. The $L_{\mu\nu}$ represent the transformations of $SO(p-1,q-1)$ and the parameter λ the dilatations in such a way that if $L_{\mu\nu}$ belongs to $SO_0(p-1,q-1)$, then $0<\lambda$; if $L_{\mu\nu}$ does not belong to the identity component of $SO(p-1,q-1)$, we have $\lambda<0$ (the latter case will not arise for $q=1$).

The Riemannian metric tensor which is invariant under (5) and (6) gives rise to the line element

$$\bar{y}^{-2}(g_{\mu\nu}\,dy_\mu\,dy_\nu+\epsilon\,d\bar{y}^2).$$

From this there follow for the invariant norm, which defines the Hilbert space in which the quasiregular representation is constructed,

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\psi^*\psi\bar{y}^{1-p-q}\,d^{p+q-2}y\,d\bar{y} \tag{7}$$

for the hyperboloids (besides H^1_p) and

$$\int_{-\infty}^{+\infty}\int_{0}^{\infty}\psi^*\psi\bar{y}^{1-p-q}\,d^{p+q-2}y\,d\bar{y} \tag{8}$$

for the cone and H^1_p. Under the action of the group

$$(y_\mu,\bar{y})\to(y'_\mu,\bar{y}')=g(y_\mu,\bar{y}),$$

ψ transforms like

$$\psi(y_\mu,\bar{y})\to\psi'(y'_\mu,\bar{y}'),$$

where

$$\psi'(y_\mu,\bar{y})\equiv U_g\psi(y_\mu,\bar{y})=\psi(g^{-1}(y_\mu,\bar{y})).$$

The last line defines the quasiregular representation U_g.

Since the reflection[3] $x'_\mu=-x_\mu$ or

$$y'_\mu=y_\mu,\quad \bar{y}'=-\bar{y},$$
$$\psi'(y_\mu,\bar{y})\equiv\mathcal{R}\psi(y_\mu,\bar{y})=\pm\psi(y_\mu,-\bar{y}),$$

commutes with all group elements and leaves the hyperboloids invariant (besides H^1_p where the homogeneous space is given by the upper sheet), we can divide the quasiregular representation into a series which contains only representations with "parity" $\mathcal{R}=+1$ and a series with "parity" $\mathcal{R}=-1$. Then the norm is in all cases given by (8).

II. THE REDUCTION OF THE MOST DEGENERATE REPRESENTATIONS OF THE PRINCIPAL SERIES OF $SO_0(p,q)$ WITH RESPECT TO $SO_0(p-1,q-1)\ \textcircled{S}\ T_{p+q-2}$

Before we proceed with the reduction of the quasiregular representations, we specify the irreducible representations of $SO_0(p-1,q-1)\ \textcircled{S}\ T_{p+q-2}$ enlarged by the dilatations D {i.e., $[SO_0(p-1,q-1)\otimes D]\ \textcircled{S}\ T_{p+q-2}$} which are relevant. Since these irreducible representations are just a direct integral over the real or imaginary mass spectrum of the irreducible representations of $SO_0(p-1,q-1)\ \textcircled{S}\ T_{p+q-2}$, the further reduction to this group is quite straightforward. The most degenerate or "spin-0" representations of mass $m\neq 0$ of $[SO_0(p-1,q-1)\otimes D]\ \textcircled{S}\ T_{p+q-2}$ can be characterized (see, for example, the analysis in Ref. 4) by the signum of $m^2=-g_{\mu\nu}p_\mu p_\nu$ and, in some cases, in addition by the signum of p_{p+q-2}, where p_μ, $\mu=1,2,\cdots,q+p-2$, are the momenta which specify a state in an irreducible representation $R^{\operatorname{sgn}m^2}_{\operatorname{sgn}p_{p+q-2}}$. (For $p=q=2$ we use besides $\operatorname{sgn}p_2$ also $[\operatorname{sgn}p_1]$ written in brackets.)

We classify the result in Table I.

Now we give the results of the decomposition of the quasiregular representations derived in Ref. 1. On the hyperboloids H^p_q the second-order Casimir operator

$$C=\nu^2-[\tfrac{1}{2}(p+q-2)]^2$$

and the reflection operator $\mathcal{R}$ assumes, in general, the eigenvalues specified in Table II (the meaning of the parameter β given there will become clear later).

TABLE II. Description of the eigenvalues of the Casimir and reflection operator.

$p+q$	Even	Even	Odd	Odd
$\mathcal{R}$	$(-1)^{\frac{1}{2}(p-q)}$	$(-1)^{\frac{1}{2}(p-q+2)}$	$(-1)^{\frac{1}{2}(p-q+1)}$	$(-1)^{\frac{1}{2}(p-q-1)}$
Discrete spectrum	$\nu=1,3,5,\cdots$	$\nu=2,4,6,\cdots$	$\nu=\frac{1}{2},\frac{5}{2},\frac{9}{2},\cdots$	$\nu=\frac{3}{2},\frac{7}{2},\frac{11}{2},\cdots$
Continuous spectrum	$i\infty>\nu>0$	$i\infty>\nu\geq 0$	$i\infty>\nu>0$	$i\infty>\nu>0$
β	0	$\frac{1}{2}\pi$	$-\frac{1}{4}\pi$	$\frac{1}{4}\pi$

All representations are irreducible and occur with multiplicity 1 in the quasiregular representation. For H_p^q, we have only to interchange p and q.

Further restrictions: For H_p^2 the irreducible representations of the discrete series are specified, in addition, by the eigenvalues $T = \pm 1$ of the signum of the eigenvalues of the $SO(2)$ subgroup, and so are the representations (in Ref. 1, the question of irreducibility has not been considered for $\nu = 0$) characterized by p even, $\nu = 0$, and $\mathfrak{R} = (-1)^{\frac{1}{2}p}$. The quasiregular representation on H_p^1 is characterized by $i\infty > \nu > 0$.

On the cone C_q^p we can use the signum of the operator

$$S = \frac{1}{i}\left[\bar{y}\frac{\partial}{\partial \bar{y}} - \frac{1}{2}(p+q-2)\right]$$

of the (genuine) similarity transformations

$$y'_\mu = y_\mu, \quad \bar{y}' = l\bar{y}, \quad 0 < l < \infty,$$
$$\psi'(y_\mu, \bar{y}) = l^{\frac{1}{2}(p+q-2)}\psi(y_\mu, l^{-1}\bar{y}),$$

to specify completely an irreducible representation. [It should be mentioned that in Ref. 1 the equivalence relation $\bar{y} \equiv (-\bar{y})$ has been omitted, and a simple parity doubling is obtained.] The operator S which commutes with all generators of $SO_0(p,q)$ is self-adjoint with respect to the norm (8) and obeys the relation

$$S^2 = -C - [\tfrac{1}{2}(p+q-2)]^2.$$

On the cone C_q^p the Casimir operator C and the operator sgn S have the eigenvalues $C = \nu^2 - [\frac{1}{2}(p+q-2)]^2$ and sgn $S = \pm 1$, $i\infty > \nu \geq 0$. Although we need the operator sgn S on the cone C_q^p to specify a representation, the unitarily inequivalent representations are determined by the different eigenvalues of C.

The next and most important step is the construction of a complete set of eigenfunctions [i.e., a complete basis of the Hilbert space defined by (8)] into which we can expand the quasiregular representation. This basis decomposes into a direct sum and integral over the bases of subspaces which are invariant under the action of a definite irreducible unitary representation of $(SO_0(p-1, q-1) \otimes D) \otimes T_{p+q-2}$. The determination of these representations solves our problem.

In the coordinate system (4), we obtain for the second-order Casimir operator C on the three manifolds

$$\left(\bar{y}^2 \epsilon g_{\mu\nu}\frac{\partial}{\partial y_\mu}\frac{\partial}{\partial y_\nu} + \bar{y}^2\frac{\partial}{\partial \bar{y}}\frac{\partial}{\partial \bar{y}} - (p+q-3)\bar{y}\frac{\partial}{\partial \bar{y}}\right)\varphi = C\varphi,$$
$$-\infty < y_\mu < +\infty, \quad 0 < \bar{y} < \infty.$$

The eigenfunctions of C and the $p+q-2$ momentum operators (a maximal commuting set of nonvanishing operators) are given by

$$\varphi = e^{ip_\mu y_\mu}\bar{y}^{\frac{1}{2}(p+q-2)}J_{\pm\nu}(\epsilon^{\frac{1}{2}}m\bar{y}), \quad \epsilon = \pm 1,$$
$$\varphi = e^{ip_\mu y_\mu}\bar{y}^{\frac{1}{2}(p+q-2)}(m\bar{y})^{\pm\nu}, \quad \epsilon = 0,$$

where $J_{\pm\nu}$ is a Bessel function of order ν. Using the methods and results of the theory of self-adjoint extensions of symmetric operators,[5] we obtain the complete set of eigenfunctions, which for the hyperboloids is characterized by a continuous real parameter, the angle β. The completeness relation for the orthogonal eigenfunctions is given on the hyperboloids as follows:

$$\frac{(\bar{y}\bar{y}')^{\frac{1}{2}(p+q-2)}}{(2\pi)^{p+q-2}}\int_{-\infty}^{+\infty} d^{p+q-2}p e^{ip_\mu(y_\mu - y'_\mu)}\left(\sum_{n=0}^{\infty} 2s_n J_{s_n}(|m|\,\bar{y}')J_{s_n}(|m|\,\bar{y})\theta(\epsilon m^2)\right.$$
$$- \int_0^\infty d\rho[J_{i\sqrt{\rho}}(|m|\,\bar{y}')f(\beta,\sqrt{\rho}) - J_{-i\sqrt{\rho}}(|m|\,\bar{y}')f(\beta,-\sqrt{\rho})][J_{i\sqrt{\rho}}(|m|\,\bar{y})f(\beta,\sqrt{\rho}) - J_{-i\sqrt{\rho}}(|m|\,\bar{y})f(\beta,-\sqrt{\rho})]$$
$$\left.\times\,[4\sinh(\pi\sqrt{\rho})]^{-1}\theta(\epsilon m^2) + \int_0^\infty d\rho K_{i\sqrt{\rho}}(|m|\,\bar{y}')K_{i\sqrt{\rho}}(|m|\,\bar{y})\frac{\sinh(\pi\sqrt{\rho})}{\pi}\theta(-\epsilon m^2)\right)$$
$$= \bar{y}^{p+q-1}\delta^{p+q-2}(y_\mu - y'_\mu)\delta(\bar{y} - \bar{y}'), \quad (9)$$

where

$$f(\beta,\sqrt{\rho}) = [\cos(\beta + \tfrac{1}{2}i\pi\sqrt{\rho})/\cos(\beta - \tfrac{1}{2}i\pi\sqrt{\rho})]^{\frac{1}{2}},$$
$$s_n(\beta) = (2/\pi)\beta + 2n + 1, \quad -\tfrac{1}{2}\pi < \beta \leq \tfrac{1}{2}\pi,$$
$$n = 0, 1, 2, \cdots, \quad \rho = +(-\nu^2)^{\frac{1}{2}},$$

and $K_{i\sqrt{\rho}}$ is the modified Bessel function of second kind of order $i\sqrt{\rho}$. In order to pick out the right expansions which form a basis of the unitary representation of the group $SO_0(p,q)$, we fix the parameter β in such a way that the discrete spectrum characterized by $s_n(\beta)$ coincides with the discrete spectrum ν derived in Ref. 1. These values of β have already been given in Table II.

TABLE III. The reduction.

		$p \geq q > 2$	$p > q = 2$	$p = q = 2$	$p > 2, q = 1$	$p = 2, q = 1$
H_q^p	discrete	R^-	R^-	$R^-_{[+]}, R^-_{[-]}$	R^-	R^-_+, R^-_-
	continuous	$R^+ + R^-$	$R^+_+ + R^+_+ + R^-$	$R^+_+ + R^+_- + R^-_{[+]} + R^-_{[-]}$	R^-	$R^-_+ + R^-_-$
H_p^q	discrete	R^+	R^+_+, R^+_-	R^+_+, R^+_-	—	—
	continuous	$R^+ + R^-$	$R^+_+ + R^+_- + R^-$	$R^+_+ + R^+_- + R^-_{[+]} + R^-_{[-]}$	R^-	$R^-_+ + R^-_-$
C_q^p	continuous	$R^+ + R^-$	$R^+_+ + R^+_- + R^-$	$R^+_+ + R^+_- + R^-_{[+]} + R^-_{[-]}$	R^-	$R^-_+ + R^-_-$

The completeness relation on the cone is given by (10):

$$\frac{(\bar{y}\bar{y}')^{\frac{1}{2}(p+q-2)}}{(2\pi)^{p+q-1}} \int_{-\infty}^{+\infty} d^{p+q-2}p e^{ip_\mu(y_\mu - y_\mu')} \times \left(\int_0^\infty ds(|m|\,\bar{y}')^{-is}(|m|\,\bar{y})^{is} + \int_0^\infty ds(|m|\,\bar{y}')^{is}(|m|\,\bar{y})^{-is} \right) = \bar{y}^{p+q-1}\delta^{p+q-2}(y_\mu - y_\mu')\delta(\bar{y} - \bar{y}'). \quad (10)$$

From the expansions (9) and (10), it is easy to see which is the maximum number of irreducible representations of $SO_0(p, q)$ which might possibly occur in the reduction for a fixed value of C, $\mathcal{R}$, or signum S, since the basis space for the unitary representations of $[SO_0(p-1, q-1) \otimes D] \mathbin{\textcircled{S}} T_{p+q-2}$ has to be irreducible. Comparing this with the number of irreducible representations which actually occur, we can derive in all cases the reduction without any ambiguity. The result is given in Table III. All representations of $SO_0(p, q)$ are irreducible.

From Table III it follows that, for the discrete spectrum, all irreducible representations of $SO_0(p, q)$ are also irreducible with respect to the subgroup $[SO_0(p-1, q-1) \otimes D] \mathbin{\textcircled{S}} T_{p+q-2}$. The continuous spectrum always contains the sum of all inequivalent irreducible most degenerate representations of the subgroup [except for $p + q$ even, $\nu = 0$, where the reduction for reflection $\mathcal{R} = (-1)^{\frac{1}{2}(p-q-2)}$ proceeds in analogy with the discrete case on both hyperboloids; this follows if we consider the limit $\rho \to 0$ in the eigenfunction expansion (9) for $\beta = \frac{1}{2}\pi$ and from the additional restriction mentioned after Table II]. This concludes our analysis.

From a physical point of view, the conformal group $SO_0(4, 2)$ of Minkowski space (considered as a group of motion of the three 5-dimensional manifolds) and its reduction with respect to the Poincaré group are of interest.[6] Since more relations always exist between the quantum numbers for the discrete representations than exist for the continuous ones, one expects the former to be of greater importance to physics. Only the discrete representations (and the representations characterized by $\nu = 0$, $\mathcal{R} = 1$) on H_4^2 decompose into nothing other than representations of the Poincaré group with real mass. As these representations have a direct physical interpretation, the more specific properties of the manifold H_4^2 may be relevant to particle physics.[7]

ACKNOWLEDGMENTS

The author is indebted to Professor A. Salam and Professor P. Budini as well as the IAEA for the kind hospitality extended to him at the International Centre for Theoretical Physics, Trieste.

* Present address: Max Planck Institut für Physik und Astrophysik, 8 München 23, Germany.

[1] N. Limić, J. Niederle, and R. Rączka, J. Math. Phys. **7** 1861, 2026 (1966); **8**, 1079 (1967).
[2] J. Niederle, J. Math. Phys. **8**, 1921 (1967).
[3] $\mathcal{R}$ belongs to the centre of $SO_0(p, q)$ if p and q are even.
[4] U. Ottoson, Arkiv Fysik **33**, 523 (1967).
[5] E. C. Titchmarsh, *Eigenfunction Expansions Associated with Second-Order Differential Equations* (Oxford U.P., Oxford, 1962), Part 1.
[6] L. Castell, Nucl. Phys. **B4**, 343 (1967).
[7] L. Castell, Nuovo Cimento **46A**, 1 (1966); Nucl. Phys. **B5**, 601 (1968).

Exact Solution of a Family of Matrix Integral Equations for Multiply Scattered Partially Polarized Radiation. II*

JUNE BUELL, JOHN CASTI, ROBERT KALABA, AND SUEO UENO
Department of Electrical Engineering, University of Southern California, Los Angeles, California 90007

(Received 20 October 1969)

In the theory of multiply scattered partially polarized radiation, a key role is played by the integral equation

$$J(t, x, z) = Ie^{-(x-t)/z} + \int_0^x K(|t-y|)J(y, x, z)\,dy, \quad 0 \le t \le x,$$

where J and K are square matrices, I is the unit matrix, z is a parameter lying in the interval $(0, 1)$, and for the interval length x, $0 \le x \le x_1$. It is shown that this family of matrix integral equations can be transformed into a Cauchy problem that is readily solved by modern computing machines.

1. INTRODUCTION

A fundamental integral equation in the theory of multiply scattered partially polarized radiation is

$$J(t, x, z) = Ie^{-(x-t)/z} + \int_0^x K(|t-y|)J(y, x, z)\,dy, \quad 0 \le t \le x \le x_1, 0 \le z \le 1.$$

The capital letters stand for $n \times n$ square matrices, I is the unit $n \times n$ matrix, and the matrix kernel K can be represented in the form

$$K(r) = \int_0^1 e^{-r/z'}W(z')\,dz', \quad r > 0.$$

The physical background is presented in Ref. 1, and references to earlier work are given there. The work by Sobolev[2] is particularly valuable.

In the first paper in this series,[3] it was shown that the solution of a certain Cauchy system provides the solution of the family of matrix integral equations for the source matrix J. There, an auxiliary matrix R, a function of an independent variable and two other parameters, is considered. The purpose of this paper is to present another Cauchy system that solves the integral equation for the matrix J. It involves two auxiliary matrices X and Y, which are functions of an independent variable and only one other parameter. This is of great significance computationally.

The Cauchy system is derived in the next section, and, following that, a validation of the initial-value method is given.[4] A discussion of remaining points to be considered is then given.

Paper III of this series will present numerical results for internal and external fields for partially polarized radiation in a finite slab with multiple scattering according to a Rayleigh phase matrix. Some numerical results for the scalar case are presented in Refs. 5 and 6. Reference 7 contains related results for the matrix case.

II. DERIVATION OF THE CAUCHY SYSTEM

Consider the family of matrix integral equations

$$J(t, x, z) = Ie^{-(x-t)/z} + \int_0^x K(|t-y|)J(y, x, z)\,dy, \quad 0 \le t \le x \le x_1, 0 \le z \le 1, \quad (1)$$

where J and K are square $n \times n$ matrices, and I is the unit $n \times n$ matrix. The kernel K can be represented in the form

$$K(r) = \int_0^1 e^{-r/z'}W(z')\,dz', \quad r > 0, \quad (2)$$

where W is a square $n \times n$ matrix. Differentiate Eq. (1) with respect to x to obtain the relation

$$J_x(t, x, z) = -z^{-1}Ie^{-(x-t)/z} + K(x-t)J(x, x, z) + \int_0^x K(|t-y|)J_x(y, x, z)\,dy. \quad (3)$$

Introduce the square matrix Φ as the solution of the integral equation

$$\Phi(t, x) = K(x-t) + \int_0^x K(|t-y|)\Phi(y, x)\,dy, \quad 0 \le t \le x. \quad (4)$$

By using the principle of superposition for linear systems, we see that

$$\Phi(t, x) = \int_0^1 J(t, x, z')W(z')\,dz', \quad 0 \le t \le x. \quad (5)$$

Equation (3) may be regarded as an integral equation for the function J_x whose solution is

$$J_x(t, x, z) = -z^{-1}J(t, x, z) + \Phi(t, x)J(x, x, z), \quad x \ge t. \quad (6)$$

By putting $t = 0$, we see that

$$J_x(0, x, z) = -z^{-1}J(0, x, z) + \Phi(0, x)J(x, x, z), \quad (7)$$

which is a differential equation for the function $J(0, x, z)$. Next, we obtain a differential equation for the matrix function $J(x, x, z)$.

Return to the integral equation (1) and replace t by $x - t$. The result is

$$J(x - t, x, z) = Ie^{-t/z} + \int_0^x K(|x - t - y|)J(y, x, z)\, dy. \quad (8)$$

By changing the variable of integration from y to $x - y$, Eq. (8) becomes

$$J(x - t, x, z) = Ie^{-t/z} + \int_0^x K(|t - y|)J(x - y, x, z)\, dy. \quad (9)$$

Through differentiation with respect to x, the above equation becomes

$$\frac{d}{dx} J(x - t, x, z) = K(x - t)J(0, x, z) + \int_0^x K(|t - y|) \frac{d}{dx} J(x - y, x, z)\, dy. \quad (10)$$

This may be regarded as an integral equation for the function $dJ(x - t, x, z)/dz$, whose solution is

$$\frac{d}{dx} J(x - t, x, z) = \Phi(t, x)J(0, x, z). \quad (11)$$

These considerations may now be summarized. First, introduce the matrices X and Y by means of the definitions

$$X(x, z) = J(x, x, z), \quad (12)$$

$$Y(x, z) = J(0, x, z), \quad 0 \leq x \leq x_1, \quad 0 \leq z \leq 1. \quad (13)$$

The differential equations (7) and (11) become

$$X_x(x, z) = \left(\int_0^1 Y(x, z')W(z')\, dz'\right) Y(x, z), \quad (14)$$

$$Y_x(x, z) = -z^{-1}Y(x, z) + \left(\int_0^1 Y(x, z')W(z')\, dz'\right) X(x, z), \quad 0 \leq x \leq x_1, 0 \leq z \leq 1. \quad (15)$$

According to the integral Eq. (1), the initial conditions at $x = 0$ are

$$X(0, z) = I, \quad (16)$$

$$Y(0, z) = I. \quad (17)$$

The differential equation for the matrix function J is

$$J_x(t, x, z) = -z^{-1}J(t, x, z) + \Phi(t, x)X(x, z), \quad x_1 \geq x \geq t, \quad (18)$$

where

$$\Phi(t, x) = \int_0^1 J(t, x, z')W(z')\, dz'. \quad (19)$$

The initial condition on the function J at $x = t$ is

$$J(t, t, z) = X(t, z), \quad 0 \leq z \leq 1. \quad (20)$$

III. VALIDATION

We now wish to show, conversely, that the solution of the Cauchy system in Eqs. (14)–(20) provides a solution of the matrix family of integral equations in Eq. (1), where the kernel K is given in terms of the matrix W by means of Eq. (2). Introduce the matrix M to be

$$M(t, x, z) = Ie^{-(x-t)/z} + \int_0^x K(|t - y|)J(y, x, z)\, dy, \quad 0 \leq t \leq x \leq x_1, 0 \leq z \leq 1. \quad (21)$$

Our task is to show that $J = M$. We accomplish this by showing that the matrices J and M satisfy the same Cauchy problem for $0 \leq t \leq x \leq x_1$. It is easily shown that the matrix M satisfies the same differential equation that the matrix J does. Verifying the initial condition at $x = t$ is more lengthy, as will be seen.

Differentiate both sides of Eq. (21) to obtain the equations

$$\begin{aligned} M_x(t, x, z) &= -z^{-1}Ie^{-(x-t)/z} + K(x - t)J(x, x, z) \\ &\quad + \int_0^x K(|t - y|)[-z^{-1}J(y, x, z) + \Phi(y, x)X(x, z)]\, dy \\ &= -z^{-1}M(t, x, z) \\ &\quad + \left(K(x - t) + \int_0^x K(|t - y|)\Phi(y, x)\, dy\right) X(x, z) \\ &= -z^{-1}M(t, x, z) + \left(\int_0^1 M(t, x, z')W(z')\, dz'\right) X(x, z), \\ &\qquad x \geq t. \quad (22) \end{aligned}$$

This shows that the matrix function M satisfies the same differential equation that the function J does for $x \geq t$. Next, we consider the initial condition at $x = t$:

$$M(t, t, z) = I + \int_0^t K(|t - y|)J(y, t, z)\, dy, \quad 0 \leq t \leq x_1, 0 \leq z \leq 1. \quad (23)$$

Introduce the nomenclature

$$\begin{aligned} P(x, z) &= M(x, x, z) \\ &= I + \int_0^x K(x - y)J(y, x, z)\, dy \\ &= I + \int_0^x K(y)J(x - y, x, z)\, dy, \\ &\qquad 0 \leq x \leq x_1, 0 \leq z \leq 1. \quad (24) \end{aligned}$$

The basic task remaining is to show that $P = X$, for $0 \leq x \leq x_1$ and $0 \leq z \leq 1$. This is done by comparing the Cauchy problem for P with that for X. Differentiate Eq. (24) with respect to x to obtain

$$P_x(x, z) = K(x)J(0, x, z) + \int_0^x K(y) \frac{d}{dx} J(x - y, x, z)\, dy. \quad (25)$$

Keeping in mind Eqs. (18) and (15) and the initial conditions on the functions $J(0, x, z)$ and $Y(x, z)$, it is seen that

$$J(0, x, z) = Y(x, z). \quad (26)$$

Introduce the auxiliary matrix function Z as the solution of the Cauchy system

$$Z_x(t, x, z) = \left(\int_0^1 J(t, x, z')W(z')\, dz'\right) Y(x, z), \quad 0 \leq z \leq 1, x \geq t, \quad (27)$$

$$Z(t, t, z) = Y(t, z). \quad (28)$$

By comparing the Cauchy systems that they satisfy, observe that

$$Z(0, x, z) = X(x, z), \quad 0 \leq x \leq x_1, \quad 0 \leq z \leq 1. \quad (29)$$

Now we obtain expressions for the partial derivatives J_t and Z_t. First, notice that the matrix function J_t satisfies the differential equation

$$[J_t(t, x, z)]_x = -z^{-1}J_t(t, x, z) + \left(\int_0^1 [J_t(t, x, z')W(z')\, dz'\right) X(x, z). \quad (30)$$

For notational precision, introduce

$$J_1 = J_t, \quad J_2 = J_x. \quad (31)$$

Then, from the equation

$$J(t, t, z) = X(t, z), \quad (32)$$

we obtain the matrix relation

$$J_1(t, t, z) + J_2(t, t, z) = \left(\int_0^1 Y(t, z')W(z')\, dz'\right) Y(t, z) \quad (33)$$

or

$$J_1(t, t, z) = z^{-1}X(t, z) - \left(\int_0^1 X(t, z')W(z')\, dz'\right) X(t, z) + \left(\int_0^1 Y(t, z')W(z')\, dz'\right) Y(t, z), \quad 0 \leq z \leq 1, \quad (34)$$

which is the desired initial condition on the function J_t at $x = t$.

Next, the function Ψ is introduced as

$$\Psi(t, x, z) = z^{-1}J(t, x, z) + \left(\int_0^1 Z(t, x, z')W(z')\, dz'\right) Y(x, z) - \left(\int_0^1 J(t, x, z')W(z')\, dz'\right) X(x, z), \quad 0 \leq t \leq x \leq x_1, 0 \leq z \leq 1. \quad (35)$$

We show that the function Ψ is actually the function J_t. By inspection, it is seen immediately that

$$J_1(t, t, z) = \Psi(t, t, z). \quad (36)$$

Differentiation of both sides of Eq. (35) yields the equation

$$\begin{aligned}\Psi_x(t, x, z) = {} & z^{-1}(-z^{-1}J(t, x, z) + \Phi(t, x)X(x, z)) \\ & + \left(\int_0^1 Z_x(t, x, z')W(z')\, dz'\right) Y(x, z) \\ & + \left(\int_0^1 Z(t, x, z')W(z')\, dz'\right) Y_x(x, z) \\ & - \left(\int_0^1 J_x(t, x, z')W(z')\, dz'\right) X(x, z) \\ & - \left(\int_0^1 J(t, x, z')W(z')\, dz'\right) X_x(x, z).\end{aligned} \quad (37)$$

Through substitution, it becomes

$$\begin{aligned}\Psi_x(t, x, z) = {} & -z^{-2}J(t, x, z) + z^{-1}\Phi(t, x)X(x, z) \\ & + \left[\int_0^1 \left(\int_0^1 J(t, x, z'')W(z'')\, dz'' \times Y(x, z')\right) W(z')\, dz'\right] Y(x, z) \\ & + \left(\int_0^1 Z(t, x, z'')W(z'')\, dz''\right)\left[-z^{-1}Y(x, z) + \left(\int_0^1 Y(x, z')W(z')\, dz'\right) X(x, z)\right] \\ & - \left(\int_0^1 (-(z')^{-1}J(t, x, z') + \Phi(t, x)X(x, z'))W(z')\, dz'\right) X(x, z) \\ & - \left(\int_0^1 J(t, x, z'')W(z'')\, dz''\right) \times \left(\int_0^1 Y(x, z')W(z')\, dz'\right) Y(x, z).\end{aligned} \quad (38)$$

Through cancellation of like terms and use of the definition of the matrix Ψ in Eq. (35), the equation

for the matrix Ψ_x becomes

$$\Psi_x(t, x, z) = -z^{-1}\Psi(t, x, z) + \left(\int_0^1 \Psi(t, x, z')W(z')\,dz'\right)X(x, z), \quad x \geq t. \tag{39}$$

Since they are solutions of identical Cauchy systems, we conclude that

$$\Psi(t, x, z) = J_1(t, x, z), \quad 0 \leq t \leq x \leq x_1, \quad 0 \leq z \leq 1. \tag{40}$$

For the matrix Z, introduce similar notation:

$$Z_1 = Z_t, \quad Z_2 = Z_x. \tag{41}$$

Observe that for the function Z_t we may write

$$[Z_t(t, x, z)]_x = \left(\int_0^1 \Psi(t, x, z')W(z')\,dz'\right)Y(x, z), \quad x \geq t. \tag{42}$$

To obtain the initial condition in the matrix function Z_t at $x = t$, we differentiate both sides of Eq. (28) to obtain the equation

$$Z_1(t, t, z) + Z_2(t, t, z) = -z^{-1}Y(t, z) + \left(\int_0^1 Y(t, z')W(z')\,dz'\right)X(t, z). \tag{43}$$

It follows that

$$Z_1(t, t, z) = -\left(\int_0^1 X(t, z')W(z')\,dz'\right)Y(t, z) - z^{-1}Y(t, z) + \left(\int_0^1 Y(t, z')W(z')\,dz'\right)X(t, z). \tag{44}$$

Introduce the new matrix Ω by means of the definition

$$\Omega(t, x, z) = -z^{-1}Z(t, x, z) + \left(\int_0^1 Z(t, x, z')W(z')\,dz'\right)X(x, z) - \left(\int_0^1 J(t, x, z')W(z')\,dz'\right)Y(x, z), \quad 0 \leq t \leq x. \tag{45}$$

By inspection, it is seen that

$$\Omega(t, t, z) = Z_1(t, t, z). \tag{46}$$

Moreover, differentiation of both sides of Eq. (45) yields the result

$$\begin{aligned}\Omega_x(t, x, z) = {} & -z^{-1}Z_x(t, x, z) \\ & + \left(\int_0^1 Z_x(t, x, z')W(z')\,dz'\right)X(x, z) \\ & + \left(\int_0^1 Z(t, x, z')W(z')\,dz'\right)X_x(x, z) \\ & - \left(\int_0^1 J_x(t, x, z')W(z')\,dz'\right)Y(x, z) \\ & - \left(\int_0^1 J(t, x, z')W(z')\,dz'\right)Y_x(x, z).\end{aligned} \tag{47}$$

By substitution, the last equation becomes

$$\begin{aligned}\Omega_x(t, x, z) = {} & -z^{-1}\Phi(t, x)Y(x, z) \\ & + \left(\int_0^1 \Phi(t, x)Y(x, z')W(z')\,dz'\right)X(x, z) \\ & + \left(\int_0^1 Z(t, x, z')W(z')\,dz'\right) \\ & \times \left(\int_0^1 Y(x, z')W(z')\,dz'\right)Y(x, z) \\ & - \left(\int_0^1 [-(z')^{-1}J(t, x, z') \right. \\ & \left. + \Phi(t, x)X(x, z')]W(z')\,dz'\right)Y(x, z) \\ & - \left(\int_0^1 J(t, x, z')W(z')\,dz'\right)\Big[-z^{-1}Y(x, z) \\ & + \left(\int_0^1 Y(x, z')W(z')\,dz'\right)X(x, z)\Big].\end{aligned} \tag{48}$$

Through cancellation, we finally obtain the relation

$$\Omega_x(t, x, z) = \left(\int_0^1 \Psi(t, x, z')W(z')\,dz'\right)Y(x, z), \quad x \geq t. \tag{49}$$

Since the functions Ω and Z_t satisfy the same Cauchy system, it is seen that

$$Z_t(t, x, z) = \Omega(t, x, z), \quad x \geq t, \quad 0 \leq z \leq 1. \tag{50}$$

We are finally in a position to evaluate $dJ(x - t, x, z)/dx$ by means of the chain rule of differentiation. The result is

$$\begin{aligned}\frac{d}{dx}J(x - t, x, z) &= \Psi(x - t, x, z) - z^{-1}J(x - t, x, z) \\ &\quad + \Phi(x - t, x)X(x, z) \\ &= z^{-1}J(x - t, x, z) \\ &\quad + \left(\int_0^1 Z(x - t, x, z')W(z')\,dz'\right)Y(x, z) \\ &\quad - \left(\int_0^1 J(x - t, x, z')W(z')\,dz'\right)X(x, z) \\ &\quad - z^{-1}J(x - t, x, z) + \Phi(x - t, x)X(x, z).\end{aligned} \tag{51}$$

The last equation simplifies to

$$\frac{d}{dx} J(x - t, x, z) = \left(\int_0^1 Z(x - t, x, z')W(z')\,dz'\right) Y(x, z). \quad (52)$$

We next wish to show that

$$Z(x - t, x, z) = J(t, x, z), \quad 0 \le t \le x \le x_1, \quad 0 \le z \le 1. \quad (53)$$

Observe that

$$Z(0, t, z) = X(t, z), \quad (54)$$

which follows from Eqs. (27), (28), (17), (14), and (16). Next differentiate both sides of Eq. (53), which yields

$$\begin{aligned}\frac{d}{dx} Z(x - t, x, z) &= \Omega(x - t, x, z) + \Phi(x - t, x)Y(x, z) \\ &= -z^{-1}Z(x - t, x, z) \\ &\quad + \left(\int_0^1 Z(x - t, x, z')W(z')\,dz'\right) X(x, z) \\ &\quad - \left(\int_0^1 J(x - t, x, z')W(z')\,dz'\right) Y(x, z) \\ &\quad + \Phi(x - t, x)Y(x, z). \end{aligned} \quad (55)$$

This becomes, finally,

$$\begin{aligned}\frac{d}{dx} Z(x - t, x, z) &= -z^{-1}Z(x - t, x, z) \\ &\quad + \left(\int_0^1 Z(x - t, x, z')W(z')\,dz'\right) X(x, z), \\ &\qquad x \ge t. \end{aligned} \quad (56)$$

Equation (53) now follows by comparing the Cauchy system for the matrix function $Z(x - t, x, z)$ in Eqs. (56) and (54) against that for the matrix function J contained in Eqs. (18)–(20).

It follows that Eq. (52) may be rewritten as

$$\frac{d}{dx} J(x - t, x, z) = \left(\int_0^1 J(t, x, z')W(z')\,dz\right) Y(x, z), \quad x \ge t. \quad (57)$$

We may now return to Eq. (25), which becomes

$$P_x(x, z) = \left(K(x) + \int_0^x K(y) \int_0^1 J(y, x, z)W(z')\,dz'\,dy\right) Y(x, z), \quad (58)$$

or

$$P_x(x, z) = \left[\int_0^1 \left(Ie^{-x/z'} + \int_0^x K(y)J(y, x, z')\,dy\right) \times W(z')\,dz'\right] Y(x, z). \quad (59)$$

Lastly, let

$$Q(x, z) = Ie^{-x/z} + \int_0^x K(y)J(y, x, z)\,dy, \quad 0 \le x \le x_1, \quad 0 \le z \le 1. \quad (60)$$

We wish to show that

$$Q(x, z) = Y(x, z), \quad 0 \le x \le x_1, \quad 0 \le z \le 1. \quad (61)$$

It is clear that

$$Q(0, z) = Y(0, z). \quad (62)$$

Also, differentiation shows that

$$\begin{aligned} Q_x(x, z) &= -z^{-1}e^{-x/z}I + K(x)X(x, z) \\ &\quad + \int_0^x K(y)(-z^{-1}J(y, x, z) + \Phi(y, x)X(x, z))\,dy \\ &= -z^{-1}Q(x, z) + \left(\int_0^1 Q(x, z')W(z')\,dz'\right) X(x, z). \end{aligned} \quad (63)$$

By comparing Eqs. (62) and (63) against the Cauchy system for the matrix Y, it is clear that Eq. (61) holds.

Eq. (59) for the matrix P becomes

$$P_x(x, z) = \left(\int_0^1 Y(x, z')W(z')\,dz'\right) Y(x, z), \quad z \ge 0, \quad (64)$$

and clearly

$$P(0, z) = I. \quad (65)$$

We have finally obtained the result

$$P(x, z) = X(x, z), \quad 0 \le x \le x_1, \quad 0 \le z \le 1. \quad (66)$$

It follows that

$$M(t, t, z) = X(t, z), \quad 0 \le t \le x_1, \quad 0 \le z \le 1, \quad (67)$$

which is the desired initial condition on the matrix M. Since the matrices M and J fulfill the same Cauchy system for $x \ge t$, it is seen that

$$J(t, x, z) = Ie^{-(x-t)/z} + \int_0^x K(|t - y|)J(y, x, z)\,dy, \quad 0 \le t \le x \le x_1, 0 \le z \le 1, \quad (68)$$

which is the desired family of matrix integral equations for the function J.

IV. DISCUSSION

Some corollaries follow readily from our earlier considerations. The matrix function Φ, defined in Eq. (19), satisfies the matrix integral equation

$$\Phi(t, x) = K(x - t) + \int_0^x K(|t - y|)\Phi(y, x)\, dy, \qquad 0 \leq t \leq x \leq x_1. \quad (69)$$

This follows readily by multiplying both sides of the integral equation (68) by the matrix W and integrating.

In Eq. (68) we replace t by $x - t$. This yields the relation

$$J(x - t, x, z) = Ie^{-t/z} + \int_0^x K(|x - t - y|)J(y, x, z)\, dy. \quad (70)$$

Then change the variable of integration from y to $x - y$. Equation (70) becomes

$$J(x - t, x, z) = Ie^{-t/z} + \int_0^x K(|t - y|)J(x - y, x, z)\, dy. \quad (71)$$

In view of Eq. (53), it is seen that the matrix Z is a solution of the integral equation

$$Z(t, x, z) = Ie^{-t/z} + \int_0^x K(|t - y|)Z(y, x, z)\, dy, \qquad 0 \leq t \leq x \leq x_1, 0 \leq z \leq 1. \quad (72)$$

Another Cauchy system for the matrices J and Z is provided by the differential equations (40), (35), (50), (45) and the initial conditions in Eqs. (26) and (29). It is assumed that the matrices X and Y have been precomputed. The computational utility of this last Cauchy system remains to be investigated.

Finally, it remains to determine wide classes of matrix functions W for which the solution of the Cauchy problem does provide the solution of the original matrix integral equation, i.e., for which the various interchanges of the orders of limiting operations are valid.

* Supported by the National Science Foundation under Grant No. GF-294, and the National Institutes of Health under Grants No. GM-16197-01 and No. GM-16437-01.

[1] T. Mullikin, in *Transport Theory*, R. Bellman, G. Birkhoff, and J. Abu-Shumays, Eds. (American Mathematical Society, Providence, R.I. 1969), Chap. 1.

[2] V. Sobolev, *A Treatise on Radiactive Transfer* (D. Van Nostrand Co., Inc., New York, 1963).

[3] J. Buell, R. Kalaba, and S. Ueno, "Exact Solution of a Family of Matrix Integral Equations for Multiply Scattered Partially Polarized Radiation. I," University of Southern California, Los Angeles, Dept. of Electrical Engineering, Technical Report 69-1, 1969.

[4] J. Casti and R. Kalaba, "Proof of the Basic Invariant Imbedding Method for Fredholm Integral Equations with Displacement Kernels. I," The Rand Corp., RM-6037-PR, 1969.

[5] J. Casti, H. Kagiwada and R. Kalaba, "The Invariant Imbedding Numeral Method for Fredholm Integral Equations with Displacement Kernels," *Proceedings of the 24th National Conference of the Association for Computing Machinery*, 1969.

[6] R. Bellman, H. Kagiwada and R. Kalaba, J. Quant. Spectry. Radiative Transfer 6, 291 (1966).

[7] C. Adams and G. Kattawar, J. Quant. Spectry. Radiative Transfer (to be published).

Some Mathematical Properties of Oscillator Phase Operators

E. C. LERNER, H. W. HUANG, AND G. E. WALTERS
Physics Department, University of South Carolina, Columbia, South Carolina 29208

(Received 17 June 1969)

A general definition of "cosine" and "sine" operators, C and S, for harmonic oscillator phase is proposed and its consequences examined. An important feature of the spectral analysis is the "chain sequence" condition which ensures that C and S have unit norm. The (nonunitary) operator $U = C + iS$ is shown to be an annihilation-type operator whose spectral properties bear a remarkable analogy to those of the standard annihilation operator, although its spectrum fills the unit circle rather than the entire complex plane. Statistical properties of the eigenstates of U are discussed briefly.

I. INTRODUCTION

This paper is concerned with an analysis of bounded Hermitian operators C and S ("cosine" and "sine"), whose properties are such that they may reasonably be associated with the classical concept of harmonic oscillator phase.[1-3] Since, as will be seen, there is a large class of such operators, it is of considerable interest to discuss them in terms of a general definition applicable to all of them.

This definition is arrived at in Sec. II via correspondence-principle-type arguments. (The operators used in Refs. 2 and 3 become a special case.) It is then shown that the spectra of C and S fill the interval $(-1, 1)$. However, an additional restriction is required to ensure that the spectra do not extend beyond this interval. This is embodied in the "chain sequence" theorem[4] whose proof completes the definition and essentially completes Sec. II, the remainder of which is devoted to some examples of phase operators which meet the specifications.

Perhaps the most interesting quantum mechanical aspects of oscillator phase are discussed in Sec. III, where the operator $U = C + iS$ is examined. Despite the fact that this operator cannot be unitary, it is shown to have many intriguing properties which are in striking analogy to those exploited by Glauber[5] in the case of the annihilation operator. Thus, while the spectral radius of U is unity (as might reasonably be expected), its spectrum fills the unit circle in the complex plane, each interior point being an eigenvalue associated with a normalizable eigenstate. These eigenstates are complete; in fact, "over-complete." Eigenvalues with magnitudes close to unity are associated with large expectation values of the number operator. As is the case with Glauber's coherent states,[6] the argument of the eigenvalue becomes identified with the phase of the oscillator in this limit.

However, one interesting way (among many) in which the "U-states" differ from the coherent states is that the projection operator formed from an eigenstate of U, when averaged over the argument of the eigenvalue, leads to a density operator which is very close to the Gaussian density operator[5] (and is, in fact, exactly the Gaussian density operator when the choice of U corresponds to the special case of Refs. 2 and 3). The significance of this result from the information theoretic viewpoint is commented on.

II. DEFINITION OF C AND S

The oscillator with Hamiltonian

$$H = (2m)^{-1}[p^2 + (m\omega q)^2] \tag{1}$$

satisfies the equation

$$\frac{d}{dt}(m\omega q + ip) = -i\omega(m\omega q + ip) \tag{2}$$

both classically and, in the Heisenberg representation, quantum mechanically. Classically, the phase of the oscillator may be defined as

$$\phi = \arg(m\omega q + ip), \tag{3}$$

and satisfies $\dot{\phi} = -\omega$. As has been noted,[2] one cannot introduce a unitary operator which is an exponential function of an Hermitian "phase operator." However, Hermitian operators C and S may be defined which satisfy commutation rules analogous to the classical Poisson bracket relations

$$\{\cos\phi, H\} = \omega\sin\phi, \quad \{\sin\phi, H\} = -\omega\cos\phi. \tag{4}$$

These commutation rules, which form the starting point of our discussion, are

$$[C, N] = iS, \quad [S, N] = -iC, \tag{5}$$

where N is the oscillator number operator. If we write

$$C = \tfrac{1}{2}(U + U^\dagger), \quad S = (2i)^{-1}(U - U^\dagger), \tag{6}$$

the basic commutation rule becomes

$$[U, N] = U. \tag{7}$$

Equation (7) says, in turn, that

$$U|n\rangle = f(n)|n-1\rangle, \tag{8a}$$

where the $|n\rangle$, $n = 0, 1, 2, \cdots$, are the eigenstates of the number operator and the $f(n)$ satisfy $f(0) = 0$ plus such other conditions as will be appended below. (It may be noted at this point that U cannot be unitary since $U|0\rangle = 0$.) We can, without loss of generality,[7] choose the $f(n)$ to be real and positive semidefinite so that

$$U^\dagger|n\rangle = f(n+1)|n+1\rangle. \tag{8b}$$

Additional properties to be assigned to the $f(n)$ are the following:

(i) $f(n)$ is bounded. This ensures the boundedness of C and S.

(ii) $f(n) \neq 0$ for $n > 0$ and approaches unity monotonically with increasing n. The nonvanishing, whose reasonableness will become clearer in context, stems from the fact that, in discussing the phase via the properties of U, we are making an "approximate polar decomposition" of the annihilation operator a.[8] Thus, we are able to write

$$\begin{aligned} a &= [(N+1)^{\frac{1}{2}}/f(N+1)]U \\ &= [(N+1)^{\frac{1}{2}}/f(N+1)](C+iS), \end{aligned} \tag{9}$$

where the operator functions of N have an obvious meaning. The argument for the approach to unity is supplied by noting that

$$C^2 + S^2 = \tfrac{1}{2}(U^\dagger U + UU^\dagger) \tag{10}$$

satisfies

$$(C^2+S^2)|n\rangle = \tfrac{1}{2}[f^2(n) + f^2(n+1)]|n\rangle, \tag{11}$$

so that we would expect $\frac{1}{2}[f^2(n) + f^2(n+1)]$ to approach unity for large n.[9] The monotonicity is a simplifying restriction which we shall see holds for physically motivated choices of the $f(n)$.

The properties enumerated above are sufficient to show that the spectra of C and S[10] contain every point on the interval $(-1, 1)$. This may be accomplished by introducing the set of normalized states

$$|\Psi_n(\phi)\rangle = \frac{1}{(n+1)^{\frac{1}{2}}} \sum_{m=0}^{n} e^{im\phi}|m\rangle, \quad n = 0, 1, 2, \cdots. \tag{12}$$

A short calculation then shows that the square of the norm of $(U - e^{i\phi})|\Psi_n(\phi)\rangle$ is given by

$$\|(U - e^{i\phi})|\Psi_n(\phi)\rangle\|^2 = \frac{1}{(n+1)}\left(1 + \sum_{m=1}^{n}[f(m)-1]^2\right). \tag{13}$$

However, $[f(m)-1]^2$, $m = 1, 2, \cdots$, is a sequence of positive numbers decreasing monotonically to zero. Therefore, given an arbitrarily small (positive) ϵ, we can always find an $n_0(\epsilon)$ such that $[f(m)-1]^2 < \frac{1}{2}\epsilon$ for $m \geq n_0$. Let $n > n_0$ and write

$$\|(U - e^{i\phi})|\Psi_n(\phi)\rangle\|^2 < \frac{1}{(n+1)}\left(1 + \sum_{m=1}^{n_0-1}[f(m)-1]^2\right) + \tfrac{1}{2}\epsilon. \tag{14}$$

We can thus choose n sufficiently large so that

$$\|(U - e^{i\phi})|\Psi_n(\phi)\rangle\|^2 < \epsilon. \tag{15}$$

It follows that

$$\|(U - e^{i\phi})|\Psi_n(\phi)\rangle\|^2 = \|(U^\dagger - e^{-i\phi})|\Psi_n(\phi)\rangle\|^2 - f^2(n+1)/(n+1) \xrightarrow[n\to\infty]{} 0 \tag{16}$$

and

$$\|(C - \cos\phi)|\Psi_n(\phi)\rangle\|^2 = \tfrac{1}{2}\|[(U - e^{i\phi}) + (U^\dagger - e^{-i\phi})]|\Psi_n(\phi)\rangle\|^2 \xrightarrow[n\to\infty]{} 0, \tag{17}$$

so that $\cos\phi$ is a point in the spectrum of C.[11] Thus the spectrum of C (and S) contains all points on the interval $(-1, 1)$.

The fact that the $f(n)$ are bounded and approach unity monotonically is obviously not in itself sufficient to ensure that the spectrum of C is confined to the interval $(-1, 1)$. This latter requirement places the most interesting condition on the $f(n)$. It is embodied in the following theorem[12]:

Chain Sequence Theorem: A necessary and sufficient condition that the spectrum of C does not extend beyond the interval $(-1, 1)$ is that the numbers $\frac{1}{4}f^2(n)$ $(n > 0)$ form a "chain sequence"; i.e., that they can be written as[13]

$$\tfrac{1}{4}f^2(n) = (1 - g_{n-1})g_n, \quad 0 < g_n < 1. \tag{18}$$

Proof: We start with the fact that the spectral radius of C is equal to its norm, which in turn can be written as

$$\|C\| = \sup\{|\langle\Psi|\, C\,|\Psi\rangle| : \|\,|\Psi\rangle\| = 1\}. \tag{19}$$

Let

$$|\Psi\rangle = \sum_{n=0}^{\infty} \rho_n e^{i\theta_n}|n\rangle \tag{20}$$

with the ρ_n positive (or zero) and $\sum_{n=0}^{\infty}\rho_n^2 = 1$. Then

$$\langle\Psi|\, C\,|\Psi\rangle = \sum_{n=1}^{\infty}\rho_{n-1}\rho_n\cos(\theta_n - \theta_{n-1})f(n), \tag{21}$$

and it is apparent that the maximum and minimum values of $\langle\Psi|\, C\,|\Psi\rangle$ have the same magnitude which

is given by the maximum value of

$$\sum_{n=1}^{\infty} \rho_{n-1}\rho_n f(n). \tag{22}$$

For the proof of sufficiency, let

$$f(n) = 2[(1 - g_{n-1})g_n]^{\frac{1}{2}},$$

and note that

$$\sum_{n=1}^{N} [\rho_{n-1}(1 - g_{n-1})^{\frac{1}{2}} - \rho_n(g_n)^{\frac{1}{2}}]^2 = \sum_{n=0}^{N} \rho_n^2 - g_0\rho_0^2 - (1 - g_N)\rho_N^2 - \sum_{n=1}^{N} \rho_{n-1}\rho_n f(n) \geq 0, \tag{23}$$

so that the partial sums $\sum_{n=1}^{N} \rho_{n-1}\rho_n f(n)$ satisfy

$$\sum_{n=1}^{N} \rho_{n-1}\rho_n f(n) \leq \sum_{n=0}^{N} \rho_n^2 \leq 1. \tag{24}$$

This proves sufficiency.

In proving necessity, it is somewhat more convenient to consider nonnormalized sequences ρ_n for which $\sum_{n=0}^{\infty} \rho_n^2 < \infty$, and to assume that

$$\sum_{n=0}^{\infty} \rho_n^2 - \sum_{n=1}^{\infty} \rho_{n-1}\rho_n f(n) \geq 0 \tag{25}$$

for any choice of the ρ_n. Our procedure consists of showing that: (a) If

$$f(n) = 2[(1 - g_{n-1})g_n]^{\frac{1}{2}}, \quad n = 1, 2, \cdots, N \geq 1, \tag{26}$$

$0 < g_n < 1$, and $g_0 = 0$, then

$$f(N + 1) = 2[(1 - g_N)g_{N+1}]^{\frac{1}{2}}$$

with $0 < g_{N+1} < 1$; (b) $f(1)$ can be written as

$$f(1) = 2(g_1)^{\frac{1}{2}}, \quad 0 < g_1 < 1.$$

Now, given Eq. (26), we have, for $g_0 = 0$,

$$\sum_{n=0}^{N+1} \rho_n^2 - \sum_{n=1}^{N+1} \rho_{n-1}\rho_n f(n) = \sum_{n=1}^{N} [\rho_{n-1}(1 - g_{n-1})^{\frac{1}{2}} - \rho_n(g_n)^{\frac{1}{2}}]^2 + [\rho_N(1 - g_N)^{\frac{1}{2}} - \rho_{N+1}]^2 + \rho_N\rho_{N+1}[2(1 - g_N)^{\frac{1}{2}} - f(N + 1)]. \tag{27}$$

The quantities $\rho_0, \cdots, \rho_{N+1}$ can be chosen so that all but the last term on the right-hand side of Eq. (27) vanish. If we also choose $\rho_{N+2} = \rho_{N+3} = \cdots = 0$, Eq. (25) then demands that $f(N + 1) \leq 2(1 - g_N)^{\frac{1}{2}}$. Therefore,

$$f(N + 1) = 2[(1 - g_N)g_{N+1}]^{\frac{1}{2}}, \quad 0 < g_{N+1} \leq 1. \tag{28}$$

However, if $g_{N+1} = 1$, it would follow from Eq. (27) with N replaced by $N + 1$ that there would be a choice of the ρ_n such that $f(N + 2) \leq \rho_{N+2}/\rho_{N+1}$ with ρ_{N+2} arbitrary. This would imply the vanishing of $f(N + 2)$ since ρ_{N+2} could be made arbitrarily small. Thus $g_{N+1} < 1$. (The same argument holds for g_1 below.) Finally, if we choose $\rho_2 = \rho_3 = \cdots = 0$, Eq. (25) says that

$$[\rho_0 - \tfrac{1}{2}f(1)\rho_1]^2 + \rho_1^2[1 - \tfrac{1}{4}f^2(1)] \geq 0. \tag{29}$$

Choosing ρ_0 and ρ_1 so as to make the first bracket vanish, we have $f^2(1) \leq 4$ so that

$$f(1) = 2[(1 - g_0)g_1]^{\frac{1}{2}}, \quad g_0 = 0, \quad 0 < g_1 < 1. \tag{30}$$

This completes the proof.

Before examining some possible choices of $f(n)$, it should be noted that any sequence (of positive numbers) which is dominated term by term by a chain sequence is itself a chain sequence as is easily seen by considering the effect on Eq. (25) of decreasing each $f(n)$.

The choice $f(n) = 1$, $n = 1, 2, \cdots$, is obtained by choosing $g_n = \frac{1}{2}$. This is, in fact, the largest constant value $f(n)$ can have and still satisfy the chain sequence conditions.[4] It follows that any sequence $f(n)$ which approaches unity monotonically from below will satisfy our conditions. A case in point is given by

$$f(n) = [n/(n + \alpha)]^{\beta}, \tag{31}$$

with α and β positive. By suitable maneuvering with α and β, $f(n)$ can be made extremely small for a large range of values of n without affecting the spectral properties of C and S.

A somewhat more realistic choice is obtained by symmetrizing the classical expression for C, i.e., by writing

$$C = \tfrac{1}{2}m\omega[(2mH)^{-\frac{1}{2}}q + q(2mH)^{-\frac{1}{2}}] = \tfrac{1}{4}[(N + \tfrac{1}{2})^{-\frac{1}{2}}(a + a^{\dagger}) + (a + a^{\dagger})(N + \tfrac{1}{2})^{-\frac{1}{2}}]. \tag{32}$$

This leads to the sequence $f_1(n)$ given by

$$f_1(n) = \tfrac{1}{2}\{[n/(n - \tfrac{1}{2})]^{\frac{1}{2}} + [n/(n + \tfrac{1}{2})]^{\frac{1}{2}}\}. \tag{33}$$

In Ref. 1, it is shown that $\frac{1}{4}f_1^2(n)$ is dominated term by term by the chain sequence $(1 - h_{n-1})h_n$ with

$$h_n = \tfrac{1}{2}[1 - (2n + 1)^{-1}]. \tag{34}$$

Another possible sequence may be deduced from the symmetrized "polar-coordinate relation"

$$m\omega q = \tfrac{1}{2}[(2mH)^{\frac{1}{2}}C + C(2mH)^{\frac{1}{2}}]. \tag{35}$$

This leads to

$$f_2(n) = 2(n)^{\frac{1}{2}}/[(n - \tfrac{1}{2})^{\frac{1}{2}} + (n + \tfrac{1}{2})^{\frac{1}{2}}]. \quad (36)$$

A simple calculation shows that

$$\frac{f_1(n)}{f_2(n)} = \frac{1}{2}\left[1 + \left(\frac{n^2}{n^2 - \frac{1}{4}}\right)^{\frac{1}{2}}\right] > 1, \quad (37)$$

so that $\frac{1}{4}f_2^2(n)$ is also a chain sequence. Both $f_1(n)$ and $f_2(n)$ approach unity monotonically from above.

III. SPECTRAL PROPERTIES OF U AND ASSOCIATED MATTERS

The operator $U = C + iS$, although not unitary, has interesting mathematical-physical properties in its own right. These stem from the following theorem:

Theorem: The spectrum of U covers the unit circle in the complex plane. All points in the interior of the unit circle are eigenvalues of U associated with normalizable eigenstates. These states are complete, in fact, over-complete. The points on the unit circle (although in the spectrum of U) are not associated with normalizable eigenstates.

Proof: Note first of all that a consequence of Eq. (16) is that the points on the unit circle in the complex plane belong to the spectrum of U.[11] Thus, the spectral radius of U is at least unity. We show that it cannot be greater than unity (and therefore is unity) by showing[14] that

$$\lim_{m\to\infty} \|U^m\|^{1/m} \leq 1. \quad (38)$$

Start with any normalized $|\Psi\rangle = \sum_{n=0}^{\infty} C_n |n\rangle$. Then

$$U^m|\Psi\rangle = \sum_{n=m}^{\infty} C_n\left(\prod_{k=0}^{m-1} f(n-k)\right)|n-m\rangle \quad (39)$$

and

$$\|U^m|\Psi\rangle\|^2 = \sum_{n=0}^{\infty} |C_{n+m}|^2 \prod_{k=n+1}^{n+m} f^2(k). \quad (40)$$

It follows that

$$\|U^m\| = \sup\left\{\prod_{k=n+1}^{n+m} f(k): n = 0, 1, 2, \cdots\right\}. \quad (41)$$

However,

$$\prod_{k=n+1}^{n+m} f(k) = 2^m[(1 - g_n)g_{n+1}(1 - g_{n+1}) \times g_{n+2}\cdots g_{n+m-1}(1 - g_{n+m-1})g_{n+m}]^{\frac{1}{2}}. \quad (42)$$

The product under the radical above contains $m - 1$ subproducts of the form $x(1 - x)$ with $0 \leq x \leq 1$. Since $\max [x(1 - x)] = \frac{1}{4}$,

$$\prod_{k=n+1}^{n+m} f(k) \leq 2^m[4^{-(m-1)}(1 - g_n)g_{n+m}]^{\frac{1}{2}} \leq 2 \quad (43)$$

and

$$\|U^m\|^{1/m} \leq 2^{1/m} \xrightarrow[m\to\infty]{} 1. \quad (44)$$

The next step is to write down the formal solution to the eigenvalue equation

$$U|z\rangle = z|z\rangle. \quad (45)$$

This is easily checked to be

$$|z\rangle = A\left[|0\rangle + \sum_{n=1}^{\infty}\left(\prod_{k=1}^{n} f(k)\right)^{-1} z^n |n\rangle\right], \quad (46)$$

where A is a normalization constant. For $|z| < 1$ the series

$$\langle z \mid z\rangle = |A|^2\left[1 + \sum_{n=1}^{\infty}\left(\prod_{k=1}^{n} f^2(k)\right)^{-1}|z|^{2n}\right] \quad (47)$$

converges by the ratio test [since the ratio of the $(n + 1)$th to the nth term eventually becomes less than unity]. However, for $|z| = 1$, Eq. (43) says that

$$\sum_{n=1}^{N}\left(\prod_{k=1}^{n} f^2(k)\right)^{-1} \geq \tfrac{1}{4}N, \quad (48)$$

so that a normalizable eigenstate does not exist.

"Over-completeness" follows from an argument of the type used by Cahill.[15] Thus, given a normalized state vector $|\psi\rangle = \sum_n C_n |n\rangle$, the function

$$A^{-1}\langle\psi \mid z\rangle = \left[C_0^* + \sum_{n=1}^{\infty}\left(\prod_{k=1}^{n} f(k)\right)^{-1} C_n^* z^n\right] \quad (49)$$

is certainly analytic and regular for $|z| < 1$. Let z_1, $z_2, \cdots, z_k, \cdots$ be a convergent sequence of complex numbers within the unit circle. If $\langle\psi \mid z_k\rangle = 0$, $k = 1, 2, \cdots$, then $\langle\psi \mid z\rangle$ must vanish identically for all z with $|z| < 1$. Therefore, $C_n = 0, n = 0, 1, 2, \cdots$, and the only vector orthogonal to all $|z_k\rangle$ is the null vector. It follows that the $|z\rangle$ states associated with any convergent sequence of the above type are complete. This completes the proof.

The relation between the argument of z and the phase of the oscillator in the state $|z\rangle$ may be brought out by noting, first of all, a trivial consequence of Eq. (45), namely,

$$\langle C\rangle \equiv \tfrac{1}{2}\langle z|(U + U^+)|z\rangle = \tfrac{1}{2}(z + z^*) = |z|\cos\phi, \quad (50)$$

where $\phi = \arg z$. Thus the expectation value of C approaches $\cos\phi$ as $|z| \to 1$.

Now, Eq. (47) shows that the normalization factor $|A|^2$, as a function of $|z|^2$, decreases monotonically from unity to a limiting value of zero as $|z|^2$ goes from $0 \to 1$. Further, a short calculation gives the following expression for the expectation value of the number

operator ("mean occupation number"):

$$\langle N\rangle = -\frac{|z|^2}{|A|^2}\frac{d(|A|^2)}{d(|z|^2)}. \tag{51}$$

This results in the not unexpected conclusion that $\langle N\rangle$ becomes indefinitely large as $|z| \to 1$. In fact, if the derivative in Eq. (51) is nonvanishing at $|z|^2 = 1$, one could infer that, for $|z|^2 \sim 1$,

$$\langle N\rangle \sim |z|^2/(1 - |z|^2). \tag{52}$$

(The arguments for these conclusions are given in Appendix A.)

The uncertainty in C may be found by writing

$$\begin{aligned}\langle C^2\rangle &= \tfrac{1}{4}\langle (U^2 + U^{\dagger 2} + 2U^\dagger U + [U, U^\dagger])\rangle \\ &= \tfrac{1}{2}|z|^2(1 + \cos 2\phi) + \tfrac{1}{4}\langle [U, U^\dagger]\rangle,\end{aligned} \tag{53}$$

so that

$$\begin{aligned}(\Delta C)^2 \equiv \langle C^2\rangle - \langle C\rangle^2 &= \tfrac{1}{4}\langle [U, U^\dagger]\rangle \\ &= \tfrac{1}{4}|A|^2\Bigg[f^2(1) + \sum_{n=1}^{\infty}(f^2(n+1) - f^2(n)) \\ &\quad \times \left(\prod_{k=1}^{n} f^2(k)\right)^{-1}|z|^{2n}\Bigg].\end{aligned} \tag{54}$$

In Appendix B it is shown that $(\Delta C)^2$ becomes indefinitely small as $|z| \to 1$. [The arguments involving $\langle C\rangle$ and $(\Delta C)^2$ hold *mutatis mutandis* for the operator S.]

For the case $f(n) = 1$ $(n > 0)$ the $|z\rangle$ states are easily seen to be

$$|z\rangle = \frac{1}{(1 + \langle N\rangle)^{\frac{1}{2}}}\sum_{n=0}^{\infty}\left(\frac{\langle N\rangle}{1 + \langle N\rangle}\right)^{\frac{1}{2}n} e^{in\phi}\,|n\rangle, \tag{55}$$

where

$$z = [\langle N\rangle/(1 + \langle N\rangle)]^{\frac{1}{2}}e^{i\phi}. \tag{56}$$

The expressions corresponding to Eqs. (50) and (54) are

$$\langle C\rangle = [\langle N\rangle/(1 + \langle N\rangle)]^{\frac{1}{2}}\cos\phi \tag{57}$$

and

$$(\Delta C)^2 = [4(1 + \langle N\rangle)]^{-1}. \tag{58}$$

If one uses these $|z\rangle$ states to form the superposition of projection operators

$$\rho = \int_0^{2\pi}\frac{d\phi}{2\pi}\,|z\rangle\langle z|, \tag{59}$$

the result is the Gaussian density operator.[5] As with the coherent states of Glauber, these $|z\rangle$ states are not orthogonal for different values of z. However, two states with the same $\langle N\rangle$ but different ϕ have the scalar product

$$\langle |z|\,e^{i\phi'}\,|\,|z|\,e^{i\phi}\rangle = (1 - |z|^2)/(1 - |z|^2\,e^{i(\phi-\phi')}) \tag{60}$$

so that they become approximately orthogonal for large $\langle N\rangle$. In this case the factor $(2\pi)^{-1}$ in Eq. (59) may be thought of as approximately representing a uniform probability density for the distribution of ϕ over the interval $(0, 2\pi)$.

One might expect that the situation would not be markedly different for the $|z\rangle$ states formed by using $f_1(n)$ or $f_2(n)$ of Eqs. (33) and (36). Since these sequences decrease monotonically to unity with increasing n, the quantities $\prod_{k=1}^{n} f(n)$ increase monotonically with n. However, we have seen in Eq. (43) that these products are bounded. Therefore, they converge as $n \to \infty$. In practical terms, the convergence is rather rapid. [For example, $f_1(1) = 1.12$, $f_1(6) = 1.00$ to three significant figures.] Thus the normalization factor $|A|^2$ should not be a substantially different function of $|z|^2$ than it is for the case of Eq. (55) where $|A|^2 = (1 - |z|^2)$. For large $\langle N\rangle$, certainly, Eq. (52) should hold quite accurately, which in turn implies the relation between $\langle N\rangle$ and $|z|$ given in Eq. (56). The density operator of Eq. (59) might then be expected to have properties very similar to those of the Gaussian density operator. In short, the inherent "statistics" of the $|z\rangle$ states are not strongly dependent on a reasonable, physically motivated choice of C and S. [This would not be the case for the $f(n)$ of Eq. (31), although they satisfy our formal requirements.]

We note also that the $|z\rangle$ states are ones for which the Schwartz inequality

$$(\Delta C)^2(\Delta S)^2 \geq |\langle (C - \langle C\rangle)(S - \langle S\rangle)\rangle|^2 \tag{61}$$

becomes an equality. However, the above uncertainty product does not reach a minimum for any $|z| < 1$, but approaches zero in the limit as $\langle N\rangle \to \infty$.[16]

From the information theoretic point of view,[17] the Gaussian density operator represents a state of the oscillator in which only the mean occupation number is specified. The fact that it is obtained by a uniform superposition of $|z\rangle\langle z|$ over $\arg z$ is suggestive, since such an averaging process "washes out" phase information. This tempts the speculation that in some sense the $|z\rangle$ states should represent pure states of maximum phase resolution for given $\langle N\rangle$. However, such speculation runs afoul of the fact that the coherent states of Glauber, for example, seem to be just as good, if not better, as far as uncertainties in C and S are concerned. Thus, for a Glauber state of large $\langle N\rangle$,[6]

$$(\Delta C)^2 \sim (\sin^2\phi)/4\langle N\rangle. \tag{62}$$

[This angle dependence of $(\Delta C)^2$, which is what one would reasonably expect, is not present in Eq. (58).] Thus the physical significance, if any, of $|z\rangle$ remains an open question.

ACKNOWLEDGMENT

The authors wish to express their appreciation to Professor J. M. Knight for informative discussions.

APPENDIX A

Let $|z|^2 = x$ and $|A|^2 = f(x)$. $f(x)$ decreases monotonically from unity at $x = 0$ to zero at $x = 1$. If we choose an arbitrarily small (positive) number ϵ, the mean value theorem says that there is a δ with $0 \leq \delta \leq \epsilon$ such that

$$-f(1-\epsilon) = f'(1-\delta)\epsilon, \tag{A1}$$

where the prime indicates differentiation. Then

$$-\frac{(1-\delta)}{f(1-\delta)} f'(1-\delta) = \frac{(1-\delta)}{\epsilon}\frac{f(1-\epsilon)}{f(1-\delta)} \geq \frac{(1-\delta)}{\epsilon}. \tag{A2}$$

[The inequality follows from the monotone decreasing property of $f(x)$.] This establishes the indefinite increase of $\langle N\rangle$ as $|z|^2 \to 1$. If $f'(x)$ is finite and does not vanish at $x = 1$, we may directly write the approximation

$$-\frac{(1-\delta)}{f(1-\delta)} f'(1-\delta) \sim \frac{(1-\delta)}{\delta}. \tag{A3}$$

This gives Eq. (52).

APPENDIX B

We show that $(\Delta C)^2$ in Eq. (54) becomes small for large $\langle N\rangle$ by considering the cases of monotone decreasing and monotone increasing $f(n)$ separately.

For $f(n)$ monotone decreasing to unity, we have immediately

$$(\Delta C)^2 \leq \tfrac{1}{4}|A|^2 f^2(1). \tag{B1}$$

The desired result follows from the behavior of $|A|^2$.

In the case of monotone increase to unity, write $f(n) = 1 - \epsilon_n$ with $0 \leq \epsilon_n < 1$. Then

$$f^2(n+1) - f^2(n) = (\epsilon_n - \epsilon_{n+1})(2 - \epsilon_n - \epsilon_{n+1}) \leq 2\epsilon_n, \tag{B2}$$

so that $f^2(n+1) - f^2(n)$ is dominated by a sequence which is monotone decreasing to zero. Equation (47) shows that, for a normalized $|z\rangle$,

$$|A|^2 \sum_{n=N_0}^{\infty} \left(\prod_{k=1}^{n} f^2(k)\right)^{-1} |z|^{2n} < 1 \tag{B3}$$

for any $N_0 \geq 1$. Given an arbitrarily small (positive) ϵ, there is an N_0 such that $\epsilon_n \leq \epsilon$ for $n \geq N_0$. It then follows that

$$(\Delta C)^2 \leq \tfrac{1}{4}|A|^2 \left[f^2(1) + 2\epsilon_1 \sum_{n=1}^{N_0-1} \left(\prod_{k=1}^{n} f^2(k)\right)^{-1}\right] + \tfrac{1}{2}\epsilon. \tag{B4}$$

Finally, by choosing $|z|$ sufficiently close to unity ($\langle N\rangle$ sufficiently large) the behavior of $|A|^2$ ensures that we can always make $(\Delta C)^2 \leq \epsilon$.

Note added in proof: A recent paper by Zak[18] discusses the phase operators which result if $f(n)$ is defined by means of the chain sequence generated by Eq. (34) above.

[1] An abbreviated, and less general, version of some of the results presented below was first published by one of us (E. C. L.) in Nuovo Cimento **56B**, 183 (1968); **57B**, 251 (1968).

[2] The idea of "cosine" and "sine" operators was developed by L. Susskind and J. Glogower, Physics **1**, 49 (1964).

[3] For a review of the quantum mechanical discussion of oscillator phase and an exhaustive list of references see P. Carruthers and M. M. Nieto, Rev. Mod. Phys. **40**, 411 (1968).

[4] H. S. Wall, *Analytic Theory of Continued Fractions* (Van Nostrand and Co., Inc., New York, 1948), p. 79ff.

[5] R. J. Glauber, Phys. Rev. **131**, 2766 (1963).

[6] P. Carruthers and M. M. Nieto, Phys. Rev. Letters **14**, 387 (1965).

[7] As far as the spectral properties of C and S are concerned, if $f(n) = |f(n)|\, e^{i\delta_n}$, we may equally well work with the unitarily equivalent operators $B^\dagger CB$ and $B^\dagger SB$, where

$$B|n\rangle = \exp\left(-i\sum_{j=0}^{n}\delta_j\right)|n\rangle$$

and δ_0 is arbitrary.

[8] Our basic commutation rules are consistent with writing $a = (N+1)^{\frac{1}{2}}U = U(N)^{\frac{1}{2}}$. This would involve the choice $f(n) = 1 - \delta_{0n}$.

[9] Although $C^2 + S^2$ is a constant of the motion (i.e., commutes with N), there is no possible choice of $f(n)$ consistent with the desired spectra such that $C^2 + S^2 = 1$ (unit operator). In fact, the only choice of $f(n)$ which would do this is $f(2n) = 0$; $f(2n+1) = (2)^{\frac{1}{2}}$. It then follows that $C|n\rangle = (2)^{-\frac{1}{2}}|n \pm 1\rangle$ depending on whether n is even or odd. The norm of C would then be $2^{-\frac{1}{2}}$.

[10] It is easily checked that $S = e^{\frac{1}{2}i\pi N} C e^{-\frac{1}{2}i\pi N}$ so that we may restrict our spectral analysis to C.

[11] See, for example, P. Halmos, *Introduction to Hilbert Space* (Chelsea, New York, 1957), 2nd ed., p. 51.

[12] Our proof of the "chain sequence" theorem is a slightly modified version of that given in Ref. 4, p. 86.

[13] g_0 can be zero.

[14] The limit on the left-hand side of Eq. (38) defines the spectral radius. See, for example, F. Riesz and B. Sz.-Nagy, *Functional Analysis* (Ungar, New York, 1955), p. 425.

[15] K. E. Cahill, Phys. Rev. **138**, B1566 (1965), Sec. VII.

[16] See, in this regard, the paper by R. Jackiw, J. Math. Phys. **9**, 339 (1968), especially Sec. IV.

[17] A. Katz, *Principles of Statistical Mechanics* (W. H. Freeman and Co., San Francisco, 1967).

[18] J. Zak, Phys. Rev. **187**, 1803 (1969). Note that Zak's $f(n)$ is not to be confused with our $f(n)$.

Electromagnetic Scattering from an Expanding Dielectric Slab*

RONALD J. POGORZELSKI
School of Engineering and Applied Science, University of California
Los Angeles, California 90024

(Received 17 November 1969)

The situation where a normally incident plane monochromatic electromagnetic wave is reflected by an expanding dielectric slab is analyzed by means of the invariant-imbedding concept. An approximate expression for the spectrum of the reflected radiation is obtained.

INTRODUCTION

In recent years considerable work has been done in the area of the electrodynamics of moving media. Much of this work involves an integrated moving region; that is, there exists a single Lorentz frame in which all of the medium in the region is at rest. If the region is surrounded by a vacuum, the problem is merely a matter of Lorentz transformation of the fields from the rest frame of the medium in the region to the laboratory frame (since it is usually easier to solve for the fields in the rest frame of the medium).[1] However, if the velocity of the medium is a continuous function of position, the number of Lorentz frames with which one must contend becomes infinite and that is the situation which arises when a plane wave is scattered by an expanding dielectric obstacle. One might then resort to solution of Maxwell's equations in the laboratory frame using Minkowski's constitutive relations.[2] This conveniently takes into account the infinity of Lorentz frames but, in general, it is mathematically quite complicated.[3] In this paper a more efficient method of dealing with this situation is used in analyzing the reflection of a plane wave by an expanding dielectric slab. The dielectric to be considered is a nonconducting, nonmagnetic, nondispersive compressible fluid. The velocity and density of the fluid are specified *a priori* and are assumed to be unmodified by the presence of the electromagnetic radiation. The formulation is quasistatic and is carried out to first order in the velocity of the medium.

DESCRIPTION OF THE PROBLEM AND THE METHOD OF SOLUTION

The configuration to be considered is shown in Fig. 1. A unit-amplitude linearly polarized plane monochromatic electromagnetic wave is normally incident from the left on a slab of dielectric fluid. The velocity of the fluid is purely z directed, and the index of refraction n and the velocity $\beta = v/c$ of the fluid are considered to be independent given functions of position z and time t. Thus, the slab is stratified in the z direction. The reflected wave is characterized by its spectral density function R. It is assumed that the properties of the slab vary slowly enough to allow a quasistatic analysis; i.e., R is a function of frequency and time but time is treated as a constant. This means that the time required for the electromagnetic wave to completely penetrate the object must be much shorter than the time during which the properties of the object change significantly. Moreover, since the analysis leads to a spectrum function which depends on time, it must be required that this time dependence be much slower than the rate at which the electromagnetic fields oscillate; otherwise, the time-varying spectrum description of the scattered wave will be of little value.

The method to be employed is known as invariant imbedding. It appears that the first application of the invariance concept to electromagnetic scattering was done by Ambarzumian in his study of the scattering of light by a foggy medium.[4] A similar approach has since been used extensively in astrophysical problems.[5] This eventually led to the development of the method now known as invariant imbedding which has been very widely applied.[6] Recently, invariant imbedding has been applied in finding the scattering from inhomogeneous nonconducting obstacles[7–9] and Kritikos, Lee, and Papas used the concept to obtain the scattering of plane waves by a jet stream which is inhomogeneous by virtue of a nonuniform velocity profile.[10]

In applying this method, it will be assumed that the reflection from that portion of the slab to the right of a given plane z is known and the change in reflection due to the addition of an infinitesimal layer of material at this plane will be found. This will result in a differential equation for R which is to be integrated from the right boundary of the slab, where the reflection is known, to the left boundary, where it is to be found. R is defined in a Lorentz frame (τ, ζ) comoving with the fluid at z at time t and is, therefore, a function of τ, ω_{out}, ω_{in}, and ζ, where ω_{in} is the frequency of the incident wave viewed in the comoving frame (a function of τ and ζ) and ω_{out} is the frequency variable in the spectrum of the reflected wave as seen in the

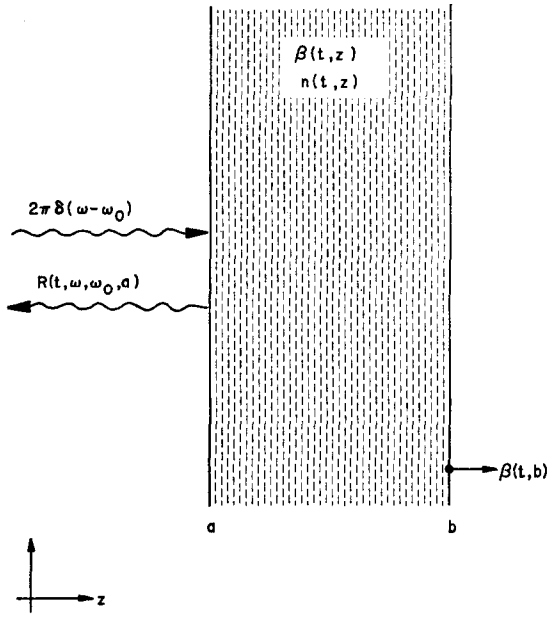


FIG. 1. Reflection from the general slab.

comoving frame (also a function of τ and ζ). Note that this approach circumvents calculation of the electromagnetic fields within the scatterer.

FORMULATION OF THE PROBLEM

Figure 2 shows the configuration used to derive the differential equation for $R(t, \omega_{\text{out}}, \omega_{\text{in}}, z)$. Recall that the laboratory frame coordinates are z and t and the comoving frame coordinates are τ and ζ; $z = \zeta$ when $t = \tau$. It is assumed that R is known at $\zeta + \Delta\zeta$ and a layer of fluid of thickness $\Delta\zeta$ having constant index $n = n(\tau, \zeta + \Delta\zeta)$ and a constant velocity gradient is added at $\zeta + \Delta\zeta$ and extends back to ζ. The material

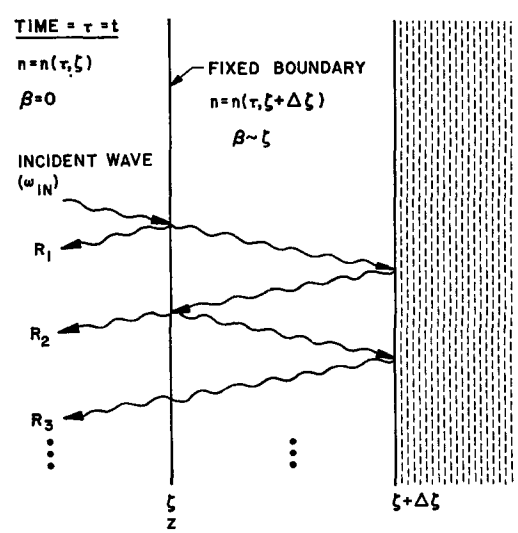


FIG. 2. Configuration for derivation of the invariant imbedding equation for the reflection function.

to the left of the added layer has a constant index $n = n(\tau, \zeta)$ and zero velocity in the comoving frame. The spectrum function of the incident wave is a δ function at ω_{in}. The R_i shown in Fig. 2 represent the spectrum functions of the results of the indicated multiple reflections.

It may be shown (by direct substitution into Maxwell's equations using the Minkowski constitutive relations for the moving medium)[2] that a plane wave solution for E within the added layer is

$$\vec{E} = \hat{e}_x E_0 \exp\left(i\frac{\omega}{c}\int_0^z \frac{n+\beta}{1+n\beta}\,dz - i\omega t\right), \quad (1)$$

where E_0 is a constant amplitude factor. Requiring continuous tangential E and H at the boundary ζ while using $R(t, \omega_{\text{out}}, \omega_{\text{in}}, \zeta + \Delta\zeta)$, which is assumed known, yields (to first order in $\Delta\zeta$)

$$R_1 = \frac{n(\tau,\zeta) - n(\tau,\zeta+\Delta\zeta)}{n(\tau,\zeta) + n(\tau,\zeta+\Delta\zeta)} 2\pi\delta(\omega_{\text{out}} - \omega_{\text{in}})$$
$$= -\frac{1}{2n}\frac{\partial n}{\partial\zeta} 2\pi\delta(\omega_{\text{out}} - \omega_{\text{in}})\Delta\zeta, \quad (2a)$$

$$R_2 = \left(\frac{2n(\tau,\zeta)}{n(\tau,\zeta) + n(\tau,\zeta+\Delta\zeta)}\right)\left[\exp\left(i\frac{n\omega'}{c}\Delta\zeta\right)\right]$$
$$\times\left[\left(\frac{1-n\Delta\beta}{1+n\Delta\beta}\right)R(\tau,\omega'',\omega',\zeta+\Delta\zeta)\right]$$
$$\times\left[\exp\left(i\frac{n\omega''}{c}\Delta\zeta\right)\right]\left(\frac{2n(\tau,\zeta+\Delta\zeta)}{n(\tau,\zeta) + n(\tau,\zeta+\Delta\zeta)}\right)$$
$$= R + \left(\frac{\partial R}{\partial\zeta} + \omega_{\text{out}} n\frac{\partial\beta}{\partial\zeta}\frac{\partial R}{\partial\omega_{\text{out}}} - \omega_{\text{in}} n\frac{\partial\beta}{\partial\zeta}\frac{\partial R}{\partial\omega_{\text{in}}}\right.$$
$$\left. + \frac{in}{c}(\omega_{\text{out}} + \omega_{\text{in}})R - 2n\frac{\partial\beta}{\partial\zeta}R\right)\Delta\zeta, \quad (2b)$$

where

$$\omega' = \omega_{\text{in}}(1 - n\Delta\beta), \quad \omega'' = \omega_{\text{out}}(1 + n\Delta\beta),$$

and

$$\Delta\beta = \frac{\partial\beta}{\partial\zeta}\Delta\zeta.$$

Similarly,

$$R_3 = \int_{-\infty}^{\infty} R(\tau,\bar{\omega},\omega_{\text{in}},\zeta)\frac{1}{2n}\frac{\partial n}{\partial\zeta}R(\tau,\omega_{\text{out}},\bar{\omega},\zeta)\frac{d\bar{\omega}}{2\pi}\Delta\zeta. \quad (2c)$$

The factors $1 \pm n\Delta\beta$ arise because R at $\zeta + \Delta\zeta$ and R at ζ are defined in different Lorentz frames. Now, R_j for $j > 3$ is of second or higher order in $\Delta\zeta$ and may be neglected. Thus $R(\zeta) = R_1 + R_2 + R_3$ and, in the

limit as $\Delta\zeta \to 0$, one obtains

$$\frac{\partial R}{\partial \zeta} + \omega_{\text{out}} n \frac{\partial \beta}{\partial \zeta}\frac{\partial R}{\partial \omega_{\text{out}}} - \omega_{\text{in}} n \frac{\partial \beta}{\partial \zeta}\frac{\partial R}{\partial \omega_{\text{in}}} = \frac{1}{2n}\frac{\partial n}{\partial \zeta}\left(2\pi\delta(\omega_{\text{out}} - \omega_{\text{in}}) - \int_{-\infty}^{\infty} R(\tau, \bar{\omega}, \omega_{\text{in}}, \zeta) R(\tau, \omega_{\text{out}}, \bar{\omega}, \zeta)\frac{d\bar{\omega}}{2\pi}\right) - i\frac{n}{c}(\omega_{\text{out}} + \omega_{\text{in}})R + 2n\frac{\partial \beta}{\partial \zeta}R. \quad (3)$$

Re-expressing the equation in terms of laboratory coordinates (quasistatically and to first order in β) gives

$$\frac{\partial R}{\partial z} + \omega_{\text{out}} n \frac{\partial \beta}{\partial z}\frac{\partial R}{\partial \omega_{\text{out}}} - \omega_{\text{in}} n \frac{\partial \beta}{\partial z}\frac{\partial R}{\partial \omega_{\text{in}}} = \frac{1}{2n}\frac{\partial n}{\partial z}\left(2\pi\delta(\omega_{\text{out}} - \omega_{\text{in}}) - \int_{-\infty}^{\infty} R(t, \bar{\omega}, \omega_{\text{in}}, z) R(t, \omega_{\text{out}}, \bar{\omega}, z)\frac{d\bar{\omega}}{2\pi}\right) - i\frac{n}{c}(\omega_{\text{out}} + \omega_{\text{in}})R + 2n\frac{\partial \beta}{\partial z}R. \quad (4)$$

That is, $\zeta \to z$ and $\tau \to t$ and the form of the equation is unchanged. It should be emphasized that R, ω_{out}, and ω_{in} are still evaluated in the comoving frame at each position z and time t.

SOLUTION OF THE PROBLEM

For a sufficiently tenuous slab, the integral term may be neglected and the equation becomes linear. The solution, valid for $\beta^2 \ll 1$, is easily obtained by the method of characteristics. The result is

$$R = \int_a^b \frac{1}{2n}\frac{\partial n}{\partial z} 2\pi\delta\left(\omega - \omega_0 + 2\omega_0\int_a^z n\frac{\partial \beta}{\partial z'}dz'\right) \times \exp\left(-\int_a^z 2n\frac{\partial \beta}{\partial z'}dz'\right) \times \exp\left(\frac{i}{c}(\omega + \omega_0)\int_a^z n\,dz'\right) dz + R(t, \omega_{\text{out}}(b), \omega_{\text{in}}(b), b)\exp\left(-\int_a^b 2n\frac{\partial \beta}{\partial z'}dz'\right) \times \exp\left(\frac{i}{c}(\omega + \omega_0)\int_a^b n\,dz'\right), \quad (5)$$

where

$$\omega_{\text{out}}(b) = \omega - \omega_0\int_b^a n\frac{\partial \beta}{\partial z'}dz',$$

$$\omega_{\text{in}}(b) = \omega_0 + \omega_0\int_b^a n\frac{\partial \beta}{\partial z'}dz'.$$

Carrying out the integration of the δ function yields

$$R(t, \omega, \omega_0, a) = -\frac{2\pi}{\omega_0}\frac{(2n)^{-1}\partial n/\partial z}{2n\,\partial\beta/\partial z}\bigg|_{z=z(\omega)} \exp\left(-\int_a^{z(\omega)} 2n\frac{\partial \beta}{\partial z'}dz'\right) \times \exp\left(\frac{i}{c}(\omega + \omega_0)\int_a^{z(\omega)} n\,dz'\right) W(\omega; \omega(b), \omega(a)) + R(t, \omega_{\text{out}}(b), \omega_{\text{in}}(b), b)\exp\left(-\int_a^b 2n\frac{\partial \beta}{\partial z'}dz'\right) \times \exp\left(\frac{i}{c}(\omega + \omega_0)\int_a^b n\,dz'\right), \quad (6)$$

where

$$\omega(z) = \omega_0\left(1 - 2\int_a^z n\frac{\partial \beta}{\partial z'}dz'\right), \quad (6'a)$$

$$W(\xi; p, q) = 1, \quad \text{for} \quad p \le \xi \le q, \quad = 0, \quad \text{otherwise}, \quad (6'b)$$

and it is assumed that (6'a) is uniquely invertible. In many cases $\exp\left[-\int_a^z 2n(\partial\beta/\partial z')\,dz'\right] \approx 1$ so that, as a "lowest-order" approximation, one may use the following formula:

$$R(t, \omega, \omega_0, a) = -\frac{2\pi}{\omega_0}\frac{(2n)^{-1}\partial n/\partial z}{2n\,\partial\beta/\partial z}\bigg|_{z=z(\omega)} \times \exp\left(\frac{i}{c}(\omega + \omega_0)\int_a^{z(\omega)} n\,dz'\right) W(\omega; \omega(b), \omega(a)) + R(t, \omega_{\text{out}}(b), \omega_{\text{in}}(b), b) \times \exp\left(\frac{i}{c}(\omega + \omega_0)\int_a^b n\,dz'\right). \quad (7)$$

This expression holds for a tenuous slab with continuous $n(t, z)$ and $\beta^2 \ll 1$. The first term gives the contribution from the interior of the slab and is band-limited due to the presence of the W function. The second term accounts for a possible discontinuity in n at the plane $z = b$ and for a possible scatterer located to the right of $z = b$. It remains for us to deal with the possibility of an index discontinuity at $z = a$.

Suppose $R(t, \omega, \omega_0, a + \epsilon)$ is known and $R(t, \omega, \omega_0, a - \epsilon)$ is desired, given that n is discontinuous at $z = a$, i.e., $\lim [n(t, a + \epsilon) - n(t, a - \epsilon)] \neq 0$, as $\epsilon \to 0$. This situation is shown in Fig. 3. The laboratory frame is comoving with the discontinuity at $z = a$. Referring to the figure, we easily see that the limit of the expression for R immediately to the left of the

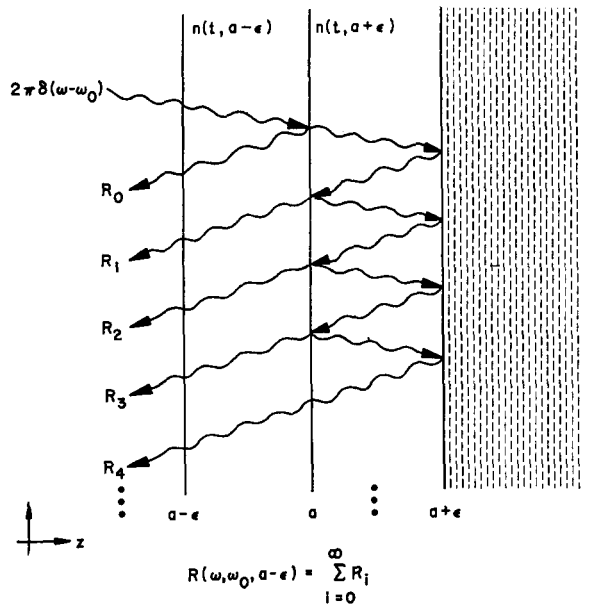


FIG. 3. Discontinuity in n at $z = a$.

discontinuity as $\epsilon \to 0$ is

$$\lim_{\epsilon\to 0} R(t, \omega, \omega_0, a - \epsilon)$$
$$= \lim_{\epsilon\to 0} \Big(\vec{r}_0(t) + \vec{t}_0(t)R(t, \omega, \omega_0, a + \epsilon)\vec{t}_0(t) + \int_{-\infty}^{\infty} \vec{t}_0(t)R(t, \omega, \omega', a + \epsilon)\vec{r}_0(t) \times R(t, \omega', \omega_0, a + \epsilon)\vec{t}_0(t)\frac{d\omega'}{2\pi} + \cdots\Big), \quad (8)$$

where

$$\vec{r}_0(t) = [n(t, a + \epsilon) - n(t, a - \epsilon)]/[n(t, a + \epsilon) + n(t, a - \epsilon)] = -\overleftarrow{r}_0(t),$$
$$\vec{t}_0(t) = 2n(t, a + \epsilon)/[n(t, a + \epsilon) + n(t, a - \epsilon)],$$
$$\overleftarrow{t}_0(t) = 2n(t, a - \epsilon)/[n(t, a + \epsilon) + n(t, a - \epsilon)],$$

where the arrows indicate the direction of the incident wave for each reflection or transmission.

Equations (7) and (8) provide a complete prescription for obtaining the approximate spectrum of the wave scattered from the slab given $n(t, z)$ and $\beta(t, z)$. Consistent with the accuracy to which R is given by (7), one need only retain the first two terms in (8).

ADDITIONAL COMMENTS

It may be observed that to first order in β the magnitude of the solution R may be written as a function of $\omega - \omega_0$ only, the ω_0's not associated with ω in this way being treated as constants. This suggests writing the differential equation for R in the form

$$\frac{\partial R}{\partial z} + 2n\omega\frac{\partial\beta}{\partial z}\frac{\partial R}{\partial\omega} = \frac{1}{2n}\frac{\partial n}{\partial z}\Big(2\pi\delta(\omega - \omega_0) - \int_{-\infty}^{\infty} R(t, \bar{\omega} - \omega_0, z)R(t, \omega - \bar{\omega}, z)\frac{d\bar{\omega}}{2\pi}\Big) - i\frac{n}{c}(\omega + \omega_0)R + 2n\frac{\partial\beta}{\partial z}R, \quad (9)$$

which facilitates taking the Fourier transform. The result is

$$\frac{\partial\hat{R}}{\partial z} - n\Big(1 + y\frac{\partial\beta}{\partial z}\Big)\frac{\partial\hat{R}}{\partial y} = (1 - \hat{R}^2)\frac{1}{2n}\frac{\partial n}{\partial z} + \Big[3n\frac{\partial\beta}{\partial z} - \frac{2in}{c}\omega_0\Big(1 + y\frac{\partial\beta}{\partial z}\Big)\Big]\hat{R}, \quad (10)$$

where

$$\hat{R}(t, y, z) = \frac{1}{2\pi}\int_{-\infty}^{\infty} R(t, \omega - \omega_0, z)\exp\Big[-i\Big(\frac{\omega - \omega_0}{c}\Big)y\Big]\,d\omega.$$

This equation, though not as accurate as Eq. (4) as far as the phase is concerned, is more convenient to deal with if one wishes to account for the nonlinear term. As a check, this equation was solved for a tenuous slab (neglecting $\hat{R}^2$) and the inverse transform of the result is as follows:

$$R(t, \omega - \omega_0, a) = -\frac{2\pi}{\omega_0}\frac{(2n)^{-1}\partial n/\partial z}{2n\,\partial\beta/\partial z}\Big|_{z=z(\omega)}\exp\Big(-\int_a^{z(\omega)} 2n\frac{\partial\beta}{\partial z'}dz'\Big) \times \exp\Big[\frac{i}{c}(\omega + \omega_0)\int_a^{z(\omega)} n\exp\Big(\int_a^{z(\omega)} n\frac{\partial\beta}{\partial z'}dz'\Big)dz\Big] \times W[\omega; \omega(b), \omega(a)] + R\Big[t, [\omega - \omega(b)]\exp\Big(\int_a^b n\frac{\partial\beta}{\partial z'}dz'\Big), b\Big] \times \exp\Big(-\int_a^b 2n\frac{\partial\beta}{\partial z'}dz'\Big) \times \exp\Big[\frac{i}{c}[2\omega(b) - \omega + \omega_0] \times \int_a^b n\exp\Big(\int_a^z n\frac{\partial\beta}{\partial z'}dz'\Big)dz\Big], \quad (11)$$

where

$$\omega(z) = \omega_0 - 2\omega_0\Big[1 - \exp\Big(-\int_a^z n\frac{\partial\beta}{\partial z'}dz'\Big)\Big],$$

which is identical (to first order in β) with (6) except for the expected discrepancy in the phase of each term.

The use of $R(t, \omega - \omega_0, z)$ as indicated above is also helpful when dealing with an index discontinuity because it converts the integrals in (8) into convolutions. Fourier transformation then leads to a geometric series which may be summed. Inverse Fourier transformation results in

$$R(t, \omega - \omega_0, a - \epsilon) = \vec{r}_0(t) + \vec{t}_0(t)\vec{t}_0(t)\int_{-\infty}^{\infty} \frac{\hat{R}(t, y, a + \epsilon)}{1 - \dot{r}_0(t)\hat{R}(t, y, a + \epsilon)} \times \exp\left[i\left(\frac{\omega - \omega_0}{c}\right)y\right] dy. \quad (12)$$

CONCLUDING REMARKS

This paper presents a method of determining the frequency spectrum of the reflected electromagnetic wave resulting when a plane monochromatic electromagnetic wave is normally incident on a stratified nonconducting slab of fluid. Details of this formulation and a similar one for the transmission problem may be found in Ref. 11. The technique employed is also applicable, in principle, to a similar problem in which an electromagnetic wave is obliquely incident on the slab. There one would expect both a frequency spectrum and an angular spectrum in the reflected wave. The formulation, however, has not been carried out in detail as yet. Work is currently proceeding on the application of the invariant imbedding concept in a study of the scattering of a plane electromagnetic wave by an expanding dielectric sphere.

ACKNOWLEDGMENTS

This research was performed with the support of the Air Force Office of Scientific Research, Office of Aerospace Research, U.S. Air Force under grant number AFOSR-68-1400/9768-02.

The author is deeply appreciative of the National Science Foundation Traineeship and the Northrop Corporation Fellowship which were held by him during this research.

The author would like to express his appreciation of the guidance, inspiration, and encouragement of his adviser, Professor C. H. Papas. Helpful discussions with Professor J. Mathews and Dr. T. C. Mo of the California Institute of Technology are gratefully acknowledged.

* This work submitted as part of a Ph.D. dissertation at the California Institute of Technology, Pasadena, California, 1970.

[1] C. Yeh, J. Appl. Phys. **36**, 3513 (1965).

[2] H. Minkowski, Nachr. Kgl. Ges. Wiss. Göttingen, II. Math. Physik. Kl. **2**, 53 (1908).

[3] J.-A. Kong and D. K. Cheng, J. Appl. Phys. **39**, 2282 (1968).

[4] V. A. Ambarzumian, Dokl. Akad. Nauk SSSR **38**, 229 (1943).

[5] S. Chandrasekhar, *Radiative Transfer* (Dover, New York, 1960).

[6] Stanley E. Lee, *Quasilinearization and Invariant Imbedding* (Academic, New York, 1968), Chap. 6, pp. 179–216.

[7] C. H. Papas, Caltech Antenna Laboratory Note, 1954.

[8] Guy W. Carlisle, Douglas Aircraft Company, Inc., Missile and Space Systems Division, Paper No. 4148, 1966.

[9] R. W. Latham, Can. J. Phys. **46**, 1463 (1968).

[10] H. N. Kritikos, K. S. H. Lee, and C. H. Papas, Radio Sci. **2**, New Series 991 (1967).

[11] R. J. Pogorzelski, Ph.D. dissertation, California Institute of Technology, 1970.

Radiation from a "Classical Spinning Electron"

MARIUS KOLSRUD
Institute of Physics, University of Oslo, Oslo, Norway

(Received 5 November 1969)

The electromagnetic field given by Cohn, according to which the radiation from a classical spinning point charge is independent of the particle's spin, is shown not to satisfy Maxwell's equations. References are given to earlier expressions, which also correspond with quantum theory.

In a recent paper,[1] Cohn tries to generalize the potentials from a static point charge with a magnetic dipole moment ("spinning" charge) to the covariant form corresponding to an arbitrary translational and "rotational" motion of the particle. He then obtains the surprising result that no term of the electromagnetic field caused by the dipole moment varies as the inverse distance, and that therefore this field does not contribute to the radiated energy and momentum. The paper contains no references to standard treatments of magnetic dipole radiation, nor to earlier papers on the problem in question.

In the present paper, we demonstrate that Cohn's expression for the field does not satisfy Maxwell's equations, because certain terms are lacking, even in the rest system.

For the special case of a charge at rest and with a time-varying magnetic moment $\boldsymbol{\mu}(t)$, Cohn's expression in his Eq. (12) for the dipole parts of the potentials reduces to the form

$$\mathbf{A}_\uparrow(\mathbf{r}, t) = \boldsymbol{\mu}(t') \times \hat{\mathbf{r}}/r^2, \quad \Phi_\uparrow = 0, \tag{1}$$

where, with $c = 1$,

$$t' = t - r, \quad \hat{\mathbf{r}} = \mathbf{r}/r. \tag{2}$$

Now consider Maxwell's equation, for $r > 0$:

$$\nabla \times \mathbf{B}_\uparrow - \frac{\partial \mathbf{E}_\uparrow}{\partial t} = 0. \tag{3}$$

The left-hand side becomes

$$\begin{aligned} \nabla \times (\nabla \times \mathbf{A}_\uparrow) + \frac{\partial}{\partial t}\left(\frac{\partial \mathbf{A}_\uparrow}{\partial t} + \nabla\Phi_\uparrow\right) &= \nabla\left(\nabla \cdot \mathbf{A}_\uparrow + \frac{\partial \Phi_\uparrow}{\partial t}\right) - \nabla^2 \mathbf{A}_\uparrow + \frac{\partial^2 \mathbf{A}_\uparrow}{\partial t^2} \\ &= \left(-\nabla^2 + \frac{\partial^2}{\partial t^2}\right)\frac{\boldsymbol{\mu}(t') \times \hat{\mathbf{r}}}{r^2} \\ &= -2\,\frac{\dot{\boldsymbol{\mu}} \times \hat{\mathbf{r}}}{r^3}, \end{aligned} \tag{4}$$

so that Eq. (3) is not satisfied.

However, the (well-known) form

$$\mathbf{A}_\uparrow = \frac{\boldsymbol{\mu}(t') \times \hat{\mathbf{r}}}{r^2} + \frac{\dot{\boldsymbol{\mu}}(t') \times \hat{\mathbf{r}}}{r} = \nabla \times \frac{\boldsymbol{\mu}(t')}{r} \tag{5}$$

satisfies

$$\left(\nabla^2 - \frac{\partial^2}{\partial t^2}\right)\mathbf{A}_\uparrow = 0. \tag{6}$$

The covariant generalization of (5) (with $\Phi_\uparrow = 0$) is given in the papers of Ellis,[2] Ward,[3] Kolsrud and Leer,[4] and Monaghan.[5] In the momentary rest system (with acceleration $\dot{\mathbf{v}} \neq 0$), they obtain

$$\mathbf{A}_\uparrow = \frac{\boldsymbol{\mu} \times \hat{\mathbf{r}}}{r^2} + \frac{\dot{\boldsymbol{\mu}} \times \hat{\mathbf{r}}}{r} + \frac{\boldsymbol{\mu} \times [(\dot{\mathbf{v}} \times \hat{\mathbf{r}}) \times \hat{\mathbf{r}}]}{r}, \tag{7}$$

$$\Phi_\uparrow = (\dot{\mathbf{v}} \times \boldsymbol{\mu}) \cdot \hat{\mathbf{r}}/r. \tag{8}$$

The integrated radiation intensity is found in invariant form in Ref. 4. When $\mathbf{v} = \dot{\mathbf{v}} = 0$, it reduces to the well-known expression $\frac{2}{3}\ddot{\boldsymbol{\mu}}^2$. In the same reference the total radiation intensity from a point particle with charge e and (rest system) magnetic moment $\boldsymbol{\mu}$ is also given. Besides the usual point charge term $\frac{2}{3}e^2\dot{u}_\mu\dot{u}_\mu$ (where $\dot{u}_\mu = d^2x_\mu/d\tau^2$, with $x_4 = it$ and $\tau =$ proper time) and the pure dipole terms, there is a cross term of the form

$$\tfrac{2}{3}eM_{\mu\nu}\dot{u}_\mu\ddot{u}_\nu. \tag{9}$$

Here $M_{\mu\nu}$ is the magnetic moment tensor, which in the rest system has the components

$$M^{(0)}_{ij} = \epsilon_{ijk}\mu_k, \quad M^{(0)}_{i4} = 0. \tag{10}$$

The invariant (9) may then be expressed by rest-system quantities as

$$\tfrac{2}{3}e\dot{\mathbf{v}} \cdot (\ddot{\mathbf{v}} \times \boldsymbol{\mu}). \tag{11}$$

This term corresponds exactly with the spin term in the low-energy and long-wave approximation of the quantum theoretical expression for spontaneous emission, as shown in Ref. 4.

[1] J. Cohn, J. Math. Phys. **10**, 799 (1969).
[2] J. R. Ellis, Proc. Cambridge Phil. Soc. **59**, 759 (1963); J. Math. Phys. **7**, 1185 (1966).
[3] G. N. Ward, Proc. Roy. Soc. (London) **279**, 562 (1964); Proc. Cambridge Phil. Soc. **61**, 547 (1965).
[4] M. Kolsrud and E. Leer, Physica Norvegica **2**, 181 (1967).
[5] J. J. Monaghan, J. Phys. **A1**, 112 (1968).

Conservation Theorem and the Method of Moments in Quantum Scattering

V. GALINDO
TRW Systems Group, Redondo Beach, Calif.

N. AMITAY
Bell Telephone Laboratories, Inc. Whippany, New Jersey

AND

Y. C. LEE
Department of Physics, State University of New York, Buffalo, New York

(Received 6 June 1969)

The unitary condition, or the constraint imposed by the conservation of probability, is examined for certain approximate solutions to physical scattering problems. The unitary condition is shown to be preserved for an approximate nonrelativistic S matrix derived by a rather generalized moments method. The unitary condition, or conservation theorem, is shown to be satisfied independent of the accuracy of the approximate solution under certain specifically defined conditions.

The unitarity condition, or the constraint imposed by the conservation of probability, has always played an important role in physical scattering problems. In recent years, many people have come to believe that the S matrix, together with the unitarity condition and certain assumed analytic and symmetry properties, may actually be sufficient to describe the dynamics of any physical system.[1] More recently, Veneziano[2] proposed a representation for the relativistic scattering amplitude, which does not satisfy the unitarity condition explicitly. Owing to the impressive successes of the Veneziano model, much effort is now being spent to make it satisfy the unitarity condition. Therefore, it is of interest to investigate, even in the nonrelativistic regime, the conditions imposed by the conservation of particles on the approximate quantum mechanical wavefunction. The related problem of scattering of electromagnetic waves in the classical regime has been studied elsewhere.[3]

We intend to show that the particle conservation theorem is satisfied by a broad class of approximations to solutions of the Schrödinger equation

$$-\frac{\hbar^2}{2m}\nabla^2\psi - \left(\frac{e\hbar}{imc}\mathbf{A}\cdot\nabla\psi\right) + \left(\frac{e^2A^2}{2mc^2} + e\phi\right)\psi + \frac{\hbar}{i}\frac{\partial\psi}{\partial t} = 0. \quad (1)$$

For simplicity, we will let $\mathbf{A} = 0$ here, although the conservation theorem is satisfied by approximate solutions of more complex forms of (1). We include scalar scattering from a static potential ϕ only.

With the particle current density given as

$$\mathbf{J} = (e\hbar/2im)(\bar{\psi}\nabla\psi - \psi\nabla\bar{\psi}),$$

where the bar denotes the complex conjugate, the conservation of particles is readily expressible as the integral

$$I = \iiint_V \nabla\cdot\mathbf{J}\,dv = \iint_S \mathbf{J}\cdot d\mathbf{s} = 0, \quad (2)$$

where S is the surface surrounding the volume V over which $\phi \neq 0$. We assume, for convenience, that the source of a steady-state stream of particles of energy E, incident upon ϕ, is exterior to V. Under these conditions, (1) becomes

$$(\nabla^2 + k^2)\psi = U\psi = U\psi_K, \quad (3)$$

where

$$U = (2me/\hbar^2)\phi, \quad k^2 = 2mE/\hbar^2,$$

and $\psi = \psi_V$ in V. The state ψ_i of the incident particles satisfies

$$(\nabla^2 + k^2)\psi_i = 0. \quad (4)$$

If G is the appropriate Green's function for $(\nabla^2 + k^2)$, then the scattered state ψ_s is given by

$$\psi_s = -\iiint_V GU\psi_V\,dv \quad (5)$$

and

$$\psi = \psi_i + \psi_s = \psi_i - \iiint_V GU\psi_V\,dv. \quad (6)$$

Now (6) is a Fredholm integral equation (of the second type) for ψ_V. If this integral equation is solved for approximate ψ_V, ψ_{V_a}, then the corresponding approximate ψ, ψ_a, is

$$\psi_a(x) = \psi_i(x) - \iiint_V G(x, x')U(x')\psi_{V_a}(x')\,dv', \quad (7)$$

where we note that $\psi_a \neq \psi_{V_a}$ in V. When this approximate ψ_a is substituted into the conservation equation (2), we obtain, noting from (7) that $(\nabla^2 + k^2)\psi_a = U\psi_{V_a}$,

$$I_a = \iiint_V \left[U\psi_{V_a}\left(\bar{\psi}_i - \iiint_V \bar{G} U \bar{\psi}_{V_a}\, dv' \right) - U\bar{\psi}_{V_a}\left(\psi_i - \iiint_V G U \psi_{V_a}\, dv' \right) \right] dv. \quad (8)$$

We now consider the conditions for which (8) will be identically zero.

An approximation for ψ_V may be given by

$$U\psi_{V_a} = \sum_{n=1}^{N} \alpha_n \Phi_n, \quad (9)$$

where the Φ_n, $n = 1, \cdots, N$, may form a basis in some subspace and the coefficients α_n are to be determined later. Substitution of (9) in (8) yields

$$I_a = \alpha \cdot [\bar{\psi}_i - \bar{\mathsf{G}}\bar{\alpha}] - \bar{\alpha} \cdot [\psi_i - \mathsf{G}\alpha], \quad (10)$$

where

$$\iiint_V U\bar{\psi}_{V_a}\, dv \iiint_V G U \psi_{V_a}\, dv' = \sum_{m=1}^{N} \bar{\alpha}_m \sum_{n=1}^{N} \alpha_n \mathsf{G}_{mn} \equiv \bar{\alpha} \cdot \mathsf{G}\alpha \quad (11)$$

and

$$\iiint_V U\bar{\psi}_{V_a}\psi_i\, dv = \sum_{m=1}^{N} \bar{\alpha}_m \psi_{im} \equiv \bar{\alpha} \cdot \psi_i. \quad (12)$$

The quantities α and ψ_i are matrix vectors, while G is a square matrix:

$$\mathsf{G}_{mn} \equiv \iiint_V dv \iiint_V dv' \bar{\Phi}_m(x) \mathsf{G}(x, x') \Phi_n(x'),$$

$$\psi_{im} \equiv \iiint_V dv \bar{\Phi}_m(x) \psi_i(x).$$

We now determine what class of α (or ψ_{V_a}) will satisfy $I_a = 0$. If, in general, the coefficients α_n are determined from a (matrix) solution of

$$\mathsf{H}\alpha = \psi_i - \mathsf{G}\alpha, \quad (13)$$

then we find that

$$I_a = \alpha \cdot \mathsf{H}\bar{\alpha} - \bar{\alpha} \cdot \mathsf{H}\alpha \to 0, \quad (14)$$

if H is Hermitian ($\mathsf{H}_{mn} = \bar{\mathsf{H}}_{nm}$). We note, in particular, that if H is the identity matrix $\mathsf{H} = 1$, then (13) is physically meaningful in that it determines α by a moments or Ritz[4] solution of the integral equation for ψ_V [Eqs. (5) and (6)]. The broad class of α defined by (13) and (14) may not, however, have direct relevance to the scattering problem.

[1] See, for example, G. F. Chew, *S-Matrix Theory of Strong Interactions* (Benjamin, New York, 1961).
[2] G. Veneziano, Nuovo Cimento **57A**, 190 (1968).
[3] N. Amitay and V. Galindo, IEEE Trans. Antennas Propagation **AP-17**, No. 6, 747 (1969).
[4] L. V. Kantorovich and V. I. Krylov, *Approximate Methods of Higher Analysis* (Interscience, New York, 1964).

Orthogonalization Methods*

H. C. SCHWEINLER
Health Physics Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee
AND
E. P. WIGNER
Princeton University, Princeton, New Jersey

(Received 20 October 1969)

An essentially unique method of forming an orthonormalized set of vectors from an independent set of n vectors is presented. The resulting set is proven to maximize a certain quartic form in the original vectors.

The best-known orthogonalization procedure, and the one most commonly used by mathematicians, is Schmidt's procedure.[1] It has the great advantage that it can be used, without further modifications, for a countably infinite set of vectors. On the other hand, the orthogonal set of vectors which it furnishes has no simple relation to the original set of vectors—the members of the set depend, in fact, on the order in which the vectors are arranged.

Physicists often use another orthogonalization method.[2] Let $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ be the vectors to be orthogonalized; they will be assumed to span a finite-dimensional complex Hilbert space. One then forms the Hermitian matrix M (Gram's matrix[1]), where

$$M_{kl} = (\mathbf{v}_k, \mathbf{v}_l). \tag{1}$$

If the vectors $\mathbf{v}$ are linearly independent, which is being assumed, M is positive definite. One forms its characteristic vectors $\mathbf{u}_\lambda$, with components $u_{k\lambda}$,

$$\sum_l M_{kl}u_{l\lambda} = p_\lambda u_{k\lambda}, \tag{2}$$

and the p are then positive. The $\mathbf{u}_\lambda$ will be assumed to be normalized so that

$$\sum_k u^*_{k\kappa}u_{k\lambda} = \delta_{\kappa\lambda}, \tag{3}$$

$$\sum_\kappa u^*_{k\kappa}u_{l\kappa} = \delta_{kl}. \tag{3'}$$

The orthogonal vectors $\mathbf{w}$ in the space of the $\mathbf{v}$ are then defined by means of the

$$\mathbf{w}^1_\kappa = \sum_k u_{k\kappa}\mathbf{v}_k. \tag{4}$$

One has

$$(\mathbf{w}^1_\kappa, \mathbf{w}^1_\lambda) = \sum_k \sum_l u^*_{k\kappa}u_{l\lambda}(\mathbf{v}_k, \mathbf{v}_l) = \sum_{kl} u^*_{k\kappa}M_{kl}u_{l\lambda}$$
$$= \sum_k u^*_{k\kappa}p_\lambda u_{k\lambda} = \delta_{\kappa\lambda}\cdot p_\lambda. \tag{5}$$

Hence, if one sets

$$\mathbf{w}_\kappa = p_\kappa^{-\frac{1}{2}}\mathbf{w}^1_\kappa = \sum_k p_\kappa^{-\frac{1}{2}}u_{k\kappa}\mathbf{v}_k, \tag{6}$$

the $\mathbf{w}$ form an orthonormal set. Since the p_κ are positive, $p_\kappa^{-\frac{1}{2}}$ can also be assumed to be positive. Conversely,

$$\mathbf{v}_k = \sum_\kappa u^*_{k\kappa}p_\kappa^{\frac{1}{2}}\mathbf{w}_k. \tag{6'}$$

As matrix equations, (6) and (6') are $\mathbf{w} = \mathbf{v}up^{-\frac{1}{2}}$ and $\mathbf{v} = \mathbf{w}p^{\frac{1}{2}}u^\dagger$, respectively, p being a diagonal matrix with diagonal elements p_κ.

The set of $\mathbf{w}$ is, evidently, independent of the order in which the $\mathbf{v}$ are arranged, and one expects it to satisfy some extremal condition which distinguishes it among all the sets of orthonormal vectors

$$\mathbf{Z}_\kappa = \sum_\lambda S_{\lambda\kappa}\mathbf{w}_\lambda, \tag{7}$$

where S is an arbitrary unitary matrix. Of course, if S is a phase matrix $\exp(i\delta) = \text{diag}\,(\exp(i\delta_1), \exp(i\delta_2), \cdots, \exp(i\delta_n))$, or a permutation matrix P, or a product $\exp(i\delta)P$ of such, the resulting set $\mathbf{Z}$ will be identical with the set of $\mathbf{w}$, except possibly for phase. One such extremal theorem is the subject of this paper. It states that the quantity

$$m = \sum_\kappa \left(\sum_k |(\mathbf{Z}_\kappa, \mathbf{v}_k)|^2\right)^2 \tag{8}$$

assumes its maximal value of $\sum p_\kappa^2$ for the $\mathbf{w}$ set of orthogonal vectors, i.e., if the S in (7) is a unit matrix or a generalized permutation matrix. In one sense, (8) is the simplest expression involving the original set of vectors $\mathbf{v}$ and the orthonormal set $\mathbf{Z}$: If one does not square the sum over k in (8), the result is simply the sum of the squares of the lengths of the $\mathbf{v}$ and, hence, independent of the S in (7). Perhaps an even more natural extremal principle would be to set

$$m' = \sum_{k\kappa} |(\mathbf{Z}_\kappa, \mathbf{v}_k)|^4, \tag{8'}$$

but this does not lead to a simple set of equations such as the preceding one.

In order to express m in terms of S and $\mathbf{w}$, we first

calculate

$$(\mathbf{w}_\lambda, \mathbf{v}_k) = \left(\mathbf{w}_\lambda, \sum_\kappa p_\kappa^{\frac{1}{2}} u_{k\kappa}^* \mathbf{w}_\kappa\right) = p_\lambda^{\frac{1}{2}} u_{k\lambda}^*, \tag{9a}$$

and hence

$$(\mathbf{Z}_\kappa, \mathbf{v}_k) = \sum_\lambda S_{\lambda\kappa}^* p_\lambda^{\frac{1}{2}} u_{k\lambda}^*, \tag{9b}$$

and, because of (3),

$$\sum_k |(\mathbf{Z}_\kappa, \mathbf{v}_k)|^2 = \sum_k \sum_{\lambda\mu} S_{\lambda\kappa} p_\lambda^{\frac{1}{2}} u_{k\lambda} S_{\mu\kappa}^* p_\mu^{\frac{1}{2}} u_{k\mu}^*$$
$$= \sum_\lambda |S_{\lambda\kappa}|^2 p_\lambda, \tag{9c}$$

so that

$$m = \sum_\kappa \left(\sum_\lambda |S_{\lambda\kappa}|^2 p_\lambda\right)^2. \tag{9d}$$

The p_λ are given by the $\mathbf{v}$, and we want to find the maximum of (9d) by varying the unitary matrix S. It is clear that the maximum exists, since (9d) is bounded and the domain of S is compact. We will assume for the present, further, that the p_λ are all different—the maximum of m is a continuous function of the p_λ, and its value for a set in which some of the p_λ are equal can be obtained as a limit of the maxima of m for a set of p all the p_λ of which are different.

The maximum is found, and the above statements about the generalized permutation matrix are verified, by observing that we can write m in the form

$$m = \sum_\kappa \{(S^\dagger p S)_{\kappa\kappa}\}^2$$
$$= \sum_\kappa \sum_\lambda (S^\dagger p S)_{\kappa\lambda}(S^\dagger p S)_{\lambda\kappa} - \sum_\kappa \sum_\lambda{}' (S^\dagger p S)_{\kappa\lambda}(S^\dagger p S)_{\lambda\kappa}$$
$$= \sum_\lambda p_\lambda^2 - \sum_\kappa \sum_\lambda{}' |(S^\dagger p S)_{\lambda\kappa}|^2. \tag{10}$$

m is bounded above by $\sum p_\lambda^2$, and it can attain this bound if $S^\dagger p S$ is a diagonal matrix d. In this case, $pS = Sd$, or

$$p_\lambda S_{\lambda\kappa} = S_{\lambda\kappa} d_\kappa, \tag{11}$$

so the set of p coincides with the set of d, and $S_{\lambda\kappa}$ vanishes unless $p_\lambda = d_\kappa$. Since we have assumed that the p_λ are all different, only one $S_{\lambda\kappa}$ is nonzero. The unitarity of S requires that the absolute value of this nonzero term be unity, and its phase is free. Thus S is a generalized permutation matrix.

The preceding discussion assumed that all the characteristic roots p of M are different. One can conclude, if some of the p are equal, by continuity arguments that no m, for any choice of the orthogonal $\mathbf{Z}$ in (8), can be larger than the m obtained for $\mathbf{Z}_\kappa = \mathbf{w}_\kappa$. However, if some p_κ are equal, the corresponding $\mathbf{w}$ can be subjected to a unitary transformation without affecting the value of m. In such a case, the S which yields the maximum value of m can contain such a unitary transformation. This was foreseeable: If some p are equal, the corresponding $\mathbf{w}$ are determined only up to a unitary transformation between them.

* Research done in part at Oak Ridge National Laboratory and sponsored by the U.S. Atomic Energy Commission under contract with Union Carbide Corporation.

[1] R. Courant and D. Hilbert, *Methoden der mathematischen Physik* (Springer, Berlin, 1931), 2nd ed., Vol. I, p. 4; see also pp. 29–32.

[2] R. Landshoff, Z. Physik **102**, 201 (1936). Actually, Landshoff's orthonormalized functions are, in our notation,

$$\sum_l (M^{-\frac{1}{2}})_{lk} \mathbf{v}_l = \sum_\kappa u_{k\kappa}^* \mathbf{w}_\kappa.$$

Inverse Scattering Problem for Dirac Particles. Explicit Expressions for the Values of the Potentials and Their Derivatives at the Origin in Terms of the Scattering and Bound-State Data

O. D. CORBELLA*
Istituto di Fisica dell'Università di Roma, Rome, Italy

(Received 13 January 1969)

The inverse scattering problem for the Dirac equation with two central potentials (one a relativistic scalar, the other the fourth component of a 4-vector) is investigated. Exact expressions are obtained for the values of the two potentials and their derivatives at the origin, as a function of the scattering data and bound-state parameters of the particle and the antiparticle for one angular momentum. The consideration is limited to the case of S and P waves ($j = \frac{1}{2}$) and to potentials satisfying a restrictive condition; explicit expressions for the values of the two potentials and their first three derivatives at the origin are given. In the case with only one potential present (the usual Dirac equation), these relationships provide explicit connections between the scattering and bound-state parameters for the particle and those for the antiparticle. For instance, an explicit and simple expression for the binding energy and for the corresponding normalization constant, in terms of the scattering phase shifts of the particle and of the antiparticle, is given in the case with only one bound state, say, for the particle, present (in S or P wave, with $j = \frac{1}{2}$). It is assumed that the two potentials are central, holomorphic, and satisfy the conditions $\int_0^\infty r^n |V_i(r)|\, dr < \infty$, $i = 1, 2$, $n = 0, 1, 2$.

1. INTRODUCTION

The inverse scattering problem for the Dirac equation with two potentials has been solved by the Gel'fand–Levitan method.[1] Specifically, the Gel'fand–Levitan equation has been established; this equation is determined once the scattering data and bound-state parameters of the particle and antiparticle at one angular momentum are assigned, and it has a unique solution. From the solution of this equation, one can find the potentials. This procedure, however, involves the solution of an integral matrix equation, that is, a system of four coupled integral equations; and this is a difficult problem. To bypass this difficulty, we have searched for another approach that will allow us to find the potentials, or at least some of their characteristics, as an explicit function of the scattering data and bound-state parameters.

To this end, we investigate an asymptotic expansion of the scattering parameters at large energy that allows us to calculate the values of the potentials and their derivatives at the origin; in principle, the whole potentials, which are by assumption holomorphic functions of r, could be reconstructed in this way. This method is similar to that developed by several authors in the case of the Schrödinger equation.[2–5] We treat with this method the S- and P-wave cases with $j = \frac{1}{2}$, and we explicitly construct expressions for the values of the two potentials and their first three derivatives at the origin in terms of the scattering and bound-state parameters.

In the previously mentioned article, Gasimov and Levitan[1] give a procedure of the Gel'fand–Levitan type to reconstruct the potentials from the scattering and bound-state data for one angular momentum (i.e., the phase shift for all values of the linear momentum k, the energies of the bound states, if any, and the corresponding normalization constants). Both the data for the particle and the antiparticle are needed as input, and two functions are produced, which play the role of local energy-independent potentials. These two functions appear in the so called "canonical form" of the Dirac equation, and Gasimov and Levitan show that all other equations of the Dirac type may be set in this form by an orthogonal transformation. Thus, without loss of generality, we choose one of the two potentials, $V_1(r)$, to transform as a relativistic scalar, and the other, $V_2(r)$, as the fourth component of a 4-vector.

With this choice, the results of the usual case are derived from our setting the scalar potential to zero. In this manner, expressions that connect the phase shifts and bound-state parameters of the particle with those of the antiparticle are also obtained.

It is interesting to note that, since it has been demonstrated that the two sets of input (for particle and antiparticle) determine univocally the two potentials, in the case of the usual Dirac equation (with only one potential), these inputs are superabundant. Therefore, there must be a relationship between them; in fact, presumably only one set of input data should be sufficient to determine the potential univocally. Therefore, solving the inverse scattering problem for

the usual Dirac equation implies finding relationships from which the scattering parameters, say, for the particle could (at least in principle) be calculated from those of the antiparticle. The explicit relationships given in this paper are not general enough to solve this problem, but they constitute a first step in this direction. Further investigation of this problem is now in progress.

The method of asymptotic expansion that we use has an intrinsic difficulty. It is necessary to impose the equality of the values of the first derivatives of the potentials at the origin (in the S-wave case; in the case of P waves, their sum must vanish), in order that the expansion at the origin of the irregular solution does not contain logarithmic terms. These logarithmic terms present a difficulty similar to that experienced in the case of the Schrödinger equation for $l > 0$, if the potential is a regular one, or for all partial waves, if the potential behaves as r^{-1} at the origin. Thus, the solution of this difficulty in our case, which would allow us to deal with more general potentials, is related to the solution of the same difficulty in the nonrelativistic case. This is still an open problem, although some progress towards its solution is implied by the nonrelativistic limit of our P-wave results; see Appendix A.

Throughout this article we consider central and holomorphic potentials that verify the restrictive condition

$$\int_0^\infty r^n |V_i(r)| \, dr < \infty, \quad i = 1, 2, n = 0, 1, 2.$$

In Sec. 2, the Dirac equation is discussed; in particular, its symmetry properties are displayed. These are subsequently used to obtain the P-wave results from those for the S-wave case.

In Sec. 3, we study the behavior at the origin of the solutions of this equation. The knowledge of this behavior allows us to find the conditions on the potentials which we need to impose to prevent the existence of logarithmic terms in the irregular solution. We will use these conditions in Sec. 7.

In Sec. 4, for completeness, we define the Jost solutions, the Jost functions, and their connections with the regular and irregular solutions.

In Sec. 5, we analyze the analyticity properties in the complex k plane of all the functions defined and their asymptotic behavior.

In Sec. 6, we construct a formula that allows us to calculate the square modulus of the Jost function in terms of the phase shifts and the bound-state energies.

In Sec. 7, we introduce a method to obtain the asymptotic expansion of the scattering parameters at large energy, which is a generalization of that given by Verde[2] for the Schrödinger equation. From this, we derive a system of recursion relations, which is useful in relating the coefficients of the asymptotic expansion of certain functions of the scattering parameters to the potentials and their derivatives.

In Sec. 8, we determine these coefficients in terms of the scattering parameters and, in this manner, we are able to calculate the values of the potentials and their derivatives at the origin.

In Sec. 9, we give the principal results of this paper, that is, the explicit exact expressions of the values of the potentials and their derivatives at the origin, in terms of the two sets of scattering parameters for the particle and the antiparticle. Expressions up to the third derivatives are given.

In Sec. 10, to study the usual Dirac equation, we particularize the results of the preceding section. We obtain formulas for the potential and its derivatives at the origin and give the relationships between the scattering parameters for the particle and those of the antiparticle. As an example, the value of the binding energy and of the corresponding normalization constant is given in terms of the (S- or P-wave) phase shift of the particle and the antiparticle, for the particular case when only one bound state for the particle (or for the antiparticle) is present.

The nonrelativistic limit of the expressions obtained is given in Appendix A. For S waves, we arrive at expressions that coincide with those of the Schrödinger case. Moreover, we obtain similar expressions in terms of the P-wave parameters, some of which constitute new results. This fact suggests a procedure to solve, at least in some cases, the Schrödinger problem for higher waves; we propose to come back to this question in a separate paper.

Several authors have derived explicit exact expressions for the values of the potential and its derivatives at the origin in the case of the Schrödinger equation. The first result of this type, consisting in an expression for the value of the potential at the origin, was given by Newton[6,7] and simultaneously by Faddeev.[8] Subsequently, Buslaev and Faddeev,[3] Percival,[9] and Roberts [4] have given an expression for the value of the second derivative at the origin. More recently, Calogero and Degasperis[5] have derived expressions for the potential and all its derivatives at the origin in terms of the S-wave parameters. In this last article, a review of the different methods of attacking this problem in the Schrödinger case is given; we refer to it for all other references in the Schrödinger case. In the case of the Dirac equation, the present work constitutes the first effort to derive this type of results; our approach is

patterned after that of Calogero and Degasperis.[5] For the Klein–Gordon equation, a similar work has been done by Degasperis.[10]

The system of units used in this article is such that $\hbar = c = 1$.

2. THE DIRAC EQUATION

The matrix form of the radial Dirac equation is

$$\frac{d}{dr}\varphi^{\lambda}(k, r) = \left[\left(\frac{\lambda}{r}\right)\rho_3 - [E - V_2(r)]i\rho_2 + [m + V_1(r)]\rho_1\right]\varphi^{\lambda}(k, r). \quad (2.1)$$

The solution is a two-component vector

$$\varphi^{\lambda}(k, r) = \begin{pmatrix} u_1^{\lambda}(k, r) \\ u_2^{\lambda}(k, r) \end{pmatrix},$$

and the matrices are the usual Pauli matrices

$$\rho_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad i\rho_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \rho_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (2.2)$$

where m is the mass of the scattering particle and λ is an integral number ($\neq 0$) related to the "orbital" and total angular momentum by

$$\text{total angular momentum} = j = |\lambda| - \tfrac{1}{2},$$

$$\text{"orbital" angular momentum} = l = |\lambda| - \tfrac{1}{2} + \tfrac{1}{2}\lambda/|\lambda|,$$

i.e.,

$$\lambda > 0, \quad l = \lambda; \quad \lambda < 0, \quad l = |\lambda| - 1. \quad (2.3)$$

Equation (2.1) has the following symmetry properties. It is invariant under the transformations

$$E \to -E, \quad V_2(r) \to -V_2(r), \quad \lambda \to -\lambda, \quad u_1^{\lambda}(k, r) \to u_2^{\lambda}(k, r), \quad u_2^{\lambda}(k, r) \to u_1^{\lambda}(k, r) \quad (2.4)$$

or

$$m \to -m, \quad V_1(r) \to -V_1(r), \quad \lambda \to -\lambda, \quad u_1^{\lambda}(k, r) \to -u_2^{\lambda}(k, r), \quad u_2^{\lambda}(k, r) \to u_1^{\lambda}(k, r). \quad (2.5)$$

Then, if the results for a positive λ are known, by the transformations

$$E \to -E, \quad V_2(r) \to -V_2(r), \quad u_1^{\lambda}(k, r) \to u_2^{\lambda}(k, r), \quad u_2^{\lambda}(k, r) \to u_1^{\lambda}(k, r) \quad (2.6)$$

or

$$m \to -m, \quad V_1(r) \to -V_1(r), \quad u_1^{\lambda}(k, r) \to -u_2^{\lambda}(k, r), \quad u_2^{\lambda}(k, r) \to u_1^{\lambda}(k, r), \quad (2.7)$$

we can find the results for the corresponding negative λ. Henceforth, we deal only with positive λ, unless we specifically indicate the contrary.

Note that the transformation (2.7) consists of a multiplication of the wavefunction by $i\rho_2$; since this matrix is not equal to its inverse, we cannot apply it twice to obtain the same results. With the transformation $-i\rho_2$, i.e.,

$$m \to -m, \quad V_1(r) \to -V_1(r), \quad u_1^{\lambda}(k, r) \to u_2^{\lambda}(k, r), \quad u_2^{\lambda}(k, r) \to -u_1^{\lambda}(k, r), \quad (2.7')$$

one can also arrive at the result with a different sign of λ. And, of course, by the successive application of (2.7) and (2.7′) we obtain the same result. Therefore, we decide by convention to use the transformation (2.7) to go from positive λ to the corresponding negative λ and (2.7′) to pass from negative λ to the corresponding positive λ.

As we see below, some of the solutions of (2.1) are double-valued functions of the complex variable k, having different values in the two Riemann sheets defined by sgn (Re E), $E = \pm(k^2 + m^2)^{\frac{1}{2}}$; therefore, we must distinguish between the two transformations (2.6) and (2.7) because, while (2.6) exchanges the Riemann sheets, (2.7) does not.

Finally, we emphasize another precaution we must take in applying transformations (2.6) or (2.7). If we apply them to functions that do not depend explicitly on the components, we have to look for their definitions to check the signs. For example, for the function $\Delta_{\lambda}(k)$ [see Eq. (4.3)], when we apply (2.7), we get

$$\Delta(k, E, V_1, V_2, m, -\lambda) = \Delta(k, E, -V_1, V_2, -m, \lambda),$$

while, applying (2.6), we get

$$\Delta(k, E, V_1, V_2, m, -\lambda) = -\Delta(k, -E, V_1, -V_2, m, \lambda).$$

3. REGULAR AND IRREGULAR SOLUTIONS

Let us now study the behavior at the origin of the solutions of Eq. (2.1). To this end we put Eq. (2.1) in the form of a system of two first-order differential coupled equations

$$u_1^{\lambda\prime}(k, r) = (\lambda/r)u_1^{\lambda}(k, r) - [E - m - V_1(r) - V_2(r)]u_2^{\lambda}(k, r),$$

$$u_2^{\lambda\prime}(k, r) = -(\lambda/r)u_2^{\lambda}(k, r) + [E + m + V_1(r) - V_2(r)]u_1^{\lambda}(k, r). \quad (3.1)$$

We introduce the expansions

$$u_1^{\lambda}(k, r) = r^A \sum_{n=0}^{\infty} a_n(k) r^n,$$

$$u_2^{\lambda}(k, r) = r^B \sum_{n=0}^{\infty} b_n(k) r^n. \quad (3.2)$$

These expressions are solutions of the system (3.1), for $A = \lambda$ and $a_1 = 0$, if $B = \lambda + 1$. The coefficients a_n and b_n are determined in terms of a_0. Since the

system (3.1) has a term r^{-1}, the other independent solution must have the form

$$Z_1^\lambda(k, r) = r^C \sum_{n=0}^{\infty} c_n(k) r^n + A_\lambda(k) \ln B(k) r \times r^\lambda \sum_{n=0}^{\infty} a_n(k) r^n,$$

$$Z_2^\lambda(k, r) = r^D \sum_{n=0}^{\infty} d_n(k) r^n + A_\lambda(k) \ln B(k) r \times r^{\lambda+1} \sum_{n=0}^{\infty} b_n(k) r^n, \quad (3.3)$$

which is a solution for $D = -\lambda$ and $d_1 = 0$, if $C = -\lambda + 1$.

The relationships between the coefficients are the following: for the solution of type (3.2), we have

$$a_{n+2} = \frac{-1}{n+2}\left[(E-m)b_n - \sum_{j=0}^{n} \frac{V_1^{(j)}(0) + V_2^{(j)}(0)}{j!} b_{n-j}\right], \quad (3.4a)$$

$$b_n = \frac{1}{n+2\lambda+1}\left[(E+m)a_n + \sum_{j=0}^{n} \frac{V_1^{(j)}(0) - V_2^{(j)}(0)}{j!} a_{n-j}\right], \quad (3.4b)$$

$$a_1 = 0 \quad (3.4c)$$

(note that a_0 is undetermined, since it depends on the normalization) and, for the solution (3.3),

$$c_n = \frac{1}{2\lambda-1-n}\left[(E-m)d_n - \sum_{j=0}^{n} \frac{V_1^{(j)}(0) + V_2^{(j)}(0)}{j!} d_{n-j}\right], \quad n < 2\lambda - 1, \quad (3.5a)$$

$$c_n = \frac{-1}{n-2\lambda+1}\left[A_\lambda a_{n-2\lambda+1} + (E-m)d_n - \sum_{j=0}^{n} \frac{V_1^{(j)}(0) + V_2^{(j)}(0)}{j!} d_{n-j}\right], \quad n > 2\lambda - 1, \quad (3.5b)$$

$$d_{n+2} = \frac{1}{n+2}\left[(E+m)c_n + \sum_{j=0}^{n} \frac{V_1^{(j)}(0) - V_2^{(j)}(0)}{j!} c_{n-j}\right], \quad n < 2\lambda - 1, \quad (3.5c)$$

$$d_{n+2} = \frac{1}{n+2}\left[-A_\lambda b_{n-2\lambda+1} + (E+m)c_n + \sum_{j=0}^{n} \frac{V_1^{(j)}(0) - V_2^{(j)}(0)}{j!} c_{n-j}\right], \quad n \geq 2\lambda - 1, \quad (3.5d)$$

$$A_\lambda = \frac{-1}{a_0}\left[(E-m)d_{2\lambda-1} - \sum_{j=0}^{2\lambda-1} \frac{V_1^{(j)}(0) + V_2^{(j)}(0)}{j!} d_{2\lambda-1-j}\right], \quad (3.5e)$$

$$d_1 = 0. \quad (3.5f)$$

d_0, $c_{2\lambda-1}$, and B are undetermined (d_0 depends on the normalization). To find these expressions, we have introduced the following power expansions of the potentials at the origin:

$$V_i(r) = \sum_{j=0}^{\infty} \frac{V_i^{(j)}(0)}{j!} r^j. \quad (3.6)$$

Here $V_i^{(j)}(0)$ indicates, of course, the j derivative of $V_i(r)$ evaluated at the origin.

In the particular case $\lambda = 1$, with which we will be concerned in Sec. 7, Eq. (3.5e) is written

$$A_1 = (d_0/a_0)[V_1'(0) + V_2'(0)]. \quad (3.7)$$

Then, if we wish to make the contribution of the logarithmic terms in the solution (3.3) vanish, we are required to have the values of the first derivatives at the origin of the two potentials be equal and of opposite signs. The presence of logarithmic terms, as we will see below, does not allow us to treat higher waves ($|\lambda| \neq 1$) with an asymptotic expansion method. (See also Appendix B.) However, it is possible to reduce a λ-wave Dirac equation to a $(\lambda - 1)$-wave Dirac equation which has a more complicated potential; then we apply the asymptotic expansion method to the last equation. We will treat this problem in a subsequent paper.

4. THE JOST SOLUTIONS AND THE JOST FUNCTIONS

In this section, for completeness, we review the definitions of all the functions we will need, which are the same as those in the case of a usual (one-potential) Dirac equation. (See, for example, Refs. 11 and 12.) In addition to (3.2) and (3.3), characterized by their behavior at the origin, we have two other solutions, the Jost solutions $f^\lambda(\pm k, r)$, characterized by their asymptotic behavior. [See formula 4 of Ref. 11 or formula (III-8/9) of Ref. 12.]

The regular and irregular solutions

$$\varphi^\lambda(k, r) = \begin{pmatrix} u_1^\lambda(k, r) \\ u_2^\lambda(k, r) \end{pmatrix}, \quad \psi^\lambda(k, r) = \begin{pmatrix} Z_1^\lambda(k, r) \\ Z_2^\lambda(k, r) \end{pmatrix} \quad (4.1)$$

can be given as linear combinations of the Jost solutions. For the regular solution we have

$$\varphi^\lambda(k, r) = \frac{1}{\Delta_\lambda(k)}[f^\lambda(-k)f^\lambda(k, r) - f^\lambda(k)f^\lambda(-k, r)], \quad (4.2)$$

where $\Delta_\lambda(k)$ is the Wronskian function of the Jost solutions $f^\lambda(k, r)$ and $f^\lambda(-k, r)$,

$$\begin{aligned}\Delta_\lambda(k) &= \Delta[f^\lambda(k, r), f^\lambda(-k, r)] \\ &= [f_1^\lambda(k, r) f_2^\lambda(-k, r) - f_1^\lambda(-k, r) f_2^\lambda(k, r)] \\ &= -2ik^{2\lambda-1}(E - m). \end{aligned} \tag{4.3}$$

Here f_i^λ indicates the ith component of the vector f^λ. With this definition of the Wronskian function, we immediately conclude that the Jost function is defined by

$$f^\lambda(k) = \Delta[\varphi^\lambda(k, r), f^\lambda(k, r)]. \tag{4.4}$$

It is a function of k only, which can be calculated taking the limit for $r \to 0$ and substituting the behavior at the origin of the regular solution

$$\begin{aligned} f^\lambda(k) &= \lim_{r\to 0} [a_0 r^\lambda f_2^\lambda(k, r) - b_0 r^{\lambda+1} f_1^\lambda(k, r)], \\ &= \lim_{r\to 0} a_0 r^\lambda f_2^\lambda(k, r). \end{aligned} \tag{4.5}$$

As usual, we define

$$f^\lambda(k) = |f^\lambda(k)| \exp [i\eta_\lambda(k)], \tag{4.6}$$

where $\eta_\lambda(k)$ is the phase shift of order λ, and from this we can define the λ element of the scattering matrix. (See formula 10 of Ref. 11 or the last formula on p. 126 of Ref. 12.)

Also, from (4.2), using (4.6), we can get the asymptotic behavior of the regular solution [See formula 9 of Ref. 11 or formula (III-13) of Ref. 12.]

Note that the irregular solution can be written in terms of the Jost solutions as

$$\psi^\lambda(k, r) = \frac{-1}{2|f^\lambda(k)|^2} [f^\lambda(-k) f^\lambda(k, r) + f^\lambda(k) f^\lambda(-k, r)] \tag{4.7}$$

and, from this and (4.2), we get

$$f^\lambda(k, r) = \frac{\Delta_\lambda(k)}{2f^\lambda(-k)} \varphi^\lambda(k, r) - f^\lambda(k)\psi^\lambda(k, r), \tag{4.8}$$

an expression that will be useful for the study of the behavior at the origin of the Jost solutions.

5. ANALYTICITY PROPERTIES. ASYMPTOTIC BEHAVIOR FOR LARGE k

We pass now to consider the above defined functions in the complex k-plane. As is well known from preceding treatments of the usual Dirac equations,[11,12] all the functions defined in the preceding sections are double-valued functions of k, since to every value of k there correspond two values of the energy

$$E = \pm(k^2 + m^2)^{\frac{1}{2}}.$$

The Jost solution $f^\lambda(k, r)$ can be extended into a double-valued function of the k complex variable that possesses branch points at $k = \pm im$. The two Riemann sheets, characterized by $\sigma = \text{sgn}\,[\text{Re}\,E] = \pm 1$, can be separated by introducing two cuts, $(-im, -i\infty)$ and $(+im, +i\infty)$. We write $f^{\lambda,\sigma}(k, r)$ for the Jost solution in the σ sheet.

If we impose the following restrictions on the potentials,

$$\int_0^\infty r^n |V_i(r)|\, dr < \infty, \quad i = 1, 2, \quad n = 0, 1, 2, \tag{5.1}$$

it is proved that $f^{\lambda,\sigma}(k, r)$ is an analytic function of k in the lower half-plane ($\text{Im}\,k \leq 0$), cut from $-im$ to $-i\infty$. The proof of this is similar to the proof given in the article of Barthélémy,[12] except that care must be taken to substitute for the function $V(r)$ the matricial function

$$\mathfrak{V}(r) = \begin{pmatrix} V_2(r) - V_1(r) & 0 \\ 0 & V_2(r) + V_1(r) \end{pmatrix}. \tag{5.2}$$

The reason for this substitution will be clear later [see (5.7) and (5.8)]. Aside from the differences implied by the introduction of this matrix, the demonstration follows the lines of the above mentioned work. Therefore, we omit this proof while we study explicitly the asymptotic behavior, in which case we will find differences because of the presence of the new potential.

In the same manner, it is proved that the regular solution $\varphi^\lambda(k, r)$ is an analytic function of k, for positive r and finite $|k|$, in the cut complex k-plane.

Let us now move on to the asymptotic behavior in k of the solutions of Eq. (3.1). Let $f_0^\lambda(k, r)$ and $f_0^\lambda(-k, r)$ be the two independent Jost solutions of the free equation (σ is fixed and is not indicated explicitly). We construct the matrix

$$\begin{aligned}\mathbf{K}(k, r, \xi) = \frac{1}{\Delta_\lambda(k)} \{&f_0^\lambda(k, r)[f_0^\lambda(-k, \xi)]^{\mathrm{T}} \\ &- f_0^\lambda(-k, r)[f_0^\lambda(k, \xi)]^{\mathrm{T}}\}. \end{aligned} \tag{5.3}$$

Note that

$$\mathbf{K}(k, r, r) = i\rho_2. \tag{5.4}$$

With these definitions, the integral matrix equation for $f^\lambda(k, r)$ is

$$f^\lambda(k, r) = f_0^\lambda(k, r) - \int_r^\infty \mathfrak{K}(k, r, \xi) f^\lambda(k, \xi)\, d\xi. \tag{5.5}$$

Here

$$\mathfrak{K}(k, r, \xi) = \mathbf{K}(k, r, \xi)\mathfrak{V}(\xi) \tag{5.6}$$

and $\mathfrak{V}$ is the matrix defined in (5.2).

Let us now define the matrix

$$L_\lambda(k, r) = \mathbf{T}_\lambda(k, r) f^\lambda(k, r), \tag{5.7}$$

with

$$\mathbf{T}_\lambda(k, r) = \begin{pmatrix} [f^\lambda_{0_1}(k, r)]^{-1} & 0 \\ 0 & [f^\lambda_{0_2}(k, r)]^{-1} \end{pmatrix} \quad (5.8)$$

so that

$$\mathbf{T}_\lambda(k, r) f^\lambda_0(k, r) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \quad (5.9)$$

Then the integral equation for $L_\lambda(k, r)$ is

$$L_\lambda(k, r) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \int_r^\infty \mathfrak{H}(k, r, \xi) L_\lambda(k, \xi)\, d\xi, \quad (5.10)$$

where

$$\mathfrak{H}(k, r, \xi) = \mathbf{T}_\lambda(k, r) \mathfrak{K}(k, r, \xi) \mathbf{T}^{-1}_\lambda(k, \xi). \quad (5.11)$$

By differentiation, we get from Eq. (5.10)

$$L'_\lambda(k, r) = \mathfrak{H}(k, r, r) L_\lambda(k, r) - \int_r^\infty \frac{\partial}{\partial r} \mathfrak{H}(k, r, \xi) L_\lambda(k, \xi)\, d\xi. \quad (5.12)$$

The asymptotic behavior for large k of $\mathfrak{H}(k, r, \xi)$ and its derivative are

$$\lim_{k\to\infty} \mathfrak{H}(k, r, r) = i \lim_{k\to\infty} \frac{k}{E} \begin{pmatrix} 0 & V_+(r) \\ V_-(r) & 0 \end{pmatrix}, \quad (5.13)$$

$$\lim_{k\to\infty} \frac{\partial}{\partial r} \mathfrak{H}(k, r, \xi) = \lim_{k\to\infty} k \cdot \frac{E}{k} \cdot e^{2ik(r-\xi)} \begin{pmatrix} V_-(\xi) & -V_+(\xi) \\ -V_-(\xi) & V_+(\xi) \end{pmatrix} \quad (5.14)$$

and, by substitution in (5.12), we obtain

$$\lim_{k\to\infty} L'_\lambda(k, r) = i \lim_{k\to\infty} \frac{k}{E} \begin{pmatrix} 0 & V_+(r) \\ V_-(r) & 0 \end{pmatrix} \lim_{k\to\infty} L_\lambda(k, r) - \int_r^\infty \lim_{k\to\infty} k \cdot \frac{k}{E} \cdot e^{2ik(r-\xi)} \begin{pmatrix} V_-(\xi) & -V_+(\xi) \\ -V_-(\xi) & V_+(\xi) \end{pmatrix} \times \lim_{k\to\infty} L_\lambda(k, \xi)\, d\xi. \quad (5.15)$$

By partial integration we get

$$\lim_{k\to\infty} L'_\lambda(k, r) = \tfrac{1}{2} i \lim_{k\to\infty} \frac{k}{E} \begin{pmatrix} V_-(r) & V_+(r) \\ V_-(r) & V_+(r) \end{pmatrix} \lim_{k\to\infty} L_\lambda(k, r) + O(k^{-1}), \quad (5.16)$$

for which the solution is

$$\lim_{k\to\infty} L_\lambda(k, r) = \exp\left[-\tfrac{1}{2} i \lim_{k\to\infty} \frac{k}{E} \int_r^\infty \begin{pmatrix} V_-(\xi) & V_+(\xi) \\ V_-(\xi) & V_+(\xi) \end{pmatrix} d\xi \right] \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \quad (5.17)$$

By diagonalizing the matrix in the exponential, we get

$$\lim_{k\to\infty} L_\lambda(k, r) = \exp\left[-i \lim_{k\to\infty} \frac{k}{E} \int_r^\infty V_2(\xi)\, d\xi \right] \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (5.18)$$

(note that the diagonalization does not change the vector). From this we infer the asymptotic behavior of $f^\lambda(k, r)$, which is

$$\lim_{\substack{k\to\infty \\ \mathrm{Im}\, k \le 0}} f^\lambda(k, r) = f^\lambda_0(k, r) \exp\left[-i \lim_{k\to\infty} \frac{k}{E} \int_r^\infty V_2(\xi)\, d\xi \right]. \quad (5.19)$$

Let us carry out the same study for the Jost function. By multiplying the second component (comp 2) of Eq. (5.5) by $r^\lambda[(2\lambda - 1)!!]^{-1}$, by taking the limit for $r \to 0$, and by using the definition (4.5) {in which we put $a_0 = [(2\lambda - 1)!!]^{-1}$} and the properties of the Bessel spherical functions, the integral expression for the Jost function is

$$f^\lambda(k) = 1 - \lim_{r\to 0} \frac{r^\lambda}{(2\lambda - 1)!!} \times \left[\int_r^\infty \mathfrak{K}(k, r, \xi) f^\lambda(k, \xi)\, d\xi \right]_{\text{comp 2}}; \quad (5.20)$$

that is,

$$f^\lambda(k) = 1 + \int_0^\infty [\varphi^\lambda_0(k, \xi)]^{\mathrm{T}} \mathfrak{V}(\xi) f^\lambda(k, \xi)\, d\xi. \quad (5.20')$$

From this equation it is possible to study the analytic properties of this function. It is proved that $f^\lambda(k)$ is an analytic function in the cut lower semiplane of the complex k-plane ($\mathrm{Im}\, k \le 0$), the points $k = 0$ and $k = -im$ excluded. The proof is not given here, for the reasons stated above, and we refer to Ref. 12 for details. In this demonstration, explicit use of the condition (5.1) is made. To obtain the asymptotic behavior, we take the limit for $k \to \infty$ in Eq. (5.20'). By introduction of the asymptotic behavior of the regular free solution

$$\lim_{k\to\infty} \varphi^\lambda_0(k, r) = k^{-\lambda} \begin{pmatrix} \cos(kr - \frac{1}{2}\lambda\pi) \\ \dfrac{k}{E} \sin(kr - \frac{1}{2}\lambda\pi) \end{pmatrix} \quad (5.21)$$

and by partial integration, we get

$$\lim_{\substack{k\to\infty \\ \mathrm{Im}\, k \le 0}} f^\lambda(k) = \exp\left[-i \lim_{k\to\infty} \frac{k}{E} \int_0^\infty V_2(r)\, dr \right]. \quad (5.22)$$

From (4.6) we deduce the asymptotic behavior of the phase shifts (the so-called Parzen theorem[13]),

$$\lim_{\substack{k\to\infty \\ \mathrm{Im}\, k \le 0}} \eta_\lambda(k) = -\lim_{k\to\infty} \frac{k}{E} \int_0^\infty V_2(r)\, dr, \quad (5.23)$$

and from (4.2), (5.19), and (5.22) we see that the regular solution verifies

$$\lim_{k\to\infty} \varphi^{\lambda}(k, r) = \frac{1}{k^{\lambda}} \begin{pmatrix} \cos\left[kr - \tfrac{1}{2}\lambda\pi - \lim_{k\to\infty} \frac{k}{E} \int_0^r V_2(\xi)\, d\xi\right] \\ \frac{k}{E} \sin\left[kr - \tfrac{1}{2}\lambda\pi - \lim_{k\to\infty} \frac{k}{E} \int_0^r V_2(\xi)\, d\xi\right] \end{pmatrix}. \tag{5.24}$$

It is proved, in a manner similar to that found in Refs. 11 and 12, that if $f^{\lambda,\sigma}(k_n^{\sigma}) = 0$, for $k_n^{\sigma} = -ip_n^{\sigma}$ with $\operatorname{Im} p_n^{\sigma} = 0$, we have a bound state of energy

$$E_n^{\sigma} = \sigma[m^2 - (p_n^{\sigma})^2]^{\frac{1}{2}}.$$

Here $f^{\lambda,\sigma}(k)$ cannot vanish to real k and its zeros are simple (except possibly at $k = 0$; see Ref. 12).

As we see from (5.19) and (5.22)–(5.24), the asymptotic behavior is the same as in the case $V_1(r) \equiv 0$. This can be justified in the following way: If we eliminate a variable in the system (3.1), an equation of the Schrödinger type is obtained [see Appendix B, formula (B1)]. In this equation we can see that the asymptotic behavior for large k is conditioned by the potential $V_2(r)$ and not by $V_1(r)$ (remembering that the potentials are not singular potentials).

6. CONSTRUCTION OF THE JOST FUNCTION IN TERMS OF THE SCATTERING PARAMETERS

We now derive an expression for the Jost function in terms of the scattering and bound-state parameters, which will be used below. Prats and Toll[11] derive an expression of this type for the case when no bound state is present. They use a modification of a theorem of Titchmarsh[14] (useful for this purpose in the nonrelativistic case) that is employed by Corinaldesi[15] for the study of this problem in the Klein–Gordon case; this is the standard procedure, as in the Schrödinger case.

We follow this same procedure in the case when bound states are present. Let us introduce a new function (λ is fixed and is not indicated explicitly)

$$\bar{f}^{\sigma}(k) = f^{\sigma}(k) \prod_n \frac{k - ip_n^{\sigma}}{k + ip_n^{\sigma}}; \tag{6.1}$$

$\bar{f}^{\sigma}(k)$ does not vanish at the energies corresponding to the bound states, because the zeros of the Jost function are simple.

The new "phase shifts," defined by

$$\bar{f}^{\sigma}(k) = |\bar{f}^{\sigma}(k)| \exp[i\bar{\eta}^{\sigma}(k)], \tag{6.2}$$

are related to the old ones by

$$\bar{\eta}^{\sigma}(k) = \eta^{\sigma}(k) - 2\sum_n \arctan \frac{p_n^{\sigma}}{k}, \quad \operatorname{Im} k = 0. \tag{6.3}$$

By writing a dispersion relation for the function $\ln \bar{f}^{+}(k)\bar{f}^{-}(k)$, we get

$$\ln |f^{+}(k)|\, |f^{-}(k)| = -\frac{P}{\pi} \int_{-\infty}^{\infty} \frac{dk'}{k' - k} [\eta^{+}(k') + \eta^{-}(k')] + \sum_{n,\sigma} \ln\left[1 + \left(\frac{p_n^{\sigma}}{k}\right)^2\right]. \tag{6.4}$$

For the function $(k/E) \ln [\bar{f}^{+}(k)/\bar{f}^{-}(k)]$, we get

$$\ln \frac{|f^{+}(k)|}{|f^{-}(k)|} = -\frac{E}{k}\frac{P}{\pi} \int_{-\infty}^{\infty} \frac{dk'}{k' - k} \left(\frac{k'}{E'} [\eta^{+}(k') - \eta^{-}(k')] + 2\mu(\infty)\right) + \sum_{n,\sigma} \sigma \ln \frac{E - E_n^{\sigma}}{E + E_n^{\sigma}} - (N^{+} - N^{-}) \ln \frac{E - m}{E + m}, \tag{6.5}$$

where N^{σ} is the number of bound states in the σ sheet. Note that this function has a cut between $-im$ and $-i\infty$ in the lower semiplane $\operatorname{Im} k \leq 0$, because of the presence of E in the denominator. Formula (6.5) is obtained by integrating the function

$$(k' - k)^{-1}\{(k'/E') \ln [\bar{f}^{+}(k')/\bar{f}^{-}(k')] + 2i\mu(\infty)\}$$

along a path like that in Fig. 1 and then by taking the imaginary part. It is easy to see that by taking the real part we immediately get a result which is analogous to the Levinson theorem for the Dirac case. (This theorem is quoted in Ref. 12.)

From (6.4) and (6.5) we obtain

$$|f^{\sigma}(k)|^2 = \left(\frac{E + \sigma m}{E - \sigma m}\right)^{N^{+} - N^{-}} \times \prod_{n,\sigma'} \left\{\left[1 + \left(\frac{p_n^{\sigma'}}{k}\right)^2\right]\left[\frac{E - \sigma E_n^{\sigma'}}{E + \sigma E_n^{\sigma'}}\right]\right\} \times \exp\left\{-\frac{P}{\pi} \int_{-\infty}^{\infty} \frac{dk'}{k' - k} \left[\eta^{+}(k') + \eta^{-}(k') + \sigma \frac{E}{k} \left(\frac{k'}{E'} [\eta^{+}(k') - \eta^{-}(k')] + 2\mu(\infty)\right)\right]\right\}. \tag{6.6}$$

In the above expression we have introduced the notation

$$\mu(r) = \int_0^r V_2(\xi)\, d\xi. \tag{6.7}$$

We note that (5.23) can be written

$$\lim_{k\to\infty} \eta^{\sigma}(k) = -\sigma\mu(\infty). \tag{5.23'}$$

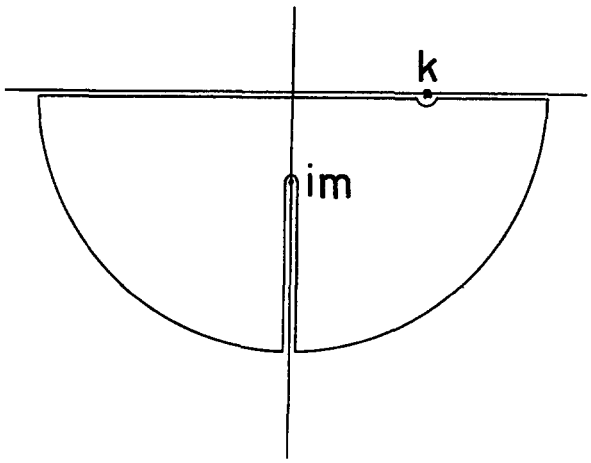


FIG. 1. The integration path of the function $(k'-k)^{-1}\{(k'/E')\times \ln[\bar{f}^{+}(k')/\bar{f}(k')]+2i\mu(\infty)\}$.

7. ASYMPTOTIC EXPANSION METHOD

Let us now study[16] the equations satisfied by the logarithmic derivative of the components of the Jost solutions. We put the system (3.1) in the following form:

$$f_1^{\lambda\prime}(k,r)=(\lambda/r)f_1^{\lambda}(k,r)-\beta(r)f_2^{\lambda}(k,r), \quad (7.1a)$$

$$f_2^{\lambda\prime}(k,r)=-(\lambda/r)f_2^{\lambda}(k,r)+\alpha(r)f_1^{\lambda}(k,r), \quad (7.1b)$$

where

$$\alpha(r)=E+m+V_1(r)-V_2(r), \quad (7.2)$$

$$\beta(r)=E-m-V_1(r)-V_2(r). \quad (7.3)$$

We define

$$y_1(k,r)=\frac{f_1^{\lambda\prime}(k,r)}{f_1^{\lambda}(k,r)}. \quad (7.4)$$

It is easy to see from (7.1) that (7.4) satisfies the equation

$$y_1'(k,r)+y_1^2(k,r)-\frac{\beta'(r)}{\beta(r)}y_1(k,r)+\frac{\beta'(r)}{\beta(r)}\frac{\lambda}{r}-\frac{\lambda(\lambda-1)}{r^2}+\alpha(r)\beta(r)=0. \quad (7.5)$$

In order to apply a method of asymptotic expansion, it is required that the equation have no terms that are divergent at the origin. Therefore, we set $\lambda=1$ and we require that

$$\beta'(0)=-[V_1'(0)+V_2'(0)]=0. \quad (7.6)$$

We note that this last condition is the same we need in order not to have logarithmic terms in the expansion at the origin of the regular solution. [See expression (3.7).]

We also define

$$y_2(k,r)=\frac{f_2^{\lambda\prime}(k,r)}{f_2^{\lambda}(k,r)}; \quad (7.7)$$

this function satisfies the equation

$$y_2'(k,r)+y_2^2(k,r)-\frac{\alpha'(r)}{\alpha(r)}y_2(k,r)-\frac{\alpha'(r)}{\alpha(r)}\frac{\lambda}{r}-\frac{\lambda(\lambda+1)}{r^2}+\alpha(r)\beta(r)=0, \quad (7.8)$$

which does not possess divergent terms at the origin, if $\lambda=-1$ and

$$\alpha'(0)=V_1'(0)-V_2'(0)=0. \quad (7.9)$$

Now we are in a position to solve the following equations: with $\lambda=1$, $V_1'(0)=-V_2'(0)$,

$$y_1'(k,r)+y_1^2(k,r)+\frac{V_1'(r)+V_2'(r)}{E-m-V_1(r)-V_2(r)}y_1(k,r)-\frac{1}{r}\frac{V_1'(r)+V_2'(r)}{E-m-V_1(r)-V_2(r)}+k^2-2EV_2(r)-2mV_1(r)-V_1^2(r)+V_2^2(r)=0; \quad (7.10)$$

with $\lambda=-1$, $V_1'(0)=V_2'(0)$,

$$y_2'(k,r)+y_2^2(k,r)-\frac{V_1'(r)-V_2'(r)}{E+m+V_1(r)-V_2(r)}y_2(k,r)+\frac{1}{r}\frac{V_1'(r)-V_2'(r)}{E+m+V_1(r)-V_2(r)}+k^2-2EV_2(r)-2mV_1(r)-V_1^2(r)+V_2^2(r)=0. \quad (7.11)$$

Note that (7.10) and (7.11) are derivable one from the other by the transformation (2.7). From (2.3) we see that (7.10) corresponds to $j=\frac{1}{2}$ and $l=1$, that is, P wave, and (7.11) to $j=\frac{1}{2}$ and $l=0$, S wave. Therefore, we will treat only this second case, which has an easier interpretation in the nonrelativistic case, while the P-wave results are derivable from this case, using the transformation (2.7').

Let us now define

$$g(k,r)=\frac{\alpha(r)}{\alpha(\infty)}\left(ik+\frac{f_2'(k,r)}{f_2(k,r)}\right), \quad (7.12)$$

where we have put $f_2^{(\lambda=-1)}(k,r)=f_2(k,r)$.

From (7.11), $g(k,r)$ satisfies the equation

$$\alpha(\infty)\alpha(r)g'(k,r)+\alpha^2(\infty)g^2(k,r)-2ik\alpha(\infty)\alpha(r)g(k,r)-2\alpha(\infty)\alpha'(r)g(k,r)=U(k,r), \quad (7.13)$$

where

$$U(k,r)=\alpha^2(r)[2EV_2(r)+2mV_1(r)+V_1^2(r)-V_2^2(r)]-\alpha'(r)\alpha(r)(ik+r^{-1}). \quad (7.14)$$

From the asymptotic behavior of the Jost solution we deduce that

$$\lim_{r\to\infty}g(k,r)=0, \quad (7.15)$$

and from the behavior at the origin of the Jost function, which is known from (4.8), we get

$$\lim_{r\to 0} g(k, r) = \text{const} = g(k). \tag{7.16}$$

Thus far, we have considered only positive energies. Since the Jost solutions are double-valued functions of complex k, the same is true of $g(k, r)$. Thus, instead of Eq. (7.13), we have a system of two equations for the functions $g^{\sigma}(k, r)$. We want to solve this system introducing an asymptotic expansion in k. But, since both equations have terms in k and E, a way to remove E is to write a system of coupled equations in the new variables

$$S(k, r) = \tfrac{1}{2}[g^{+}(k, r) + g^{-}(k, r)], \tag{7.17}$$

$$D(k, r) = (2E)^{-1}[g^{+}(k, r) - g^{-}(k, r)]. \tag{7.18}$$

We note that

$$U(k, r) = P(k, r) + EQ(k, r).$$

In this way the system obtained is

$$\begin{aligned} &a_1S'(k, r) + a_2D'(k, r) + a_3S^2(k, r) + a_4D^2(k, r) \\ &\quad + a_52S(k, r)D(k, r) + a_6S(k, r) + a_7D(k, r) \\ &\qquad = P(k, r), \\ &b_1S'(k, r) + b_2D'(k, r) + b_3S^2(k, r) + b_4D^2(k, r) \\ &\quad + b_52S(k, r)D(k, r) + b_6S(k, r) + b_7D(k, r) \\ &\qquad = Q(k, r), \end{aligned} \tag{7.19a}$$

where the "potentials" are

$$\begin{aligned} P(k, r) = {} &\{4V_2(r)[m + V_1(r) - V_2(r)] \\ &+ 2mV_1(r) + V_1^2(r) - V_2^2(r)\}k^2 \\ &- [V_1'(r) - V_2'(r)][m + V_1(r) - V_2(r)]ik \\ &+ \{4m^2V_2(r) - r^{-1}[V_1(r) - V_2(r)]\} \\ &\times [m + V_1(r) - V_2(r)] \\ &+ [2mV_1(r) + V_1^2(r) - V_2^2(r)] \\ &\times \{m^2 + [m + V_1(r) - V_2(r)]^2\}, \end{aligned} \tag{7.19b}$$

$$\begin{aligned} Q(k, r) = {} &2V_2(r)k^2 - [V_1'(r) - V_2'(r)]ik \\ &+ 2[2mV_1(r) + V_1^2(r) - V_2^2(r)] \\ &\times [m + V_1(r) - V_2(r)] - r^{-1}[V_1'(r) - V_2'(r)] \\ &+ 2V_2(r)\{m^2 + [m + V_1(r) - V_2(r)]^2\} \end{aligned} \tag{7.19c}$$

and the coefficients are

$$\begin{aligned} &b_1 = 2m + V_1(r) - V_2(r), && a_1 = b_2 = k^2 + mb_1, \\ &a_2 = (k^2 + m^2)b_1, && a_3 = b_5 = k^2 + 2m^2, \\ &a_4 = (k^2 + m^2)a_3, && b_3 = 2m, \\ &a_5 = b_4 = (k^2 + m^2)b_3, \\ &a_6 = b_7 = (-2ik)a_1 - [V_1'(r) - V_2'(r)]b_3, \\ &b_6 = (-2ik)b_1 - 2[V_1'(r) - V_2'(r)], \\ &a_7 = (k^2 + m^2)b_6. \end{aligned} \tag{7.19d}$$

By introducing the following asymptotic expansions for large k,

$$S(k, r) = \sum_{n=0}^{N} S_n(r)(-2ik)^{-n-1} + O(k^{-N-2}), \tag{7.20}$$

$$D(k, r) = \sum_{n=0}^{N} D_n(r)(-2ik)^{-n-1} + O(k^{-N-2}), \tag{7.21}$$

in the system (7.19), we obtain the system of coupled recursion relations

$$\begin{aligned} &\tfrac{1}{16}\sum_{j=0}^{n+4} D_j(r)D_{n-j+4}(r) \\ &- \tfrac{1}{4}\{S_{n+3}'(r) + S_{n+4}(r) + [2m + V_1(r) - V_2(r)] \\ &\times [D_{n+3}'(r) + D_{n+4}(r)] - 2[V_1'(r) - V_2'(r)]D_{n+3}(r)\} \\ &- \tfrac{1}{4}\sum_{j=0}^{n+2}[S_j(r)S_{n-j+2}(r) + 3m^2D_j(r)D_{n-j+2}(r) \\ &+ 4mS_j(r)D_{n-j+2}(r)] + m[2m + V_1(r) - V_2(r)] \\ &\times \{S_{n+1}'(r) + S_{n+2}(r) + m[D_{n+1}'(r) + D_{n+2}(r)]\} \\ &- 2m[V_1'(r) - V_2'(r)][S_{n+1}(r) + mD_{n+1}(r)] \\ &+ 2m^2\sum_{j=0}^{n}[S_j(r) + mD_j(r)][S_{n-j}(r) + mD_{n-j}(r)] = 0, \end{aligned} \tag{7.22a}$$

$$\begin{aligned} &- \tfrac{1}{4}[D_{n+3}'(r) + D_{n+4}(r)] \\ &- \tfrac{1}{2}\sum_{j=0}^{n+2}[S_j(r) + mD_j(r)]D_{n-j+2}(r) \\ &+ [2m + V_1(r) - V_2(r)] \\ &\times \{S_{n+1}'(r) + S_{n+2}(r) + m[D_{n+1}'(r) + D_{n+2}(r)]\} \\ &- 2[V_1'(r) - V_2'(r)][S_{n+1}(r) + mD_{n+1}(r)] \\ &+ 2m\sum_{j=0}^{n}[S_j(r) + mD_j(r)][S_{n-j}(r) + mD_{n-j}(r)] = 0, \end{aligned} \tag{7.22b}$$

where $n \geq 0$. The first coefficients are given by

$$D_0(r) = 2V_2(r), \tag{7.23a}$$

$$D_1(r) = -2V_1'(r), \tag{7.23b}$$

$$\begin{aligned} D_2(r) = {} &2V_1''(r) + 4r^{-1}[V_1'(r) - V_2'(r)] \\ &+ 8m[V_1(r)V_2(r) - V_1^2(r) - V_2^2(r)] - 4V_1^3(r), \end{aligned} \tag{7.23c}$$

$$\begin{aligned} D_3(r) = {} &-2V_1'''(r) \\ &+ 4\left(\frac{V_1'(r) - V_2'(r)}{r^2} - \frac{V_1''(r) - V_2''(r)}{r}\right) \\ &+ 8[m + V_1(r)][V_1(r) + V_2(r)] \\ &\times [V_1'(r) + V_2'(r)], \end{aligned} \tag{7.23d}$$

$$S_0(r) = 2mV_1(r) + V_1^2(r) + 2V_1(r)V_2(r) - 2V_2^2(r), \tag{7.24a}$$

$$S_1(r) = -2mV_2'(r) - 2V_1(r)[V_1'(r) + V_2'(r)], \quad (7.24b)$$

$$\begin{aligned} S_2(r) = {} & 2mV_2''(r) + 2V_1(r)[V_1''(r) + V_2''(r)] \\ & + 2V_1''(r)V_2(r) + 6V_1'(r)V_2'(r) - [V_1'(r)]^2 \\ & + 4\left(\frac{V_1'(r) - V_2'(r)}{r}\right)[V_2(r) - m] - V_1^4(r) \\ & - 4V_1^3(r)V_2(r) - 12mV_1^2(r)V_2(r) \\ & - 16m^2V_1(r)V_2(r) + 4m^2[V_1^2(r) + V_2^2(r)], \end{aligned} \quad (7.24c)$$

$$\begin{aligned} S_3(r) = {} & -2mV_2'''(r) - 4\left(\frac{V_1'(r) - V_2'(r)}{r}\right) \\ & \times [3V_2'(r) - V_1'(r)] - 4[m - 2V_2(r)] \\ & \times \left(\frac{V_1'(r) - V_2'(r)}{r^2} - \frac{V_1''(r) - V_2''(r)}{r}\right) \\ & - 2\{V_1(r)[V_1'''(r) + V_2'''(r)] + 2V_1'''(r)V_2(r)\} \\ & - 2[6V_1''(r)V_2'(r) + 4V_1'(r)V_2''(r) - V_1'(r)V_1''(r)] \\ & + 8V_1(r)[V_1(r) + V_2(r)]^2[V_1'(r) + V_2'(r)] \\ & + 8m\{V_2(r)[3V_1(r) + V_2(r)][V_1'(r) + V_2'(r)] \\ & + V_1^2(r)[3V_2'(r) + V_1'(r)]\} \\ & + 8m^2[V_1(r) + V_2(r)][V_1'(r) + V_2'(r)]. \end{aligned} \quad (7.24d)$$

The system (7.22), together with the initial conditions (7.23a), (7.23b), (7.24a), and (7.24b), the asymptotic conditions

$$S_n(\infty) = 0, \quad D_n(\infty) = 0, \quad (7.25)$$

and the remaining conditions (7.24c), (7.24d), (7.25c), and (7.25d), allows us, in principle, to calculate by recursion all the functions $S_n(r)$ and $D_n(r)$. We note that, since we chose $V_1'(0) = V_2'(0)$, then

$$\lim_{r\to 0} \frac{V_1'(r) - V_2'(r)}{r} = V_1''(0) - V_2''(0),$$

$$\lim_{r\to 0} \left(\frac{V_1'(r) - V_2'(r)}{r^2} - \frac{V_1''(r) - V_2''(r)}{r}\right) = -\tfrac{1}{2}[V_1'''(0) - V_2'''(0)];$$

therefore, at the origin, the relationships between the coefficients of (7.20) and (7.21) and the potentials and their derivatives are

$$D_0 = 2V_2(0), \quad (7.26a)$$

$$D_1 = -2V_1'(0), \quad (7.26b)$$

$$\begin{aligned} D_2 = {} & 6V_1''(0) - 4V_2''(0) \\ & + 8m[V_1(0)V_2(0) - V_1^2(0) - V_2^2(0)] - 4V_1^3(0), \end{aligned} \quad (7.26c)$$

$$\begin{aligned} D_3 = {} & -4V_1'''(0) + 2V_2'''(0) \\ & + 8[m + V_1(0)][V_1(0) + V_2(0)][V_1'(0) + V_2'(0)], \end{aligned} \quad (7.26d)$$

$$S_0 = 2V_2(0)[V_1(0) - V_2(0)] + [2m + V_1(0)]V_1(0), \quad (7.27a)$$

$$S_1 = -2mV_2'(0) - 2V_1(0)[V_1'(0) + V_2'(0)], \quad (7.27b)$$

$$\begin{aligned} S_2 = {} & 2[3m + V_1(0) - 2V_2(0)]V_2''(0) \\ & + 2[V_1(0) + 3V_2(0) - 2m]V_1''(0) \\ & + 6V_1'(0)V_2'(0) - [V_1'(0)]^2 \\ & - V_1^4(0) - 4V_1^3(0)V_2(0) - 12mV_1^2(0)V_2(0) \\ & - 16m^2V_1(0)V_2(0) + 4m^2[V_1^2(0) + V_2^2(0)], \end{aligned} \quad (7.27c)$$

$$\begin{aligned} S_3 = {} & 2[2V_2(0) - V_1(0) - 2m]V_2'''(0) \\ & + 2[m - 4V_2(0) - V_1(0)]V_1'''(0) \\ & + 6V_1''(0)[V_1'(0) - 4V_2'(0)] \\ & + 8mV_2(0)[3V_1(0) + V_2(0)][V_1'(0) + V_2'(0)] \\ & + 8mV_1^2(0)[3V_2'(0) + V_1'(0)] \\ & + 8m^2[V_1(0) + V_2(0)][V_1'(0) + V_2'(0)] \\ & + 8V_1(0)[V_1(0) + V_2(0)]^2[V_1'(0) + V_2'(0)]. \end{aligned} \quad (7.27d)$$

Here we have set $S_n(0) = S_n$ and $D_n(0) = D_n$.

The subsequent coefficients are determined by the recursion relations. In principle, one could obtain in this manner all the coefficients S_n and D_n and from these calculate the potentials and all their derivatives at the origin, that is, obtain, in fact, both potentials, since we assumed they are both holomorphic functions of r.

8. DETERMINATION OF THE COEFFICIENTS S_n AND D_n

From the definition of $g(k, r)$ [see (7.12)], from the analyticity properties of the Jost function and the function $f_2(k, r)$, and from the value and derivative at the origin of this function calculated from (4.8), namely,

$$f_2(k, 0) = -f(k)\alpha(0), \quad (8.1a)$$

$$f_2'(k, 0) = -\Delta(k)/2f(-k) + f(k)c_1, \quad (8.1b)$$

we infer that the function $g(k)$ [see (7.16)] is a meromorphic function of k in the cut lower semiplane, Im $k \leq 0$. In the first Riemann sheet, $g^+(k)$ has simple poles at $k = -ip_n^+$ because of the zeros of the function $f^+(k)$. The same occurs for $g^-(k)$ at $k = -ip_n^-$ in the second sheet.

The residues at the poles are

$$\text{Res}\,[g^\sigma(k), k = -ip_n^\sigma] = \frac{-2p_n^\sigma}{f^\sigma(ip_n^\sigma)\dot{f}^\sigma(-ip_n^\sigma)}. \quad (8.2)$$

Here $\dot{f}^\sigma(k)$ indicates the derivative of $f^\sigma(k)$ with respect to k. To obtain this result, we take $f_2^+(-ip_n^+, 0)$ from

(8.1a), while we evaluate $f_2^{+\prime}(-ip_n^+, 0)$, differentiating Eq. (4.2) and setting $k = -ip_n^+$ and $r = 0$.

Let us now define the functions

$$\bar{S}(k) = S(k) + \sum_{n,\sigma} \frac{E_n^\sigma C_n^\sigma}{(E_n^\sigma + \sigma m)} \frac{1}{[k^2 + (p_n^\sigma)^2]}, \quad (8.3)$$

$$\bar{D}(k) = D(k) + \sum_{n,\sigma} \frac{\sigma C_n^\sigma}{(E_n^\sigma + \sigma m)} \frac{1}{[k^2 + (p_n^\sigma)^2]}, \quad (8.4)$$

where we denote by C_n^σ the normalization coefficient of the bound-state wavefunctions

$$C_n^\sigma = \delta_{nj}\left(\int_0^\infty [\varphi(E_n^\sigma, r)]^{\mathrm{T}} \varphi(E_j^\sigma, r)\, dr\right)^{-1} = \frac{-2i(p_n^\sigma)^2(E_n^\sigma + \sigma m)}{E_n^\sigma f^\sigma(ip_n^\sigma) f^{\dot{\sigma}}(-ip_n^\sigma)}. \quad (8.5)$$

The functions $\bar{S}(k)$ and $\bar{D}(k)$ are holomorphic in k in the lower semiplane, $\operatorname{Im} k \leq 0$, since the sums cancel out the possible contribution of the poles in $S(k)$ and $D(k)$ because of the presence of bound states.

By comparing the expansions (7.20) and (7.21) with the similar ones for $\bar{S}(k)$ and $\bar{D}(k)$,

$$\bar{S}(k) = \sum_{n=0}^{N} \bar{S}_n \cdot (-2ik)^{-n-1} + O(k^{-N-2}), \quad (8.6)$$

$$\bar{D}(k) = \sum_{n=0}^{N} \bar{D}_n \cdot (-2ik)^{-n-1} + O(k^{-N-2}), \quad (8.7)$$

we obtain

$$S_{2j} = \bar{S}_{2j}, \quad (8.8)$$

$$S_{2j+1} = \bar{S}_{2j+1} + \sum_{n,\sigma} \frac{2^{2j+2} E_n^\sigma C_n^\sigma (p_n^\sigma)^{2j}}{(E_n^\sigma + \sigma m)}, \quad (8.9)$$

$$D_{2j} = \bar{D}_{2j}, \quad (8.10)$$

$$D_{2j+1} = \bar{D}_{2j+1} + \sum_{n,\sigma} \frac{\sigma 2^{2j+2} C_n^\sigma (p_n^\sigma)^{2j}}{(E_n^\sigma + \sigma m)}. \quad (8.11)$$

For real k, it is possible to express the imaginary parts of $\bar{S}(k)$ and $\bar{D}(k)$ as follows:

$$\operatorname{Im} \bar{S}(k) = -\frac{m[V_1(0) - V_2(0)]}{k} + \tfrac{1}{2}k\left(2 - \frac{1}{|f^+(k)|^2} - \frac{1}{|f^-(k)|^2}\right), \quad \operatorname{Im} k = 0, \quad (8.12)$$

$$\operatorname{Im} \bar{D}(k) = \frac{V_1(0) - V_2(0)}{k} + \frac{k}{2E}\left(\frac{1}{|f^-(k)|^2} - \frac{1}{|f^+(k)|^2}\right), \quad \operatorname{Im} k = 0. \quad (8.13)$$

The real parts can be expressed by the dispersion relations

$$\operatorname{Re} \bar{S}(k) = -\frac{P}{\pi}\int_{-\infty}^{\infty} dk' \frac{k'}{k'^2 - k^2} \operatorname{Im} \bar{S}(k'), \quad \operatorname{Im} k = 0, \quad (8.14)$$

$$\operatorname{Re} \bar{D}(k) = -\frac{P}{\pi}\int_{-\infty}^{\infty} dk' \frac{k'}{k'^2 - k^2} \operatorname{Im} \bar{D}(k'), \quad \operatorname{Im} k = 0. \quad (8.15)$$

By looking at the expansions (8.6) and (8.7), we note that from (8.12) and (8.13) we can get the coefficients $\bar{S}_n$ and $\bar{D}_n$ of even order as generalized moments of the phase shifts [see also (6.6)], while from (8.14) and (8.15) we can obtain the odd coefficients in the form of generalized moments of $|f^\sigma(k)|^2$. Then from (8.8)–(8.11) we can get S_n and D_n.

Note that the even coefficients depend on the phase shifts and the bound-state energies, but not on the bound-state normalization coefficients, while the odd coefficients depend on all the parameters, in close relation with the results of the nonrelativistic case.[5]

We are now ready to calculate from (7.26), (7.27), and (7.22) the potentials and their derivatives at the origin.

9. VALUE OF THE POTENTIALS AND THEIR DERIVATIVES AT THE ORIGIN

In this section, we deduce expressions for the values of the potentials and their first three derivatives at the origin in terms of the scattering phase shifts and the binding energies and normalization constants of the bound states. The results are given by formulas (9.20)–(9.26). In (9.27) we give the conditions on the scattering parameters, implied by the assumed equality at the origin of the first derivatives of the potentials.

To deduce these expressions, we simply make an expansion for large k of the function $|f^\sigma(k)|^2$ and then compare it with expressions (7.26) and (7.27) without deriving general formulas, which are very complicated.

In order to simplify the description of this computation, we consider a case where no bound state is present, and we give at the end the corresponding expressions in the general case with bound states.

When no bound state exists, (6.6) becomes

$$|f^\sigma(k)|^2 = \exp[N_1(k) + \sigma E N_2(k)], \quad (9.1)$$

where

$$N_1(k) = -\frac{P}{\pi}\int_{-\infty}^{\infty} \frac{dk'}{k' - k} \eta^S(k'), \quad (9.2)$$

$$\eta^S(k) = \eta^+(k) + \eta^-(k), \quad (9.3)$$

$$N_2(k) = -\frac{P}{\pi}\int_{-\infty}^{\infty} \frac{dk'}{k' - k} [\eta^D(k') + 2\mu(\infty)], \quad (9.4)$$

$$\eta^D(k) = (k/E)[\eta^+(k) - \eta^-(k)]. \quad (9.5)$$

From (8.13) and (9.1), we have

$$\operatorname{Im} \bar{D}(k) = \frac{V_1(0) - V_2(0)}{k} + \frac{k}{E} \exp[-N_1(k)] \sinh[EN_2(k)]. \quad (9.6)$$

By expanding for large k, we get

$$\text{Im } \bar{D}(k) = [V_1(0) - V_2(0) + d_1]k^{-1} + [d_2 + \tfrac{1}{6}d_1^3 - s_1 d_1]k^{-3} + O(k^{-5}), \quad (9.7)$$

where the coefficients are given by

$$d_n = \frac{-1}{\pi(2n-1)!}\int_{-\infty}^{\infty} dk \mathfrak{D}_k^{2n-1}[k^{-1}\eta^D(k)], \quad (9.8)$$

$$s_n = \frac{-1}{\pi(2n-1)!}\int_{-\infty}^{\infty} dk \mathfrak{D}_k^{2n-1}\eta^S(k). \quad (9.9)$$

In these formulas $n \geq 1$ and $\mathfrak{D}_k^n$ is the operator

$$\mathfrak{D}_k^n = k^n \frac{d^n}{dk^n} k^n. \quad (9.10)$$

The coefficients were obtained by using the mathematical results given in the article by Calogero and Degasperis[5] (see also Buslaev and Faddeev[3]), who have performed similar expressions in the Schrödinger case (see Appendix C).

In a similar way, we obtain

$$\text{Im } \bar{S}(k) = [-mV_1(0) + mV_2(0) - \tfrac{1}{2}d_1^2 + s_1]k^{-1} + [s_2 - \tfrac{1}{2}m^2 d_1^2 - d_1 d_2 - \tfrac{1}{24}d_1^4 + \tfrac{1}{2}s_1 d_1^2 - \tfrac{1}{2}s_1^2]k^{-3} + O(k^{-5}). \quad (9.11)$$

In this manner, from (9.7) and (9.11) we obtain the first two even coefficients, i.e., $\bar{D}_0 = D_0$, $\bar{D}_2 = D_2$, $\bar{S}_0 = S_0$, and $\bar{S}_2 = S_2$ [see (8.8) and (8.10)].

In order to evaluate the odd coefficients, we expand (8.14) and (8.15) and then compare the coefficients obtained in such a way with those of the expansions (8.6) and (8.7); remembering (8.9) and (8.11), we obtain (see Appendix C)

$$S_{2j+1} = \frac{(-)^j 2^{2j+1}}{\pi(2j+1)!}\int_{-\infty}^{\infty} dk \mathfrak{D}_k^{2j+1}\left[k\left(2 - \frac{1}{|f^-(k)|^2} - \frac{1}{|f^+(k)|^2}\right)\right] + \sum_{n,\sigma}\frac{2^{2j+2}E_n^\sigma C_n^\sigma (p_n^\sigma)^{2j}}{(E_n^\sigma + \sigma m)}, \quad (9.12)$$

$$D_{2j+1} = \frac{(-)^j 2^{2j+1}}{\pi(2j+1)!}\int_{-\infty}^{\infty} dk \mathfrak{D}_k^{2j+1}\left[\frac{k}{E}\left(\frac{1}{|f^-(k)|^2} - \frac{1}{|f^+(k)|^2}\right)\right] + \sum_{n,\sigma}\frac{\sigma 2^{2j+2}C_n^\sigma (p_n^\sigma)^{2j}}{(E_n^\sigma + \sigma m)}. \quad (9.13)$$

By relating the coefficients obtained from (9.7) and (9.11) with the expressions (7.26a) and (7.27a), we get the value of the potentials at the origin (no bound states present),

$$V_1(0) = -m + [(d_1 - m)^2 + \tfrac{4}{3}(2md_1 - d_1^2 + s_1)]^{\frac{1}{2}}, \quad (9.14)$$

$$V_2(0) = \tfrac{1}{2}(d_1 - m) + \tfrac{1}{2}[(d_1 - m)^2 + \tfrac{4}{3}(2md_1 - d_1^2 + s_1)]^{\frac{1}{2}}. \quad (9.15)$$

By comparing (9.12) and (9.13) with (7.26b) and (7.27b), we find the value of the first derivative at the origin

$$V_1'(0) = V_2'(0) = -\frac{1}{\pi}\int_{-\infty}^{\infty} dk \mathfrak{D}_k^1\left[\frac{k}{E}\left(\frac{1}{|f^-(k)|^2} - \frac{1}{|f^+(k)|^2}\right)\right] \quad (9.16)$$

and the condition on the scattering parameters, which follows from the requirement that the first derivatives of the two potentials be equal at the origin,

$$\int_{-\infty}^{\infty} dk \mathfrak{D}_k^1\left[k\left(2 - \frac{E + m - 2[(d_1 - m)^2 + \frac{4}{3}(2md_1 - d_1^2 + s_1)]^{\frac{1}{2}}}{E|f^+(k)|^2} - \frac{E - m + 2[(d_1 - m)^2 + \frac{4}{3}(2md_1 - d_1^2 + s_1)]^{\frac{1}{2}}}{E|f^-(k)|^2}\right)\right] = 0. \quad (9.17)$$

The second derivatives are calculated by comparing (7.26c) and (7.27c) with the coefficients of order k^{-3} of the expansions (9.7) and (9.11):

$$V_1''(0) = \tfrac{1}{6}[V_1(0) + m]^{-1}\{4m^2 d_1^2 + 8d_1 d_2 + \tfrac{1}{3}d_1^4 - 4d_1^2 s_1 - 8s_2 + 4s_1^2 + \tfrac{1}{2}[V_1(0) - 2V_2(0) + 3m][8d_1 s_1 - \tfrac{4}{3}d_1^2 - 8d_2] - 5[V_1'(0)]^2 + 12mV_1(0)V_2^2(0) + 10mV_1^3(0) - 8mV_2^3(0) + 4m^2[V_1(0)V_2(0) + 2V_1^2(0) + 2V_2^2(0)] + 3V_1^4(0)\}, \quad (9.18)$$

$$V_2''(0) = \tfrac{3}{10}[V_1(0) + m]^{-1}\{4m^2 d_1^2 + 8d_1 d_2 + \tfrac{1}{3}d_1^4 - 4d_1^2 s_1 - 8s_2 + 4s_1^2 - \tfrac{1}{3}[V_1(0) + 3V_2(0) - 2m][8d_1 s_1 - \tfrac{4}{3}d_1^2 - 8d_2] - 5[V_1'(0)]^2 + \tfrac{16}{3}mV_1(0)V_2^2(0) - 8mV_2^3(0) + \tfrac{20}{3}mV_1^2(0)V_2(0) + \tfrac{4}{3}m^2[8V_1(0)V_2(0) + V_1^2(0) + V_2^2(0)] - \tfrac{1}{3}V_1^4(0)\}. \quad (9.19)$$

We find the third derivatives by relating the expressions (9.12) and (9.13) to (7.26d) and (7.27d). Their values are given later [see (9.25) and (9.26)], in the general case with bound states present.

In order to compute the values of higher-order derivatives at the origin, it is necessary to continue this procedure. That is, one should compare higher-order coefficients from the expansions (9.7) and (9.11) and the coefficients given by (9.12) and (9.13) with the values at the origin of the coefficients found with the system of recursion relations (7.22).

If, instead of the expansion of $|f^{\sigma}(k)|^2$ given by the formula (9.1), we expand the expression given by (6.6), we obtain the values at the origin of the potentials and their derivatives when bound states are present. We give here the results of these computations:

$$V_1(0) = -m + [m^2 - \tfrac{4}{3}(\tfrac{1}{2}d_1 - c_1)(\tfrac{1}{2}d_1 - c_1 - m) + \tfrac{4}{3}(s_1 + h_1)]^{\frac{1}{2}}, \tag{9.20}$$

$$V_2(0) = \tfrac{1}{2}\{d_1 - 2c_1 - m + [m^2 - \tfrac{4}{3}(\tfrac{1}{2}d_1 - c_1)(\tfrac{1}{2}d_1 - c_1 - m) + \tfrac{4}{3}(s_1 + h_1)]^{\frac{1}{2}}\}, \tag{9.21}$$

$$V_1'(0) = V_2'(0) = -\frac{1}{\pi}\int_{-\infty}^{\infty} dk \mathfrak{D}_k^1\left[\frac{k}{E}\left(\frac{1}{|f^-(k)|^2} - \frac{1}{|f^+(k)|^2}\right)\right] - \sum_{n,\sigma}\frac{2\sigma C_n^{\sigma}}{(E_n^{\sigma} + \sigma m)}, \tag{9.22}$$

$$\begin{aligned} V_1''(0) = {} & \tfrac{1}{5}[V_1(0) + m]^{-1}\{S_2 + \tfrac{1}{2}[V_1(0) - 2V_2(0) + 3m]D_2 - 5[V_1'(0)]^2 - 12mV_1(0)V_2^2(0) + 10mV_1^3(0) \\ & - 8mV_2^3(0) + 4m^2[V_1(0)V_2(0) + 2V_1^2(0) + 2V_2^2(0)] + 3V_1^4(0)\}, \end{aligned} \tag{9.23}$$

$$\begin{aligned} V_2''(0) = {} & \tfrac{3}{10}[V_1(0) + m]^{-1}\{S_2 - \tfrac{1}{3}[V_1(0) + 3V_2(0) - 2m]D_2 - 5[V_1'(0)]^2 + \tfrac{16}{3}mV_1(0)V_2^2(0) - 8mV_2^3(0) \\ & + \tfrac{20}{3}mV_1^2(0)V_2(0) + \tfrac{4}{3}m^2[8V_1(0)V_2(0) + V_1^2(0) + V_2^2(0)] - \tfrac{1}{3}V_1^4(0)\}, \end{aligned} \tag{9.24}$$

where

$$D_2 = 8(d_1 - 2c_1)(s_1 + h_1) - \tfrac{4}{3}d_1^3 - 8d_2 + 8(2c_3 - 2m^2c_1 - 2c_1^2d_1 + c_1d_1^2), \tag{9.24'}$$

$$\begin{aligned} S_2 = {} & 4m^2d_1^2 + 8d_1d_2 + \tfrac{1}{3}d_1^4 - 4(d_1 - 2c_1)^2(s_1 + h_1) - 8(s_2 - h_2) + 4(s_1 + h_1)^2 \\ & - 4h_1^2 + 16(c_4 - m^2c_1^2 - c_3d_1 - c_1d_2 + \tfrac{1}{2}c_1^2d_1^2 - \tfrac{1}{6}c_1d_1^3), \end{aligned} \tag{9.24''}$$

$$\begin{aligned} V_1'''(0) = {} & \tfrac{1}{6}[V_1(0) + m]^{-1}(-S_3 + [2V_2(0) - V_1(0) - 2m]D_3 \\ & + 2V_1'(0)\{-9V_1''(0) + 8V_1(0)[V_1(0)V_2(0) + 5mV_1(0) + 4mV_2(0) + 2V_1^2(0)] \\ & - 8V_2^2(0)[m + V_1(0)] + 24m^2[V_1(0) + V_2(0)]\}), \end{aligned} \tag{9.25}$$

$$\begin{aligned} V_2'''(0) = {} & \tfrac{1}{3}[V_1(0) + m]^{-1}(-S_3 + \tfrac{1}{2}[V_1(0) + 4V_2(0) - m]D_3 \\ & + 2V_1'(0)\{-9V_1''(0) + 4V_1(0)[-V_1(0)V_2(0) + 4mV_1(0) + 2mV_2(0) + V_1^2(0)] \\ & - 8V_2^2(0)[m + V_1(0)] + 12m^2[V_1(0) + V_2(0)]\}), \end{aligned} \tag{9.26}$$

where S_3 and D_3 are given by (9.12) and (9.13).

The condition $V_1'(0) = V_2'(0)$ gives the relationship

$$\begin{aligned} & \frac{1}{\pi}\int_{-\infty}^{\infty} dk \mathfrak{D}_k^1\left[k\left(2 - \frac{E - m - 2V_1(0)}{E|f^+(k)|^2} \right.\right. \\ & \qquad \left.\left. - \frac{E + m + 2V_1(0)}{E|f^-(k)|^2}\right)\right] \\ & \quad + \sum_{n,\sigma}\frac{2\{E_n^{\sigma} - \sigma[m + 2V_1(0)]\}C_n^{\sigma}}{(E_n^{\sigma} + \sigma m)} = 0. \end{aligned} \tag{9.27}$$

In all these formulas the coefficients s_n and d_n are given by (9.8), (9.9), and (9.10), while the bound-state coefficients are the following:

$$c_1 = \sum_{n,\sigma}(E_n^{\sigma} - \sigma m), \tag{9.28a}$$

$$c_3 = \tfrac{2}{3}c_1^3 + \tfrac{1}{3}\sum_{n,\sigma}[(E_n^{\sigma})^3 - \sigma m^3], \tag{9.28b}$$

$$c_4 = \tfrac{1}{3}c_1^4 + \tfrac{2}{3}c_1\sum_{n,\sigma}[(E_n^{\sigma})^3 - \sigma m^3], \tag{9.28c}$$

$$h_1 = \sum_{n,\sigma}(p_n^{\sigma})^2, \tag{9.29a}$$

$$h_2 = \sum_{n,\sigma}(p_n^{\sigma})^4 + \sum_{n<m,\sigma}(p_n^{\sigma}p_m^{\sigma})^2 + \sum_{n,m}(p_n^{+}p_m^{-})^2. \tag{9.29b}$$

All the above results are in the case $\lambda = -1$, i.e., S wave. To find the corresponding expressions for P wave, we simply change the sign of m and $V_1^{(n)}(0)$, $n = 0, 1, 2, \cdots$ [see (2.7') and remember that $g(k, r)$ is the ratio between a component and its derivative], and use the corresponding P-wave data.

10. RELATIONSHIPS BETWEEN PARTICLE AND ANTIPARTICLE SCATTERING AND BOUND-STATE DATA

In this section, we particularize the results of the last section to the usual Dirac case, that is, $V_2(r) = V(r)$ and $V_1(r) \equiv 0$, providing expressions for the values of the potential and its first three derivatives at

the origin and also giving conditions on the scattering parameters.

As previously noted, these expressions constitute a first step in finding relationships between the scattering and bound-state data of the particle and the antiparticle. These conditions, however, have a restricted validity, because they are deduced in the particular case of a potential whose first derivative vanishes at the origin. Later, we hope to find more general expressions in order to solve the inverse scattering problem for particles that satisfy the Dirac equation.

From (9.21), we get the value of the potential

$$V(0) = \tfrac{1}{2}d_1 - c_1 = -\frac{1}{2\pi}\int_{-\infty}^{\infty} dk \mathfrak{D}_k^1[k^{-1}\eta^D(k)] - \sum_{n,\sigma}(E_n^\sigma - \sigma m) \tag{10.1}$$

and, from (9.20), the first condition, linking scattering and bound-state data,

$$s_1 + h_1 = (\tfrac{1}{2}d_1 - c_1)(\tfrac{1}{2}d_1 - c_1 - m) = V(0)[V(0) - m]. \tag{10.2}$$

The requirement that the first derivative of the potential vanish at the origin gives the two conditions

$$\frac{1}{\pi}\int_{-\infty}^{\infty} dk \mathfrak{D}_k^1\left[\frac{k}{E}\left(\frac{1}{|f^-(k)|^2} - \frac{1}{|f^+(k)|^2}\right)\right] + \sum_{n,\sigma}\frac{\sigma C_n^\sigma}{(E_n^\sigma + \sigma m)} = 0 \tag{10.3}$$

and

$$\frac{1}{\pi}\int_{-\infty}^{\infty} dk \mathfrak{D}_k^1\left[k\left(2 - \frac{1}{|f^-(k)|^2} - \frac{1}{|f^+(k)|^2}\right)\right] + \sum_{n,\sigma}\frac{E_n^\sigma C_n^\sigma}{(E_n^\sigma + \sigma m)} = 0, \tag{10.4}$$

which are obtained from (9.22) and (9.27).

For the second derivative, from (9.23) and (9.24), imposing (10.2), we obtain

$$V''(0) = -\tfrac{1}{6}d_1^3 + \tfrac{1}{2}md_1^2 + 2d_2 + c_1d_1^2 - 2mc_1d_1 - 2c_1^2d_1 + 2mc_1^2 + 4m^2c_1 + 4c_1^3 - 4c_3 \tag{10.5}$$

and the condition

$$s_2 - h_2 = -\tfrac{3}{2}V^4(0) + 3mV^3(0) + \tfrac{1}{2}m^2V^2(0) + \tfrac{3}{2}[2V(0) - m]d_2 + 6m^2c_1V(0) - 3[2V(0) - m](c_3 - \tfrac{2}{3}c_1^3) - \tfrac{1}{2}h_1^2 - 3m^3c_1. \tag{10.6}$$

From (9.25) and (9.26), we find the third derivative

$$V'''(0) = \tfrac{1}{2}D_3 = -\frac{2}{3\pi}\int_{-\infty}^{\infty} dk \mathfrak{D}_k^3\left[\frac{k}{E}\left(\frac{1}{|f^-(k)|^2} - \frac{1}{|f^+(k)|^2}\right)\right] + \sum_{n,\sigma}\frac{8\sigma C_n^\sigma (p_n^\sigma)^2}{(E_n^\sigma + \sigma m)} \tag{10.7}$$

and the condition

$$S_3 = (d_1 - 2c_1 - 2m)D_3 = 2[V(0) - m]D_3. \tag{10.8}$$

In all the above formulas, the coefficients are given by (9.8)–(9.10), (9.12), (9.13), (9.28), and (9.29).

The conditions found may be used to derive formulas for the energy and normalization constants of the bound states in terms of the scattering phase shifts. For instance, in the particular case when there exists only one bound state for the particle (or for the antiparticle), we get, from (10.2),

$$N^+ = 1, \quad N^- = 0,$$

$$E_1^+(S) = \tfrac{1}{4}(d_1 + m) + \{[\tfrac{1}{4}(d_1 + m)]^2 - \tfrac{1}{8}d_1^2 - \tfrac{1}{4}md_1 + \tfrac{1}{2}m^2 + \tfrac{1}{2}s_1\}^{\frac{1}{2}} \tag{10.9}$$

and, from (10.3), with (9.13) and (8.11),

$$C_1^+(S) = -\tfrac{1}{2}[E_1^+(S) + m]\bar{D}_1(S). \tag{10.10}$$

It should be noted that these results have been obtained using only the conditions $V_1(0) = 0$ and $V_1'(0) = 0$, without explicitly using the condition $V_2'(0) = V'(0) = 0$. It is therefore reasonable to conjecture that these relationships hold for more general potentials, without any restriction on the value at the origin of their first derivative.

The expressions (10.9) and (10.10) give us the energy level and the normalization constant in the case of one bound state for the particle in S wave. The corresponding results for P wave are

$$E_1^+(P) = \tfrac{1}{4}(d_1 + 3m) + \{[\tfrac{1}{4}(d_1 + 3m)]^2 - \tfrac{1}{8}d_1^2 - \tfrac{3}{4}md_1 - \tfrac{1}{2}m^2 + \tfrac{1}{2}s_1\}^{\frac{1}{2}} \tag{10.11}$$

and

$$C_1^+(P) = -\tfrac{1}{2}[E_1^+(P) + m]\bar{D}_1(P). \tag{10.12}$$

If there exists only a bound state for the antiparticle in S wave, we get

$$N^+ = 0, \quad N^- = 1,$$

$$E_1^-(S) = \tfrac{1}{4}(d_1 - 3m) - \{[\tfrac{1}{4}(d_1 - 3m)]^2 - \tfrac{1}{8}d_1^2 + \tfrac{3}{4}md_1 - \tfrac{1}{2}m^2 + \tfrac{1}{2}s_1\}^{\frac{1}{2}}, \tag{10.13}$$

and

$$C_1^-(S) = \tfrac{1}{2}[E_1^-(S) - m]\bar{D}_1(S) \tag{10.14}$$

and, correspondingly, in P wave we find

$$E_1^-(P) = \tfrac{1}{4}(d_1 - m) - \{[\tfrac{1}{4}(d_1 - m)]^2 - \tfrac{1}{8}d_1^2 + \tfrac{1}{4}md_1 + \tfrac{1}{2}m^2 + \tfrac{1}{2}s_1\}^{\frac{1}{2}} \tag{10.15}$$

and

$$C_1^-(P) = \tfrac{1}{2}[E_1^-(P) - m]\bar{D}_1(P). \tag{10.16}$$

ACKNOWLEDGMENTS

The author wishes to thank Professor F. Calogero who guided the general lines of this work, as well as

Dr. A. Degasperis for the illuminating conversations. Also, he wants to acknowledge the hospitality of the Istituto di Fisica dell'Università di Roma and the financial support of the CNICT of Argentina.

APPENDIX A: THE NONRELATIVISTIC LIMIT

Taking the nonrelativistic limit of our results (for the usual Dirac case), we must recover the results of the Schrödinger case. Therefore, the nonrelativistic limit of our results constitutes an excellent test of their validity in the S-wave case; moreover, in the P-wave case it yields expressions, valid in the Schrödinger case, for the second and third derivatives of the potential at the origin in terms of the P-wave scattering and bound-state data, which are new results.

To avoid confusion, we indicate in all functions the explicit dependence on λ and denote by a subscript NR, appended to a function, that its nonrelativistic limit has been performed.

Let us now study the equations that the components of the Jost solutions satisfy. Introducing the auxiliary function

$$F_2^+(k, r, +1) = [\alpha(r)\beta(\infty)]^{-\frac{1}{2}} f_2^+(k, r, +1), \quad \text{i.e.,} \quad \lambda = +1, \tag{A1}$$

we see that it satisfies the (Schrödinger-type) equation

$$F_2^{+\prime\prime}(k, r, +1) + \left(k^2 - \frac{2}{r^2}\right) F_2^+(k, r, +1) = \left[2EV(r) - V^2(r) + \frac{1}{r}\frac{\alpha'(r)}{\alpha(r)} + \frac{3}{4}\left(\frac{\alpha'(r)}{\alpha(r)}\right)^2 - \frac{1}{2}\frac{\alpha''(r)}{\alpha(r)}\right] F_2^+(k, r, +1) \tag{A2}$$

and that

$$\lim_{r\to\infty} F_2^+(k, r, +1) = ie^{-ikr}. \tag{A3}$$

From the definition (4.5) of the Jost function and from (4.8) and (3.3), with $V_1(r) \equiv 0$ and $V_2'(0) = V'(0) = 0$, one sees that

$$f^+(k, +1) = \lim_{r\to 0} [r f_2^+(k, r, +1)] = \frac{f_1^+(k, 0, +1)}{\beta(0)} = |f^+(k, +1)| \exp [i\eta^+(k, +1)]. \tag{A4}$$

Thus, in the nonrelativistic limit,

$$F_{2_{\mathrm{NR}}}^+(k, r, +1) = k^{-1} f_{2_{\mathrm{NR}}}^+(k, r, +1), \tag{A5}$$

$$F_{2_{\mathrm{NR}}}^{+\prime\prime}(k, r, +1) + (k^2 - 2r^{-2}) F_{2_{\mathrm{NR}}}^+(k, r, +1) = 2mV(r) F_{2_{\mathrm{NR}}}^+(k, r, +1), \tag{A6}$$

$$\lim_{r\to 0} F_{2_{\mathrm{NR}}}^+(k, r, +1) = ie^{-ikr}, \tag{A7}$$

and

$$f_{\mathrm{NR}}^+(k, +1) = \lim_{r\to 0} [r f_{2_{\mathrm{NR}}}^+(k, r, +1)] = \lim_{r\to 0} [krF_{2_{\mathrm{NR}}}^+(k, r, +1)]. \tag{A8}$$

Therefore, $f_{\mathrm{NR}}^+(k, +1)$ is the Jost function in the Schrödinger case for $l = 1$, and $F_{2_{\mathrm{NR}}}^+(k, r, +1)$ is the Jost solution for this case. Correspondingly, one gets for the phase shift

$$\eta_{\mathrm{NR}}^+(k, +1) = \delta(k, V, l = 1), \tag{A9}$$

where $\delta(k, V, l)$ is the Schrödinger phase shift for linear momentum k, angular momentum l, and potential $+V(r)$.

Repeating the argument for the other components of the Dirac Jost solutions, for $\lambda = \pm 1$ and for the two signs of the energy, one gets

$$\eta_{\mathrm{NR}}^-(k, +1) = \delta(k, -V, l = 0), \tag{A10}$$

$$\eta_{\mathrm{NR}}^+(k, -1) = \delta(k, V, l = 0), \tag{A11}$$

$$\eta_{\mathrm{NR}}^-(k, -1) = \delta(k, -V, l = 1), \tag{A12}$$

that is, the Dirac phase shifts,for the particle and for the antiparticle for "angular momentum" zero and one, goes into the phase shifts for the Schrödinger problem with potentials $V(r)$ and $-V(r)$ and $l = 0$ and 1.

Let us now consider the expression of the Jost function in terms of the phase shifts and bound-state energies [see (6.6)]. We observe that in the nonrelativistic limit

$$\frac{E - E_n^+}{E + E_n^+}\frac{E + m}{E - m} = 1 + \left(\frac{p_n^+}{k}\right)^2, \tag{A13}$$

$$\frac{E - E_n^-}{E + E_n^-}\frac{E - m}{E + m} = \left[1 + \left(\frac{p_n^-}{k}\right)^2\right]^{-1}, \tag{A14}$$

so that

$$|f_{\mathrm{NR}}^+(k, +1)| = \prod_{n=1}^{N}\left(1 - \frac{E_n(l = 1)}{k^2}\right) \times \exp\left(-\frac{P}{\pi}\int_{-\infty}^{\infty} dk' \frac{k'}{k'^2 - k^2}\delta(k', V, l = 1)\right) = |f(k, V, l = 1)|, \tag{A15}$$

a formula that reproduces the well-known formula for the Schrödinger case [see, e.g., Eq. (2.29) of Ref. 5]. Here $E_n\ (l = 1) = -[p_n^+(\lambda = +1)]^2$ indicates the energy of the nth bound state produced by the potential $V(r)$ for angular-momentum one.

In the same way, we see that

$$|f^{+}_{\mathrm{NR}}(k,-1)| = |f\,(k, V, l=0)|, \tag{A16}$$

$$|f^{-}_{\mathrm{NR}}(k,+1)| = |f\,(k,-V, l=0)|, \tag{A17}$$

$$|f^{-}_{\mathrm{NR}}(k,-1)| = |f\,(k,-V, l=1)|. \tag{A18}$$

In these equations, $f(k, V, l)$ is the Jost function of the Schrödinger problem with linear momentum k, angular momentum l, and potential $+V(r)$.

For the bound-state normalization constant [see (8.5)], the nonrelativistic limit gives

$$C_n(l) = \frac{-4i[p_n(l)]^2}{f[ip_n(l)]f'[-ip_n(l)]}, \tag{A19}$$

this being also a well-known result for the Schrödinger case [see Eq. (A10) of Ref. 5].

From (7.12) with $V_1(r) \equiv 0$ and from (7.13) and (7.14), we also see that in the nonrelativistic limit

$$g^{+}_{\mathrm{NR}}(k, r, -1) = g\,(k, r, V, l=0), \tag{A20}$$

where $g(k, r, V, l)$ is the function of Eq. (4.1) of Ref. 5. We note also that

$$g^{+}_{\mathrm{NR}}(k, r, -1) = S_{\mathrm{NR}}(k, r, -1) + mD_{\mathrm{NR}}(k, r, -1). \tag{A21}$$

From these equations it is easily seen that the nonrelativistic limit of (7.22) reproduces the recursion relations of the Schrödinger case [see Eq. (4.6) of Ref. 5].

In the same manner, we obtain

$$\begin{aligned} g^{-}_{\mathrm{NR}}(k, r, +1) &= S_{\mathrm{NR}}(k, r, +1) - mD_{\mathrm{NR}}(k, r, +1) \\ &= g\,(k, r, -V, l=0). \end{aligned} \tag{A22}$$

Therefore, the nonrelativistic limit of the Dirac problem for the particle, with $\lambda = -1$ ("S wave"), and for the antiparticle, with $\lambda = +1$ ("P wave"), corresponds to two Schrödinger problems with potentials $V(r)$ and $-V(r)$, respectively, both in S wave.

This correspondence may be verified by performing the nonrelativistic limit of the expressions for the values of the potential and its derivatives at the origin.

From (10.2) we see that, in the nonrelativistic limit,

$$(s_1)_{\mathrm{NR}} + (h_1)_{\mathrm{NR}} = -m[\tfrac{1}{2}(d_1)_{\mathrm{NR}} - (c_1)_{\mathrm{NR}}] \tag{A23}$$

and, by substitution in (10.1),

$$V(0) = (d_1)_{\mathrm{NR}} - 2(c_1)_{\mathrm{NR}} + m^{-1}[(s_1)_{\mathrm{NR}} + (h_1)_{\mathrm{NR}}]. \tag{A24}$$

From (9.28a), (9.29a), (9.8), and (9.9), we obtain

$$(c_1)_{\mathrm{NR}} = -\frac{1}{2m}\left(\sum_{n=1}^{N^+}(p_n^+)^2 - \sum_{n=1}^{N^-}(p_n^-)^2\right),$$

$$(h_1)_{\mathrm{NR}} = \sum_{n=1}^{N^+}(p_n^+)^2 + \sum_{n=1}^{N^-}(p_n^-)^2,$$

$$(d_1)_{\mathrm{NR}} + \frac{1}{m}(s_1)_{\mathrm{NR}} = -\frac{2}{\pi m}\int_{-\infty}^{\infty} dk\mathfrak{D}_k^1\delta\,(k, V, l=0),$$

and from these we get

$$2mV(0) = -\frac{8}{\pi}\int_0^{\infty} dk\mathfrak{D}_k^1\delta\,(k, V, l=0) - 4\sum_{n=1}^{N}E_n(l=0). \tag{A25}$$

This is the well-known result first obtained by Newton[6] and independently by Faddeev.[8]

The first derivative of the potential is, by assumption, zero at the origin in our case, as we stated in Sec. 10. Nevertheless, it is interesting to note that by combining the nonrelativistic limit of conditions (10.3) and (10.4) we obtain

$$0 = \frac{4}{\pi}\int_{-\infty}^{\infty} dk\mathfrak{D}_k^1\left[k\left(1 - \frac{1}{|f\,(k, V, l=0)|^2}\right)\right] + 4\sum_{n=1}^{N}C_n\,(l=0), \tag{A26}$$

whose right-hand side reproduces the expression for $V'(0)$ in the Schrödinger case [see Eq. (6.5) of Ref. 5].

For the second derivative, one can see that, from the nonrelativistic limit of (10.5),

$$V''(0) = 2mV^2(0) + 2(d_2)_{\mathrm{NR}} + 4m^2(c_1)_{\mathrm{NR}} - 4(c_3)_{\mathrm{NR}} + \tfrac{8}{3}(c_1^3)_{\mathrm{NR}} \tag{A27}$$

and, from (10.6),

$$(s_2)_{\mathrm{NR}} = (h_2)_{\mathrm{NR}} + \tfrac{1}{2}m^2V^2(0) - \tfrac{3}{2}m(d_2)_{\mathrm{NR}} + 3m(c_3)_{\mathrm{NR}} - 2m(c_1^3)_{\mathrm{NR}} - \tfrac{1}{2}(h_1^2)_{\mathrm{NR}} - 3m^3(c_1)_{\mathrm{NR}}. \tag{A28}$$

By combining these expressions, one gets

$$2mV''(0) = 8m^2V^2(0) + \frac{2^4}{\pi 3!}\int_{-\infty}^{\infty} dk\mathfrak{D}_k^3\,\delta\,(k, V, l=0) + 2^3\sum_{n=1}^{N}[E_n\,(l=0)]^2. \tag{A29}$$

This coincides again with the formula for the Schrödinger case [see Eq. (6.6) of Ref. 5].

The nonrelativistic limit of the third derivative is easily found. Note that by combining (10.7) and (10.8) we can write

$$2mV'''(0) = -m(D_3)_{\mathrm{NR}} - (S_3)_{\mathrm{NR}}, \tag{A30}$$

from which we obtain

$$2mV'''(0) = \frac{2^4}{\pi 3!}\int_{-\infty}^{\infty} dk\mathfrak{D}_k^3\{k[1-|f(k,V,l=0)|^{-2}]\} + 2^4\sum_{n=1}^{N} C_n(l=0)\,E_n(l=0),$$

$$= -\frac{2^4}{\pi 5!}\int_{-\infty}^{\infty} dk\mathfrak{D}_k^5[k^{-1}|f(k,V,l=0)|^{-2}] + 2^4\sum_{n=1}^{N} C_n(l=0)\,E_n(l=0). \quad \text{(A31)}$$

This equation coincides with the formula obtained in the Schrödinger case [see Eq. (6.7) of Ref. 5], provided that $V'(0)=0$. (To obtain the last expression, we have performed two partial integrations.)

Applying the transformation (2.7′) to the expressions (10.1) and (10.2), that is, simply changing the sign of m since $V_1(r)\equiv 0$, and introducing the scattering and bound-state parameters for $\lambda=1$, one obtains the results for the Dirac "P wave," $\lambda=+1$, and in the nonrelativistic limit the results become

$$2mV(0) = -\frac{8}{3\pi}\int_0^{\infty} dk\mathfrak{D}_k^1\delta(k,V,l=1) - \tfrac{4}{3}\sum_{n=1}^{N} E_n(l=1). \quad \text{(A32)}$$

This is the formula for the Schrödinger case for $l=1$. [See Newton, Ref. 7, formula (8.14′).]

For the second and third derivative, one gets

$$2mV''(0) = -\tfrac{2^4}{5}m^2V^2(0) - \frac{2^5}{3!\,5\pi}\int_0^{\infty} dk\mathfrak{D}_k^3\delta(k,V,l=1) - \tfrac{8}{5}\sum_{n=1}^{N}[E_n(l=1)]^2, \quad \text{(A33)}$$

$$2mV'''(0) = -\frac{2^5}{3!\,3\pi}\int_0^{\infty} dk\mathfrak{D}_k^3\{k[1-|f(k,V,l=1)|^{-2}]\} - \frac{2^4}{3}\sum_{n=1}^{N} C_n(l=1)\,E_n(l=1). \quad \text{(A34)}$$

The last two expressions constitute new results. These formulas could not be obtained with the method of asymptotic expansions[5] because of the singularity of the centrifugal potential. The results obtained here indicate, however, that this difficulty can be overcome, provided that the potential satisfies the additional condition $V'(0)=0$. The technique of doing this is suggested by the procedure which we have followed in obtaining these results here; it proceeds essentially through the transformation of the P-wave Schrödinger equation into an S-wave Schrödinger equation with an "equivalent" potential, which is complicated and energy dependent, but finite [if $V'(0)=0$]. A more detailed discussion of this method and its application in the Schrödinger case in obtaining results for higher partial waves will be treated in a separate paper.

Finally, we note that from (A23) and (A24) one can also obtain (A32) for the potential $-V(r)$, while by repeating the procedure to obtain (A32) one can obtain (A25), again with $-V(r)$ in place of $V(r)$. This is a confirmation of the fact that the nonrelativistic limit of the Dirac problems with $\lambda=\pm 1$, and $\pm E$ goes into the Schrödinger problems with $l=0$, 1 and potentials $\pm V(r)$.

APPENDIX B

Here we put forward some considerations on the presence of logarithmic terms in the expansion at the origin of the solutions of Eq. (3.1).

From the system of equations (3.1), by elimination of one variable, we get

$$v_\lambda''(k,r) + \left(k^2 - \frac{\lambda(\lambda-1)}{r^2}\right)v_\lambda(k,r) = \left[2EV_2(r) + 2mV_1(r) + V_1^2(r) - V_2^2(r) + \frac{\lambda}{r}\frac{V_1'(r)+V_2'(r)}{E-m-V_1(r)-V_2(r)} + \frac{3}{4}\left(\frac{V_1'(r)+V_2'(r)}{E-m-V_1(r)-V_2(r)}\right)^2 + \frac{1}{2}\frac{V_1''(r)+V_2''(r)}{E-m-V_1(r)-V_2(r)}\right]v_\lambda(k,r), \quad \text{(B1)}$$

where

$$v_\lambda(k,r) = \{[E-m-V_1(r)-V_2(r)][E-m]\}^{-\frac{1}{2}} \times u_1^\lambda(k,r).$$

Equation (B1) can be interpreted as a Schrödinger equation with an energy-dependent potential, which is singular at the origin (since $\lambda\neq 0$). Therefore, the behavior of the solution at the origin will be (for the case $\lambda=1$) of the type

$$\lim_{r\to 0} v_1(k,r) = \text{const} + A\ln Br, \quad \text{(B2)}$$

where

$$A \propto V_1'(0) + V_2'(0).$$

[For the case $\lambda=-1$, it is proved that $A\propto V_1'(0) - V_2'(0)$; see (2.6)′ or (2.7).]

Note that, if A vanishes, the singularity in Eq. (B1) disappears. In the Schrödinger case, one meets logarithmic terms if there are singularities ($l\neq 0$ or no regular potentials) and these terms impede the use of a method of asymptotic expansion such as that used in

this paper. Therefore, in order to employ a similar method, one must choose the potentials in such a way that they exclude the presence of logarithmic terms.

APPENDIX C

Calogero and Degasperis[5] introduced the following asymptotic expansion for the scattering phase shift (numbers on the left denote the equations of their paper):

$$\text{(2.41)} \qquad \eta(k) = \sum_{n=0}^{N} a_n k^{-2n-1} + O(k^{-2N-3}).$$

They have the expression

$$\text{(2.50)} \qquad N(k) = -\frac{P}{\pi}\int_{-\infty}^{\infty} dk' \frac{k'}{k'^2 - k^2}\eta(k')$$

and prove that it can be developed in the asymptotic expansion

$$\text{(3.18)} \qquad N(k) = \sum_{n=1}^{N} d_n k^{-2n} + O(k^{-2N-2}),$$

in which the coefficients are given by

$$\text{(3.17)} \qquad d_n = -\frac{1}{\pi(2n-1)!}\int_{-\infty}^{\infty} dk \mathfrak{D}_k^{2n-1}\eta(k).$$

In our case, we are in a similar situation. Let us introduce the following asymptotic expansions:

$$\eta^+(k) = \sum_{n=0}^{N} a_n k^{-2n-1} + O(k^{-2N-3}) + E\sum_{n=0}^{N'} b_n k^{-2n-1} + O(k^{-2N'-3}), \quad \text{(C1)}$$

$$\eta^-(k) = \sum_{n=0}^{N} a_n k^{-2n-1} + O(k^{-2N-3}) - E\sum_{n=0}^{N'} b_n k^{-2n-1} + O(k^{-2N'-3}), \quad \text{(C2)}$$

in which we choose $b_0 = -2\mu(\infty)$; obviously these expansions are consistent with Eqs. (4.9) and (5.30′). The absence of logarithmic terms is a consequence of the conditions imposed on the potentials (see Sec. 3). From these we define

$$\eta^S(k) = \eta^+(k) + \eta^-(k) = 2\sum_{n=0}^{N} a_n k^{-2n-1} + O(k^{-2N-3}), \quad \text{(C3)}$$

$$\eta^D(k) = \frac{k}{E}[\eta^+(k) - \eta^-(k)] = 2\sum_{n=0}^{N'} b_n k^{-2n} + O(k^{-2N'-2}), \quad \text{(C4)}$$

$$\eta^D(k) + 2\mu(\infty) = 2\sum_{n=0}^{N'} b_{n+1} k^{-2n-2} + O(k^{-2N'-4}). \quad \text{(C5)}$$

Then we can write (9.2) as

$$N_1(k) = -\frac{P}{\pi}\int_{-\infty}^{\infty} dk' \frac{k'}{k'^2 - k^2}\eta^S(k') \quad \text{(C6)}$$

and, introducing the asymptotic expansion

$$N_1(k) = \sum_{n=1}^{N} s_n k^{-2n} + O(k^{-2N-2}), \quad \text{(C7)}$$

we get for the coefficients (making use of the results of the Schrödinger case)

$$s_n = \frac{-1}{\pi(2n-1)!}\int_{-\infty}^{\infty} dk \mathfrak{D}_k^{2n-1}\eta^S(k). \quad \text{(C8)}$$

For the function $N_2(k)$, from (9.4), we get

$$N_2(k) = -\frac{P}{\pi}\int_{-\infty}^{\infty} dk' \frac{1}{k'^2 - k^2}[\eta^D(k') + 2\mu(\infty)]; \quad \text{(C9)}$$

therefore, an asymptotic expansion of the type

$$N_2(k) = \sum_{n=1}^{N} d_n k^{-2n} + O(k^{-2N-2}) \quad \text{(C10)}$$

implies for the coefficients the expression

$$d_n = \frac{-1}{\pi(2n-1)!}\int_{-\infty}^{\infty} dk \mathfrak{D}_k^{2n-1}\{k^{-1}[\eta^D(k) + 2\mu(\infty)]\} = \frac{-1}{\pi(2n-1)!}\int_{-\infty}^{\infty} dk \mathfrak{D}_k^{2n-1}[k^{-1}\eta^D(k)]. \quad \text{(C11)}$$

In the same way, having

$$\text{Re } \bar{S}(k) = -\frac{P}{\pi}\int_{-\infty}^{\infty} dk' \frac{k'}{k'^2 - k^2}\text{Im } \bar{S}(k') \quad \text{(C12)}$$

and taking the real part in (8.6)

$$\text{Re } \bar{S}(k) = \sum_{n=1}^{N}(-)^n \frac{\bar{S}_{2n-1}}{2^{2n}} k^{-2n} + O(k^{-2N-2}), \quad \text{(C13)}$$

we obtain

$$(-)^n 2^{-2n}\bar{S}_{2n-1} = \frac{-1}{\pi(2n-1)!}\int_{-\infty}^{\infty} dk \mathfrak{D}_k^{2n-1}\text{Im } \bar{S}(k). \quad \text{(C14)}$$

From this and (8.9), we obtain (9.12). The same procedure may be used to go from (8.15) to (9.13).

* Supported by a fellowship of the Consejo Nacional de Investigaciones Científicas y Técnicas (Argentina). On leave from the Instituto de Matemática, Astronomía y Física, Córdoba, Argentina.

[1] M. G. Gasimov and B. M. Levitan, Dokl. Akad. Nauk SSSR **167**, 967 (1966).

[2] M. Verde, Nuovo Cimento **2**, 1001 (1955).
[3] V. S. Buslaev and L. D. Faddeev, Dokl. Akad. Nauk SSSR **132**, 13 (1960) [Sov. Math. Dokl. **1**, 451 (1960)].
[4] M. J. Roberts, Proc. Phys. Soc. (London) **84**, 825 (1964).
[5] F. Calogero and A. Degasperis, J. Math. Phys. **9**, 90 (1968).
[6] R. G. Newton, Phys. Rev. **101**, 1588 (1956).
[7] R. G. Newton, J. Math. Phys. **1**, 319 (1960).
[8] L. D. Faddeev, Dokl. Akad. Nauk SSSR **115**, 878 (1957).
[9] I. C. Percival, Proc. Phys. Soc. (London) **80**, 1290 (1962).
[10] A. Degasperis, J. Math. Phys. **11**, 551 (1970).
[11] F. Prats and J. S. Toll, Phys. Rev. **113**, 363 (1969).
[12] M. C. Barthélémy, Ann. Inst. Henri Poincaré **7**, 115 (1967).
[13] G. Parzen, Phys. Rev. **80**, 261 (1950).
[14] E. C. Titchmarsh, *Theory of Fourier Integrals* (Clarendon, Oxford, 1948), p. 128.
[15] E. Corinaldesi, Nuovo Cimento **11**, 468 (1954).
[16] We use a procedure similar to that used by Verde, Ref. 2, in the nonrelativistic case.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 5 MAY 1970

Scalar-Tensor Theory in Canonical Form. I. The Question of Mach's Principle*†

EDWARD T. TOTON‡
Physics Department, University of Maryland, College Park, Maryland

(Received 12 December 1969)

By first casting the scalar-tensor theory into canonical form, the constraint equations are obtained from which it is possible to discuss the self-energy problem of the neutral particle and the electron. With the assumptions of Minkowski space and a constant scalar field as boundary conditions at infinity, it is shown by variational arguments that the inertia of these objects is unchanged from the corresponding general relativity values. Accordingly, if the boundary value of the scalar is altered by a change of matter at infinity, the mass of the electron is unchanged. From this point of view, the scalar-tensor theory is no more compatible with Mach's principle than is general relativity.

INTRODUCTION

In 1916 Einstein, influenced by the writings of Mach, proposed a new theory of gravitation which was to be compatible with the concept of the relativity of all motion. The special theory of relativity was partially successful in excising the Newtonian absolute space from physics by establishing an equivalence of inertial frames; the general theory made another step towards the elimination of the concept of absolute space by establishing the fundamental validity of essentially any coordinate frame.

However, since that time there have been many criticisms of general relativity because of what is referred to as Mach's principle. Owing to the large number of formulations of Mach's principle, it has itself been criticized as being too vague, but there are essentially two dominant themes common to all its formulations. One form emphasizes the intrinsic inertial properties of matter by demanding that the inertia of a body is due to the presence and distribution of matter in the universe; for example: "... matter has inertia only because there is other matter in the universe" (Sciama)[1]; and "From Mach's point of view, the inertial reaction may be considered to be gravitational in origin" (Dicke).[2] The second form emphasizes the properties of coordinate systems themselves by demanding that inertial frames must be completely determined by the presence and distribution of matter in the universe; for example: "According to Mach, then, inertial frames are those which are unaccelerated relative to the 'fixed stars'—i.e., relative to some suitably defined average over all the matter in the universe" (Sciama)[1]; and "By Mach's principle the inertial coordinate system ought to be uniquely determined by the mass distribution (in the universe)" (Dicke).[3]

The first form of Mach's principle is not substantiated by general relativity and can be shown by a particular solution of the Einstein field equations. We imagine a thin, spherically symmetric mass shell situated in an otherwise empty universe. Furthermore, we suppose that the shell is at rest relative to the inertial frame at infinity. It is then well-known that an inertial frame can be constructed everywhere within the shell, and thus all local physics conducted at the center of the shell and all physical parameters, including the masses of elementary particles, must be independent of its mass or radius; the mass of a single electron placed at the center of the shell is unaffected even if the radius of the shell is increased without limit. Thus, in general relativity the inertia of bodies is not due to the distribution of matter in the universe.

[2] M. Verde, Nuovo Cimento **2**, 1001 (1955).
[3] V. S. Buslaev and L. D. Faddeev, Dokl. Akad. Nauk SSSR **132**, 13 (1960) [Sov. Math. Dokl. **1**, 451 (1960)].
[4] M. J. Roberts, Proc. Phys. Soc. (London) **84**, 825 (1964).
[5] F. Calogero and A. Degasperis, J. Math. Phys. **9**, 90 (1968).
[6] R. G. Newton, Phys. Rev. **101**, 1588 (1956).
[7] R. G. Newton, J. Math. Phys. **1**, 319 (1960).
[8] L. D. Faddeev, Dokl. Akad. Nauk SSSR **115**, 878 (1957).
[9] I. C. Percival, Proc. Phys. Soc. (London) **80**, 1290 (1962).
[10] A. Degasperis, J. Math. Phys. **11**, 551 (1970).
[11] F. Prats and J. S. Toll, Phys. Rev. **113**, 363 (1969).
[12] M. C. Barthélémy, Ann. Inst. Henri Poincaré **7**, 115 (1967).
[13] G. Parzen, Phys. Rev. **80**, 261 (1950).
[14] E. C. Titchmarsh, *Theory of Fourier Integrals* (Clarendon, Oxford, 1948), p. 128.
[15] E. Corinaldesi, Nuovo Cimento **11**, 468 (1954).
[16] We use a procedure similar to that used by Verde, Ref. 2, in the nonrelativistic case.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 5 MAY 1970

Scalar-Tensor Theory in Canonical Form. I. The Question of Mach's Principle*†

EDWARD T. TOTON‡
Physics Department, University of Maryland, College Park, Maryland

(Received 12 December 1969)

By first casting the scalar-tensor theory into canonical form, the constraint equations are obtained from which it is possible to discuss the self-energy problem of the neutral particle and the electron. With the assumptions of Minkowski space and a constant scalar field as boundary conditions at infinity, it is shown by variational arguments that the inertia of these objects is unchanged from the corresponding general relativity values. Accordingly, if the boundary value of the scalar is altered by a change of matter at infinity, the mass of the electron is unchanged. From this point of view, the scalar-tensor theory is no more compatible with Mach's principle than is general relativity.

INTRODUCTION

In 1916 Einstein, influenced by the writings of Mach, proposed a new theory of gravitation which was to be compatible with the concept of the relativity of all motion. The special theory of relativity was partially successful in excising the Newtonian absolute space from physics by establishing an equivalence of inertial frames; the general theory made another step towards the elimination of the concept of absolute space by establishing the fundamental validity of essentially any coordinate frame.

However, since that time there have been many criticisms of general relativity because of what is referred to as Mach's principle. Owing to the large number of formulations of Mach's principle, it has itself been criticized as being too vague, but there are essentially two dominant themes common to all its formulations. One form emphasizes the intrinsic inertial properties of matter by demanding that the inertia of a body is due to the presence and distribution of matter in the universe; for example: "... matter has inertia only because there is other matter in the universe" (Sciama)[1]; and "From Mach's point of view, the inertial reaction may be considered to be gravitational in origin" (Dicke).[2] The second form emphasizes the properties of coordinate systems themselves by demanding that inertial frames must be completely determined by the presence and distribution of matter in the universe; for example: "According to Mach, then, inertial frames are those which are unaccelerated relative to the 'fixed stars'—i.e., relative to some suitably defined average over all the matter in the universe" (Sciama)[1]; and "By Mach's principle the inertial coordinate system ought to be uniquely determined by the mass distribution (in the universe)" (Dicke).[3]

The first form of Mach's principle is not substantiated by general relativity and can be shown by a particular solution of the Einstein field equations. We imagine a thin, spherically symmetric mass shell situated in an otherwise empty universe. Furthermore, we suppose that the shell is at rest relative to the inertial frame at infinity. It is then well-known that an inertial frame can be constructed everywhere within the shell, and thus all local physics conducted at the center of the shell and all physical parameters, including the masses of elementary particles, must be independent of its mass or radius; the mass of a single electron placed at the center of the shell is unaffected even if the radius of the shell is increased without limit. Thus, in general relativity the inertia of bodies is not due to the distribution of matter in the universe.

On the other hand, solutions of the Einstein equations can be found which do support the second version of Mach's principle. If the mass shell is set in slow rotation relative to the inertial frame at infinity, we find, as shown by Brill and Cohen,[4] that an inertial frame can still be constructed everywhere within the shell, but it is now also rotating with respect to the inertial frame at infinity. Furthermore, this "dragging" of the inertial frame by the shell is perfect if the radius of the shell equals its Schwarzschild radius. Under this condition, observers within the shell could not detect any rotation of the shell relative to their local inertial frame. This, in fact, can be offered as an explanation as to why all visible stars in the universe are, in the large, rotationally at rest relative to a local inertial frame.

So we see that, although general relativity does not suggest a determination of inertia by the presence of matter, it does allow a determination of inertial *frames* by matter.

The scalar-tensor theory of gravity was proposed in 1961 by Brans and Dicke[5] in order to remedy the incompatibility of general relativity with Mach's principle. Since this theory also enables one to determine the metrical properties of space and may even be cast into a form in which the Einstein equations are satisfied,[6] it is clear that the main thrust towards a full integration of Mach's principle must be through its inertial implications rather than through the influence of matter on inertial frames.

Brans and Dicke have pointed out an important consideration with regard to the inertia of bodies and to whether inertia can vary.[5] If it is said that the mass of an electron at one point of space is the same as its mass at another point, this statement is either a definition or it is meaningless. In addition to the generally accepted units of length and time, we must also make a choice for the unit of mass. If we choose the electron mass as this unit of measure, then, of course, the mass of an electron is constant by definition. On the other hand, mass ratios, such as the ratio of the proton mass to the electron mass, are dimensionless numbers. Such quantities cannot be altered by adjusting units and so may well be spatially dependent.

A case in point is the dimensionless combination $(G/\hbar c)^{\frac{1}{2}}m_e$, where G is the gravity constant and m_e is the mass of an electron. Such a combination enables us to define the mass of the electron in terms of the quantities G, $\hbar$, and c or, if we wish to choose m_e as a fundamental unit of measure, the quantity G may be expressed in terms of m_e, $\hbar$, and c.

Such consideration might be specious but for the fact that the combination $(G/\hbar c)^{\frac{1}{2}}m_e$ is an exceedingly small number, on the order of 10^{-23}. It is generally believed that in a complete theory of physics the fundamental dimensionless numbers would be of the order of unity and constant, whereas either exceedingly small or exceedingly large numbers must in some way be connected with the actual distribution of matter in the universe. It is the assumption of the Brans–Dicke theory that the quantity $(G/\hbar c)^{\frac{1}{2}}m_e$ is such a number. The original form of the theory is written such that the gravity constant G absorbs all the variable behavior of this number, but it must be recognized that this behavior can be apportioned in arbitrary amounts to each of the quantities appearing in the combination.

In order to test the inertial implications of the scalar-tensor theory, it is then evident that we must choose the representation in which the quantities G, $\hbar$, and c are constants.

The scalar-tensor theory of gravity was first introduced in the form[5,7]

$$\delta\int d^4x(-{}^4g)^{\frac{1}{2}}\left(\phi\,{}^4R-\omega\frac{\phi_{,\mu}\phi^{,\mu}}{\phi}+\frac{16\pi G_0}{c^4}L_M\right)=0. \tag{1}$$

It is understood that L_M depends in no way upon ϕ and so any mass parameter appearing in L_M is assumed to be constant by definition. On the other hand, the presence of ϕ as the coefficient of ${}^4R$ shows that the gravity constant is assumed to be a variable; in fact, $G=\phi^{-1}G_0$. This action principle can be cast into the representation in which the gravity constant is constant by definition by means of the conformal map $g_{\mu\nu}\rightarrow\phi^{-1}g_{\mu\nu}$. The result is[6]

$$\delta\int d^4x(-{}^4g)^{\frac{1}{2}}\times\left({}^4R-(\tfrac{3}{2}+\omega)\frac{\phi_{,\mu}\phi^{,\mu}}{\phi^2}+\frac{16\pi G_0}{c^4}\frac{1}{\phi^2}\bar{L}_M\right)=0. \tag{2}$$

Here, $\bar{L}_M$ is obtained from L_M by replacing $g_{\mu\nu}$ in L_M by $g_{\mu\nu}/\phi$.

If we refer back to the large matter shell, we can see how the scalar field may lead to the inertial implications of Mach's principle. Whereas in general relativity the masses of elementary particles are unaffected by the presence of the large shell because the interior inertial frame is independent of its size, we find that the scalar field in the Brans–Dicke theory is capable of reaching into the interior of the shell and affecting the elementary particles' masses therein. The scalar field will, in general, depend upon the size of the shell, and

so any change in it can lead to changes in the local physics at its center.

Specifically, the action principle (2) leads to an equation for the scalar field[6]

$$(-{}^4g)^{-\frac{1}{2}}\frac{\partial}{\partial x^\alpha}\left((-{}^4g)^{\frac{1}{2}}{}^4g^{\alpha\beta}\frac{\partial \ln \phi}{\partial x^\beta}\right)=\frac{8\pi G_0}{c^4(3+2\omega)}T. \tag{3}$$

If we look for the static solution in isotropic coordinates for the interior solution, we find that (3) becomes

$$\nabla\cdot(\chi\lambda\nabla\ln\phi)=0, \tag{4}$$

where ∇ is the flat-space gradient operator, where $ds^2=-\lambda^2(r)\,dt^2+\chi^4(r)[dx^2+dy^2+dz^2]$, and where T is zero everywhere within the shell. Since (4) must hold everywhere within the shell, we find that $\phi=$ const; this condition clearly yields a flat-space metric in the interior of the shell. If we have at our disposal a detailed model for an elementary particle, we can then compute the mass of the particle by demanding that the space become Minkowskian "at infinity" and that the scalar field take on the value given by the shell's interior solution. It is then possible to ask how the mass of the elementary particle is affected as the size or matter content of the large shell is altered.

When dealing with localized matter distributions that include such fields as the tensor and scalar fields, the use of the canonical formalism as developed by Arnowitt, Deser, and Misner[8] is most suitable. Its use is particularly enhanced if Minkowskian space is a boundary condition. The Hamiltonian of the system gives us its energy; the generators of space translations give us the momentum of the system when they are evaluated for the actual field configuration. But most important, these four quantities form a 4-vector relative to the Minkowskian space at infinity.[8] The localized system will behave like a particle in special relativistic mechanics; in the presence of an external field, it will respond with an inertia given by the magnitude of this 4-vector.

When evaluated in the rest frame of the system—i.e., in the frames in which the three spatial components of the 4-momentum are zero—the energy becomes equivalent to the magnitude of this vector. We are thus able to compute the inertia of a system by means of the self-energy; the two are equal to factors of c which, by prior choice of units, is taken to be constant.

1. THE SCALAR-TENSOR THEORY IN CANONICAL FORM

A model for the electron with the corresponding self-energy calculation has already been discussed by Arnowitt, Deser, and Misner within the framework of general relativity.[9] With minor modification one can also discuss the neutral "point" particle. Both are characterized by a δ-function source and a bare mass. Both are clothed by the tensor gravitational field where the boundary condition of Minkowskian space is assumed at infinity. In the scalar-tensor theory these models would be further characterized by clothing by the scalar field whose boundary value is determined by the matter at infinity.

Referring to Eq. (2), we see that, in the representation in which the masses of fundamental particles can vary, the scalar field behaves much as a matter field; it behaves as a source for the Einstein field (the metric) and, in fact, the Einstein form of the field equations is obtained, since only metric components comprise the curvature term of the action. This facilitates the reduction to canonical form of the scalar-tensor theory and the comparison of inertial implications between it and general relativity; the (3 + 1)-dimensional decomposition of the action, the imposition of coordinate conditions, the orthogonal decomposition of the metric, and the treatment of the constraints can be managed in the scalar-tensor theory exactly as in general relativity.

In general relativity the action

$$I_G=\int d^4\tau(-{}^4g)^{\frac{1}{2}}{}^4g^{\mu\nu}R_{\mu\nu}, \tag{5}$$

where

$$R_{\mu\nu}=\Gamma^\alpha_{\mu\nu,\alpha}-\Gamma^\alpha_{\mu\alpha,\nu}+\Gamma^\alpha_{\mu\nu}\Gamma^\beta_{\alpha\beta}-\Gamma^\alpha_{\mu\beta}\Gamma^\beta_{\nu\alpha} \tag{6}$$

will yield the Einstein equations

$$R_{\mu\nu}-\tfrac{1}{2}{}^4g_{\mu\nu}R=0 \tag{7}$$

and the relations

$$(-{}^4g)^{-\frac{1}{2}}[(-{}^4g)^{\frac{1}{2}}{}^4g^{\mu\nu}]_{,\alpha}+{}^4g^{\mu\beta}\Gamma^\nu_{\beta\alpha}+{}^4g^{\nu\beta}\Gamma^\mu_{\beta\alpha}-{}^4g^{\mu\nu}\Gamma^\beta_{\beta\alpha}=0 \tag{8}$$

(from which the affine connection can be expressed in terms of the ${}^4g_{\mu\nu}$ and their derivatives) upon independent variation of the ${}^4g^{\mu\nu}$ and the $\Gamma^\alpha_{\mu\nu}$. An equivalent form of this action principle can be obtained by means of a (3 + 1)-dimensional decomposition in which we define[8]

$$g_{ij}={}^4g_{ij},\quad N=(-{}^4g^{00})^{-\frac{1}{2}},\quad N_i={}^4g_{0i}, \tag{9a}$$

$$\pi^{ij}=(-{}^4g)^{\frac{1}{2}}(\Gamma^0_{pq}-g_{pq}g^{rs}\Gamma^0_{rs})g^{ip}g^{jq}. \tag{9b}$$

Latin indices run from 1 through 3, and the quantities g^{ij} form the matrix reciprocal to g_{ij}. Insertion of these definitions into I_G yields an equivalent Hamiltonian form[8]

$$I_G=\int d^4\tau\left(\pi^{ij}\frac{\partial g_{ij}}{\partial\tau^0}-NR^0-N^iR^i\right), \tag{10}$$

where

$$R^0 = -g^{\frac{1}{2}}[{}^3R + g^{-1}(\tfrac{1}{2}\pi_i^i\pi_j^j - \pi^{ij}\pi_{ij})], \quad (11a)$$

$$R^i = -2\pi^{ij}|_j, \quad (11b)$$

and where both a total time derivative and a divergence have been discarded. Also, "$|_j$" means the covariant derivative with respect to the jth coordinate. It is now understood that the π^{ij}, g_{ij}, N, and N^i are independently varied quantities and that indices are raised and lowered with the 3-space metric g_{ij}.

This decomposition of the action can be carried over into the scalar-tensor theory because only metric components appear in the curvature term. For a complete reduction of the scalar-tensor theory to canonical form, it is also necessary to cast the scalar part of the action into Hamiltonian form. If we take the Lagrangian density to be

$$\mathsf{L}_\phi = -(\tfrac{3}{2} + \omega)(-{}^4g)^{\frac{1}{2}}\frac{\phi_{,\mu}\phi^{,\mu}}{\phi^2}, \quad (12)$$

then an insertion of the (3 + 1)-dimensional decomposition of the metric (9) yields

$$\mathsf{L}_\phi = -\frac{(\frac{3}{2} + \omega)(g)^{\frac{1}{2}}}{N\phi^2} \times [-(\phi_{,0})^2 + 2N^i\phi_{,i}\phi_{,0} + (N^2g^{ij} - N^iN^j)\phi_{,i}\phi_{,j}]. \quad (13)$$

A conversion to Hamiltonian form can be made in the usual way by means of the definitions

$$\pi = \frac{\partial \mathsf{L}}{\partial \phi_{,0}}, \quad h = \pi\phi_{,0} - \mathsf{L}, \quad (14)$$

with the result that

$$\mathsf{L}_\phi = \pi\frac{\partial\phi}{\partial\tau^0} - N\pi^i\phi_{,i} - N\left(\frac{\phi^2\pi^2}{(6 + 4\omega)(g)^{\frac{1}{2}}} + (\tfrac{3}{2} + \omega)g^{\frac{1}{2}}g^{ij}\frac{\phi_{,i}\phi_{,j}}{\phi^2}\right), \quad (15)$$

and so the action for the tensor and scalar fields can be written

$$I_{G+\phi} = \int d^4\tau\left(\pi^{ij}\frac{\partial g_{ij}}{\partial\tau^0} + \pi\frac{\partial\phi}{\partial\tau^0} - NR^0 - N_iR^i\right), \quad (16)$$

where now

$$R^0 = -(g)^{\frac{1}{2}}[{}^3R + g^{-1}(\tfrac{1}{2}\pi_i^i\pi_j^j - \pi^{ij}\pi_{ij})] + \frac{\phi^2\pi^2}{(6 + 4\omega)g^{\frac{1}{2}}} + (\tfrac{3}{2} + \omega)g^{\frac{1}{2}}g^{jk}\frac{\phi_{,j}\phi_{,k}}{\phi^2}, \quad (17a)$$

$$R^i = -2\pi^{ij}|_j + \pi\phi^{,i}. \quad (17b)$$

If we were in fact considering the source-free scalar-tensor field, $I_{G+\phi}$ would be the full action in Hamiltonian form. A reduction to canonical form requires the use of the constraints obtained from (16)—namely, $R^0 = R^i = 0$—to eliminate the constrained variables from the action. Substitution of these constraints (obtained by extremizing $I_{G+\phi}$ with respect to N and N^i independently) into $I_{G+\phi}$ yields

$$I_{G+\phi} = \int d^4\tau\left(\pi^{ij}\frac{\partial g_{ij}}{\partial\tau^0} + \pi\frac{\partial\phi}{\partial\tau^0}\right), \quad (18)$$

where it is understood that four of the quantities in the integrand are expressed in terms of the others through the constraints.

It is further possible to make an orthogonal decomposition of the metric components. For any symmetric array f_{ij}, it is possible to write

$$f_{ij} = f_{ij}^{\mathrm{TT}} + f_{ij}^{\mathrm{T}} + (f_{i,j}^{\mathrm{T}} + f_{j,i}^{\mathrm{T}}) + f_{,ij}^{\mathrm{L}} \quad (19)$$

if f_{ij}^{T} is defined according to

$$f_{ij}^{\mathrm{T}} = \tfrac{1}{2}[\delta_{ij}f^{\mathrm{T}} - (1/\nabla^2)f_{,ij}^{\mathrm{T}}]. \quad (20)$$

Here, $(1/\nabla^2)$ is the inverse of the flat-space Laplacian operator with suitable boundary conditions. Also, by definition $f_{ij,j}^{\mathrm{TT}} = f_{ii}^{\mathrm{TT}} = f_{ij,j}^{\mathrm{T}} = f_{ii}^{\mathrm{T}} = 0$. Insertion of this decomposition into (18) allows us to write in one of several possible forms:

$$I_{G+\phi} = \int d^4\tau\left(\pi\frac{\partial\phi}{\partial\tau^0} + \pi^{ij^{\mathrm{TT}}}\frac{\partial g_{ij}^{\mathrm{TT}}}{\partial\tau^0} + [-\nabla^2(\pi^{i\mathrm{T}} + \pi_{,i}^{\mathrm{L}})]\frac{\partial}{\partial\tau^0}[-(1/2\nabla^2)g_{,i}^{\mathrm{T}}] - (-\nabla^2g^{\mathrm{T}})\frac{\partial}{\partial\tau^0}[(-1/2\nabla^2)(\pi^{\mathrm{T}} + \nabla^2\pi^{\mathrm{L}})]\right). \quad (21)$$

If we choose the coordinate conditions

$$x^i = 2g_i - (1/2\nabla^2)g_{,i}^{\mathrm{T}}, \quad (22a)$$

$$t = -(1/2\nabla^2)(\pi^{\mathrm{T}} + \nabla^2\pi^{\mathrm{L}}), \quad (22b)$$

then the action reduces to

$$I_{G+\phi} = \int d^3x\,dt\left(\pi\frac{\partial\phi}{\partial t} + \pi^{ij^{\mathrm{TT}}}\frac{g_{ij}^{\mathrm{TT}}}{\partial t} - h\right), \quad (23)$$

where

$$h = -\nabla^2g^{\mathrm{T}}. \quad (24)$$

Once the variables g^{T}, $\pi^{i\mathrm{T}}$, and π^{L} are expressed in terms of the other variables by means of the constraints, the action $I_{G+\phi}$ is in canonical form with h a functional of the variables π, ϕ, $\pi^{ij^{\mathrm{TT}}}$, and g_{ij}^{TT}; if these quantities are specified initially, then the Hamilton equations

obtained from (23) guarantee their unique time development. Furthermore, the previous claim that the orthogonal decomposition of the metric and the imposition of coordinate conditions can be carried out exactly as in general relativity is substantiated in the above equations. The main alteration of the treatment of the physical system appears, rather, in the new constraints [Eqs. (17)].

In order to discuss the self-energy of the neutral particle and the electron, we must also include in the action those quantities related to the coupling with the electromagnetic and mechanical variables. The Lagrangian density for a point charge in general relativity is given by[8]

$$\begin{aligned}\mathsf{L}_M &= \tfrac{1}{2}(A_{\mu,\nu} - A_{\nu,\mu})F^{\mu\nu} + \tfrac{1}{4}(-{}^4g)^{-\frac{1}{2}}F^{\mu\nu}F^{\alpha\beta}\,{}^4g_{\mu\alpha}\,{}^4g_{\mu\beta} \\ &\quad + \int ds e \frac{dx^\mu}{ds}(s)A_\mu(\tau)\delta^4(\tau - x(s)) \\ &\quad + \int ds\left[\pi_\mu \frac{dx^\mu}{ds} - \tfrac{1}{2}\lambda'(s)(\pi_\mu\pi_\nu\,{}^4g^{\mu\nu} + m_0^2)\right] \\ &\quad \times \delta^4(\tau - x(s)). \qquad (25)\end{aligned}$$

The appropriate quantities for varying are the vector potential A_μ, the field $F^{\mu\nu}$, the coordinate position of the particle x^μ, and the momentum of the particle π_μ. We introduce $\lambda'(s)$ as a Lagrange multiplier in order to maintain the arbitrary-parameterization of the action in the parameter s. This density can be expressed in the scalar-tensor theory by means of the prescription given in Eq. (2) with the result

$$\begin{aligned}\frac{1}{\phi^2}\bar{\mathsf{L}}_M &= \tfrac{1}{2}(A_{\mu,\nu} - A_{\nu,\mu})F^{\mu\nu} + \tfrac{1}{4}(-{}^4g)^{-\frac{1}{2}}F^{\mu\nu}F^{\alpha\beta}\,{}^4g_{\mu\nu}\,{}^4g_{\nu\beta} \\ &\quad + \int ds e \frac{dx^\mu}{ds}(s)A_\mu(\tau)\delta^4(\tau - x(s)) \\ &\quad + \int ds\left[\pi_\mu \frac{dx^\mu}{ds} - \tfrac{1}{2}\lambda'(s)(\pi_\mu\pi_\nu\phi\,{}^4g^{\mu\nu} + m_0^2)\right] \\ &\quad \times \delta^4(\tau - x(s)), \qquad (26)\end{aligned}$$

where ϕ is absorbed into the variables $F^{\mu\nu}$.

In order to reduce to canonical form the complete system it is necessary to effect a (3 + 1)-dimensional decomposition of the "interaction" Lagrangian density as well. In 3-dimensional notation

$$E^i = F^{0i}, \quad B^i = \tfrac{1}{2}\epsilon^{ijk}(A_{k,j} - A_{j,k}). \qquad (27)$$

In terms of these quantities and the 3-dimensional notation of the metric, the above Lagrangian becomes

$$\begin{aligned}&\frac{1}{\phi^2}\bar{\mathsf{L}}_M \\ &= -E^i\frac{\partial A_i}{\partial\tau^0} + \delta^3(\tau - x(\tau^0))\left\{[\pi_i(\tau^0) + eA_i]\frac{dx^i}{d\tau^0} + \pi^0\right\} \\ &\quad + A_0[e\delta^3(\tau - x(\tau^0)) - E^i_{,i}] \\ &\quad - \tfrac{1}{2}\lambda[\phi g^{ij}\pi_i\pi_j - \phi N^{-2}(\pi_0 - N^i\pi_i)^2 + m_0^2] \\ &\quad \times \delta^3(\tau - x(\tau^0)) \\ &\quad - \tfrac{1}{2}Ng^{-\frac{1}{2}}g_{ij}(E^iE^j + B^iB^j) + N^i\epsilon_{ijk}E^jB^k. \qquad (28)\end{aligned}$$

In obtaining the above, an integration over the arbitrary parameter s has been made where $\lambda = \lambda'(s)\,ds/dx^0(s)$ evaluated at $x^0(s) = \tau^0$.

It is now convenient to eliminate the nongravitational constraints. Varying the total action with respect to A_0 and λ will produce a set of constraints because these quantities appear nowhere differentiated in the above Lagrangian. The constraints are found to be

$$E^i_{,i} = e\delta^3(\tau - x(\tau^0)), \qquad (29a)$$

$$\phi N^{-2}(\pi_0 - N^i\pi_i)^2 = \phi g^{ij}\pi_i\pi_j + m_0^2, \qquad (29b)$$

where solutions are

$$E^{\mathrm{L}} = -\nabla\frac{e}{4\pi\,|\tau - x(\tau^0)|}, \qquad (30a)$$

$$\pi^0 = N^i\pi_i - N(g^{ij}\pi_i\pi_j + \phi^{-1}m_0^2)^{\frac{1}{2}}. \qquad (30b)$$

$E^{i\mathrm{L}}$ is the longitudinal part of the vector E^i; i.e., we have implicitly assumed the orthogonal decomposition $E^i = E^{i\mathrm{T}} + E^{i\mathrm{L}}$. If we insert these relations into (28), bearing in mind the orthogonal decomposition of the vectors E^i and A^i, we find

$$\begin{aligned}\frac{1}{\phi^2}\bar{\mathsf{L}}_M &= (-E^{i\mathrm{T}})\frac{\partial A_i^{\mathrm{T}}}{\partial\tau^0} + p_i\frac{dx^i}{d\tau^0}\delta^3(\tau - x(\tau^0)) \\ &\quad - \tfrac{1}{2}Ng^{-\frac{1}{2}}g_{ij}(E^iE^j + B^iB^j) \\ &\quad - N[g^{ij}(p_i - eA_i^{\mathrm{T}})(p_j - eA_j^{\mathrm{T}}) + \phi^{-1}m_0^2]^{\frac{1}{2}} \\ &\quad \times \delta^3(\tau - x(\tau^0)) \\ &\quad + N^i[\epsilon_{ijk}E^jB^k + (p_i - eA_i^{\mathrm{T}})\delta^3(\tau - x(\tau^0))], \qquad (30c)\end{aligned}$$

where $p_i = \pi_i + eA_i^{\mathrm{T}}$.

The independent excitations of the electromechanical system can now be seen to be A_i^{T} and its conjugate momentum $(-E^{i\mathrm{T}})$, the x^i, and the p_i. At this point the variables N, N^i, and ϕ are unspecified, but their significance becomes complete when this Lagrangian is added to those corresponding to the scalar and

tensor components of the gravitational field. The total action then becomes[10]

$$I = \int d^4\tau \Big[\pi^{ij} \frac{\partial g_{ij}}{\partial \tau^0} + \pi \frac{\partial \phi}{\partial \tau^0} + (-E^{i\mathrm{T}}) \frac{\partial A_i^{\mathrm{T}}}{\partial \tau^0} + p_i \frac{dx^i}{d\tau^0} \delta^3(\tau - x(\tau^0)) - NR^0 - N_i R^i \Big], \quad (31)$$

where now

$$R^0 = -(g)^{\frac{1}{2}} [{}^3R + g^{-1}(\tfrac{1}{2}\pi_i^i \pi_j^j - \pi^{ij}\pi_{ij})] - \frac{\phi^2 \pi^2}{(6+4\omega) g^{\frac{1}{2}}} + (\tfrac{3}{2} + \omega) g^{\frac{1}{2}} g^{jk} \frac{\phi_{,j}\phi_{,k}}{\phi^2} + \frac{1}{2g^{\frac{1}{2}}} g_{ij}(E^i E^j + B^i B^j) + [g^{ij}(p_i - eA_i^{\mathrm{T}})(p_j - eA_j^{\mathrm{T}}) + \phi^{-1} m_0^2]^{\frac{1}{2}} \times \delta^3(\tau - x(\tau^0)), \quad (32a)$$

$$R^i = -2\pi^{ij}\big|_j + \pi\phi^{,i} - \epsilon_{njk} g^{in} E^j B^k - (p^i - eA^{i\mathrm{T}}) \delta^3(\tau - x(\tau^0)). \quad (32b)$$

Again, as for the free scalar-tensor field, the orthogonal decomposition of the metric and the choice of coordinates can be made in the above action since these manipulations are most directly related to the first term in the integrand. If the same choice of coordinates is made as Eqs. (22), then the action reduces to

$$I = \int d^3x\, dt \Big[\pi^{ij^{\mathrm{TT}}} \frac{\partial g_{ij}^{\mathrm{TT}}}{\partial t} + \pi \frac{\partial \phi}{\partial t} + (-E^{i\mathrm{T}}) \frac{\partial A_i^{\mathrm{T}}}{\partial t} + p_i \frac{dx^i}{dt} \delta^3(x - x(t)) - h \Big], \quad (33)$$

where h is a functional of the quantities $\{\pi^{ij^{\mathrm{TT}}}, g_{ij}^{\mathrm{TT}}; \pi, \phi; (-E^{i\mathrm{T}}), A_i^{\mathrm{T}}; p_i, x^i\}$ and is obtained by solving the constraint relations from Eqs. (32); i.e., $R^0 = R^i = 0$.

I is now in canonical form; the set of quantities $\{\pi^{ij^{\mathrm{TT}}}, g_{ij}^{\mathrm{TT}}; \pi, \phi; (-E^{i\mathrm{T}}), A_i^{\mathrm{T}}; p_i, x^i\}$ comprises the minimal set of Cauchy data for the charged particle in the scalar-tensor theory; arbitrary specification of these quantities initially will lead to their unique time evolution by means of Hamilton's equations generated by extremizing I.

2. SELF-ENERGY OF THE ELECTRON AND NEUTRAL PARTICLE IN THE SCALAR-TENSOR THEORY; COMPARISON WITH GENERAL RELATIVITY

From the canonical form of a theory we can recognize the energy of a system as the value of the Hamiltonian for the actual fields present. The self-energy of a particle is characterized by the static fields associated with the particle. All independent excitations related to the presence of waves are excluded. In the case of either the neutral particle or the electron, the independent excitations of the tensor field and the electromagnetic field may immediately be set equal to zero. The scalar field cannot be so treated since no gauge condition on ϕ exists and there is no natural decomposition of the scalar into that part describing waves and that related to the source. The static condition is, however, characterized by a minimum energy configuration of the fields since any higher energy will in general decay to the minimum energy configuration by the emission of waves. The self-energy of a particle can then be characterized by setting all independent excitations equal to zero except for the scalar field and by inserting into the Hamiltonian the scalar field appropriate to the minimum energy configuration.

From the action in (33) we see that the Hamiltonian is

$$H = \int d^3x h = -\int d^3x \nabla^2 g^{\mathrm{T}}, \quad (34)$$

where it is assumed that g^{T} is expressed in terms of the independent excitations by solving the constraints. Instead of directly expressing g^{T} in terms of the independent excitations, it is more convenient to first insert the static self-energy conditions into the constraints and then solve for the appropriate g^{T}. If we place the charge at rest at the origin of coordinates and leave ϕ for the moment unspecified, the static condition is characterized by $g_{ij}^{\mathrm{TT}} = \pi^{ij^{\mathrm{TT}}} = p_i = \pi = E^{i\mathrm{T}} = A_i^{\mathrm{T}} = 0$. With the coordinate conditions (22) we are dealing with the time-symmetric situation $\pi^{ij} = 0$. We see, then, that the constraints from (32b)—namely, $R^i = 0$—are automatically satisfied; the "energy" constraint $R^0 = 0$ leads to

$$g^{\frac{1}{2}}\, {}^3R = (\tfrac{3}{2} + \omega) g^{\frac{1}{2}} g^{ij} \frac{\phi_{,i}\phi_{,j}}{\phi^2} + \frac{1}{2g^{\frac{1}{2}}} g_{ij} E^{i\mathrm{L}} E^{j\mathrm{L}} + \frac{m_0 \delta^3(x)}{\phi^{\frac{1}{2}}}. \quad (35)$$

A special advantage of the choice of coordinates is the following: Substitution of the coordinate conditions into the orthogonal decomposition of the metric leads to

$$g_{ij} = g_{ij}^{\mathrm{TT}} + \delta_{ij}(1 + \tfrac{1}{2} g^{\mathrm{T}}). \quad (36)$$

Thus, in the above static models, the coordinates reduce to isotropic coordinates. If we assume $g_{ij} = \chi^4(r)\delta_{ij}$, then the constraint (35) reduces to

$$g^{\frac{1}{2}}\, {}^3R = -8\chi\nabla^2\chi = (\tfrac{3}{2} + \omega)\chi^2 \Big(\frac{\nabla\phi}{\phi}\Big)^2 + \frac{(E^{\mathrm{L}})^2}{2\chi^2} + \frac{m_0 \delta^3(x)}{\phi^{\frac{1}{2}}}. \quad (37)$$

In terms of χ, the self-energy of the particle can be written

$$m = -\int d^3x \nabla^2 g^{\mathrm{T}} = -2\int d^3x \nabla^2 \chi^4. \tag{38}$$

Because we always wish to maintain the boundary condition of Minkowskian space at infinity, the condition of $\chi \to 1$ as $r \to \infty$ is imposed. We can thus write the identity

$$\int d^3x \nabla^2 \chi^n = \int_\infty d\mathbf{s} \cdot \nabla \chi^n = n \int d\mathbf{s} \cdot \nabla \chi = n \int d^3x \nabla^2 \chi. \tag{39}$$

The self-energy can therefore be written

$$m = -8\int d^3x \nabla^2 \chi. \tag{40}$$

From this it is clear that the self-energy of the system is determined as 32π times the coefficient of the r^{-1} term appearing in the asymptotic expansion of χ.

The determination of the mass of either the neutral point particle or the electron is not yet complete since the scalar field is as yet unspecified in (37). But, because we expect the minimum energy situation to arise in the static case for some configuration of the scalar field, we need only consider all scalar field configurations which are initially static and choose that configuration which gives the minimum value of m.

Consider, first, the neutral particle. The constraint equation is

$$-8\chi \nabla^2 \chi = (\tfrac{3}{2} + \omega)\chi^2 \left(\frac{\nabla \phi}{\phi}\right)^2 + \frac{m_0 \delta^3(x)}{\phi^{\frac{1}{2}}}, \tag{41}$$

where that value of ϕ is to be inserted in (41) which minimizes m. An easy comparison can be made with general relativity by considering the constraint equation arising in the Einstein theory:

$$-8\chi \nabla^2 \chi = m_0 \delta^3(x). \tag{42}$$

A rigorous solution of this equation leads to not only a zero mass but also to a Minkowski space everywhere away from the origin.[9] Furthermore, χ becomes infinitely large at the origin. If n is chosen equal to -2 in Eq. (39), then the mass can be written as

$$\begin{aligned} m &= -8\int d^3x \nabla^2 \chi = 4\int d^3x \nabla^2 \chi^{-2} \\ &= 4\int d^3x \left(\frac{6(\nabla \chi)^2}{\chi^4} - \frac{2\nabla^2 \chi}{\chi^3}\right) \\ &= \int d^3x \left(\frac{24(\nabla \chi)^2}{\chi^4} + (\tfrac{3}{2} + \omega)\frac{(\nabla \phi)^2}{\phi^2 \chi^2} + \frac{m_0 \delta^3(x)}{\phi^{\frac{1}{2}} \chi^4}\right), \end{aligned} \tag{43}$$

the last statement being true by virtue of the constraint equation. Because each of the terms are positive or zero, we see that the mass must be greater than or equal to zero whatever the scalar field and that the mass is zero for a constant scalar field, since a constant scalar in (41) reproduces (42) (except for an alteration of the bare mass).

Even though zero is the least value m can attain, does this correspond to a minimum energy configuration of the scalar field? By elimination of $\nabla^2 \chi$ in (40), by means of the constraint, we have

$$m = \int d^3x \left[(\tfrac{3}{2} + \omega)\chi \left(\frac{\nabla \phi}{\phi}\right)^2 + \frac{m_0 \delta^3(x)}{\phi^{\frac{1}{2}} \chi}\right]. \tag{44}$$

A first variation of m with respect to ϕ gives

$$\delta m = \int d^3x \left[(\tfrac{3}{2} + \omega)[(\nabla \ln \phi)^2 \delta\chi + 2\chi \nabla \ln\phi \cdot \delta \nabla \ln \phi] + \frac{m_0 \delta^3(x)}{\phi^{\frac{1}{2}} \chi}\left(-\frac{\delta\phi}{2\phi} - \frac{\delta\chi}{\chi}\right)\right], \tag{45}$$

where $\delta\chi$ is related to $\delta\phi$ by maintaining the constraint. It is clear that, if ϕ is chosen to be constant, then the first bracketed term vanishes because $\nabla \ln \phi$ vanishes and that the contribution of the second term is zero because at the origin χ becomes infinitely large as the solution of (41) for ϕ a constant.

We see, then, that, if ϕ is a constant, the minimum energy configuration is obtained and the mass is zero. The first variation of m is zero at the point where m is zero. Since m can be no smaller than zero, this point must be a minimum.

In general relativity the mass of such an object is zero even if placed near the center of a large matter shell. Even if we, for the moment, entertain the hope that the neutral particle will gain a mass as a consequence of the presence of the scalar and the coupling to the shell by the scalar, we see that the mass of the neutral particle is unchanged from the general relativity result.

A structureless particle, such as the neutral particle just discussed, is perhaps unrealistic. If nongravitational interactions are added to the particle, such as the electromagnetic field associated with a charge, the possibility of differences in the inertial implications of the two theories can result. A distribution of energy throughout space in the form of the electromagnetic energy density about the charge gives us this structure. But under such circumstances general relativity already assigns an inertia to such systems; it appears unlikely that the scalar-tensor theory can modify this situation to the point where the inertia of the object is totally due to all matter in the universe, but it is more likely to cause an influence on the mass of the object.

We consider the constraint equation for the point charge (37),

$$-8\chi\nabla^2\chi = (\tfrac{3}{2}+\omega)\chi^2\left(\frac{\nabla\phi}{\phi}\right)^2 + \frac{(E^{\mathrm{L}})^2}{2\chi^2} + \frac{m_0\delta^3(x)}{\phi^{\frac{1}{2}}}.$$

Here the electric field is given as the solution of the nongravitational constraint Eqs. (29a) and (30a). The corresponding system in general relativity is obtained from the above equation by eliminating those terms related to the scalar field

$$-8\chi_{\mathrm{E}}\nabla^2\chi_{\mathrm{E}} = \frac{(E^{\mathrm{L}})^2}{2\chi_{\mathrm{E}}^2} + m_0\delta^3(x). \tag{46}$$

The subscript E is added to χ to designate its particular relation to the Einstein theory. Arnowitt, Deser, and Misner[9] have shown that a rigorous solution of (46) leads to

$$\chi_{\mathrm{E}} = 1 + \frac{|e|^{\frac{1}{2}}}{8\pi r} \tag{47}$$

and a singularity at the origin of the form $\epsilon^{-\frac{1}{2}}$ as $\epsilon \to 0$ and that the total mass of the electron is $m = 2|e|$. We can compare this result with an expected result in the scalar-tensor theory by examining variations of the mass integral in the vicinity of the point where the Einstein constraint (46) is satisfied in a manner similar to the neutral particle case.

We first show that the mass of the charged particle is positive definite. According to Eq. (43) and the constraint (37) m can be written

$$\begin{aligned} m &= 4\int d^3x\left(\frac{6(\nabla\chi)^2}{\chi^4} - \frac{2\nabla^2\chi}{\chi^3}\right) \\ &= \int d^3x\left(\frac{24(\nabla\chi)^2}{\chi^4} + (\tfrac{3}{2}+\omega)\frac{(\nabla\phi)^2}{\chi^2\phi^2} \right. \\ &\quad \left. + \frac{(E^{\mathrm{L}})^2}{2\chi^6} + \frac{m_0\delta^3(x)}{\phi^{\frac{1}{2}}\chi^4}\right). \end{aligned} \tag{48}$$

Because of the variable behavior of E^{L} in x, all of the terms in (48) cannot be simultaneously zero; their positive behavior thus shows the positive definite nature of m. But whether the mass is increased or decreased from the general relativity value is not *a priori* arguable.

According to the definition of the mass integral, its first variation in ϕ about ϕ equal to a constant (say $\bar{\phi}$) is

$$\delta m = -8\int d^3x\nabla^2\delta\chi, \tag{49}$$

where we assume variations commute with derivatives and the constraint is maintained throughout the variation. Thus, from (37),

$$\begin{aligned} -8\nabla^2\delta\chi &= (\tfrac{3}{2}+\omega)\left(\frac{\nabla\phi}{\phi}\right)^2\delta\chi + \chi\delta\left(\frac{\nabla\phi}{\phi}\right)^2 - \frac{3(E^{\mathrm{L}})^2}{2\chi^4}\delta\chi \\ &\quad + \delta^3(x)\left[-\left(\frac{m_0}{\phi^{\frac{1}{2}}}\right)\frac{\delta\chi}{\chi^2} + \frac{1}{\chi}\delta\left(\frac{m_0}{\phi^{\frac{1}{2}}}\right)\right], \end{aligned} \tag{50}$$

where this equation is to be taken at $\phi = \bar{\phi}$ everywhere. At $\phi = \bar{\phi}$, Eq. (50) becomes

$$\begin{aligned} -8\nabla^2\delta\chi &= -\frac{3(E^{\mathrm{L}})^2}{2\chi^4}\delta\chi \\ &\quad + \delta^3(x)\left[-\left(\frac{m_0}{\phi^{\frac{1}{2}}}\right)\frac{\delta\chi}{\chi^2} + \frac{1}{\chi}\delta\left(\frac{m_0}{\phi^{\frac{1}{2}}}\right)\right]_{\phi=\bar{\phi}}. \end{aligned} \tag{51}$$

χ appearing in (51) is now the function obtained by solving the original constraint equation (37) for the condition $\phi = \bar{\phi}$

$$-8\chi\nabla^2\chi = \frac{(E^{\mathrm{L}})^2}{2\chi^2} + \frac{m_0\delta^3(x)}{\phi^{\frac{1}{2}}}. \tag{52}$$

A comparison of this equation with the Einstein constraint (46) shows that the equations are identical but for a change in the bare mass.

We consider, for the moment, the Einstein constraint alone. The mass of the electron is $m_{\mathrm{E}} = 2|e|$ where $m_{\mathrm{E}} = -8\int d^3x\nabla^2\chi_{\mathrm{E}}$ with χ_{E} given by (47). By how much does the mass change if the bare mass is altered slightly? Formally, the change in m is $\delta m_{\mathrm{E}} = -8\int d^3x\nabla^2\delta\chi_{\mathrm{E}}$ where $\delta\chi_{\mathrm{E}}$ is determined by a variation of (46),

$$\begin{aligned} -8\nabla^2\delta\chi_{\mathrm{E}} &= -\frac{3(E^{\mathrm{L}})^2}{2\chi_{\mathrm{E}}^4}\delta\chi_{\mathrm{E}} \\ &\quad + \delta^3(x)\left(-m_0\frac{\delta\chi_{\mathrm{E}}}{\chi_{\mathrm{E}}^2} + \frac{1}{\chi_{\mathrm{E}}}\delta m_0\right). \end{aligned} \tag{53}$$

But, because the mass is independent of the bare mass, the solution of (53) must be such that δm_{E} is zero.

The problem posed by Eqs. (49) and (51) is identical to that of the previous paragraph except that the bare mass is replaced by $m_0/\bar{\phi}^{\frac{1}{2}}$. The solution must then be that δm vanishes, for ϕ a constant, in the scalar-tensor theory. The mass is extremized at this point; it remains to be shown that it is also minimized.

The second variation of the mass integral is given by

$$\delta^2 m = -8\int d^3x\nabla^2\delta^2\chi, \tag{54}$$

and $\delta^2\chi$ can be obtained from

$$\left(8\nabla^2 - \frac{3(E^{\mathrm{L}})^2}{2\chi^4}\right)\delta^2\chi = f_1 + f_2 + f_3\delta^3(x), \tag{55}$$

where

$$f_1 = -\frac{6(E^L)^2}{\chi^5}\delta\chi^2, \tag{56a}$$

$$f_2 = -(3+2\omega)\chi(\delta\nabla \ln \phi)^2, \tag{56b}$$

$$f_3 = \frac{m_0}{\bar{\phi}^{\frac{1}{2}}\chi}\Big[-2\Big(\frac{\delta\chi}{\chi}\Big)^2 - \frac{\delta\chi}{\chi}\frac{\delta\phi}{\phi} - \frac{3}{4}\Big(\frac{\delta\phi}{\phi}\Big)^2 + \frac{\delta^2\chi}{\chi}\Big]_{\phi=\bar{\phi}}. \tag{56c}$$

It is assumed that every term is evaluated at $\phi = \bar{\phi}$; χ in the preceding equations is thus given by Eq. (47). Note that f_1 and f_2 are less than zero.

It is shown in the Appendix that a Green's function corresponding to the operator $8\nabla^2 - 3(E^L)^2/2\chi^4$ can be obtained to invert (55):

$$\delta^2\chi(x) = \int d^3x' G(x', x)[f_1(x') + f_2(x') + f_3(x')\delta^3(x')]. \tag{57}$$

Because we are only interested in the sign of $\delta^2 m$, we can investigate

$$\delta^2 m = -8\iint d^3x\, d^3x' \nabla_x^2 G(x', x) \times [f_1(x') + f_2(x') + f_3(x')\delta^3(x')]. \tag{58}$$

Furthermore, it is shown in the Appendix that

$$\delta^2 m = \tfrac{8}{5}\int d^3x \Big(1 + \frac{|e|}{8\pi|x|}\Big)^{-2}[-f_1(x) - f_2(x)], \tag{59}$$

so that it is evident that $\delta^2 m$ is greater than zero. Consequently, ϕ being a constant is a minimum energy configuration for the electron. The static point charge solution is thus the same as in general relativity and the inertia of the electron is unchanged from the general relativity result $m = 2|e|$.

CONCLUSIONS

It was pointed out earlier that there are a variety of representations of the scalar-tensor theory; in particular, there are two representations in which the one permits a variable behavior to appear in the gravitational constant G and the other permits a variable behavior to appear in the mass of a fundamental particle such as an electron. The one in which the latter occurs is distinguished by a choice of the unit of mass $(\hbar c/G)^{\frac{1}{2}}$. It is this representation which most clearly enables us to investigate the inertial implications of Mach's principle.

The self-energy calculation was made for both a neutral particle and an electron to see how the self-energy depends upon the boundary value of the scalar field. By extremization arguments, it was concluded that ϕ, equal to a constant everywhere outside the particle, is the minimum energy configuration, that as a consequence the self-energy of either the neutral particle or the electron is unchanged from the corresponding general relativity value. Justification for the boundary conditions of Minkowskian space and a constant scalar field "at infinity" is obtained from the interior solution of a large matter shell in which the space is, in fact, Minkowskian and the scalar field constant. Finally, justification for the interpretation of the self-energy as the inertia of the system arises from the boundary condition of Minkowski space; an electron with its electromagnetic field and its gravity field, when treated as a composite object, will respond to an external field with an inertia corresponding to its self-energy (up to a factor of c^2) because the self-energy is equal to the magnitude of the 4-momentum describing the system.

The conclusion must be, then, that the scalar-tensor theory affords no substantial improvement over general relativity in the compatibility with Mach's principle.

Just as the above particle models, in a sense, force the scalar-tensor theory to reduce to general relativity, the problem of the rotating matter shell and its effect on the interior inertial frame has similar results. Brill[11] has shown that the interior inertial frame is "dragged" along by the rotating shell in the scalar-tensor theory as in general relativity. And, for the limit in which the density of matter in the universe approaches zero, the "dragging" is perfect. This is just the condition under which the radius of the shell equals its Schwarzschild radius, the same condition for perfect "dragging" by the shell in general relativity. Although the limit procedures to obtain this model are different, it must be concluded that here, too, the scalar-tensor theory reduces to general relativity.

The objection may arise that the solution of a constant scalar field does not represent a minimum energy configuration since it is hardly a static solution of $\Box\phi \sim T$, where T is not identically zero. In Paper II exact solutions of a charged spherically symmetric mass distribution will be given which show that the scalar field can be zero everywhere outside the particles, but that an essential singularity appears at the particle to maintain $\Box\phi \sim T$.

ACKNOWLEDGMENT

I wish to express my sincerest gratitude to Professor J. Weber for his many stimulating suggestions and fruitful conversations on Mach's principle and the scalar-tensor theory.

APPENDIX

In this appendix we discuss the Green's function related to the operator $8\nabla^2 - 3(E^L)^2/2\chi^4$, where E^L and χ are the rigorous solutions to the self-energy problem of the electron in general relativity. Reference to Eqs. (30a) and (47) shows that these quantities are given by

$$(E^L)^2 = \alpha^2/r^4 \tag{A1}$$

and

$$\chi^2 = 1 + \alpha/2r, \tag{A2}$$

where

$$\alpha = |e|/4\pi.$$

If we let $\omega^2(|x|) = 3(E^L)^2/2\chi^4$, then the Green's function is defined as the solution to the differential equation

$$[8\nabla_x^2 - \omega^2(|x|)]G(x, x') = \delta^3(x - x'). \tag{A3}$$

The equation we wish to integrate is [see Eq. (55)]

$$[8\nabla_x^2 - \omega^2(|x|)]\delta^2\chi = f_1 + f_2 + f_3\delta^3(x) \equiv F(x). \tag{A4}$$

If Eq. (A3) is multiplied by $\delta^2\chi$ and Eq. (A4) by $G(x, x')$ and the equations are subtracted, the following equation arises:

$$8G(x, x')\nabla_x^2\delta^2\chi - 8\delta^2\chi\nabla_x^2G(x, x') = F(x)G(x, x') - \delta^2\chi(x)\delta^3(x - x'). \tag{A5}$$

An integration on the variable x over all space will yield, by virtue of Gauss' theorem and because of the δ function,

$$\delta^2\chi(x') = \int d^3xF(x)G(x, x') + 8\int d\mathbf{s}\cdot[\delta^2\chi\nabla_xG(x, x') - G(x, x')\nabla_x\delta^2\chi], \tag{A6}$$

where the surface integral is taken at infinity. Because all variations vanish at infinity, the first term in the integrand of the surface integral is zero; an appropriate choice of boundary condition on G, then, is that $G(x, x')$ vanish in the limit $|x| \to \infty$. With this choice Eq. (A6) becomes

$$\delta^2\chi(x) = \int d^3x'F(x')G(x', x) \tag{A7}$$

by an interchange of labels.

Let us now consider an expansion of $G(x, x')$ in spherical harmonics. G is to be the solution of Eq. (A3). An expansion of the δ function is given by

$$\delta^3(x - x') = \frac{1}{r^2}\delta(r - r')\delta(\phi - \phi')\delta(\cos\theta - \cos\theta') = \frac{1}{r^2}\delta(r - r')\sum_{l=0}^{\infty}\sum_{m=-l}^{l}Y_{lm}^*(\theta', \phi')Y_{lm}(\theta, \phi). \tag{A8}$$

Because of this special angular behavior in (θ, ϕ) and (θ', ϕ'), it is clear that G is expanded in the form

$$G(x, x') = \sum_{l=0}^{\infty}\sum_{m=l}^{l}g_{ml}(r, r')Y_{lm}^*(\theta', \phi')Y_{lm}(\theta, \phi). \tag{A9}$$

Now the Laplacian can be written as

$$\nabla_x^2 = R^2 - \frac{L^2}{r^2},$$

where

$$R^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)$$

and

$$L^2 = -\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) - \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}.$$

More importantly, L^2 has the property that $L^2Y_{lm} = l(l + 1)Y_{lm}$. A substitution of these expressions into the differential equation (A3) thus leads to

$$\left(R^2 - \frac{l(l + 1)}{r^2} - \omega^2(r)\right)g_{ml}(r, r') = \frac{1}{r^2}\delta(r - r'), \tag{A10}$$

because of the orthogonality of the Y_{lm}. Once the g_{ml} are determined from this equation, (A9) represents the solution to be used in (A7).

For our purposes, however, knowledge of $\delta^2\chi$ is not as important as knowledge of the sign of δ^2m. The latter is given by

$$\delta^2m = -8\int d^3x\nabla^2\delta^2\chi. \tag{A11}$$

According to (A7), it is true that

$$\nabla^2\delta^2\chi = \int d^3x'F(x')\nabla_x^2G(x', x) \tag{A12}$$

and, thus,

$$\delta^2m = -8\int d^3x\, d^3x'F(x')\nabla_x^2G(x', x). \tag{A13}$$

Note that the integrand depends on x only through $G(x', x)$. If the order of integrations are interchanged, the integration on x can readily be performed. From

the expression for $G(x, x')$ given by (A9), we find that

$$\nabla_x^2 G(x', x) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} Y_{lm}^*(\theta, \phi) Y_{lm}(\theta', \phi') \left(R - \frac{l(l+1)}{r^2} \right) g_{ml}(r', r). \tag{A14}$$

Writing d^3x as $r^2\, dr\, d\Omega$ and noting that

$$\int d\Omega Y_{lm}^*(\theta, \phi) = (4\pi)^{\frac{1}{2}} \delta_{l0} \delta_{m0},$$

we readily see that $\delta^2 m$ is given by

$$\delta^2 m = -8 \int d^3x' F(x') \int r^2\, dr R^2 g_{00}(r', r). \tag{A15}$$

This is because the summation only contributes at $l = 0$ and $m = 0$ and because $Y_{00}(\theta, \phi) = (4\pi)^{-\frac{1}{2}}$.

Finally, because of the form of R^2, $dr r^2 R^2$ is an exact differential; $\delta^2 m$ can thus be written

$$\delta^2 m = -8 \int d^3x' F(x') \left(r^2 \frac{dg_{00}(r', r)}{dr} \right)_0^{\infty}. \tag{A16}$$

It is clear, then, that we need only the quantity g_{00} from Eq. (A10). For $l = 0$, Eq. (A10) becomes

$$\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{dg_{00}}{dr} \right) - \frac{3\alpha^2 g_{00}}{2r^4(1 + \alpha/2r)^2} = \frac{1}{r^2} (r - r'), \tag{A17}$$

where this equation generates $g_{00}(r, r')$. A more convenient form is found by multiplying (A17) by r^4. For, then,

$$r^2 \frac{d}{dr} \left(r^2 \frac{d}{dr} g_{00} \right) - \frac{3\alpha^2}{2(1 + \alpha/2r)} g_{00} = r^2 \delta(r - r'). \tag{A18}$$

Equation (A18) is in a form amenable to the substitution $z = 1 + \alpha/2r$ since

$$\frac{d}{dr} = \frac{dz}{dr} \frac{d}{dz} = -\frac{\alpha}{2r^2} \frac{d}{dz}.$$

Thus,

$$r^2 \frac{d}{dr} = -\tfrac{1}{2}\alpha \frac{d}{dz}$$

and (A18) becomes, away from $r = r'$,

$$\tfrac{1}{4}\alpha^2 \frac{d^2 g_{00}}{dz^2} - \frac{3\alpha^2 g_{00}}{2z^2} = 0. \tag{A19}$$

If it is assumed that g_{00} is of the form $g_{00} = z^\sigma$, then (19) is satisfied if σ obeys

$$\sigma(\sigma - 1) - 6 = (\sigma - 3)(\sigma + 2) = 0. \tag{A20}$$

Clearly, $\sigma = 3, -2$. In terms of r, then, the general solution of (18) away from $r = r'$ is of the form

$$g_{00} = C_1(1 + \alpha/2r)^3 + C_2(1 + \alpha/2r)^{-2}. \tag{A21}$$

Several conditions are yet to be imposed on g_{00}. The Green's function must be continuous everywhere, it must reduce to zero as $r \to \infty$, and it must satisfy its differential equation everywhere. The condition that $G(x, x') \to 0$ as $r \to \infty$ demands that g_{00} be written, for $r > r'$, as

$$g_{00}(r, r') = A(1 + \alpha/2r)^3 - A(1 + \alpha/2r)^{-2}, \quad r > r'. \tag{A22}$$

For $r < r'$, we can write

$$g_{00}(r, r') = B(1 + \alpha/2r)^{-2}, \quad r < r', \tag{A23}$$

if we ask that g be well behaved at the origin.

The continuity of G demands from g_{00} that $g_{00}(r \to r', r')$ for $r > r'$ equals $g_{00}(r \to r', r')$ for $r < r'$. This implies from (A22) and (A23) that

$$A(1 + \alpha/2r')^3 - A(1 + \alpha/2r')^{-2} = B(1 + \alpha/2r')^{-2}. \tag{A24}$$

The final condition on g_{00} is that it satisfy the differential equation (A18) everywhere. If we divide (A18) by r^2 and then integrate across $r = r'$, we find that

$$r^2 \frac{dg_{00}(r > r')}{dr}\bigg|_{r'} - r^2 \frac{dg_{00}(r < r')}{dr}\bigg|_{r'} = 1. \tag{A25}$$

The second term in (A18) does not contribute because of continuity. A substitution of the expressions for g_{00} given by (A22) and (A23) thus gives

$$-\tfrac{3}{2}\alpha A(1 + \alpha/2r')^2 - \alpha A(1 + \alpha/2r')^{-3} - \alpha B(1 + \alpha/2r')^{-3} = 1. \tag{A26}$$

Equations (A26) and (A24) imply that A and B are given by

$$A = -(2/5\alpha)(1 + \alpha/2r')^{-2}, \tag{A27a}$$

$$B = -(2/5\alpha)[(1 + \alpha/2r')^3 - (1 + \alpha/2r')^{-2}], \tag{A27b}$$

and so $g_{00}(r, r')$ is given by

$$g_{00}(r, r') = -\frac{2}{5\alpha} \left(1 + \frac{\alpha}{2r'} \right)^{-2} \left[\left(1 + \frac{\alpha}{2r} \right)^3 - \left(1 + \frac{\alpha}{2r} \right)^{-2} \right], \quad r > r', \tag{A28a}$$

$$g_{00}(r, r') = -\frac{2}{5\alpha} \left(1 + \frac{\alpha}{2r} \right)^{-2} \left[\left(1 + \frac{\alpha}{2r'} \right)^3 - \left(1 + \frac{\alpha}{2r'} \right)^{-2} \right], \quad r < r'. \tag{A28b}$$

Note that $g_{00}(r, r') = g_{00}(r', r)$.

Again referring to Eq. (A16), we see that the sign of $\delta^2 m$ depends on evaluating the derivative of $g_{00}(r', r)$.

For fixed r' and $r > r'$,

$$r^2 \frac{dg_{00}}{dr} = \tfrac{1}{5}\left(1 + \frac{\alpha}{2r'}\right)^{-2}\left[3\left(1 + \frac{\alpha}{2r}\right)^{2} - 2\left(1 + \frac{\alpha}{2r}\right)^{-3}\right], \tag{A29}$$

and so

$$r^2 \frac{dg_{00}}{dr}\bigg|_{r\to\infty} = \tfrac{1}{5}\left(1 + \frac{\alpha}{2r'}\right)^{-2}. \tag{A30}$$

For fixed r' and $r < r'$,

$$r^2 \frac{dg_{00}}{dr} = -\tfrac{2}{5}\left(1 + \frac{\alpha}{2r}\right)^{-3}\left[\left(1 + \frac{\alpha}{2r'}\right)^{3} - \left(1 + \frac{\alpha}{2r'}\right)^{-2}\right], \tag{A31}$$

and so

$$r^2 \frac{dg_{00}}{dr}\bigg|_{r\to 0} = 0. \tag{A32}$$

Because of these equations, (A16) becomes

$$\delta^2 m = -\tfrac{8}{5}\int d^3x' F(x')\left(1 + \frac{\alpha}{2\,|x'|}\right)^{-2}. \tag{A33}$$

As pointed out in Eq. (A4), $F(x')$ is written as

$$F(x') = f_1(x') + f_2(x') + f_3(x')\delta^3(x'),$$

where f_1 and f_2 are less than zero. We thus write $\delta^2 m$ as

$$\delta^2 m = \tfrac{8}{5}\int d^3x'\left(1 + \frac{\alpha}{2\,|x'|}\right)^{-2}[-f_1(x') - f_2(x')] \tag{A34}$$

because $(1 + \alpha/2\,|x'|)^{-2} \to 0$ as $|x'| \to 0$ and because $f_3(0)$ is zero. Thus, $\delta^2 m > 0$. Accordingly, the mass of the point electron is minimized about $\phi = \text{const}$.

* This work is supported in part by the National Science Foundation and NASA Grant NGR 21-002-010.

† Based on the author's thesis: E. Toton, Ph.D. thesis, University of Maryland, 1969 (unpublished).

‡ Present address: Institute for Advanced Study, Princeton, New Jersey 08540.

[1] D. W. Sciama, *The Unity of the Universe* (Doubleday and Co., Garden City, N.Y., 1959).

[2] H. Y. Chiu and W. F. Hoffman, *Gravitation and Relativity* (W. A. Benjamin, New York, 1964).

[3] R. H. Dicke, *Proceedings of the International School of Physics "Enricho Fermi,"* Course XX (Academic, New York, 1961).

[4] D. Brill and J. Cohen, Phys. Rev. **143**, 1011 (1966).

[5] R. H. Dicke and C. Brans, Phys. Rev. **124**, 925 (1961).

[6] R. H. Dicke, Phys. Rev. **125**, 2163 (1962).

[7] The left superscript 4 is used to designate the specifically 4-dimensional character of the quantities so marked.

[8] L. Witten, *Gravitation: An Introduction to Current Research* (John Wiley & Sons, New York, 1962).

[9] R. Arnowitt, S. Deser, and C. W. Misner, Phys. Rev. **120**, 313 (1960).

[10] A constant choice of units is made so that $16\pi G_0/c^4 = 1$.

[11] D. Brill, Z. Naturforsch. **22a**, 1336 (1967).

Surface Layer in General Relativity and Integral Form of Field Law

AMRITAVA GUPTA
Department of Applied Mathematics, Calcutta University, Calcutta, India

(Received 22 June 1967; Second Copy Received 4 August 1969)

Assuming that the metric tensor g_{ij} depends on a continuous parameter ϵ, we write Einstein's field equations of general relativity in the form of a divergence equation. The theory of surface layer in general relativity is then worked out. Finally, an integral flux law is proposed from which both the field equations and the boundary conditions for a surface layer can be deduced.

INTRODUCTION

In this paper we assume that the metric tensor g_{ij} depends on a continuous parameter ϵ; that is, instead of a single metric tensor we consider a family of metric tensors. Under this hypothesis we first write Einstein's field equations of general relativity in the form of a divergence equation. The theory of surface layers in general relativity is then worked out, which consists principally of the generalization of the classical boundary condition, viz., that the jump of the normal derivative of the potential is proportional to the surface density of matter together with the surface analog of the Bianchi identities. All these are found to give correct classical approximations. Finally, an integral flux law is proposed from which both the field equations and the boundary conditions for a surface layer can be deduced.

1. THE PARAMETER ϵ; DIVERGENCE FORM OF FIELD LAW

Postulate 1.1: We assume that the metric tensor g_{ij} depends on a real parameter ϵ such that $g_{ij} = g_{ij}(\epsilon)$ are analytic functions of ϵ (i.e., representable in power series of ϵ) in some neighborhood of $\epsilon = 0$, and for $\epsilon = 0$, $g_{ij} = {}^0g_{ij}$, where ${}^0g_{ij}$ denotes the Euclidean metric tensor in any coordinate system.

Notation: In the following discussions, an overhead bar will indicate differentiation with respect to ϵ and a left superscript zero the value of a quantity for $\epsilon = 0$. As usual, a comma in the subscript will denote partial differentiation, and a semicolon in the subscript will denote covariant differentiation.

Remark: $\bar{\Gamma}^k_{ij}$ is a mixed tensor of rank three.

The Ricci tensor is given by

$$R_{ij} = \Gamma^\alpha_{i\alpha,j} - \Gamma^\alpha_{ij,\alpha} + \Gamma^\alpha_{\beta j}\Gamma^\beta_{i\alpha} - \Gamma^\alpha_{\beta\alpha}\Gamma^\beta_{ij}. \quad (1.1)$$

The terms on the rhs containing the second derivatives only are

$$\Gamma^\alpha_{i\alpha,j} - \Gamma^\alpha_{ij,\alpha} = \delta^k_j\Gamma^\alpha_{i\alpha,k} - \Gamma^k_{ij,k},$$

which is in the form of an ordinary divergence. We note that $\delta^k_j\Gamma^\alpha_{i\alpha} - \Gamma^k_{ij}$ is not a tensor, but that $\delta^k_j\Gamma^\alpha_{i\alpha} - \bar{\Gamma}^k_{ij}$ is a tensor.

Definition 1.1:

$$P^k_{ij} = \delta^k_j\Gamma^\alpha_{i\alpha} - \bar{\Gamma}^k_{ij}.$$

The tensor P^k_{ij} has the following remarkable property:

$$P^k_{ij;k} = \bar{R}_{ij}. \quad (1.2)$$

P^k_{ij} is unsymmetric in i, j, and we define Q^k_{ij} as follows.

Definition 1.2:

$$Q^k_{ij} = \tfrac{1}{2}(P^k_{ij} + P^k_{ji}),$$

which is symmetric in i, j.

From (1.2) we have

$$Q^k_{ij;k} = \bar{R}_{ij}. \quad (1.3)$$

Definition 1.3: Let T_{ij} denote the energy-momentum tensor. Then

$$T^*_{ij} = T_{ij} - \tfrac{1}{2}g_{ij}T$$

will be called the modified energy–momentum tensor.

The field equations of general relativity are known to be

$$R_{ij} = -\kappa T^*_{ij}, \quad \kappa = \text{const.} \quad (1.4)$$

Postulate 1.2: For $\epsilon = 0$, $T^*_{ij} = {}^0T^*_{ij} = 0$.

Remark: If we neglect pressure or stresses, $T_{ij} = \rho_0 u_i u_j$, where $u^i = dx^i/ds$ and ρ_0 is the proper mass density. Then

$$T = \rho_0 = -T^*, \quad T^*_{ij} = \rho_0(u_i u_j - \tfrac{1}{2}g_{ij})$$

so that Postulate 1.2 amounts to saying that $\rho_0 = \rho_0(\epsilon)$ is such that $\rho_0(0) = 0$.

We note that for $\epsilon = 0$, $R_{ij} = {}^0R_{ij} = 0$, and, by virtue of Postulate 1.2, (1.4) is equivalent to

$$\bar{R}_{ij} = -\kappa \bar{T}^*_{ij}$$

or by (1.3) to

$$Q^k_{ij;k} = -\kappa \bar{T}^*_{ij}, \tag{1.5}$$

which is the required divergence form of the field equations.

2. SURFACE LAYER: CONSERVATION IDENTITY AND BOUNDARY CONDITION

We assume that there exists a coordinate system in which the Euclidean metric tensor

$${}^0g_{ij} = \text{diag}\,(-1, -1, -1, 1).$$

Consider a smooth hypersurface

$$S\colon f(x^1, x^2, x^3, x^4) = 0 \tag{2.1}$$

(in any coordinate system), where the function f is independent of ϵ and the unit vector normal to S is spacelike, i.e.,

$$g^{\alpha\beta} n_\alpha n_\beta = -1 \tag{2.2}$$

so that

$$n_i = f_{,i}/(-g^{\alpha\beta} f_{,\alpha} f_{,\beta})^{\frac{1}{2}}. \tag{2.3}$$

Definition 2.1: A hypersurface S is said to be parametrically stationary if the covariant components of the unit normal vector n_i are independent of ϵ, i.e., $\tilde{n}_i = 0$.

By (2.3) S will be parametrically stationary if $g_{\alpha\beta} f_{,\alpha} f_{,\beta}$ is independent of ϵ. In particular, the hypersurface $x^1 = 0$ will be parametrically stationary if g^{11} is independent of ϵ, i.e., $g^{11} = {}^0g^{11}$; in fact, we can always make a coordinate transformation $(x^i) \to (y^i)$ in which $y^1 = f(x^1, x^2, x^3, x^4)$.

Postulate 2.1: A hypersurface S represents a surface layer if and only if it is parametrically stationary.

Postulate 2.2: The energy-momentum tensor S_{ij} for a surface layer will be given by

$$S_{ij} = \sigma_0 u_i u_j,$$

where σ_0 is the proper surface density of mass, the pressure or stresses being neglected. Further, $\sigma_0 = \sigma_0(\epsilon)$ is such that $\sigma_0(0) = 0$.

Definition 2.2: The modified energy-momentum tensor for a surface layer is given by

$$S^*_{ij} = S_{ij} - \tfrac{1}{2} g_{ij} S.$$

By Postulate 2.2,

$$S = \sigma_0, \quad S^*_{ij} = \sigma_0(u_i u_j - \tfrac{1}{2} g_{ij}). \tag{2.4}$$

It follows that for $\epsilon = 0$,

$$S^*_{ij} = {}^0S^*_{ij} = 0. \tag{2.5}$$

In the following discussions the symbol Δ will denote the jump of a quantity across a given hypersurface.

Remark: $\Delta\Gamma^k_{ij}$ is a tensor.

Definition 2.3:

$$H^k_{ij} = \delta^k_j \Delta\Gamma^\alpha_{i\alpha} - \Delta\Gamma^k_{ij}.$$

The tensor H^k_{ij} satisfies the identity

$$H^k_{kj} = 0. \tag{2.6}$$

We assert that at a surface layer this identity would express the conservation law or the physical boundary condition and as such is the surface analog of the Bianchi identities. Noting that H^k_{ij} is unsymmetric in i, j, we make the following postulate.

Postulate 2.3: g_{ij} are continuous throughout 4-space and, at a surface layer S,

$$\frac{\partial}{\partial\epsilon}(\sigma_0^{-1} H^k_{ij}) = \kappa\left(2u_i \bar{u}_j n^k - \frac{1}{2}\frac{\partial}{\partial\epsilon}(g_{ij} n^k)\right).$$

Therefore,

$$\frac{\partial}{\partial\epsilon}(\sigma_0^{-1} H^k_{kj}) = \kappa(2u_k \bar{u}_j n^k - \tfrac{1}{2}\bar{n}_j) = 2\kappa \bar{u}_j u_k n^k,$$

since n_i are independent of ϵ. Hence the identity (2.6) leads to the condition at S

$$u_k n^k = 0 = u^k n_k. \tag{2.7}$$

Interchanging i, j in Postulate 2.3 and adding the result to the same postulate, we obtain

$$\frac{\partial}{\partial\epsilon}[\tfrac{1}{2}\sigma_0^{-1}(H^k_{ij} + H^k_{ji})] = \kappa\left(\frac{\partial}{\partial\epsilon}(u_i u_j) n^k - \frac{1}{2}\frac{\partial}{\partial\epsilon}(g_{ij} n^k)\right).$$

Multiplying by n_k, which are independent of ϵ, we have

$$\frac{\partial}{\partial\epsilon}[\tfrac{1}{2}\sigma_0^{-1}(H^k_{ij} + H^k_{ji}) n_k] = -\kappa\left(\frac{\partial}{\partial\epsilon}(u_i u_j) - \tfrac{1}{2}\bar{g}_{ij}\right)$$

$$= -\kappa \frac{\partial}{\partial\epsilon}(\sigma_0^{-1} S^*_{ij})$$

or

$$\frac{\partial}{\partial\epsilon}[\sigma_0^{-1}(\tfrac{1}{2} n_j \Delta\Gamma^\alpha_{i\alpha} + \tfrac{1}{2} n_i \Delta\Gamma^\alpha_{j\alpha} - n_k \Delta\Gamma^k_{ij} + \kappa S^*_{ij})] = 0,$$

which are satisfied if

$$\tfrac{1}{2} n_j \Delta\Gamma^\alpha_{i\alpha} + \tfrac{1}{2} n_i \Delta\Gamma^\alpha_{j\alpha} - n_k \Delta\Gamma^k_{ij} = -\kappa S^*_{ij}. \tag{2.8}$$

These are the required boundary conditions at a surface layer S. We note that (2.8) is free from derivatives with respect to ϵ and as such provides a satisfactory criterion.

Since $\Delta^0\Gamma^k_{ij} = 0$, ${}^0S^*_{ij} = 0$, and n_i are independent of ϵ, Eq. (2.8) is equivalent to

$$n_k \Delta Q^k_{ij} = -\kappa S^*_{ij}. \tag{2.9}$$

3. CLASSICAL APPROXIMATIONS

In the linear approximation we assume

$$g_{ij} \simeq {}^0g_{ij} + \epsilon h_{ij}, \tag{3.1}$$

where ϵ is a small quantity whose square is negligible, and we take that coordinate system in which

$${}^0g_{ij} = {}^0g^{ij} = \text{diag}\,(-1, -1, -1, 1). \tag{3.2}$$

We set

$$h^{ij} = {}^0g^{i\alpha}\,{}^0g^{j\beta}h_{\alpha\beta}, \quad h^i_j = {}^0g^{i\alpha}h_{\alpha j}, \quad h = {}^0g^{\alpha\beta}h_{\alpha\beta}, \tag{3.3}$$

$$g = \det(g_{ij}) \simeq -1 - \epsilon h, \tag{3.4}$$

$$g^{ij} \simeq {}^0g^{ij} - \epsilon h^{ij}. \tag{3.5}$$

If $(ij, k)_h$ denote the Christoffel symbols of the first kind by treating h_{ij} as the metric tensor, then

$$\Gamma^k_{ij} \simeq \epsilon\, {}^0g^{k\alpha}(ij, \alpha)_h, \tag{3.6}$$

$$R_{ij} \simeq \tfrac{1}{2}\epsilon({}^0g^{\alpha\beta}h_{ij,\alpha\beta} + h_{,ij} - h^\alpha_{i,\alpha j} - h^\alpha_{j,\alpha i}). \tag{3.7}$$

For convenience, let us write $(x^i) = (x, y, z, t)$. In the classical approximation, in addition to the linear approximation, we have to make two other approximations, viz., (i) the quasisteady approximation in which the derivatives with respect to t are treated as small compared with the derivatives with respect to x, y, z and (ii) the approximation that the velocity of a particle is small compared with the velocity of light which is taken to be unity, so that we may take u^1, u^2, u^3 to be small quantities and $u^4 \simeq 1$.

Neglecting pressure or stresses which are usually small compared with the matter-density, we have $T^{ij} = \rho_0 u^i u^j$. Since $\rho_0 = 0$ for $\epsilon = 0$, $\rho_0 = O(\epsilon)$. The only component of T^{ij} of order ϵ is $T^{44} \simeq \rho_0$; all other components are negligible. Hence, $T = \rho_0$, $T_{44} \simeq \rho_0$, and

$$T^*_{44} \simeq T_{44} - \tfrac{1}{2}\,{}^0g_{44}T \simeq \tfrac{1}{2}\rho_0.$$

From (3.7)

$$R_{44} \simeq -\tfrac{1}{2}\epsilon\nabla^2 h_{44}.$$

Hence, the equation $R_{44} = -\kappa T^*_{44}$ approximates to $\nabla^2\phi = \frac{1}{2}\kappa\rho_0$ if

$$\phi = \tfrac{1}{2}\epsilon h_{44}. \tag{3.8}$$

This equation reduces to the classical Poisson equation if

$$\kappa = 8\pi G, \tag{3.9}$$

G being the Newtonian constant of gravitation. This deduction, however, is well known.

In the case of a surface layer we similarly have $S^*_{44} \simeq \frac{1}{2}\sigma_0$. In the quasisteady approximation f_t will be taken to be small compared with the space derivatives f_x, f_y, f_z. By (2.3)

$$n_i \simeq f_{,i}/(f_x^2 + f_y^2 + f_z^2)^{\frac{1}{2}}. \tag{3.10}$$

If $\mathbf{n} = (n_x, n_y, n_z)$ denotes unit vector normal to the spatial surface $f(x, y, z, t) = 0$ at time t, then $\mathbf{n} \simeq (n_1, n_2, n_3)$ and n_4 is small compared with n_1, n_2, n_3. For $k = 1, 2, 3$,

$$\Gamma^k_{44} \simeq \tfrac{1}{2}\epsilon h_{44,k} = \phi_{,k}.$$

Hence, the 44-component of (2.8) reduces to the approximate equation

$$\Delta(n_x\phi_x + n_y\phi_y + n_z\phi_z) = \tfrac{1}{2}\kappa\sigma_0$$

or

$$\Delta\left(\frac{\partial\phi}{\partial n}\right) = 4\pi G\sigma_0,$$

which is the classical boundary condition at a surface layer.

We set

$$\mathbf{q} = (q_x, q_y, q_z) = \left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}\right),$$

$$\mathbf{q} \simeq (u^1, u^2, u^3), \quad u^4 \simeq 1.$$

The condition (2.7) approximates to

$$q_x f_x + q_y f_y + q_z f_z + f_t = 0,$$

which is the well-known classical condition that the equation of the surface layer at any time is $f(x, y, z, t) = 0$. We note that q_x, q_y, q_z, and f_t are small quantities.

4. INTEGRAL FIELD LAW

The field equations (1.5) and the boundary conditions (2.8) can be simultaneously deduced from an integral flux law.

Definition 4.1: Let V be a region of 4-space. Then the integral

$$-\int_V \bar{T}^*_{ij}\, dV,$$

where dV denotes the invariant 4-volume element, will be called the parametric rate of decrease or simply the parametric decay of energy-momentum in the region V.

Similarly for a surface layer S, the parametric decay of energy-momentum in the layer S is defined

to be

$$-\int_S \bar{S}_{ij}^* \, dS,$$

where dS is the invariant hypersurface element.

Postulate 4.1 (Integral Field Law): Let Σ be any closed hypersurface given by $F(x^1, x^2, x^3, x^4) = 0$ where the function F does not involve the parameter ϵ and N_i, the unit normal to Σ, is spacelike, i.e., $g^{\alpha\beta}N_\alpha N_\beta = -1$. Let V denote the region of 4-space bounded by Σ and N_i point out of V across Σ. We postulate that

$$\int_\Sigma Q_{ij}^k N_k \, d\Sigma$$

($d\Sigma$-invariant hypersurface element), which represents the flux of the tensor Q_{ij}^k across Σ, is κ times the parametric decay of energy-momentum in V.

Thus, for any region V,

$$\int_\Sigma Q_{ij}^k N_k \, d\Sigma = -\kappa \int_V \bar{T}_{ij}^* \, dV.$$

By Gauss's theorem we get

$$\int_V (Q_{ij;k}^k + \kappa \bar{T}_{ij}^*) \, dV = 0,$$

which gives (1.5).

We take a fixed point on a surface layer S, at which the unit normal is n_i, and take Σ to be a small closed cylinder across S, enclosing the point whose generators are parallel to n_i. Applying the integral law to Σ, we get

$$n_k \Delta Q_{ij}^k \, dS = -\kappa \bar{S}_{ij}^* \, dS,$$

which gives (2.9).

If S is any hypersurface which is not a surface layer (and, hence, S is not necessarily parametrically stationary), then the integral law gives the following boundary condition at S:

$$n_k \Delta Q_{ij}^k = 0$$

or, by (2.3),

$$f_{,k} \Delta Q_{ij}^k = 0.$$

Integrating with respect to ϵ from 0 to ϵ, we have

$$\tfrac{1}{2} f_{,j} \Delta \Gamma_{i\alpha}^\alpha + \tfrac{1}{2} f_{,i} \Delta \Gamma_{j\alpha}^\alpha - f_{,k} \Delta \Gamma_{ij}^k = 0$$

or

$$\tfrac{1}{2} n_j \Delta \Gamma_{i\alpha}^\alpha + \tfrac{1}{2} n_i \Delta \Gamma_{j\alpha}^\alpha - n_k \Delta \Gamma_{ij}^k = 0. \tag{4.1}$$

5. APPLICATIONS TO SCHWARZSCHILD'S SPHERE

We shall now apply the criterion (4.1) to the field of a stationary sphere of perfect fluid[1] and show that it leads to physically significant results.

The field being static and spherically symmetric, we assume

$$ds^2 = -e^\lambda \, dr^2 - r^2 \, d\theta^2 - r^2 \sin^2\theta \, d\phi^2 + e^\nu \, dt^2, \tag{5.1}$$

where $\lambda = \lambda(r)$, $\nu = \nu(r)$.

Let $r = r_1$ be the boundary of the sphere at which we have $\Delta g_{ij} = 0$. Further $\lambda, \nu \to 0$ as $r \to \infty$.

The nonvanishing Christoffel symbols have the following values:

$$\begin{aligned} &\Gamma_{11}^1 = \tfrac{1}{2}\lambda', && \Gamma_{12}^2 = \Gamma_{13}^3 = r^{-1}, \\ &\Gamma_{14}^4 = \tfrac{1}{2}\nu', && \Gamma_{23}^3 = \cot\theta, \\ &\Gamma_{33}^1 = -r\sin^2\theta e^{-\lambda}, && \Gamma_{33}^2 = -\sin\theta\cos\theta, \\ &\Gamma_{22}^1 = -re^{-\lambda}, && \Gamma_{44}^1 = \tfrac{1}{2}\nu' e^{\nu-\lambda}, \end{aligned} \tag{5.2}$$

where the prime denotes differentiation with respect to r.

External region: $r > r_1$: The equations are

$$\begin{aligned} e^{-\lambda}(r^{-1}\nu' + r^{-2}) - r^{-2} &= 0, \\ e^{-\lambda}(\tfrac{1}{2}\nu'' - \tfrac{1}{4}\lambda'\nu' + \tfrac{1}{4}\nu'^2 + \tfrac{1}{2}r^{-1}\nu' - \tfrac{1}{2}r^{-1}\lambda') &= 0, \\ e^{-\lambda}(r^{-1}\lambda' - r^{-2}) + r^{-2} &= 0. \end{aligned} \tag{5.3}$$

The solution of these equations under the given condition at infinity is the well-known Schwarzschild's exterior solution

$$ds^2 = -(1 - 2Gm/r)^{-1} \, dr^2 - r^2 \, d\theta^2 - r^2 \sin^2\theta \, d\phi^2 + (1 - 2Gm/r) \, dt^2, \tag{5.4}$$

where m is a constant.

Internal region: $r < r_1$: The energy-momentum tensor is given by

$$T^{ij} = (\rho_0 + p_0)u^i u^j - g^{ij} p_0,$$

where ρ_0 and p_0 are the proper density and pressure of the fluid and we take ρ_0 to be constant. The equations are

$$\begin{aligned} e^{-\lambda}(r^{-1}\nu' + r^{-2}) - r^{-2} &= \kappa p_0, \\ e^{-\lambda}(\tfrac{1}{2}\nu'' - \tfrac{1}{4}\lambda'\nu' + \tfrac{1}{4}\nu'^2 + \tfrac{1}{2}r^{-1}\nu' - \tfrac{1}{2}r^{-1}\lambda') &= \kappa p_0, \\ e^{-\lambda}(r^{-1}\lambda' - r^{-2}) + r^{-2} &= \kappa \rho_0. \end{aligned} \tag{5.5}$$

It follows that

$$p_0' + \tfrac{1}{2}(\rho_0 + p_0)\nu' = 0. \tag{5.6}$$

The interior solution is of the form

$$ds^2 = -(1 - r^2/R^2)^{-1} \, dr^2 - r^2 \, d\theta^2 - r^2 \sin^2\theta \, d\phi^2 + [A - B(1 - r^2/R^2)^{\frac{1}{2}}]^2 \, dt^2, \tag{5.7}$$

where $R^2 = 3/\kappa\rho_0$.

$\Delta g_{11} = 0$ at $r = r_1$ gives

$$m = \tfrac{4}{3}\pi r_1^3 \rho_0 . \tag{5.8}$$

The first equation is of (5.5) gives

$$\kappa R^2 p_0 = \frac{3B(1 - r^2/R^2)^{\frac{1}{2}} - A}{A - B(1 - r^2/R^2)^{\frac{1}{2}}} . \tag{5.9}$$

From the first equations of (5.3) and (5.5), we get

$$\Delta\nu' = -\kappa r_1 e^{\lambda(r_1)} p_0(r_1) \quad \text{at} \quad r = r_1 . \tag{5.10}$$

By (5.2), the only nonvanishing components of $\Delta\Gamma^k_{ij}$ at $r = r_1$ are

$$\Delta\Gamma^1_{11} = \tfrac{1}{2}\Delta\lambda', \quad \Delta\Gamma^4_{14} = \tfrac{1}{2}\Delta\nu', \quad \Delta\Gamma^1_{44} = \tfrac{1}{2}e^{\nu-\lambda}\Delta\nu'. \tag{5.11}$$

For the boundary $r = r_1$, n_2, n_3, n_4 are all zero. Then (4.1) simply gives $\Delta\nu' = 0$ and, hence, by (5.10),

$$p_0(r_1) = 0. \tag{5.12}$$

This is, in fact, a physical necessity and was assumed by Schwarzschild on physical grounds.

Now (5.12) and $\Delta g_{44} = 0$ at $r = r_1$ give

$$A = \tfrac{3}{2}(1 - r_1^2/R^2)^{\frac{1}{2}}, \quad B = \tfrac{1}{2}, \tag{5.13}$$

and the solution is complete.

Remark: Note that in the above solution the conditions (4.1) are satisfied although dg_{11}/dr is discontinuous at the boundary.

ACKNOWLEDGMENT

I wish to express my heartfelt thanks to Professor J. A. Wheeler of the Physics Department, Princeton University, for his kind interest in the paper and for communicating the same for publication.

[1] R. C. Tolman, *Relativity, Thermodynamics and Cosmology* (Oxford U. P., London, 1934).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 5 MAY 1970

Symmetric Dirac Bracket in Classical Mechanics

WILHELM H. FRANKE* AND ANDRÉS J. KÁLNAY†
Facultad de Ciencias Físicas y Matemáticas, Universidad Nacional de Ingeniería, Lima, Perú

(Received 22 July 1969)

It is known that, in classical systems that have second-class constraints which relate the canonical coordinates and momentum, the ordinary skew-symmetric Poisson bracket must be replaced by the skew-symmetric Dirac bracket. It is also known that in the process of quantization of such systems, the Dirac bracket replaces the Poisson bracket in its correspondence with the quantum commutators. In this paper we obtain the symmetric partner of the Dirac bracket, which is of interest not only to classical mechanics but also in regard to the quantization procedure; i.e., the quantization rules for systems which are restricted by second-class constraints such that the commutation rules involve anticommutators (instead of commutators, as in certain fields with Fermi–Dirac statistics) can be given in terms of this new symmetric bracket. This symmetric bracket is related to the Poisson–Droz-Vincent symmetric bracket.

1. INTRODUCTION

There are important physical systems that are restricted by second-class constraints.[1,2] In the quantization of these systems, the Poisson bracket *does not* correspond to the classical limit of the commutator and, therefore, the quantization procedure must be modified. This was first shown by Dirac in the skew-symmetric case.[1–4] There are important systems such as the vector field[5] and the gravitational field[4,6] which belong to this case.

As a review, let us remember that usually in Hamiltonian mechanics the generalized coordinates q_r and their canonically conjugate momenta

$$p_r = \frac{\partial L}{\partial(dq_r/dt)} \tag{1.1}$$

are independent variables! Also there are some less known cases in which it fails and the Eqs. (1.1) are identities in q and p only, even before Lagrange equations are used. Equations (1.1) and others derived from it by consistency relations are called constraints and denoted by $f_a(q, p) \approx 0$.[1–7] The weak equality sign, $\approx$ introduced by Dirac, means that the expression can be used only after computing all the Poisson brackets and partial differentiations in an equation (see, e.g., Ref. 2). The above mentioned second-class constraints[1–3] belong to a certain subset of the former set (where, eventually, some constraints are replaced by linear combinations of them[1,2,8]). We denote the second-class constraints as[9]

$$\theta_a(q, p) \approx 0, \quad a = 1, 2, \cdots, N_\theta. \tag{1.2}$$

$\Delta g_{11} = 0$ at $r = r_1$ gives

$$m = \tfrac{4}{3}\pi r_1^3 \rho_0 . \tag{5.8}$$

The first equation is of (5.5) gives

$$\kappa R^2 p_0 = \frac{3B(1 - r^2/R^2)^{\frac{1}{2}} - A}{A - B(1 - r^2/R^2)^{\frac{1}{2}}} . \tag{5.9}$$

From the first equations of (5.3) and (5.5), we get

$$\Delta\nu' = -\kappa r_1 e^{\lambda(r_1)} p_0(r_1) \quad \text{at} \quad r = r_1 . \tag{5.10}$$

By (5.2), the only nonvanishing components of $\Delta\Gamma^k_{ij}$ at $r = r_1$ are

$$\Delta\Gamma^1_{11} = \tfrac{1}{2}\Delta\lambda', \quad \Delta\Gamma^4_{14} = \tfrac{1}{2}\Delta\nu', \quad \Delta\Gamma^1_{44} = \tfrac{1}{2}e^{\nu-\lambda}\Delta\nu'. \tag{5.11}$$

For the boundary $r = r_1$, n_2, n_3, n_4 are all zero. Then (4.1) simply gives $\Delta\nu' = 0$ and, hence, by (5.10),

$$p_0(r_1) = 0. \tag{5.12}$$

This is, in fact, a physical necessity and was assumed by Schwarzschild on physical grounds.

Now (5.12) and $\Delta g_{44} = 0$ at $r = r_1$ give

$$A = \tfrac{3}{2}(1 - r_1^2/R^2)^{\frac{1}{2}}, \quad B = \tfrac{1}{2}, \tag{5.13}$$

and the solution is complete.

Remark: Note that in the above solution the conditions (4.1) are satisfied although dg_{11}/dr is discontinuous at the boundary.

ACKNOWLEDGMENT

I wish to express my heartfelt thanks to Professor J. A. Wheeler of the Physics Department, Princeton University, for his kind interest in the paper and for communicating the same for publication.

[1] R. C. Tolman, *Relativity, Thermodynamics and Cosmology* (Oxford U. P., London, 1934).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 5 MAY 1970

Symmetric Dirac Bracket in Classical Mechanics

WILHELM H. FRANKE* AND ANDRÉS J. KÁLNAY†
Facultad de Ciencias Físicas y Matemáticas, Universidad Nacional de Ingeniería, Lima, Perú

(Received 22 July 1969)

It is known that, in classical systems that have second-class constraints which relate the canonical coordinates and momentum, the ordinary skew-symmetric Poisson bracket must be replaced by the skew-symmetric Dirac bracket. It is also known that in the process of quantization of such systems, the Dirac bracket replaces the Poisson bracket in its correspondence with the quantum commutators. In this paper we obtain the symmetric partner of the Dirac bracket, which is of interest not only to classical mechanics but also in regard to the quantization procedure; i.e., the quantization rules for systems which are restricted by second-class constraints such that the commutation rules involve anticommutators (instead of commutators, as in certain fields with Fermi–Dirac statistics) can be given in terms of this new symmetric bracket. This symmetric bracket is related to the Poisson–Droz-Vincent symmetric bracket.

1. INTRODUCTION

There are important physical systems that are restricted by second-class constraints.[1,2] In the quantization of these systems, the Poisson bracket *does not* correspond to the classical limit of the commutator and, therefore, the quantization procedure must be modified. This was first shown by Dirac in the skew-symmetric case.[1–4] There are important systems such as the vector field[5] and the gravitational field[4,6] which belong to this case.

As a review, let us remember that usually in Hamiltonian mechanics the generalized coordinates q_r and their canonically conjugate momenta

$$p_r = \frac{\partial L}{\partial(dq_r/dt)} \tag{1.1}$$

are independent variables! Also there are some less known cases in which it fails and the Eqs. (1.1) are identities in q and p only, even before Lagrange equations are used. Equations (1.1) and others derived from it by consistency relations are called constraints and denoted by $f_a(q, p) \approx 0$.[1–7] The weak equality sign, $\approx$ introduced by Dirac, means that the expression can be used only after computing all the Poisson brackets and partial differentiations in an equation (see, e.g., Ref. 2). The above mentioned second-class constraints[1–3] belong to a certain subset of the former set (where, eventually, some constraints are replaced by linear combinations of them[1,2,8]). We denote the second-class constraints as[9]

$$\theta_a(q, p) \approx 0, \quad a = 1, 2, \cdots, N_\theta. \tag{1.2}$$

The procedure of finding this subset is explained in Dirac's papers[1,2,8] and it will be shortly reviewed in Sec. 4. Here we only mention that the constraints θ_a are such that the matrix c_{ab}^- defined by[1,2,9,10]

$$c_{ab}^-\{\theta^b, \theta^c\}_- = \delta_a^c \tag{1.3}$$

exists. It is clear that, when the constraints exist, the generalized velocities dq^r/dt cannot be solved in terms of the coordinates and momenta, so that it is impossible to go from the Lagrangian to the Hamiltonian form of classical mechanics by the standard procedure. All this suggests that for constrained systems the Hamiltonian classical mechanics and the standard quantization procedure must be reformulated. Moreover, if the quantization rule is used in the standard form

$$i\{\ ,\ \} \to [\ ,\ ] \tag{1.4a}$$

when there are constraints like (1.2), there will be inconsistencies in the quantum theory.[1-3] In order to solve those difficulties Dirac found the following bracket[1]:

$$\{f, g\}^* \equiv \{f, g\}_-^* = \{f, g\}_- - \{f, \theta^a\}_- c_{ab}^- \{\theta^b, g\}_-, \tag{1.5}$$

now called the *Dirac bracket*,[3] which replaces the Poisson bracket in the theory (see Refs. 2-4). The quantization rule is no longer (1.4a), but

$$i\{\ ,\ \}_-^* \to [\ ,\ ]_-. \tag{1.6a}$$

This bracket has the same formal properties as the minus Poisson bracket has (for example, it obeys Jacobi's identity[1,3,11]). Bergmann and Goldberg[3] showed their relation to the infinitesimal Lie bracket of a special group of transformations, and Mukunda and Sudarshan[11] clarified their structure and properties in an algebraic way.

But for unconstrained classical systems it was shown by Droz-Vincent[12] *that classical mechanics has a richer structure than was previously supposed: Not only the ordinary* (i.e., *the skew-symmetric*) *Poisson bracket exists, but also its symmetrical partner. To a certain extent the bracket theory can be developed in equal footing for both brackets* [cf. Sec. 3, Remark (2)]. Not only does the old Poisson bracket play an important role in classical mechanics, but so does the Droz-Vincent's new symmetric Poisson bracket. While it is well known that the Hamiltonian equations of motion can be transformed in such a way that the time evolution of the system is expressed purely in terms of the antisymmetric Poisson bracket, Droz-Vincent proved[12] that the same can be done in terms of the new symmetric bracket he found. With regard to the *analogy with quantum mechanics*, it was known that the transformation theory for classical systems allows us to introduce commutators in classical mechanics and to relate them directly to the minus Poisson bracket. Droz-Vincent showed that anticommutators can be also introduced in classical mechanics and that they have a similar relation to the symmetric Poisson bracket. This closes a gap for the quantization procedure: the quantization rule will be expressed as

$$i\{\ ,\ \}_\pm \to [\ ,\ ]_\pm, \tag{1.4b}$$

where the + sign is for systems to be quantized by anticommutators and the − sign is for systems to be quantized by commutators. The quantization rule with the + sign is of interest in regards to the quantization of fields which obey Fermi–Dirac statistics.

We have briefly reviewed that for classical unconstrained systems symmetric and antisymmetric Poisson brackets appear in almost equal footing, and that for constrained systems there exists an antisymmetric Dirac bracket $\{\ ,\ \}_-^*$ which plays the role that the antisymmetric Poisson bracket does for unconstrained systems. Therefore, *we should expect that, when constraints are present, then a* (*yet unknown*) *symmetric partner* $\{\ ,\ \}_+^*$ *which we shall call plus Dirac bracket exists. It is a basic purpose of this paper to obtain this bracket.* In particular, it should be expected that instead of (1.4) the quantization rule should be

$$i\{\ ,\ \}_\pm^* \to [\ ,\ ]_\pm. \tag{1.6b}$$

We shall follow the following plan: In Sec. 2 we will state the notation and the conventions. A further discussion of the basic notions will be given in Sec. 3, and in Sec. 4 we shall introduce the symmetric partner of the Dirac bracket $\{\ ,\ \}_+^*$, and we shall study some of its properties.

2. NOTATIONS AND CONVENTIONS

The sum convention is used in any place, as well as the natural system of unities.

Commutators (or minus *commutators*) and *anticommutators* (or plus *commutators*) are denoted by $[\ ,\ ]_\pm$. The curly bracket is reserved for *the ordinary Poisson bracket* (or minus *Poisson bracket*) and *the plus Poisson bracket*[13]:

$$\{f, g\}_\pm \equiv \{f, g\}_+(q, p) \equiv \{f, g\}_\pm(\omega) = \frac{\partial f}{\partial q^r}\frac{\partial g}{\partial p^r} \pm \frac{\partial f}{\partial p^r}\frac{\partial g}{\partial q^r}, \tag{2.1}$$

where

$$q = (q^1, q^2, \cdots, q^N), \quad p = (p^1, p^2, \cdots, p^N)$$

includes *all* variables of the phase space of the system considered. As in Ref. 11, we use $\omega^1 = q^1, \cdots, \omega^N = q^N, \omega^{N+1} = p^1, \cdots, \omega^{2N} = p^N, \omega = (\omega^1, \cdots, \omega^{2N}) = (q, p)$. The ordinary *Dirac bracket* (or minus *Dirac bracket*) and the new plus *Dirac bracket* are denoted respectively $\{\ ,\ \}^*_-$ and $\{\ ,\ \}^*_+$. [They are defined in Eqs. (1.5) and (4.10).] To distinguish entities constructed from (or for) the plus and minus brackets, we attach the corresponding + and − sign to them. For example, ± second-class constraints will be introduced for the theory of ± Dirac brackets.

Indices for coordinates in phase space: For q and p we use r, s, and t as in q^r; for ω we use $\mu, \nu, \cdots, \rho, \sigma$.

Other indices: For the functions ψ to be introduced in Eq. (4.5) we shall use the indices $m, n, \cdots, q$. For the functions θ previously introduced (and as a particular case, for second-class constraints) we reserve a, b, and c. Our index conventions are defined in such a way that do not contradict those used in Ref. 11.

3. REMARKS

(1) Droz-Vincent[12] introduces the plus Poisson bracket (as well as a generalization of the usual minus Poisson bracket) in terms of covariant derivatives in a connection Γ. We use $\Gamma = 0$.

(2) It is surprising the parallelism that exists between the properties of the plus and minus Poisson brackets. For example, both brackets are related (in a symmetric form) to infinitesimal transformations. Moreover, the standard Poisson bracket form of Hamilton's equations is almost the same for both brackets[12]:

$$\frac{dq_r}{dt} = \pm\{H, q_r\}_\pm, \quad \frac{dp_r}{dt} = -\{H, p_r\}_\pm. \tag{3.1}$$

(3) However, the old minus Poisson bracket seems to maintain some privileges with respect to Droz-Vincent's plus Poisson bracket. For example, Eqs. (3.1) are particular cases of the well-known formula for the minus Poisson bracket

$$\frac{df}{dt} = -\{H, f\}_-, \quad \text{for all } f = f(q, p), \tag{3.2}$$

and it seems that Eq. (3.2) has no counterpart for the plus Poisson bracket. This is all right with regard to the quantization procedure since the pre-eminence of the minus commutator exists also in quantum mechanics. There the time evolution of any quantum operator is always computed in terms of the minus commutator,

$$\frac{df_{\text{op}}}{dt} = i[H_{\text{op}}, f_{\text{op}}]_-, \tag{3.3}$$

("op" means operator), even for those fields quantized by anticommutators.

(4) Besides Eqs. (3.1) we also have

$$\frac{\partial f}{\partial q_r} = \{f, p_r\}_\pm \quad \text{and} \quad \frac{\partial f}{\partial p_r} = \pm\{f, q_r\}_\pm. \tag{3.4}$$

(5) While minus Poisson (or Dirac) bracket and minus commutators have the same formal laws, the same does not happen for plus Poisson (or Dirac) bracket and plus commutators. For example,

$$\{f, gh\}_\pm = g\{f, h\}_\pm + \{f, g\}_\pm h \tag{3.5}$$

[the same will happen with the ± Dirac brackets; cf. Eq. (4.19d)] and

$$[f, gh]_- = g[f, h]_- + [f, g]_- h \tag{3.6}$$

have the same form, but, in general, $[f, gh]_+$ and $g[f, h]_+ + [f, g]_+ h$ are different.

(6) For a classical system, the antisymmetric structure obtained from the minus bracket can coexist with the symmetric one obtained from the plus bracket. The well-known rule that integral spin fields must be quantized with minus commutators while half-integral spin fields must be quantized with plus commutators shows that the quantum level (where only one structure type is allowed for each field) is more restrictive in this sense than the classical level (where both structures can coexist) (cf. Ref. 12).

(7) The well-known canonical rules for minus Poisson brackets are also right for plus Poisson brackets:

$$\{q_r, q_s\}_\pm = 0, \quad \{q_r, p_s\}_\pm = \delta_{rs}, \quad \text{and} \quad \{p_r, p_s\}_\pm = 0. \tag{3.7}$$

This gives further confidence to rule (1.4) for the symmetric case.

(8) It follows from Eq. (1.5) that the minus Dirac bracket can not obey simultaneously relations of the type (3.7). This fact is related to the existence of constraints. The same happens to the plus Dirac bracket to be introduced in Sec. 4.

(9) The Dirac bracket can be developed in an abstract form in which the θ^a are arbitrary functions not necessarily equal to zero. The work of Mukunda and Sudarshan was done with this consideration.[11]

4. THE NEW SYMMETRIC DIRAC BRACKET

A. Extension of Mukunda and Sudarshan's Approach

Mukunda and Sudarshan[11] gave a new method to obtain the ordinary (or minus type, or antisymmetric) Dirac bracket $\{\ ,\ \}^*_-$ of Eq. (1.5). We will use the same method to introduce its symmetric partner $\{\ ,\ \}^*_+$. In

order to stress the similarity with the antisymmetric case, we will present both cases at once, with the following convention: In every term of the type $F_{\pm}$ which appears in this section, the upper sign (+) will correspond to the obtainment of $\{\ ,\ \}_{+}^{*}$ and the lower one (−) to $\{\ ,\ \}_{-}^{*}$. Of course, only expressions involving the upper sign will be new; those involving the lower one were considered by Mukunda and Sudarshan.[11] In this way we also emphasize that no new procedure of computation is developed by us; we will symmetrize in a natural manner Mukunda and Sudarshan's expressions. Our contribution is only to realize that a symmetrization can be done step by step.

Let us consider a classical system with canonical variables

$$\omega = \omega(q, p) = (\omega^1, \omega^2, \cdots, \omega^{2N}) = (q^1, q^2, \cdots, q^N, p^1, p^2, \cdots, p^N).$$

If $\mathbf{1}$ is the $N \times N$ unit matrix, the matrices

$$\|\epsilon_{\pm}^{\mu\nu}\| = \begin{Vmatrix} 0 & \mathbf{1} \\ \pm\mathbf{1} & 0 \end{Vmatrix}, \quad \|\epsilon_{\mu\nu}^{\pm}\| = \begin{Vmatrix} 0 & \pm\mathbf{1} \\ \mathbf{1} & 0 \end{Vmatrix}, \tag{4.1a}$$

$$\|\epsilon_{\pm}^{\mu\nu}\|^{-1} = \|\epsilon_{\mu\nu}^{\pm}\|, \tag{4.1b}$$

allow us to give the form

$$\{f, g\}_{\pm} \equiv \{f, g\}_{\pm}(\omega) = \epsilon_{\pm}^{\mu\nu} \frac{\partial f(\omega)}{\partial \omega^{\mu}} \frac{\partial g(\omega)}{\partial \omega^{\nu}} \tag{4.2a}$$

to the ordinary (or minus type) Poisson bracket $\{f, g\} \equiv \{f, g\}_{-}$ and to the simplest form of the plus Poisson bracket $\{f, g\}_{+}$ introduced by Droz-Vincent[12] [cf. Eq. (2.1)]. Here the functional form (and not the value) of $f(\omega)$ and $g(\omega)$ are important; the partial derivatives are computed as if all the variables ω^{μ} were independent, irrespectively of the existence (or nonexistence) of the identities (1.2) (cf., e.g., Ref. 2 or 11). In the standard q, p notation, Eqs. (4.2a) read

$$\{f, g\}_{\pm} \equiv (f, g)_{\pm}(q, p) = \frac{\partial f}{\partial q^r} \frac{\partial g}{\partial p^r} \pm \frac{\partial f}{\partial p^r} \frac{\partial g}{\partial q^r}. \tag{4.2b}$$

Let us consider now a set ϕ of $2N$ independent functions $\phi^1(\omega), \phi^2(\omega), \cdots, \phi^{2N}(\omega)$. We introduce the symmetric counterpart $L_{\rho\sigma}^{+}(\phi)$ of the ordinary Lagrange bracket $-L_{\rho\sigma}^{-}(\phi)$:

$$L_{\rho\sigma}^{\pm}(\phi) = \epsilon_{\mu\nu}^{\pm} \frac{\partial \omega^{\mu}}{\partial \phi^{\rho}} \frac{\partial \omega^{\nu}}{\partial \phi^{\sigma}}. \tag{4.3a}$$

In the standard q, p notation this means

$$L_{\rho\sigma}^{\pm}(\phi) = \frac{\partial p^r}{\partial \phi^{\rho}} \frac{\partial q^r}{\partial \phi^{\sigma}} \pm \frac{\partial q^r}{\partial \phi^{\rho}} \frac{\partial p^r}{\partial \phi^{\sigma}}, \tag{4.3b}$$

which provides the inverse matrix of the corresponding ±Poisson bracket (as it is well known for the minus case)

$$L_{\mu\nu}^{\pm}(\phi)\{\phi^{\nu}, \phi^{\xi}\}_{\pm} = \delta_{\mu}^{\xi}, \quad \mu, \nu = 1, 2, \cdots, 2N. \tag{4.4}$$

Now we decompose the set ϕ into disjoint subsets

$$\phi = \psi \cup \theta, \quad \psi \cap \theta = \varnothing, \tag{4.5a}$$

$$\psi(\omega) \equiv \psi = (\psi_1, \psi_2, \cdots, \psi_m, \cdots, \psi_{N_{\psi}^{\pm}}), \tag{4.5b}$$

$$\theta(\omega) \equiv \theta = (\theta_1, \theta_2, \cdots, \theta_a, \cdots, \theta_{N_{\theta}^{\pm}}), \tag{4.5c}$$

$$N_{\psi}^{\pm} + N_{\theta}^{\pm} = 2N. \tag{4.6}$$

In the antisymmetric case N_{θ}^{-} must be even, but it is not necessary in the symmetric case.

In Mukunda and Sudarshan's approach, both sets ψ and θ are arbitrarily chosen, *irrespectively* of whether or not θ_a is a constraint (i.e., irrespectively of whether or not $\theta_a \approx 0$). The only restrictions are that they obey Eqs. (4.5) and (4.6), that the set ϕ be a set of independent functions, and that the matrix $c_{ab}^{\pm}$ exist, defined by

$$c_{ab}^{\pm}\{\theta^b, \theta^c\}_{\pm} = \delta_a^c, \quad a, b, c = 1, 2, \cdots, N_{\theta}^{\pm} \tag{4.7}$$

[cf. Ref. 11 and Eq. (1.3)].

We shall prove below that a matrix $\eta_{\pm}^{mp}(\theta, \psi)$ exists defined by

$$\eta_{\pm}^{mp} L_{pn}^{\pm} = \delta_n^m, \quad m, n, p = 1, 2, \cdots, N_{\psi}. \tag{4.8}$$

[Notice the difference with Eq. (4.4).] This matrix is such that

$$\eta_{\pm}^{mp}(\theta, \psi) = \pm\eta_{\pm}^{pm}(\theta, \psi). \tag{4.9}$$

Now we can introduce a new bracket, which we call the *plus Dirac bracket* $\{\ ,\ \}_{+}^{*}$ because it is the symmetric partner of the definition of the minus Dirac bracket $\{\ ,\ \}_{-}^{*}$ in Mukunda and Sudarshan's approach:

$$\{f, g\}_{\pm}^{*}(\theta, \psi) = \eta_{\pm}^{mn}(\theta, \psi) \frac{\partial f}{\partial \psi^m} \frac{\partial g}{\partial \psi^n}, \quad m, n = 1, 2, \cdots, N_{\psi}, \tag{4.10}$$

for arbitrary $f(\theta, \psi)$ and $g(\theta, \psi)$.[14]

We shall transform it to a different form. From Eqs. (4.4) and (4.5) we find that

$$\{\theta^a, \theta^b\}_{\pm} L_{bc}^{\pm}(\theta, \psi) + \{\theta^a, \psi^m\}_{\pm} L_{mc}^{\pm}(\theta, \psi) = \delta_c^a, \tag{4.11a}$$

$$\{\theta^a, \theta^b\}_{\pm} L_{bn}^{\pm}(\theta, \psi) + \{\theta^a, \psi^m\}_{\pm} L_{mn}^{\pm}(\theta, \psi) = 0, \tag{4.11b}$$

$$\{\psi^m, \theta^a\}_{\pm} L_{ab}^{\pm}(\theta, \psi) + \{\psi^m, \psi^n\}_{\pm} L_{nb}^{\pm}(\theta, \psi) = 0, \tag{4.11c}$$

$$\{\psi^m, \theta^a\}_{\pm} L_{an}^{\pm}(\theta, \psi) + \{\psi^m, \psi^p\}_{\pm} L_{pn}^{\pm}(\theta, \psi) = \delta_n^m, \tag{4.11d}$$

$a, b, c = 1, 2, \cdots, N_{\theta}^{\pm}, m, n, p = 1, 2, \cdots, N_{\psi}^{\pm}$. By

combining Eqs. (4.7), (4.11b), and (4.11d), we deduce

$$[\{\psi^m, \psi^p\}_\pm - \{\psi^m, \theta^a\}_\pm c^\pm_{ab}\{\theta^b, \psi^p\}_\pm]L^\pm_{pn}(\theta, \psi) = \delta^m_n. \tag{4.12}$$

This shows that $\eta^{mp}_\pm$ exists and that

$$\eta^{mn}_\pm = \{\psi^m, \psi^n\}_\pm - \{\psi^m, \theta^a\}_\pm c^\pm_{ab}\{\theta^b, \psi^n\}_\pm, \tag{4.13}$$

so that

$$\{f, g\}^*_\pm(\theta, \psi) = \frac{\partial f}{\partial \psi^m}[\{\psi^m, \psi^n\}_\pm - \{\psi^m, \theta^a\}_\pm c^\pm_{ab}\{\theta^b, \psi^n\}_\pm]\frac{\partial g}{\partial \psi^n}. \tag{4.14}$$

On the other hand, in

$$\{f, \psi^n\}_\pm = \frac{\partial f}{\partial q^r}\frac{\partial \psi^n}{\partial p^r} \pm \frac{\partial f}{\partial p^r}\frac{\partial \psi^n}{\partial q^r} \tag{4.15}$$

[cf. Eq. (4.2b)] we express $\partial f/\partial q^r$ and $\partial f/\partial p^r$ in terms of $\partial f/\partial \psi^m$ and $\partial f/\partial \theta^c$. After some elementary calculations we obtain

$$\frac{\partial f}{\partial \psi^m}\{\psi^m, \psi^n\}_\pm = \{f, \psi^n\}_\pm - \frac{\partial f}{\partial \theta^c}\{\theta^c, \psi^n\}_\pm. \tag{4.16a}$$

Similarly,

$$\frac{\partial f}{\partial \psi^m}\{\psi^m, \theta^a\}_\pm = \{f, \theta^a\}_\pm - \frac{\partial f}{\partial \theta^c}\{\theta^c, \theta^a\}_\pm. \tag{4.16b}$$

We combine Eqs. (4.16) with Eq. (4.14):

$$\{f, g\}^*_\pm(\theta, \psi) = \{f, \psi^n\}_\pm \frac{\partial g}{\partial \psi^n} - \{f, \theta^a\}_\pm c^\pm_{ab}\{\theta^b, \psi^n\}_\pm \frac{\partial g}{\partial \psi^n}. \tag{4.17}$$

By transforming the $\partial g/\partial \psi^n$ terms in the same form as it was done above with the $\partial f/\partial \psi^m$ terms, we obtain

$$\{f, g\}^*_\pm = \{f, g\}_\pm - \{f, \theta^a\}_\pm c^\pm_{ab}\{\theta^b, g\}_\pm, \tag{4.18}$$

$a, b = 1, 2, \cdots, N^\pm_\theta$. In this form (for the $-$ sign) Mukunda and Sudarshan[11] proved that the Dirac bracket (4.10) defined in their approach is identical with the original one given by Dirac [cf. Eq. (1.5)].

With regard to the new plus case, if one tries to write directly the symmetric partner $\{f, g\}^*_+$ of the minus Dirac bracket $\{f, g\}^*_-$ given in its original form [i.e., as in Eq. (1.5)], problems will arise because, whereas

$$\{f, g\}^*_+ = \{f, g\}_+ - \{f, \theta^a\}_+ c^+_{ab}\{\theta^b, g\}_+$$

is a possible symmetric partner of Eq. (1.5), it does not seem obvious that this must be the right answer. For example,

$$-\{f, \theta^a\}_+ c^+_{ab}\{\theta^b, g\}_+$$

can be thought as well as

$$-\{f, \theta^a\}_- c^+_{ab}\{\theta^b, g\}_-$$

to be the symmetric partner of the term

$$-\{f, \theta^a\}_- c^-_{ab}\{\theta^b, g\}_-$$

of Eq. (1.5). On the other hand, Mukunda and Sudarshan's approach for the minus case has only one natural symmetric partner, as it can be seen in comparison of the $+$ and the $-$ signs in Eqs. (4.2b), (4.3b), (4.7), (4.8), and (4.10). That is why we use Mukunda and Sudarshan's approach in this section.

B. Properties

We know[1–3,11] that the minus Dirac bracket has the properties it is expected to have, if it is a generalization of the minus Poisson bracket and if it should be replaced while quantizing (beside an i factor) by a commutator. In the same way, Eq. (4.10) implies immediately that the new plus Dirac bracket has the properties it is expected to have if it is a generalization of the plus Poisson bracket and if it should be replaced while quantizing (beside an i factor) by an anticommutator. For example, given the arbitrary functions $f = f(\theta, \psi)$, $g = g(\theta, \psi)$, and $h = h(\theta, \psi)$ and any constant α, it follows that

$$\{f, g\}^*_\pm = \pm\{g, f\}^*, \tag{4.19a}$$

$$\{f, \alpha g\}^*_\pm = \alpha\{f, g\}^*_\pm, \tag{4.19b}$$

$$\{f, g + h\}^*_\pm = \{f, g\}^*_\pm + \{f, h\}^*_\pm, \tag{4.19c}$$

and

$$\{f, gh\}^*_\pm = g\{f, h\}^*_\pm + \{f, g\}^*_\pm h. \tag{4.19d}$$

An important property that the $\pm$ Poisson brackets do not have is

$$\{f, \theta^a\}^*_\pm = 0, \quad \text{for } a = 1, 2, \cdots, N^\pm_\theta. \tag{4.20}$$

This last result avoids inconsistencies when quantizing according to rule (1.6b), when constraints $\theta^a \approx 0$ are present.

C. Constraints

We extend the classification developed by Dirac,[1] Anderson and Bergmann[7] for the skew-symmetric case to the symmetric one.

The set of constraints (if any) deduced *directly* from Eq. (1.1) are called *primary constraints*. If $\Phi_a(q, p) \approx 0$ is a primary constraint, then it must remain valid for all time. Since Eq. (3.2) is valid for the plus and minus cases, *we have*

$$\{\Phi_a, H_T\}_- \approx 0, \tag{4.21}$$

(H_T is the total Hamiltonian of the system)[2] *not only when looking for the minus Dirac bracket but also when*

looking for the plus Dirac bracket. It may be that from Eqs. (4.21) new independent constraints of the type $\chi_a(q, p)$ are deduced, in which case they are called secondary constraints. The procedure is iterated, getting more secondary constraints, until no new constraint can be obtained. Let us call $\zeta_a(q, p)$ any of the Φ_a or χ_a: $\zeta_a(q, p) \approx 0$. We say that ζ_a is a first-class constraint if

$$\{\zeta_a, \zeta_b\}_\pm = 0, \text{ for all } b, \tag{4.22}$$

and that ζ_a is of second-class constraints if there exists b such that

$$\{\zeta_a, \zeta_b\}_\pm \neq 0. \tag{4.23}$$

It is intended in these expressions for only one sign to be used systematically for each classification (+ or −).

The constraints ζ_a are replaced by nonsingular linear combinations of these in such a way that we get as many $\pm$ first-class constraints as possible. The remaining ones result in $\pm$ second-class constraint functions which will be denoted by θ_a. They are such that

$$\theta_a(q, p) \approx 0, \quad a = 1, 2, \cdots, N_\theta^\pm. \tag{4.24}$$

In the same way as for the antisymmetric case it is proved that

$$\det |\{\theta_a, \theta_b\}_\pm| \neq 0. \tag{4.25}$$

D. The $\pm$ Dirac Brackets in the Theory of Constrained Systems

In Sec. 4A the Dirac bracket is defined in terms of any set of functions ψ_m and θ_a that satisfy the conditions which are imposed there. It is not required that the functions θ_a be constraint functions: this was not required either in Dirac's[1] first definition of $\{\ ,\ \}^*$ or in Mukunda and Sudarshan's[11] approach to them. But the principal interest of $\pm$ Dirac bracket is when the functions θ_a considered in subsections A and B are just the $\pm$ second-class constraints θ_a obtained in subsection C (cf. Refs. 1–6 for the antisymmetric case). This selection of the set θ can be done because Eq. (4.25) insures the existence of $c_{ab}^\pm$ [cf. Eq. (4.7)].

ACKNOWLEDGMENTS

We acknowledge useful discussions with Professor D. Finkelstein and Professor J. Tiomno held at the International Centre for Theoretical Physics at Miramare-Trieste during the visit there of one of us (A. J. K.). The Centre's hospitality is also warmly approciated.

* Mail address: Programa de Ciencias, Universidad Nacional de Ingeniería, Apartado Postal 1301, Lima, Perú.

† New address: Departamento de Física, Escuela de Física y Matemáticas, Facultad de Ciencias, Universidad Central de Venezuela, Caracas, Venezuela.

[1] P. A. M. Dirac, Can. J. Math. **2**, 129 (1950).

[2] P. A. M. Dirac, *Lectures on Quantum Mechanics*, Belfer Graduate School of Science Monograph Series, No. 2 (Yeshiva University, New York, 1964).

[3] P. G. Bergmann and I. Goldberg, Phys. Rev. **98**, 531 (1955).

[4] P. G. Bergmann and A. B. Komar, "Status Report on Quantization of the Gravitational Field," in *Recent Developments in General Relativity* (Pergamon Press, Oxford and Polish Scientific Publishers, Warsaw, 1962).

[5] P. A. M. Dirac, Can. J. Math. **3**, 1 (1951).

[6] P. A. M. Dirac, Phys. Rev. **114**, 924 (1959).

[7] J. L. Anderson and P. G. Bergmann, Phys. Rev. **83**, 1018 (1951).

[8] P. A. M. Dirac, Proc. Roy. Soc. (London), Ser. A, **246**, 326 (1958).

[9] The notation will be stated in Sec. 2.

[10] We use the sum convention.

[11] N. Mukunda and E. C. G. Sudarshan, J. Math. Phys. **9**, 411 (1968).

[12] P. Droz-Vincent, Ann. Inst. Henri Poincaré, Sec. A, **5**, 257 (1966).

[13] The plus Poisson bracket here defined is the same as that introduced by Droz-Vincent (Ref. 12) if in his paper we put $\Gamma = 0$ [see Sec. 3, Remark (1)] and use the simplest choice of his matrix M.

[14] It follows that $\{f, g\}_\pm^*(\theta, \psi)$ is the particular kind of the Droz-Vincent's Poisson bracket (Ref. 12) (given in its general form, not in the elementary one of our Sec. 2), which is obtained if in Droz-Vincent's paper we choose the ψ-space as V_m and set $M^{mn} = \eta_\pm^{mn}(\theta, \psi)$ and $\Gamma = 0$ there.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 5 MAY 1970

Analytic Solutions to Two Albedo Problems

PETER D. NOERDLINGER
Department of Physics, New Mexico Institute of Mining and Technology, Socorro, New Mexico 87801

(Received 24 November 1969)

A random walk approach leads to rigorous solution of certain albedo problems usually solved by Monte Carlo techniques or other numerical methods.

I. INTRODUCTION

In the problem of "diffuse reflection," light is incident at an angle $\cos^{-1}\mu_0$ on material that can scatter or absorb it, with the absorption probability a (often[1] $1-a$ is denoted ω_0) per scattering. The general problem, in various geometries, is to determine the intensity and angular distribution of the transmitted and reflected light. The problem may be treated by use of integral equations[1] or by Monte Carlo techniques,[2] but no purely analytic solutions (free of numerical procedures) have been given for any of the quantities. In this paper, analytic expressions are derived for the total albedo from a semi-infinite slab when the albedo per scattering, $b = 1 - a$, is known and the scattering is isotropic. Random walk techniques are used, the first case being in one dimension, with walk allowed only in a direction normal to the slab surface; in the second case, random walk is allowed in any of the six principal directions. Only the total albedo,[3] not the angular distribution of the reflected light, can be obtained in this manner, since all the photons emerge normally to the slab. However, the results appear in simple closed form, offering great ease in use and great advantage for physical interpretation. For example, when a is large, we would expect our best chances to get a photon back are if it scatters only once, with negligible chance of survival otherwise. Thus, in one dimension, the slab albedo R should be approximately $\frac{1}{2}b$ and in three dimensions $\frac{1}{6}b$; these limits are easily extracted from our answers. Furthermore, our method illustrates a way of using generating functions to obtain averages weighted by the path length of the photon in the slab; these might be of use in a variety of other problems, as in finding the mean time delay for exit or evaluating effects of the photons on slab material. Our results agree extremely well with those obtained using the X and Y functions of Chandrasekhar.[1] In Sec. II we evaluate the slab albedo in one dimension and in Sec. III in three dimensions.

II. 1-DIMENSIONAL CASE

This case is absurd as a photon-scattering problem, as may be seen from the factor-3 difference in R when a is large (Sec. I). However, there are interesting systems that are intrinsically of a more 1-dimensional nature, such as the propagation of cosmic rays along a fairly regular magnetic field.[4] This case serves, then, as an introduction to the 3-dimensional case and as a case of possible interest in other fields. Suppose that the photons arrive from the left (empty) half-space at the first of an infinite row of scatterers occupying the right half-space and that at each scatterer a photon can do any of the following three things:

be absorbed (probability a),
scatter to the left (probability $\frac{1}{2}b$),
scatter to the right (probability $\frac{1}{2}b$).

The problem is basically a 1-dimensional random walk, and we follow Feller's[5] treatment; but we must change the initialization and use certain symmetry ideas to take care of the fact that half the space is empty, and we must take into account the survival probability per scattering, b. Clearly, if $b = 1$, all the photons return to the surface and escape, since[5] a random walk in one dimension leads to unit probability of return to the origin. The interesting question in regard to absorption arises when one remembers that the *expectation* of the recurrence time in this problem is *infinite*, which suggests that perhaps a small chance of absorption per scattering will really destroy a lot of photons. In fact, it turns out that only very few photons contribute the exceedingly long paths in the material that lead to infinite expectation for the path length; hence, the slab albedo falls only linearly with b at small b, and approaches unity rapidly when b is near 1.

We define a *chain* of scatterings as a sequence of possible scatterings leading to re-emission at the last member, but not before. Thus, if scattering to the right is denoted R and to the left L, a sample of the

simpler chains allowed is the following.

Scattering	Probability	
L	$(\frac{1}{2}b)$	
RLL	$(\frac{1}{8}b^3)$	
$RRLLL$		
$RLRLL$	$(\frac{1}{16}b^5)$	(1)
$RRRLLLL$		
$RRLRLLL$		
$RRLLRLL$	$(\frac{5}{128}b^7)$	
$RLRRLLL$		
$RLRLRLL$		

We have grouped the chains by length and have listed with each group the associated contribution to R, the slab albedo. Each chain must have one more L than R, must end with an L, and must contain no more L's than R's *until* the last L. By comparison with Feller,[5] it is easy to see that, if we prefix an extra R to each chain, we get exactly all those 1-dimensional random walks that start to the *right* and return to the origin at the end of the chain, which is half of all the recurrent walks considered by Feller (whose lattice of scattering centers was infinite, not semi-infinite). Since our sample space contains only random walks that start to the right, our probabilities for return (when $b = 1$) are identical with Feller's.

One may think of the missing initial "R" as the process of directing the photon at the slab, which is taken here to have probability 1. Clearly, if we included the initial "R" *and* assigned it (wrongly) probability b, the sum of the terms analogous to those in the second column of Eq. (1) would be Feller's generating function [Eq. (3.18) of Ref. 5], with $s = b$, viz.,

$$F(b) = 1 - (1 - b^2)^{\frac{1}{2}}, \tag{2}$$

and this would be the slab albedo. Since the insertion of the photon is *demanded*, not assigned a probability $b < 1$, we must remove the extra factor b from Eq. (2) to get the true albedo of the slab $R(b)$, namely

$$R(b) = b^{-1}[1 - (1 - b^2)^{\frac{1}{2}}] \quad \text{(1-dimensional)}. \tag{3}$$

This completes the solution of the 1-dimensional problem, there being no point in rederiving here Feller's generating function for recurrence times. In the next section such a function must be derived in three dimensions for recurrence at a plane, since Feller treated only recurrence at a point.

In closing, we note that (3) gives $R \simeq \frac{1}{2}b$ when b is small, as promised, and that, when a is small, (3) yields $R \simeq 1 - (2a)^{\frac{1}{2}}$, which supports the claim that the total albedo rapidly approaches unity when a does, the long chains evidently contributing little.

III. 3-DIMENSIONAL CASE

In this case, a photon arrives from the left and strikes the slab, which occupies the right half-space. At each scattering, it can do any of seven things:

be absorbed (probability a),
scatter to the left (probability $\frac{1}{6}b$),
scatter to the right (probability $\frac{1}{6}b$),
scatter sidewise any of four ways (total probability $\frac{2}{3}b$).

Again, $a + b = 1$. Clearly, we do not care which of the four sidewise directions is involved, and we may lump them together, using the symbol S on a par with R and L. Thus, chains may be formed just as before, but it is now a little problem to arrange them so as to easily count up the probabilities. Note that we again expect unit probability of return to the original *surface* if $b = 1$, although the probability of return to the starting *point* is only[6] 0.35.

We form chains as before, but with an arbitrary number of S's inserted at any point save the end of the chain. (The first or any intermediate scattering could be sidewise, but the last is assumed to be *out* of the slab.) We group the chains just as was done in Eq. (1), but with subgroups containing a given number of S's, this number being denoted in the sequel as r:

$$L, \quad SL, \quad SSL, \quad SSSL, \cdots \tag{4a}$$

$$\begin{array}{llll} RLL, & SRLL, & SSRLL, & SSSRLL, \cdots \\ & RSLL, & RSSLL, & RSSSLL, \cdots \\ & RLSL, & RLSSL, & RLSSSL, \cdots \\ & & RSLSL, & RSLSSL, \cdots. \\ 0 & 1 & 2 & 3 \end{array} \tag{4b}$$

In (4a) r takes on the successive values 0, 1, 2, 3, and in (4b) the chains are grouped by r values shown beneath the columns. If the original chain (before insertion of S's) was n units long, the number of ways that rS's can be inserted is a well-known problem in Bose–Einstein statistics,[7] since the n possible spaces before each R or L form "boxes" into which the S's may be put like Bose particles. The resulting number of terms is[7]

$$\binom{n + r - 1}{r}. \tag{5}$$

We denote by $P(n)$ the probabilities for the original 1-dimensional chains (1), but with $\frac{1}{2}b$ replaced by $\frac{1}{6}b$ as is appropriate in three dimensions. The chain length is n (not counting the S's), and P is supposed to be already summed over all original chains of length n.

Let $R(n; r)$ denote the partial albedo due to chains with nR's or L's, and rS's. Then

$$R(b) = \sum_{n,r} R(n; r), \quad n \text{ odd}. \tag{6}$$

We perform the sum over r first. By inspection, we see that

$$\sum R(n; r) = P(n) \sum_{r=0}^{\infty} \binom{n + r - 1}{r} (\tfrac{2}{3}b)^r, \tag{7a}$$

$$= P(n)(1 - \tfrac{2}{3}b)^{-n}, \tag{7b}$$

where we have used an identity in binomial coefficients [Ref. 5, Eq. (9.1)] and the definition of binomial series. Now the appearance of a simple power law in (7b) tremendously facilitates the final summation over n, because the quantities $P(n)$ are already simple coefficients times $(\frac{1}{6}b)^n$. Therefore, the final summation is *exactly* as in Sec. II, except that $\frac{1}{2}b$ is to be replaced by $(\frac{1}{6}b)/(1 - \frac{2}{3}b) = \frac{1}{2}b/(3 - 2b)$. The final answer for the slab albedo is then

$$R(b) = x^{-1}[1 - (1 - x^2)^{\frac{1}{2}}], \quad x \equiv b/(3 - 2b). \tag{8}$$

If b is small, $R \to \frac{1}{6}b$, as asserted, and if b is large, we find

$$R \simeq 1 - (6a)^{\frac{1}{2}}, \quad a \text{ small}, \tag{9}$$

a nice analytical limit that could never be obtained by numerical techniques.[8] Again, in this limit, the influence of long paths must be small, in *spite* of the infinite expectation of the recurrence time! To conclude, we give in Table I a comparison of our results with those obtainable from Chandrasekhar's X, Y, and H functions; $R = 1 - H(1; b)(1 - b)^{\frac{1}{2}}$, where H is tabulated.[1]

TABLE I. Comparison of analytic expression (8) with Chandrasekhar's numerical results.

b	R, from Eq. (8)	R, from Ref. 1	b	R, from Eq. (8)	R, from Ref. 1
0.1	0.0179	0.0164	0.8	0.315	0.2853
0.2	0.0385	0.0352	0.9	0.453	0.4149
0.3	0.0627	0.0572	0.925	0.504	0.4665
0.4	0.0916	0.0834	0.95	0.574	0.5355
0.5	0.127	0.1152	0.975	0.677	0.641
0.6	0.172	0.1554	1.0	1.0	1.0
0.7	0.231	0.2087			

ACKNOWLEDGMENTS

The author is indebted to Dr. W. Irvine and Dr. S. Chandrasekhar for private communications and to the National Science Foundation for partial support under Grant GU 2921.

[1] S. Chandrasekhar, *Radiative Transfer* (Dover, New York, 1960); W. Irvine, Astrophys. J. **144**, 1140 (1966).

[2] G. N. Plass and G. W. Kattawar, Appl. Opt. **7**, 361, 415, and 699 (1968).

[3] The fraction of incident light reflected from the slab.

[4] J. R. Jokipii, Astrophys. J. **146**, 480 (1966); P. Noerdlinger, Phys. Rev. Letters **20**, 1513 (1968).

[5] W. Feller, *Probability Theory and its Applications* (Wiley, New York, 1950) Vol. I, pp. 238–246.

[6] Ref. 5, p. 297.

[7] Ref. 5, p. 52.

[8] It has been kindly pointed out to the author that similar results were obtained by other techniques, as in H. van de Hulst and F. G. Terhoeve, Bull. Astron. Neth. **18**, 377 (1966); V. Sobolev, Astron. Zh. **46**, 419 (1969) [Sov.Astron. AJ **13**, 330 (1969)].

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 5 MAY 1970

Explosive Instabilities in the Well-Defined Phase Description

H. WILHELMSSON AND L. STENFLO
Institute of Physics, University of Uppsala, Uppsala, Sweden
AND
F. ENGELMANN
Laboratori Gas Ionizzati, Associazione Euratom-CNEN, Frascati (Rome), Italy

(Received 14 July 1969)

The paper presents a general theoretical investigation of 3-wave interaction in the well-defined phase description with special regard to nonlinear explosive instabilities in the presence of linear damping or growth.

1. INTRODUCTION

In recent years, there has arisen an almost explosively increasing interest in the possibility of having unbounded solutions to nonlinear equations describing wave interaction in plasmas. The corresponding instabilities may be responsible, e.g., for certain astrophysical phenomena of catastrophic nature,[1] as well as for enhanced losses in some kinds of laboratory plasmas. Several specific examples have been considered.[2–5] They are alike in that there is free energy available to be converted into wave motion, expressed by the fact that one or two of the interacting waves have negative energy in any frame of reference. A typical example may be a plasma containing groups of particles with different drift velocities. Such a system may be linearly stable, but may exhibit nonlinear explosive instabilities. A case of this kind occurs, e.g., when a weak ion beam penetrates a nondrifting electron–ion plasma in a strong magnetic field.[3,4]

Nonlinear wave interaction has been investigated assuming well-defined phases of the waves[6–9] or by means of a random-phase approach.[3–5,10–12] The explosive case has been studied mainly using the latter method. It has been pointed out, however, that phase effects lead to a significant modification of details of the dynamics of the interaction.[13] It is the purpose of the present paper to discuss, within the framework of the well-defined phase description, the formal conditions under which a nonlinear 3-wave interaction is explosive, allowing also for the presence of linear damping of the interacting waves.

2. EQUATIONS OF MOTION

The equations of motion of waves of well-defined phases interacting nonlinearly can be conveniently deduced using a normal mode approach.[7,9] A linear analysis yields, for every normal mode j, a complex eigenfrequency ω_j, and a positive or negative Im (ω_j) corresponding to damping or growth, respectively. In the special case of a system which, in the absence of wave excitation, is homogeneous, the normal modes correspond to different wave vectors $\mathbf{k}_j$.

To lowest nonlinear order, the equations of motion contain bilinear interaction terms between normal modes. If a pure 3-wave interaction is possible, as, e.g., in the case of a homogeneous system in which waves with wave vectors $\mathbf{k}_j$, $j = 0, 1, 2$, satisfying the matching condition

$$\mathbf{k}_0 = \mathbf{k}_1 + \mathbf{k}_2,$$

are excited, one has the closed system of equations

$$\begin{aligned} \frac{\partial a_0}{\partial t} - i\omega_0 a_0 &= c_{12}^* a_1 a_2, \\ \frac{\partial a_1}{\partial t} - i\omega_1 a_1 &= c_{02} a_0 a_2^*, \\ \frac{\partial a_2}{\partial t} - i\omega_2 a_2 &= c_{01} a_0 a_1^*, \end{aligned} \tag{1}$$

where the a_j are the amplitudes of the normal modes and the star describes the complex conjugate. In the following, we assume that the frequencies of the interacting waves fulfill the matching condition

$$\text{Re}\,(\omega_0) = \text{Re}\,(\omega_1) + \text{Re}\,(\omega_2).$$

Explicit evaluations of the coupling constants c_{ij} show that, in the absence of linear damping or growth [Im $(\omega_j) = 0$], the c_{ij} can be taken to be real.[14] For longitudinal waves, as well as for transverse waves, their signs then depend on that of factors of the form $[\partial(\omega^2\epsilon_j)/\partial\omega]_{\omega=\omega_j}$, where $\epsilon_j(\omega)$ is the dielectric constant associated with the mode j. It is a characteristic feature of Eqs. (1) that the relative signs of the c_{ij} cannot be changed by changing the normalization of the a_j, i.e., by transformations of the kind $a_j \to \alpha_j a_j$ with α_j real. This is closely related to the fact that, for a suitable normalization of the a_j, the energy of a wave can be written

$$|a_j|^2\, \omega_j\, \text{sgn} \left(\frac{\partial(\omega^2\epsilon_j)}{\partial\omega}\right)_{\omega=\omega_j}.$$

In the presence of linear damping or growth [Im $(\omega_j \neq 0)$], the coupling coefficients are essentially complex. The special case of situations where collisional damping is important has been considered by other authors.[15]

Equations (1) can be reformulated in different ways. Introducing

$$a_j(t) = A_j(t)e^{i[\mathrm{Re}(\omega_j)]t},$$

and using also the matching condition for the frequencies, one obtains

$$\begin{aligned}
&\frac{\partial A_0}{\partial t} + [\mathrm{Im}\,(\omega_0)]A_0 = c_{12}^* A_1 A_2,\\
&\frac{\partial A_1}{\partial t} + [\mathrm{Im}\,(\omega_1)]A_1 = c_{02} A_0 A_2^*,\\
&\frac{\partial A_2}{\partial t} + [\mathrm{Im}\,(\omega_2)]A_2 = c_{01} A_0 A_1^*.
\end{aligned} \tag{2}$$

Note that $|\partial[\ln A_j(t)]/\partial t| \ll |\omega_j|$ holds if the linear damping or growth is small ($|\mathrm{Im}\,(\omega_j)| \ll |\mathrm{Re}\,(\omega_j)|$) and the nonlinear interaction is a small correction to the linear motion.

Often, it is more convenient to consider, rather than the $A_j \equiv |A_j|\, e^{i\phi_j}$, the real quantities

$$\begin{aligned}
u_0 &= (|c_{01}|\,|c_{02}|)^{\frac{1}{2}}|A_0|,\\
u_1 &= (|c_{01}|\,|c_{12}|)^{\frac{1}{2}}|A_1|,\\
u_2 &= (|c_{02}|\,|c_{12}|)^{\frac{1}{2}}|A_2|,
\end{aligned} \tag{3}$$

and

$$\Phi = \phi_0 - \phi_1 - \phi_2. \tag{4}$$

Putting $c_{ij} \equiv |c_{ij}|\, e^{i\theta_{ij}}$, one obtains, from Eq. (2),

$$\begin{aligned}
&\frac{\partial u_0}{\partial t} + [\mathrm{Im}\,(\omega_0)]u_0 = u_1 u_2 \cos(\Phi + \theta_{12}),\\
&\frac{\partial u_1}{\partial t} + [\mathrm{Im}\,(\omega_1)]u_1 = u_0 u_2 \cos(\Phi + \theta_{02}),\\
&\frac{\partial u_2}{\partial t} + [\mathrm{Im}\,(\omega_2)]u_2 = u_0 u_1 \cos(\Phi + \theta_{01}),
\end{aligned} \tag{5}$$

and

$$\begin{aligned}
\frac{\partial \Phi}{\partial t} = &-\frac{u_1 u_2}{u_0}\sin(\Phi + \theta_{12})\\
&-\frac{u_0 u_2}{u_1}\sin(\Phi + \theta_{02})\\
&-\frac{u_0 u_1}{u_2}\sin(\Phi + \theta_{01}).
\end{aligned} \tag{6}$$

For comparison, we note that in the random-phase description one has, instead, equations of the form

$$\frac{\partial u_j^2}{\partial t} + 2[\mathrm{Im}\,(\omega_j)]u_j^2 = \alpha_{j01}u_0^2u_1^2 + \alpha_{j02}u_0^2u_2^2 + \alpha_{j12}u_1^2u_2^2, \tag{7}$$

with the α_{jkl} being real and tending to ± 1 in the limit of vanishing linear damping or growth.[13]

3. DISCUSSION

In the absence of linear damping or growth [Im $(\omega_j) = 0$], the possibility of taking the coupling coefficients c_{ij} real implies $\theta_{ij} = 0, \pm\pi$. If all $\theta_{ij} = 0$ or all $\theta_{ij} = \pm\pi$, the system given by (5) and (6) has unbounded solutions, which diverge after a finite time t_∞ like $u_j \sim (t_\infty - t)^{-1}$, describing an "explosive" instability. In all other cases, the solutions of the system are oscillating and, hence, stable. In the random-phase approach the situation is analogous, the explosively unstable case corresponding to all $\alpha_{jkl} = +1$. However, the divergence is of lower order in this case, namely, $u_j \sim (\bar{t}_\infty - t)^{-\frac{1}{2}}$ where the time $\bar{t}_\infty$, after which the divergence occurs, is in general different from the corresponding time t_∞ in the well-defined phase description. Here exist, hence, important differences between the two complementary descriptions. It may be noted in this context that, if the waves have well-defined phases, an explosive instability may be preceded by a finite time interval during which the amplitudes of all interacting waves decrease, whereas this is impossible when the phases are random.[13] The case mentioned appears when, initially,

$$\cos(\Phi + \theta_{ij}) < 0$$

holds.

We now deduce a *necessary condition* for the presence of an explosive instability in the general case Im $(\omega_j) \neq 0$. Let us put

$$u_j = f_j(t)/(t_\infty - t), \tag{8}$$

where $f_j(t)$, depending on all parameters of the problem including the initial conditions, is positive and varies "slowly" in the neighborhood of the divergence of the u_j, assumed to occur for $t \to t_\infty$, i.e.,

$$\frac{d(\ln f_j)}{dt} \ll (t_\infty - t)^{-1}, \quad \text{for } t \to t_\infty. \tag{9}$$

Then Eqs. (5) yield

$$\begin{aligned}
f_0^2(t_\infty) &= [\cos(\Phi(t_\infty) + \theta_{02})\cos(\Phi(t_\infty) + \theta_{01})]^{-1},\\
f_1^2(t_\infty) &= [\cos(\Phi(t_\infty) + \theta_{01})\cos(\Phi(t_\infty) + \theta_{12})]^{-1},\\
f_2^2(t_\infty) &= [\cos(\Phi(t_\infty) + \theta_{02})\cos(\Phi(t_\infty) + \theta_{12})]^{-1},
\end{aligned} \tag{10a}$$

with the constraint that all

$$\cos(\Phi(t_\infty) + \theta_{ij}) > 0. \tag{10b}$$

Hence, the complex vectors having phase angles $\Phi(t_\infty) + \theta_{ij}$ all point in the right half-plane. This implies that, for an explosive instability to occur, it is

necessary that *the phases θ_{ij} of the coupling coefficients c_{ij} define complex vectors which all point in the same half-plane.*

Furthermore, from Eqs. (6), (8), and (10a), one obtains

$$\tan(\Phi(t_\infty) + \theta_{12}) + \tan(\Phi(t_\infty) + \theta_{02}) + \tan(\Phi(t_\infty) + \theta_{01}) = 0. \quad (10c)$$

This relation, together with the constraint (10b), determines the asymptotic phase $\Phi(t_\infty)$ uniquely.

More information about the time t_∞ can be obtained, considering the special case $\mathrm{Im}(\omega_j) = \nu$, where all $\mathrm{Im}(\omega_j)$ are equal. Using the transformation

$$\tau = \nu^{-1}(1 - e^{-\nu t}), \quad (11)$$

$$U_j = u_j e^{\nu t}, \quad (12)$$

Eqs. (5) and (6) then yield

$$\frac{\partial U_0}{\partial \tau} = U_1 U_2 \cos(\Phi + \theta_{12}),$$

$$\frac{\partial U_1}{\partial \tau} = U_0 U_2 \cos(\Phi + \theta_{02}), \quad (13)$$

$$\frac{\partial U_2}{\partial \tau} = U_0 U_1 \cos(\Phi + \theta_{01}),$$

and

$$\frac{\partial \Phi}{\partial \tau} = -\frac{U_1 U_2}{U_0}\sin(\Phi + \theta_{12}) - \frac{U_0 U_2}{U_1}\sin(\Phi + \theta_{02}) - \frac{U_0 U_1}{U_2}\sin(\Phi + \theta_{01}). \quad (14)$$

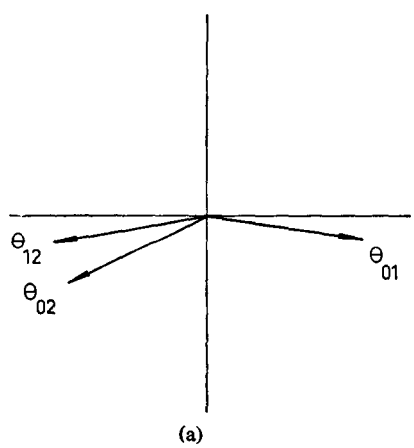
The preceding equations are identical with Eqs. (5) and (6) with the terms $\mathrm{Im}(\omega_j)u_j$ suppressed. Hence, it is possible to express t_∞ in terms of a fictitious $\hat{t}_\infty$, corresponding to an explosively unstable situation where the coupling coefficients are the same as in the actual problem, but the linear damping or growth is disregarded. Using Eq. (11), one obtains

$$t_\infty = \nu^{-1} \ln(1 - \nu \hat{t}_\infty)^{-1}. \quad (15)$$

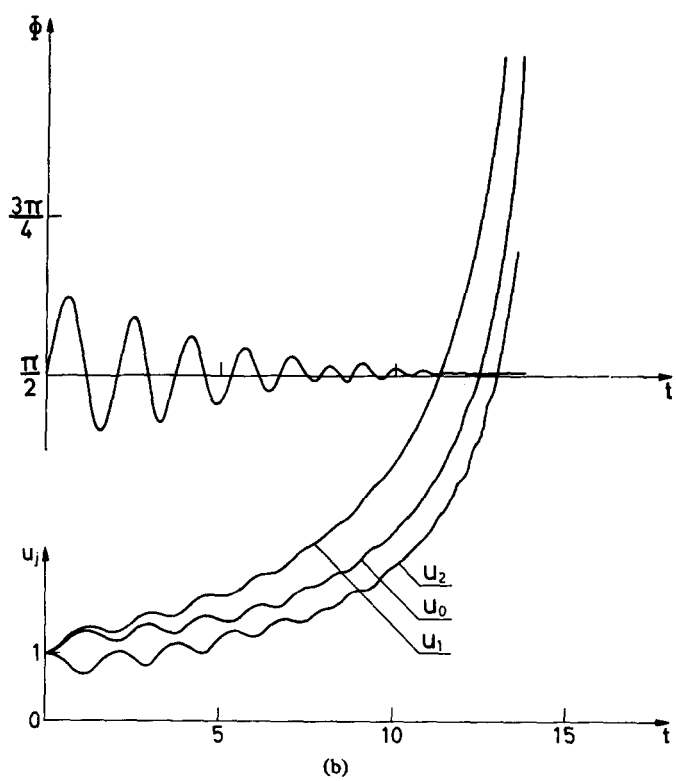
This shows that the tendency brought in by the term describing linear damping ($\nu > 0$) or growth ($\nu < 0$) is, in the first case, to increase t_∞, and, in the second, to decrease it, as is intuitive.

In particular, if ν is larger than a critical value given by $\nu_c = \hat{t}_\infty^{-1} > 0$, the time t_∞ is no longer real and, consequently, an explosive instability does not exist: it is suppressed by the effect of linear damping.[16] Physically, such a situation occurs when the initial excitation of the waves is so small that, at $t = 0$, the nonlinear terms are dominated by the terms describing linear damping. These results are analogous to those obtained by a random-phase approach.[17]

The opposite case $\nu < \nu_c$, which appears for higher initial excitations of the waves, comprises the interesting possibility that a plasma, which in the absence of dissipation would be stable, becomes nonlinearly (i.e., explosively) unstable by the effect of even a weak



(a)



(b)

FIG. 1(a). Vectors $e^{i\theta_{ij}}$ in the same half-plane (qualitative plot for destabilized case). (b) Time dependence of the wave amplitudes u_j and the phase Φ for $\theta_{12} = -\pi + \pi/60$; $\theta_{02} = -\pi + \pi/20$; $\theta_{01} = -\pi/60$, and $\nu = 0.01$.

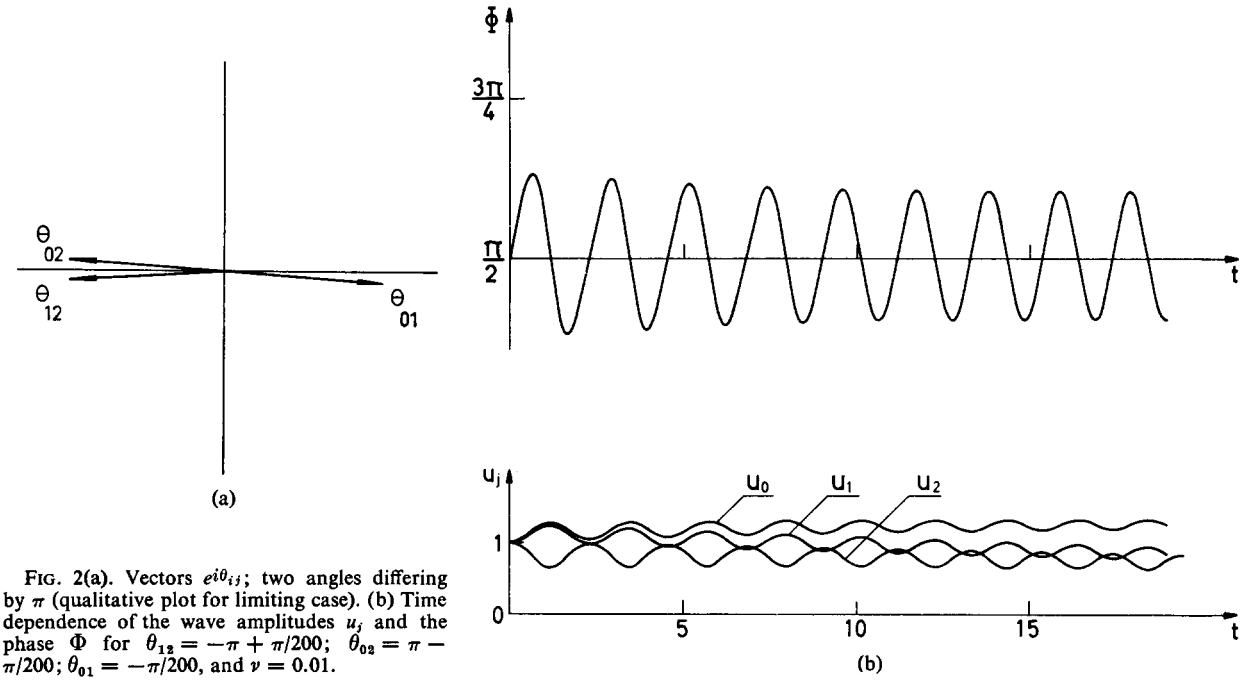


FIG. 2(a). Vectors $e^{i\theta_{ij}}$; two angles differing by π (qualitative plot for limiting case). (b) Time dependence of the wave amplitudes u_j and the phase Φ for $\theta_{12} = -\pi + \pi/200$; $\theta_{02} = \pi - \pi/200$; $\theta_{01} = -\pi/200$, and $\nu = 0.01$.

dissipation.[18] This occurs when the dissipative effects modify the phases θ_{ij} in such a way that the corresponding complex vectors all point in the same half-plane (cf. Fig. 1), the limiting case of vanishing dissipation corresponding to a situation where two θ_{ij} are equal to 0, and the remaining one equal to $\pm\pi$, or vice versa.

The three typical cases which can appear according to the relative values of θ_{ij} are reported in Figs. 1–3. The given time dependencies of the wave amplitudes u_j and the phase Φ have been obtained by numerical integration of Eqs. (5) and (6) in the limit where the effect of linear damping is unimportant. All cases with $e^{i\theta_{ij}}$, being all in the same half-plane which have been

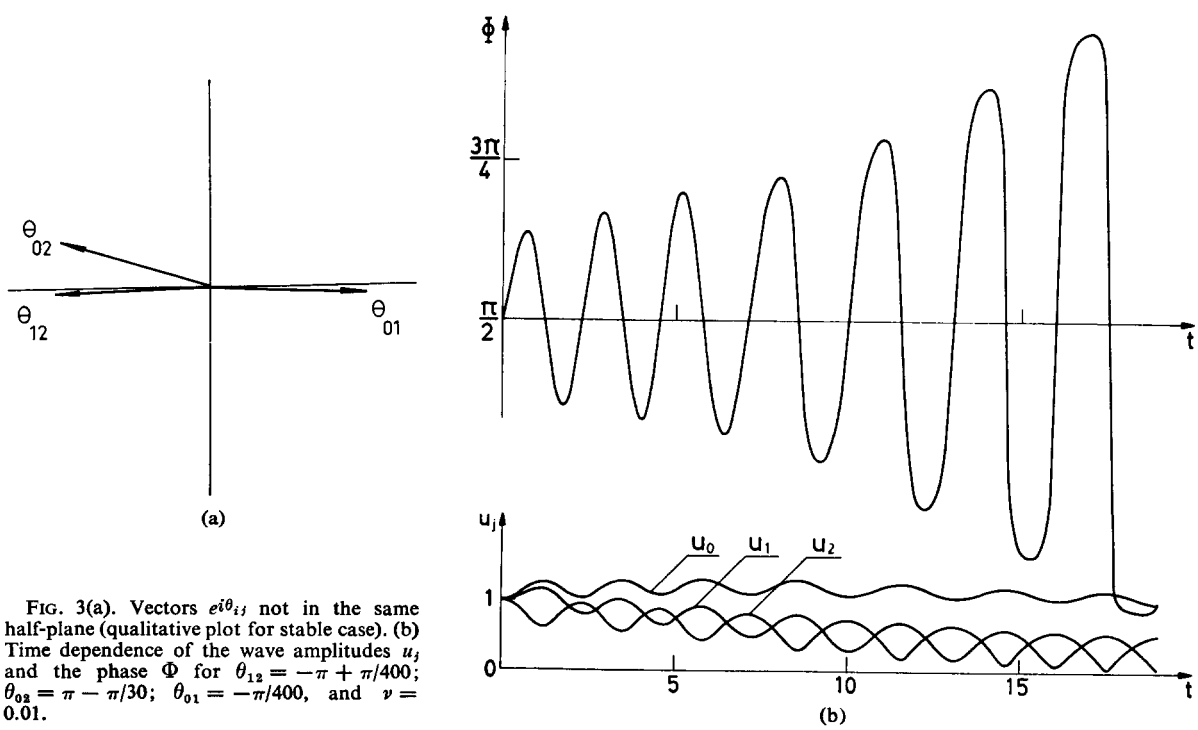


FIG. 3(a). Vectors $e^{i\theta_{ij}}$ not in the same half-plane (qualitative plot for stable case). (b) Time dependence of the wave amplitudes u_j and the phase Φ for $\theta_{12} = -\pi + \pi/400$; $\theta_{02} = \pi - \pi/30$; $\theta_{01} = -\pi/400$, and $\nu = 0.01$.

treated numerically, were explosively unstable. Therefore, the necessary condition for explosive instability given above seems sufficient for these cases, also.

We note that, for $\theta_{ij} = 0, \pm\pi$, Eqs. (5) and (6) can also be integrated analytically [if all Im (ω_j) are equal]. This is readily seen from Eqs. (13) and (14), which then become formally identical with the equations of motion in the absence of linear damping or growth [Im $(\omega_j) = 0$] and, hence, have simple integrals.[6,8,13] In the general case where $\theta_{ij} \neq 0, \pm\pi$, the problem becomes considerably more complicated, though one has the new constant of motion

$$U_0^2 \sin(\theta_{02} - \theta_{01}) + U_1^2 \sin(\theta_{01} - \theta_{12}) + U_2^2 \sin(\theta_{12} - \theta_{02}). \quad (16)$$

Generalizations of the constants of motion, previously known for the case Im $(\omega_j) = 0$, have not been found, however.

ACKNOWLEDGMENTS

This work has been supported by the Swedish Board for Technical Development. One of the authors (H. W.) acknowledges the Swedish Atomic Research Council for travelling grants to visit Laboratori Gas Ionizzati, Frascati (Rome). The authors' thanks are due to Dr. A. Jarmén for his aid in handling the computer calculations.

[1] P. A. Sturrock, Phys. Rev. Letters **16**, 270 (1966).
[2] B. B. Kadomtsev, A. B. Mikhailovskii, and A. V. Timofeev, Zh. Eksp. Teor. Fiz. **47**, 2266 (1964) [Sov. Phys.—JETP **20**, 1517 (1965)].
[3] V. M. Dikasov, L. I. Rudakov, and D. D. Ryutov, Zh. Eksp. Teor. Fiz. **48**, 913 (1965) [Sov. Phys.—JETP **21**, 608 (1965)].
[4] R. E. Aamodt and M. Sloan, Phys. Rev. Letters **19**, 1227 (1967); Phys. Fluids **11**, 2218 (1968).
[5] M. N. Rosenbluth, B. Coppi, and R. N. Sudan, Third Conference on Plasma Physics and Controlled Fusion Research, Novosibirsk Paper CN-24/E-13, 1968.
[6] J. A. Armstrong, N. Bloembergen, J. Ducuing, and P. S. Pershan, Phys. Rev. **127**, 1918 (1962).
[7] A. Sjölund and L. Stenflo, Physica **35**, 499 (1967); Z. Physik **204**, 211 (1967).
[8] R. Sugihara, Phys. Fluids **11**, 178 (1968).
[9] H. Wilhelmsson, J. Plasma Phys. **3**, 215 (1969); Proceedings of the Conference on Fusion and Plasma Physics, Beam-Plasma Interactions, Utrecht, 1969.
[10] V. N. Tsytovich, Usp. Fiz. Nauk **90**, 435 (1966) [Sov. Phys.—Usp. **9**, 805 (1967)].
[11] M. Bornatici, A. Cavaliere, and F. Engelmann, Phys. Fluids **12**, 2362 (1969).
[12] R. C. Davidson and A. N. Kaufman, J. Plasma Phys. **3**, 97 (1969).
[13] F. Engelmann and H. Wilhelmsson, Z. Naturforsch. **24a**, 206 (1969).
[14] See, e.g., a recent work by Wilhelmsson (Ref. 9), where the case of a 1-dimensional plasma, containing groups of particles with different drift velocities, is considered. The result can also be deduced from the postulate that in the interaction the numbers of "quanta" corresponding to each wave change by the same amount.
[15] V. G. Makhankov and V. N. Tsytovich, Zh. Tekh. Fiz. **37**, 1381 (1967) [Sov. Phys.—Tech. Phys. **12**, 1003 (1968)]; Zh. Eksp. Teor. Fiz. **53**, 1789 (1967) [Sov. Phys.—JETP **26**, 1023 (1968)]; L. Stenflo, Plasma Phys. (1970), in press.
[16] Coupling between more than three waves may introduce explosive instabilities for which the wave amplitudes diverge, like $u_j \sim (t_\infty^{(p)} - t)^p$ with $-1 < p < 0$. The stabilization of such cases may be determined in a similar way. See also Ref. 17.
[17] L. Stenflo and H. Wilhelmsson, Phys. Letters **29A**, 217 (1969).
[18] A. Jarmén, L. Stenflo, H. Wilhelmsson, and F. Engelmann, Phys. Letters **28A**, 748 (1969).

Application of the Method of Matched Asymptotic Expansions to a Problem in Linear Transport Theory*

W. L. HENDRY
Los Alamos Scientific Laboratory, University of California, Los Alamos, New Mexico 87544

(Received 18 December 1969)

The method of matched asymptotic expansions is used to reduce the time-dependent, one-velocity neutron transport equation to a set of more tractable equations. This reduction is accomplished subject to general initial conditions on both the neutron flux and delayed neutron precursors.

1. INTRODUCTION

The use of matched asymptotic expansions in solving singular perturbation problems is a well-known technique in fluid mechanics.[1-3] There the equations are usually nonlinear, and the goal is often an analytic approximation to the solution that is uniformly valid in some independent variable of the problem. By a uniformly valid solution, we mean that the absolute value of the difference between the exact solution and the approximate solution has a uniform bound (with respect to an independent variable), and this bound can be made as small as we please by choosing some characteristic parameter of the system sufficiently close to a limit, usually zero or infinity. The approximation is thus valid in an asymptotic sense.

In this paper, the method of matched asymptotic expansions is applied to the initial value problem for the time-dependent, one-velocity neutron transport equation with delayed neutrons. The analysis is unusual in the sense that the problem is linear, and the goal is not an analytic approximation to the exact solution (which seems to be out of the question for any realistic formulation of the problem), but is rather the reduction of the transport equation to a set of asymptotically equivalent equations that are more easily solved on a digital computer.

Very little exists in the way of a rigorous foundation for the method of matched asymptotic expansions (see the book by Kaplun[3] for some first steps in this direction), and the analysis which follows is largely formal.

Our goal is solutions to the following system of equations:

$$\epsilon \frac{\partial \psi}{\partial t} + B\psi = \frac{1}{4\pi} \sum_{i=1}^{N} \lambda_i C_i, \tag{1.1}$$

$$\frac{\partial C_i}{\partial t} + \lambda_i C_i = \beta_i K\psi, \quad i = 1, \cdots, N, \tag{1.2}$$

for the neutron distribution function $\psi(\mathbf{r}, \mathbf{\Omega}, t)$ and delayed neutron precursor densities $C_i(\mathbf{r}, t)$. Here $\mathbf{r} = (x_1, x_2, x_3)$ is a vector specifying the position in a convex bounded volume V with surface S, within which the solutions to (1.1) and (1.2) are sought; $\mathbf{\Omega}$ is a unit vector pointing in the direction of the neutron velocity; t is time. The magnitude of the neutron speed is assumed constant and, because it is usually quite large, is denoted by $v = 1/\epsilon$. B, L, S, and K are operators defined by the relations

$$B = L - S, \tag{1.3a}$$

$$L\psi = \mathbf{\Omega} \cdot \nabla\psi + \Sigma(\mathbf{r})\psi, \tag{1.3b}$$

$$S\psi = \int_{\Omega} d\mathbf{\Omega}' \left(\Sigma_s(\mathbf{r}; \mathbf{\Omega}' \cdot \mathbf{\Omega}) + \frac{1}{4\pi}(1 - \beta)\nu\Sigma_f(\mathbf{r}) \right) \psi, \tag{1.3c}$$

$$K\psi = \nu\Sigma_f(\mathbf{r}) \int_{\Omega} d\mathbf{\Omega}' \psi. \tag{1.3d}$$

The symbol $d\mathbf{\Omega}$ is shorthand for a differential element of solid angle, and the integrals are understood to run over the unit sphere. We denote the range of the variable $\mathbf{\Omega}$ by Ω. The neutron distribution function $\psi(\mathbf{r}, \mathbf{\Omega}, t)$ is defined to be the speed times the time-dependent density of neutrons in the 5-dimensional phase space $(\mathbf{r}, \mathbf{\Omega})$. The delayed neutron precursor densities $C_i(\mathbf{r}, t)$ may be looked upon as the time- and space-dependent densities in the volume V of the ith $(i = 1, \cdots, N)$ isotope produced in fission and decaying with the emission of one or more delayed neutrons (actually, C_i is a group of such isotopes sharing approximately the same half-life). Typically, $N \leq 6$. Other quantities appearing in the above equations are the total and fission cross sections Σ and Σ_f (reciprocals of the corresponding mean free paths); the scattering kernel $\Sigma_s(\mathbf{r}; \mathbf{\Omega}' \cdot \mathbf{\Omega})$, defined such that $\Sigma_s(\mathbf{r}; \mathbf{\Omega}' \cdot \mathbf{\Omega})\, d\mathbf{\Omega}$ is the reciprocal of the mean free path between scattering collisions of a neutron at $\mathbf{r}$ with velocity direction $\mathbf{\Omega}'$ which sends the neutron into $d\mathbf{\Omega}$ about $\mathbf{\Omega}$; the fraction of fission neutrons produced by the ith delayed neutron precursor β_i; $\beta = \beta_1 + \cdots + \beta_N$; the total number of neutrons released per fission ν; and the delayed neutron precursor decay constants λ_i.

Equation (1.1) expresses a balance between neutrons being lost from an element of phase space by collisions and streaming and being added by in-scattering, prompt fission, and decay of delayed neutron precursors. Equation (1.2) is the usual radioactive decay relation with a source term due to fissions. We shall seek solutions to (1.1) and (1.2) subject to the boundary condition of no incoming neutrons

$$\psi(\mathbf{r}, \mathbf{\Omega}, t) = 0, \quad \mathbf{r} \in S, \quad \mathbf{n} \cdot \mathbf{\Omega} < 0, \tag{1.4}$$

where $\mathbf{n}$ is the outward normal to S. We shall also specify initial conditions on both ψ and the C_i, namely

$$\psi(\mathbf{r}, \mathbf{\Omega}, 0) = f(\mathbf{r}, \mathbf{\Omega}), \tag{1.5}$$

$$C_i(\mathbf{r}, 0) = g_i(\mathbf{r}), \quad i = 1, \cdots, N. \tag{1.6}$$

The last condition generalizes earlier work,[4] in which the formalism made essential use of the conditions $g_i(\mathbf{r}) = 0$, $i = 1, \cdots, N$. For a proof of the existence and uniqueness of positive solutions to (1.1) and (1.2) subject to (1.4), (1.5), and (1.6) in the more general velocity-dependent case, we refer the reader to a paper by Marti.[5]

The formal perturbation method described below involves the expansion of ψ and the C_i in powers of ϵ and the assumption that these expansions are valid asymptotically, as $\epsilon \to 0$. The parameter ϵ is equal to the reciprocal of the neutron speed, and for most systems of practical interest is much smaller than the other factors appearing in (1.1) and (1.2). This approach can be applied to the velocity-dependent case with no changes whatever, but in this case the factor $1/v$ is a variable and it becomes necessary to *prefix* the time derivative with a fictitious parameter ϵ and let $\epsilon \to 1$ at the end of the analysis. This is a procedure similar to that used in the Chapman–Enskog–Hilbert expansion in statistical mechanics.[6] In principle, it should be possible to scale (1.1) and (1.2) in the velocity-dependent case by introducing a lower bound on the range of v and using this lower bound to make the equations dimensionless with respect to velocity. The difficulty with this approach is that in the velocity-dependent case the operators B and K involve integrals over v and the scaling process causes cross sections to become functions of ϵ. The form that cross section data usually takes makes the required asymptotic expansions in ϵ unobtainable. To keep the analysis presented here on as firm ground as possible, we limit the discussion to the one-velocity case.

2. KNOWN RESULTS

The method of matched asymptotic expansions, as we shall apply it below, makes considerable use of *a priori* knowledge of the transport equation and its solutions; in fact, the method almost amounts to a transformation of qualitative information into quantitative information. If the general features of the required solutions are known, then the method can yield very accurate results. Numerical tests have been made[4,7] that indicate the results in the present case are essentially exact. Further verification may be found in Sec. 4, where it is shown that in at least one case the mathematical structure of our solution agrees with that given by exact analysis.

The operator B has been studied in considerable generality by Vladimirov[8] (many of the results we shall attribute to Vladimirov were stated by him in much more general form). His results confirmed what has been "known" by reactor physicists for a long time, namely that positive bounded solutions to the steady state equation

$$(L - \gamma S)\varphi_\gamma(\mathbf{r}, \mathbf{\Omega}) = Q(\mathbf{r}, \mathbf{\Omega}) \tag{2.1}$$

exist for $Q > 0$ in a bounded volume surrounded by vacuum as long as $0 \leq \gamma < \gamma_0$. The number γ_0 is the first "criticality" eigenvalue for the system. It is simple, and to it there corresponds an eigenfunction φ_0 satisfying the homogeneous equation

$$(L - \gamma_0 S)\varphi_0(\mathbf{r}, \mathbf{\Omega}) = 0, \tag{2.2}$$

which is unique to within multiplication by a scalar. These facts were proven subject to restrictions that in effect precluded the possibility $S = 0$, and for a linear manifold of differentiable functions in the Hilbert space h of functions square integrable in the Lebesque sense on the set of all $\mathbf{r} \in V$, $\mathbf{\Omega} \in \Omega$.

Equation (2.2) describes a self-sustaining nuclear reaction or a "prompt critical" nuclear reactor, neglecting delayed neutrons and feedback effects. The word "prompt" indicates that criticality is achieved with prompt neutrons alone. Furthermore, there exists a denumerably infinite set of eigenvalues γ_k and eigenfunctions φ_k satisfying

$$(L - \gamma_k S)\varphi_k(\mathbf{r}, \mathbf{\Omega}) = 0. \tag{2.3}$$

The set of functions $\{S\varphi_k\}$ is complete in the range of S (in particular, they are complete for functions of r only), but the set $\{\varphi_k\}$ is not complete in h. The inverse operator $(L - \gamma S)^{-1}$ exists and is bounded on h for $\gamma \neq \gamma_k$, $k = 1, 2, \cdots$.

No equally complete theory exists for the time-dependent equation, but the following observations have been made for many applications of the transport equation to simple systems[9–12] and for various approximations to the neutron transport equation.[13] Below

prompt critical ($\gamma < \gamma_0$), solutions to the equation

$$\epsilon \frac{\partial \varphi}{\partial t} + (L - \gamma S)\varphi = 0 \tag{2.4}$$

(subject to a given initial distribution in $\mathbf{r}$ and $\mathbf{\Omega}$ for φ and the same boundary conditions as before) decay at least exponentially to zero with increasing time. At prompt critical they settle down to a positive steady state solution. Above prompt critical they grow with time. Solutions to (1.1) and (1.2) may grow for systems below prompt critical (when we choose the parameters in the operator S in such a manner that $\gamma = 1 < \gamma_0$) because of the contribution of the delayed neutrons. The dividing line between solutions to (1.1) and (1.2) which grow with time or decay with time is called "delayed critical."

3. MATCHED ASYMPTOTIC EXPANSIONS

The philosophy underlying the application of the method of matched asymptotic expansions to this problem has been discussed before,[4] and we will stress here only those points which differ from the analysis presented earlier. The short-time and long-time solutions to (1.1) and (1.2) are denoted by ψ^{I}, C_i^{I} and ψ^{III}, C_i^{III}, respectively. These solutions have a common time range of validity, and within this overlap the solutions are denoted by ψ^{II}, C_i^{II}. We shall also refer to the solutions in the three time ranges I, II, and III as inner, intermediate, and outer solutions, respectively. Evidently,

$$\psi = \psi^{\mathrm{I}} + \psi^{\mathrm{III}} - \psi^{\mathrm{II}}, \tag{3.1a}$$

$$C_i = C_i^{\mathrm{I}} + C_i^{\mathrm{III}} - C_i^{\mathrm{II}}, \quad i = 1, \cdots, N, \tag{3.1b}$$

since addition of the inner to the outer solution gives two times the correct solution in the intermediate region.

Our perturbation method breaks down in the neighborhood of prompt criticality. This is because the operator $L - \gamma_0 S$ is "small" on that region of Hilbert space which is close to the first criticality eigenfunction. Consequently, the parameter ϵ, which is given beforehand in practical problems, cannot be considered small compared to other terms. The smaller the ϵ, the closer prompt criticality can be approached. The perturbation method works for all systems sufficiently below prompt critical and, in particular, describes systems both at and above delayed critical. It works above prompt critical in the trivial sense that the inner solution dominates for all time. This is of little interest [because there is little advantage in solving the inner equations instead of the full set of transport equations, (1.1) and (1.2)], and further analysis will be based on the assumption that the parameters in the operator S are chosen such that $\gamma = 1 < \gamma_0$.

Inner solution: In order to satisfy the initial condition on ψ, the time coordinate is stretched in such a manner that the time derivative is retained in the zeroth-order perturbation theory. We put $\tau = t/\epsilon$, and substitute into (1.1) and (1.2). The result is

$$\left(\frac{\partial}{\partial \tau} + B\right)\psi^{\mathrm{I}}(\mathbf{r}, \mathbf{\Omega}, \tau) = \frac{1}{4\pi}\sum_{i=1}^{N} \lambda_i C_i^{\mathrm{I}}(\mathbf{r}, \tau), \tag{3.2}$$

$$\left(\frac{\partial}{\partial \tau} + \epsilon\lambda_i\right)C_i^{\mathrm{I}}(\mathbf{r}, \tau) = \epsilon\beta_i(K\psi^{\mathrm{I}})(\mathbf{r}, \tau). \tag{3.3}$$

Now, we introduce the expansions

$$\psi^{\mathrm{I}}(\mathbf{r}, \mathbf{\Omega}, \tau) = \psi_0^{\mathrm{I}}(\mathbf{r}, \mathbf{\Omega}, \tau) + \epsilon\psi_1^{\mathrm{I}}(\mathbf{r}, \mathbf{\Omega}, \tau) + O(\epsilon^2), \tag{3.4}$$

$$C_i^{\mathrm{I}}(\mathbf{r}, \tau) = C_{i0}^{\mathrm{I}}(\mathbf{r}, \tau) + \epsilon C_{i1}^{\mathrm{I}}(\mathbf{r}, \tau) + O(\epsilon^2), \quad i = 1, \cdots, N, \tag{3.5}$$

into (3.2) and (3.3), equate equal powers of ϵ, and neglect terms $O(\epsilon^2)$. This leads to the following equations:

$$\left(\frac{\partial}{\partial \tau} + B\right)\psi_0^{\mathrm{I}}(\mathbf{r}, \mathbf{\Omega}, \tau) = \frac{1}{4\pi}\sum_{i=1}^{N} \lambda_i C_{i0}^{\mathrm{I}}(\mathbf{r}, \tau), \tag{3.6}$$

$$\left(\frac{\partial}{\partial \tau} + B\right)\psi_1^{\mathrm{I}}(\mathbf{r}, \mathbf{\Omega}, \tau) = \frac{1}{4\pi}\sum_{i=1}^{N} \lambda_i C_{i1}^{\mathrm{I}}(\mathbf{r}, \tau), \tag{3.7}$$

$$\frac{\partial}{\partial \tau} C_{i0}^{\mathrm{I}}(r, \tau) = 0, \quad i = 1, \cdots, N, \tag{3.8}$$

$$\frac{\partial}{\partial \tau} C_{i1}^{\mathrm{I}}(\mathbf{r}, \tau) = \beta_i(K\psi_0^{\mathrm{I}})(\mathbf{r}, \tau) - \lambda_i C_{i0}^{\mathrm{I}}(\mathbf{r}, \tau), \quad i = 1, \cdots, N. \tag{3.9}$$

These are to be solved subject to the initial conditions

$$\psi_0^{\mathrm{I}}(\mathbf{r}, \mathbf{\Omega}, 0) = f(\mathbf{r}, \mathbf{\Omega}), \tag{3.10}$$

$$\psi_j^{\mathrm{I}}(\mathbf{r}, \mathbf{\Omega}, 0) = 0, \quad j > 0, \tag{3.11}$$

$$C_{i0}^{\mathrm{I}}(\mathbf{r}, 0) = g_i(\mathbf{r}), \tag{3.12}$$

$$C_{ij}^{\mathrm{I}}(\mathbf{r}, 0) = 0, \quad j > 0, \tag{3.13}$$

which result from substitution of (3.4) and (3.5) into (1.5) and (1.6). Equations (3.8) and (3.9) may be integrated directly:

$$C_{i0}^{\mathrm{I}}(\mathbf{r}, \tau) = g_i(\mathbf{r}), \quad i = 1, \cdots, N, \tag{3.14}$$

$$C_{i1}^{\mathrm{I}}(\mathbf{r}, \tau) = \beta_i \int_0^{\tau} (K\psi_0^{\mathrm{I}})(\mathbf{r}, \tau')\, d\tau' - \lambda_i g_i(\mathbf{r})\tau. \tag{3.15}$$

The rhs of (3.6) is thus a known function. To solve (3.6), we break ψ_0^{I} into the sum of two functions, a solution to the homogeneous form of (3.6), and a particular solution to the inhomogeneous equation.

The need for this manipulation will become evident when the outer solution is matched to the inner solution for intermediate times. We put

$$\psi_0^{\mathrm{I}} = \psi_{0h}^{\mathrm{I}} + \psi_{0p}^{\mathrm{I}}, \tag{3.16}$$

where ψ_{0p}^{I} is the solution to the steady-state equation

$$B\psi_{0p}^{\mathrm{I}}(\mathbf{r}, \boldsymbol{\Omega}) = \frac{1}{4\pi}\sum_{i=1}^{N} \lambda_i g_i(\mathbf{r}) \tag{3.17}$$

and ψ_{0h}^{I} is the solution to the time-dependent equation

$$\left(\frac{\partial}{\partial\tau} + B\right)\psi_{0h}^{\mathrm{I}}(\mathbf{r}, \boldsymbol{\Omega}, \tau) = 0, \tag{3.18}$$

subject to the initial condition

$$\psi_{0h}^{\mathrm{I}}(\mathbf{r}, \boldsymbol{\Omega}, 0) = f(\mathbf{r}, \boldsymbol{\Omega}) - \psi_{0p}^{\mathrm{I}}(\mathbf{r}, \boldsymbol{\Omega}). \tag{3.19}$$

That the ψ_0^{I} so defined solves (3.6) subject to (3.10) follows immediately from the linearity of B. Since we have confined our attention to systems below prompt critical, the operator inverse to B exists and is bounded, and the solution to (3.17) may be written

$$\psi_{0p}^{\mathrm{I}}(\mathbf{r}, \boldsymbol{\Omega}) = \frac{1}{4\pi}\sum_{i=1}^{N} \lambda_i B^{-1} g_i(\mathbf{r}). \tag{3.20}$$

The operator B^{-1} may be found for slabs with the help of Case's method.[12] For more complex systems one must generally resort to approximate representations of B. Computer codes that can invert approximate representations of B exist for a variety of systems, and the principal contribution of this analysis is to show how time-dependent solutions to the transport equations may be obtained (in the intermediate and outer regions) by solving several steady state problems.

Substituting from (3.16) into (3.15) and making use of (3.20), we have

$$C_{i1}^{\mathrm{I}}(\mathbf{r}, \boldsymbol{\Omega}) = \int_0^{\tau} \beta_i K\psi_{0h}^{\mathrm{I}}(\mathbf{r}, \boldsymbol{\Omega}, \tau')\, d\tau' + \left(\frac{\beta_i}{4\pi}\sum_{j=1}^{N} \lambda_i K B^{-1} g_i - \lambda_i g_i\right)\tau. \tag{3.21}$$

The rhs of (3.7) is now known. Again, to facilitate the matching procedure described below, the solution to (3.7) is written as the sum of three particular solutions

$$\psi_1^{I}(\mathbf{r}, \boldsymbol{\Omega}, \tau) = \psi_{1p}^{\mathrm{I}}(\mathbf{r}, \boldsymbol{\Omega}, \tau) + \psi_{1r}^{\mathrm{I}}(\mathbf{r}, \boldsymbol{\Omega}) + \psi_{1s}^{\mathrm{I}}(\mathbf{r}, \boldsymbol{\Omega}, \tau), \tag{3.22}$$

where ψ_{1p}^{I} and ψ_{1r}^{I} may be written formally

$$\psi_{1p}^{\mathrm{I}}(\mathbf{r}, \boldsymbol{\Omega}, \tau) = \frac{1}{4\pi}\sum_{i=1}^{N} \lambda_i \left(\int_0^{\tau} \beta_i B^{-1} K\psi_{0h}^{\mathrm{I}}(\mathbf{r}, \boldsymbol{\Omega}, \tau')\, d\tau' + \frac{\beta_i \tau}{4\pi}\sum_{j=1}^{N} \lambda_j B^{-1} K B^{-1} g_j - \lambda_i B^{-1} g_i \tau\right), \tag{3.23}$$

$$\psi_{1r}^{\mathrm{I}}(\mathbf{r}, \boldsymbol{\Omega}) = -\frac{1}{4\pi}\sum_{i=1}^{N} \lambda_i \left(\frac{\beta_i}{4\pi}\sum_{j=1}^{N} \lambda_j B^{-2} K B^{-1} g_j - \lambda_i B^{-2} g_i\right), \tag{3.24}$$

and ψ_{1s}^{I} is the solution to

$$\left(\frac{\partial}{\partial\tau} + B\right)\psi_{1s}^{\mathrm{I}}(\mathbf{r}, \boldsymbol{\Omega}, \tau) = -\frac{1}{4\pi}\sum_{i=1}^{N} \lambda_i \beta_i B^{-1} K\psi_{0h}^{\mathrm{I}}(\mathbf{r}, \boldsymbol{\Omega}, \tau), \tag{3.25}$$

subject to the initial condition

$$\psi_{1s}^{\mathrm{I}}(\mathbf{r}, \boldsymbol{\Omega}, 0) = \frac{1}{4\pi}\sum_{i=1}^{N} \lambda_i \left(\frac{\beta_i}{4\pi}\sum_{j=1}^{N} \lambda_j B^{-2} K B^{-1} g_j - \lambda_i B^{-2} g_i\right). \tag{3.26}$$

That (3.22) satisfies (3.7) subject to (3.11) is easily verified by operating on both sides of (3.22) with $\partial/\partial\tau + B$ and using (3.23)–(3.26).

Outer solution: We introduce the expansions

$$\psi^{\mathrm{III}}(\mathbf{r}, \boldsymbol{\Omega}, t) = \psi_0^{\mathrm{III}}(\mathbf{r}, \boldsymbol{\Omega}, t) + \epsilon\psi_1^{\mathrm{III}}(\mathbf{r}, \boldsymbol{\Omega}, t) + O(\epsilon^2), \tag{3.27}$$

$$C_i^{\mathrm{III}}(\mathbf{r}, t) = C_{i0}^{\mathrm{III}}(\mathbf{r}, t) + \epsilon C_{i1}^{\mathrm{III}}(\mathbf{r}, t) + Q(\epsilon^2) \tag{3.28}$$

into (1.1) and (1.2), equate equal powers of ϵ, and neglect terms $O(\epsilon^2)$:

$$B\psi_0^{\mathrm{III}}(\mathbf{r}, \boldsymbol{\Omega}, t) = \frac{1}{4\pi}\sum_{i=1}^{N} \lambda_i C_{i0}^{\mathrm{III}}(\mathbf{r}, t), \tag{3.29}$$

$$B\psi_1^{\mathrm{III}}(\mathbf{r}, \boldsymbol{\Omega}, t) = \frac{1}{4\pi}\sum_{i=1}^{N} \lambda_i C_{i1}^{\mathrm{III}}(\mathbf{r}, t) - \frac{\partial\psi_0^{\mathrm{III}}(\mathbf{r}, \boldsymbol{\Omega}, t)}{\partial t}, \tag{3.30}$$

$$\left(\frac{\partial}{\partial t} + \lambda_i\right) C_{ij}^{\mathrm{III}}(\mathbf{r}, t) = \beta_i (K\psi_j^{\mathrm{III}})(\mathbf{r}, t), \quad j = 0, 1. \tag{3.31}$$

The solution to (3.31) may be written immediately:

$$C_{ij}^{\mathrm{III}}(\mathbf{r}, t) = C_{ij}^{\mathrm{III}}(\mathbf{r}, 0)\exp(-\lambda_i t) + \beta_i \int_0^t \exp[-\lambda_i(t - t')](K\psi_j^{\mathrm{III}})(\mathbf{r}, t')\, dt', \quad j = 0, 1. \tag{3.32}$$

Note that numerical solution of (3.30) does not necessitate numerical differentiation of ψ_0^{III}; $\partial\psi_0^{\mathrm{III}}/\partial t$

may be obtained in terms of ψ_0^{III} and its integrals by differentiating (3.29), with the result

$$\frac{\partial \psi_0^{III}(\mathbf{r}, \mathbf{\Omega}, t)}{\partial t} = -\frac{1}{4\pi}\sum_{i=1}^{N}\lambda_i\Big(B^{-1}C_{i0}^{III}(\mathbf{r}, 0)\lambda_i \exp(-\lambda_i t) - \beta_i B^{-1}K\psi_0^{III}(\mathbf{r}, \mathbf{\Omega}, t) + \lambda_i\beta_i\int_0^t \exp[-\lambda_i(t-t')]B^{-1} \times K\psi_0^{III}(\mathbf{r}, \mathbf{\Omega}, t')\,dt'\Big). \quad (3.33)$$

Matching: The arbitrary functions $C_{ij}^{III}(\mathbf{r}, 0)$ are fixed by the matching process. The matching is accomplished by forcing the short-time outer solution (expanded in a power series about $t = 0$) to agree with the long-time inner solution. That portion of the solutions held in common is the intermediate solution. It is in determining the long-time behavior of the inner solution that the virtues of the decomposition into particular solutions becomes apparent, for the behavior of these solutions can be seen by inspection. Because the system is assumed to be below prompt critical, both ψ_{0h}^{I} and $\psi_{1s}^{I} \to 0$ for large τ. Both ψ_{0p}^{I} and $\psi_{1\tau}^{I}$ are constant in τ and ψ_{1p}^{I} grows like τ. Moreover, the integral $\int_0^\infty \psi_{0h}^{I}(\mathbf{r}, \mathbf{\Omega}, \tau)\,d\tau$ exists because of the exponential decrease of ψ_{0h}^{I}. Making use of (3.19), we may integrate (3.18) over all time to obtain

$$\int_0^\infty \beta_i K\psi_{0h}^{I}(\mathbf{r}, \mathbf{\Omega}, \tau)\,d\tau = \beta_i KB^{-1}\Big(f(\mathbf{r}, \mathbf{\Omega}) - \frac{1}{4\pi}\sum_{j=1}^{N}\lambda_j B^{-1}g_j(\mathbf{r})\Big). \quad (3.34)$$

The long-time inner solution may therefore be written

$$\psi^{I}(\mathbf{r}, \mathbf{\Omega}, \tau) \simeq \frac{1}{4\pi}\sum_{i=1}^{N}\lambda_i B^{-1}g_i(\mathbf{r}) + \frac{\epsilon}{4\pi}\sum_{i=1}^{N}\lambda_i\Big[\beta_i B^{-1}KB^{-1}\Big(f(\mathbf{r}, \mathbf{\Omega}) - \frac{1}{4\pi}\sum_{j=1}^{N}\lambda_j B^{-1}g_j(\mathbf{r})\Big) + (\tau - B^{-1})\Big(\frac{\beta_i B^{-1}KB^{-1}}{4\pi}\sum_{j=1}^{N}\lambda_j g_j(\mathbf{r}) - \lambda_i B^{-1}g_i(\mathbf{r})\Big)\Big] + O(\epsilon^2), \quad \tau \text{ large}. \quad (3.35)$$

Note that the rhs of (3.35) is also the intermediate solution $\psi^{II}(\mathbf{r}, \mathbf{\Omega}, \tau)$ to $O(\epsilon^2)$. The short-time outer solution, expanded in powers of $t = \epsilon\tau$ about $t = 0$, is

$$\psi^{III}(\mathbf{r}, \mathbf{\Omega}, \epsilon\tau) \simeq \frac{1}{4\pi}\sum_{i=1}^{N}\lambda_i B^{-1}C_{i0}^{III}(\mathbf{r}, 0) + \frac{\epsilon}{4\pi}\sum_{i=1}^{N}\lambda_i\Big[B^{-1}C_{i1}^{III}(\mathbf{r}, 0) + (\tau - B^{-1}) \times \Big(\frac{\beta_i B^{-1}KB^{-1}}{4\pi}\sum_{j=1}^{N}\lambda_j C_{j0}^{III}(\mathbf{r}, 0) - \lambda_i B^{-1}C_{i0}^{III}(\mathbf{r}, 0)\Big)\Big] + O(\epsilon^2), \quad \epsilon\tau \text{ small}. \quad (3.36)$$

The required matching is obviously accomplished if we put

$$C_{i0}^{III}(\mathbf{r}, 0) = g_i(\mathbf{r}), \quad (3.37)$$

$$C_{i1}^{III}(\mathbf{r}, 0) = \beta_i KB^{-1}\Big(f(\mathbf{r}, \mathbf{\Omega}) - \frac{1}{4\pi}\sum_{j=1}^{N}\lambda_j B^{-1}g_j(\mathbf{r})\Big). \quad (3.38)$$

In summary, Eq. (3.1) gives a representation for ψ to $O(\epsilon^2)$. The inner solution ψ^{I} is given by Eq. (3.4), where ψ_0^{I} is obtained from (3.16) and ψ_1^{I} is obtained from (3.22). The intermediate solution ψ^{II} is equal to the rhs of (3.35). The outer solution ψ^{III} is given by (3.27), where ψ_0^{I} and ψ_1^{I} are the solutions to (3.29) and (3.30). The rhs's of the latter two equations have now been fixed by the matching procedure, and the equations may be rewritten in the following form:

$$B\psi_0^{III}(\mathbf{r}, \mathbf{\Omega}, t) = \frac{1}{4\pi}\sum_{i=1}^{N}\lambda_i\Big(g_i(\mathbf{r})\exp(-\lambda_i t) + \beta_i\int_0^t \exp[-\lambda_i(t-t')](K\psi_0^{III})(\mathbf{r}, t')\,dt'\Big), \quad (3.39)$$

$$B\psi_1^{III}(\mathbf{r}, \mathbf{\Omega}, t) = \frac{1}{4\pi}\sum_{i=1}^{N}\lambda_i\Big[\beta_i KB^{-1}\Big(f(\mathbf{r}, \mathbf{\Omega}) - \frac{1}{4\pi}\sum_{j=1}^{N}\lambda_j B^{-1}g_j(\mathbf{r})\Big) \times \exp(-\lambda_i t) + \beta_i\int_0^t \exp[-\lambda_i(t-t')] \times (K\psi_1^{III})(\mathbf{r}, t')\,dt' + \lambda_i B^{-1}g_i(\mathbf{r}) \times \exp(-\lambda_i t) - \beta_j B^{-1}K\psi_0^{III}(\mathbf{r}, \mathbf{\Omega}, t) - \lambda_i\beta_i\int_0^t \exp[-\lambda_i(t-t')]B^{-1}K\psi_0^{III}(\mathbf{r}, \mathbf{\Omega}, t')\,dt'\Big]. \quad (3.40)$$

An interesting feature of these equations is that the initial condition on the precursor densities is a source for ψ_0^{III}, whereas the initial condition on the neutron distribution function is a source only for ψ_1^{III}. Mathematically, there is nothing inconsistent: ψ_1^{III} is fed by C_{i1}^{III}, and the initial condition on C_{i1}^{III} results from the

delayed neutron precursors produced during the decay of the prompt pulse from its initial condition. The duration of the prompt pulse is $O(\epsilon)$; hence, f should appear multiplied by ϵ. Physically, our results make sense whenever f is of the same order of magnitude as the g_i. This corresponds, for example, to initial conditions which result from a previously established steady state [add a source to the rhs of (1.1) and set derivatives in (1.1) and (1.2) equal to zero]. The results also apply whenever only one of the initial conditions is nonzero, since an initial condition of any magnitude can then be handled by scaling it before and after the analysis. Our results do not apply, however, to a case in which f is very much larger than the g_i (for instance, if the actual density of neutrons, rather than the speed-multiplied density, is initially comparable with the g_i). Such a case might be handled by allowing the initial condition on ψ or C_i to be a function of ϵ; thus, one might put $\psi(\mathbf{r}, \mathbf{\Omega}, 0) = f(\mathbf{r}, \mathbf{\Omega})/\epsilon$. Of course, the perturbation analysis would then differ in its details from that presented above.

Equations (3.39) and (3.40) are Volterra integral equations in time and are readily solved by Neumann series. Calculations in the diffusion theory approximation indicate that only the first few terms need to be retained even for systems above delayed critical. Of course, convergence is speeded the farther the system is below critical and the smaller t. Note that all time quadratures may be performed analytically and that the solution of the time-dependent outer problem has been reduced to the solution of a number of steady-state problems. This number goes up rapidly with the number of iterations, but for well-subcritical problems the number of iterations is generally ≤ 1 [zero iterations corresponding to neglecting the integral terms in (3.39) and (3.40)]. The number of required inversions of B is also greatly reduced by taking only one or two groups of delayed neutrons and in problems in which $g_i = 0$, $i = 1, \cdots, N$.

Whereas the solution in the outer time domain has been reduced to the solution of several steady-state problems, it is clear that the solution in the inner domain still requires the solution of the monoenergetic time-dependent transport equation. This problem has been the subject of a considerable number of investigations,[9–11,14–18] but with limited success except in the case of an unreflected slab. The contribution of the present analysis is that given a method for solving the time-dependent equation, such as a computer code, the known source terms required to introduce delayed neutrons are specified. The time at which the computation of the inner solution can be abandoned in favor of the much simpler intermediate solution is well defined, namely when the difference between them goes below a preset error bound.

In addition to providing a useful method for calculating numerical solutions uniformly valid in time, the decomposition into inner, intermediate, and outer solutions makes the study of the mathematical structure of the solution much simpler. Of course, there is no guarantee that the application of perturbation theory will not change this structure radically. This is especially true in the case of the spectra of the operators involved, since these are known to be very sensitive to small changes in the physical model in question.[19] We shall show, however, that in at least one case this structure remains essentially the same as that which is obtained by exact analysis. In the next section, we shall obtain the outer solution as an eigenfunction expansion. It is necessary to restrict this analysis to the case of isotropic scattering, constant cross sections, and $g_i(\mathbf{r}) = 0$, $i = 1, \cdots, N$, in order to make use of the results of Vladimirov. In addition, we discuss the case of one group of delayed neutrons. The latter restriction is made for clarity in the discussion only, and the generalization to N groups is easy.

4. EIGENFUNCTION EXPANSIONS

The constant cross section, isotropic scattering approximation may be specified by the following revised definitions of the operators L, S, and K:

$$L\psi = \mathbf{\Omega} \cdot \nabla\psi + \Sigma\psi, \tag{4.1}$$

$$S\psi = \frac{c_p\Sigma}{4\pi}\int_\Omega d\mathbf{\Omega}\psi, \tag{4.2}$$

$$K\psi = c_d\Sigma\int_\Omega d\mathbf{\Omega}\psi = \frac{4\pi c_d}{c_p}S\psi, \tag{4.3}$$

where c_p and c_d are the mean number of prompt and delayed secondary neutrons, respectively, and Σ is the reciprocal mean free path for collisions of all types. The Hilbert space h is defined by the real inner product

$$(f, g) = \int_V d\mathbf{r}\int_\Omega d\mathbf{\Omega} fg, \tag{4.4}$$

and norm

$$\|f\| = (f, f)^{\frac{1}{2}}. \tag{4.5}$$

It is easy to show that the following orthogonality relation is valid for appropriately normalized φ_k satisfying (2.3), with L and S defined by (4.1) and (4.2):

$$(\varphi_n, S\varphi_m) = \delta_{mn}. \tag{4.6}$$

The operator S is self-adjoint, and it is unnecessary to consider adjoint eigenfunctions. If we subtract $S\varphi_k$ from both sides of (2.3), we obtain

$$B\varphi_k = (\gamma_k - 1)S\varphi_k. \tag{4.7}$$

This equation, plus the completeness of the set $\{S\varphi_n\}$

in the range of S, permits the solution of (3.39) and (3.40) by eigenfunction expansions for the revised L, S, and K.

It is convenient to work with the Laplace transforms of ψ_0^{III} and ψ_1^{III}. We use the notation

$$\bar{f}(p) = \int_0^\infty e^{-pt} f(t)\, dt. \tag{4.8}$$

Equation (3.39) has the unique solution $\psi_0^{III} = 0$, since $g_i = 0$. The Laplace transform of (3.40) (by use of the fact that $\psi_0^{III} = 0$) is

$$B\bar{\psi}_1^{III}(\mathbf{r}, \boldsymbol{\Omega}, p) = \frac{\lambda\beta c_d}{c_p(p+\lambda)}\left(\frac{1}{4\pi} SB^{-1}f(\mathbf{r}, \boldsymbol{\Omega}) + (S\bar{\psi}_1^{III})(\mathbf{r}, p)\right). \tag{4.9}$$

We have stated that the set $\{\varphi_n\}$ are not complete in h (see Davison[20] for a simple proof); nevertheless, $\bar{\psi}_1^{III}$ can be represented as an expansion in the φ_n. To see this, we note that the rhs of (4.9) is a function of $\mathbf{r}$ and p only and call it $F(\mathbf{r}, p)$. Then

$$L\bar{\psi}_1^{III} = S\bar{\psi}_1^{III} + F(\mathbf{r}, p) \tag{4.10}$$

and, since the rhs of (4.10) is also a function of $\mathbf{r}$ and p only, we can expand in the $S\varphi_n$:

$$L\bar{\psi}_1^{III} = \sum_{n=1}^{\infty} \gamma_n \bar{a}_n(p) S\varphi_n, \tag{4.11}$$

where

$$\bar{a}_n(p) = \gamma_n^{-1}(S\bar{\psi}_1^{III} + F, \varphi_n). \tag{4.12}$$

The operator L^{-1} exists as a bounded transformation of h to h. The solution to (4.11) is thus

$$\bar{\psi}_1^{III} = \sum_{n=1}^{\infty} \bar{a}_n(p)\varphi_n. \tag{4.13}$$

Substituting from (4.13) into (4.9), using (4.6) and (4.7), we can solve for the $\bar{a}_n$:

$$\bar{a}_n(p) = \frac{\lambda\beta c_d (SB^{-1}f, \varphi_n)}{4\pi c_p(\gamma_n - 1)}\left(p + \lambda - \frac{\lambda\beta c_d}{c_p(\gamma_n - 1)}\right)^{-1}. \tag{4.14}$$

Applying the inverse Laplace transformation to (4.13), we obtain

$$\psi_1^{III}(\mathbf{r}, \boldsymbol{\Omega}, t) = \frac{\lambda\beta c_d}{4\pi c_p} \sum_{n=1}^{\infty} \frac{(SB^{-1}f, \varphi_n)}{\gamma_n - 1} \times \exp\left[-\lambda\left(1 - \frac{\beta c_d}{c_p(\gamma_n - 1)}\right)t\right]\varphi_n(\mathbf{r}, \boldsymbol{\Omega}). \tag{4.15}$$

5. DISCUSSION

The γ_n are a sequence of real numbers that increase with n and have no finite accumulation point.[8] Therefore, the $\bar{a}_n(p)$ have poles which accumulate at $-\lambda$ with increasing n. This behavior has been noted in the exact analysis of this problem in slab geometry.[21,22] In addition, comparison of the eigenvalues

$$-\lambda(1 - \beta c_d/c_p(\gamma_n - 1))$$

with the "delayed eigenvalues" predicted by the exact analysis in slabs for a variety of cases indicates that these numbers are essentially exact.[7]

The solution (4.15) complements the Neumann series solution to the original Volterra equation in the sense that, for systems near critical, the series (4.15) may be well approximated by the first mode, whereas several terms may be needed in the Neumann series. The practical value of this observation is not great, however, since, when the system is too near prompt critical, the perturbation method itself breaks down.

ACKNOWLEDGMENTS

The author wishes to thank G. Bell and N. DeMuth for many useful discussions.

* Work performed under the auspices of the U.S. Atomic Energy Commission.

[1] M. Van Dyke, *Perturbation Methods in Fluid Mechanics* (Academic, New York, 1964), p. 77.
[2] J. D. Cole, *Perturbation Methods in Applied Mathematics* (Blaisdell, Waltham, 1968), p. 4.
[3] S. Kaplun, *Fluid Mechanics and Perturbations*, P. A. Lagerstrom *et. al.*, Eds. (Academic, New York, 1967).
[4] W. L. Hendry and G. I. Bell, Nucl. Sci. Eng. **35**, 240 (1969).
[5] J. T. Marti, Nukleonik **8**, 159 (1966).
[6] H. P. McKean, Jr., J. Math. Phys. **8**, 547 (1967).
[7] N. S. DeMuth and W. L. Hendry, Trans. Am. Nucl. Soc. **12**, 729 (1969).
[8] V. S. Vladimirov, Atomic Energy of Canada Limited, Report No. AECL-1661, 1963.
[9] J. Lehner and G. M. Wing, Commun. Pure Appl. Math. **8**, 217 (1955).
[10] J. Lehner and G. M. Wing, Duke Math. J. **23**, 125 (1956).
[11] K. Jorgens, Commun. Pure Appl. Math. **11**, 219 (1958).
[12] K. M. Case and P. F. Zweifel, *Linear Transport Theory* (Addison-Wesley, Reading, Mass., 1967).
[13] B. Davison, *Neutron Transport Theory* (Oxford U. P., London, 1958).
[14] I. Vidav, J. Math. Anal. Appl. **22**, 144 (1968).
[15] A. Bianco, Nukleonika **12**, 459 (1967).
[16] R. Van Norton, Commun. Pure Appl. Math. **15**, 149 (1962).
[17] R. L. Bowden and C. D. Williams, J. Math. Phys. **5**, 1527 (1964).
[18] I. Kuščer and P. F. Zweifel, J. Math. Phys. **6**, 1125 (1965).
[19] G. M. Wing, *An Introduction to Transport Theory* (Wiley, New York, 1962), p. 152.
[20] B. Davison, Ref. 13, p. 435.
[21] H. G. Kaper, J. Math. Anal. Appl. **19**, 207 (1967).
[22] J. Mika, Nukleonik **9**, 46 (1967).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 5 MAY 1970

Analyticity of the Scattering Amplitude in Field Theories with Exponentially Decreasing Commutators in the Spacelike Region*

PETER N. DOBSON, JR.
Department of Physics and Astronomy, University of Hawaii, Honolulu, Hawaii 96822

(Received 2 July 1969)

The implications of weakening the axiom of local commutativity to a condition of exponential decrease are studied. A strong operator limit on the behavior of the commutator for large spacelike separations leads to fields which are almost local in the sense of Haag. A further condition on the spatial behavior of matrix elements of the commutator is introduced, which leads to analyticity of the elastic scattering amplitude at fixed momentum transfer in a striplike region in the complex-energy plane.

I. INTRODUCTION

In most discussions of quantum-field theory from an axiomatic point of view, it is assumed that field operators defined at two points of space–time commute (or anticommute) if the separation between the points is spacelike. This assumption—the axiom of local commutativity, or microcausality—can be justified by the plausible argument that fields at spacelike separated points should be simultaneously measurable. However, in the absence of a consistent theory of measurement with a physical interpretation of the results of such measurements, the argument is not compelling. Furthermore, even for measurable fields there is no experimental evidence excluding precursor effects in the very small spacelike region. From a practical point of view, local commutativity has proven to be a very powerful tool with important consequences for the analytic properties of vacuum expectation values of products of field operators[1] and of scattering amplitudes in theories with an appropriate asymptotic condition. It is the purpose of this paper to explore the consequences for the latter of a weakening of local commutativity to a condition of exponential decrease.

We will treat a model field theory consisting of a single Hermitian scalar field $A(x)$ describing particles of nonzero mass m: We make the usual assumptions,[2] omitting only the local commutativity axiom; namely:

(a) In order to satisfy the requirements of ordinary *quantum mechanics*, we assume that the physical world is represented by a Hilbert space of physical states. The field $A(x)$ is then considered to be an operator-valued distribution defined on a dense subset of this space.

(b) The requirements of *special relativity* are met by assuming that the inhomogeneous Lorentz group is represented on the Hilbert space by a group of unitary operators $U(a, \Lambda)$ and that the field transforms according to the rule

$$U(a, \Lambda)A(x)U^{-1}(a, \Lambda) = A(\Lambda x + a).$$

(c) To insure the absence of states in the Hilbert space with unphysical energy momentum, we consider the *spectrum* of the momentum operator P^μ. As is well known, this is the generator of translations; i.e.,

$$U(a, I) = \exp(iP \cdot a),$$

where a is any 4-vector. We assume that the Hilbert space contains a unique state $|0\rangle$ invariant under all Lorentz transformations. This, of course, implies

$$P^\mu |0\rangle = 0.$$

The state $|0\rangle$ will be interpreted as the "vacuum" or no-particle state. On all other states of the Hilbert space we require that the spectrum of P^μ include only timelike momenta with positive energy. Furthermore, the spectrum of

$$P^2 = (P^0)^2 - |\mathbf{P}|^2$$

contains on these states only a discrete contribution at m^2 and a continuum beginning at $4m^2$. The interpretation of these states is clear: Those with $P^2 = m^2$ are 1-particle states, and those with $P^2 \geq m^2$ contain two or more particles.

In Sec. II, we introduce a condition of strong spacelike exponential decrease for the commutator $[A(x), A(x')]$. We show that this implies the field is almost local in the sense of Haag.[3] This allows us to apply the result of Hepp[4] and introduce the asymptotic condition of Lehmann, Symanzik, and Zimmermann[5] which leads to the usual reduction formulas. In Sec. III, it is pointed out that the strong commutator condition is insufficient in itself to imply analyticity of the scattering amplitude, and a further condition of weak spatial exponential decrease is developed to remedy this. The method of Bremermann, Oehme,

and Taylor[6] is then applied in Sec. IV to give an analytic continuation of the scattering amplitude for fixed momentum transfer into a striplike region in the energy-variable plane.

II. STRONG SPACELIKE DECREASE

The condition of local commutativity will be replaced by the condition that the commutator decrease exponentially in the spacelike region. More precisely, let u be a dimensionless 4-vector, normalized so that $u^2 = -1$. Then we require that there exist some positive number M, such that

$$\lim_{R\to\infty} e^{\mu R}[A(x + Ru), A(x')] = 0, \tag{1}$$

for all $\mu < M$, where R is a parameter with the dimensions of length. The limit in (1) is to be taken in the strong operator sense, and we will refer to the condition as one of strong spacelike decrease.

We wish to show that fields obeying (1) are almost local in the sense of Haag.[3] Almost-local fields obey a condition on the spacelike asymptotic behavior of the truncated vacuum expectation values. The latter are formed from the usual vacuum expectation values by subtracting out contributions due to vacuum intermediate states in a symmetric way.[7] To get an explicit statement of the spacelike asymptotic condition, let us define

$$T(x + a) = \langle 0|\, A(x_1 + a_1)\cdots A(x_n + a_n)\,|0\rangle_{\mathrm{T}},$$

where the subscript T indicates truncation. Next define

$$F_\phi(a) = \int dx\phi(x)T(x + a),$$

where $\phi(x)$ is a test function defined on $4n$-dimensional space and $dx = dx_1\cdots dx_n$. Let the a_i be purely spatial; i.e.,

$$a_i = (0, \mathbf{a}_i).$$

Then the spacelike asymptotic condition takes the form

$$\lim_{\lambda\to\infty} \lambda^N F_\phi(a) = 0, \tag{2}$$

for any positive integer N, where λ is the diameter of the configuration of the a_i. The limit must be taken in such a way that the relative positions of the a_i within the configuration are left unaltered.[8] Ruelle[8] and Araki, Hepp, and Ruelle[9] have demonstrated that strictly-local fields are almost local.

In order to avoid introducing the cumbersome notation required for dealing with the most general case, let us consider here an n-fold truncated vacuum expectation value, wherein only one of the fields undergoes spacelike displacement; i.e., all the a_i are zero but one. In particular, define

$$F_\phi(a) = \int dx\phi(x)\,\langle 0|\, A(x_1 + a)A(x_2)\cdots A(x_n)\,|0\rangle_{\mathrm{T}}$$

and

$$G_\phi(a) = \int dx\phi(x)\,\langle 0|\, A(x_2)\cdots A(x_n)A(x_1 + a)\,|0\rangle_{\mathrm{T}}.$$

We first prove the following lemma.

Lemma: For any positive integer N,

$$\lim_{|\mathbf{a}|\to\infty} |\mathbf{a}|^N [F_\phi(a) - G_\phi(a)] = 0.$$

Proof: First note that

$$F_\phi(a) - G_\phi(a) = \int dx\phi(x)\,\langle 0|\, [A(x_1 + a), A(x_2)\cdots A(x_n)]\,|0\rangle_{\mathrm{T}}. \tag{3}$$

The commutator may be expanded:

$$[A(x_1 + a), A(x_2)\cdots A(x_n)] = \sum_{j=2}^{n} A(x_2)\cdots A(x_{j-1})[A(x_1 + a), A(x_j)] \times A(x_{j+1})\cdots A(x_n).$$

Every commutator appearing in this expansion will involve fields at spacelike separated points if

$$[(x_1 - x_j) + a]^2 < 0,\ j = 2, 3, \cdots, n. \tag{4}$$

Define the Euclidean norm:

$$\|x_j\|^2 = (x_j^0)^2 + |\mathbf{x}_j|^2.$$

Then (4) will be satisfied if

$$\|x_1 - x_j\|^2 < \tfrac{1}{2}|\mathbf{a}|^2$$

and, consequently, if

$$\|x\|^2 \equiv \sum_{j=1}^{n} \|x_j\|^2 < \tfrac{1}{4}|\mathbf{a}|^2.$$

Indeed, this last condition implies

$$[(x_1 - x_j) + a]^2 < -\tfrac{1}{4}|\mathbf{a}|^2.$$

We may write the integral on the right-hand side of Eq. (3) as a sum of two parts. One part is the integral over the exterior of the hypersphere $\|x\| = \frac{1}{2}|\mathbf{a}|$. Ruelle, in his proof for local fields, shows that this must decrease more rapidly than any power of $|\mathbf{a}|$, as $|\mathbf{a}| \to \infty$. The remainder is the integral over the interior of the hypersphere, which is identically zero in the case of local fields. In our case, the strong spacelike decrease of the commutators involved in the

expansion of the integrand is sufficient to give rapid decrease of the integral with increasing $|\mathbf{a}|$.

We can now proceed exactly as Ruelle does to show that the lemma and the spectral conditions of the theory imply

$$\lim_{|\mathbf{a}|\to\infty} |\mathbf{a}|^N F_\phi(a) = 0. \tag{5}$$

Let us introduce a sum over intermediate states into the definition of $F_\phi(a)$:

$$F_\phi(a) = \sum_{|\alpha\rangle} \int dx \phi(x) \langle 0| A(x_1 + a) |\alpha\rangle \times \langle \alpha| A(x_2) \cdots A(x_n) |0\rangle.$$

It must be remembered that the matrix elements appearing here have had vacuum intermediate states subtracted out. Translation invariance immediately reduces the right-hand side to

$$\int dx \phi(x) e^{-ip_\alpha \cdot (x_1 + a)} \langle 0| A(0) |\alpha\rangle \langle \alpha| A(x_2) \cdots A(x_n) |0\rangle.$$

If we define the partial Fourier transform of $\phi(x)$ by

$$\tilde{\phi}(p, x_2, \cdots, x_n) = \int dx_1 e^{-ip\cdot x_1} \phi(x),$$

then

$$F_\phi(a) = \sum_{|\alpha\rangle} \int dx_2 \cdots dx_n \tilde{\phi}(p_\alpha, x_2, \cdots, x_n) e^{ip_\alpha \cdot \mathbf{a}} \times \langle 0| A(0) |\alpha\rangle \langle \alpha| A(x_2) \cdots A(x_n) |0\rangle. \tag{6}$$

The spectral conditions require that either $p_\alpha = 0$ or $p_\alpha^0 > 0$ and $p_\alpha^2 \geq m^2$. The first of these alternatives is ruled out because the truncated vacuum expectation value has no vacuum intermediate states and hence $|\alpha\rangle$ cannot be $|0\rangle$. Let us define the test function

$$\tilde{\psi}(p, x_2, \cdots, x_n) = h(p) \tilde{\phi}(p, x_2, \cdots, x_n),$$

where $h(p) = 1$, for $p^0 > 0$ and $p^2 \geq m^2$, but $h(p) = 0$ outside the forward light cone. Equation (6) and the spectral conditions imply

$$F_\psi(a) = F_\phi(a). \tag{7}$$

The lemma requires that

$$\lim_{|\mathbf{a}|\to\infty} |\mathbf{a}|^N [F_\psi(a) - G_\psi(a)] = 0; \tag{8}$$

so let us examine $G_\psi(a)$ more closely. With the introduction of a sum over intermediate states, we have

$$G_\psi(a) = \sum_{|\alpha\rangle} \int dx_2 \cdots dx_n \tilde{\psi}(-p_\alpha, x_2, \cdots, x_n) e^{-ip_\alpha \cdot \mathbf{a}} \times \langle 0| A(x_2) \cdots A(x_n) |\alpha\rangle \langle \alpha| A(0) |0\rangle.$$

Since $p_\alpha^0 > 0$, by the spectral conditions, we must have

$$\tilde{\psi}(-p_\alpha, x_2, \cdots, x_n) = 0,$$

and hence $G_\psi(a) = 0$. Inserting this result and Eq. (7) in Eq. (8) proves Eq. (5), which is the desired spacelike asymptotic condition for the case at hand.

The general case follows from an extension of the arguments above. In the most general case, all fields in the n-fold truncated vacuum expectation value may be undergoing displacement. We consider all partitions of the displacements into two subgroups and choose that partition which maximizes the distance between the subgroups. We define $F_\phi(a)$ as above, and let $G_\phi(a)$ be the same except for having a different permutation of the fields, the permutation being chosen so as to keep the same relative order of the fields within the two subgroups. The expansion of $F_\phi(a) - G_\phi(a)$ involves only commutators between a field in one subgroup and a field in the other. As in the special case treated above, the expression may be divided into two parts. One is an integral over the exterior of a hypersphere whose radius tends to infinity as the displacements tend to infinity. This integral vanishes faster than any power of the diameter of the configuration, as shown by Ruelle. The other part is an integral over the interior of the hypersphere and involves only commutators of fields at points which are separated by spacelike 4-vectors which grow as the diameter of the configuration is increased. The strong spacelike decrease of these commutators is sufficient to damp the integral faster than any power of the diameter of the configuration. By transforming to momentum space and using the support properties given there by the spectral conditions, we can always find some ψ such that (7) is true and for which $G_\psi(a) = 0$. Thus, fields obeying Eq. (1) are almost-local fields.

The significance of this identification is that we may now employ the result of Hepp[4] to define asymptotic fields and apply the Lehmann, Symanzik, and Zimmerman[5] (LSZ) reduction technique. The reduction formula for the Lorentz-invariant scattering amplitude is

$$T(p_1, p_2; p_3, p_4) = i \int d^4x e^{ip_1 \cdot x} \{\theta(x^0) \langle p_2| [j(x), j(0)] |p_3\rangle + \delta(x^0) \langle p_2| [A^0(x), j(0)] |p_3\rangle\}. \tag{9}$$

In Eq. (9), $j(x) = (\Box + m^2)A(x)$ is the usual current operator, and $A^0(x) = \partial A(x)/\partial x^0$. Since the fields under consideration here do not have a vanishing spacelike commutator, the right-hand side of (9) is not manifestly covariant. However, T is covariant;

thus we will assume that any apparent lack of covariance cancels out on integration. We also ignore the possible difficulties associated with the multiplication of the matrix elements by step functions such as $\theta(x^0)$. Hepp[10] has investigated cases in which ambiguities due to this point can be avoided, and we simply assume the same is true in all cases treated here, while recognizing the fact that it may be technically very difficult to justify this assumption even if it is correct.

An additional reduction formula we will need for our later discussion is

$$\int d^4x e^{ip_1\cdot x}\langle p_2|\, j(x)j(0)\,|p_3\rangle = \tfrac{1}{16} g_{ra}(k_1, k_2, k_3), \quad (10)$$

where

$$\begin{aligned} g_{ra}(k_1, k_2, k_3) &= \sum_{i,j=1}^{2} \int d^4x_1\, d^4x_2\, d^4x_3 e^{i(k_1\cdot x_1 + k_2\cdot x_2 + k_3\cdot x_3)} \\ &\times \rho_i(x^0)\rho_j(-x^0)\langle 0|\, [B^i(\tfrac{1}{2}x_1 + \tfrac{1}{4}x_3), j(-\tfrac{1}{2}x_1 + x_3)] \\ &\times [B^j(-\tfrac{1}{2}x_2 - \tfrac{1}{4}x_3), j(\tfrac{1}{2}x_2 - \tfrac{1}{4}x_3)]\,|0\rangle, \end{aligned}$$

where

$$\begin{aligned} \rho_1(x^0) = \theta(x^0), \quad \rho_2(x^0) = \delta(x^0), \\ B^1(x) = j(x), \quad B^2(x) = A^0(x), \\ k_1 = \tfrac{1}{2}(p_2 - p_1), \quad k_2 = \tfrac{1}{2}(p_3 - p_4), \end{aligned}$$

and

$$k_3 = \tfrac{1}{2}(p_1 + p_2).$$

Equations (9) and (10) are the starting point for the analytic continuation of the scattering amplitude. As we discover in the next section, however, strong spacelike decrease is insufficient in itself for this purpose.

III. WEAK SPATIAL DECREASE

Using the transformation properties of the fields under displacements, we can rewrite Eq. (9) in the form

$$\begin{aligned} T(p_1, p_2; p_3, p_4) = i\int d^4x e^{ip_1\cdot x} e^{i(p_2 - p_3)\cdot \frac{1}{2}x} \\ \times \{\theta(x_0)\langle p_2|\, [j(\tfrac{1}{2}x), j(-\tfrac{1}{2}x)]\,|p_3\rangle \\ + \delta(x^0)\langle p_2|\, [A^0(\tfrac{1}{2}x), j(-\tfrac{1}{2}x)]\,|p_3\rangle\}. \end{aligned}$$

We evaluate the right-hand side of this expression in the "brick wall" reference frame defined by

$$\begin{aligned} p_1 = (\omega, \boldsymbol{\Delta} + \rho\mathbf{e}), \quad p_3 = (E_\Delta, \boldsymbol{\Delta}), \\ p_2 = (E_\Delta, -\boldsymbol{\Delta}), \quad p_4 = (\omega, -\boldsymbol{\Delta} + \rho\mathbf{e}), \end{aligned}$$

where

$$\begin{aligned} |\mathbf{e}| = 1, \quad \mathbf{e}\cdot\boldsymbol{\Delta} = 0, \quad \rho^2 = \omega^2 - E_\Delta^2, \\ E_\Delta^2 = \boldsymbol{\Delta}^2 + m^2. \end{aligned}$$

Now

$$\begin{aligned} p_1\cdot x + (p_2 - p_3)\cdot \tfrac{1}{2}x &= (p_1 + p_4)\cdot \tfrac{1}{2}x \\ &= \omega x^0 - \rho\mathbf{e}\cdot\mathbf{x}. \end{aligned}$$

Thus we have

$$\begin{aligned} T(p_1, p_2; p_3, p_4) = i\int d^4x e^{i(\omega x^0 - \rho\mathbf{e}\cdot\mathbf{x})} \\ \times \{\theta(x^0)\langle p_2|\, [j(\tfrac{1}{2}x), j(-\tfrac{1}{2}x)]\,|p_3\rangle \\ + \delta(x^0)\langle p_2|\, [A^0(\tfrac{1}{2}x), j(-\tfrac{1}{2}x)]\,|p_3\rangle\} \\ \equiv M_r(\omega, \Delta^2), \end{aligned}$$

where the subscript r refers to the presence of retarded commutators in the integrand. Following Bremermann, Oehme, and Taylor,[6] we introduce a new parameter β and define a generalized amplitude

$$\begin{aligned} M_r(\omega, \Delta^2, \beta) = i\int d^4x \exp\{i[\omega x^0 - (\omega^2 - \beta)^{\frac{1}{2}}\mathbf{e}\cdot\mathbf{x}]\} \\ \times \{\theta(x^0)\langle p_2|\, [j(\tfrac{1}{2}x), j(-\tfrac{1}{2}x)]\,|p_3\rangle \\ + \delta(x^0)\langle p_2|\, [A^0(\tfrac{1}{2}x), j(-\tfrac{1}{2}x)]\,|p_3\rangle\}. \end{aligned} \quad (11)$$

In the case of local fields, the integrand is nonzero only in the forward light cone. In our case, we will write it as the sum of two parts: a "temporal part," involving an integral over the forward light cone, and a "spatial part" consisting of the remainder. The arguments used for local fields give an analytic continuation of the temporal part M_r^t into the region

$$\{(\omega, \beta)\,|\,\operatorname{Im}\omega > \operatorname{Im}(\omega^2 - \beta)^{\frac{1}{2}}\}. \quad (12)$$

The branch of the square root has been chosen so that the imaginary part is nonnegative. The usual assumption that the integrand is a tempered distribution also implies that the temporal part is of at most polynomial increase as $|\omega| \to \infty$ inside the specified region. For β real and negative, we have a continuation to the entire upper-half ω-plane.

The spatial part M_r^s has no continuation to values in the region (12). The difficulty arises in the fact that the exponential damping of the commutators can only give convergence inside a hyperboloid $x^2 < -a^2$. The integral, however, includes an infinite volume outside this hyperboloid, in which the explicit exponential factors are increasing for ω and β in (12). What is needed, evidently, is for the integrand to provide exponential decrease with increasing $|\mathbf{x}|$, independent of x^0.

A condition such as suggested above clearly lacks manifest covariance. For this reason, it could not be stated in the form of a strong operator limit on the commutator without a strong presumption that the latter would violate the Lorentz invariance of the

theory. However, there seems no such objection to stating such a condition on matrix elements of the commutator, that is, in terms of a weak operator limit. The matrix elements of the commutator must involve state vectors which themselves are not Lorentz invariant, with the exception of the vacuum state. The vacuum state can give no difficulty, for it is well known that the vacuum expectation value of the commutator vanishes for spacelike separation as a consequence of the other axioms even when local commutativity is not assumed. Hence, the commutator, as an operator, could extract coordinate-system-dependent information from the state vectors involved in its nonvanishing matrix elements.

As an example of what is meant here, consider the matrix element

$$F(x,p) = \langle 0|\,[A(x), A(0)]\,|p\rangle,$$

where $|p\rangle$ has 4-momentum $(E, \mathbf{p})$. Lorentz invariance requires only that $F(x,p)$ be a function of the invariants x^2, p^2, and $x\cdot p$. Suppose that for spacelike x,

$$F(x,p) = e^{f(x,p)},$$

where

$$f(x,p) = x^2p^2 - (x\cdot p)^2.$$

Maximizing f over spacelike values of x gives

$$f(x,p) \leq -(E - |\mathbf{p}|)^2\,|\mathbf{x}|^2.$$

Thus

$$\lim_{|\mathbf{x}|\to\infty} e^{\mu|\mathbf{x}|}F(x,p) = 0,$$

for any μ, and the limit is uniform in x^0 (provided $x^0 < |\mathbf{x}|$).

A condition appropriate for our purposes is the following. Let Φ and Ψ be any two states in the domain of the field operator, and let u be a dimensionless spacelike 4-vector with $|\mathbf{u}| = 1$. Then we require that there exist some positive number M, such that

$$\lim_{R\to\infty} e^{\mu R}\langle\Phi|\,[A(x+Ru), A(x')]\,|\Psi\rangle = 0, \quad (13)$$

for all $\mu < M$. We require the limit to be uniform with respect to u^0; it is this requirement that introduces a lack of manifest covariance and which gives the desired exponential decrease with increasing spatial displacement. We will call (13) the condition of weak spatial decrease and require that derivatives of the commutator obey the same condition.

With Eq. (12) we can now obtain a continuation of the spatial part M_r^s to the region

$$\{(\omega,\beta)\mid \operatorname{Im}\omega \geq 0, \operatorname{Im}(\omega^2+\beta)^{\frac{1}{2}} < M\}.$$

In particular, we see that this includes

$$\{(\omega,\beta)\mid \omega\in D_+(\Delta^2,\delta), \beta\in S_\beta(\delta)\}, \quad (14)$$

where

$$D_+(\Delta^2,\delta) = \{\omega=\omega_1+\omega_2\mid 4M^2 + 4M^2(\omega_1^2-\omega_2^2-E_\Delta^2-\delta) > 4\omega_1^2\omega_2^2 + 4|\omega_1|\,|\omega_2|\,\delta+\delta^2;\ \omega_2\geq 0\} \quad (15)$$

and

$$S_\beta(\delta) = \{\beta\mid -R < \operatorname{Re}\beta < E_\Delta^2+\delta;\ |\operatorname{Im}\beta| < \delta;\ R>0\}. \quad (16)$$

For our later discussion we will want the entire real axis to lie in $D_+(\Delta^2,\delta)$, so we impose the restrictions

$$M > E_\Delta, \quad |\delta| < 2M(M-E_\Delta), \quad (17)$$

where M_r^s will be of at most polynomial increase as $|\omega|\to\infty$ inside the region defined by Eqs. (14)–(16). For simplicity, we assume that both M_r^t and M_r^s actually vanish as $|\omega|\to\infty$, and the more general case follows from this one by the usual subtraction technique.

Consider β real and in $S_\beta(\delta)$. Then we can write an integral representation for $M_r(\omega,\Delta^2,\beta)$ by deforming the contours of Cauchy integral representations of M_r^t and M_r^s. The result is

$$M_r(\omega,\Delta^2,\beta) = \frac{1}{2\pi i}\int_{-\infty}^{+\infty} d\omega^1\,\frac{M_r(\omega^1,\Delta^2,\beta)}{\omega^1-\omega} + \frac{1}{2\pi i}\int_{C_+} d\omega^1\,\frac{M_r^s(\omega^1,\Delta^2,\beta)}{\omega^1-\omega}.$$

The contour C_+ in the second integral is the upper boundary of $D_+(\Delta^2,\delta)$. The representation holds for ω in $D_+(\Delta^2,\delta)$.

We now define functions M_a^s and M_a^t as the spatial and temporal parts of $M_a(\omega^a,\Delta^2,\beta)$, which is defined by the replacement $\theta(x^0)\to-\theta(-x^0)$ in (11). The a refers to the advanced commutator which this introduces. By the same reasoning as used for M_r, we find

$$M_a(\omega,\Delta^2,\beta) = \frac{1}{2\pi i}\int_{-\infty}^{+\infty} d\omega^1\,\frac{M_a(\omega^1,\Delta^2,\beta)}{\omega^1-\omega} + \frac{1}{2\pi i}\int_{C_-} d\omega^1\,\frac{M_a^s(\omega^1,\Delta^2,\beta)}{\omega^1-\omega},$$

for ω in $D_-(\Delta^2,\delta)$, which is the reflection of $D_+(\Delta^2,\delta)$ in the real axis. C_- is the lower boundary of $D_-(\Delta^2,\delta)$.

Let us consider the absorptive part of the amplitude, defined by

$$A(\omega,\Delta^2,\beta) = \frac{1}{2i}\lim_{\epsilon\to 0}[M_r(\omega+i\epsilon,\Delta^2,\beta) - M_a(\omega-i\epsilon,\Delta^2,\beta)]$$
$$= \tfrac{1}{2}\int d^4x\,\exp\{i[\omega x^0 - (\omega^2-\beta)^{\frac{1}{2}}\mathbf{e}\cdot\mathbf{x}]\} \times \langle p_2|\,[j(\tfrac{1}{2}x), j(-\tfrac{1}{2}x)]\,|p_3\rangle. \quad (18)$$

If we insert a sum over a complete set of states, make use of translation invariance, and do the integral in (18), we find

$$A(\omega, \Delta^2, \beta) = \tfrac{1}{2}(2\pi)^4 \sum_{|\alpha\rangle} \langle p_2| j(0) |\alpha\rangle \langle \alpha| j(0) |p_3\rangle \times \{\delta(p_\alpha - Q - \tfrac{1}{2}p_2 - \tfrac{1}{2}p_3) - \delta(p_\alpha + Q - \tfrac{1}{2}p_2 - \tfrac{1}{2}p_3)\},$$

where $Q = [\omega, (\omega^2 - \beta)^{\frac{1}{2}}\mathbf{e}]$.

Since $|\alpha\rangle$ must be at least a 1-particle state, with $p_\alpha^2 \geq m^2$, we have $A(\omega, \Delta^2, \beta) = 0$, provided $\beta < -\Delta^2$ and $|\omega| < -(\beta + \Delta^2)/2E_\Delta^2$. Thus, we really have a single analytic function $M(\omega, \Delta^2, \beta)$ represented by M_r, when $\operatorname{Im}\omega > 0$, and by M_a, when $\operatorname{Im}\omega < 0$. $M(\omega, \Delta^2, \beta)$ itself has the integral representation

$$M(\omega, \Delta^2, \beta) = \frac{1}{\pi}\int_{-\infty}^{+\infty} d\omega^1 \frac{A(\omega^1, \Delta^2, \beta)}{\omega^1 - \omega} + \frac{1}{2\pi i}\int_{C_+} d\omega^1 \frac{M_r^s(\omega^1, \Delta^2, \beta)}{\omega^1 - \omega} + \frac{1}{2\pi i}\int_{C_-} d\omega^1 \frac{M_a^s(\omega^1, \Delta^2, \beta)}{\omega^1 - \omega}. \quad (19)$$

Again following Ref. 6, we define

$$\rho(k^2, \beta - \Delta^2, \Delta^2) = \tfrac{1}{2}(2\pi)^4 \sum_{|\alpha\rangle} \langle p_2| j(0) |\alpha\rangle \langle \alpha| j(0) |p_3\rangle \delta(k - p_\alpha)$$

and write the spectral representation

$$A(\omega, \Delta^2, \beta) = \int dk^2 \rho(k^2, \beta - \Delta^2, \Delta^2) \times \{\delta(k^2 - \beta - E_\Delta^2 - 2\omega E_\Delta) - \delta(k^2 - \beta - E_\Delta^2 + 2\omega E_\Delta)\}.$$

The spectral conditions applied to the state $|\alpha\rangle$ require $p_\alpha^2 = m^2$ or $p_\alpha^2 \geq 4m^2$, so that

$$\rho(k^2, \beta - \Delta^2, \Delta^2) = \pi m^2 g^2(\beta - \Delta^2)\delta(k^2 - m^2) + \theta(k^2 - 4m^2)\sigma(k^2, \beta - \Delta^2, \Delta^2).$$

Inserting this into Eq. (19) gives

$$M(\omega, \Delta^2, \beta) = \frac{g^2(\beta - \Delta^2)h(\beta, \Delta^2)}{\omega_\beta^2(m) - \omega^2} + \frac{1}{\pi}\int_{\omega_\beta(2m)+\beta/2E_\Delta}^{\infty} d\omega^1 \sigma(2E_\Delta\omega^1 + E_\Delta^2, \beta - \Delta^2, \Delta^2) \times \left[\left(\omega^1 - \frac{\omega - \beta}{2E_\Delta}\right)^{-1} + \left(\omega^1 + \frac{\omega + \beta}{2E_\Delta}\right)^{-1}\right] + \frac{1}{2\pi i}\int_{C_+} d\omega^1 \frac{M_r^s(\omega^1, \Delta^2, \beta)}{\omega^1 - \omega} + \frac{1}{2\pi i}\int d\omega^1 \frac{M_a^s(\omega^1, \Delta^2, \beta)}{\omega^1 - \omega}, \quad (20)$$

where we have defined

$$h(\beta, \Delta^2) = -m^2(\beta + \Delta^2)/4E_\Delta^2$$

and

$$\omega_\beta(k) = (k^2 - E_\Delta^2 - \beta)/2E_\Delta.$$

The lower limit of the integral over the cut has been shifted by a change of variables, so as to be independent of β, for

$$\omega_\beta(2m) + \beta/2E_\Delta = (4m^2 - E_\Delta^2)/2E_\Delta.$$

In summary, the representation (20) is valid if ω lies in the region bounded above by C_+ and below by C_-, with cuts from $\pm\omega_\beta(2m)$ to $\pm\infty$ on the real axis, if the restrictions (17) are imposed. To make use of this representation, we will need to explore the analyticity of the spectral function in β.

IV. ANALYTICITY OF THE SCATTERING AMPLITUDE

In the last section, we introduced a representation for the function $M(\omega, \Delta^2, \beta)$ which explicitly exhibits its analyticity in a portion of the ω plane. The next step in demonstrating analyticity of the scattering amplitude consists of continuing this representation in the variable β back to the physical value E_Δ^2. It is sufficient for this purpose that the integrands in (20) be analytic in the strip $S_\beta(\delta)$ defined by Eq. (16). The required analyticity of $g^2(\beta - \Delta^2)$ follows from this by arguments essentially those of Ref. 6. We have already shown in Sec. III that M_r^s and M_a^s have the required analyticity in β. We now demonstrate it for the spectral function.

Since $k^2 \geq m^2$, we may write

$$(k^2 - m^2)\sigma(k^2, \beta - \Delta^2, \Delta^2) = (k^2 - m^2)\rho(k^2, \beta - \Delta^2, \Delta^2) = (k^2 - m^2)\bar{A}(p_1 + p_4, p_2, p_3),$$

where

$$\bar{A}(p_1 + p_4, p_2, p_3) = \tfrac{1}{2}\int d^4x e^{i(p_1+p_4)\cdot\frac{1}{2}x} \langle p_2| j(\tfrac{1}{2}x) j(-\tfrac{1}{2}x) |p_3\rangle.$$

Applying the reduction formula (10), we have

$$\bar{A}(p_1 + p_4, p_2, p_3) = 16 g_{ra}(k_1, k_2, k_3).$$

The subscripts r, a refer to the presence of a retarded commutator in x_1 and an advanced commutator in x_2 in the integral defining g_{ra}. We now define analogous functions g_{rr}, g_{aa}, and g_{ar} by the indicated replacements $\theta(x^0) \leftrightarrow -\theta(-x^0)$ in the definition of g_{ra}.

The spectral conditions imply a number of "joining" properties of the g_{ij}. These are given in Ref. 6 and are

$$g_{rj} - g_{aj} = 0, \quad \text{for} \quad (k_1 + k_3)^2,$$

and

$$(k_1 - k_3)^2 < 4m^2;$$

$$g_{jr} - g_{ja} = 0, \quad \text{for} \quad (k_2 + k_3)^2,$$

and

$$(k_2 - k_3)^2 < 4m^2;$$

$$g_{ij} = 0, \quad \text{for} \quad k_3^2 < m^2 \quad \text{or} \quad k_3^0 < 0.$$

Let $k_3 = (t, \mathbf{0})$ with $t \geq m$ real. Then the representation for g_{ra} plus the condition of weak spatial decrease defines a function analytic for

$$(\operatorname{Im} k_1^0) > |\operatorname{Im} \mathbf{k}_1|, \quad |\operatorname{Im} \mathbf{k}_1| < M.$$

Similarly, g_{aa} is analytic for

$$(\operatorname{Im} k_1^0) < -|\operatorname{Im} \mathbf{k}_1|, \quad |\operatorname{Im} \mathbf{k}_1| < M.$$

Thus, as functions of k_1, g_{ra} and g_{aa} define a function analytic in a "flattened wedge" given by

$$W_f = \{k_1 \mid |\operatorname{Im} \mathbf{k}_1| < \min(|\operatorname{Im} k_1^0|, M)\}.$$

The joining property for fixed t gives us an "E-limiting" property on the set

$$E = \{k_1 \mid \operatorname{Im} k_1 = 0, (\operatorname{Re} k_1 \pm t)^2 < 4m^2\}.$$

Simply stated, this means that if p is in E, then $g_{ra}(k_1, k_2, t)$ and $g_{aa}(k_1, k_2, t)$ have the same limit as k_1 approaches p along paths in the respective domains of analyticity.

Consideration of the functions g_{ar} and g_{rr} leads to a similar result in the variable k_2. So we actually have a function $g(k_1, k_2, t)$ for each fixed $t \geq m$, analytic in $W_f \times W_f$ with an E-limiting property on the set $E \times E$. By virtue of the "edge-of-the-wedge" theorem,[11] $g(k_1, k_2, t)$ will also be analytic in a neighborhood of $E \times E$.

By Lorentz invariance, we can express g as a function of invariants only. Using the notation of Ref. 6, we define

$$\chi(z_1, z_2, z_3, z_4, z_5, z_6) = g(k_1, k_2, k_3),$$

where

$$z_1 = (k_1 + k_3)^2 = p_2^2, \quad z_2 = (k_2 + k_3)^2 = p_3^2,$$
$$z_3 = (k_1 - k_3)^2 = p_1^2, \quad z_4 = (k_2 - k_3)^2 = p_4^2,$$
$$z_5 = (k_1 - k_2)^2 = (p_1 - p_4)^2,$$

and $z_6 = 4k_3^2 = 4t^2$. Then we have the following:

Theorem: Considered as a function of the five complex variables $z_1, \cdots, z_5$, the function $\chi(z_1, \cdots, z_6)$ is analytic for each real $t \geq m$ in the region

$$\{z_1, \cdots, z_5 \mid |z_{1,2} - m^2| < \delta, |z_{3,4} - \gamma| < \delta, \quad |z_5 + 4\Delta^2| < \delta\},$$

for all real γ such that $-R \leq \gamma \leq m^2$ and $\Delta < \Delta_{\max}$. To satisfy (17), $\Delta_{\max}$ can be no larger than $(M^2 - m^2)^{\frac{1}{2}}$. We will find a further restriction on its value in the course of indicating the proof of the theorem. Note that

$$\tfrac{1}{4}(p_1 + p_4)^2 = \tfrac{1}{2}(z_3 + z_4) - \tfrac{1}{4}z_5.$$

Thus the theorem implies that the spectral function is analytic for all β satisfying

$$|\beta - \gamma - \Delta^2| < \tfrac{5}{4}\delta,$$

which includes the strip defined by (16).

Proof of the theorem: We must show that for $z_1 = z_2 = m^2$, $z_3 = z_4 = \gamma$, $z_5 = -4\Delta^2$, and $z_6 = 4t^2$, we can always find k_1, k_2 in the domain of analyticity of $g(k_1, k_2, t)$. We have

$$k_1^0 = k_2^0 = (m^2 - \gamma)/4t,$$

$$\mathbf{k}_1 = \rho(t, \gamma)\mathbf{e}_1 + \Delta\mathbf{e}_2, \quad \text{and} \quad \mathbf{k}_2 = \rho(t, \gamma)\mathbf{e}_1 - \Delta\mathbf{e}_2,$$

where

$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij} \quad \text{and} \quad \rho^2(t, \gamma) = \left(t + \frac{m^2 - \gamma}{4t}\right)^2 - E_\Delta^2.$$

If we have $\rho^2(t, \gamma) > -\bar{\delta}^2$, then

$$|\operatorname{Im} \mathbf{k}_i| < \bar{\delta} \quad \text{and} \quad (\operatorname{Re} k_i^0 \pm t)^2 - |\operatorname{Re} \mathbf{k}_i|^2 \leq m^2.$$

But this means that if $\bar{\delta}$ is sufficiently small, k_1 and k_2 will be in a small neighborhood of the set $E \times E$, where $g(k_1, k_2, t)$ is analytic by virtue of the edge-of-the-wedge theorem. We can see that $\Delta^2 < \bar{\delta}^2$ guarantees this; so for sufficiently small momentum transfer the theorem holds. To continue to larger Δ^2, we use the semitube method as developed by Bremermann.[12] First note that, for $\rho^2(t, \gamma)$ positive, we have analyticity via edge of the wedge, as indicated above. In the (t, γ) plane, the curve $\rho^2 = 0$ is as shown in Fig. 1. Since we need only consider $\gamma \leq m^2$ and $t \geq m$, the shaded region in the figure is all that remains to be accounted for. As shown in Ref. 6, for $\Delta^2 < 3m^2$ and $m \leq t \leq E_\Delta$, the set E contains the set

$$E' = \{k \mid \operatorname{Im} k = 0, |\operatorname{Re} \mathbf{k}| < \infty, |\operatorname{Re} k^0| < \eta(t)\},$$

with $\eta(t) = 2m - t$. Applying Bremermann's theorem[13] for $\eta < M$, we have that $g(k_1, k_2, t)$ is analytic in the semitube domain $H_0 \times H_0$, where

$$H_0 = \{k \mid k^0 \in B, |\operatorname{Im} \mathbf{k}| < v(k^0), |\operatorname{Re} \mathbf{k}| < \infty\}$$

and where B is the k^0 plane cut from $\pm\eta$ to $\pm\infty$. The

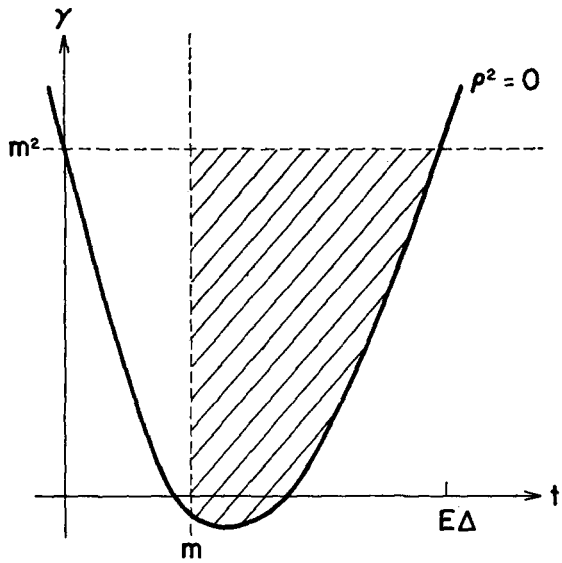


FIG. 1. Region in the (γ, t) plane not covered by the edge-of-the-wedge theorem.

function $v(k^0)$ is

$$\max\{\tfrac{1}{4}(\eta - |\mathrm{Re}\, k^0|), \min(|\mathrm{Im}\, k^0|, M)\}.$$

The holomorphy envelope of this domain is $H \times H$, where

$$H = \{k \mid k^0 \in B, |\mathrm{Im}\, \mathbf{k}| < V(k^0), |\mathrm{Re}\, \mathbf{k}| < \infty\}$$

and $V(k^0)$ is the smallest superharmonic majorant of $v(k^0)$.[14] We find

$$V(k^0) = \frac{2M}{\pi}\,\mathrm{Im}\cosh^{-1}\left(\frac{\cosh(\pi k^0/2M)}{\cosh(\pi\eta/2M)}\right), \quad |\mathrm{Im}\, k^0| < M,$$
$$= M, \quad |\mathrm{Im}\, k^0| > M.$$

If for k_1 and k_2 we have

$$|\mathrm{Re}\, k^0| < \eta(t)$$

and

$$|\mathrm{Im}\, \mathbf{k}| < V(k^0),$$

then $g(k_1, k_2, t)$ is analytic. The first condition is simply $(m^2 - \gamma)/4t < 2m - t$, which is fulfilled for all γ and t in the shaded region of Fig. 1, if $E_\Delta < 2m$. Now,

$$\mathrm{Im}\, \mathbf{k}_1 = \mathrm{Im}\, \mathbf{k}_2 = \left[E_\Delta^2 - \left(t + \frac{m^2 - \gamma}{4t}\right)^2\right]^{\frac{1}{2}},$$

so the second condition implies

$$\frac{2M}{\pi}\tan^{-1}\left[\left(\frac{\cosh^2[\pi(2m - t)/2M]}{\cosh^2[\pi(m^2 - \gamma)/2M]} - 1\right)^{\frac{1}{2}}\right] - \left[E_\Delta^2 - \left(t + \frac{m^2 - \gamma}{4t}\right)^2\right]^{\frac{1}{2}} > 0.$$

The left-hand side of this is always greater than

$$(2M/\pi)\tan^{-1}[\sinh \pi m/2M] - \Delta,$$

for t and γ in the shaded region. Thus, if we take

$$\Delta_{\max} \leq (2M/\pi)\tan^{-1}(\sinh \pi m/2M),$$

then the theorem holds. Combining this with the restriction from (17), we have

$$\Delta_{\max} = \min\left\{(M^2 - m^2)^{\frac{1}{2}}, \frac{2M}{\pi}\tan^{-1}\left(\sinh\frac{\pi m}{2M}\right)\right\}. \tag{21}$$

A plot of $\Delta_{\max}/m$ vs M/m is given in Fig. 2.

The theorem gives us the analyticity needed to carry out a continuation of (20) to the physical value of β, but we still need to show that the spectral function retains its temperateness in the relevant part of the envelope of holomorphy. This step was omitted in Ref. 6 and is implicit in modern proofs of dispersion relations for local fields. However, the author has demonstrated that the necessary extension of temperateness can be carried out in the Bremermann, Oehme, and Taylor framework,[15] and essentially the same arguments go through in the present case.

With the analyticity demonstrated above, the details of the continuation of (20) to $\beta = E_\Delta^2$ are as in the local case and need not be repeated here. We find a continuation into the domain $D(\Delta^2, \delta) = D_+(\Delta^2, \delta) \cup D_-(\Delta^2, \delta)$ except for the usual cuts and poles on the real axis, where δ is a small but otherwise arbitrary parameter. If we define

$$D(\Delta^2) = \{\omega \mid (\mathrm{Im}\, \omega/M)^2 < 1 - E_\Delta^2/(M^2 + \mathrm{Re}\, \omega^2)\}, \tag{22}$$

then we see that every interior point of $D(\Delta^2)$ lies in some $D(\Delta^2, \delta)$. Thus our result may be summarized as follows:

The elastic scattering amplitude may be analytically continued into the domain $D(\Delta^2)$ except for cuts

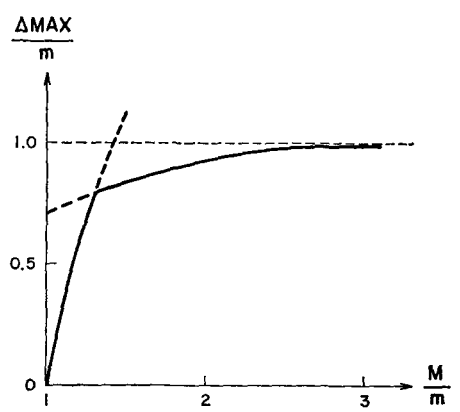


FIG. 2. Variation of $\Delta_{\max}$ with M.

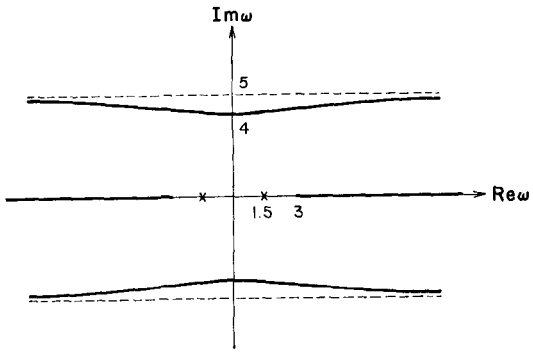


FIG. 3. The domain $D(\Delta^2)$ for $M = 5$, $m = 3$, and $\Delta = 0$ (forward scattering).

running from $\pm(m^2 - \Delta^2)/E_\Delta$ to $\pm\infty$ and poles at $\pm(E_\Delta - m^2/2E_\Delta)$, provided that $M < m$ and $\Delta < \Delta_{\max}$, where $\Delta_{\max}$ and $D(\Delta^2)$ are defined by Eqs. (21) and (22). For an example of a domain $D(\Delta^2)$, see Fig. 3.

In determining the domain of analyticity of the scattering amplitude we made use of the analyticity of the absorptive part in several complex variables. To make maximum use of this analyticity, we would need to know the envelope of holomorphy of the primitive domain constructed from the spectral conditions, the condition of weak spatial decrease, and Lorentz invariance. Instead, we have employed the semitube method which gives only a subset of the envelope of holomorphy. In the local field case, extending the domain of analyticity further into the holomorphy envelope permits the proof of dispersion relations for larger Δ and eventually for $\Delta < 2m$. It seems reasonable, if somewhat less than urgent, that we might be able to prove analyticity in our case for Δ larger than the $\Delta_{\max}$ of Eq. (21); the present result should be considered a minimum of analytic continuation.

ACKNOWLEDGMENTS

The author would like to acknowledge the support of the National Science Foundation, during the period when the dissertation on which the present work is based was being written, and the support of the U.S. Atomic Energy Commission during its revision. Particular gratitude is expressed to J. S. Toll for his guidance and assistance while he was Chairman of the Department of Physics and Astronomy at the University of Maryland, and to him and K. Hepp for several subsequent suggestions and criticisms.

* This work is based on a dissertation submitted to the University of Maryland in partial fulfillment of the requirements for the Ph.D. degree.

[1] See, for example, A. S. Wightman, Phys. Rev. **101**, 860 (1956), and a related paper in the reprint collection *Dispersion Relations and the Abstract Approach to Field Theory*, L. Klein, Ed. (Gordon & Breach, New York, 1961).

[2] For a more general discussion of possible axioms for a quantum-field theory, see R. Haag and B. Schroer, J. Math. Phys. **3**, 248 (1962).

[3] R. Haag, Nuovo Cimento, Suppl., **14**, 131 (1959).

[4] K. Hepp, Commun. Math. Phys. **1**, 95 (1965).

[5] H. Lehmann, K. Symanzik, and W. Zimmermann, Nuovo Cimento **1**, 205 (1955).

[6] H. J. Bremermann, R. Oehme, and J. G. Taylor, Phys. Rev. **109**, 2178 (1958).

[7] See Ref. 3 for a precise definition of truncation.

[8] D. Ruelle, Helv. Phys. Acta **35**, 147 (1962).

[9] H. Araki, K. Hepp, and D. Ruelle, Helv. Phys. Acta **35**, 164 (1962).

[10] K. Hepp, Helv. Phys. Acta **37**, 639 (1964).

[11] R. F. Streater and A. S. Wightman, *PCT, Spin & Statistics, and All That* (Benjamin, New York, 1964), pp. 74–84.

[12] H. J. Bremermann, Math. Ann. **127**, 406 (1954).

[13] H. J. Bremermann, Rev. Mat. Hispano-Am. **27**, 175 (1957).

[14] For a discussion of superharmonic functions, see R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience, New York, 1962), Vol. II, pp. 306–312.

[15] P. N. Dobson, Jr., J. Math. Phys. **8**, 315 (1967).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 5 MAY 1970

Unitary Representations of the Group $SL(3, C)$ in an $SU(3)$ Basis

N. MUKUNDA
Tata Institute of Fundamental Research, Bombay, India

(Received 5 September 1969)

Unitary irreducible representations of the group $SL(3, C)$ are obtained in an $SU(3)$ basis. The principal nondegenerate and principal degenerate series of representations are considered, and the matrix elements of the generators are worked out by the use of expansion formulas expressing them as functions of the generators of the group $SU(3) \times T_8$. The forms of the Casimir operators of $SL(3, C)$ and their values in these representations are also given.

INTRODUCTION

The purpose of this paper is to obtain in explicit form the unitary irreducible representations (UIR's) of the noncompact group $SL(3, C)$ in a basis made up of vectors belonging to finite-dimensional UIR's of the maximal compact subgroup $SU(3)$. We concern ourselves only with the principal nondegenerate and the principal degenerate series of UIR's of $SL(3, C)$, and not with the supplementary series.[1] Furthermore, we deal only with the problem of obtaining the matrix elements of the generators of the Lie algebra of $SL(3, C)$ in the above-mentioned basis and do not deal with finite elements of the group.

The most severe difficulty that one faces in trying to solve the problem outlined above is the problem of internal multiplicity. In general, a UIR of $SL(3, C)$ contains a given finite-dimensional UIR of $SU(3)$ several times. The difficulty lies in the choice of state labels that take account of this multiplicity and serve to distinguish the several occurrences of one UIR of $SU(3)$, and then in evaluating the dependence of the matrix elements of the generators on these labels. Such problems are present for many families of semisimple noncompact Lie groups.[2] In a previous paper, we outlined a method by which these problems can be solved in the case of $SL(3, C)$.[3] The group $SL(3, C)$ can be related via group contraction to the group $SU(3) \times T_8$ which is also noncompact but not semisimple. For suitably related UIR's of these two groups, the rules for reduction under $SU(3)$ turn out to be identical. In the case of the representations of $SU(3) \times T_8$, an elegant and simple method exists for choosing labels to take care of the multiplicity problem as well as for evaluating the generator matrix elements.[4] In the paper referred to above, we showed that, in both the principal nondegenerate and the principal degenerate series of UIR's of $SL(3, C)$, the generators of $SL(3, C)$ are written as polynomial functions of the generators of suitably chosen UIR's of $SU(3) \times T_8$, and we refer to these formulas as expansion formulas. Thus, by combining the known methods for setting up UIR's of $SU(3) \times T_8$ with the previously derived expansion formulas, we obtain the UIR's of $SL(3, C)$ in an $SU(3)$ basis.

Section 1 contains a brief description of the Lie algebras of $SL(3, C)$ and $SU(3) \times T_8$, the corresponding Casimir invariants, and the classes of UIR's of these groups that we are interested in. We also write down the expressions for the matrix elements of the $SU(3) \times T_8$ generators, describe the correspondence between the UIR's of $SL(3, C)$ and $SU(3) \times T_8$ as revealed by the common $SU(3)$ reduction characteristics, and give the expansion formulas leading from the $SU(3) \times T_8$ generators to those of $SL(3, C)$. Sections 2 and 3 deal with the evaluation of the matrix elements of two characteristic combinations of the $SU(3) \times T_8$ generators, which appear in the expansion formulas. In Sec. 4, we put these results together to obtain the final expressions for the matrix elements of the $SL(3, C)$ generators, for both the nondegenerate and the degenerate principal series. Some of the calculations necessary to obtain the results of Secs. 2 and 3 have been explained in an appendix.

1. UIR's OF $SL(3, C)$, $SU(3) \times T_8$ AND THE EXPANSION FORMULAS

We begin with a description of the Lie algebra of $SL(3, C)$. The sixteen generators of $SL(3, C)$ are made up of the eight generators of the subgroup $SU(3)$, written as J^α_β, and eight "noncompact" generators K^α_β that transform according to the 8-dimensional adjoint representation of $SU(3)$ (as do the J^α_β). [As in I, where convenient, we use tensor notation with respect to $SU(3)$, with the indices $\alpha, \beta, \cdots$ taking the values 1, 2, 3.] The Hermiticity, trace, and commutation properties in a unitary representation of $SL(3, C)$ are[5]

$$(J^\alpha_\beta)^\dagger = J^\beta_\alpha, \quad (K^\alpha_\beta)^\dagger = K^\beta_\alpha, \quad J^\alpha_\alpha = K^\alpha_\alpha = 0, \tag{1.1a}$$

$$[J^\alpha_\beta, J^\lambda_\mu] = \delta^\alpha_\mu J^\lambda_\beta - \delta^\lambda_\beta J^\alpha_\mu, \tag{1.1b}$$

$$[J^\alpha_\beta, K^\lambda_\mu] = \delta^\alpha_\mu K^\lambda_\beta - \delta^\lambda_\beta K^\alpha_\mu, \tag{1.1c}$$

$$[K^\alpha_\beta, K^\lambda_\mu] = \delta^\lambda_\beta J^\alpha_\mu - \delta^\alpha_\mu J^\lambda_\beta. \tag{1.1d}$$

In the choice of Casimir invariants for $SL(3, C)$, we can be guided by the fact that, disregarding Hermiticity properties for the moment, the two sets of operators

$$M_\beta^\alpha = \tfrac{1}{2}(J_\beta^\alpha + iK_\beta^\alpha), \quad N_\beta^\alpha = \tfrac{1}{2}(J_\beta^\alpha - iK_\beta^\alpha) \tag{1.2}$$

obey the commutation rules of two commuting $SU(3)$ groups. It follows that $SL(3, C)$ possesses four Casimir invariants which may be obtained from the expressions for the two invariants of an $SU(3)$ Lie algebra. We define the four Hermitian Casimir invariants for $SL(3, C)$ as follows:

$$\mathcal{C}_1 = \mathrm{Re}\, M_\beta^\alpha M_\alpha^\beta = \tfrac{1}{4}(J_\beta^\alpha J_\alpha^\beta - K_\beta^\alpha K_\alpha^\beta); \tag{1.3a}$$

$$\mathcal{C}_2 = \mathrm{Im}\, M_\beta^\alpha M_\alpha^\beta = \tfrac{1}{2}J_\beta^\alpha K_\alpha^\beta; \tag{1.3b}$$

$$\begin{aligned}\mathcal{C}_3 &= \mathrm{Re}\, \tfrac{1}{2}\{M_\beta^\alpha, M_\gamma^\beta\}_+ M_\alpha^\gamma \\ &= \tfrac{1}{16}[(\{J_\beta^\alpha, J_\gamma^\beta\}_+ - \{K_\beta^\alpha, K_\gamma^\beta\}_+)J_\alpha^\gamma \\ &\quad - (\{J_\beta^\alpha, K_\gamma^\beta\}_+ + \{K_\beta^\alpha, J_\gamma^\beta\}_+)K_\alpha^\gamma];\end{aligned} \tag{1.3c}$$

$$\begin{aligned}\mathcal{C}_4 &= \mathrm{Im}\, \tfrac{1}{2}\{M_\beta^\alpha, M_\gamma^\beta\}_+ M_\alpha^\gamma \\ &= \tfrac{1}{16}[(\{J_\beta^\alpha, J_\gamma^\beta\}_+ - \{K_\beta^\alpha, K_\gamma^\beta\}_+)K_\alpha^\gamma \\ &\quad + (\{J_\beta^\alpha, K_\gamma^\beta\}_+ + \{K_\beta^\alpha, J_\gamma^\beta\}_+)J_\alpha^\gamma].\end{aligned} \tag{1.3d}$$

(In these expressions, $\{\,,\}_+$ denotes the anticommutator.)

The UIR's of the principal nondegenerate series of $SL(3, C)$ are labeled by four real parameters, in the form $(m_2, m_3, \rho_2, \rho_3)$. Of these, m_2 and m_3 are integers (positive, negative, or zero), while ρ_2 and ρ_3 are two arbitrary real numbers. In every UIR of this series, there is a multiplicity in the reduction with respect to $SU(3)$. To exhibit the Casimir operators as functions of the labeling parameters, we introduce the expressions for the values of the quadratic and cubic invariants of $SU(3)$[6]:

$$\begin{aligned}I_2(\lambda, \mu) &= \tfrac{1}{9}(\lambda^2 + \lambda\mu + \mu^2 + 3\lambda + 3\mu), \\ I_3(\lambda, \mu) &= \tfrac{1}{162}(\lambda - \mu)(\lambda + 2\mu + 3)(2\lambda + \mu + 3).\end{aligned} \tag{1.4}$$

Then one finds that, in the UIR $(m_2, m_3, \rho_2, \rho_3)$ of $SL(3, C)$, the Casimir invariants have the following values[7]:

$$\begin{aligned}\mathcal{C}_1 + i\mathcal{C}_2 &= 6I_2(-1 + \tfrac{1}{2}(i\rho_3 - m_3), \\ &\qquad -1 + \tfrac{1}{2}(m_2 - i\rho_2)), \\ \mathcal{C}_3 + i\mathcal{C}_4 &= 18I_3(-1 + \tfrac{1}{2}(i\rho_3 - m_3), \\ &\qquad -1 + \tfrac{1}{2}(m_2 - i\rho_2)).\end{aligned} \tag{1.5}$$

Of course, the $\mathcal{C}_i$ are real numbers, and are obtained by separating the right-hand sides into real and imaginary parts.

In the principal degenerate series of $SL(3, C)$, each UIR is labeled by just two parameters (m_3, ρ_3), with m_3 once again being an integer (positive, negative, or zero) and ρ_3 an arbitrary real number. These UIR's are free of multiplicity in their reduction under $SU(3)$. The values of the Casimir invariants are now given by[8]

$$\begin{aligned}\mathcal{C}_1 + i\mathcal{C}_2 &= 6I_2(-\tfrac{3}{2} + \tfrac{1}{2}(i\rho_3 - m_3), 0), \\ \mathcal{C}_3 + i\mathcal{C}_4 &= 18I_3(-\tfrac{3}{2} + \tfrac{1}{2}(i\rho_3 - m_3), 0).\end{aligned} \tag{1.6}$$

Let us next give a brief account of the structure and representations of $SU(3) \times T_8$. This group also has sixteen generators, which we write as J_β^α, P_β^α. Equations (1.1a), (1.1b), and (1.1c) continue to hold, with K replaced in all of them by P.[9] However, in place of Eq. (1.1d), we now have

$$[P_\beta^\alpha, P_\mu^\lambda] = 0. \tag{1.7}$$

Thus, $SU(3) \times T_8$ has the structure of a semidirect product. We define the Casimir invariants of $SU(3) \times T_8$ in such a way that they are the limits of those of $SL(3, C)$ under group contraction.[10] Writing them as $\mathcal{C}_i^{(0)}$, we have

$$\begin{aligned}\mathcal{C}_1^{(0)} &= -\tfrac{1}{4}P_\beta^\alpha P_\alpha^\beta, \\ \mathcal{C}_2^{(0)} &= \tfrac{1}{2}J_\beta^\alpha P_\alpha^\beta, \\ \mathcal{C}_3^{(0)} &= -\tfrac{3}{8}J_\beta^\alpha P_\gamma^\beta P_\alpha^\gamma, \\ \mathcal{C}_4^{(0)} &= -\tfrac{1}{8}P_\beta^\alpha P_\gamma^\beta P_\alpha^\gamma.\end{aligned} \tag{1.8}$$

The expression for $\mathcal{C}_3^{(0)}$ has been simplified by making use of the commutation rules.

The UIR's of $SU(3) \times T_8$ are divided in Paper I into three families, denoted as types I, II, and III.[11] Here we need to use only the first two types. A UIR of type I is specified by four parameters in the form (M_0, Y_0, p_0, p_1). Here M_0 and Y_0 are real discrete-valued variables, the possible values of the pair (M_0, Y_0) being all possible "weights" within representations of $SU(3)$. [A "weight" within $SU(3)$ is any allowed choice of eigenvalues for the ordered pair of operators $(\tfrac{1}{2}(J_1^1 - J_2^2), -J_3^3)$.[12]] Then, p_0 and p_1 are two arbitrary real numbers subject only to the condition that p_1 should neither vanish nor be equal to $\pm(3)^{\frac{1}{2}}p_0$.[13] Within such a UIR of $SU(3) \times T_8$, the invariants $\mathcal{C}_i^{(0)}$ have the following values[14]:

$$\begin{aligned}\mathcal{C}_1^{(0)} &= -\tfrac{3}{2}(p_0^2 + p_1^2), \\ \mathcal{C}_2^{(0)} &= (3)^{\frac{1}{2}}p_1M_0 - \tfrac{3}{2}p_0Y_0, \\ \mathcal{C}_3^{(0)} &= \tfrac{3}{2}(3)^{\frac{1}{2}}p_0p_1M_0 + \tfrac{3}{8}(p_0^2 - p_1^2)Y_0, \\ \mathcal{C}_4^{(0)} &= -\tfrac{3}{4}p_0(p_0^2 - 3p_1^2).\end{aligned} \tag{1.9}$$

For the UIR's of $SU(3) \times T_8$ belonging to type II, we need only two parameters, namely, (Y_0, p_0). The allowed values of Y_0 are just the possible eigenvalues of the operator $-J_3^3$ within representations of $SU(3)$, while p_0 is an arbitrary nonzero real number. The values of $\mathcal{C}_i^{(0)}$ are now obtained from (1.9) by just

setting $p_1 = 0$:

$$\mathcal{C}_1^{(0)} = -\tfrac{3}{2}p_0^2, \quad \mathcal{C}_2^{(0)} = -\tfrac{3}{2}p_0 Y_0,$$
$$\mathcal{C}_3^{(0)} = \tfrac{9}{8}p_0^2 Y_0, \quad \mathcal{C}_4^{(0)} = -\tfrac{3}{4}p_0^3. \tag{1.10}$$

We now describe these representations in more detail. Let us first consider nondegenerate UIR's of $SL(3, C)$ and type I UIR's of $SU(3) \times T_8$. The UIR $(m_2, m_3, \rho_2, \rho_3)$ of $SL(3, C)$ and the UIR (M_0, Y_0, p_0, p_1) of $SU(3) \times T_8$ contain the same UIR's of $SU(3)$ with the same multiplicities, if

$$m_2 = 2M_0, \quad 2m_3 - m_2 = 3Y_0. \tag{1.11}$$

This common $SU(3)$ content is given by the following rule: The UIR (λ, μ) of $SU(3)$ occurs as many times as it contains vectors with the weight (M_0, Y_0). Note that the rule does not involve the continuous parameters ρ_2, ρ_3 or p_0, p_1. Now within a UIR of $SU(3)$, we use three labels I, M, and Y to specify a basis vector. The last two constitute the weight of the vector, while I distinguishes different vectors with the same weight and is related to the quadratic invariant of an $SU(2)$ subgroup of $SU(3)$. (In applications to particle physics, I is the total isotopic spin.) It is no surprise then that we can set up an orthonormal basis in the space of the UIR (M_0, Y_0, p_0, p_1) of $SU(3) \times T_8$ with the vectors labeled

$$|\lambda, \mu; IMY; I_0\rangle, \tag{1.12}$$

where (λ, μ) denotes the representation of $SU(3)$ for which this vector forms a basis element and IMY is the triplet of quantities distinguishing vectors within this UIR of $SU(3)$. For given λ and μ, the possible values of I_0 are just the values that I could assume, if M and Y were set equal to M_0 and Y_0, respectively. In writing the basis vectors for the UIR of $SU(3) \times T_8$ in this way, we have indicated in (1.12) only those labels which change from vector to vector in the representation space, omitting the parameters M_0, Y_0 and p_0, p_1 which characterize the entire representation. It should be clear that transformations of the subgroup $SU(3)$ affect the quantities IMY only, and leave I_0 unaltered. The UIR of $SU(3) \times T_8$ is completely specified if one gives the expressions for the $SU(3)$-reduced matrix elements (RME's) of J and P in the basis (1.12).[15] These are

$$\langle \lambda'\mu'; I_0' \| J \| \lambda\mu; I_0 \rangle_\gamma = \delta_{\lambda'\lambda}\delta_{\mu'\mu}\delta_{I_0'I_0}\delta_{\gamma,1}[I_2(\lambda, \mu)]^{\frac{1}{2}}, \tag{1.13a}$$

$$\langle \lambda'\mu'; I_0' \| P \| \lambda\mu; I_0 \rangle_\gamma = [d(\lambda\mu)/d(\lambda'\mu')]^{\frac{1}{2}} \times \sum_{I_0''=0,1} p_{I_0''} C(\lambda\mu\ 11\ \lambda'\mu'\gamma; I_0 M_0 Y_0\ I_0''00\ I_0' M_0 Y_0). \tag{1.13b}$$

Note that the RME's of P are themselves given in terms of Clebsch–Gordan (CG) coefficients. The quantity $I_2(\lambda, \mu)$ was defined in Eq. (1.4); $d(\lambda\mu)$ stands for the dimension of the UIR (λ, μ) of $SU(3)$.[16] It may be remarked that the RME's of J suffer no change when we go from representations of $SU(3) \times T_8$ to those of $SL(3, C)$. It is also clear that with the range of values of I_0 described above, the states (1.12) serve as an $SU(3)$ basis for the UIR $(m_2, m_3, \rho_2, \rho_3)$ of $SL(3, C)$, provided Eqs. (1.11) hold.

The analogous description of the degenerate UIR's of $SL(3, C)$ and the type II UIR's of $SU(3) \times T_8$ is much simpler. If the relation

$$2m_3 = 3Y_0 \tag{1.14}$$

holds, the UIR (m_3, ρ_3) of $SL(3, C)$ and the UIR (Y_0, p_0) of $SU(3) \times T_8$ reduce in the following common multiplicity-free manner under $SU(3)$: The UIR (λ, μ) of $SU(3)$ appears once if it contains the state $IMY = (0, 0, Y_0)$ and is absent otherwise. Once again omitting labels which do not change within a representation, we see that an $SU(3)$ basis for the UIR (Y_0, p_0) is given by the vectors

$$|\lambda\mu; IMY\rangle, \tag{1.15}$$

with the values of λ and μ chosen according to the above rule. Again, these vectors just suffice to form a basis for the UIR (m_3, ρ_3) of $SL(3, C)$, with (1.14) being valid. The RME's of the $SU(3) \times T_8$ generators now are

$$\langle \lambda'\mu' \| J \| \lambda\mu \rangle_\gamma = \delta_{\lambda'\lambda}\delta_{\mu'\mu}\delta_{\gamma,1}[I_2(\lambda, \mu)]^{\frac{1}{2}}, \tag{1.16a}$$

$$\langle \lambda'\mu' \| P \| \lambda\mu \rangle_\gamma = [d(\lambda\mu)/d(\lambda'\mu')]^{\frac{1}{2}} p_0 \times C(\lambda\mu\ 11\ \lambda'\mu'\gamma; 00 Y_0\ 000\ 00 Y_0). \tag{1.16b}$$

Now we come to the expansion formulas. Let a nondegenerate UIR $(m_2, m_3, \rho_2, \rho_3)$ of $SL(3, C)$ be given, and introduce the parameter

$$s = -1 + (\rho_3/\rho_2). \tag{1.17}$$

We say that this UIR of $SL(3, C)$ and the UIR (M_0, Y_0, p_0, p_1) of $SU(3) \times T_8$ are "corresponding representations" of the two groups if the two pairs of discrete parameters are related by Eqs. (1.11), while p_0 and p_1 are given by

$$p_0 = \tfrac{1}{6}(2s + 1), \quad p_1 = -\tfrac{1}{2}(3)^{-\frac{1}{2}}. \tag{1.18}$$

Thus, this is a many-to-one correspondence from $SL(3, C)$ to $SU(3) \times T_8$. Then the generators K for the UIR $(m_2, m_3, \rho_2, \rho_3)$ of $SL(3, C)$ are constructed from the generators of the corresponding UIR of

$SU(3) \times T_8$ according to

$$K_\beta^\alpha = \rho_2 P_\beta^\alpha + \tfrac{2}{5} i \frac{(s^2 + 4s + 1)}{s(s+1)} (PJ - JP)_\beta^\alpha + \tfrac{2}{3} i \frac{(1-s)}{s(s+1)} (QJ - JQ)_\beta^\alpha + \frac{i}{s(s+1)} (QJP - PJQ)_\beta^\alpha, \tag{1.19}$$

while the $SU(3)$ generators J are taken over unchanged. Here, in addition to the operators J and P, we have introduced another octet of operators Q given by[17]

$$Q_\beta^\alpha = P_\gamma^\alpha P_\beta^\gamma - \tfrac{1}{3} \delta_\beta^\alpha P_\mu^\lambda P_\lambda^\mu. \tag{1.20}$$

At the same time, we have used matrix notation in (1.19), treating K, J, P, and Q as 3×3 matrices with the superscript (subscript) being identified as the row (column) index. Keeping s fixed, we see that P is the contracted form of K:

$$P_\beta^\alpha = \lim_{\rho_2 \to \infty} \frac{1}{\rho_2} K_\beta^\alpha. \tag{1.21}$$

The choice of Casimir invariants for the two groups has been made in such a way that the values of the $SL(3, C)$ invariants go over under contraction, i.e., $\rho_2 \to \infty$ at a fixed value of s, into the values of the $SU(3) \times T_8$ invariants for the corresponding representation. Thus one may check

$$\lim_{\rho_2 \to \infty} (\rho_2)^{-n_i} \mathcal{C}_i(m_2, m_3, \rho_2, \rho_3) = \mathcal{C}_i^{(0)}(M_0, Y_0, p_0, p_1), \quad n_i = 2, 1, 2, 3, \quad \text{for } i = 1, 2, 3, 4, \text{ respectively}, \tag{1.22}$$

where the dependences of the $\mathcal{C}_i$ and $\mathcal{C}_i^{(0)}$ on their arguments are given by Eqs. (1.5) and (1.9), while the parameters M_0, Y_0, p_0, p_1 are fixed by Eqs. (1.11) and (1.18).

For the degenerate representations of $SL(3, C)$, the situation is as follows. The UIR (m_3, ρ_3) of $SL(3, C)$ being given, the "corresponding" (type II) UIR (Y_0, p_0) of $SU(3) \times T_8$ is that for which Y_0 is given by Eq. (1.14), and

$$p_0 = \tfrac{1}{3}. \tag{1.23}$$

Once again, this is a many-to-one relationship from $SL(3, C)$ to $SU(3) \times T_8$. Then the generators K, for the UIR (m_3, ρ_3) of $SL(3, C)$, can be built up from the generators of the corresponding UIR of $SU(3) \times T_8$ by the formula

$$K_\beta^\alpha = \rho_3 P_\beta^\alpha + i(PJ - JP)_\beta^\alpha. \tag{1.24}$$

The $SU(3)$ generators are, of course, unchanged. Equation (1.21) now holds with ρ_3 written in place of ρ_2; for the values of the Casimir invariants one can check the following relations:

$$\lim_{\rho_3 \to \infty} (\rho_3)^{-n_i} \mathcal{C}_i(m_3, \rho_3) = \mathcal{C}_i^{(0)}(Y_0, p_0), \quad Y_0 = \tfrac{2}{3} m_3, \quad p_0 = \tfrac{1}{3}. \tag{1.25}$$

[For the values of $\mathcal{C}_i$ and $\mathcal{C}_i^{(0)}$, see Eqs. (1.6) and (1.10). The indices n_i are the same as in Eq. (1.22).]

We close this section with a few remarks. The two expansion formulas (1.19) and (1.24) make no explicit reference to the discrete parameters, (m_2, m_3) and m_3, respectively, in the problem. This means that their validity is not restricted to *irreducible* representations of $SU(3) \times T_8$. We can take a reducible representation of $SU(3) \times T_8$ made up of any number of type I UIR's of $SU(3) \times T_8$, all of which have the same values for p_0 and p_1 [given by Eq. (1.18) but possibly corresponding to different values of M_0 and Y_0]. If we then apply formula (1.19) to the generators of this reducible representation of $SU(3) \times T_8$, then we end up with operators J and K that generate a reducible representation of $SL(3, C)$. The irreducible parts of this representation operate precisely on those subspaces that carry the irreducible representations of $SU(3) \times T_8$. All the UIR's of $SL(3, C)$ are of non-degenerate type and have the same values for ρ_2 and ρ_3 (we do not mean $\rho_2 = \rho_3$!), while the various values of m_2 and m_3 are determined by the various values of M_0 and Y_0. A similar statement can be made in the case of formula (1.24). This feature of the expansion formulas simplifies the calculation of the matrix elements of the last term in (1.19) considerably.

From the way Q has been constructed, it follows that J and Q obey the commutation rules of $SU(3) \times T_8$, as do J and P. In a general type I UIR of $SU(3) \times T_8$, where the RME's of P are given by Eq. (1.13b), the RME's of Q are given by exactly the same expressions with p_0 and p_1 replaced by q_0 and q_1, where

$$q_0 = p_0^2 - p_1^2, \quad q_1 = -2p_0 p_1. \tag{1.26}$$

(However, the set of operators J, Q is not always irreducible.) Therefore, a single calculation is sufficient to handle both the term $(PJ - JP)$ and the term $(QJ - JQ)$, while the term $(QJP - PJQ)$ has to be treated somewhat differently. In the following section, we deal with the former kind of term and in Sec. 3 with the latter.

2. TREATMENT OF THE TERM $(PJ - JP)$

It is clear that the RME's of the tensor operator $(PJ - JP)$ between specified $SU(3)$ representations must be proportional to those of P between the same representations, the factor of proportionality depending on the representations involved.[18] This is so

because J does not change either the $SU(3)$ representation or any $SU(3)$ invariant labels the vectors may carry. Of the eight possible RME's of $(PJ - JP)$, we see that the two that are representation diagonal in $SU(3)$ must vanish. For, if we form the two $SU(3)$ scalar operators by combining $(PJ - JP)$ with J and (JJ), use of the commutation relations shows that they both vanish:

$$J^{\beta}_{\alpha}(PJ - JP)^{\alpha}_{\beta} = J^{\beta}_{\gamma}J^{\gamma}_{\alpha}(PJ - JP)^{\alpha}_{\beta} = 0. \quad (2.1)$$

These two combinations essentially project out two independent linear combinations of the representation-diagonal RME's of $(PJ - JP)$; hence, the stated result

$$\langle \lambda\mu; I_0' \| \, PJ - JP \, \| \lambda\mu; I_0 \rangle_{\gamma} = 0, \quad \gamma = 1, 2. \quad (2.2)$$

We are then left with the six RME's in which the initial and final $SU(3)$ representations are different, and we can write these as

$$\langle \lambda'\mu'; I_0' \| \, PJ - JP \, \| \lambda\mu; I_0 \rangle = a(\lambda\mu; \lambda'\mu') \langle \lambda'\mu'; I_0' \| \, P \, \| \lambda\mu; I_0 \rangle, \quad (\lambda'\mu') \neq (\lambda\mu). \quad (2.3)$$

With the usual $SU(3)$ phase conventions, all the $a(\lambda\mu; \lambda'\mu')$ are real. In the evaluation of the coefficients $a(\lambda\mu; \lambda'\mu')$, we use the properties following from the behavior of the tensors P and $(PJ - JP)$ under Hermitian conjugation and R conjugation.[19] From the former, we get the relation

$$a(\lambda\mu; \lambda'\mu') = -a(\lambda'\mu'; \lambda\mu), \quad (2.4)$$

while the latter implies

$$a(\lambda\mu; \lambda'\mu') = -a(\mu'\lambda'; \mu\lambda). \quad (2.5)$$

Combining these leads to a third useful relation,

$$a(\lambda\mu; \lambda'\mu') = a(\mu\lambda; \mu'\lambda'). \quad (2.6)$$

As a result, it is enough to compute $a(\lambda\mu; \lambda'\mu')$ for the two cases $(\lambda'\mu') = (\lambda + 1, \mu + 1)$ and $(\lambda + 2, \mu - 1)$ only. In both these cases, the simplest procedure is to compute explicitly the matrix element

$$\langle \lambda'\mu'; I_m I_m Y_m; I_0' | \, (PJ - JP)_{000} \, | \lambda\mu; I_m I_m Y_m; I_0 \rangle, \quad (2.7)$$

by putting in the relevant intermediate states between P and J. (Note that one does not really need to know the value of the RME's of P for this calculation.)

Here we choose the component of the tensor $PJ - JP$ with labels $IMY = 000$, while the labels $I_m I_m Y_m$ denote the highest weight state in the UIR $(\lambda\mu)$.[20] States with these labels do exist in the two UIR's $(\lambda'\mu')$ we are interested in. After obtaining (2.7), we factor out an $SU(3)$ CG coefficient, compare with the right-hand side of (2.3), and deduce the value of $a(\lambda\mu; \lambda'\mu')$. In this way, we obtain $a(\lambda\mu; \lambda'\mu') = (\lambda + 2), -(\lambda + \mu + 1), -\lambda, (\lambda + \mu + 3), (\mu + 2), -\mu$, for

$$(\lambda'\mu') = (\lambda + 2, \mu - 1), (\lambda - 1, \mu - 1), (\lambda - 2, \mu + 1), (\lambda + 1, \mu + 1), (\lambda - 1, \mu + 2), (\lambda + 1, \mu - 2),$$

respectively. (2.8)

This completes the evaluation of the nonvanishing RME's of $(PJ - JP)$. It is obvious that Eqs. (2.1)–(2.3) are valid if in all of them we replace the tensor P by the tensor Q introduced in the previous section. Also, Eq. (2.3) can be used with the $SU(3)$ invariant labels I_0, I_0' omitted on both sides. Therefore, the results of this section suffice for obtaining the matrix elements of the $SL(3, C)$ generators K in the principal degenerate series of representations.

3. THE OPERATOR $(QJP - PJQ)$

We turn now to the evaluation of RME's of the operator $(QJP - PJQ)$, which occurs only in the principal nondegenerate UIR's of $SL(3, C)$. At first sight, it appears that this requires a lot more work than the last section, since it seems necessary to introduce two sets of intermediate states in the calculation of the matrix elements. However, this is not so, as a slight rewriting of the operator $(QJP - PJQ)$ reduces its treatment to the same level of simplicity as the treatment of the operator $(PJ - JP)$.

Let us write

$$U^{\alpha}_{\beta} = (QJP - PJQ)^{\alpha}_{\beta}. \quad (3.1)$$

Using the commutation relations, we can bring the $SU(3)$ generators J to the extreme left in each term, and we then find that

$$U^{\alpha}_{\beta} = J^{\lambda}_{\mu} T^{\mu\alpha}_{\beta\lambda},$$
$$T^{\mu\alpha}_{\beta\lambda} = P^{\mu}_{\beta} Q^{\alpha}_{\lambda} - P^{\alpha}_{\lambda} Q^{\mu}_{\beta}. \quad (3.2)$$

In arriving at this result, we used the following two obvious properties:

$$[Q^{\alpha}_{\beta}, P^{\lambda}_{\mu}] = 0, \quad Q^{\alpha}_{\gamma} P^{\gamma}_{\beta} = P^{\alpha}_{\gamma} Q^{\gamma}_{\beta}. \quad (3.3)$$

Of course, the different components of Q commute with one another, so that the different components of T also commute with one another. Now, we analyze the tensor T. It has the antisymmetry property

$$T^{\mu\alpha}_{\beta\lambda} = -T^{\alpha\mu}_{\lambda\beta} \quad (3.4)$$

which, coupled with the vanishing of the traces

$$T^{\mu\alpha}_{\mu\lambda} = 0, \quad T^{\mu\alpha}_{\beta\alpha} = 0, \quad (3.5)$$

implies that T can be expressed as a linear combination of irreducible tensor operators belonging to those UIR's of $SU(3)$ that are contained in the antisymmetric part of the direct product of the adjoint representation with itself. This direct product contains the UIR's (3, 0), (0, 3), and (1, 1) of $SU(3)$, once each. However, the trace operation which would lead to the representation (1, 1) also gives a vanishing result on account of Eq. (3.3):

$$T^{\mu\alpha}_{\beta\mu} = 0. \tag{3.6}$$

Consequently, T can be written as the sum of two tensor operators, one belonging to the representation (3, 0) of $SU(3)$ and the other to the representation (0, 3). The part that transforms by the representation (3, 0) is given by a completely symmetric third-rank tensor $S_{\alpha\beta\gamma}$ obtained from T as follows:

$$S_{\alpha\beta\gamma} = \epsilon_{\alpha\rho\sigma}T^{\rho\sigma}_{\beta\gamma} + \epsilon_{\beta\rho\sigma}T^{\rho\sigma}_{\alpha\gamma} + \epsilon_{\gamma\rho\sigma}T^{\rho\sigma}_{\beta\alpha}; \tag{3.7}$$

the (0, 3) component is a tensor $\bar{S}^{\alpha\beta\gamma}$ given by

$$\bar{S}^{\alpha\beta\gamma} = \epsilon^{\alpha\rho\sigma}T^{\beta\gamma}_{\rho\sigma} + \epsilon^{\beta\rho\sigma}T^{\alpha\gamma}_{\rho\sigma} + \epsilon^{\gamma\rho\sigma}T^{\beta\alpha}_{\rho\sigma}. \tag{3.8}$$

These two tensors are related by Hermitian conjugation, for from

$$(T^{\mu\alpha}_{\beta\lambda})^\dagger = T^{\beta\lambda}_{\mu\alpha} \tag{3.9}$$

it follows that

$$(S_{\alpha\beta\gamma})^\dagger = \bar{S}^{\alpha\beta\gamma}. \tag{3.10}$$

We can reconstruct T in terms of S and $\bar{S}$ and thereby also write U in terms of these tensors. We have

$$\begin{aligned} T^{\mu\alpha}_{\beta\lambda} &= \tfrac{1}{6}\epsilon^{\mu\alpha\rho}S_{\beta\lambda\rho} + \tfrac{1}{6}\epsilon_{\beta\lambda\rho}\bar{S}^{\mu\alpha\rho}, \\ U^\alpha_\beta &= \tfrac{1}{6}\epsilon^{\alpha\rho\mu}J^\lambda_\mu S_{\beta\lambda\rho} - \tfrac{1}{6}\epsilon_{\beta\rho\mu}\bar{S}^{\alpha\lambda\rho}J^\mu_\lambda. \end{aligned} \tag{3.11}$$

In the second term in U, we have interchanged the order of the operators $\bar{S}$ and J and arranged the indices to give a minus sign. One can check by using the commutation relations between J and $\bar{S}$ that this interchange introduces no extra terms. The point of writing U in this form is that one now recognizes one term to be essentially the Hermitian conjugate of the other; i.e., one has

$$\begin{aligned} U^\alpha_\beta &= \tfrac{1}{6}[V^\alpha_\beta - (V^\beta_\alpha)^\dagger], \\ V^\alpha_\beta &= \epsilon^{\alpha\rho\mu}J^\lambda_\mu S_{\beta\lambda\rho}. \end{aligned} \tag{3.12}$$

Therefore, the RME's of U can be written in terms of those of V in the following way[21]:

$$\begin{aligned} &\langle\lambda'\mu'; I'_0\| \ U \ \|\lambda\mu; I_0\rangle_\gamma \\ &\quad = \tfrac{1}{6}[\langle\lambda'\mu'; I'_0\| \ V \ \|\lambda\mu; I_0\rangle_\gamma \\ &\qquad - (-1)^{\lambda'-\lambda}[d(\lambda\mu)/d(\lambda'\mu')]^{\frac{1}{2}} \langle\lambda\mu; I_0\| \ V \ \|\lambda'\mu'; I'_0\rangle_\gamma]. \end{aligned} \tag{3.13}$$

Since the RME's of V turn out to be real, the complex conjugation of the second term here has been omitted.

The job of relating the RME's of V to those of S is only slightly more complicated than finding a similar relation between $(PJ - JP)$ and P. First we can check whether any of the $SU(3)$ representation-diagonal RME's of V vanish. In contrast to the last section, now we find only one relation

$$J^\beta_\alpha V^\alpha_\beta = 0. \tag{3.14}$$

The invariant formed from V and JJ cannot be shown to vanish. We can only conclude that the $\gamma = 1$ representation-diagonal RME's of V is zero, since the matrix elements of J are essentially the CG coefficients with $\gamma = 1$, and these are orthogonal to the coefficients with $\gamma = 2$. So we have

$$\langle\lambda\mu; I'_0\| \ V \ \|\lambda\mu; I_0\rangle_\gamma = 0 \quad \text{for} \quad \gamma = 1. \tag{3.15}$$

For the seven remaining possibilities, we write[22]

$$\begin{aligned} &\langle\lambda'\mu'; I'_0\| \ V \ \|\lambda\mu; I_0\rangle_\gamma \\ &\qquad = b(\lambda\mu; \lambda'\mu')\,\langle\lambda'\mu'; I'_0\| \ S \ \|\lambda\mu; I_0\rangle, \end{aligned} \tag{3.16}$$

it being understood that $\gamma = 2$, in case $(\lambda'\mu') = (\lambda\mu)$. Next we can look for conditions on $b(\lambda\mu; \lambda'\mu')$ analogous to Eqs. (2.4) and (2.5). Since Hermitian conjugation does not take V into itself, this can give no relation. However, R conjugation does yield the relation

$$b(\lambda\mu; \lambda'\mu') = \eta_{\lambda\mu,\lambda'\mu',\gamma}(-1)^{\frac{1}{3}(\lambda'-\mu'+\mu-\lambda)}b(\mu'\lambda'; \mu\lambda). \tag{3.17}$$

Here, $\eta_{\lambda\mu,\lambda'\mu',\gamma}$ is a phase factor that appears in the conjugation relation for CG coefficients relating to the UIR (1, 1) of $SU(3)$; its precise definition and values are given in the Appendix. It suffices to evaluate $b(\lambda\mu; \lambda'\mu')$ for four cases, the others following from (3.17). The technique of calculation is the same as in the previous section: Namely, one takes a suitable component of V, evaluates its matrix element between carefully chosen initial and final states by putting in a complete set of intermediate states between the components of J and S, factors out the proper $SU(3)$-CG coefficient, and then reads off the value of the coefficient $b(\lambda\mu; \lambda'\mu')$.[23] Note that one need not know the reduced matrix elements of S for this work. In this way, one gets

$$\begin{aligned} b(\lambda\mu; \lambda'\mu') = &-\tfrac{1}{6}[2(\lambda+\mu+1)(\lambda+\mu+4)]^{\frac{1}{2}}, \\ &\tfrac{1}{6}[2(\mu-1)(\mu+2)]^{\frac{1}{2}}, \ \tfrac{1}{6}[2\mu(\mu+3)]^{\frac{1}{2}}, \\ &-\tfrac{1}{6}[2\lambda(\lambda+3)]^{\frac{1}{2}}, \ -\tfrac{1}{6}[2(\lambda-1)(\lambda+2)]^{\frac{1}{2}}, \\ &-\tfrac{1}{6}[2(\lambda+\mu)(\lambda+\mu+3)]^{\frac{1}{2}}, \\ &-[I_2(\lambda,\mu)/2]^{\frac{1}{2}}, \end{aligned}$$

for

$$(\lambda'\mu') = (\lambda+2, \mu-1), (\lambda-1, \mu-1), (\lambda-2, \mu+1), (\lambda+1, \mu+1), (\lambda-1, \mu+2), (\lambda+1, \mu-2), (\lambda, \mu). \quad (3.18)$$

To complete the analysis of this section, it is necessary to know the structure of the RME's of the tensor S.[24] This is worked out in the Appendix. The expression is

$$\langle \lambda'\mu'; I_0' \| S \| \lambda\mu; I_0 \rangle = 18(2)^{\frac{1}{2}}(p_0q_1 - p_1q_0)[d(\lambda\mu)/d(\lambda'\mu')]^{\frac{1}{2}} \times C(\lambda\mu\, 30\, \lambda'\mu'; I_0M_0Y_0\, 100\, I_0'M_0Y_0). \quad (3.19)$$

The analysis that leads to this form for the RME's of S is very similar to that which led to Eq. (1.13b). Just like P, S is an irreducible tensor with commuting components, the difference being that it belongs to the UIR (3, 0) while P belongs to the UIR (1, 1). Thus, while the RME of P are given in terms of CG coefficients involving the UIR (1, 1), those of S involve the CG coefficients for the UIR (3, 0).

Combining Eqs. (3.13), (3.16), and (3.19), we have the final expression for the RME's of U as

$$\langle \lambda'\mu'; I_0' \| U \| \lambda\mu; I_0 \rangle_\gamma = 3(2)^{\frac{1}{2}}(p_0q_1 - p_1q_0)[b(\lambda\mu; \lambda'\mu')[d(\lambda\mu)/d(\lambda'\mu')]^{\frac{1}{2}} \times C(\lambda\mu\, 30\, \lambda'\mu'; I_0M_0Y_0\, 100\, I_0'M_0Y_0) - (-1)^{\lambda'-\lambda}b(\lambda'\mu'; \lambda\mu) \times C(\lambda'\mu'\, 30\, \lambda\mu; I_0'M_0Y_0\, 100\, I_0M_0Y_0)]. \quad (3.20)$$

We remind ourselves that, for $(\lambda'\mu') = (\lambda\mu)$, only $\gamma = 2$ is relevant, the RME's with $\gamma = 1$ being zero.

4. MATRIX ELEMENTS OF K

We have at hand all the information necessary to write down the RME's of all the $SL(3, C)$ generators in an $SU(3)$ basis. The precise definition of the basis states has been given in Sec. 1, and the RME's of the $SU(3)$ generators J have also been written down in Eqs. (1.13a) and (1.16a). All that now remains is to use Eqs. (1.19) and (1.24) and the results of the last two sections, to work out the RME's of the "noncompact" generators K. Since this is simpler for the principal degenerate series, we tackle this case first, and the principal nondegenerate series later.

Case (I): Principal Degenerate Series. Here we need only the results of Sec. 2. Putting together Eqs. (1.16b), (1.23), (1.24), and (2.3), we have the result

$$\langle \lambda'\mu' \| K \| \lambda\mu \rangle_\gamma = \tfrac{1}{3}[\rho_3 + ia(\lambda\mu; \lambda'\mu')] \times [d(\lambda\mu)/d(\lambda'\mu')]^{\frac{1}{2}} \times C(\lambda\mu\, 11\, \lambda'\mu'\gamma; 00\, Y_0\, 000\, 00\, Y_0). \quad (4.1)$$

It must be remembered that $a(\lambda\mu; \lambda\mu)$ vanishes so that, though the $SU(3)$ off-diagonal RME's of K are complex, the $SU(3)$ diagonal ones are real.[25] However, by making suitable phase transformations on the basic states, one can pass to a new basis in which all the RME of K become real. For this, we remember that the $SU(3)$ content of a UIR of the principal degenerate series is very simply stated: Each UIR (λ, μ) for which $\mu - \lambda = \frac{3}{2}Y_0$ appears once. Therefore, it is possible to order the $SU(3)$ UIR's in a definite sequence, in the form

$$(\lambda_{\min}, \mu_{\min}), (\lambda_{\min}+1, \mu_{\min}+1), \cdots, (\lambda, \mu), (\lambda+1, \mu+1), \cdots.$$

If Y_0 is positive (negative), then $\lambda_{\min}(\mu_{\min})$ is zero and $\mu_{\min}(-\lambda_{\min})$ is $\frac{3}{2}Y_0$. In Eq. (4.1), a given UIR (λ, μ) is connected by K only to the three UIR's $(\lambda+1, \mu+1)$, (λ, μ), and $(\lambda-1, \mu-1)$, because along with (λ, μ) these are the only UIR's (λ', μ') appearing in the space of the representation. The only values of $a(\lambda\mu; \lambda'\mu')$ that are needed are

$$a(\lambda\mu; \lambda+1, \mu+1) = (\lambda+\mu+3), \quad a(\lambda\mu; \lambda-1, \mu-1) = -(\lambda+\mu+1). \quad (4.2)$$

Therefore, we can define a set of angles $\psi(\lambda, \mu)$ by

$$\rho_3 + ia(\lambda\mu; \lambda+1, \mu+1) = [\rho_3^2 + (\lambda+\mu+3)^2]^{\frac{1}{2}} \exp i\psi(\lambda, \mu),$$
$$\rho_3 + ia(\lambda\mu; \lambda-1, \mu-1) = [\rho_3^2 + (\lambda+\mu+1)^2]^{\frac{1}{2}} \exp[-i\psi(\lambda-1, \mu-1)]. \quad (4.3)$$

For the two $SU(3)$ off-diagonal RME's of K, we then have the expressions

$$\langle \lambda+1, \mu+1 \| K \| \lambda\mu \rangle = \exp i\psi(\lambda, \mu) \times \tfrac{1}{3}\{d(\lambda\mu)[\rho_3^2 + (\lambda+\mu+3)^2]/d(\lambda+1, \mu+1)\}^{\frac{1}{2}} \times C(\lambda\mu\, 11\, \lambda+1\, \mu+1; 00\, Y_0\, 000\, 00\, Y_0),$$
$$\langle \lambda-1, \mu-1 \| K \| \lambda\mu \rangle = \exp[-i\psi(\lambda-1, \mu-1)] \times \tfrac{1}{3}[d(\lambda\mu)\{\rho_3^2 + (\lambda+\mu+1)^2\}/d(\lambda-1, \mu-1)]^{\frac{1}{2}} \times C(\lambda\mu\, 11\, \lambda-1\, \mu-1; 00\, Y_0\, 000\, 00\, Y_0). \quad (4.4)$$

It is now a trivial matter to transform away these phase factors. We define a new set of basis vectors by

$$|\lambda\mu; IMY) = \exp[-i\phi(\lambda\mu)] \times |\lambda\mu; IMY\rangle, \quad (4.5)$$

the phases $\phi(\lambda\mu)$ being determined by the difference equation

$$\phi(\lambda+1, \mu+1) - \phi(\lambda\mu) = \psi(\lambda\mu), \quad \phi(\lambda_{\min}, \mu_{\min}) = 0. \quad (4.6)$$

This is really a single-variable difference equation since $\lambda - \mu$ is constant. In the new basis, all the RME's of K are real, and we have, in place of Eq. (4.1), the final result

$$(\lambda'\mu' \| K \| \lambda\mu)_\gamma = \tfrac{1}{3}\{d(\lambda\mu)[\rho_3^2 + a(\lambda\mu; \lambda'\mu')^2]/d(\lambda'\mu')\}^{\frac{1}{2}} \times C(\lambda\mu\ 11\ \lambda'\mu'\gamma; 00Y_0\ 000\ 00Y_0), \quad (4.7)$$

for the RME's of K in the UIR (m_3, ρ_3) of $SL(3, C)$. It should be mentioned that the RME's of J are unaltered by this change of basis, since J only has representation-diagonal matrix elements. Thus, Eq. (1.16a) is unchanged with round brackets substituted for angular ones.

We have now completely solved the problem of obtaining the RME's of the $SL(3, C)$ generators in an $SU(3)$ basis, for the principal degenerate series. Such a calculation had previously been performed for the particular case when $Y_0 = 0$, i.e., for the subset of degenerate UIR's $(0, \rho_3)$, by directly solving the $SL(3, C)$ commutation rules and taking advantage of the fact that these UIR's do not have multiplicity problems.[26] Our results could also have been obtained by this method.

Case (2): Principal Nondegenerate Series. We now need the results of both Secs. 2 and 3. From Eqs. (1.13b), (1.18), (1.19), (1.26), (2.3), and (3.20), we have, for the RME's of K in the UIR $(m_2, m_3, \rho_2, \rho_3)$, the result

$$\begin{aligned}
&\langle \lambda'\mu'; I_0' \| K \| \lambda\mu; I_0 \rangle_\gamma \\
&\quad = \tfrac{1}{6}[d(\lambda\mu)/d(\lambda'\mu')]^{\frac{1}{2}}[2\rho_3 - \rho_2 + 2ia(\lambda\mu; \lambda'\mu')] \\
&\qquad \times C(\lambda\mu\ 11\ \lambda'\mu'\gamma; I_0M_0Y_0\ 000\ I_0'M_0Y_0) \\
&\qquad - \tfrac{1}{6}(3)^{-\frac{1}{2}}[d(\lambda\mu)/d(\lambda'\mu')]^{\frac{1}{2}}[3\rho_2 + 2ia(\lambda\mu; \lambda'\mu')] \\
&\qquad \times C(\lambda\mu\ 11\ \lambda'\mu'\gamma; I_0M_0Y_0\ 100\ I_0'M_0Y_0) \\
&\qquad + i(6)^{-\frac{1}{2}}[b(\lambda\mu; \lambda'\mu')\{d(\lambda\mu)/d(\lambda'\mu')\}^{\frac{1}{2}} \\
&\qquad \times C(\lambda\mu\ 30\ \lambda'\mu'; I_0M_0Y_0\ 100\ I_0'M_0Y_0) \\
&\qquad - (-1)^{\lambda'-\lambda} b(\lambda'\mu'; \lambda\mu) \\
&\qquad \times C(\lambda'\mu'\ 30\ \lambda\mu; I_0'M_0Y_0\ 100\ I_0M_0Y_0)]. \quad (4.8)
\end{aligned}$$

In arriving at this form, we have collected together and simplified the coefficients of the CG coefficients, which originally involve the parameter s, and then expressed the result in terms of ρ_2 and ρ_3. It is worth noting that the worrisome factor of $s(s+1)$, which appears in the denominator in the terms of the expansion formula (1.19), has disappeared in Eq. (4.8) after combination of all factors. A more symmetrical way of writing the result is to express the last CG coefficient in Eq. (4.8) in terms of the CG coefficients involving the UIR (0, 3) instead of (3, 0).[27] Then we have

$$\begin{aligned}
&6[d(\lambda'\mu')/d(\lambda\mu)]^{\frac{1}{2}} \langle \lambda'\mu'; I_0' \| K \| \lambda\mu; I_0 \rangle_\gamma \\
&\quad = [2\rho_3 - \rho_2 + 2ia(\lambda\mu; \lambda'\mu')] \\
&\qquad \times C(\lambda\mu\ 11\ \lambda'\mu'\gamma; I_0M_0Y_0\ 000\ I_0'M_0Y_0) \\
&\qquad - (3)^{-\frac{1}{2}}[3\rho_2 + 2ia(\lambda\mu; \lambda'\mu')] \\
&\qquad \times C(\lambda\mu\ 11\ \lambda'\mu'\gamma; I_0M_0Y_0\ 100\ I_0'M_0Y_0) \\
&\qquad + i(6)^{\frac{1}{2}}[b(\lambda\mu; \lambda'\mu') \\
&\qquad \times C(\lambda\mu\ 30\ \lambda'\mu'; I_0M_0Y_0\ 100\ I_0'M_0Y_0) \\
&\qquad - b(\lambda'\mu'; \lambda\mu) C(\lambda\mu\ 03\ \lambda'\mu'; I_0M_0Y_0\ 100\ I_0'M_0Y_0)]. \quad (4.9)
\end{aligned}$$

In connection with Eqs. (4.8) and (4.9), one must remember that for $(\lambda'\mu') = (\lambda\mu)$, $a(\lambda\mu; \lambda'\mu')$ vanishes while $b(\lambda\mu; \lambda'\mu')$ contributes only for $\gamma = 2$.

The problem of obtaining the $SL(3, C)$ generator matrices in an $SU(3)$ basis is thereby solved for the principal nondegenerate series of UIR's. One can ask whether by a suitable alteration of the phases of the basis states, one can render all the RME's of K real, as was possible in the case of the degenerate series.[28] We do not attempt to do so here, since that would make this paper unduly long. The situation is a little more complicated now, since it might require that one perform a unitary transformation, for each fixed λ, μ and with respect to I_0, in which one constructs nontrivial superpositions of states with different values of I_0. If this becomes necessary, then the dependence of the RME's of K on the labels I_0, I_0' may no longer be isolated in the CG coefficients, as has been possible in Eqs. (4.8) and (4.9). We hope to discuss this question elsewhere, leaving the RME's of K complex for the present.

5. CONCLUSION

We wish to make some comments on the results of Paper I and the present paper.

Prior to the evaluation of the RME's of K, the structure of the expansion formula (1.19) might have suggested that it would break down in some cases. Certainly, the individual terms of the formula are singular at the point $s = 0$ and $s = -1$, corresponding to the conditions $\rho_3 = \rho_2$ and $\rho_3 = 0$, respectively. However, one would be surprised if, after evaluation of the matrix elements of K, these singularities persisted, because there is nothing to single out these values of ρ_3 from all others when one contemplates the UIR's of $SL(3, C)$ as a totality. Therefore, it is gratifying to see that there are cancellations between various terms of the expansion formula, and we are left with very simple and nonsingular dependences on ρ_2 and ρ_3.

Since the RME's of the $SL(3, C)$ generators have found their simplest expression in terms of various CG coefficients of $SU(3)$, from a practical point of view it is necessary that one employ CG coefficients evaluated within the framework of a single, consistent phase convention for the UIR's of $SU(3)$. Fortunately, mutually consistent tables of the CG coefficients for the UIR's (1, 1), (3, 0), and (0, 3) of $SU(3)$, all of them employing the Biedenharn phase convention for $SU(3)$, do exist in the literature, and these are just the coefficients needed here.[29]

An interesting property of the $SU(3)$ multiplicity label I_0 deserves mention. If one had tried to solve the $SL(3, C)$ commutation relations directly by working from the Lie algebra of $SL(3, C)$ and functions of the elements of the algebra, a natural thing to have done would have been to first construct some operator in terms of J and K which would have been a scalar under $SU(3)$, but which would not commute with K.[30] The eigenvalues of such an operator could possibly be used to distinguish multiple occurrences of $SU(3)$ representations. A possible choice is, for example, the operator

$$\{J^{\alpha}_{\beta}, J^{\beta}_{\gamma}\}_{+}K^{\gamma}_{\alpha}. \tag{5.1}$$

This is an $SU(3)$ scalar, but not a Casimir operator of $SL(3, C)$. However, it is quite a difficult matter to determine the possible eigenvalues of such an operator, and then harder still to determine the RME's of K between eigenvectors of this operator, if all along one uses just the commutation relations of $SL(3, C)$.[31] In the first place, of course, one must make sure that the operator has enough distinct eigenvalues to "tell apart" different vectors behaving identically under $SU(3)$. Thus, it may be easy to write down such operators, but it is difficult to make use of them. On the other hand, the multiplicity label I_0 used by us has a range of values which is easy to describe, and the RME's of K have a fairly simple and easily expressed dependence on it. But it is quite difficult to construct an operator function of J and K which is diagonal in the basis we have chosen and whose eigenvalues are, say, $I_0(I_0 + 1)$ or some nontrivial function thereof! Undoubtedly, such an operator does exist, since J and K form an irreducible set, but we suspect that no finite polynomial function of J and K has the properties described above.

The multiplicity label I_0 originated in the reduction of the regular representation of $SU(3)$ in the Hilbert space $\mathcal{H}$ of square integrable functions on $SU(3)$. According to general theorems, the space $\mathcal{H}$ is large enough to accommodate any UIR of $SL(3, C)$; more generally, the regular representation of the maximal compact subgroup of a semisimple noncompact Lie group can accommodate any UIR of the whole group. These facts confer greater validity to the use of multiplicity labels of the nature of I_0 in general.[32] Some of the practical advantages of working with the space $\mathcal{H}$ and, therefore, with many UIR's of $SL(3, C)$ at the same time in the form of a direct sum have been alluded to before and are more fully explained in the Appendix.

We mentioned in Sec. 1 that the non-Hermitian combinations M, N of the Hermitian generators J and K obey the commutation relations of $SU(3) \otimes SU(3)$. Therefore, one must be able to reach nontrivial finite-dimensional unitary representations of $SU(3) \otimes SU(3)$ in a basis in which the vectors belong to definite UIR's of the "total" $SU(3)$, by a process of analytic continuation of the RME's of K in the principal nondegenerate series of $SL(3, C)$. The parameters that have to be varied to effect this continuation are ρ_2 and ρ_3. This process could probably be carried out with ease only after we have made all the RME's of K real, by a suitable transformation. If this can be done, then one might see a way to defining a complete set of orthonormalized CG coefficients for the most general case in $SU(3)$. It is interesting to observe in this connection that Biedenharn's lemma shows that the way one counts outer multiplicities in the Clebsch–Gordan series for $SU(3)$ is similar to the counting of $SU(3)$ representation multiplicities within UIR's of $SL(3, C)$.[33]

In connection with the possibility of analytic continuation of representation matrices, we remark that the RME's of K in the principal degenerate series have already been brought to the form where we can pass over to the supplementary series of UIR's of $SL(3, C)$, by the method of the master analytic representation.[34] The need for rendering all the RME's real in the principal nondegenerate series, which was mentioned in the previous paragraph, is reinforced by the need to see exactly how the master analytic representation method goes through in a case involving multiplicity.

Turning to possible physical applications of the UIR's of $SL(3, C)$, we refer to recent attempts to incorporate the Lie algbras of $SL(3, C)$ and the Poincaré group in a common associative-algebra framework to describe the spectrum of elementary particle states.[35] Here $SL(3, C)$ acts as an agency for providing a sequence of $SU(3)$ multiplets. If the multiplicity-free degenerate representations of $SL(3, C)$ prove inadequate in this connection, one may then have to utilize the nondegenerate series UIR's worked out here. Needless to say, in such applications the use of an $SU(3)$ basis is absolutely essential!

Last of all, there is the question of using similar methods for other noncompact groups. One way is to proceed in more or less the path we follow, by discovering expansion formulas. However, possibly a more useful (and therefore harder!) way might be to try to understand precisely what there is in the structure of the group $SL(3, C)$ that allows us to express the RME's of K in terms of the CG coefficients of $SU(3)$ corresponding to the UIR's (1, 1), (3, 0), and (0, 3). What is desirable is the connection of this fact to something intrinsic to $SL(3, C)$, without making mention of the expansion formula at all. If one examines the UIR's of $SL(2, C)$ (in which case there are no multiplicity problems), it turns out that the RME's of the noncompact generator K can be expressed in terms of the CG coefficients of $SU(2)$ corresponding to the single 3-dimensional UIR of $SU(2)$. (This is clear from the discussion in Appendix A of I.) If such an understanding can be achieved, then that might be the most efficient way of extending the present methods to other groups of interest. We hope to come back to this question elsewhere.

ACKNOWLEDGMENTS

The author thanks Dr. P. P. Divakaran and Dr. A. K. Rajagopal for useful discussions. He also thanks Professor A. Böhm for having provided some of the motivation for completion of this work.

APPENDIX

Here we outline the manner in which the RME's of the tensor operator S which occurred in Sec. 3 have been evaluated. We also give the details of the R-conjugation operation which leads to useful results like Eqs. (2.5) and (3.17), and record some useful identities obeyed by those $SU(3)$ CG coefficients that have appeared in this paper. The formalism used is a specialization to $SU(3)$ of the general one given in Appendix B of I.

Let us refer to elements of $SU(3)$ by $a, a', \cdots$, etc., and to the corresponding 3×3 unitary unimodular matrices by $A^{\alpha}_{\beta}, A'^{\alpha}_{\beta}, \cdots$, etc. The group composition corresponds to multiplication of these matrices with the superscript (subscript) being the row (column) index. We define the Hilbert space $\mathcal{H}$ to be the space of all square-integrable functions on $SU(3)$. A continuous basis for $\mathcal{H}$ consists of the vectors $|a\rangle$, with normalization properties as given in Eqs. (B5) and (B6) of I. The regular representation of $SU(3)$ associates, with each a, the unitary operator $U(a)$, which acts as follows:

$$U(a)\,|a'\rangle = |aa'\rangle. \tag{A1}$$

The reduction of this representation is accomplished by passing to the discrete orthonormal basis

$$|\lambda\mu;\, IMY;\, I_0M_0Y_0\rangle, \tag{A2}$$

related to the continuous basis by Eq. (B9) of I.[8] Since we need it later, we write down the expansion

$$|a\rangle = \sum_{\lambda\mu IMY I_0M_0Y_0} [d(\lambda\mu)]^{\frac{1}{2}}(a)^{(\lambda\mu)}_{IMY,I_0M_0Y_0} \times |\lambda\mu;\, IMY;\, I_0M_0Y_0\rangle. \tag{A3}$$

Apart from the dimensionality factor, the coefficient of the vector on the right-hand side of (A3) is the matrix representing the element a in the UIR $(\lambda\mu)$, expressed in the canonical basis labeled by IMY. We define a set of orthogonal subspaces $\mathcal{H}(M_0, Y_0)$ of $\mathcal{H}$ by defining the vectors (A2) for fixed M_0, Y_0 and all possible values of the other labels to form a basis for $\mathcal{H}(M_0, Y_0)$. So we have the decomposition

$$\mathcal{H} = \sum_{M_0Y_0} \oplus\, \mathcal{H}(M_0, Y_0). \tag{A4}$$

Since we wish to be able to deal with tensor operators labeled with upper and lower tensor indices or with their canonical components labeled by IMY, interchangeably, we write down the transformation laws in both languages. For a tensor operator belonging to the UIR (1, 1), the transformation law

$$U(a)P^{\alpha}_{\beta}U(a^{-1}) = (A^{\lambda}_{\alpha})^{*}A^{\mu}_{\beta}P^{\lambda}_{\mu} \tag{A5}$$

is exactly equivalent to the statement

$$U(a)P^{(1,1)}_{IMY}U(a^{-1}) = \sum_{I'M'Y'} (a)^{(1,1)}_{I'M'Y',IMY}P^{(1,1)}_{I'M'Y'}, \tag{A6}$$

which is Eq. (B4b) of I for this case. For the UIR (3, 0), the two equations are

$$\begin{aligned} U(a)S_{\alpha\beta\gamma}U(a^{-1}) &= A^{\lambda}_{\alpha}A^{\mu}_{\beta}A^{\nu}_{\gamma}S_{\lambda\mu\nu}, \\ U(a)S^{(3,0)}_{IMY}U(a^{-1}) &= \sum_{I'M'Y'} (a)^{(3,0)}_{I'M'Y',IMY}S^{(3,0)}_{I'M'Y'}. \end{aligned} \tag{A7}$$

The relation between the two sets of components for the UIR (1, 1) is given in Eq. (3.3) of I; for the UIR (3, 0), it is

$$\begin{aligned} &S_{\frac{3}{2}\frac{3}{2}1} = S_{111},\ S_{\frac{3}{2}\frac{1}{2}1} = (3)^{\frac{1}{2}}S_{112},\ S_{\frac{3}{2}-\frac{1}{2}1} = (3)^{\frac{1}{2}}S_{122}, \\ &S_{\frac{3}{2}-\frac{3}{2}1} = S_{222},\ S_{110} = (3)^{\frac{1}{2}}S_{113},\ S_{100} = (6)^{\frac{1}{2}}S_{123}, \\ &S_{1-10} = (3)^{\frac{1}{2}}S_{223},\ S_{\frac{1}{2}\frac{1}{2}-1} = (3)^{\frac{1}{2}}S_{133},\ S_{\frac{1}{2}-\frac{1}{2}-1} \\ &\quad = (3)^{\frac{1}{2}}S_{233},\ S_{00-2} = S_{333}. \end{aligned} \tag{A8}$$

We can now define a reducible representation of $SU(3) \times T_8$ in $\mathcal{H}$ by setting

$$P^{\alpha}_{\beta}\,|a\rangle = p^{\lambda}_{\mu}A^{\alpha}_{\lambda}(A^{\beta}_{\mu})^{*}\,|a\rangle. \tag{A9}$$

This is Eq. (B8) of I, use having been made of the unitarity of the matrix A. We choose the 3×3 traceless Hermitian matrix p_μ^λ to be diagonal, with elements

$$p_1^1 = -p_0 + (3)^{\frac{1}{2}}p_1,\ p_2^2 = -p_0 - (3)^{\frac{1}{2}}p_1,$$
$$p_3^3 = 2p_0. \tag{A10}$$

Assuming that p_0, p_1 obey the conditions for a type I UIR of $SU(3) \times T_8$, this representation of $SU(3) \times T_8$ breaks up into a sum of type-I UIR's, the subspace $\mathcal{H}(M_0, Y_0)$ carrying the UIR (M_0, Y_0, p_0, p_1). Note that p_0, p_1 are the same for all the UIR's in the direct sum. Let us now calculate the values of the Casimir operator $C_2^{(0)}$ of $SU(3) \times T_8$. If e denotes the identity element of $SU(3)$, then from Eq. (A3) we have

$$|e\rangle = \sum_{\lambda\mu I_0 M_0 Y_0} [d(\lambda\mu)]^{\frac{1}{2}} |\lambda\mu; I_0 M_0 Y_0; I_0 M_0 Y_0\rangle. \tag{A11}$$

Thus, this vector $|e\rangle$ is a superposition of contributions from all the subspaces $\mathcal{H}(M_0, Y_0)$. Now the operator $C_2^{(0)}$ reduces within each $\mathcal{H}(M_0, Y_0)$ to a multiple of the identity, the magnitude being a function of M_0, Y_0 (and, of course, p_0, p_1). We can easily get the value of $C_2^{(0)}$ in each subspace by applying the operator $C_2^{(0)}$ to the vector $|e\rangle$ and seeing by what numerical factor each term on the right-hand side of (A11) gets multiplied. Using the fact that p_μ^λ is diagonal, we get

$$\begin{aligned} C_2^{(0)} |e\rangle &\equiv \tfrac{1}{2} J_\beta^\alpha P_\alpha^\beta |e\rangle \\ &= \tfrac{1}{2}(J_1^1 p_1^1 + J_2^2 p_2^2 + J_3^3 p_3^3) |e\rangle \\ &= \tfrac{1}{2}[p_1(3)^{\frac{1}{2}}(J_1^1 - J_2^2) + 3p_0 J_3^3] |e\rangle \\ &= \tfrac{1}{2} \sum_{\lambda\mu I_0 M_0 Y_0} [d(\lambda\mu)]^{\frac{1}{2}} [2(3)^{\frac{1}{2}} p_1 M_0 - 3p_0 Y_0] \\ &\quad \times |\lambda\mu; I_0 M_0 Y_0; I_0 M_0 Y_0\rangle. \end{aligned} \tag{A12}$$

Comparing (A11) and (A12), we read off the value of $C_2^{(0)}$ in the UIR (M_0, Y_0, p_0, p_1), and get the expression given in Eq. (1.9). A similar calculation leads to the value of $C_3^{(0)}$.

The ease with which we obtain the values of these Casimir operators demonstrates the advantages of working with a sum of infinitely many UIR's in the space $\mathcal{H}$. The point is that many calculations are most efficiently performed in the continuous basis in $\mathcal{H}$; specialization to the various UIR's only involves restriction of the results to the subspaces $\mathcal{H}(M_0, Y_0)$.

Following the remarks made in Sec. 1, just after Eq. (1.25), we define a set of operators K_β^α in terms of the generators J, P of the reducible representation of $SU(3) \times T_8$ in $\mathcal{H}$, by Eq. (1.19), after imposing Eq. (1.18) on the parameters p_0 and p_1. Then J and K give a reducible representation of $SL(3, C)$ in $\mathcal{H}$, the subspace $\mathcal{H}(M_0, Y_0)$ carrying the UIR $(m_2, m_3, \rho_2, \rho_3)$ of $SL(3, C)$ [cf. Eq. (1.11)]. The values of ρ_2 and ρ_3 do not change from one UIR to another in this direct sum. Now any function of J and P and, therefore, each of the terms on the right-hand side of Eq. (1.19) leave each subspace $\mathcal{H}(M_0, Y_0)$ invariant. This is also true of the tensor S belonging to the UIR $(3, 0)$, introduced in Sec. 3. We get the RME's of S by first finding out its action on the states $|a\rangle$. The tensor Q obeys Eq. (A9) with the matrix p_μ^λ replaced by another traceless Hermitian diagonal matrix q_μ^λ with diagonal elements

$$\begin{aligned} q_1^1 = -q_0 + (3)^{\frac{1}{2}}q_1,\quad q_2^2 = -q_0 - (3)^{\frac{1}{2}}q_1, \\ q_3^3 = 2q_0, \\ q_0 = p_0^2 - p_1^2,\quad q_1 = -2p_0 p_1. \end{aligned} \tag{A13}$$

Since both P and Q are diagonal in the basis $|a\rangle$, it follows that the tensors T and S given in Eqs. (3.2) and (3.7) are also diagonal in this basis. In fact, we have

$$\begin{aligned} T_{\beta\lambda}^{\mu\alpha} |a\rangle &= t_{\beta'\lambda'}^{\mu'\alpha'} A_{\mu'}^{\mu} (A_{\beta'}^{\beta})^* A_{\alpha'}^{\alpha} (A_{\lambda'}^{\lambda})^* |a\rangle, \\ S_{\alpha\beta\gamma} |a\rangle &= s_{\alpha'\beta'\gamma'} (A^{-1})_\alpha^{\alpha'} (A^{-1})_\beta^{\beta'} (A^{-1})_\gamma^{\gamma'} |a\rangle, \end{aligned} \tag{A14}$$

where t is obtained from Eq. (3.2), by replacing the operators P, Q by the numerical matrices p and q, and in the same way s comes from Eq. (3.7) after replacing T by t. The second of Eqs. (A14) is of the same form as Eq. (B8) of I; written in terms of canonical components, it reads

$$S_{IMY}^{(3,0)} |a\rangle = \sum_{I'M'Y'} s_{I'M'Y'} (a^{-1})_{I'M'Y',IMY}^{(3,0)} |a\rangle. \tag{A15}$$

[cf. Eq. (A8) for going from tensor to canonical components.] We can therefore use Eq. (B17) of I to directly write down the RME's of S. The only nonzero component of s turns out to be

$$s_{100} = 18(2)^{\frac{1}{2}}(p_0 q_1 - p_1 q_0), \tag{A16}$$

and this immediately leads to Eq. (3.19) of the text. The simplification achieved by working in the basis $|a\rangle$ for $\mathcal{H}$, rather than restricting oneself from the start to the space of a single UIR of $SU(3) \times T_8$, is once again evident.

We describe next the operation of R conjugation. We find it simplest to define R to be antilinear, antiunitary, and diagonal in the basis $|a\rangle$ in $\mathcal{H}$:

$$R |a\rangle = |a\rangle. \tag{A17}$$

From this and Eq. (A1) follows the characteristic behavior of the $SU(3)$ generators

$$R J_\beta^\alpha R^{-1} = -J_\alpha^\beta. \tag{A18}$$

To know the action of R on the discrete basis states (A2), it is necessary to have the precise connection between the representation matrices for the UIR's $(\lambda\mu)$ and $(\mu\lambda)$. With the standard Biedenharn phase

conventions for $SU(3)$, this is[36]

$$[(a)^{(\lambda\mu)}_{IMY,I'M'Y'}]^* = (-1)^{M'-M-\frac{1}{2}(Y'-Y)}(a)^{(\mu\lambda)}_{I-M-Y,I'-M'-Y'}. \quad \text{(A19)}$$

From this and Eq. (B9) of I,[3] we get

$$R\,|\lambda\mu; IMY; I_0M_0Y_0\rangle = (-1)^{M-M_0-\frac{1}{2}(Y-Y_0)} \times |\mu\lambda; I-M-Y; I_0-M_0-Y_0\rangle. \quad \text{(A20)}$$

We next work out the effect of R on various tensor operators. From Eqs. (A9), (A17), and the reality of the diagonal elements of p^λ_μ, we get

$$RP^\alpha_\beta R^{-1} = (P^\alpha_\beta)^\dagger. \quad \text{(A21)}$$

In terms of the canonical components, the relation is

$$RP_{IMY}R^{-1} = P^\dagger_{IMY}. \quad \text{(A22)}$$

Combining Eqs. (A18) and (A21), we get

$$R(PJ-JP)^\alpha_\beta R^{-1} = -(PJ-JP)^{\alpha\dagger}_\beta. \quad \text{(A23)}$$

From the definitions of the tensors S and V in Sec. 3, we can work out the effect of R on them:

$$RS_{\alpha\beta\gamma}R^{-1} = S^\dagger_{\alpha\beta\gamma}, \quad RV^\alpha_\beta R^{-1} = -V^{\alpha\dagger}_\beta. \quad \text{(A24)}$$

These equations imply specific consequences for the RME's of these operators. To arrive at these consequences, we need two relations obeyed by the $SU(3)$ CG coefficients pertaining to the UIR's (1, 1) and (3, 0). These relations are[37]

$$C(\lambda\mu\ 11\ \lambda'\mu'\gamma; IMY\ JNZ\ I'M'Y') = \eta_{\lambda\mu\lambda'\mu'\gamma}(-1)^{\frac{1}{2}Z-N}[d(\lambda'\mu')/d(\lambda\mu)]^{\frac{1}{2}} \times C(\mu'\lambda'\ 11\ \mu\lambda\gamma; I'-M'-Y'\ JNZ\ I-M-Y) \quad \text{(A25)}$$

and

$$C(\lambda\mu\ 30\ \lambda'\mu'; IMY\ JNZ\ I'M'Y') = -(-1)^{\frac{1}{2}(\lambda'-\mu'+\mu-\lambda)+\frac{1}{2}Z-N}[d(\lambda'\mu')/d(\lambda\mu)]^{\frac{1}{2}} \times C(\mu'\lambda'\ 30\ \mu\lambda; I'-M'-Y'\ JNZ\ I-M-Y). \quad \text{(A26)}$$

Here η is a phase factor with the value $+1$ for $(\lambda'\mu') = (\lambda-2, \mu+1)$, $(\lambda-1, \mu+2)$, and (λ, μ), $\gamma = 2$, and the value -1 for the other cases.

Since we know the effect of R both on the states and on the operators, we can work out the relations obeyed by the RME's. For P, for instance, we get

$$\langle\lambda'\mu'; I'_0M_0Y_0\|\ P\ \|\lambda\mu; I_0M_0Y_0\rangle_\gamma = \eta_{\lambda\mu\lambda'\mu'\gamma}[d(\lambda\mu)/d(\lambda'\mu')]^{\frac{1}{2}} \times \langle\mu\lambda; I_0-M_0-Y_0\|\ P\ \|\mu'\lambda'; I'_0-M_0-Y_0\rangle_\gamma. \quad \text{(A27)}$$

Here we have taken the same values for M_0 and Y_0 in both states, since in our application these lead to the only nonvanishing RME's of P; we have also used the reality of the RME's. For the RME's of the operator $(PJ-JP)$, we get a relation which differs from (A27) only in having an extra minus sign on the right because of the extra minus sign in (A23) as compared to (A21):

$$\langle\lambda'\mu'; I'_0M_0Y_0\|\ PJ-JP\ \|\lambda\mu; I_0M_0Y_0\rangle_\gamma = -\eta_{\lambda\mu\lambda'\mu'\gamma}[d(\lambda\mu)/d(\lambda'\mu')]^{\frac{1}{2}} \times \langle\mu\lambda; I_0-M_0-Y_0\|\ PJ-JP\ \|\mu'\lambda'; I'_0-M_0-Y_0\rangle_\gamma. \quad \text{(A28)}$$

From Eqs. (A27) and (A28), it is easy to see that Eq. (2.5) follows. Since the operators V and $(PJ-JP)$ behave in the same way under R, as shown by Eqs. (A23) and (A24), Eq. (A28) holds for the RME's of V also. (These RME's are also real.) On the other hand, for the RME's of S, we have the relation

$$\langle\lambda'\mu'; I'_0M_0Y_0\|\ S\ \|\lambda\mu; I_0M_0Y_0\rangle = -(-1)^{\frac{1}{2}(\lambda'-\mu'+\mu-\lambda)}[d(\lambda\mu)/d(\lambda'\mu')]^{\frac{1}{2}} \times \langle\mu\lambda; I_0-M_0-Y_0\|\ S\ \|\mu'\lambda'; I'_0-M_0-Y_0\rangle. \quad \text{(A29)}$$

The condition (3.17) on the coefficients $b(\lambda\mu; \lambda'\mu')$, now follows from Eq. (A28), written for V, and Eq. (A29).

The last item to be mentioned is the relation imposed by Hermiticity conditions on the RME's of an operator belonging to the UIR (1, 1). This relation was used in going from Eq. (3.12) to Eq. (3.13), as well as in getting Eq. (2.4).

Let T^α_β be an arbitrary octet tensor operator obeying the condition

$$T^{\alpha\dagger}_\beta = T^\beta_\alpha. \quad \text{(A30)}$$

This can also be written as

$$T^{(1,1)\dagger}_{IMY} = (-1)^{M-\frac{1}{2}Y}T^{(1,1)}_{I-M-Y}. \quad \text{(A31)}$$

To derive the consequences of this property, we need a corresponding relation for the CG coefficients, and this turns out to be[38]

$$C(\lambda\mu\ 11\ \lambda'\mu'\gamma; IMY\ JNZ\ I'M'Y') = (-1)^{\lambda'-\lambda+\frac{1}{2}Z-N}[d(\lambda'\mu')/d(\lambda\mu)]^{\frac{1}{2}} \times C(\lambda'\mu'\ 11\ \lambda\mu\gamma; I'M'Y'\ J-N-Z\ IMY). \quad \text{(A32)}$$

We then get

$$\langle\lambda'\mu'; \beta\|\ T\ \|\lambda\mu; \alpha\rangle^*_\gamma = (-1)^{\lambda'-\lambda}[d(\lambda\mu)/d(\lambda'\mu')]^{\frac{1}{2}} \times \langle\lambda\mu; \alpha\|\ T\ \|\lambda'\mu'; \beta\rangle_\gamma. \quad \text{(A33)}$$

Here α and β are any $SU(3)$ invariant labels the vectors may carry, and no assumption has been made about the reality of the RME's of T. This relation can easily be generalized to an operator like V^{α}_{β} in Sec. 3, which does not obey any relation like (A30). In that case, we get a connection between the RME's of V and those of the (distinct) tensor operator $V^{\dagger}$, whose components are defined by

$$(V^{\dagger})^{\alpha}_{\beta} = (V^{\beta}_{\alpha})^{\dagger}, \tag{A34}$$

the relation being

$$\langle \lambda'\mu'; \beta \| V \| \lambda\mu; \alpha \rangle^{*}_{\gamma} = (-1)^{\lambda'-\lambda}[d(\lambda\mu)/d(\lambda'\mu')]^{\frac{1}{2}} \times \langle \lambda\mu; \alpha \| V^{\dagger} \| \lambda'\mu'; \beta \rangle_{\gamma}. \tag{A35}$$

In our use of Eq. (A35), we have operators V, $V^{\dagger}$ with real RME's, which allow us to omit the complex conjugation on the left-hand side. Finally, the RME's of $(PJ - JP)$ obey Eq. (A33), with an extra minus sign on the right (and no complex conjugation on the left), because the corresponding Hermiticity relation differs by a sign from Eq. (A30). This explains the origin of Eq. (2.4).

[1] For the construction of the unitary representations of the groups $SL(n, C)$, the standard reference is I. M. Gel'fand and M. A. Neumark, *Unitäre Darstellungen der Klassischen Gruppen* (Akademie-Verlag, Berlin, 1957).

[2] An investigation of cases where rather simple solutions to this problem exist has recently been carried out by D. W. Joseph and C. J. Hieggelke, J. Math. Phys. **11**, 272 (1970).

[3] N. Mukunda, J. Math. Phys. **10**, 897 (1969). We refer to this paper as I. Further references may be found there. We take this opportunity to inform the reader that, in Eq. (B9) of this paper, the subscripts m and n on the right-hand side should be interchanged.

[4] C. J. Goebel, Phys. Rev. Letters **16**, 1130 (1966). See also Appendix B of I.

[5] Repeated indices are to be summed over.

[6] The Casimir invariants of $SU(3)$ and their values in the UIR (λ, μ) are $J^{\alpha}_{\beta}J^{\beta}_{\alpha} = 6I_2(\lambda, \mu)$, $\frac{1}{2}\{J^{\alpha}_{\beta}, J^{\beta}_{\gamma}\}_{+}J^{\gamma}_{\alpha} = 18I_3(\lambda, \mu)$.

[7] For evaluating the C_i, we have used Eqs. (4.13), (4.17), and (4.18) of I.

[8] Here we have used Eqs. (5.9), (5.12), and (5.13) of I.

[9] When there is no danger of confusion, we omit tensor indices on tensor operators when they are referred to in the text.

[10] The precise connection is contained in Eqs. (1.22) and (1.25).

[11] We do not consider the trivial representations in which $P = 0$. This was not explicitly mentioned in I.

[12] For details regarding UIR's of $SU(3)$, see, for instance, N. Mukunda and L. K. Pandit, J. Math. Phys. **6**, 746 (1965), and other papers quoted therein.

[13] If we take the formulas defining a type-I UIR of $SU(3) \times T_8$ and set p_1 equal to zero or $\pm(3)^{\frac{1}{2}}p_0$, the representation breaks up into a direct sum of UIR's of types II and III.

[14] The values of $C_1^{(0)}$ and $C_4^{(0)}$ are given by Eq. (4.30) of I. A simple way to evaluate $C_2^{(0)}$ and $C_3^{(0)}$ is described in the Appendix.

[15] For tensor operators belonging to the UIR (1, 1) of $SU(3)$, such as P, K, and J, the $SU(3)$ RME are defined, as in I, by

$$\langle \lambda'\mu'; I'M'Y'; b | P_{JMZ} | \lambda\mu; IMY; a \rangle = \sum_{\gamma} C(\lambda\mu\, 11\, \lambda'\mu'\gamma; IMY\, JMZ\, I'M'Y') \langle \lambda'\mu'; b \| P \| \lambda\mu; a \rangle_{\gamma}$$

and similarly for K and J. Tables of the Clebsch-Gordan coefficients appearing here are given by J. G. Kuriyan, D. Lurié, and A. J. Macfarlane, J. Math. Phys. **6**, 722 (1965). To go from the tensor components of P to the canonical ones, use Eq. (3.3) of I.

[16] $d(\lambda\mu) = \frac{1}{2}(\lambda + 1)(\mu + 1)(\lambda + \mu + 2)$.

[17] To avoid confusion, the operator written as k in I is here called Q.

[18] This factor is essentially a Racah coefficient of $SU(3)$.

[19] R conjugation is discussed by D. Lurié and A. J. Macfarlane, J. Math. Phys. **5**, 565 (1964). A detailed treatment is also included in the Appendix to the present paper.

[20] $I_m = \frac{1}{2}(\lambda + \mu)$, $Y_m = \frac{1}{3}(\lambda - \mu)$.

[21] For details, see the Appendix, Eq. (A35).

[22] The coefficients $b(\lambda\mu; \lambda'\mu')$ are essentially Racah coefficients of $SU(3)$.

[23] The CG coefficients of $SU(3)$ relating to the UIR (3, 0), needed for writing out the various matrix elements of S, have been compiled by L. K. Pandit and N. Mukunda, J. Math. Phys. **6**, 1547 (1965).

[24] The RME's of S are defined by an equation exactly like that for P in Footnote 15, except that the subscript γ is no longer necessary, and in place of the CG coefficients for the UIR (1, 1), we substitute those for the UIR (3, 0).

[25] The reality of the $SU(3)$-diagonal RME's of K is required by the Hermiticity properties of K.

[26] S. K. Bose, Phys. Rev. **150**, 1231 (1966).

[27] The connection between these two sets of CG coefficients is given in L. K. Pandit and N. Mukunda, Ref. 23.

[28] For $(\lambda'\mu') = (\lambda\mu)$ and $I_0' = I_0$, the expression for the RME's of K becomes real for both values of γ, as required by the Hermiticity properties of K.

[29] J. G. Kuriyan, D. Lurié, and A. J. Macfarlane, quoted in Footnote 15, and L. K. Pandit and N. Mukunda, Ref. 23. The phase convention used in these papers is that of L. C. Biedenharn, Phys. Letters **3**, 69 (1962).

[30] An analogous approach was followed by V. Bargmann and M. Moshinsky, Nucl. Phys. **23**, 177 (1961), in their work on the representations of $SU(3)$ in an $R(3)$ basis.

[31] Using our results for the RME's of K, we can see easily that the operator expression (5.1) has nonzero matrix elements connecting different values of the multiplicity label I_0.

[32] Such an approach was used by A. Kihlberg, V. F. Müller, and F. Halbwachs [Commun. Math. Phys. **3**, 194 (1966)] in their work on the representations of the group $SU(2, 2)$.

[33] L. C. Biedenharn, Phys. Letters **3**, 254 (1963). A nice discussion of these questions is given by A. J. Macfarlane, L. O'Raifeartaigh, and P. S. Rao, J. Math. Phys. **8**, 536 (1967).

[34] J. G. Kuriyan, N. Mukunda, and E. C. G. Sudarshan, Commun. Math. Phys. **8**, 204 (1968); J. Math. Phys. **9**, 2100 (1968).

[35] A. Böhm and E. C. G. Sudarshan, Syracuse University, Syracuse, N.Y., Preprint NYO-3399-188. References to earlier work of these authors are given here.

[36] We use here the results of L. K. Pandit and N. Mukunda, Ref. 23, Sec. 2.

[37] The existence of relations of this type follows from general arguments; see J. J. deSwart, Rev. Mod. Phys. **35**, 916 (1963), and A. J. Macfarlane, N. Mukunda, and E. C. G. Sudarshan, J. Math. Phys. **5**, 576 (1964). The first of Eqs. (II.16) in L. K. Pandit and N. Mukunda, Ref. 23, is incorrect, as an extra minus sign is needed on the right-hand side. The correct form is given here in Eq. (A26).

[38] The existence of a relation of this general nature again follows on general grounds. The exact form is determined by inspection of the tables compiled by J. G. Kuriyan, D. Lurié, and A. J. Macfarlane, quoted in Footnote 15.

Quantum Corrections to the Second Virial Coefficient for a Square-Well Potential

ALBA THEUMANN
Institutt for Teoretisk Fysikk, Norges Tekniske Høgskole, Trondheim, Norway

(Received 17 June 1969)

The first two quantum corrections to the classical second virial coefficient are evaluated in the case that the intermolecular potential is

$$\begin{aligned} V(\mathbf{r}) &= \infty, && 0 < r \leq a, \\ &= -U < 0, && a < r \leq b, \\ &= 0, && r > b. \end{aligned}$$

The first terms of the expansion of $B(T)$ in powers of $\lambda = (h^2/2\pi mkT)^{\frac{1}{2}}$ are

$$\begin{aligned} B(T) = {} & \tfrac{2}{3}\pi Na^3 e^{U\beta}[1 + (\sqrt{\tfrac{9}{8}})\lambda/a + (\pi)^{-1}(\lambda/a)^2] + \tfrac{2}{3}\pi Nb^3(1 - e^{U\beta} + (\sqrt{\tfrac{9}{8}})(\lambda/b)[1 + e^{U\beta} - 2e^{\frac{1}{2}U\beta}I_0(\tfrac{1}{2}U\beta)] \\ & + (\pi)^{-1}(\lambda/b)^2\{\tfrac{1}{2}(e^{U\beta} - 1) - \tfrac{3}{8}(\pi U\beta)^{\frac{1}{2}}e^{\frac{1}{2}U\beta}[3I_{\frac{1}{2}}(\tfrac{1}{2}U\beta) + I_{\frac{3}{2}}(\tfrac{1}{2}U\beta)] \\ & + [1 + \tfrac{1}{2}(U\beta)^{-1}]3e^{U\beta}(U\beta)^{-\frac{1}{2}}\,\mathrm{Erf}\,[(U\beta)^{\frac{1}{2}}] - [1 - \tfrac{1}{2}(U\beta)^{-1}]3(U\beta)^{-\frac{1}{2}}\,\mathrm{Erfi}\,[(U\beta)^{\frac{1}{2}}] - \tfrac{3}{2}(U\beta)^{-1} \\ & \times (1 + e^{U\beta})\}) + O(\lambda^3). \end{aligned}$$

The contribution of the bound states to the first quantum correction is evaluated and the behavior of $B_1^{(b)}/B_1$ vs T is plotted.

I. INTRODUCTION

The second virial coefficient for a monatomic gas can be expressed as a sum of a direct and exchange contribution. For nonanalytic potentials, the direct part can be expanded in powers of the thermal wavelength $\lambda = (h^2/2\pi mkT)^{\frac{1}{2}}$, where h is the Planck's constant, k the Boltzmann's constant, m the mass of the particle, and T the absolute temperature. The exchange contribution was shown to decrease faster than any power of λ in the presence of a repulsive hard core.[1,2] In the following, we consider only the direct part, which we denote by $B(T)$.

The first terms in the expansion of $B(T)$ for a gas of hard spheres were studied by several authors.[3–7] Our purpose here is to calculate the first two quantum corrections to the second virial coefficient when the interacting potential is a hard core plus an attractive square well.

We follow the method introduced by Hill,[7] which consists of relating the thermal Green's function at high temperatures to the Green's function of the Schrödinger equation at large negative energies by means of a Laplace transform. We refer the reader to Ref. 7, in the following indicated by H, for details in the mathematical justification of the method.

$B(T)$ is given by

$$B(T) = \tfrac{1}{2}N\int d^3\mathbf{r}[1 - 2^{\frac{3}{2}}\lambda^3 G(\mathbf{r}, \mathbf{r}, \beta)], \qquad (1)$$

where N is Avogadro's number, $\beta = 1/kT$, and λ the thermal wavelength. $G(\mathbf{r}, \mathbf{r}', \beta)$ is the thermal Green's function and its Laplace transform $G(\mathbf{r}, \mathbf{r}', s)$ satisfies the equation

$$[s - (\hbar^2/m)\nabla^2 + V(\mathbf{r})]G(\mathbf{r}, \mathbf{r}', s) = \delta(\mathbf{r} - \mathbf{r}'), \qquad (2)$$

where

$$G(\mathbf{r}, \mathbf{r}', s) = \int_0^\infty e^{-\beta s}G(\mathbf{r}, \mathbf{r}', \beta)\, d\beta \qquad (3)$$

and $V(\mathbf{r})$ is the intermolecular potential; in our case

$$\begin{aligned} V(\mathbf{r}) &= +\infty, && 0 < r \leq a, \\ &= -U < 0, && a < r \leq b, \\ &= 0, && r > b. \end{aligned} \qquad (4)$$

We find in Sec. II the solution to Eq. (2), and in Sec. III we evaluate explicitly the first two quantum corrections to the classical $B(T)$. The first term was already calculated by Nilsen,[8] considering that to lowest order in λ the boundary conditions on a sphere can be replaced by boundary conditions on a plane.

In Sec. IV the contribution of the bound states of the attractive potential to the first quantum correction is analyzed.

If we write

$$B(T) = B_c + B_1\lambda + B_2\lambda^2 + O(\lambda^3), \qquad (5)$$

and indicate by $B_1^{(b)}$ the bound state's contribution to B_1, we find that the ratio $B_1^{(b)}/B_1$ is always larger than unity at low temperatures and vanishes slowly as $T \to \infty$. In Fig. 1, we plot $B_1^{(b)}/B_1$ vs kT in the special case $b = 2a$. It shows a maximum at $kT \simeq \frac{1}{2}U$ and it becomes less than unity when $kT \gtrsim 3.47\ U$, but it has still an appreciable value at finite temperatures when Eq. (5) is a valid expansion.

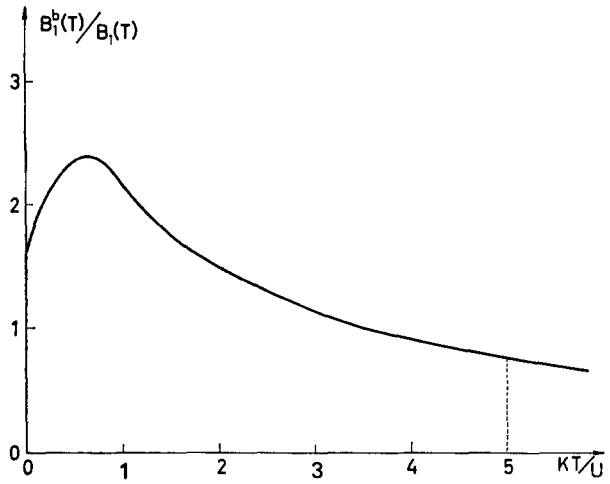


FIG. 1. Relative contribution from bound states to the first quantum correction to the second virial coefficient. The range b of the potential is here chosen to be twice the hard core diameter a.

II. GREEN'S FUNCTION

We introduce the following notation:

$$\gamma^2 = sm/\hbar^2,$$
$$\alpha^2 = (s - U)m/\hbar^2, \tag{6}$$
$$g(\mathbf{r}, \mathbf{r}', s) = 4\pi(\hbar^2/m)G(\mathbf{r}, \mathbf{r}', s). \tag{7}$$

It follows from Eqs. (2) and (4) that $g(\mathbf{r}, \mathbf{r}', s)$ can be written as

$$g(\mathbf{r}, \mathbf{r}', s) = g_<(\mathbf{r}, \mathbf{r}', s), \quad r < b,$$
$$= g_>(\mathbf{r}, \mathbf{r}', s), \quad r > b, \tag{8}$$

where $g_<$ and $g_>$ are the solutions to the pair of differential equations

$$(\nabla^2 - \alpha^2)g_<(\mathbf{r}, \mathbf{r}', s) = -4\pi\delta(\mathbf{r} - \mathbf{r}'), \quad r < b, \tag{9}$$
$$(\nabla^2 - \gamma^2)g_>(\mathbf{r}, \mathbf{r}', s) = -4\pi\delta(\mathbf{r} - \mathbf{r}'), \quad r > b, \tag{10}$$

with the boundary conditions

$$g_<(\mathbf{r}, \mathbf{r}', s) = 0, \quad 0 < r \leq a, \tag{11}$$
$$[g_<(\mathbf{r}, \mathbf{r}', s)]_{r=b} = [g_>(\mathbf{r}, \mathbf{r}', s)]_{r=b}, \tag{12}$$
$$\left[\frac{\partial}{\partial r} g_<(\mathbf{r}, \mathbf{r}', s)\right]_{r=b} = \left[\frac{\partial}{\partial r} g_>(\mathbf{r}, \mathbf{r}', s)\right]_{r=b}. \tag{13}$$

In the case of a spherically symmetric potential, each one of the functions $g_<(\mathbf{r}, \mathbf{r}', s)$ and $g_>(\mathbf{r}, \mathbf{r}', s)$ can be expressed as a sum over partial waves:

$$g_<(\mathbf{r}, \mathbf{r}', s) = \sum_{\nu=\frac{1}{2}}^{\infty} 2\nu P_{\nu-\frac{1}{2}}(\cos\Theta) g_<^{(\nu)}(r, r', s),$$
$$g_>(\mathbf{r}, \mathbf{r}', s) = \sum_{\nu=\frac{1}{2}}^{\infty} 2\nu P_{\nu-\frac{1}{2}}(\cos\Theta) g_>^{(\nu)}(r, r', s), \tag{14}$$

where Θ is the angle between $\mathbf{r}$ and $\mathbf{r}'$, $\nu = l + \frac{1}{2}$, and $g_<^{(\nu)}$, $g_>^{(\nu)}$ satisfy the equations

$$\left[\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d}{dr}\right) - \frac{\nu^2 - \frac{1}{4}}{r^2} - \alpha^2\right] g_<^{(\nu)}(r, r', s) = -\frac{\delta(r - r')}{r^2}, \quad r < b, \tag{15}$$

$$\left[\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d}{dr}\right) - \frac{\nu^2 - \frac{1}{4}}{r^2} - \gamma^2\right] g_>^{(\nu)}(r, r', s) = -\frac{\delta(r - r')}{r^2}, \quad r > b, \tag{16}$$

with the boundary conditions given in Eqs. (11)–(13). The linearly independent solutions to the homogeneous equations (15) and (16) are $r^{-\frac{1}{2}}I_\nu(\alpha r)$ and $r^{-\frac{1}{2}}K_\nu(\alpha r)$ when $r < b$, and $r^{-\frac{1}{2}}I_\nu(\gamma r)$ and $r^{-\frac{1}{2}}K_\nu(\gamma r)$ when $r > b$. $I_\nu(z)$ and $K_\nu(z)$ are the modified Bessel functions,[9] regular at zero and infinity, respectively.

To find a solution to Eqs. (9) and (10) we have to consider separately the cases $a < r' \leq b$ and $r' > b$.

A. $a < r' \leq b$

$$g_<(\mathbf{r}, \mathbf{r}', s) = g_0(\mathbf{r}, \mathbf{r}', \alpha) - (rr')^{-\frac{1}{2}} \sum_{\nu=\frac{1}{2}}^{\infty} 2\nu P_{\nu-\frac{1}{2}}(\cos\Theta) \frac{I_\nu(\alpha a)}{K_\nu(\alpha a)} K_\nu(\alpha r')K_\nu(\alpha r)$$
$$- (rr')^{-\frac{1}{2}} \sum_{\nu=\frac{1}{2}}^{\infty} 2\nu P_{\nu-\frac{1}{2}}(\cos\Theta) A_\nu \Delta_\nu(a, r')\Delta_\nu(a, r), \quad a < r \leq b, \tag{17}$$

$$g_>(\mathbf{r}, \mathbf{r}', s) = -(rr')^{-\frac{1}{2}} \sum_{\nu=\frac{1}{2}}^{\infty} 2\nu P_{\nu-\frac{1}{2}}(\cos\Theta) B_\nu K_\nu(\gamma r), \quad r > b, \tag{18}$$

where

$$\Delta_\nu(a, r) = K_\nu(\alpha a)I_\nu(\alpha r) - I_\nu(\alpha a)K_\nu(\alpha r) \tag{19}$$

and

$$g_0(\mathbf{r}, \mathbf{r}', \alpha) = (|\mathbf{r} - \mathbf{r}'|)^{-1}e^{-\alpha|\mathbf{r}-\mathbf{r}'|} \tag{20}$$

is the free space solution to Eq. (9). It is shown in H that $g_0(\mathbf{r}, \mathbf{r}', \alpha)$ can be expanded in partial waves:

$$g_0(\mathbf{r}, \mathbf{r}', \alpha) = (rr')^{-\frac{1}{2}} \sum_{\nu=\frac{1}{2}}^{\infty} 2\nu P_{\nu-\frac{1}{2}}(\cos\Theta) I_\nu(\alpha r')K_\nu(\alpha r), \quad r > r',$$
$$= (rr')^{-\frac{1}{2}} \sum_{\nu=\frac{1}{2}}^{\infty} 2\nu P_{\nu-\frac{1}{2}}(\cos\Theta) K_\nu(\alpha r')I_\nu(\alpha r), \quad r < r'. \tag{21}$$

The first sum in Eq. (17) is a solution of the homogeneous equation, and it can be seen from Eq. (21) [bottom] that it cancels the value of g_0 at $r = a$. The second sum in Eq. (18) is an arbitrary solution of the

homogeneous equation that vanishes at $r = a$, while $g_>(\mathbf{r}, \mathbf{r}', s)$ is a solution of the homogeneous Eq. (10) with the appropriate behavior at infinity.

The constants A_ν and B_ν are determined by fitting the boundary conditions at $r = b$, Eqs. (12) and (13), and using Eq. (21) [top] for $g_0(\mathbf{r}, \mathbf{r}', s)$.

B. $r' > b$

In this case the source is outside the well. Therefore we need the solution of the homogeneous equation (9), when $a < r \leq b$,

$$g_<(\mathbf{r}, \mathbf{r}', s) = -(rr')^{-\frac{1}{2}} \sum_{\nu=\frac{1}{2}}^{\infty} 2\nu P_{\nu-\frac{1}{2}}(\cos\Theta) C_\nu \Delta_\nu(a, r), \quad a < r \leq b, \quad (22)$$

$$g_>(\mathbf{r}, \mathbf{r}', s) = g_0(\mathbf{r}, \mathbf{r}', \gamma) - (rr')^{-\frac{1}{2}} \sum_{\nu=\frac{1}{2}}^{\infty} 2\nu P_{\nu-\frac{1}{2}}(\cos\Theta) D_\nu K_\nu(\gamma r') K_\nu(\gamma r), \quad r > b. \quad (23)$$

Using the expression (21), for $r < r'$, for $g_0(\mathbf{r}, \mathbf{r}', \gamma)$ and the boundary conditions (12) and (13) we determine the coefficients C_ν and D_ν.

To evaluate $B(T)$ from Eq. (1), we need $g(\mathbf{r}, \mathbf{r}, s)$; hence, we need only the value of $g(\mathbf{r}, \mathbf{r}', s)$ when $\mathbf{r}$ and $\mathbf{r}'$ are both inside or both outside the well.

We obtain

$$g(\mathbf{r}, \mathbf{r}', s) = g_0(\mathbf{r}, \mathbf{r}', \alpha) + g_{\text{h.c.}}(\mathbf{r}, \mathbf{r}', \alpha) + g_{w<}(\mathbf{r}, \mathbf{r}', s), \quad a \leq r, r' \leq b,$$

$$g(\mathbf{r}, \mathbf{r}', s) = g_0(\mathbf{r}, \mathbf{r}', \gamma) + g_{w>}(\mathbf{r}, \mathbf{r}', s), \quad r, r' > b,$$

$$g(\mathbf{r}, \mathbf{r}', s) = 0, \quad 0 < r \leq a, \quad (24)$$

where

$$g_{\text{h.c.}}(\mathbf{r}, \mathbf{r}', \alpha) = -(rr')^{-\frac{1}{2}} \sum_{\nu=\frac{1}{2}}^{\infty} 2\nu P_{\nu-\frac{1}{2}}(\cos\Theta) \frac{I_\nu(\alpha a)}{K_\nu(\alpha a)} K_\nu(\alpha r') K_\nu(\alpha r), \quad (25)$$

$$g_{w<}(\mathbf{r}, \mathbf{r}', s) = -(rr')^{-\frac{1}{2}} \sum_{\nu=\frac{1}{2}}^{\infty} 2\nu P_{\nu-\frac{1}{2}}(\cos\Theta) A_\nu \Delta_\nu(a, r) \Delta_\nu(a, r'), \quad (26)$$

$$g_{w>}(\mathbf{r}, \mathbf{r}', s) = -(rr')^{-\frac{1}{2}} \sum_{\nu=\frac{1}{2}}^{\infty} 2\nu P_{\nu-\frac{1}{2}}(\cos\Theta) D_\nu K_\nu(\gamma r') K_\nu(\gamma r), \quad (27)$$

$$A_\nu = \frac{\alpha K_\nu'(\alpha b) K_\nu(\gamma b) - \gamma K_\nu'(\gamma b) K_\nu(\alpha b)}{\alpha \Delta_\nu'(a, b) K_\nu(\gamma b) - \gamma \Delta_\nu(a, b) K_\nu'(\gamma b)} \frac{1}{K_\nu(\alpha a)}, \quad (28)$$

$$D_\nu = \frac{\alpha \Delta_\nu'(a, b) I_\nu(\gamma b) - \gamma I_\nu'(\gamma b) \Delta_\nu(a, b)}{\alpha \Delta_\nu'(a, b) K_\nu(\gamma b) - \gamma K_\nu'(\gamma b) \Delta_\nu(a, b)}, \quad (29)$$

$$\Delta_\nu'(a, b) = K_\nu(\alpha a) I_\nu'(\alpha b) - I_\nu(\alpha a) K_\nu'(\alpha b). \quad (30)$$

The series for $g_{w<}$ and $g_{w>}$ are convergent for every value of r, but $g_{\text{h.c.}}$ has a pole[7] at $r = r' = a$, while g_0 diverges at $r = r'$.

III. QUANTUM CORRECTIONS

We add and subtract a term to $g_{\text{h.c.}}$ that has the same residue at the pole and gives the first quantum correction for a gas of hard spheres[7]:

$$g(\mathbf{r}, \mathbf{r}', s) = g_0(\mathbf{r}, \mathbf{r}', \alpha) + g_1(\mathbf{r}, \alpha) + g_2(\mathbf{r}, \mathbf{r}', \alpha) + g_{w<}(\mathbf{r}, \mathbf{r}', s), \quad a < r, r' < b, \quad (31)$$

where

$$g_1(r, \alpha) = -a^2 e^{-2\alpha(r-a)}/2r^2(r - a), \quad (32)$$

$$g_2(\mathbf{r}, \mathbf{r}', \alpha) = g_{\text{h.c.}}(\mathbf{r}, \mathbf{r}', \alpha) - g_1(r, \alpha). \quad (33)$$

We can now find the inverse Laplace transform of $g_0(\mathbf{r}, \mathbf{r}', \alpha)$, $g_1(\mathbf{r}, \alpha)$, and $g_0(\mathbf{r}, \mathbf{r}', \gamma)$, with α and γ defined by Eq. (5), and integrate afterwards over $r = r'$, remembering that $g(\mathbf{r}, \mathbf{r}', s) = 0$ when $r \leq a$.

The contribution of these terms to $B(T)$ are

$$B^0(T) + B^{(1)}(T) = \tfrac{2}{3}\pi N a^3 e^{U\beta}[1 + (\sqrt{\tfrac{9}{8}})\lambda/a] + \tfrac{2}{3}\pi N b^3(1 - e^{U\beta}) + O(\exp[-(b - a)^2/\lambda^2]). \quad (34)$$

Now we can put $r' = r$ in $g_2(\mathbf{r}, \mathbf{r}', \alpha)$ and integrate over r. We define

$$g_3(\mathbf{r}, \mathbf{r}, s) = g_2(\mathbf{r}, \mathbf{r}, \alpha) + g_{w<}(\mathbf{r}, \mathbf{r}, s), \quad a \leq r \leq b, \qquad = g_{w>}(\mathbf{r}, \mathbf{r}, s), \quad r > b, \quad (35)$$

and obtain from Eqs. (25)–(27) and (32), using Eq. (37) of H,

$$\int_a^\infty g_3(\mathbf{r}, \mathbf{r}, s)\, d^3\mathbf{r} = \lim_{L\to\infty}\left[2\pi a^2\left(\log\frac{L}{\alpha a} + O(e^{-2\alpha(b-a)})\right) - 4\pi \sum_{\nu=\frac{1}{2}}^{L+\frac{1}{2}} 2\nu \frac{I_\nu(\alpha a)}{K_\nu(\alpha a)} \int_a^b K_\nu^2(\alpha r) r\, dr\right] - 4\pi \sum_{\nu=\frac{1}{2}}^{\infty} 2\nu A_\nu \int_a^b \Delta_\nu^2(a, r) r\, dr - 4\pi \sum_{\nu=\frac{1}{2}}^{\infty} 2\nu D_\nu \int_b^\infty K_\nu^2(\gamma r) r\, dr. \quad (36)$$

The indefinite integrals of the modified Bessel functions are[10]

$$\int W_\nu^2(\alpha r) r\, dr = \tfrac{1}{2} r^2[(1 + \nu^2/\alpha^2 r^2) W_\nu^2(\alpha r) - W_\nu'^2(\alpha r)], \quad (37)$$

where $W_\nu(\alpha r)$ is either $I_\nu(\alpha r)$ or $K_\nu(\alpha r)$ and

$$\int I_\nu(\alpha r)K_\nu(\alpha r)r\,dr = \frac{r^2}{2}\left\{I_\nu(\alpha r)K_\nu(\alpha r)\left(1+\frac{\nu^2}{\alpha^2 r^2}\right) - K'_\nu(\alpha r)I'_\nu(\alpha r)\right\}. \quad (38)$$

Taking into account that $\Delta_\nu(a, a) = 0$ and $\Delta'_\nu(a, a)$ is the Wronskian

$$\Delta'_\nu(a, a) = K_\nu(\alpha a)I'_\nu(\alpha a) - I_\nu(\alpha a)K'_\nu(\alpha a) = -(\alpha a)^{-1},$$

we obtain, for Eq. (36),

$$\begin{aligned}\int_a^\infty g_3(\mathbf{r}, \mathbf{r}, s)\,d^3\mathbf{r} &= \lim_{L\to\infty}\Bigg[2\pi a^2\left(\log\frac{L}{\alpha a} + O(e^{-2\alpha(b-a)})\right) \\ &+ 2\pi a^2\sum_{\nu=\frac{1}{2}}^{L+\frac{1}{2}} 2\nu\,\frac{I_\nu(\alpha a)}{K_\nu(\alpha a)} \\ &\times\left(\left(1+\frac{\nu^2}{\alpha^2 a^2}\right)K_\nu^2(\alpha a) - K_\nu'^2(\alpha a)\right) \\ &- 2\pi b^2\sum_{\nu=\frac{1}{2}}^{L+\frac{1}{2}} 2\nu\,\frac{I_\nu(\alpha a)}{K_\nu(\alpha a)} \\ &\times\left(\left(1+\frac{\nu^2}{\alpha^2 b^2}\right)K_\nu^2(\alpha b) - K_\nu'^2(\alpha b)\right)\Bigg] \\ &- 2\pi b^2\sum_{\nu=\frac{1}{2}}^{\infty} 2\nu A_\nu\left(\left(1+\frac{\nu^2}{\alpha^2 b^2}\right)\Delta_\nu^2(a, b) - \Delta_\nu'^2(a, b)\right) \\ &- 2\pi a^2\sum_{\nu=\frac{1}{2}}^{\infty} 2\nu A_\nu\,\frac{1}{\alpha^2 a^2} \\ &+ 2\pi b^2\sum_{\nu=\frac{1}{2}}^{\infty} 2\nu D_\nu\left(\left(1+\frac{\nu^2}{\alpha^2 b^2}\right)K_\nu^2(\gamma b) - K_\nu'^2(\gamma b)\right).\end{aligned} \quad (39)$$

For large values of the Laplace variable s, that is, for large values of α or γ, Debye's series[9] provide us with an asymptotic expansion of $I_\nu(x)$, $K_\nu(x)$, $I'_\nu(x)$, and $K'_\nu(x)$ uniformly valid in ν for large values of x.

The Debye expansions read

$$\begin{aligned} I_\nu(x) &= (2\pi)^{-\frac{1}{2}}(\nu^2+x^2)^{-\frac{1}{4}}e^{\mu(x)}\sum_{k=0}^{\infty}\nu^{-k}u_k(t_x), \\ K_\nu(x) &= \left(\frac{\pi}{2}\right)^{\frac{1}{2}}(\nu^2+x^2)^{-\frac{1}{4}}e^{-\mu(x)}\sum_{k=0}^{\infty}(-\nu)^{-k}u_k(t_x), \\ I'_\nu(x) &= (2\pi)^{-\frac{1}{2}}x^{-1}(\nu^2+x^2)^{\frac{1}{4}}e^{\mu(x)}\sum_{k=0}^{\infty}\nu^{-k}v_k(t_x), \\ K'_\nu(x) &= -\left(\frac{\pi}{2}\right)^{\frac{1}{2}}x^{-1}(\nu^2+x^2)^{\frac{1}{4}}e^{-\mu(x)}\sum_{k=0}^{\infty}(-\nu)^{-k}v_k(t_x),\end{aligned} \quad (40)$$

where

$$\mu(x) = (\nu^2+x^2)^{\frac{1}{2}} - \nu\sinh^{-1}(\nu/x), \qquad t_x = \nu(\nu^2+x^2)^{-\frac{1}{2}}. \quad (41)$$

The first five polynomials $u_k(t_x)$ and $v_k(t_x)$ and their recurrence relations are explicitly given in Ref. 9.

We can see from (40) that

$$\begin{aligned} K_\nu(\alpha a)I_\nu(\alpha b) &\simeq O(\exp[\mu(\alpha b) - \mu(\alpha a)]), \\ K_\nu(\alpha b)I_\nu(\alpha a) &\sim O(\exp[-\mu(\alpha b) + \mu(\alpha a)]).\end{aligned} \quad (42)$$

Hence, we can approximate Eqs.(19) and (30) by

$$\begin{aligned}\Delta_\nu(a, b) &\simeq K_\nu(\alpha a)I_\nu(\alpha b) \\ &\quad\times\{1 + O(\exp[-2\mu(\alpha b) + 2\mu(\alpha a)])\}, \\ \Delta'_\nu(a, b) &\simeq K_\nu(\alpha a)I'_\nu(\alpha b) \\ &\quad\times\{1 + O(\exp[-2\mu(\alpha b) + 2\mu(\alpha a)])\}.\end{aligned} \quad (43)$$

All the terms $O(\exp[-2\alpha(b-a)])$ give contributions to $B(T) \simeq O(\exp[-(b-a)^2/\lambda^2])$, and do not contribute to any finite power of λ.

Introducing (43) in Eqs. (28), (29), and (39) we obtain

$$\begin{aligned}\int_a^\infty g_3(\mathbf{r}, \mathbf{r}, s)4\pi r^2\,dr &= \lim_{L\to\infty}\Bigg[2\pi a^2\log\frac{L}{\alpha a} + 2\pi a^2\sum_{\nu=\frac{1}{2}}^{L+\frac{1}{2}} 2\nu\,\frac{I_\nu(\alpha a)}{K_\nu(\alpha a)} \\ &\quad\times\left(\left(1+\frac{\nu^2}{\alpha^2 a^2}\right)K_\nu^2(\alpha a) - K_\nu'^2(\alpha a)\right)\Bigg] \\ &\quad - 2\pi b^2\sum_{\nu=\frac{1}{2}}^{\infty} 2\nu F_\nu(x, y) + O(\exp[-2\alpha(b-a)]),\end{aligned} \quad (44)$$

where

$$\begin{aligned}F_\nu(x, y) &= \{[yK'_\nu(y)K_\nu(x) - xK'_\nu(x)K_\nu(y)] \\ &\quad\times[(1+\nu^2/y^2)I_\nu^2(y) - I_\nu'^2(y)] \\ &\quad - [yI'_\nu(y)I_\nu(x) - xI'_\nu(x)I_\nu(y)] \\ &\quad\times[(1+\nu^2/x^2)K_\nu^2(x) - K_\nu'^2(x)]\} \\ &\quad\times[yI'_\nu(y)K_\nu(x) - xK'_\nu(x)I_\nu(y)]^{-1},\end{aligned} \quad (45)$$

$$x = \gamma b, \quad y = \alpha b. \quad (46)$$

Aside from exponentially small terms, what we obtain in (44) are the separate contributions of a hard core at $r = a$, given by the first two terms in parenthesis, and a new contribution corresponding to a square well for $0 < r \leq b$.

The limit in (44) was already evaluated in H up to $O((\lambda/a)^6)$ and taking into account that $\alpha^2 = (s-U)m/\hbar^2$ the second-order contribution from this term to $B(T)$ is

$$B^{(2)}(T) = \tfrac{2}{3}N\lambda^2 a e^{U\beta}. \quad (47)$$

From Eq. (40) we find

$$(1 + \nu^2/x^2)I_\nu^2(x) - I_\nu'^2(x) = (2\pi)^{-1}x^{-2}(\nu^2 + x^2)^{\frac{1}{2}}e^{2\mu(x)}\sum_{k=0}^{\infty}a_k\nu^{-k}, \quad (48)$$

$$(1 + \nu^2/x^2)K_\nu^2(x) - K_\nu'^2(x) = \tfrac{1}{2}\pi x^{-2}(\nu^2 + x^2)^{\frac{1}{2}}e^{-2\mu(x)}\sum_{k=0}^{\infty}a_k(-\nu)^{-k}, \quad (49)$$

$$yI_\nu'(y)I_\nu(x) - xI_\nu'(x)I_\nu(y) = (2\pi)^{-1}e^{[\mu(x)+\mu(y)]}(\nu^2 + x^2)^{-\frac{1}{4}}(\nu^2 + y^2)^{-\frac{1}{4}}\sum_{k=0}^{\infty}\nu^{-k}b_k, \quad (50)$$

$$yK_\nu'(y)K_\nu(x) - xK_\nu'(x)K_\nu(y) = -\tfrac{1}{2}\pi e^{-[\mu(x)+\mu(y)]}(\nu^2 + x^2)^{-\frac{1}{4}}(\nu^2 + y^2)^{-\frac{1}{4}}\sum_{k=0}^{\infty}(-\nu)^{-k}b_k, \quad (51)$$

$$yI_\nu'(y)K_\nu(x) - xK_\nu'(x)I_\nu(y) = \tfrac{1}{2}e^{[\mu(y)-\mu(x)]}(\nu^2 + x^2)^{-\frac{1}{4}}(\nu^2 + y^2)^{-\frac{1}{4}}\sum_{k=0}^{\infty}c_k\nu^{-k}, \quad (52)$$

where

$$a_k(x) = \sum_{l=0}^{k}[u_l(t_x)u_{k-l}(t_x) - v_l(t_x)v_{k-l}(t_x)], \quad (53)$$

$$b_k(x, y) = (\nu^2 + y^2)^{\frac{1}{2}}\sum_{l=0}^{k}v_l(t_y)u_{k-l}(t_x) - (\nu^2 + x^2)^{\frac{1}{2}}\sum_{l=0}^{k}v_l(t_x)u_{k-l}(t_y), \quad (54)$$

$$c_k(x, y) = (\nu^2 + y^2)^{\frac{1}{2}}\sum_{l=0}^{k}(-)^{k-l}v_l(t_y)u_{k-l}(t_x) + (\nu^2 + x^2)^{\frac{1}{2}}\sum_{l=0}^{k}(-)^{l}v_l(t_x)u_{k-l}(t_y). \quad (55)$$

Introducing Eqs. (48)–(55) in (45), we obtain an expansion for $F_\nu(x, y)$:

$$F_\nu(x, y) = \tfrac{1}{2}\sum_{n=0}^{\infty}\xi_n\nu^{-n}, \quad (56)$$

where the ξ_n can be found from the recurrence relation

$$\xi_n = \frac{f_n}{c_0} - \frac{1}{c_0}\sum_{l=0}^{n-1}c_{n-l}\xi_l, \quad (57)$$

and the f_n are given by

$$f_n = y^{-2}(\nu^2 + x^2)^{\frac{1}{2}}\sum_{l=0}^{n}(-)^l b_l(x, y)a_{n-l}(y) + (-1)^n x^{-2}(\nu^2 + x^2)^{\frac{1}{2}}\sum_{l=0}^{n}(-1)^l b_l(x, y)a_{n-l}(x). \quad (58)$$

The first ξ_n are

$$\xi_0 = 0,$$

$$\xi_1 = \frac{(\nu^2 + x^2)^{\frac{1}{2}} - (\nu^2 + y^2)^{\frac{1}{2}}}{(\nu^2 + x^2)^{\frac{1}{2}} + (\nu^2 + y^2)^{\frac{1}{2}}}\left(\frac{\nu}{\nu^2 + y^2} - \frac{\nu}{\nu^2 + x^2}\right),$$

$$\xi_2 = \nu^2\frac{[(\nu^2 + x^2)^{\frac{1}{2}} - (\nu^2 + y^2)^{\frac{1}{2}}]^2}{(x^2 - y^2)} \times \left[\frac{5}{12}\frac{1}{x^2 - y^2}\left(\frac{\nu^4}{(\nu^2 + y^2)^{\frac{5}{2}}} - \frac{\nu^4}{(\nu^2 + x^2)^{\frac{5}{2}}}\right) + \frac{\nu^2}{(\nu^2 + y^2)^{\frac{5}{2}}}\left(-\frac{7}{6} + \frac{5}{12}\frac{x^2}{x^2 - y^2}\right) + \frac{\nu^2}{(\nu^2 + x^2)^{\frac{5}{2}}}\left(-\frac{7}{6} - \frac{5}{12}\frac{y^2}{x^2 - y^2}\right) + \frac{1}{4}\left(\frac{y^2}{(\nu^2 + y^2)^{\frac{5}{2}}} + \frac{x^2}{(\nu^2 + x^2)^{\frac{5}{2}}}\right) + \frac{1}{3}\frac{1}{x^2 - y^2}\left(\frac{\nu^2}{(\nu^2 + y^2)^{\frac{3}{2}}} - \frac{\nu^2}{(\nu^2 + x^2)^{\frac{3}{2}}}\right) - \frac{1}{4}\frac{1}{x^2 - y^2}\left(\frac{x^2}{(\nu^2 + y^2)^{\frac{3}{2}}} - \frac{y^2}{(\nu^2 + x^2)^{\frac{3}{2}}}\right) - \frac{1}{x^2 - y^2}\left(\frac{y^2}{(\nu^2 + y^2)(\nu^2 + x^2)^{\frac{1}{2}}} - \frac{x^2}{(\nu^2 + x^2)(\nu^2 + y^2)^{\frac{1}{2}}}\right) - \frac{3}{4}\frac{1}{x^2 - y^2}\left(\frac{1}{(\nu^2 + y^2)^{\frac{1}{2}}} - \frac{1}{(\nu^2 + x^2)^{\frac{1}{2}}}\right)\right]. \quad (59)$$

The sum over ν in Eq. (44) can be expressed as an integral with the help of the Euler–MacLaurin sum formula.[7] The corrections to the integral can be shown to give contributions of $O(\lambda^3)$, and the lower limit can be extended to $\nu = 0$, the corrections being also of higher order.

After integrating over ν, we find

$$\int_a^\infty [g_3(\mathbf{r}, \mathbf{r}, s) - g_2(\mathbf{r}, \mathbf{r}, \alpha)]4\pi r^2\,dr \simeq -2\pi b^2\left(\tfrac{1}{2}\log\frac{s - U}{s} - 2\log\frac{2(s - U)^{\frac{1}{2}}}{(s)^{\frac{1}{2}} + (s - U)^{\frac{1}{2}}}\right) - 2\pi b\left(\frac{\hbar^2}{m}\right)^{\frac{1}{2}}\left[\frac{2}{3}\left(\frac{1}{(s)^{\frac{1}{2}}} - \frac{1}{(s - U)^{\frac{1}{2}}}\right) + \frac{1}{[(s)^{\frac{1}{2}} + (s - U)^{\frac{1}{2}}]^2}\left(\frac{s - U}{(s)^{\frac{1}{2}}} - \frac{s}{(s - U)^{\frac{1}{2}}}\right) + \frac{1}{U}\left(\frac{s - U}{(s)^{\frac{1}{2}}} + \frac{s}{(s - U)^{\frac{1}{2}}}\right) + \left(\frac{1}{2}\frac{s - U}{U} - 1\right)\frac{1}{(U)^{\frac{1}{2}}}\log\frac{(s)^{\frac{1}{2}} - (U)^{\frac{1}{2}}}{(s)^{\frac{1}{2}} + (U)^{\frac{1}{2}}} - \left(\frac{1}{2}\frac{s}{U} + 1\right)\frac{i}{(U)^{\frac{1}{2}}}\log\frac{(s - U)^{\frac{1}{2}} - i(U)^{\frac{1}{2}}}{(s - U)^{\frac{1}{2}} + i(U)^{\frac{1}{2}}}\right]. \quad (60)$$

Expressing the logarithmic terms as a power series in $U/[(s)^{\frac{1}{2}} + (s-U)^{\frac{1}{2}}]^2$, $(U)^{\frac{1}{2}}/(s)^{\frac{1}{2}}$, and $(U)^{\frac{1}{2}}/(s-U)^{\frac{1}{2}}$, respectively, we find that the contribution of $g_3 - g_2$ to $B(T)$ is

$$\begin{aligned} B^{(3)}(T) = {} & N\pi b^2[\lambda/(2)^{\frac{1}{2}}][1 + e^{U\beta} - 2e^{\frac{1}{2}U\beta}I_0(\tfrac{1}{2}U\beta)] \\ & + Nb\lambda^2\{\tfrac{1}{3}(e^{U\beta}-1) - \tfrac{1}{4}(\pi U\beta)^{\frac{1}{2}} \\ & \times e^{\frac{1}{2}U\beta}[3I_{\frac{1}{2}}(\tfrac{1}{2}U\beta) + I_{\frac{3}{2}}(\tfrac{1}{2}U\beta)] \\ & + [2 + (U\beta)^{-1}]e^{U\beta}(U\beta)^{-\frac{1}{2}}\,\mathrm{Erf}\,[(U\beta)^{\frac{1}{2}}] \\ & - [2 - (U\beta)^{-1}](U\beta)^{-\frac{1}{2}}\,\mathrm{Erfi}\,[(U\beta)^{\frac{1}{2}}] \\ & - (U\beta)^{-1}(1 + e^{U\beta})\}, \end{aligned} \tag{61}$$

where[10]

$$\begin{aligned} \mathrm{Erf}\,[(x)^{\frac{1}{2}}] &= e^{-x}\sum_{n=0}^{\infty}\frac{2^n[(x)^{\frac{1}{2}}]^{2n+1}}{(2n+1)\cdots 3\cdot 1} \\ &= (\tfrac{1}{2}\pi)^{\frac{1}{2}}\sum_{n=0}^{\infty}(-)^{\frac{1}{2}n}I_{n-\frac{1}{2}}(x), \\ \mathrm{Erfi}\,[(x)^{\frac{1}{2}}] &= e^{x}\sum_{n=0}^{\infty}\frac{(-2)^n[(x)^{\frac{1}{2}}]^{2n+1}}{(2n+1)(2n-1)\cdots 3\cdot 1} \\ &= (\tfrac{1}{2}\pi)^{\frac{1}{2}}\sum_{n=0}^{\infty}(-)^{\frac{1}{2}n}I_{n+\frac{1}{2}}(x). \end{aligned} \tag{62}$$

The first three terms of the expansion of $B(T)$ in powers of λ, from Eqs. (34), (47), and (61) are

$$\begin{aligned} B(T) = {} & \tfrac{2}{3}\pi Na^3\left[1 + \frac{3}{2(2)^{\frac{1}{2}}a}\frac{\lambda}{a} + \frac{1}{\pi}\left(\frac{\lambda}{a}\right)^2\right]e^{U\beta} \\ & + \tfrac{2}{3}\pi Nb^3\Bigg(1 - e^{U\beta} \\ & + \frac{3}{2(2)^{\frac{1}{2}}b}\frac{\lambda}{b}[1 + e^{U\beta} - 2e^{\frac{1}{2}U\beta}I_0(\tfrac{1}{2}U\beta)] \\ & + \frac{1}{\pi}\left(\frac{\lambda}{b}\right)^2\{\tfrac{1}{2}(e^{U\beta}-1) - \tfrac{3}{8}(\pi U\beta)^{\frac{1}{2}} \\ & \times e^{\frac{1}{2}U\beta}[3I_{\frac{1}{2}}(\tfrac{1}{2}U\beta) + I_{\frac{3}{2}}(\tfrac{1}{2}U\beta)] \\ & + [1 + (2U\beta)^{-1}]3e^{U\beta}(U\beta)^{-\frac{1}{2}}\,\mathrm{Erf}\,[(U\beta)^{\frac{1}{2}}] \\ & - [1 - (2U\beta)^{-1}]3(U\beta)^{-\frac{1}{2}}\,\mathrm{Erfi}\,[(U\beta)^{\frac{1}{2}}] \\ & - 3(2U\beta)^{-1}(1 + e^{U\beta})\}\Bigg) + O(\lambda^3). \end{aligned} \tag{63}$$

It can be verified that, when $U \to 0$, the only contribution left is that of a hard core of diameter a, while when $U \to -\infty$ we obtain $B(T)$ for a hard core of diameter b.

IV. CONTRIBUTION OF THE BOUND STATES TO $B_1(T)$

The first quantum correction can be obtained from the analogous 1-dimensional problem considering that, to the lowest order in λ, the spherical surface can be replaced by a plane.[3,5]

In this approximation we can also consider $r = b$ in the volume element of Eq. (1) and write for the linear term in Eq. (5)

$$B_1(T) = 4\pi b^2 B_{(O.D.)1}, \tag{64}$$

where $B_{(O.D.)1}$ is the first quantum correction to the second virial coefficient for the 1-dimensional potential

$$\begin{aligned} V(x) &= \infty, \quad x < a, \\ &= -U, \quad a \le x \le b, \\ &= 0, \quad x > b, \end{aligned} \tag{65}$$

in the half-space $x > 0$.

The 1-dimensional second virial coefficient $B_{O.D.}$ can be expressed[3,7] in the form

$$B_{O.D.} = -\frac{N}{2^{\frac{1}{2}}}\lambda\left(\sum_n e^{-\beta\epsilon_n} + \frac{1}{\pi}\int_0^{\infty}dk\,\frac{d\eta(k)}{dk}\,e^{-k^2\lambda^2/2\pi}\right), \tag{66}$$

where the ϵ_n form the spectrum of discrete negative energies, $-U \le \epsilon_n < 0$, and satisfy the equation

$$(-\epsilon_n)^{\frac{1}{2}}\tan\,[(U+\epsilon_n)(b-a)^2m/\hbar^2]^{\frac{1}{2}} + (U+\epsilon_n)^{\frac{1}{2}} = 0, \tag{67}$$

while $\eta(k)$ is the phase shift of the wave with wave vector k.

In terms of the new variable,

$$\rho^2 = \beta(U+\epsilon), \quad 0 \le \rho^2 \le U\beta, \tag{68}$$

Eq. (67) can be written

$$\sin\,(\rho(2\pi)^{\frac{1}{2}}(b-a)/\lambda + \varphi(\rho)) = 0, \tag{69}$$

where

$$\varphi(\rho) = \tan^{-1}\,(\rho(U\beta - \rho^2)^{-\frac{1}{2}}). \tag{70}$$

It follows from (69) that the ρ_n are given by the successive values of n in the equation

$$\rho_n[(b-a)\sqrt{2\pi}/\lambda] + \varphi(\rho_n) = n\pi, \quad n = 1, 2, 3\cdots. \tag{71}$$

The term with $n = 0$ should be discarded because it gives the trivial solution $\rho_n = 0$. In the limit $\lambda \to 0$, $(b-a)\sqrt{2\pi}/\lambda \to \infty$, more solutions appear and they become closer together, forming a continuum.[3]

Therefore, in this limit we can write from Eqs. (66) and (68)

$$\begin{aligned} B^{(b)}_{O.D.} &= -2^{-\frac{1}{2}}N\lambda e^{U\beta}\left(\sum_{n=0}e^{-\rho_n^2} - 1\right) \\ &\simeq -2^{-\frac{1}{2}}N\lambda e^{U\beta}\left(\int_0^{(U\beta)^{\frac{1}{2}}}e^{-\rho^2}g(\rho)\,d\rho - \tfrac{1}{2}\right), \end{aligned} \tag{72}$$

where $g(\rho)$ is the density of states

$$g(\rho) = \frac{(b-a)\sqrt{2\pi}}{\lambda} + \frac{d\varphi(\rho)}{d\rho}, \tag{73}$$

and we approximate the sum by an integral with the help of the Euler–MacLaurin sum formula. Keeping only the term linear in λ in Eq. (72), we obtain from Eq. (64),

$$B_1^{(b)} = 2\pi b^2 N 2^{-\frac{1}{2}}\lambda[e^{U\beta} - e^{\frac{1}{2}U\beta}I_0(\tfrac{1}{2}U\beta)]. \tag{74}$$

From Eqs. (74) and (81), the ratio $B_1^{(b)}/B_1$ is

$$\frac{B_1^{(b)}}{B_1} = \frac{2b^2[1 - e^{-\frac{1}{2}U\beta}I_0(\frac{1}{2}U\beta)]}{a^2 + b^2[e^{-U\beta} + 1 - 2e^{-\frac{1}{2}U\beta}I_0(\frac{1}{2}U\beta)]}. \tag{75}$$

We show in Fig. (1) the variation of $B_1^{(b)}/B_1$ with temperature for the special value of the parameter $b = 2a$.

Note added in proof: The quantum corrections to $B(T)$ have been evaluated previously by José J. D'Arruda in first order of an expansion in powers of βU. [Bull. Am. Phys. Soc. **13,** 646 (1968); J. J. D'Arruda and R. N. Hill, preprint]. We thank Professor Hill for this communication and for valuable comments.

ACKNOWLEDGMENTS

I am grateful to Professor P. C. Hemmer for introducing me to the problem and for discussions that were essential for its understanding. I acknowledge discussions with Lic. techn. T. S. Nilsen.

[1] S. Larsen, J. Kilpatrick, E. Lieb, and H. Jordan, Phys. Rev. **140,** A129 (1965).
[2] E. Lieb, J. Math. Phys. **8,** 43 (1967).
[3] G. E. Uhlenbeck and E. Beth, Physica **3,** 129 (1936).
[4] F. Mohling, Phys. Fluids **6,** 1097 (1963).
[5] R. A. Handelsman and J. B. Keller, Phys. Rev. **148,** 94 (1966).
[6] P. C. Hemmer and K. J. Mork, Phys. Rev. **158,** 114 (1967).
[7] R. N. Hill, J. Math. Phys. **9,** 1534 (1968), and references therein.
[8] T. S. Nilsen, Phys. Rev. **186,** 262 (1969).
[9] M. Abramovitz and I. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables* (Dover Publications, Inc., New York, 1965), Secs. 9 and 10.
[10] *Higher Transcendental Functions* (McGraw-Hill Book Company, Inc., New York, 1954), Vol. II. Equation (37) is tabulated, while Eq. (38) can be deduced from the equivalent integral for Bessel functions, 7.14.1(11). Erf (x) as it is defined here differs from a factor $2\pi^{-\frac{1}{2}}$ from the definition in Ref. 9.

Expansion of the Inhomogeneous Pseudo-unitary Symplectic Lie Algebras $IUSp(2\nu_1, 2\nu_2)$ to $USp(2\nu_1 + 2, 2\nu_2)$ or to $USp(2\nu_1, 2\nu_2 + 2)$

JOHN G. NAGEL
Institute of Theoretical Physics, Chalmers University of Technology, Göteborg, Sweden

(Received 4 August 1969; Revised Manuscript Received 18 December 1969)

We show that the inhomogeneous pseudo-unitary symplectic Lie algebra $IUSp(2\nu_1, 2\nu_2)$ can be expanded to $USp(2\nu_1 + 2, 2\nu_2)$ or to $USp(2\nu_1, 2\nu_2 + 2)$. We also identify a $(2\nu_1 + 2\nu_2 + 2)$-dimensional *linear* Lie group $\mathfrak{G}(2\nu_1 + 2\nu_2 + 2, R) \cong \mathfrak{IUS}\not{p}(2\nu_1, 2\nu_2)$, which under this expansion is being globally expanded to a $(2\nu_1 + 2\nu_2 + 2)$-dimensional linear group $\mathfrak{E}(2\nu_1 + 2, 2\nu_2; R) \cong \mathfrak{US}\not{p}(2\nu_1 + 2, 2\nu_2)$ or to $\mathfrak{E}(2\nu_1, 2\nu_2 + 2; R) \cong \mathfrak{US}\not{p}(2\nu_1, 2\nu_2 + 2)$.

INTRODUCTION

Let us first define the concept of expansion of a Lie algebra G over the field k, where k is either the real field R or the complex field C. By expansion of G one understands a process, which can be roughly described as one of replacing some of the elements of G by new operators, which are certain functions of the elements of G, and also adding further operators of this form such that the new set of operators close under commutation. The new Lie algebra $E(G)$ which is spanned by these operators is called an expansion of G. More precisely, one can say that $E(G)$ is a Lie algebra, the elements of which are elements of an algebraic extension of the quotient-division algebra[1-4] of the enveloping algebra of G.

The expansion problem was first treated by Melvin[5] who expanded the Poincaré Lie algebra to the de Sitter Lie algebra. Later, several other authors expanded various special cases of the inhomogeneous pseudo-orthogonal Lie algebras[6-16] $ISO(n_1, n_2; R)$, and some authors expanded the inhomogeneous pseudo-unitary Lie algebras[11,12,15,16] $IU(n_1, n_2)$. Sankaranarayanan[7] obtained the expansion of $ISO(n; R)$ to $SO(n, 1; R)$, Rosen and Roman[11] and Rosen[12] of $ISO(n_1, n_2; R)$ to $SO(n_1 + 1, n_2; R)$ or to $SO(n_1, n_2 + 1; R)$ and, furthermore, of $IU(n_1, n_2)$ to $U(n_1 + 1, n_2)$ or to $U(n_1, n_2 + 1)$.

For the expansion problem for the general linear Lie algebras, it was found by Nagel[17] that the semidirect sum $T(2n, k) \overleftarrow{+} GL(n, k)$, with $T(2n, k)$ the ideal, can be expanded to $GL(n + 1, k)$ in a similar way as the expansion of $IU(n_1, n_2)$.

For the symplectic Lie algebras it was shown by Nagel and Shah[18] that $T(2n, k) \overleftarrow{+} Sp(n, k)$ can be expanded to $Sp(n + 2, k)$.

The method of expansion has been applied to various areas of physics. In relativistic quantum mechanics it has been used to obtain a relativistic position operator.[6] In connection with particle physics, it has been utilized to construct dynamical groups from symmetry groups.[5,8,9,10,15] In representation theory for Lie algebras, this method has been applied to obtain representations of the expanded Lie algebra from representations of the original Lie algebra. In particular, it has been used to expand representations of the homogeneous Galilei Lie algebra to representations of the Lorentz Lie algebra and to expand representations of the Poincaré Lie algebra to representations of the de Sitter Lie algebra.[14] It has also been utilized to obtain representations of noncompact Lie algebras from representations of inhomogenizations of compact Lie algebras.[13,16] The method has also been applied in representation theory for Lie groups. Philips and Wigner[19] thus obtained representations of the Lie group $\mathcal{O}(n, 1; R)$ from its Lie algebra $O(n, 1; R)$ considered as an expansion of a Lie algebra isomorphic to $IO(n; R)$.

In this paper we shall solve the last and more complicated of the expansion problems for the classical Lie algebras, namely that of the pseudo-unitary symplectic Lie algebras. We show that $IUSp(2\nu_1, 2\nu_2)$ can be expanded to $USp(2\nu_1 + 2, 2\nu_2)$ or to $USp(2\nu_1, 2\nu_2 + 2)$. We also identify a $(2\nu_1 + 2\nu_2)$-dimensional *linear* Lie group $\mathfrak{G}(2\nu_1 + 2\nu_2 + 2, R) \cong \mathfrak{IUS}\not{p}(2\nu_1, 2\nu_2)$, which under this expansion is being globally expanded to a $(2\nu_1 + 2\nu_2 + 2)$-dimensional linear Lie group $\mathfrak{E}(2\nu_1 + 2, 2\nu_2; R) \cong \mathfrak{US}\not{p}(2\nu_1 + 2, 2\nu_2)$ or to $\mathfrak{E}(2\nu_1, 2\nu_2 + 2; R) \cong \mathfrak{US}\not{p}(2\nu_1, 2\nu_2 + 2)$.

1. THE GROUPS $\mathfrak{IUS}\not{p}(2\nu_1, 2\nu_2)$

Let us start by giving a few facts about the inhomogeneous pseudo-unitary symplectic groups $\mathfrak{IUS}\not{p}(2\nu_1, 2\nu_2)$. The latter is the *real* Lie group of those *complex* transformations

$$x'_{\alpha\rho} = \mathfrak{S}^{\beta\sigma}_{\alpha\rho} x_{\beta\sigma} + a_{\alpha\rho}, \quad a_{\alpha\rho} = b_{\alpha\rho} + ic_{\alpha\rho},$$
$$\alpha, \beta = 1, -1, \quad \rho, \sigma = 1, \cdots, \nu$$

of a 2ν-dimensional complex vector space

$$U_{2\nu} = V'_\nu \otimes X_2, \quad \nu = \nu_1 + \nu_2,$$

where

$$S^{\beta\sigma}_{\alpha\rho} \in \mathfrak{US}\not{p}(2\nu_1, 2\nu_2) \equiv \mathfrak{U}(2\nu_1, 2\nu_2) \cap \mathfrak{S}\not{p}(2\nu, C),$$
$$b_{\alpha\rho}, c_{\alpha\rho} \in R.$$

Here $\mathfrak{US}\not{p}(2\nu_1, 2\nu_2)$, $\mathfrak{U}(2\nu_1, 2\nu_2)$, and $\mathfrak{S}\not{p}(2\nu, C)$ are the 2ν-dimensional pseudo-unitary symplectic, pseudo-unitary and complex symplectic groups,[20] respectively. $\otimes$ and $\cap$ denote tensor product and intersection, respectively. We have that

$$\mathfrak{J}\mathfrak{US}\not{p}(2\nu_1, 2\nu_2) \cong \mathfrak{J}\mathfrak{U}(2\nu_1, 2\nu_2) \cap \mathfrak{JS}\not{p}(2\nu, C)$$
$$\cong \mathfrak{C}(4\nu, R) \overleftarrow{\times} \mathfrak{US}\not{p}(2\nu_1, 2\nu_2),$$

the semidirect product of $\mathfrak{US}\not{p}(2\nu_1, 2\nu_2)$ by $\mathfrak{C}(4\nu, R)$, and leaves invariant the nondegenerate antisymmetric bilinear form

$$(\bar{x} - x, \bar{y} - y) \equiv G^{\alpha\beta} g^{\rho\sigma} (\bar{x}_{\alpha\rho} - x_{\alpha\rho})(\bar{y}_{\beta\sigma} - y_{\beta\sigma})$$

and the pseudo-Hermitian form

$$\langle \bar{x} - x, \bar{y} - y \rangle \equiv \delta^{\alpha\beta} g^{\rho\sigma} (\bar{x}^*_{\alpha\rho} - x^*_{\alpha\rho})(\bar{y}_{\beta\sigma} - y_{\beta\sigma}).$$

Here the asterisk denotes complex conjugation and the metric $G^{\alpha\beta}$ of X_2 has the property

$$G^{\beta\alpha} = -G^{\alpha\beta}.$$

Since

$$G^{\alpha\beta} = G^{\alpha-\alpha}\delta^{-\alpha\beta},$$

we also have

$$G^{-\alpha-\beta} = -G^{\alpha\beta}.$$

The metric $g^{\rho\sigma}$ of V'_ν has the property

$$g^{\sigma\rho} = g^{\rho\sigma} \sim \begin{array}{c|c|c} & 1 \cdots \nu_1 & \nu_1 + 1 \cdots \nu \\ \hline 1 & 1 & \\ \vdots & \ddots & \\ \nu_1 & 1 & \\ \hline \nu_1 + 1 & & -1 \\ \vdots & & \ddots \\ \nu & & -1 \end{array}.$$

The generators

$$\left.\begin{aligned} N_{\rho\sigma} &= -N_{\sigma\rho} \\ N^r_{\rho\sigma} &= N^r_{\sigma\rho}, \quad r = 1, 2, 3 \end{aligned}\right\} \sim S^{\beta\sigma,\alpha\rho},$$

of $\mathfrak{US}\not{p}(2\nu_1, 2\nu_2)$ and the generators

$$Q_\sigma \sim b^{\beta\sigma}, \quad Q^1_\sigma \sim b^\sigma_\beta,$$
$$Q^2_\sigma \sim c^{-\beta\,\sigma}, \quad Q^3_\sigma \sim c^\sigma_{-\beta},$$
$$\forall\beta = 1, -1,$$

of $T(4\nu, R)$ form a basis for the real Lie algebra

$$IUSp(2\nu_1, 2\nu_2) \cong T(4\nu, R) \overleftarrow{+} USp(2\nu_1, 2\nu_2),$$

of $\mathfrak{JUS}\not{p}(2\nu_1, 2\nu_2)$. The generators of $\mathfrak{JUS}\not{p}(2\nu_1, 2\nu_2)$ satisfy the perhaps not so well-known commutation relations

$$[N_{\rho\sigma}, N_{\rho'\sigma'}] = g_{\sigma\rho'}N_{\rho\sigma'} - g_{\rho'\rho}N_{\sigma\sigma'} + g_{\sigma\sigma'}N_{\rho'\rho} - g_{\sigma'\rho}N_{\rho'\sigma}, \quad (1a)$$

$$[N_{\rho\sigma}, N^r_{\rho'\sigma'}] = g_{\sigma\rho'}N^r_{\rho\sigma'} - g_{\rho'\rho}N^r_{\sigma\sigma'} + g_{\sigma\sigma'}N^r_{\rho'\rho} - g_{\sigma'\rho}N^r_{\rho'\sigma}, \quad (1b)$$

$$[N^r_{\rho\sigma}, N^s_{\rho'\sigma'}] = \delta^{rs}(-g_{\sigma\rho'}N_{\rho\sigma'} - g_{\rho'\rho}N_{\sigma\sigma'} + g_{\sigma\sigma'}N_{\rho'\rho} + g_{\sigma'\rho}N_{\rho'\sigma}) - \epsilon^{rst}(g_{\sigma\rho'}N^t_{\rho\sigma'} + g_{\rho'\rho}N^t_{\sigma\sigma'} + g_{\sigma\sigma'}N^t_{\rho'\rho} + g_{\sigma'\rho}N^t_{\rho'\sigma}), \quad (1c)$$

$$[N_{\rho\sigma}, Q_{\sigma'}] = g_{\sigma\sigma'}Q_\rho - g_{\sigma'\rho}Q_\sigma, \quad (1d)$$

$$[N_{\rho\sigma}, Q^r_{\sigma'}] = g_{\sigma\sigma'}Q^r_\rho - g_{\sigma'\rho}Q^r_\sigma, \quad (1e)$$

$$[N^r_{\rho\sigma}, Q_{\sigma'}] = -g_{\sigma\sigma'}Q^r_\rho - g_{\sigma'\rho}Q^r_\sigma, \quad (1f)$$

$$[N^r_{\rho\sigma}, Q^s_{\sigma'}] = \delta^{rs}(g_{\sigma\sigma'}Q_\rho + g_{\sigma'\rho}Q_\sigma) - \epsilon^{rst}(g_{\sigma\sigma'}Q^t_\rho + g_{\sigma'\rho}Q^t_\sigma), \quad (1g)$$

$$[Q_\sigma, Q_{\sigma'}] = [Q_\sigma, Q^r_{\sigma'}] = [Q^r_\sigma, Q^s_{\sigma'}] = 0. \quad (1h)$$

Here and in the following the summation convention is also applied for the Latin indices, even though these only appear in the upper position. The second-order Casimir operator of $USp(2\nu_1, 2\nu_2)$ is given by

$$C_2(USp(2\nu_1, 2\nu_2)) = N_{\rho\sigma}N^{\rho\sigma} + N^r_{\rho\sigma}N^{r\rho\sigma}.$$

2. EXPANSION OF $IUSp(2\nu_1, 2\nu_2)$ to $USp(2\nu_1 + 2, 2\nu_2)$ OR TO $USp(2\nu_1, 2\nu_2 + 2)$

Our aim is to show that we can expand $IUSp(2\nu_1, 2\nu_2)$ to the Lie algebra of the *linear* group $\mathfrak{US}\not{p}(2\nu_1 + 2, 2\nu_2)$ or to $\mathfrak{US}\not{p}(2\nu_1, 2\nu_2 + 2)$. Instead of considering the 2ν-dimensional complex vector space $V_{2\nu}$ and the group $\mathfrak{JUS}\not{p}(2\nu_1, 2\nu_2)$, which is *nonlinear* because of $\mathfrak{C}(4\nu, R) \subset \mathfrak{JUS}\not{p}(2\nu_1, 2\nu_2)$, let us therefore extend V'_ν to

$$U'_{\nu+1} = V'_\nu + W'_1,$$

and consider the $(2\nu + 2)$-dimensional complex vector space

$$U_{2\nu+2} = U'_{\nu+1} \otimes X_2 = V_{2\nu} + W_2, \quad \nu = \nu_1 + \nu_2,$$

and the *real* Lie group $\mathfrak{S}(2\nu+2, R)$ of all those *complex linear* transformations

$$x'_{\alpha\rho} = \mathcal{S}^{\beta\sigma}_{\alpha\rho} x_{\beta\sigma} + b_{\alpha\rho} x_{1\,\nu+1} + i c_{\alpha\rho} x_{-1\,\nu+1},$$
$$\alpha, \beta = 1, -1, \quad \rho, \sigma = 1, \cdots, \nu,$$
$$x'_{\alpha\,\nu+1} = x_{\alpha\,\nu+1},$$

of $U_{2\nu+2}$, where

$$\mathcal{S}^{\beta\sigma}_{\alpha\rho} \in \mathfrak{US}\not{p}(2\nu_1, 2\nu_2), \quad b_{\alpha\rho}, c_{\alpha\rho} \in R.$$

It is seen that

$$\mathfrak{S}(2\nu+2, R) \cong \mathfrak{JUS}\not{p}(2\nu_1, 2\nu_2)$$

and, therefore, also has the Lie algebra given in Eqs. (1).

We shall now show that, out of the generators of $\mathfrak{S}(2\nu+2, R)$, we can construct certain new operators which obey the commutation relations of $USp(2\nu_1+2, 2\nu_2)$ or of $USp(2\nu_1, 2\nu_2+2)$.

To this end we shall extend the metric $g^{\rho\sigma}$ of V'_ν to all of $U_{\nu+1}$ by the following definitions:

$$g_{\nu+1\,\sigma} \equiv 0 \equiv g_{\sigma\,\nu+1},$$
$$g_{\nu+1\,\nu+1} \equiv \gamma \in R.$$

Let us now define the operators

$$Q^2 \equiv \gamma^{-1}(Q_\sigma Q^\sigma + Q^r_\sigma Q^{r\sigma}),$$
$$\bar{Q}_\sigma \equiv \tfrac{1}{4}[C_2(USp(2\nu_1, 2\nu_2)), Q_\sigma] = \{Q^\rho N_{\rho\sigma} - Q^{r\rho} N^r_{\rho\sigma}\},$$
$$\bar{Q}^r_\sigma \equiv \tfrac{1}{4}[C_2(USp(2\nu_1, 2\nu_2)), Q^r_\sigma]$$
$$= \{Q^\rho N^r_{\rho\sigma} + Q^{r\rho} N_{\rho\sigma} - \epsilon^{rst} Q^{s\rho} N^t_{\rho\sigma}\},$$
$$L^r \equiv \{\bar{Q}^r_\sigma Q^\sigma - \bar{Q}_\sigma Q^{r\sigma} + \epsilon^{rst} \bar{Q}^s_\sigma Q^{t\sigma}\}$$
$$= \{2Q^{r\rho}(Q^{s\sigma} N^s_{\rho\sigma} + Q^\sigma N_{\rho\sigma}) - (Q^{s\rho} Q^{s\sigma} - Q^\rho Q^\sigma) N^r_{\rho\sigma}$$
$$- \epsilon^{rst} Q^{s\rho}(2Q^\sigma N^t_{\rho\sigma} - Q^{t\sigma} N_{\rho\sigma})\},$$

where $\{\ \}$ denotes that the expression inside the bracket has been symmetrized with respect to the position of the Q^σ and $Q^{r\sigma}$ relative to $N_{\rho\sigma}$ and $N^r_{\rho\sigma}$ and to $\bar{Q}_\sigma$ and $\bar{Q}^r_\sigma$ (but not relative to one another as they commute) and then divided by the number of terms.

In terms of the above operators we define

$$\bar{N}'_{\rho\sigma} \equiv N_{\rho\sigma},$$
$$\bar{N}'^r_{\rho\sigma} \equiv N_{\rho\sigma},$$
$$\bar{N}'_{\nu+1\,\sigma} \equiv Q_\sigma + \lambda \bar{Q}_\sigma / \sqrt{-Q^2},$$
$$\bar{N}'^r_{\nu+1\,\sigma} \equiv Q^r_\sigma + \lambda \bar{Q}^r_\sigma / \sqrt{-Q^2},$$
$$\bar{N}'^r_{\nu+1\,\nu+1} \equiv -L^r / Q^2,$$

where λ is a free parameter. For these operators we obtain the following commutation relations:

$$[\bar{N}'_{\rho\sigma}, \bar{N}'_{\rho'\sigma'}]$$
$$= g_{\sigma\rho'} \bar{N}'_{\rho\sigma'} - g_{\rho'\rho} \bar{N}'_{\sigma\sigma'} + g_{\sigma\sigma'} \bar{N}'_{\rho'\rho} - g_{\sigma'\rho} \bar{N}'_{\rho'\sigma}, \tag{2a}$$

$$[\bar{N}'_{\rho\sigma}, \bar{N}'^r_{\rho'\sigma'}]$$
$$= g_{\sigma\rho'} \bar{N}'^r_{\rho\sigma'} - g_{\rho'\rho} \bar{N}'^r_{\sigma\sigma'} + g_{\sigma\sigma'} \bar{N}'^r_{\rho'\rho} - g_{\sigma'\rho} \bar{N}'^r_{\rho'\sigma}, \tag{2b}$$

$$[\bar{N}'^r_{\rho\sigma}, \bar{N}'^s_{\rho'\sigma'}]$$
$$= \delta^{rs}(-g_{\sigma\rho'} \bar{N}'_{\rho\sigma'} - g_{\rho'\rho} \bar{N}'_{\sigma\sigma'} + g_{\sigma\sigma'} \bar{N}'_{\rho'\rho} + g_{\sigma'\rho} \bar{N}'_{\rho'\sigma})$$
$$- \epsilon^{rst}(g_{\sigma\rho'} \bar{N}'^t_{\rho\sigma'} + g_{\rho'\rho} \bar{N}'^t_{\sigma\sigma'} + g_{\sigma\sigma'} \bar{N}'^t_{\rho'\rho} + g_{\sigma'\rho} \bar{N}'^t_{\rho'\sigma}), \tag{2c}$$

$$[\bar{N}'_{\rho\sigma}, \bar{N}'_{\nu+1\,\sigma'}] = g_{\sigma\sigma'} \bar{N}'_{\nu+1\,\rho} - g_{\sigma'\rho} \bar{N}'_{\nu+1\,\rho}, \tag{2d}$$

$$[\bar{N}'_{\rho\sigma}, \bar{N}'^r_{\nu+1\,\sigma'}] = g_{\sigma\sigma'} \bar{N}'^r_{\nu+1\,\rho} - g_{\sigma'\rho} \bar{N}'^r_{\nu+1\,\sigma}, \tag{2e}$$

$$[\bar{N}'^r_{\rho\sigma}, \bar{N}'_{\nu+1\,\sigma'}] = -g_{\sigma\sigma'} \bar{N}'^r_{\nu+1\,\rho} + g_{\sigma'\rho} \bar{N}'^r_{\nu+1\,\sigma}, \tag{2f}$$

$$[\bar{N}'^r_{\rho\sigma}, \bar{N}'^s_{\nu+1\,\sigma'}] = \delta^{rs}(g_{\sigma\sigma'} \bar{N}'_{\nu+1\,\rho} + g_{\sigma'\rho} \bar{N}'_{\nu+1\,\sigma})$$
$$- \epsilon^{rst}(g_{\sigma\sigma'} \bar{N}'^t_{\nu+1\,\rho} + g_{\sigma'\rho} \bar{N}'^t_{\nu+1\,\sigma}), \tag{2g}$$

$$[\bar{N}'_{\rho\sigma}, \bar{N}'^r_{\nu+1\,\nu+1}] = 0, \tag{2h}$$

$$[\bar{N}'^r_{\rho\sigma}, \bar{N}'^s_{\nu+1\,\nu+1}] = 0, \tag{2i}$$

$$[\bar{N}'_{\nu+1\,\sigma}, \bar{N}'_{\nu+1\,\sigma'}] = -\lambda^2 g_{\nu+1\,\nu+1} \bar{N}'_{\sigma\sigma'}, \tag{2j}$$

$$[\bar{N}'_{\nu+1\,\sigma}, \bar{N}'^r_{\nu+1\,\sigma'}] = \lambda^2(-g_{\nu+1\,\nu+1} \bar{N}'^r_{\sigma\sigma'} + g_{\sigma\sigma'} \bar{N}'^r_{\nu+1\,\nu+1}), \tag{2k}$$

$$[\bar{N}'^r_{\nu+1\,\sigma}, \bar{N}'^s_{\nu+1\,\sigma'}] = -\lambda^2(\delta^{rs} g_{\nu+1\,\nu+1} \bar{N}'_{\sigma\sigma'}$$
$$+ \epsilon^{rst}(g_{\nu+1\,\nu+1} \bar{N}'^t_{\sigma\sigma'} + g_{\sigma\sigma'} \bar{N}'^t_{\nu+1\,\nu+1})), \tag{2l}$$

$$[\bar{N}'^r_{\nu+1\,\nu+1}, \bar{N}'_{\nu+1\,\sigma}] = 0, \tag{2m}$$

$$[\bar{N}'^r_{\nu+1\,\nu+1}, \bar{N}'^s_{\nu+1\,\sigma}] = -2g_{\nu+1\,\nu+1}(\delta^{rs} \bar{N}'_{\nu+1\,\sigma} + \epsilon^{rst} \bar{N}'^t_{\nu+1\,\sigma}), \tag{2n}$$

$$[\bar{N}'^r_{\nu+1\,\nu+1}, \bar{N}'^s_{\nu+1\,\nu+1}] = -4\epsilon^{rst} g_{\nu+1\,\nu+1} \bar{N}'^t_{\nu+1\,\nu+1}. \tag{2o}$$

In order to remove the parameter λ from the above commutation relations, we finally define the operators

$$\bar{N}_{\rho\sigma} \equiv \bar{N}'_{\rho\sigma},$$
$$\bar{N}^r_{\rho\sigma} \equiv \bar{N}'^r_{\rho\sigma},$$
$$\bar{N}_{\nu+1\,\sigma} \equiv \lambda^{-1} \bar{N}'_{\nu+1\,\sigma} \equiv \bar{N}_{\sigma\,\nu+1},$$
$$\bar{N}^r_{\nu+1\,\sigma} \equiv \lambda^{-1} \bar{N}'^r_{\nu+1\,\sigma} \equiv \bar{N}^r_{\nu+1\,\sigma},$$
$$\bar{N}^r_{\nu+1\,\nu+1} \equiv \bar{N}'^r_{\nu+1\,\nu+1},$$

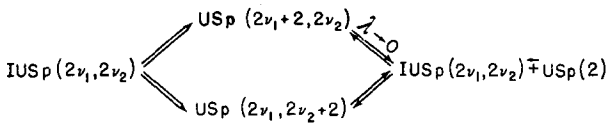


FIG. 1. Expansion, contraction, and deformation diagram for the pseudo-unitary symplectic Lie algebras. Here $\Rightarrow$, $\rightarrow$, and $\leftarrow$ denote expansion, contraction, and deformation, respectively.

which then satisfy

$$
\begin{aligned}
[\bar{N}_{AB}, \bar{N}_{A'B'}] &= g_{BA'}\bar{N}_{AB'} - g_{A'A}\bar{N}_{BB'} \\
&\quad + g_{BB'}\bar{N}_{A'A} - g_{B'A}\bar{N}_{A'B}, \\
[\bar{N}_{AB}, \bar{N}^r_{A'B'}] &= g_{BA'}\bar{N}^r_{AB'} - g_{A'A}\bar{N}^r_{BB'} \\
&\quad + g_{BB'}\bar{N}^r_{A'A} - g_{B'A}\bar{N}^r_{A'B}, \\
[\bar{N}^r_{AB}, \bar{N}^s_{A'B'}] &= \delta^{rs}(-g_{BA'}\bar{N}_{AB'} - g_{A'A}\bar{N}_{BB'} \\
&\quad + g_{BB'}\bar{N}_{A'A} + g_{B'A}\bar{N}_{A'B}) \\
&\quad - \epsilon^{rst}(g_{BA'}\bar{N}^t_{AB'} + g_{A'A}\bar{N}^t_{BB'} \\
&\quad + g_{BB'}\bar{N}^t_{A'A} + g_{B'A}\bar{N}^t_{A'B}), \\
&\qquad A, B = 1, \cdots, \nu + 1,
\end{aligned}
$$

where

$$\bar{N}_{BA} = -\bar{N}_{AB}, \quad \bar{N}^r_{BA} = \bar{N}^r_{AB}, \text{ and } \quad g_{BA} = g_{AB}.$$

These commutation relations are evidently those of $USp(2\nu_1 + 2\nu_1', 2\nu_2 + 2\nu_2')$ where

$$\nu_1' + \nu_2' = 1, \quad \nu_1' - \nu_2' = \gamma/|\gamma| = \pm 1.$$

We have shown that $IUSp(2\nu_1, 2\nu_2)$ *can be expanded to* $USp(2\nu_1 + 2, 2\nu_2)$ *or to* $USp(2\nu_1, 2\nu_2 + 2)$. *These expansions correspond globally to the expansions of* $\mathfrak{g}(2\nu + 2, R) \cong \mathfrak{IUS}\not{p}(2\nu_1, 2\nu_2)$ *to a linear Lie group* $\mathcal{E}(2\nu_1 + 2, 2\nu_2; R) \cong \mathfrak{US}\not{p}(2\nu_1 + 2, 2\nu_2)$ or to $\mathcal{E}(2\nu_1, 2\nu_2 + 2; R) \cong \mathfrak{US}\not{p}(2\nu_1, 2\nu_2 + 2)$.

By analyzing the commutation relations (2), we find that, when taking the limit $\lambda \to 0$, $USp(2\nu_1 + 2\nu_1', 2\nu_2 + 2\nu_2')$ contracts[21–23] to

$$
\begin{aligned}
T(4\nu, R) \overset{\leftarrow}{+} (USp(2\nu_1, 2\nu_2) + USp(2)) \\
\cong IUSp(2\nu_1, 2\nu_2) \overset{\leftarrow}{+} USp(2),
\end{aligned}
$$

where $+$ denotes direct sum, i.e., $USp(2\nu_1 + 2\nu_1', 2\nu_2 + 2\nu_2')$ is an expansion of $IUSp(2\nu_1, 2\nu_2)$ but a deformation[24–27] of $IUSp(2\nu_1, 2\nu_2) + USp(2)$ (see Fig. 1).

ACKNOWLEDGMENT

It is a pleasure for the author to thank Dr. M. Levy-Nahas for a stimulating discussion and Professor M. Flato for an interesting comment on the expansion problem. Furthermore, the author would like to express his gratitude to Professor N. Svartholm and to Professor J. Nilsson for the hospitality extended to him at Chalmers University of Technology.

[1] N. Bourbaki, *Algèbre* (Hermann, Paris, 1958), Chap. I, p. 145.
[2] I. M. Gel'fand and A. A. Kirillov, Dokl. Akad. Nauk SSSR **167**, 503 (not 407 as quoted in Ref. 4) (1966); Publ. Math. France **31**, 509 (1966).
[3] Y. Chow, J. Math. Phys. **10**, 975 (1969).
[4] A. Simoni and F. Zaccaria, Nuovo Cimento **59A**, 280 (1969).
[5] M. A. Melvin, Bull. Am. Phys. Soc. **7**, 493 (1962).
[6] A. Sankaranarayanan and R. H. Good Jr., Phys. Rev. **140**, B509 (1965).
[7] A. Sankaranarayanan, Nuovo Cimento **38**, 1441 (1965).
[8] A. Böhm, Phys. Rev. **145**, 1212 (1966).
[9] M. Y. Han, Nuovo Cimento **42B**, 367 (1966).
[10] A. Sankaranarayanan, Nuovo Cimento **44B**, 193 (1966).
[11] J. Rosen and P. Roman, J. Math. Phys. **7**, 2072 (1966).
[12] J. Rosen, Nuovo Cimento **46B**, 1 (1966); J. Math. Phys. **9**, 1305 (1968).
[13] A. Sankaranarayanan, J. Math. Phys. **9**, 611 (1968).
[14] A. Sankaranarayanan, Nuovo Cimento **52A**, 91 (1967); Douglas Advanced Research Laboratories, Huntington Beach, California, Preprint, April, 1969.
[15] Y. Ne'eman, *Algebraic Theory of Particle Physics* (W. A. Benjamin, New York, 1967).
[16] A. Chakrabarti, J. Math. Phys. **9**, 2087 (1968).
[17] J. G. Nagel, Institute of Theoretical Physics, Chalmers University of Technology, Göteborg, Sweden, Preprint 69-13, 1969, to appear in Ann. Inst. Henri Poincaré.
[18] J. G. Nagel and K. T. Shah, J. Math. Phys. **11**, 1483 (1970).
[19] T. O. Philips and E. P. Wigner, in *Group Theory and Its Applications*, E. M. Loebl, Ed. (Academic Press, New York, 1968).
[20] M. Gourdin, *Unitary Symmetries* (North-Holland Publishing Co., Amsterdam, 1967), Chap. XIII.
[21] I. E. Segal, Duke Math. J. **18**, 221 (1951).
[22] E. Inönü and Wigner, Proc. Natl. Acad. Sci. U.S. **39**, 510 (1953).
[23] E. J. Saletan, J. Math. Phys. **2**, 1 (1961).
[24] M. Gerstenhaber, Ann. Math. **79**, 59 (1964).
[25] R. Hermann, Commun. Math. Phys. **2**, 251 (1966); **3**, 53, 75 (1966); **5**, 131, 157 (1967); **6**, 205 (1967).
[26] M. Levy-Nahas, J. Math. Phys. **8**, 1211 (1967), to appear (in French) in Compt. Rend. Colloque Gif-sur-Yvette.
[27] M. Levy-Nahas and R. Seneor, Commun. Math. Phys. **9**, 242 (1968).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 5 MAY 1970

Recurrence Relations for the Multiplicities in the Classical Groups

B. GRUBER
*Research Institute for Theoretical Physics, University of Helsinki**
and
Department of Physics, St. Louis University, St. Louis, Missouri 63103†

(Received 14 July 1969)

Recurrence relations for the "outer multiplicity" (multiplicity of irreducible representations occurring in the reduction of the direct product of two irreducible representations) as well as the "branching multiplicity" (multiplicity of irreducible representations into which an irreducible representation of a group decomposes, if this group is restricted to a subgroup) are obtained. These recurrence relations are shown to be closely related to each other. Racah has obtained a recurrence relation for the "inner multiplicity" (multiplicity of weights). It turns out that the recurrence relation for the inner multiplicity is a special case of the recurrence relation for the outer multiplicity which, in turn, can be looked upon as a special case of recurrence relation for the branching multiplicity. Moreover, Kostant's recurrence relation for the partition function as well as Racah's recurrence relation for the inner multiplicity have been generalized. This generalization introduces a parameter which can be used to simplify actual calculations. In an appendix, finally, a formula is given which allows one to calculate the branching multiplicity from known inner multiplicities of the subgroup to which it is restricted.

I. INTRODUCTION

The multiplicity of irreducible representations occurring in the decomposition of tensor products of two irreducible representations is called "outer multiplicity" throughout this paper. By "branching multiplicity" is meant the multiplicity of irreducible representations of a subgroup into which an irreducible representation of a group decays if the group is restricted to this subgroup. Finally, by "inner multiplicity" is meant the multiplicity of weights.

The groups for which these multiplicities are studied are the semisimple compact and connected Lie groups, namely $SU(l+1)$, $SO(2l+1)$, $Sp(2l)$, and $SO(2l)$, and the five exceptional groups G_2, F_4, E_6 E_7, and E_8.[1,2]

A. Review of Present Situation

Before giving an outline of the contents of this paper, it might be of interest to review the present status of these multiplicities.

There already exist several explicit formulas for the three multiplicities. For the inner multiplicity there are available Kostant's formula[3] and Freudental's formula.[4] Kostant's formula is given as a sum over the Weyl group W in terms of the so-called partition function.[5] Freudenthal's formula is a recurrence relation which involves a summation over all positive multiples of the positive roots. Neither of these formulas is very easy to work with in actual calculations. Racah[6] obtained a recurrence relation for the inner multiplicity. This recurrence relation seems to be the most convenient formula for actual calculations.

An explicit formula also exists for the outer multiplicity, namely Steiberg's formula.[7] This formula has a structure similar to Kostant's formula, namely it is given as a sum over the same partition function which appears in Kostant's formula. However, it involves a double summation over the Weyl group W. There are yet other types of formulas for the outer multiplicity, namely relationships which express the outer multiplicity in terms of the inner multiplicity. Such a formula, relating the outer multiplicity to a sum which is quadratic in inner multiplicities, was obtained by Straumann[8] and by Gruber.[9]

A simpler formula which relates the outer multiplicity to the inner multiplicity was obtained by Racah[6] (see also Ref. 9). This formula is linear and is an explicit expression for the so-called Racah–Speiser lemma.[10] Of the formulas mentioned, this formula seems to be the easiest one to use in actual calculations. This is true in particular for the case of the $SU(n)$ groups for which the inner multiplicities can be obtained easily.[11,8]

Finally, we consider the branching multiplicity. Mandel'tsveig[12] has obtained an expression for the branching multiplicity with respect to subgroups which are "naturally embedded" and maximal. A naturally embedded subgroup is a subgroup whose algebra is reductive (a direct sum of a semisimple Lie algebra and a commutative algebra), such that its root system is part of the root system of the algebra of the group. Maximal means that the algebra of the subgroup must be of the same rank as the algebra of the group. Then Mandel'tsveig's formula is given as a sum over the Weyl group W of the group in terms of

a partition function (this partition function is different from the one in Kostant's and Steinberg's formula). As in the case of outer multiplicity, there also exists in the case of branching multiplicity a formula relating the branching multiplicity to the inner multiplicity. This formula was first obtained by Straumann[8] (Kostant's formula for the inner multiplicity is written out explicitly in his expression), then rederived independently by Klymyk[13] and also by Delaney and Gruber.[11] This formula holds for a wider class of subgroups than Mandel'tsveig's formula does, namely for all subgroups such that the toroid of the subgroup is part of the toroid of the group (the Cartan algebra of the subgroup is part of the Cartan algebra of the group).

B. Outline of Paper

In this paper, *recurrence relations* for the outer and branching multiplicity are obtained. These recurrence relations are analogous to the one obtained by Racah[6] for the inner multiplicity. In fact, it is shown that Racah's recurrence relation for the inner multiplicity is a special case of the recurrence relation for the outer multiplicity. This special case appears under the conditions of the Biedenharn lemma,[10] i.e., when there are no cancellations in the Clebsch–Gordan series as generally happens by virtue of the Racah–Speiser lemma.[10] Moreover, the recurrence relation for the branching multiplicity can be seen to contain in turn the recurrence relation for the outer multiplicity. This again holds under certain circumstances, namely when in this recurrence relation the subgroup to which the group is restricted is chosen to be the group itself. Thus, the recurrence relation for the branching multiplicity turns out to be the most fundamental recurrence relation, containing both the recurrence relations for outer and inner multiplicities as special cases. However, this is not too surprising. After all, the Biedenharn lemma is a special case of the Racah–Speiser lemma (the limits for the validity of the Biedenharn lemma have been determined by Zaccaria[14] and Vitale[15]), and the fact that the outer multiplicity can be looked upon as a special case of the branching multiplicity has been noticed already by Straumann.[8]

In addition to recurrence relations for outer and branching multiplicities, this paper contains a generalization of Racah's recurrence relation for the inner multiplicity as well as a generalization of Kostant's recurrence relation for the partition function.[16] These generalizations are of particular value in actual calculations. Due to the introduction of a parameter into the recurrence relations (the parameter is a dominant weight satisfying a condition), it is generally possible to reduce in an actual calculation the number of successive steps considerably.

In the Appendix, a formula for the branching multiplicity is given. This formula expresses the branching multiplicity in terms of inner multiplicities of the subgroup and the Weyl group of the group. It is clear that this formula is of particular interest in the case when the inner multiplicities of the subgroup are known. Then this formula reduces to a system of linear equations for the unknown branching multiplicities.

II. NOTATION AND DEFINITION

Let G denote a semisimple compact and connected Lie group, and Q its (semisimple compact) Lie algebra of rank l. With $H_1, H_2, \cdots, H_l$, we denote the l independent generators of its Cartan subalgebra. The eigenvalues of this set $H_1, H_2, \cdots, H_l$, corresponding to a simultaneous eigenvector, are denoted by $m_1, m_2, \cdots, m_l$. Thus, a general weight is denoted by $m = (m_1, m_2, \cdots, m_l)$. A dominant weight (the highest weight of a set of equivalent weights Sm, $S \in W$, where S is an arbitrary element of the Weyl group W) is denoted by $M = (M_1, M_2, \cdots, M_l)$. Thus, in particular, the highest weight of an irreducible representation is denoted by $M = (M_1, M_2, \cdots, M_l)$.

The inner multiplicity is denoted by $\gamma(m)$ [or $\gamma^M(m)$], i.e., $\gamma(m)$ is the multiplicity of the weight m [which is contained in the irreducible representation $D(M)$]. Similarly, $\tilde{\gamma}(M)$ denotes the outer multiplicity, i.e., the multiplicity of the irreducible representation $D(M)$ with highest weight M. All symbols referring to a subgroup G_s of G are characterized by a suffix s. Thus, m_s, M_s, S_s, and W_s stand for a general weight, a dominant weight, an element of the Weyl group, and the Weyl group of the subgroup, respectively. Then $\tilde{\gamma}(M_s)$ denotes the branching multiplicity, i.e., the multiplicity of the irreducible representation $D(M_s)$ of the subgroup.

$[m, \varphi] = m_1\varphi_1 + m_2\varphi_2 + \cdots + m_l\varphi_l$ is a linear form in terms of the parameters φ_i of the toroid of the group G (Cartan subgroup of the group G). If the group G is restricted to a subgroup G_s, then

$$[m, \varphi] \to [m_s, \varphi_s] \equiv m_{s1}\varphi_{s1} + \cdots + m_{sl'}\varphi_{sl'},$$

where $l' \leq l$ is the rank of the subgroup G_s. We look at this restriction as a mapping L such that[11]

$$L(m) = m_s,$$

$$[m, \varphi]_{\text{restricted}} = [L(m), \varphi_s] = [m_s, \varphi_s]. \qquad (1)$$

R and R_s denote half the sum over the positive roots of the group G and subgroup G_s, respectively. $\delta_S = -1$ if $S \in W$ is a reflection; otherwise $\delta_S = +1$. $\delta_{x,y}$

denotes the Kronecker symbol; i.e., $\delta_{x,y} = 1$ if the two weights x and y are equal, and $\delta_{x,y} = 0$ otherwise.

The positive roots are denoted by α and, if necessary to distinguish them, by α_i, $i = 1, 2, \cdots$. By convention, the first l of the α_i are taken to be the simple positive roots. Similarly, β denotes the negative roots, and $\beta_1, \beta_2, \cdots, \beta_l$ the simple negative roots.

III. THE PARTITION FUNCTION

From what was said in Sec. I.A, it can easily be recognized that all three multiplicities, namely inner, outer, and branching multiplicity, can be expressed in terms of one and the same partition function. The respective formulas are Kostant's formula,[3] Steinberg's formula,[7] and Straumann's formula.[8] Thus the partition function is of considerable interest and it might be worthwhile to say a few words about it.

The partition function $P(m)$ with

$$m = k_1'\beta_1 + \cdots + k_l'\beta_l, \quad \text{all } k_i' \text{ integers},$$

is defined as follows:

$P(m)$ = number of solutions of

$$\sum_{\beta<0} k_\beta \beta = m, \quad k_\beta \text{ nonnegative integers}, \tag{2}$$

where the sum extends over *all* negative roots β. It is clear from definition (2) that

$$P(m) = 0, \quad \text{if any } k_i' < 0, \quad i = 1, 2, \cdots, l. \tag{3}$$

An example will make the meaning of the partition function easily understandable, and we choose the partition function of $SU(3)$ for this purpose. The weights of $SU(3)$ are taken to be of the form

$$m = (m_1, m_2, m_3), \quad m_1 + m_2 + m_3 = 0,$$
$$m_i = \tfrac{1}{3} \times (\text{integer})$$

(hypercharge Y and third component of isotopic spin T_3 are then given as $Y = -m_2$, $T_3 = m_1 + \frac{1}{2}m_2$). Then $P(m)$, where

$$P(m) = P(k_1'\beta_1 + k_2'\beta_2) \equiv P(k_1', k_2'),$$

is the number of solutions of

$$k_1' = k_1 + k_2,$$
$$k_2' = k_1 + k_3, \quad k_i', k_j \text{ nonnegative integers};$$

i.e.,

$$P(k_1', k_2') = \min(k_1', k_2') + 1, \quad k_1', k_2' \geq 0,$$
$$= 0, \quad \text{for any } k_i' < 0.$$

In general, however, it is very complicated to obtain an explicit expression for the values of the partition function. Tarski[17] has studied this problem for $SU(4)$ and the rank-two groups. Explicit formulas for the partition function of $SU(4)$ and the rank-two groups, except G_2, have been given in Refs. 18 and 19. Kostant[16] has obtained a very elegant and useful recurrence relation for the partition function in general. It is the aim of the present section to generalize Kostant's recurrence relation.

Kostant's recurrence relation is given as

$$P(m) = -\sum_{\substack{S\in W \\ S\neq 1}} \delta_S P(m + R - SR), \tag{4}$$

where the sum extends over all elements S of the Weyl group W, except for the identity. A generalization of Eq. (4) is obtained as follows:

Kostant's formula for the inner multiplicity reads as[3]

$$\gamma^M(m) = \sum_{S\in W} \delta_S P(m + R - S(M + R)). \tag{5}$$

Now $\gamma^M(m) = 0$ if $m \notin D(M)$, by virtue of Eq. (3). Thus, if $|m| > |M|$, we obtain the generalized Kostant recurrence relation as[20]

$$P(m - M) = -\sum_{\substack{S\in W \\ S\neq 1}} \delta_S P(m + R - S(M + R)),$$
$$|m| > |M|. \tag{6}$$

For $M = 0$, the original equation is obtained. Equation (6) has the advantage that, through suitable choice of the parameter M, the number of steps in a calculation of some $P(m)$ can be reduced considerably, except for very small weights m ("small" and "large" are meant with respect to the length of the weight). In order to illustrate this point, a simple example is given, again for the partition function $P(m)$ of $SU(3)$. The value of the weight m is chosen to be $m = (-3, 0, 3)$. Then one obtains, from Eq. (4),

$$\begin{aligned} P(-3, 0, 3) = -[&-P(-2, -1, 3) - P(-3, 1, 2) \\ &+ P(-1, -1, 2) + P(-2, 1, 1) \\ &- P(-1, 0, 1)] \\ = -[&-3 - 3 + 2 + 2 - 2] = 4, \end{aligned}$$

if the values on the right-hand side of the equation are assumed to be already known. On the other hand, from Eq. (6) follows, with $M = (1, 0, -1)$,

$$\begin{aligned} P(-3, 0, 3) &= -[-P(-1, -2, 3) - P(-3, 2, 1)] \\ &= -[-2 - 2] = 4. \end{aligned}$$

See Fig. 1.

IV. INNER MULTIPLICITY

In this section, Racah's recurrence relation for the inner multiplicity will be generalized. Racah's recurrence relation is[6]

$$\gamma^M(m) = -\sum_{\substack{S\in W \\ S\neq 1}} \delta_S \gamma^M(m + R - SR), \quad \gamma^M(M) = 1. \tag{7}$$

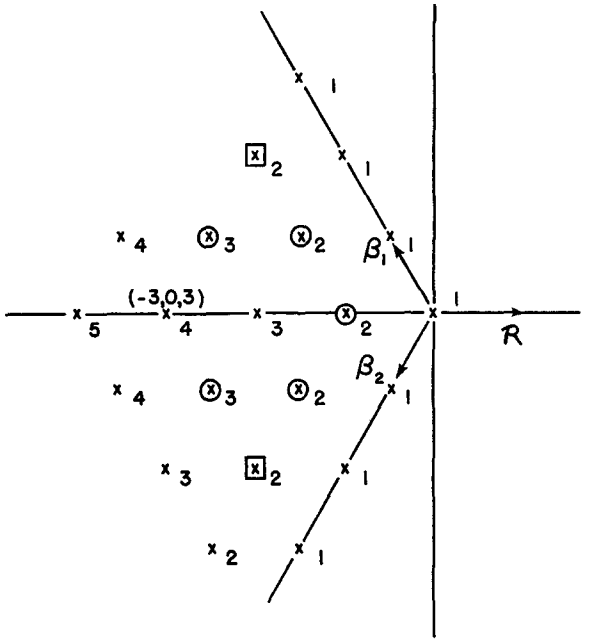


FIG. 1. $\beta_1 = (0, -1, 1)$, $\beta_2 = (-1, 1, 0)$, $R = -(\beta_1 + \beta_2)$; the five encircled points correspond to the example given for Kostant's recurrence relation and add up (taking care of δ_s) to 4. The two points characterized by squares correspond to the example given for the generalized Kostant recurrence relation. Again, they add up to 4, the value of $P(-3, 0, 3)$.

In order to generalize Eq. (7), the generalized Kostant recurrence relation for the partition function, namely Eq. (6), is inserted in Kostant's formula (5) for the inner multiplicity. This yields

$$\gamma^M(m) = \sum_{S\in W} \delta_S (-1) \sum_{\substack{S'\in W \\ S'\neq 1}} \delta_{S'} P(m + R - S(M + R) + R + \bar{M} - S'(\bar{M} + R)) = -\sum_{\substack{S'\in W \\ S'\neq 1}} \delta_{S'} \gamma^M(m + R + \bar{M} - S'(\bar{M} + R)), \tag{8a}$$

with the condition

$$|m + R + \bar{M} - S(M + R)| > |\bar{M}|, \quad S \in W, \quad m \neq M. \tag{8b}$$

For $m = M$, we have

$$\gamma^M(M) = 1. \tag{8c}$$

Equations (8) are the desired relations and, for $\bar{M} = 0$, Racah's original formula is obtained. For reasons of illustration, a simple example is given in the following. The $SU(3)$ representation $D(2, 0, -2)$ is considered, and the multiplicity of the weight $(-1, -1, 2)$ determined. For the parameter $\bar{M}$ the value $\bar{M} = (1, 0, -1)$ $(= R)$ is chosen. For this value of $\bar{M}$ Eq. (8b) is satisfied. Then, it follows from Eq. (8a) that

$$\gamma(-1, -1, 2) = -[\gamma(-1, 1, 0) + \gamma(1, 1, -2)] = -[-2 + 1] = 1,$$

where the values on the right side of this equation were

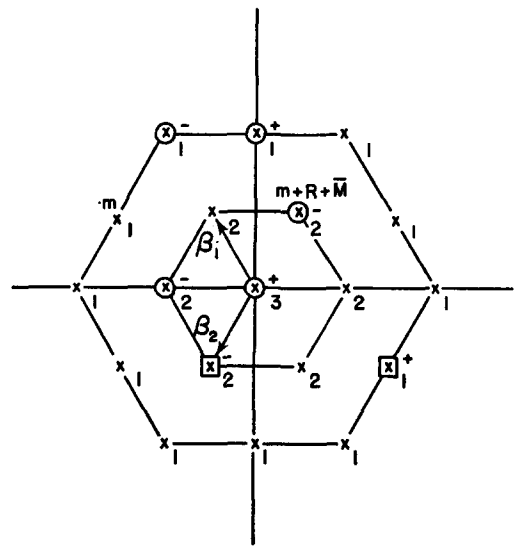


FIG. 2. The representation $D(2, 0, -2)$; the numbers denote the multiplicities of the weights, the signs the values of δ_s. The points characterized by squares correspond to the example given, the encircled points correspond to a straight application of Racah's recurrence relation. β_1 and β_2 are as in Fig. 1.

assumed to be already known. This compares with six terms on the right-hand side when Racah's relation is being used. See Fig. 2. Thus, for large representations and when only the multiplicity of a particular weight is of interest, Eqs. (8) can save considerable work.

V. OUTER MULTIPLICITY

In this section, a recurrence relation for the outer multiplicity is obtained in a very simple manner. Let $\chi(M)$ denote the character of the irreducible representation $D(M)$. Then the product of two characters is given as[21]

$$\chi(M) \cdot \chi(M') = \sum_{\bar{M}''} \bar{\gamma}(M'')\chi(M''). \tag{9}$$

Inserting Weyl's formula[22] for all three characters in Eq. (9), one obtains

$$\sum_{\substack{S\in W \\ S'\in W}} \delta_S \delta_{S'} e^{i[S(M+R)+S'(M'+R),\varphi]} = \sum_{M''} \bar{\gamma}(M'') \sum_{\substack{S''\in W \\ S\in W}} \delta_S \delta_{S''} e^{i[S''(M''+R)+SR,\varphi]}. \tag{10}$$

From Eq. (10) (using the orthogonality of the trigonometric functions), by multiplying both sides with $e^{-i[\bar{M}+2R,\varphi]}$ and by integrating over all parameters φ_i, one easily obtains the recurrence relation for the outer multiplicity,[21]

$$\bar{\gamma}(\bar{M}) = -\sum_{\substack{S,S''\in W \\ S=S''\neq 1}} \delta_S \bar{\gamma}(S''(\bar{M} + 2R) - SR - R) + \sum_{S,S'\in W} \delta_S \delta_{S'} \delta_{S(M+R)+S'(M'+R),\bar{M}+2R}. \tag{11}$$

In deriving Eq. (11), the fact has been used that $S(m + m') = Sm + Sm'$,

$$\delta_S = \delta_{S'}, \quad \delta_{SS'} = \delta_S \delta_{S'},$$

and

$$\sum_{S \in W} S''SR = \sum_{S \in W} SR,$$

together with the fact that the values of the integrals involved are invariant under Weyl reflections.

Looking at Eq. (11), one realizes immediately that the first line of it is just Racah's recurrence relation for the inner multiplicity if $S'' = 1$. On the other hand, it is known that outer and inner multiplicity are the same under certain conditions, namely when one of the two representations forming the direct product "dominates" the other.[14,15] This is just the content of the Biedenharn lemma.[10] These considerations suggest finding the conditions under which Eq. (11) goes over into Racah's recurrence relation (7). It is shown that this happens under the condition that

$$M + S'M' + SR = \text{dominant weight}, \quad \forall S, S' \in W. \tag{12}$$

It should be noted that this condition is stronger than the condition for an irreducible representation $D(M)$ to dominate the irreducible representation $D(M')$, namely that

$$M + S'M' = \text{dominant weight}, \quad \forall S' \in W. \tag{13}$$

Assuming that Eq. (12) holds, we now show that the second term of Eq. (11) disappears, except for the case of $M'' = M + M'$, where this term equals one, and $S'' = 1$ only.

The weights M'' which occur as highest weights of the irreducible representations $D(M'')$ in the Clebsch–Gordan series $D(M) \otimes D(M') = \sum_{M''} \bar{\gamma}(M'')D(M'')$ are contained in the polygon $\mathfrak{P}(M, M')$ with extremal points $M + S'M'$, $S' \in W$. Thereby, the weights M'' can be reached from any one of these extremal points by means of the root vectors. Thus, one can write $M'' \in \mathfrak{P}(M, M')$, meaning that M'' is a weight contained in the polygon $\mathfrak{P}(M, M')$. Then, if Eq. (12) is assumed to hold, the weights $M + R + S'M' + SR$, $S, S' \in W$, are not only dominant, but lie *inside* the fundamental domain (i.e., *not* on a singular hyperplane). The set of weights $M + R + S'(M' + R)$, $S' \in W$, forms a subset of this set of weights. Therefore, the weights

$$S(M + R + S'(M' + R)), \quad S, S' \in W, \quad S \neq 1,$$

are all nondominant. The same set of weights is also given as $S(M + R) + S'(M' + R)$, $S, S' \in W$, $S \neq 1$. However, it should be true that

$$S(M + R) + S'(M' + R) = \bar{M} + 2R. \tag{14}$$

Since the right-hand side of this equation is a dominant weight, it follows that this relation can hold, at most, for $S = 1$. In the following, it is also shown that $S' = 1$ is the only possibility and, thus, $\bar{M} = M + M'$.

In the equation

$$M + S'(M' + R) = \bar{M} + R,$$

the weights $\bar{M}$ are contained, by assumption, in the polygon $\mathfrak{P}(M, M')$. The polygon $\mathfrak{P}(M, M')$ is, however, contained *inside* the polygon $\mathfrak{P}(M, M' + R)$ with extremal points $M + S'(M' + R)$, $S' \in W$, since

$$|S'M'| < |S'(M' + R)|.$$

Shifting the polygon $\mathfrak{P}(M, M')$ by the vector R brings into concidence the two extremal points $M + M'$ of $\mathfrak{P}(M, M')$ and $M + M' + R$ of $\mathfrak{P}(M, M' + R)$, while all other extremal points do not coincide with each other. From this, it follows that the relation

$$M + R + S'(M' + R) = \bar{M} + 2R, \quad S' \in W, \quad \bar{M} \in \mathfrak{P}(M, M'), \tag{15}$$

is valid only for

$$S' = 1 \quad \text{and} \quad \bar{M} = M + M'.$$

What is left to be shown is that, under condition (12), the sum over $S'' \in W$ in Eq. (11) is trivial, i.e., $S'' = 1$ only. If Eq. (11) is assumed to hold, then

$$M'' + SR$$

is a dominant weight for all $S \in W$, $M'' \in \mathfrak{P}(M, M'')$, and again the weights $M'' + R + SR$ lie *inside* the fundamental domain for all $S \in W$, $M'' \in \mathfrak{P}(M, M')$. Thus, $S''(M'' + R + SR)$ is a nondominant weight for all $S, S'' \in W$, $S'' \neq 1$, $M'' \in \mathfrak{P}(M, M'')$. However, the two sets of weights $S''(M'' + R) + S''SR$ and $S''(M'' + R) + SR$, $S, S'' \in W$, $S'' \neq 1$, are equal. It should hold that

$$S''(M'' + R) + SR = \bar{M} + 2R. \tag{16}$$

However, this is impossible unless $S'' = 1$ since the right-hand side of Eq. (16) is a dominant weight while the left-hand side is nondominant for $S, S'' \in W$, $S'' \neq 1$.

This shows that Racah's recurrence relation (7) for the inner multiplicity is a special case of the recurrence relation for the outer multiplicity. Accordingly, to the property

$$\gamma(M) = 1 \tag{17}$$

in the case of Racah's recurrence relation, there

corresponds

$$\bar{\gamma}(M + M') = 1 \tag{18}$$

in the case of the recurrence relation for the outer multiplicity. Two examples are given for Eq. (11). The first example is chosen such that the Racah–Speiser lemma applies, while in the second example the Biedenharn lemma applies. [Thus, the second example can also be taken as an example for Racah's recurrence relation (7) for the inner multiplicity.]

Example 1: $D(\frac{1}{3}(2, 2, -4)) \otimes D(1, 0, -1)$: One obtains, from Eq. (11),

$$\begin{aligned}
&\bar{\gamma}(\tfrac{1}{3}(5, 2, -7)) = 1,\\
&\bar{\gamma}(\tfrac{1}{3}(2, -1, -1)) = \bar{\gamma}(\tfrac{1}{3}(2, 2, -4))\\
&\qquad\qquad -\bar{\gamma}(\tfrac{1}{3}(5, 2, -7)) + 1,\\
&\bar{\gamma}(\tfrac{1}{3}(2, 2, -4)) = \bar{\gamma}(\tfrac{1}{3}(5, -1, -4)),\\
&\bar{\gamma}(\tfrac{1}{3}(5, -1, -4)) = \bar{\gamma}(\tfrac{1}{3}(5, 2, -7)),
\end{aligned}$$

and thus

$$\begin{aligned}
\bar{\gamma}(\tfrac{1}{3}(5, 2, -7)) &= \bar{\gamma}(\tfrac{1}{3}(2, -1, -1))\\
&= \bar{\gamma}(\tfrac{1}{3}(2, 2, -4))\\
&= \bar{\gamma}(\tfrac{1}{3}(5, -1, -4)) = 1.
\end{aligned} \tag{19a}$$

Example 2: $D(3, 0, -3) \otimes D(1, 0, -1)$: One obtains, from Eq. (11),

$$\begin{aligned}
&\bar{\gamma}(4, 0, -4) = 1,\\
&\bar{\gamma}(2, 0, -2) = \bar{\gamma}(4, 0, -4) + \bar{\gamma}(3, -1, -2)\\
&\qquad + \bar{\gamma}(2, 1, -3) - \bar{\gamma}(3, 1, -4)\\
&\qquad - \bar{\gamma}(4, -1, -3),\\
&\bar{\gamma}(3, -1, -2) = -\bar{\gamma}(4, 0, -4) + \bar{\gamma}(3, 0, -3),\\
&\bar{\gamma}(3, 0, -3) = \bar{\gamma}(4, -1, -3) + \bar{\gamma}(3, 1, -4),\\
&\bar{\gamma}(4, -1, -3) = \bar{\gamma}(4, 0, -4),\\
&\bar{\gamma}(3, 1, -4) = \bar{\gamma}(4, 0, -4),\\
&\bar{\gamma}(2, 1, -3) = \bar{\gamma}(3, 0, -2) - \bar{\gamma}(4, 0, -4),
\end{aligned}$$

and thus

$$\begin{aligned}
\bar{\gamma}(4, 0, -4) &= \bar{\gamma}(4, -1, -3) = \bar{\gamma}(3, 1, -4)\\
&= \bar{\gamma}(2, 1, -3) = \bar{\gamma}(3, -1, -2)\\
&= \bar{\gamma}(2, 0, -2) = 1,\\
\bar{\gamma}(3, 0, -3) &= 2.
\end{aligned} \tag{19b}$$

VI. BRANCHING MULTIPLICITY

A recurrence relation for the branching multiplicity of an irreducible representation of the group G under the restriction of this group to a subgroup G_s can be obtained by methods similar to those used in the previous section. The subgroups G_s for which this formula for the branching multiplicity holds are limited by the methods used for its derivation. It turns out that the subgroups G_s for which this recurrence relation is valid are semisimple Lie groups with the property that the toroid (Cartan subgroup) of the subgroup G_s is contained in the toroid of the group G. The reason for this requirement is that under this condition the elements of the toroid of the subgroup G_s and the elements of the toroid of the group G can be diagonalized simultaneously. In other words, the "weight space of the subgroup" then forms a subspace of the "weight space of the group." This condition on the subgroups G_s can be recognized by following the proof of the recurrence relation. Similarly, it can be verified that a generalization of the formula is not difficult for restriction of the group G with respect to subgroups G', which are direct products of semisimple Lie groups (each satisfying the condition stated above) and to subgroups which are groups of the form $G' \times U(1) \times \cdots \times U(1)$ (or locally isomorphic to it), where G' is a semisimple Lie group (satisfying the condition stated above).

If a group G is restricted to a subgroup G_s, an irreducible representation $D(M)$ of the group G remains in general not irreducible, but decomposes into several irreducible representations of the subgroup G_s. In terms of characters, this reads as

$$[\chi^{(M)}]_{\text{restricted}} = \sum_{M_s'} \bar{\gamma}(M_s')\chi(M_s'), \tag{20}$$

where $\bar{\gamma}(M_s)$ is the multiplicity of the irreducible representation $D(M_s)$ of the subgroup G_s. Substituting Weyl's formula for the character on both sides of Eq. (20) yields

$$\begin{aligned}
&\left[\left(\sum_{S\in W}\delta_S e^{i[S(M+R),\varphi]}\right)\Big/\left(\sum_{S\in W}\delta_S e^{i[SR,\varphi]}\right)\right]_{\text{restricted}}\\
&\quad = \sum_{M_s'}\bar{\gamma}(M_s')\left(\sum_{S_s\in W_s}\delta_{S_s}e^{i[S_s(M_s'+R_s),\varphi_s]}\right)\Big/\\
&\qquad\qquad \left(\sum_{S_s\in W_s}\delta_{S_s}e^{i[S_sR_s,\varphi_s]}\right).
\end{aligned} \tag{21}$$

Using Eq. (1), Eq. (21) goes over, after some rearrangements, into

$$\begin{aligned}
&\sum_{\substack{S\in W\\ S_s\in W_s}}\delta_S\delta_{S_s}e^{i[LS(M+R)+S_sR_s,\varphi_s]}\\
&\quad = \sum_{M_s'}\bar{\gamma}(M_s')\sum_{\substack{S\in W\\ S_s\in W_s}}\delta_S\delta_{S_s}e^{i[S_s(M_s'+R_s)+LSR,\varphi_s]}.
\end{aligned} \tag{22}$$

Multiplying both sides of Eq. (22) by $e^{-i[M_s+R_s+LR,\varphi_s]}$ and using the orthogonality properties of the trigonometric functions, one obtains the recurrence relation

for the branching multiplicity

$$\tilde{\gamma}(M_s) = -\sum_{\substack{S\in W \\ S_s\in W_s \\ S=S_s\neq 1}} \delta_S\delta_{S_s}\tilde{\gamma}(S_s\{M_s+R_s+L(R-SR)\}-R_s) + \sum_{\substack{S\in W \\ S_s\in W_s}} \delta_S\delta_{S_s}\delta_{LS(M+R)+S_s(\bar{M}_s+R_s),M_s+R_s+LR}, \quad (23)$$

with $\bar{M}_s = 0$.

In deriving Eq. (23), the fact has been used that a reflection or a product of reflections (i.e., the application of an element $S \in W$) does not change the content of the relation

$$S_s(M_s' + R_s) = M_s + R_s + L(R - SR), \quad (24)$$

and thus the equation

$$M_s' + R_s = S_s^{-1}[M_s + R_s + L(R - SR)] \quad (25)$$

is equivalent to Eq. (24). In fact, in Eq. (24) a succession of reflections brings the weight $M_s' + R_s$ into the vector given on the right-hand side of the equation, while in Eq. (25) a succession of reflections brings the vector $M_s + R_s + L(R - SR)$ into the vector $M_s' + R_s$. (Note that it is summed over $S_s \in W_s$, and that $\delta_{S_s} = \delta_{S_s^{-1}}$.) Writing S_s again, instead of S_s^{-1}, we see that Eq. (23) follows.

Looking at Eq. (23) more closely, we note the interesting fact that Eq. (23) contains the recurrence relation for the outer multiplicity as a special case; that is, when the subgroup G_s to which the group G is restricted is chosen to be the group G itself, then Eq. (23) goes into the recurrence relation (11) for the special case of $D(M) \otimes D(M' = 0)$. [This explains why the $\bar{M}_s = 0$ has been introduced into Eq. (23).] That this is true can be verified easily. If $G_s = G$, then $L = 1$, $W_s = W$, and $S_s = S'$. Moreover, all weights lose the s suffix and $M_s \to \bar{M}$.

Thus, the following chain has been obtained. The recurrence relation (23) for the branching multiplicity contains as special case the recurrence relation (11) for the outer multiplicity [for $D(M) \otimes D(0)$], which in turn contains as a special case the recurrence relation (7) for the inner multiplicity. In concluding this section, two examples are given for the recurrence relation (23).

Example 1: $G = SU(3)$, $G_s = SU(2)$; $R = \alpha$, the positive root of $SU(2)$. In this case $R_s = \frac{1}{2}R$ and

$$L(m) = m_s = (m_1 + \tfrac{1}{2}m_2, -(m_1 + \tfrac{1}{2}m_2)),$$

with $m = (m_1, m_2, m_3)$, $m_1 + m_2 + m_3 = 0$, and $m_{s1} + m_{s2} = 0$. Moreover, in

$$S_s(M_s + R_s + L(R - SR)) - R_s, \quad S_s \in W_s,$$

only $S_s = 1$ is possible, since $M_s + R_s + L(R - SR)$ is dominant. Then one obtains the following for $D(M) = D(1, 0, -1)$ (the octet):

$$\tilde{\gamma}(1, -1) = 1,$$
$$\tilde{\gamma}(\tfrac{1}{2}, -\tfrac{1}{2}) = -[-2\tilde{\gamma}(1, -1)] = 2,$$
$$\tilde{\gamma}(0, 0) = -[2\tilde{\gamma}(\tfrac{1}{2}, -\tfrac{1}{2})] - 3 = 1.$$

For $\tilde{\gamma}(1, -1)$, the first term on the right-hand side of Eq. (23) is equal to zero, while for $\tilde{\gamma}(\frac{1}{2}, -\frac{1}{2})$, the second term is equal to zero. Both terms contribute to $\tilde{\gamma}(0, 0)$.

Example 2: $G = SU(3)$, $G_s = SO(3)$; $L(m) = m_s = m_1 - m_2$, $m = (m_1, m_2, m_3)$, $m_1 + m_2 + m_3 = 0$, $R = (1, 0, -1)$, $R_s = \frac{1}{2}$. Again, $S_s = 1$ is the only possibility for the first term of Eq. (23). One obtains

$$2\tilde{\gamma}(1) = \tilde{\gamma}(0) + 2,$$
$$2\tilde{\gamma}(2) = \tilde{\gamma}(1) + 1,$$
$$\tilde{\gamma}(0) = \tilde{\gamma}(2) - 1,$$

and from these equations there follows

$$\tilde{\gamma}(2) = 1, \quad \tilde{\gamma}(1) = 1, \quad \tilde{\gamma}(0) = 0.$$

ACKNOWLEDGMENT

The author wishes to express his thanks for the hospitality extended to him at the Research Institute for Theoretical Physics, Helsinki, during July and August, 1968, where the first version of this paper was completed.

APPENDIX

By the same simple methods as used throughout this paper a formula for the branching multiplicity $\tilde{\gamma}(M_s)$ can be derived which expresses the branching multiplicity in terms of inner multiplicities of the *subgroup* G_s and the Weyl group W of the group G. Thus, this formula is of particular interest, if the subgroup should happen to be very familiar, i.e., if the inner multiplicities are either known or easily accessible (the Weyl group W presents no problem). For properties of the subgroup G_s and other details the reader is referred to Sec. VI of this paper.

The formula is obtained as

$$\sum_{M_s}\sum_{S\in W} \delta_S\tilde{\gamma}(M_s)\gamma^{M_s}(m_s + L(R - SR)) = \sum_{S\in W} \delta_S\delta_{LS(M+R),m_s+LR}, \quad (A1)$$

where the sum extends over all $D(M_s)$ occurring under the restriction of the group G to the subgroup G_s.

Assuming the multiplicities $\gamma^{M_s}(m_s + L(R - SR))$ of the subgroup G_s to be known, then Eq. (A1) provides a set of linear equations for the $\tilde{\gamma}(M_s)$, one equation for each $m_s = M_s^{(1)}, \cdots, M_s^{(n)}$, where the

$M_s^{(1)}, \cdots, M_s^{(n)}$ are the highest weights of the irreducible representations $D(M_s)$ occurring under the restriction of the group G to the subgroup G_s.

A simple example is given in order to illustrate Eq. (A1). The group G is taken to be $SU(3)$, the group G_s is taken to be $SU(2)$. $R = (1, 0, -1)$, $L(m) = m_s = (m_1 + \frac{1}{2}m_2, -(m_1 + \frac{1}{2}m_2))$, $m = (m_1, m_2, m_3)$, $m_1 + m_2 + m_3 = 0$. $D(M) = D(1, 0, -1)$, i.e., the octet of $SU(3)$. Then [denoting the weights of $SU(2)$ by the first component $m_1 + \frac{1}{2}m_2$ only] one obtains, for $m_s = 1$,

$$\tilde{\gamma}(1)\gamma^1(1) = 1,$$

for $m_s = \frac{1}{2}$,

$$-2\tilde{\gamma}(1)\gamma^1(1) + \tilde{\gamma}(\tfrac{1}{2})\gamma^{\frac{1}{2}}(\tfrac{1}{2}) = 0,$$

and, for $m_s = 0$,

$$\tilde{\gamma}(1)\gamma^1(0) - 2\tilde{\gamma}(\tfrac{1}{2})\gamma^{\frac{1}{2}}(\tfrac{1}{2}) + \tilde{\gamma}(0)\gamma^0(0) = -2.$$

Since, however, $\gamma^{M_s}(m_s) = 1$ if $m_s \in D(M_s)$, it follows that

$$\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & -2 & 1 \end{pmatrix} \begin{pmatrix} \tilde{\gamma}(1) \\ \tilde{\gamma}(\frac{1}{2}) \\ \tilde{\gamma}(0) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix},$$

and, solving this set of equations, one obtains

$$\tilde{\gamma}(1) = \tilde{\gamma}(0) = 1, \quad \tilde{\gamma}(\tfrac{1}{2}) = 2.$$

It might be worthwhile pointing out that, for $G_s = G$, Eq. (A1) goes over into Racah's formula (7).

* Visiting member.
† On leave of absence.
[1] L. S. Pontrjagin, *Topologische Gruppen* (Teubner, Leipzig, 1958), Vol. II.
[2] C. Chevalley, *Theory of Lie Groups* (Princeton U.P., Princeton, N.J. 1941), Vol. I.
[3] N. Jacobson, *Lie Algebras* (Interscience, New York, 1962), p. 261.
[4] See. Ref. 3, p. 247.
[5] See Ref. 3, p. 260.
[6] G. Racah, *Group Theoretical Concepts and Methods in Elementary Particle Physics*, F. Gursey, Ed., (Gordon and Breach, New York, 1964).
[7] See Ref. 3, p. 262.
[8] N. Straumann, Helv. Phys. Acta **38**, 481 (1965).
[9] B. Gruber, Ann. Inst. Henri Poincaré **8A**, 43 (1968).
[10] A. J. Macfarlane, L. O'Raifeartaigh, and P. S. Rao, J. Math. Phys. **8**, 536 (1967); G. E. Baird and L. C. Biedenharn, J. Math. Phys. **5**, 1730 (1964).
[11] R. M. Delaney and B. Gruber, J. Math. Phys. **10**, 252 (1969).
[12] V. B. Mandel'tsveig, Yad. Fiz. **1**, 1106 (1965) [Sov. J. Nucl. Phys. **1**, 787 (1965)].
[13] A. U. Klymyk, Academy of Science, Ukranian SSR, Kiev, Institute for Theoretical Physics, Preprint No. 67-17, 1967.
[14] F. Zaccaria, J. Math. Phys. **7**, 1548 (1966); **8**, 953 (1967).
[15] B. Vitale, Nuovo Cimento **44**, 291 (1966).
[16] See Ref. 3, p. 263.
[17] J. Tarski, J. Math. Phys. **3**, 569 (1963).
[18] B. Gruber, Nuovo Cimento **48**, 23 (1967).
[19] B. Gruber and F. Zaccaria, Nuovo Cimento, Suppl., **5**, 914 (1967); B. Gruber and H. J. Weber, Proc. Roy. Irish Acad. **66A**, 31 (1968); B. Gruber and H. J. Weber, "Recurrence Relation for Kostant's Formula," Proc. Roy. Irish Acad., in press. In these papers use has been made of recurrence relations for the partition function.
[20] Equation (6) is, in fact, valid whenever $\gamma^M(m) = 0$, which does not necessarily imply that it holds only for $|m| > |M|$.
[21] B. Gruber, Lecture Notes on Multiplicities in the Classical Groups, Instituto di Fisica Teorica, Universita di Napoli, Napoli, Italy (Fall, 1968); D. Radhakrishnan, J. Math. Phys. **10**, 2129 (1969).
[22] See Ref. 3, p. 255.